Climate
of the Past

Patterns in data of extreme droughts/floods and harvest grades derived from historical documents in eastern China during 801–1910

Zhixin Hao[1,2], Maowei Wu[1], Jingyun Zheng[1,2], Jiewei Chen[1,2], Xuezhen Zhang[1,2], and Shiwei Luo[3]

[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]University of Chinese Academy of Sciences, Beijing, 100049, China
[3]School of Geography and Tourism, Chongqing Normal University, Chongqing, 401331, China

Correspondence: Jingyun Zheng (zhengjy@igsnrr.ac.cn) and Shiwei Luo (lsw@cqnu.edu.cn)

Abstract. In China, historical documents record a large quantity of information related to climate change and grain harvest. This information can help to explore the impacts of extreme drought or flood on crop production, which can provide implications for the adaptation of agriculture to higher-probability extreme climate in the context of global warming. In this paper, reported extreme drought/flood chronologies and reconstructed grain harvest series derived from historical documents were adopted in order to investigate the association between the reported frequency of extreme drought/flood in eastern China and reconstructed poor harvests during 801–1910. The results show that extreme droughts were reported more often in 801–870, 1031–1230, 1481–1530, and 1581–1650 over the whole of eastern China. On a regional scale, extreme droughts were reported more often in 1031–1100, 1441–1490, 1601–1650, and 1831–1880 in the North China Plain, 801–870, 1031–1120, 1161–1220, and 1471–1530 in Jianghuai, and 991–1040, 1091–1150, 1171–1230, 1411–1470, and 1481–1530 in Jiangnan. The grain harvest was reconstructed to be generally poor in 801–940, 1251–1650, and 1841–1910, but the reconstructed harvests were bumper in 951–1250 and 1651–1840, approximately. During the entire period from 801 to 1910, the frequency of reporting of extreme droughts in any subregion of eastern China was significantly associated over the long term with lower reconstructed harvests. The association between reported frequency of extreme floods and reconstructed low harvests appeared to be much weaker, while reconstructed harvest was much worse when extreme drought and extreme flood in different subregions were reported in the same year. The association between reconstructed poor harvests and reported frequency of regional extreme droughts was weak during the warm epoch of 920–1300 but strong during the cold epoch of 1310–1880, which could imply that a warm climate could weaken the impact of extreme drought on poor harvests; yet other historical factors may also contribute to these different patterns extracted from the two datasets.

1 Introduction

Extreme drought is the most damaging climate-related hazard to agriculture, as it leads to crop failure due to a reduction in water supply. Numerous studies, based on observational data, show that many regions of the world (in particular southern Europe and western Africa) have been experiencing trends toward longer and more intense droughts since the 1950s (IPCC, 2012; Dai, 2011, 2013). The climate in eastern China is dominated by the Asian monsoon, with a large precipitation variability, leading to drought and flood on a regional, and occasionally larger, scale (Ding et al., 2013). Data show that severe and extreme droughts have become more frequent in northeast China, northern China, and the eastern part of northwest China since the late 1990s and in southwest China between 2006 and 2013 (Zou et al., 2005, 2010; Zhai et al., 2010; and Yu et al., 2014). Assessment reports show that droughts affected, on average, 20.9×10^4 km^2

of cropland annually from 1949 to 2013, accounting for one-sixth of China's total arable land. From 1991 to 2013, the loss of grain crop yields as a result of drought was, on average, $26.75 \times 10^9 \, \text{kg yr}^{-1}$. From 2004 to 2013, direct economic losses caused by drought averaged CNY 63.67 billion (approximately USD 10 billion) annually. In addition, projections suggest that the risk of extreme droughts will increase in most of China along with future global warming (D. Qin et al., 2015).

Recently, a number of studies have focused on long-term drought event reconstruction and extreme drought identification through the use of high-resolution proxy data, such as historical documents and tree rings, at local, regional, and national scales. These studies have found that historical extreme droughts had tremendous impacts on agriculture, livelihoods, and socioeconomic systems (e.g., Hao et al., 2010a, b; Shen et al., 2007, 2008; Zhang, 2000, 2005; Zheng et al., 2006, 2014a; Gou et al., 2015a, b; Deng et al., 2016; Liu et al., 2017; Yang et al., 2014; C. Qin et al., 2015; Li et al., 2018; Xiao et al., 2017; Zhang et al., 2011; and Liang et al., 2006). For example, using Chinese historical documents, Hao et al. (2010b) found that droughts in 1876 and 1877 led to a 45 % and 50 % harvest reduction in each of those years, respectively, which caused the price of rice to increase by 5–10 times its normal value and resulted in the deaths of more than 13 million people, due to famine and plague. Zheng et al. (2014a) found that severe droughts in 1627–1643 triggered and also significantly promoted widespread peasant uprising, which finally resulted in the collapse of the Ming dynasty.

Meanwhile, several studies also focused on the correlation between reconstructed climate and grain harvest data over the past 2000 years, expecting to build a possible connection between climate and harvest fluctuations at the macroscale. For example, Su et al. (2014) investigated the fluctuations in reconstructed climate and grain harvest data in China from 206 BCE to 960 CE and found a high positive correlation between reconstructed temperature and grain yield data. According to their datasets, bumper harvest decades corresponded to warm–normal or warm–wet climate reconstructions, while poor harvest decades often corresponded to cold–dry climate reconstructions. Yin et al. (2015, 2016) further demonstrated that reconstructed harvests were mostly poor between 210 BCE and 1910 CE under cold conditions when the climate changed from warm to cold while dry or from wet to dry; reconstructed grain harvest was better during the warm phase. It should be noticed that these conclusions are based on reconstructed datasets and do not necessarily reflect actual historical connections due to several reasons addressed in the Discussion section. These studies focused on the connection between agriculture and long-term climate change, while the effect on harvest induced by short-term extreme events (such as extreme droughts/floods) might be different. Therefore, this study aims to explore patterns in the data of extreme droughts/floods and harvests
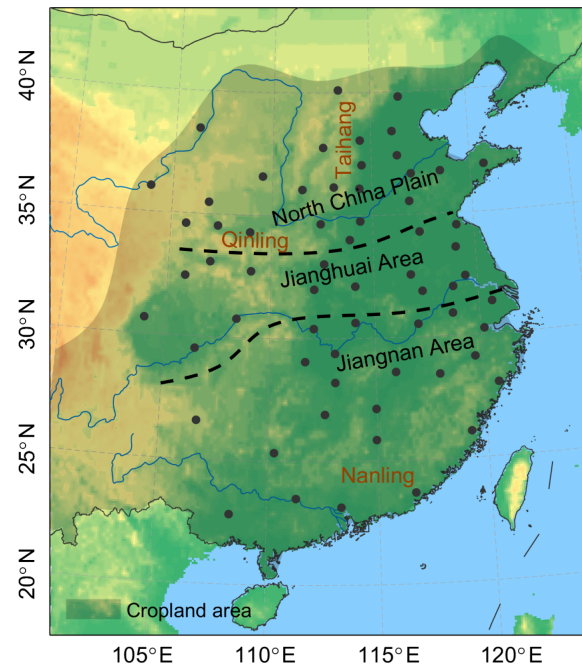
Figure 1. Locations of 63 drought/flood grade data stations and subregion divisions. The shaded area indicates approximate cropland distributed from 801 to 1910, approximately.

in eastern China from 801 to 1910, using reconstructions of regional grain harvest grades and extreme drought/flood events derived from Chinese historical documents. The results from historical datasets could provide implications to improve the current understanding of the relationship between poor harvests and extreme drought/flood and how cold and warm periods, such as the Medieval Climate Anomaly (MCA, 950–1250) and the Little Ice Age (LIA, 1450–1850) (IPCC, 2013), contributed to differences in that relationship.

2 Data and method

2.1 Data

Two datasets were used, including a chronology of regional extreme drought and flood and a series of grain harvest grades. Both were derived from Chinese historical documents.

1. The chronology of regional extreme drought and flood dataset is described below. This chronology was reconstructed based on the annual drought/flood grade at 63 stations in eastern China collected since 137 BCE (east of 105° E 25–40° N approximately). This area was divided into three subregions: the North China Plain, Jianghuai, and Jiangnan (Fig. 1) (Hao et al., 2010a, and renewed by Zheng et al., 2014b). Records for the reconstruction of the annual drought/flood grade at 63 stations were extracted from such historical documents as
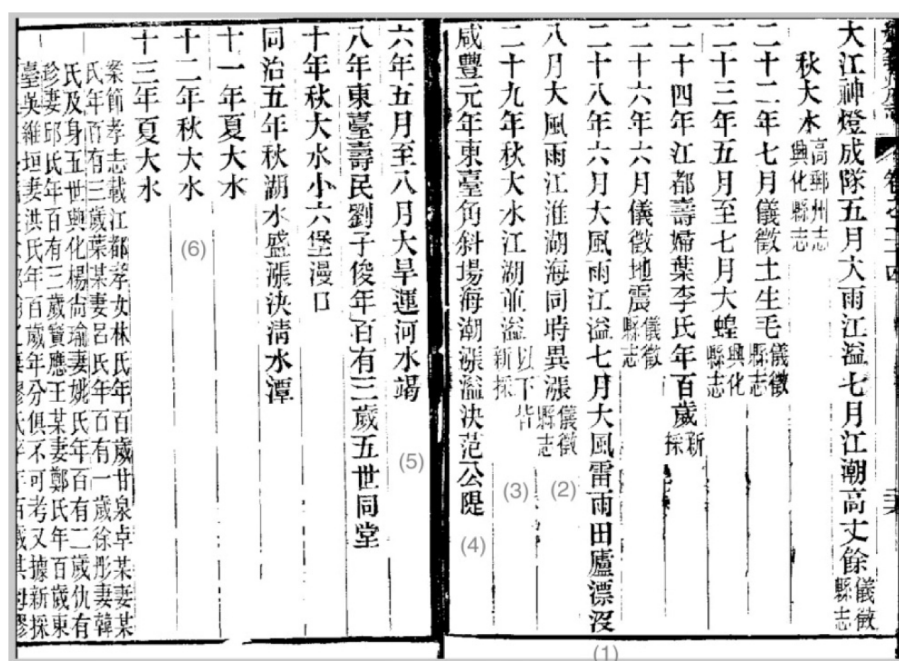
Figure 2. An example of anomalous climate information recorded in local gazettes (quoted from *Gazettes of Yangzhou Prefecture*, compiled in 1874). The two pages list anomalous climate information in that region from 1842 to 1874, which is read from right to left, and dated according to the Chinese lunar calendar. Numbers in parenthesis are added by the authors to show meaning. (1) In the 28th year (of the Daoguang Emperor, 1848), the sixth month, due to the strong wind and heavy rain, the Yangtze River overflowed; the seventh month, because of strong wind and thunderstorms, fields and houses were submerged. (2) In the eighth month there was strong wind and heavy rain; the level of the Yangtze River, the Huai River, the Gaoyou Lake, and the sea abnormally rose at the same time; the citation from *Gazettes of Yizheng County* (a county in the prefecture), where records of "strong wind" and "the sea level abnormally rose" also suggested heavy rain induced by a typhoon. (3) There was a flood in the autumn of the 29th year (1849); the Yangtze River and the (Gaoyou) Lake overflowed at the same time. Notes of "below are newly collected", suggest that these records were quoted from earlier local gazettes. (4) In the first year of Emperor Xianfeng (1851), in Jiaoxiechang (the name of a salt field) in Dongtai County (a county in the prefecture), the tide from the sea overflowed and broke the seawall of Mr. Fan. (5) The sixth year (1856), the fifth to eighth months, there was severe drought and the Grand Canal dried up. (6) Autumn of the 12th year of Emperor Tongzhi (1873), there was a flood. This picture is taken by the author from the original file, which is stored in the library of the Institute of Geographic Sciences and Natural Resources Research, CAS. We obtained the rights to use and publish this picture.

Twenty-Four Histories and *Qing History Draft* (official histories of each Chinese dynasty), chronicles, miscellaneous historical books, local gazettes, and others. For example, in both *The Book of Tang* and *New Book of Tang* (two books of *Twenty-Four Histories*), documents in the 19th year of Emperor Zhenyuan in the Tang Dynasty (803 CE) record:

> In day 25 of the sixth month [of the Chinese lunar calendar or 17 July 803 in the solar calendar; same hereinafter], no rain fell in the Guanzhong area [now Xi'an and the surrounding areas] from the first month [27 January to 24 February] until now. Hundreds of officials and many people are praying for rain. In day 26 of the seventh month [18 August], rain. In day 17 of the eighth month [6 September], heavy rain. From mid-autumn, drought occurred again.

Descriptions such as these clearly provide information about the duration, spatial coverage, and intensity of drought. The drought in that year was also recorded in *The Corpus of Huangpu Chizheng* (written by Huangpu Chizheng, 777–835) and in *The Complete Prose Works of the Tang Dynasty* (edited by Dong Gao, 1740–1818). These texts provide different sources for cross-validation, thus increasing the reliability of the records. Figure 2 shows a copy of a local gazette, which regularly recorded anomalous climate information at the county, prefectural, or provincial level. From this sample comes information about the duration and intensity of drought or flood occurrence in Yangzhou Prefecture. For example, in 1848, there were three instances of heavy rain, causing flooding from the sixth to the eighth month. In 1856, severe drought from the fifth to the eighth month led to the drying up of the Grand Canal (see Fig. 2 caption). As this sample sug-

gests, historical documents were usually focused on events that were due to no or less precipitation than usual and, thus, these could be regarded as meteorological droughts rather than hydrological or agricultural droughts, although some records also report impacts on the hydrosphere (e.g., rivers drying up) or agriculture (e.g., wilting of crops). Historical documents, therefore, appear to report all droughts equally, rather than only those affecting crops. Similarly, floods recorded in the historical documents could be regarded as more rain or heavier rain than usual, rather than in the context of rivers bursting their banks or tsunamis, although some records also report the impacts of overflowing or bursting lakes and rivers due to more or heavier precipitation.

Based on these records, Zhang (1996) reconstructed a dataset of annual drought/flood grades at 63 stations from 137 BCE. Each station consisted of a local area of approximately 20 counties with the same climate. Grades were classified using ideal frequency criteria of 10 % (grade 1, severe drought), 20 % (grade 2, drought), 40 % (grade 3, normal), 20 % (grade 4, flood), and 10 % (grade 5, heavy flood) for the whole area and the whole study period with available data since 137 BCE. These grades were calibrated based on descriptions of duration, intensity, and area of the drought/flood event during the wet season (usually May to September) and its impact (Table S1 in the Supplement). Thus, the season of the drought/flood grade data overlaps with critical agricultural activities and phases of crop growth. The drought/flood grade data are unevenly spatially distributed across the 2000-year period. For example, drought/flood grade data for south China (south of 30° N approximately) were limited for the period before 760 CE, and there were even fewer data for the region south of the Huai River (approximately 34° N) before 300 CE (Zhang, 1996). However, coverage of this dataset has extended to south China since 760 CE and, therefore, covered the whole study area. There also exist missing data before 1470, as fewer historical documents have survived from these earlier times (Zhang, 1996; Hao et al., 2016). Statistics show that the mean percentage of available data was 44.1 % for 800–1469 and only 20 % or lower for periods around 850 and the 880s–920s, 1230s–1250s, 1360s, and 1390s. During the period of 800–1469, the mean percentage of available data reporting "disasters or extremes" (i.e., grade 1, 2, 4, 5) was 41.8 %, and the mean percentage of data reporting "normal" (i.e., grade 3) was 2.3 % (Fig. S1 in the Supplement). Moreover, there was a period of 520 years when no normal record existed. This means that most of the available grade data recorded disasters and extremes following the principle of recording the unusual rather than the normal in the compilation of Chinese history. Even so, it is reasonable to infer that a con-

siderable proportion of extreme events was recorded in 800–1469 by comparing the percentage of records reporting "disasters and extremes" in that period and in the ideal frequency (41.8 % compared with 60 %, respectively). Compared with other datasets, this dataset has certain spatiotemporal advantages. For example, it spans a longer period than the dataset for yearly dry–wet grades (1: very wet; 2: wet; 3: normal; 4: dry; and 5: very dry) for the 120 stations across China from 1470 to 1979 (Academy of Meteorological Science of China Central Meteorological Administration, 1981). In addition, this dataset also covers more areas with higher spatial resolution for local stations (each consisting of around 20 counties), compared with the series in Zhang et al. (1997) of regional dry/wet grades in six areas (it consisted of more than 100 counties) from the Lower Yangtze Valley to the North China Plain during 960–1992. Therefore, this dataset provides a valuable proxy and has already been used to study characteristics of precipitation change in eastern China over the past 2000 years. For example, Zheng et al. (2006) used this dataset to reconstruct a 1500-year regional dry/wet index series for the North China Plain (approximately 34–40° N), the Jianghuai area (approximately 31–34° N), and the Jiangnan area (approximately 25–31° N). Hao et al. (2016) used the dataset to investigate spatial precipitation anomaly patterns in eastern China during centennial cold and warm periods over the past 2000 years. However, this dataset also has weaknesses. For example, the reconstruction derived from the historical documents relies highly on the accuracy of the compilation and is only available in a form in which the retracing of the original sources and the reconstruction of previous steps are not the same as reconstructions from natural evidences.

Since there existed missing data for each of the subregions (i.e., North China Plain, Jianghuai, and Jiangnan) shown in Fig. 1, extreme drought and flood years were identified according to both the coherence of drought/flood grade for each individual station and the percentage of each drought/flood grade occurrence within the available data in each subregion. In detail, an extreme drought (or flood) year in any subregion was defined with two terms: (1) grade 3 occurred in less than 25 % of all stations (i.e., disasters or extremes were reported for more than 75 % of all stations with available data); (2) grade 1 and grade 2 (or grade 5 and grade 4) occurred in more than 80 % of stations with records of disasters or extreme conditions (stations with grade 3 or missing data were excluded) and grade 1 (or grade 5) occurred in two or more stations. Extreme drought or extreme flood years were defined in this way under the assumption that the probabilities for omitting drought and flood events in reporting and transmission of his-

torical records were random and unbiased, despite the greater frequency of missing data in the earlier records. In other words, if one period had a large number of documents, it was expected to be rich in both drought and flood records and vice versa. Therefore, the amount of missing data should not have a significant effect on the relative drought-to-flood ratio within the available data. To verify the rationality of this method and these criteria, validation was conducted in Hao et al. (2010a), based on 10 extreme events identified from a series of precipitation observations in each subregion according to a threshold of probabilities of 10 % and 90 % occurrence. In this validation, all or part of grade 3 stations were deliberately omitted, and only 40 % or 60 % of stations with a disaster or extreme grade were reserved without changing the drought-to-flood ratio within the available data. The results show that, with one exception, years of extreme drought and extreme flood, identified according to this method and these criteria, closely matched those extreme events identified by precipitation data, demonstrating that the method and criteria were reasonable. The reason for the close match is that precipitation variability in eastern China is dominated by the East Asian summer monsoon (EASM). Therefore, when extreme drought or flood events occur, the precipitation variations for stations within each subregion usually share similar relative magnitudes. Moreover, this validation also demonstrates that the criteria used for identifying historical subregional extreme drought or flood years are equivalent to the definition of an extreme climate event with an occurrence probability of less than 10 % from 1951 to 2000, which was suggested by IPCC (2012) and has often been adopted in other research focusing on extreme climate.

Furthermore, extreme drought or flood years for the whole of eastern China were defined by reconstructed extreme drought or flood years from the three subregions combined and the dry–wet index series over the whole of eastern China, reconstructed and renewed by Zheng et al. (2006, 2014c), who synthesized annual drought/flood grades from all 63 stations. The criteria to define an extreme drought or flood year for the whole of eastern China were as follows: extreme drought (or flood) in all three subregions, extreme drought (or flood) in two subregions and no extreme flood (or drought) in the other subregion, or extreme drought (or flood) in only one subregion with an annual dry–wet index for the whole of eastern China lower (or greater) than the mean of that series by at least 1.282 times the standard deviation (i.e., less than 10 % probability of drought or flood occurrence over the last 2000 years). Moreover, if both reconstructed extreme drought years and extreme flood years for different subregions overlapped in the same year, that year was defined as an extreme year with

the co-reporting of both conditions. In addition, to illustrate the uncertainty of regional extreme drought/flood reconstructions, the confidence levels of these reconstructions were also assessed based on the percentage of years with data available in 50-year intervals and are defined as follows: extremely high confidence is defined as more than 99 %, very high confidence: > 90 %, high confidence: > 80 %, medium confidence: > 50 %, and low confidence: > 33.3 %.

2. The grain harvest grade series is described below. This series was reconstructed based on historical records collected from *Twenty-Four Histories* and *Qing History Draft*, which included descriptions of yearly grain yield estimates (e.g., a golden bumper year, abundant, plentiful, not bad, slightly poor, poor, and very poor) and related information regarding national food security (e.g., enough, insufficient, starving, famine, and widespread begging), features of tax remission induced by agricultural disasters, people's livelihoods, grain prices, and grain storage status at the country level from 206 BCE to 1910 CE (Su et al., 2014; Yin et al., 2015). In Chinese historical documents, the yearly harvest was usually recorded as a relative level compared to an expected maximum yield, rather than crop yield per hectare, although some records also report impacts of harvest fluctuation on food availability, tax remissions, livelihoods, and so on. Therefore, these harvest records exclude differences in absolute yield between subregions with different climates, soil fertility and type, crop varieties, etc., as well as the difference between historical periods with changing agricultural centers, farming technologies, staple crops, and so on (Su et al., 2014). As argued by Marks (1998), although it is unclear how yearly harvest percentages in historical times were estimated and what the standard harvest yield was, the ratings were also probably impressionistic; these harvest rating estimates in Chinese history were commonly used for food supply management and the administration of the state granary system by officials, and these estimates also served as public information for tenants who use the ratings to reduce the amount of rent they were expected to pay landlords. Therefore, it is believed that the harvest rating descriptions recorded throughout Chinese history are a good indication of fluctuations in yearly grain yields (Marks, 1998).

In the dataset, yearly harvest was classified into six grades: 1 – very poor, 2 – poor, 3 – slightly poor, 4 – average, 5 – near-bumper, and 6 – bumper. It should be noted that these reconstructed grades do not represent absolute grain yield but rather the relative percentage in the production of staple crops and reflect their interannual variability. The criteria and methods for year-by-year grading of the documentary records (i.e., grain yield descriptions and related information) were presented by Su et al. (2014) and summarized

by Yin et al. (2015). The classification of the yearly harvest grade and descriptions recorded in historical documents are shown in Table S2. Comparison and validation in their research (Su et al., 2014; Yin et al., 2015) show that this grade system corresponds to a percentage harvest rating system defined by the government as a code in *The Collected Statutes of the Qing Dynasty* (the "*Da-Qing Hui-Dian*" in Chinese was finally compiled in 1899, and the definition of the harvest rating system was recorded in Vol. 21), in which "very poor" corresponds to "below 40 %" of harvest yield, "poor" corresponds to 40 %–50 %, "slightly poor" corresponds to 50 %–60 %, "average" corresponds to 60 %–70 %, "near-bumper" corresponds to 70 %–80 %, and "bumper" corresponds to above 80 %.

It is worth noting that the further back in time one examines, the more records are missing. This is especially the case prior to 760 CE, regarding both the dataset of annual drought/flood grades for regional extreme drought/flood chronology and the grain harvest grade series. Thus, the study period chosen for investigating the associated variation in reported extreme droughts/floods and reconstructed harvests in eastern China was 801 to 1910. During this study period, several social factors existed that could have influenced China's total yield, such as changing borders of empires, the expansion of agricultural area (e.g., uplands in the south and southwest colonized by Han settlers during the late Ming and Qing periods), the updated crop varieties introduced from the New World (e.g., peanuts and sweet potatoes), the advanced agricultural management technology, and so on. However, such social factors should only have a limited influence on the yearly harvest grade dataset, since the harvest in the documents was reported as a relative level rather than an absolute yield; also, the main grain product area, the staple crop, and the cropping system was relatively stable throughout the study period (Yin et al., 2015).

In addition, a 2000-year temperature series, with decadal resolution for the whole of China (Ge et al., 2013), was also adopted to identify long-term cold and warm period fluctuations. This temperature series was reconstructed based on 28 temperature proxies using principal component (PC) regression and partial least squares (PLS) regression (Ge et al., 2013). The comparison showed that the pattern of long-term temperature change over China is roughly consistent with that shown by the IPCC (2013) for the Northern Hemisphere as an overlap of multi-reconstructions since 850.

2.2 Method

Four kinds of data processing methods were used in this study, including the moving average, the Wilcoxon rank sum test, the two-sample t test, and the contingency table with the chi-square test (χ^2).

1. To illustrate variations in the reported frequency of regional extreme drought and flood from 801 to 1910,

the moving-window frequency of reconstructed extreme drought/flood years in three subregions and in the whole of eastern China was first calculated with windows of 50 years and steps of 10 years. A smoothed series was made up of means of 801–850, 811–860, 821–970, and so on. Next, the Wilcoxon rank sum test was applied to examine which interval had significantly more or fewer drought/flood years compared to all the other intervals. By labeling extreme drought (or flood) years as 1 and nonextreme years as 0, the chronology of extreme drought (or flood) years could be transformed into a rank series, with the mean of this rank series equivalent to the frequency of drought (or flood) years. Therefore, those intervals with significantly more or fewer reconstructed drought (or flood) years could be recognized through a Wilcoxon rank sum test performed on the rank series.

2. To examine whether the reported frequency of extreme drought/flood in each subregion and in the whole of eastern China was associated with reconstructed poor harvest, harvest grade data for 801–1910 were divided into three categories: extreme drought years, extreme flood years, and nonextreme years; this is done according to the chronology of regional extreme drought and flood for each subregion or for the whole of eastern China. The two-sample t test was then applied to examine whether the means of the harvest grade in extreme drought (or flood) years and nonextreme years showed a significant difference. Meanwhile, a contingency table and the chi-square test (χ^2) were adopted to examine the association between reported regional extreme drought/flood and each grade of reconstructed harvest. For example, corresponding to the 720 years with available harvest grade data, there were 97 extreme drought years, 82 extreme flood years, and 541 non-extreme years in the reconstruction for the North China Plain. To examine whether there was a significantly higher frequency of harvest reconstructed as poor (grade 2) in extreme drought years compared to nonextreme years for this subregion, a contingency table was made by counting extreme drought years with grade 2, extreme drought years with other grades, nonextreme years with grade 2, and nonextreme years with other grades. The chi-square test (χ^2) was then adopted to test the significance of each.

3. To illustrate whether there were any differences in the frequency of reporting extreme drought/flood between cold and warm periods, the frequency of reconstructed extreme drought/flood years was calculated and the Wilcoxon rank sum test was performed for cold and warm periods in the three subregions and the whole of eastern China. Similarly, the contingency table and chi-square test (χ^2) were adopted to examine if there

Table 1. Frequency of reported extreme drought and flood in each subregion and the whole of eastern China, 801–1910.

Subregion	Extreme event	Number of years	Frequency (%)
North China Plain	Drought	133	12.0
	Flood	113	10.2
Jianghuai area	Drought	95	8.6
	Flood	118	10.6
Jiangnan area	Drought	90	8.1
	Flood	119	10.7
Whole of eastern China	Drought	126	11.4
	Flood	122	11.0
	Co-reporting of drought and flood	20	1.8

existed a significantly different pattern in the association between the reported frequency of regional extreme drought/flood and reconstructed harvests during cold and warm periods.

3 Results

3.1 The variation in reported regional extreme drought/flood and reconstructed grain harvest grade during 801–1910

Figure 3 shows the chronology of extreme drought/flood years and variation in each subregion (i.e., North China Plain, Jianghuai, and Jiangnan area) as well as the whole of eastern China from 801 to 1910. According to this chronology, the North China Plain had 133 extreme drought years and 113 extreme flood years, Jianghuai had 95 extreme drought and 118 extreme flood years, and Jiangnan had 90 extreme drought and 119 extreme flood years (Table 1). The comparison shows that the reporting of extreme droughts was a bit more frequent than extreme flood in the North China Plain due to its subhumid climate, while both Jianghuai and Jiangnan were reported to have slightly fewer extreme drought than extreme flood years due to the humid climate in these regions. The whole of eastern China had 126 extreme drought years, 122 extreme flood years, and 20 extreme years when extreme drought and flood were co-reported in different subregions.

Moreover, moving-window frequencies of extreme drought/flood years during each of the 50 years with a decadal step show remarkable multi-decadal variation in the frequency of reporting both extreme drought and flood in each subregion and the whole of eastern China from 801 to 1910 (Fig. 3). In the North China Plain, extreme droughts were reported more frequently in 1031–1100,

1441–1490, 1601–1650, and 1831–1880 but less frequently in 1151–1200, 1381–1430, 1651–1710, and 1721–1780. Extreme floods were reported more frequently in 941–1000, 1021–1070, 1721–1790, 1801–1830, and 1841–1900 but less frequently in 821–880, 1191–1260, and 1461–1550. In Jianghuai, extreme droughts were reported more frequently in 801–870, 1031–1120, 1161–1220, and 1471–1530, but less frequently in 881–960, 1221–1290, 1301–1350, 1381–1430, 1531–1580, and 1671–1760. Extreme floods were reported more frequently in 1551–1600, 1641–1680, 1691–1710, 1721–1770, and 1811–1890 but less frequently in 1261–1310, 1321–1370, and 1461–1550. In Jiangnan, extreme droughts were reported more frequently in 991–1040, 1091–1150, 1171–1230, 1411–1470, and 1481–1530, but less frequently in 1231–1320, 1361–1420, 1691–1750, and 1841–1890. Extreme floods were reported more frequently in 991–1040, 1151–1230, 1241–1300, 1400–1430, and 1821–1880 but less frequently in 851–970 and 1300–1370. For the whole of eastern China, extreme droughts were reported more frequently in 801–870, 1031–1230, 1481–1530, and 1581–1650 but less frequently in 891–1000, 1231–1320, 1381–1430, 1531–1580, 1651–1780, 1791–1850, and 1881–1910. Extreme floods were reported more frequently in 811–840, 951–990, 1051–1070, 1381–1430, 1641–1670, 1721–1770, and 1810–1870 but less frequently in 841–900, 1321–1380, 1431–1560, 1581–1640, and 1671–1700. Moreover, there were significantly more frequent co-reported extreme drought and flood years in 1291–1300, 1481–1500 (both with 2 years), and 1581–1600 (with 3 years).

Figure 4a shows the series of reconstructed grain harvest grades from 801 to 1910, with 390 years (35.1 %) of missing data. Percentages for each harvest grade were 4.9 % (grade 1), 16.0 % (grade 2), 18.4 % (grade 3), 14.4 % (grade 4), 6.9 % (grade 5), and 4.2 % (grade 6) (Table 2). In total, the percentage of all poor harvest years (grades 1, 2, and 3) was 39.3 %, while the percentage of bumper harvest years (grades 5 and 6) was only 11.2 %. Moreover, the moving-window percentage for each harvest grade and data missing for each 50 years with a decadal step (Fig. 4b) show an evident jump around the 1640s, when there was an increase in grade 4 years and a decrease in missing data years. This is mainly because *Qing History Draft* recorded more average harvest years, while the *Twenty-Four Histories* excluded most average years, according to the principle of "recording unusual rather than common events", as argued in Yin et al. (2015). Therefore, to avoid the effect of missing data and to maintain homogeneity in the statistics, data from missing years and average (grade 4) years were excluded in the analysis, as adopted in Zheng et al. (2006). Relative percentages for years with each anomalous harvest grade (i.e., 1, 2, 3, 5, and 6) compared with total anomalous years were calculated for each 50-year moving window (Fig. 4c) to illustrate variation in reconstructed grain harvest during 801–1910. By excluding grade 4 and missing data years, the relative percentages for each harvest grade for the whole period were
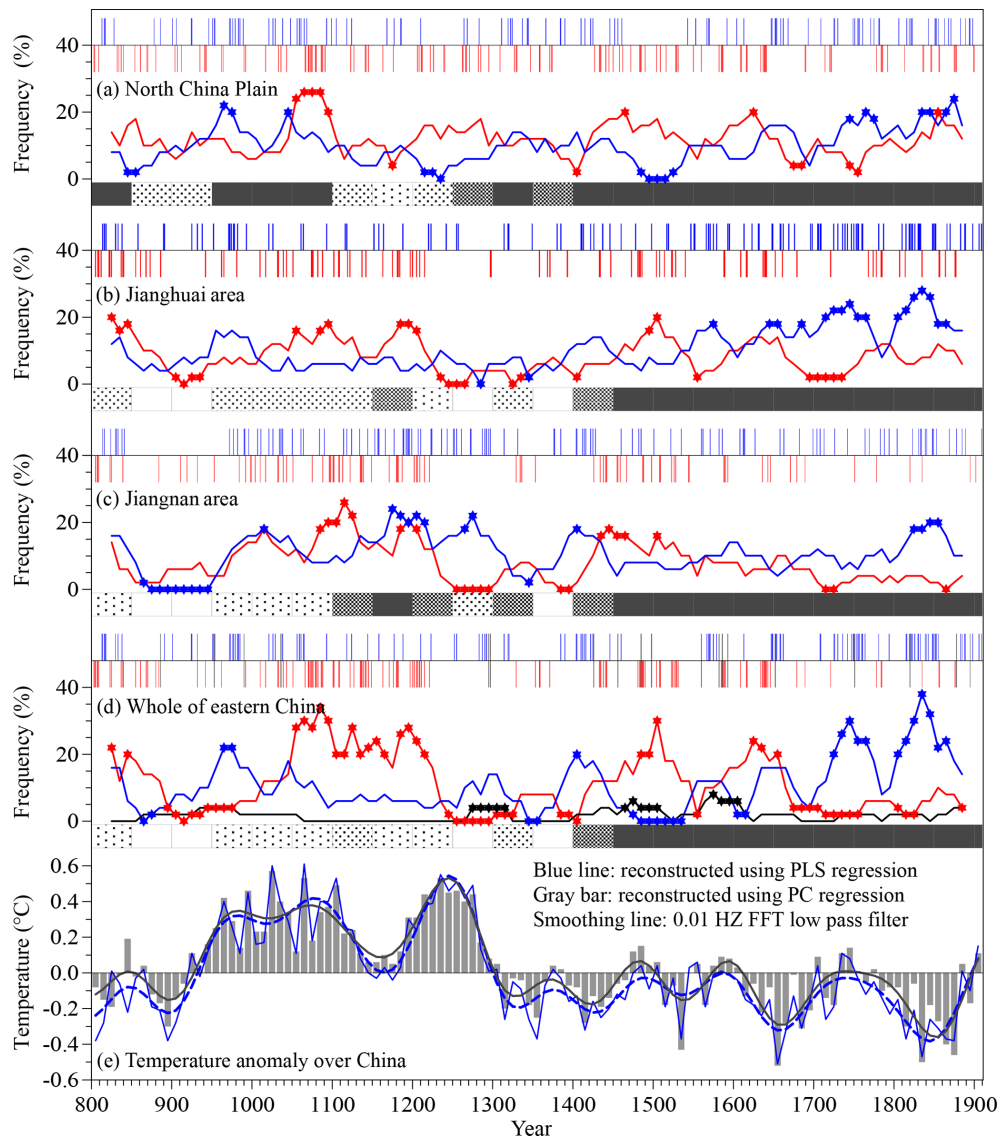
Figure 3. The chronology (vertical line in the top row of each panel) and frequency (star–line, 50-year moving window with a decadal time step) of extreme drought (red) and flood (blue) in the **(a)** North China Plain, **(b)** Jianghuai area, **(c)** Jiangnan area, and **(d)** whole of eastern China from 801 to 1910. The black line in panel **(d)** indicates the same frequency of co-reporting of extreme drought and flood for the whole of eastern China. The stars represent intervals with significantly more or less extreme drought and flood compared with all other intervals, based on the Wilcoxon rank sum test. The bars on the bottom row of each plate illustrate confidence levels (probability of being correct, PBC) for each reconstruction at 50-year intervals. Extremely high confidence (PBC > 99 %): dark; very high confidence (PBC > 90 %): 50 % shaded dark; high confidence (PBC > 80 %): 25 % shaded dark; medium confidence (PBC > 50 %): 12.5 % shaded dark; and low confidence (PBC > 33.3 %): blank. **(e)** The temperature series over China (at a decadal resolution) for comparison, which was reconstructed based on 28 temperature proxies using principal component (PC) regression and partial least squares (PLS) regression (Ge et al., 2013).

Table 2. Number of years and percentage for each harvest grade and missing data, 801–1910.

Harvest grade	1	2	3	4	5	6	Data missing
Number of years	54	178	204	160	77	47	390
% in all years	4.9	16.0	18.4	14.4	6.9	4.2	35.1
% in years excluding grade 4 and missing data	9.6	31.8	36.4	–	13.8	8.4	–
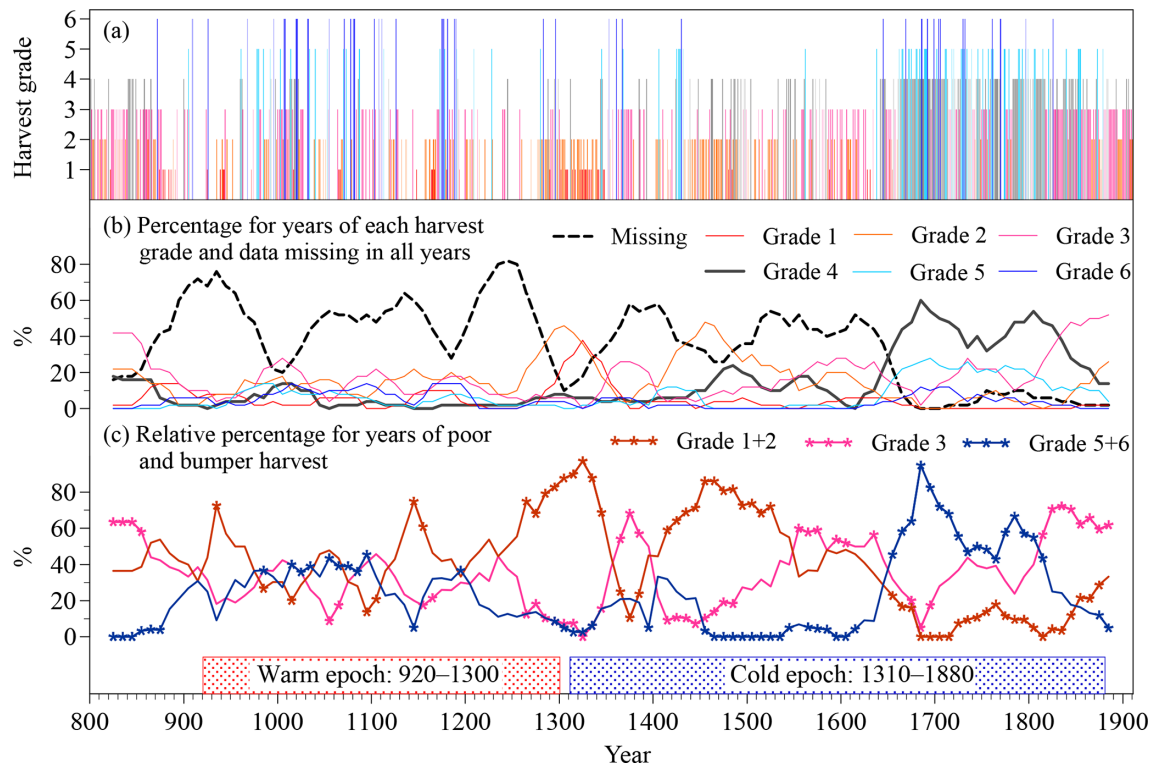
Figure 4. (a) The reconstructed yearly harvest grade from 801–1910 are as follows: 1 – very poor (red), 2 – poor (orange), 3 – slightly poor (pink), 4 – average (black), 5 – near-bumper (light blue), and 6 – bumper (blue). **(b)** The variation in percentage for the years with each harvest grade and missing data, using a moving window of 50 years with a decadal step. **(c)** Relative percentage for years of poor (grade 1 + 2), slightly poor (grade 3), and bumper (grade 5 + 6) harvest compared with total years, excluding average harvests and missing data years, using a moving window of 50 years with a decadal step. Stars: intervals with a significantly higher or lower percentage compared with all other intervals, using the Wilcoxon rank sum test. Boxes: the warm and cold epochs identified from the temperature reconstruction for all of China, shown in Fig. 2e.

9.6 % (grade 1), 31.8 % (grade 2), 36.4 % (grade 3), 13.8 % (grade 5), and 8.4 % (grade 6) (Table 2). Note that in Fig. 4c, grades 1 and 2 (hereinafter "grade 1 + 2") were combined as the poor group, and grades 5 and 6 (hereinafter "grade 5 + 6") were combined as the bumper group, because grade 1 and grade 6 were only 4.9 % and 4.2 %, respectively, in all years. For the entire period, the relative percentages of grade 1 + 2 and grade 5 + 6 were 41.4 % and 22.2 %, respectively.

According to the variation in the frequency of reconstructed poor and bumper harvests (Fig. 4c), the grain harvest for 801–1910 can be roughly divided into five periods. From 801 to 940, the reconstructed grain harvest was generally poor with fewer bumper harvests (only 7.1 % for grade 5 + 6) but more slightly poor harvests (51.4 % for grade 3). From 951 to 1250, there were more reconstructed bumper harvests (31.9 % for grade 5 + 6) with relatively fewer poor harvests (37.7 % for grade 1 + 2; 30.4 % for grade 3), with the exception of 1121–1170, when the frequencies of harvest grades were similar to that of the ninth century. From 1251 to 1650, there were noticeably more reconstructed poor harvests (63.1 % for grade 1 + 2) but very few bumper harvests (9.2 % for grade 5+6). In 1651–1840, there were more recon-

structed bumper harvests (53.5 % for grade 5 + 6) and fewer poor harvests (9.3 % for grade 1 + 2). After 1841, the reconstructed grain harvest became generally poor again with fewer grade 5 + 6 (9.1 %) but more slightly poor (63.6 % for grade 3) harvest grades. On the whole, the reconstructed harvest was generally better in 951–1250, corresponding to a warmer climate, and lower in 1450–1650, corresponding to a colder climate, as argued in Yin et al. (2015, 2016). However, the reconstructed harvests were again higher in 1651–1840, corresponding to a cold climate. This can also be confirmed by other datasets from independent sources. For example, according to harvest reports in the *Archives of the Qing Dynasty*, the mean harvest percentage over eastern China for 1730–1820 was even greater than 70 % (i.e., near-bumper) (Ge and Wang, 1995). In Guangdong Province, southern China, this was over 75 % for 1707–1800 (Marks, 1998).

Table 3. Comparison between mean harvest grades in extreme drought/flood years and nonextreme years for the three subregions and the whole of eastern China.

	Nonextreme	Extreme drought	Extreme flood	Co-reporting[a]
North China Plain	3.31	**2.84*****	3.20	–
Jianghuai area	3.27	**2.99***	3.23	–
Jiangnan area	3.28	**2.89****	3.21	–
Whole of eastern China	3.30	**2.95****	3.27	**2.53****

Note the significance levels: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.05$. [a] The co-reporting of extreme drought and extreme flood in different subregions.

3.2 Association between reported frequency of regional extreme drought/flood and reconstructed harvest grades during 801–1910

Table 3 shows the mean harvest grade for all reconstructed extreme drought/flood years and nonextreme years for the three subregions and the whole of eastern China from 801 to 1910. The mean was 3.3 for all nonextreme years over the whole of eastern China and similar in the three subregions. The two-sample t test showed that the mean harvest grade for all extreme drought years, which was 2.95 in the whole of eastern China, 2.84 for the North China Plain, 2.89 for Jiangnan, and 2.99 for Jianghuai, was significantly lower than for nonextreme years. However, no significant differences were detected for extreme flood years in any subregion. However, the mean harvest grade for years of co-reporting of extreme drought and flood for the whole of eastern China was a mere 2.53 (Table 3), which is much lower than the mean of 3.30 for nonextreme years.

To illustrate the association between reconstructed harvest and reported frequency of regional extreme drought/flood, Table 4 shows the frequency of each harvest grade (i.e., the percentage of years with each harvest grade, accounting for all years without missing data) for extreme drought/flood years and nonextreme years at each subregion and for the whole of eastern China. It was found that the frequency of poor harvests (grade $1 + 2$) for extreme drought years in the North China Plain was 47.4 %, which is significantly larger than 29.2 % for nonextreme years, with frequencies of grade 1 and grade 2 significantly increasing from 6.3 % and 22.9 % to 13.4 % and 34.0 %, respectively. Compared with nonextreme years, although extreme drought in Jianghuai was not accompanied by significant increases in the frequency of grade 1 or grade 2, there was a significant increase in the frequency of grade $1 + 2$ (poor harvest) from 31.5 % to 41.9 %, and a significant decrease in the frequency of grade 5 (near-bumper harvest) from 12.2 % to 5.4 %. For extreme drought years in Jiangnan, the frequency of grade $1 + 2$ (poor harvest) significantly increased from 30.3 % to 51.5 %, with the frequency of grade 2 significantly increasing from 22.9 % to 42.4 % and the frequency of grade 4 (average harvest) significantly decreasing from 23.8 % to 12.1 %. When extreme droughts occurred in the whole of eastern China, it was found

that not only did the frequencies of grade 1 and grade 2 increase from 6.8 % and 22.4 % to 11.7 % and 36.2 %, respectively, but also the frequency of grade 4 significantly decreased from 24.6 % to 13.8 %. Moreover, it was also found that the co-reporting of extreme drought and flood in the whole of eastern China, though only in limited number of years, was associated with the greatest shift in reconstructed harvest grades, in which the frequency of grade 2 significantly increased from 22.4 % to 60.0 %. In summary, reported extreme drought and reconstructed poor harvests (grade $1 + 2$) tend to occur simultaneously, and this pattern could be found in any subregion of eastern China.

However, the association between reconstructed harvest and reported frequency of extreme floods seems to be much weaker. No significant difference was found between the frequency of poor harvests in extreme flood years and nonextreme years, although extreme floods were accompanied by a significantly decreased bumper harvest frequency in some cases. Extreme floods in Jianghuai and Jiangnan, for instance, were both associated with significantly decreased frequency of grade 5 (near-bumper) harvests, from 12.2 % and 11.9 % to 5.7 % and 5.6 %, respectively. It is possible that rainfall in these two humid areas was already sufficient for crop growth; thus, extreme floods had a negative impact on harvest potential (Zheng and Huang, 1998).

3.3 The different patterns in the association between reported frequency of regional extreme drought and reconstructed grain harvest between warm and cold periods

To explore whether the reported frequency of regional extreme drought/flood was differently associated with reconstructed grain harvests in warm and cold periods, a multiple-proxy-based temperature reconstruction with a decadal resolution for all of China is presented in Fig. 3e. It is shown that temperature across China from 801 to 1910 experienced two long-term anomalous epochs, both of which contained several multi-decadal fluctuations. These two anomalous epochs were the warm epoch during 920–1300 and the cold epoch during 1310–1880 (Fig. 4), in which the first epoch covered the MCA (950–1250) and the latter covered the LIA (1450–1850).

Table 4. Comparison of the frequency of each harvest grade between extreme drought/flood years and nonextreme years, 801–1910.

Region	Harvest grade		1	2	3	4	5	6	1 + 2[a]	5 + 6[b]	Total
North China Plain	Nonextreme	Years	34	124	158	129	57	39	158	96	541
		%	6.3	22.9	29.2	23.8	10.5	7.2	29.2	17.7	
	Extreme drought	Years	13	33	22	17	9	3	46	12	97
		%	**13.4****	**34.0****	22.7	17.5	9.3	3.1	**47.4*****	12.4	
	Extreme flood	Years	7	21	24	14	11	5	28	16	82
		%	8.5	25.6	29.3	17.1	13.4	6.1	34.1	19.5	
Jianghuai area	Nonextreme	Years	39	137	153	127	68	34	176	102	558
		%	7	24.6	27.4	22.8	12.2	6.1	31.5	18.3	
	Extreme drought	Years	8	23	21	12	4	6	31	10	74
		%	10.8	31.1	28.4	16.2	*5.4* *	8.1	**41.9***	13.5	
	Extreme flood	Years	7	18	30	21	5	7	25	12	88
		%	8	20.5	34.1	23.9	*5.7* *	8	28.4	13.6	
Jiangnan area	Nonextreme	Years	42	129	158	134	67	34	171	101	564
		%	7.4	22.9	28	23.8	11.9	6	30.3	17.9	
	Extreme drought	Years	6	28	14	8	5	5	34	10	66
		%	9.1	**42.4*****	21.2	*12.1* **	7.6	7.6	**51.5*****	15.2	
	Extreme flood	Years	6	21	32	18	5	8	27	13	90
		%	6.7	23.3	35.6	20	*5.6* *	8.9	30	14.4	
Whole of eastern China	Nonextreme	Years	35	115	146	126	59	32	150	91	513
		%	6.8	22.4	28.5	24.6	11.5	6.2	29.2	17.7	
	Extreme drought	Years	11	34	21	13	7	8	45	15	94
		%	**11.7***	**36.2*****	22.3	*13.8* **	7.4	8.5	**47.9*****	16	
	Extreme flood	Years	7	20	35	19	10	7	27	17	98
		%	7.1	20.4	35.7	19.4	10.2	7.1	27.6	17.3	
	Co-occurrence[c]	Years	1	9	2	2	1	0	10	1	15
		%	6.7	**60.0*****	13.3	13.3	6.7	0	**66.7*****	6.7	

Values in bold with stars denote significance, more or less (also in italic), of reported extremes, using a chi-square test (χ^2), at the following levels: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. [a] For grade 1 and grade 2 in total. [b] For grade 5 and grade 6 in total. [c] For the co-reporting of extreme drought and extreme flood in different subregions.

Table 5. Comparison of frequency of reported extreme drought/flood between warm and cold epochs.

Subregion		Frequency (%)	
		Warm epoch 920–1300	Cold epoch 1310–1880
North China Plain	Extreme drought	12.6	11.7
	Extreme flood	9.4	11.2
Jianghuai area	Extreme drought	9.2	7.9
	Extreme flood	**7.1*****	**13.3*****
Jiangnan area	Extreme drought	**10.8****	**6.5****
	Extreme flood	**13.6***	**9.8***
whole of eastern China	Extreme drought	**13.9****	**9.6****
	Extreme flood	9.7	11.9

Values in bold with stars denote significantly more or less (also in italic) frequent reported extreme droughts/floods in 920–1300 compared with 1310–1880, using a rank sum test at the following levels: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

The comparison of the reported frequencies of extreme drought/flood between the warm (920–1300) and cold (1310–1880) epochs (Table 5) shows only a small difference in the reported frequency of extreme droughts between warm and cold epochs for the North China Plain and Jianghuai. However, in Jiangnan, reconstructed extreme drought years were more frequent in 920–1300 (10.8 %) than in 1310–1880 (6.5 %). Therefore, the reported extreme droughts for the whole of eastern China in 920–1300 (13.9 %) were significantly more frequent than in 1310–1880 (9.6 %). As for the reported frequency of extreme floods, significant differences were only found in the south. Compared with 1310–1880, there were significantly fewer extreme floods in Jianghuai but more from 920 to 1300 in Jiangnan. After the composition of their opposite tendency, no significant differences in the frequency of extreme floods between 920–1300 and 1310–1880 were found for the whole of eastern China. Thus, there were slightly more reported extreme droughts during the warm period than the cold period, while no differences were found in the frequency of reported extreme floods over eastern China.

As found in Sect. 3.2, higher frequency of reporting of extreme droughts in eastern China was significantly associated with reconstructed poor harvests (grade 1 + 2) when compared with nonextreme years. Since there were more extreme droughts over eastern China in 920–1300 than in 1310–1880 in the reconstructed drought/flood dataset, the reconstructed harvest in the warm epoch could be expected to be worse than in the cold epoch. However, as Yin et al. (2015, 2016) found, the reconstructed harvest in a warm epoch was better than that in a cold epoch. This suggests that the patterns in the association between reported frequency of regional extreme drought and reconstructed grain harvest differed between warm and cold epochs. Table 6 illustrates this difference, by showing the frequency of poor harvests (grade 1 + 2) in extreme drought/flood years and nonextreme years for warm and cold epochs, i.e., 920–1300 and 1310–1880. Note that there exists a shift in the distribution of each harvest grade after the 1640s due to an increase in grade 4 years and a decrease in missing data years. Meanwhile, there existed an evidently high frequency of bumper harvests (grade 5 + 6) from the 1650s to the 1810s. Therefore, the patterns in the association between reconstructed poor harvests (grade 1 + 2) and reported frequency of regional extreme droughts for 1650–1880 were also investigated.

The results show that during the warm epoch of 920–1300 the reconstructed poor harvest was not significantly associated with regional reported frequency of extreme droughts, although the frequency of poor harvest in extreme drought years was slightly higher than in nonextreme years for each subregion. In contrast, during the cold epoch of 1310–1880, the frequency of poor harvest in extreme drought years was significantly higher than in nonextreme years, which indicates a significant association in reconstructed poor harvest and reported frequency of extreme droughts. Moreover, sim-ilar characteristics were found for the latter half of the cold period from 1650 to 1880, which indicates that the shift of the harvest grade distribution did not affect the association between reconstructed poor harvest and reported frequency of extreme droughts/floods during the cold epoch.

3.4 Discussion

Chinese historical documents provide a unique proxy for reconstructions of climate change and annual harvest grades, which helps to explore the impacts of climate change on grain yields up to 1000 years ago. Compared with previous research (e.g., Su et al., 2014; Yin et al., 2015, 2016), which focused on the correlation between reconstructed temperature variation and grain harvest fluctuations at decadal resolution during historical times, this study presented an examination of the association between reported frequency of extreme drought/flood and reconstructed grain harvests and its different pattern between warm and cold periods year by year. The statistical results from Sect. 3.2 indicate that the reported frequency of regional extreme droughts is closely associated with poor harvests in the reconstruction, while no significant pattern is found between the reported frequency of extreme floods and reconstructed poor harvests. This phenomenon might be explained by the fact that extreme droughts usually cover an immense area dominated by a single large-scale air mass, leading to significant and extensive impacts on agriculture, economics, and many other social factors. On the other hand, extreme floods were mostly caused by rainstorms induced by the confrontation of air masses, which usually occur across a relatively narrow belt. Meanwhile, rainstorms could irrigate agricultural land in areas surrounding the extreme floods and, thus, improve grain yields, leading to limited impacts of extreme floods on harvests over an immense area (Zhang, 1982). It should be noted that these meteorological interpretations only apply for short-term extreme flood events. When it comes to long-term wet periods, it could as well be responsible for bad harvest, due to the low temperature and short sunlight.

Section 3.3 shows the statistical results for the different patterns in the association between reported frequency of regional extreme droughts and reconstructed grain harvests between warm and cold periods. These results suggest that the simultaneous occurrence of reported extreme droughts and reconstructed poor harvests tends to be weakened in the case of a warm climate. Possible causes might be that there was low and limited adaptability during that period, and the warm climate provided more thermal resources and extended the growing season, thus increasing the multiple cropping index and providing more thermal-limited lands for growing crops. This gave people more options to adapt to climatic variation and mitigated the impacts of extreme droughts on harvest yield. As assessed by Zhang (1982), the harvest may change by approximately 10 % if the temperature changed by 1 °C on a national scale, based on the data from 1909 to

Table 6. Comparison of frequency of reconstructed poor harvest (grade $1+2$) for extreme drought years and nonextreme years for warm and cold epochs.

Region	Anomalous temperature epochs		Warm epoch 920–1300		Cold epoch 1310–1880		1650–1880	
	Harvest grade		$1+2$	others	$1+2$	others	$1+2$	others
North China Plain	Nonextreme	Years	53	90	78	233	6	159
		%	37.1	62.9	25.1	74.9	3.6	96.4
	Extreme drought	Years	13	13	24	31	5	17
		%	50	50	**43.6*****	56.4	**22.7*****	77.3
Jianghuai area	Nonextreme	Years	57	90	86	237	10	154
		%	38.8	61.2	26.6	73.7	6.1	93.9
	Extreme drought	Years	12	12	16	20	4	12
		%	50	50	**44.4****	55.6	**25.0*****	75
Jiangnan area	Nonextreme	Years	49	79	86	254	15	173
		%	38.3	61.7	25.3	74.7	8	92
	Extreme drought	Years	13	14	19	11	1	5
		%	48.2	51.8	**63.3*****	36.7	16.7	83.3
Whole eastern China	Nonextreme	Years	48	79	76	229	8	156
		%	37.8	62.2	24.9	75.1	4.9	95.4
	Extreme drought	Years	15	20	23	20	3	8
		%	42.9	57.1	**53.5*****	46.5	**27.3****	72.7

Values in bold with stars denote significantly more reported droughts/floods, using a chi-square test (χ^2) at the following levels: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

1979, in which the harvest increased significantly in 7 out of 8 warm years. However, a cold climate could limit multiple cropping and shrink the area of arable land, leading to the harvest becoming more vulnerable to extreme drought. Moreover, as reported by Zhang et al. (2007), limited resources could also cause social turbulence, such as famine, peasant uprising, the outbreak of war, and population decline, all of which may further increase agricultural vulnerability. Therefore, even though the frequency of reporting of extreme droughts was slightly higher in the warm period of 920–1300, the frequency of reconstructed poor harvests did not increase significantly.

However, these inferences are purely based on those two reconstructed datasets and are insufficient to reveal actual historical connections. One of the reasons is that both datasets used *Twenty-Four Histories* and *Qing History Draft* as their record resources in the reconstruction, which might induce an artifact into the databases and lead to spurious correlations between extreme drought/flood and harvest. Another reason is that there existed major shifts in many social factors (such as political, economic, demographic, or cultural change) in the study period, which means that the different patterns in the association between reported extreme drought and reconstructed poor harvests between warm and cold periods might be cocreated by those confounding factors along with climate factors. Although the datasets for extreme

droughts/floods used more historical documents and were not limited to the *Twenty-Four Histories* and *Qing History Draft* as resources, and both datasets have been validated by other proxy data in previous studies, the independence and consistency problems could not be eliminated entirely. The solution to these two problems requires further research using more independent datasets from multiple proxies and many other study fields, and it could be of great importance for improving the understanding of climatic impacts on agriculture and adaptation to future global warming along with a higher extreme climate probability.

Meanwhile, several issues still exist that could lead to uncertainty in the statistics. First, due to missing records before 1400, the reconstructed chronology of regional extreme drought and flood may not include all events. For example, as shown in Fig. 3, the reconstructed chronologies of regional extreme drought and flood for several periods (e.g., 850–950, 1350–1400 in both Jianghuai and Jiangnan, and 1250–1300 in Jianghuai) had low confidence levels. Moreover, the chronologies for 1150–1200 in the North China Plain, 1200–1250 in Jianghuai, and 800–850 and 950–1100 in Jiangnan also had medium confidence levels. Therefore, most of the reconstructed chronology of extreme drought and flood for the whole of eastern China before 1400 was only in medium or low confidence. Second, there may also be limitations in the grading of annual harvests for some years. For example,

in 1181, the harvest was assessed as grade 6, according to the description, "It is a plentiful year after sufficient rainfall killed locusts" in the *History of Song* (one of the *Twenty-Four Histories*), which usually recorded countrywide events. Thus, this description was regarded as a record at the national scale. However, *Compilation of Essential Regulations of the Song Dynasty* (*Song Hui-Yao Ji-Gao* in Chinese; edited by Xu Song, Vol. 66, Vol. 69) reported widespread drought in southern China that year, which led to poor or slightly poor harvests at the subregional scale. Similar inconsistencies in the historical records have also been found in 1032, 1073, 1353, and 1638. These cases suggest that the grain harvest descriptions recorded in *Twenty-Four Histories* may not indicate anomalous harvests for all years at the national scale. In addition, around the 1650–60s, there was a clear jump to grade $5 + 6$ (bumper harvest), yet this period is commonly recognized in previous literature as having poor harvests and famine, as it is in the coldest interval of the Little Ice Age. Also, the periods 1130–1150 and 1210–1270 (the early and later Southern Song dynasty), 880–980 (late Tang dynasty to late Ten Kingdoms), and around 1400 have remarkably more missing harvest grade data. These limitations introduce uncertainty by affecting the statistical relationship between harvest and extreme drought/flood in certain periods. Therefore, further research is required, especially in the search for more historical records from the other scattered documents to supplement the evidence recorded in *Twenty-Four Histories*.

4   Conclusions

Variation in the reported frequency of regional extreme drought/flood and reconstructed grain harvest, and the possible association between the two, were investigated for 801–1910, using a reconstructed chronology of regional extreme drought/flood and a series of grain harvest grades. The results reveal significantly higher reported frequency of extreme droughts in 801–870, 1031–1230, 1481–1530, and 1581–1650 and higher reported frequency of extreme floods in 811–840, 951–990, 1051–1070, 1381–1430, 1641–1670, 1721–1770, and 1810–1870 over the whole of eastern China. Regionally, more extreme droughts were reported in different intervals, including 1031–1100, 1441–1490, 1601–1650, and 1831–1880 for the North China Plain, 801–870, 1031–1120, 1161–1220, and 1471–1530 for Jianghuai, and 991–1040, 1091–1150, 1171–1230, 1411–1470, and 1481–1530 for Jiangnan. The grain harvest was reconstructed to be generally poor in, approximately, 801–940, 1251–1650, and 1841–1910 but bumper in 951–1250 and 1651–1840. Both the *t* test and chi-square test (χ^2) demonstrated that a higher frequency of reporting of extreme droughts in any subregion of eastern China is significantly associated with reconstructed poor harvests. Meanwhile, the harvest was reconstructed to be much worse when extreme drought and extreme flood were reported in the same year in different sub-

regions of eastern China. However, the association between reported frequency of extreme flood and reconstructed harvest yield was not detected in the long-term average. Moreover, the comparison showed the reported frequency of extreme droughts over the whole of eastern China was higher in the warm epoch of 920–1300 than that in the cold epoch of 1310–1380. However, during the warm epoch, the association between reconstructed poor harvests and reported frequency of regional extreme droughts was weak, though the frequency of poor harvests in extreme drought years was still slightly higher than in nonextreme years for each subregion. This could be explained by the provision of more thermal resources and the extended growing season in the warm climate, which increased the multiple cropping index and created more thermal-limited lands for agricultural crops, allowing for more options to mitigate the impacts of extreme drought on the harvest. Thus, the study provides further historical evidence to help explore the implications for agricultural adaptation to future global warming along with a higher probability of extreme climate.

References

Academy of Meteorological Science of China Central Meteorological Administration: Yearly Charts of Dryness/Wetness in China for the Last 500 years, Cartographic Publishing House, Beijing, China, 332 pp., 1981.

Dai, A.: Drought under global warming: A review, Wires Clim. Change, 2, 45–65, https://doi.org/10.1002/wcc.81, 2011.

Dai, A.: Increasing drought under global warming in observations and models, Nat. Clim. Change, 3, 52–58, https://doi.org/10.1038/nclimate1633, 2013.

Deng, Y., Gou, X., Gao, L., Zhang, F., Xu, X., and Yang, M.: Tree-ring recorded drought variability in the northern Min Mountains of northwestern China, Int. J. Climatol., 36, 3550–3560, https://doi.org/10.1002/joc.4575, 2016.

Ding, Y., Wang, S., Zheng, J., Wang, H., and Yang, X.: Climate in China, Science Press, Beijing, China, 558 pp., 2013 (in Chinese).

Ge, Q. and Wang, W.-C.: Population Pressure, climate change and Taiping Rebellion, Geogr. Res., 14, 32–42, 1995 (in Chinese).

Ge, Q., Hao, Z., Zheng, J., and Shao, X.: Temperature changes over the past 2000 yr in China and comparison with the Northern Hemisphere, Clim. Past, 9, 1153–1160, https://doi.org/10.5194/cp-9-1153-2013, 2013.

Gou, X., Gao, L., Deng, Y., Chen, F., Yang, M., and Still, C.: An 850-year tree-ring-based reconstruction of drought history in the western Qilian Mountains of northwestern China, Int. J. Climatol., 35, 3308–3319, https://doi.org/10.1002/joc.4208, 2015a.

Gou, X., Deng, Y., Gao, L., Chen, F., Cook, E., Yang, M., and Zhang, F.: Millennium tree-ring reconstruction of drought variability in the eastern Qilian Mountains, northwest China, Clim. Dynam., 45, 1761–1770, https://doi.org/10.1007/s00382-014-2431-y, 2015b.

Hao, Z., Ge, Q., and Zheng, J.: Variations of extreme drought/flood events over eastern China during the past 2000 years, Climatic and Environmental Research, 15, 388–394, 2010a (in Chinese).

Hao, Z., Zheng, J., Wu, G., Zhang, X., and Ge, Q.: 1876–1878 severe drought in North China: facts, impacts and climatic background, Chinese Sci. Bull., 55, 3001–3007, https://doi.org/10.1007/s11434-010-3243-z, 2010b.

Hao, Z., Zheng, J., Zhang, X., Liu, H., Li, M., and Ge, Q.: Spatial patterns of precipitation anomalies in eastern China during centennial cold and warm periods of the past 2000 years, Int. J. Climatol., 36, 467–475, https://doi.org/10.1002/joc.4367, 2016.

IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M., Cam-

bridge University Press, Cambridge, UK and New York, USA, 582 pp., 2012.

IPCC: 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, UK and New York, USA, 1535 pp., 2013.

Li, J., Shao, X., Qin, N., and Li, Y.: Runoff variations at the source of the Yangtze River over the past 639 years based on tree-ring data, Clim. Res., 75, 131–142, https://doi.org/10.3354/cr01510, 2018.

Liang, E., Liu, X., Yuan, Y., Qin, N., Fang, X., Huang, L., Zhu, H., Wang, L., and Shao, X.: The 1920S drought recorded by tree rings and historical documents in the semi-arid and arid areas of Northern China, Climatic Change, 79, 403–432, https://doi.org/10.1007/s10584-006-9082-x, 2006.

Liu, Y., Zhang, X., Song, H., Cai, Q., Li, Q., Zhao, B., Liu, H., and Mei, R.: Tree-ring-width-based PDSI reconstruction for central Inner Mongolia, China over the past 333 years, Clim. Dynam., 48, 867–879, https://doi.org/10.1007/s00382-016-3115-6, 2017.

Marks, R. B.: It never used to snow: climatic variability and harvest yields in late-imperial south China, 1650–1680, in: Sediments of Time, Part 1: Environment and Society in Chinese History, edited by: Mark, E. and Liu, T. J., Cambridge and New York, Cambridge University Press, Cambridge, UK and New York, USA, 411–446, 1998.

Qin, C., Yang, B., Brauning, A., Griessginger, J., and Wernicke, J.: Drought signals in tree-ring stable oxygen isotope series of Qilian juniper from the arid northeastern Tibetan Plateau, Global Planet. Change, 125, 48–59, https://doi.org/10.1016/j.gloplacha.2014.12.002, 2015.

Qin, D., Zhang J., Shan C., and Song L.: China national assessment report on risk management and adaptation of climate extremes and disasters (Refined edition), Science Press, Beijing, China, 124 pp., 2015.

Shen, C., Wang, W.-C., Hao, Z., and Gong, W.: Exceptional drought events over eastern China during the last five centuries, Climatic Change, 85, 453–471, https://doi.org/10.1007/s10584-007-9283-y, 2007.

Shen, C., Wang, W.-C., Hao, Z., and Gong, W.: Characteristics of anomalous precipitation events over eastern China during the past five centuries, Clim. Dynam., 31, 463–476, https://doi.org/10.1007/s00382-007-0323-0, 2008.

Su, Y., Fang, X., and Yin, J.: Impact of climate change on fluctuations of grain harvests in China from the Western Han Dynasty to the Five Dynasties (206 BC–960 AD), Sci. China Earth Sci., 57, 1701–1712, https://doi.org/10.1007/s11430-013-4795-y, 2014.

Xiao, D., Shao, X., Qin, N., and Huang, X.: Tree-ring-based reconstruction of streamflow for the Zaqu River in the Lancang River source region, China, over the past 419 years, Int. J. Biochem., 61, 1173–1189, https://doi.org/10.1007/s00484-016-1297-6, 2017.

Yang, B., Kang, S., Ljungqvist, F. C., He, M., Zhao, Y., and Qin, C.: Drought variability at the northern fringe of the Asian summer monsoon region over the past millennia, Clim. Dynam., 43, 845–859, https://doi.org/10.1007/s00382-013-1962-y, 2014.

Yin, J., Su, Y., and Fang, X.: Relationships between temperature change and grain harvest fluctuations in China from 210 BC to 1910 AD, Quatern. Int., 355, 153–163, https://doi.org/10.1016/j.quaint.2014.09.037, 2015.

Yin, J., Fang, X., and Su, Y.: Correlation between climate and grain harvest fluctuations and the dynastic transitions and prosperity in China over the past two millennia, Holocene, 26, 1914–1923, https://doi.org/10.1177/0959683616646186, 2016.

Yu, M., Li, Q., Michael, J. H., Mark, D. S., and Richard, R. H.: Are droughts becoming more frequent or severe in China based on the Standardized Precipitation Evapotranspiration Index: 1951–2010?, Int. J. Climatol., 34, 545–558, https://doi.org/10.1002/joc.3701, 2014.

Zhai, J., Su, B., Gao, C., and Jiang, T.: Spatial variation and trends in PDSI and SPI indices and their relation to streamflow in 10 large regions of China, J. Climate, 23, 649–663, https://doi.org/10.1175/2009JCLI2968.1, 2010.

Zhang, D.: 1784–1787 drought occurrence over east China in a warm climatic background, Acta Geographica Sinica, 55, 106–112, 2000 (in Chinese).

Zhang, D.: Severe drought events as revealed in the climate records of China and their temperature situations over the last 1000 years, Acta Meteorol. Sin., 19, 485–491, 2005.

Zhang, D., Liu, C., and Jiang, J.: Reconstruction of six regional dry/wet series and their abrupt changes during the last 1000 year in east China, Quat. Sci., 17, 1–11, 1997 (in Chinese).

Zhang, D., Brecke, P, Lee, H., He, Y., and Zhang, J.: Global climate change, war, and population decline in recent human history, P. Natl. Acad. Sci. USA, 104, 19214–19219, https://doi.org/10.1073/pnas.0703073104, 2007.

Zhang, J.: Possible impacts of climatic variation on agriculture in China, Geogr. Res., 1, 8–15, 1982 (in Chinese).

Zhang, P. (Ed.): Historical climate change in China, Shandong Science and Technology Press, Jinan, Shandong, China, 1996 (in Chinese).

Zhang, Y., Shao, X., Yin, Z., Liang, E., Tian, Q., and Xu, Y.: Characteristics of extreme droughts inferred from tree-ring data in the Qilian Mountains, 1700–2005, Clim. Res., 50, 141–159, https://doi.org/10.3354/cr01051, 2011.

Zheng, J. and Huang, J.: An estimation of grain loss caused by natural disasters in China, 1950–1990, Acta Geogr. Sin., 53, 501–510, 1998 (in Chinese).

Zheng, J., Wang, W.-C., Ge, Q., Man, Z., and Zhang, P.: Precipitation variability and extreme events in eastern China during the past 1500 years, Terr. Atmos. Ocean. Sci., 17, 579–592, 2006.

Zheng, J., Xiao, L., Fang, X., Hao, Z., Ge, Q., and Li, B.: How climate change impacted the collapse of the Ming dynasty, Climatic Change, 127, 169–182, https://doi.org/10.1007/s10584-014-1244-7, 2014a.

Zheng, J., Hao, Z., Fang, X., and Ge, Q.: Changing characteristics of extreme climate events during past 2000 years in China, Prog. Geogr., 33, 3–12, 2014b.

Zheng, J., Ge, Q., Hao, Z., Liu, H., Man, Z., Hou, Y., and Fang, X., Paleoclimatology proxy recorded in historical documents and method for reconstruction on climate change, Quat. Sci., 34, 1186–1196, 2014c.

Zou, X., Zhai, P., and Zhang, Q.: Variations in drought over China: 1951–2003, Geophys. Res. Lett., 32, L04707, https://doi.org/10.1029/2004GL021853, 2005.

Zou, X., Ren, G., and Zhang, Q.: Droughts Variations in China Based on a Compound Index of Meteorological Drought, Climatic and Environmental Research, 15, 371–378, 2010 (in Chinese).