Dear editor:

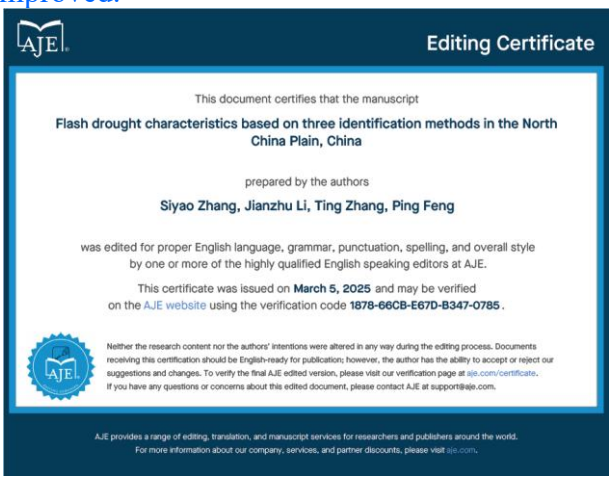
On behalf of all the contributing authors, I would like to express our sincere appreciations of your letter and the constructive comments from Referee #2 concerning our article entitled "Flash drought characteristics based on three identification methods in the North China Plain, China". All the comments are very helpful for revising our paper. We have studied and discussed all the comments point-by-point carefully, and accordingly made substantial revisions to our paper. All the changes we have made were in the red-colored text. Our point-by-point responses to all the comments are provided below in the blue-colored texts.

****

Zhang et al, have studied flash drought characteristics in China based on three indicators. This manuscript does not meet the requirements for publication in the HESS journal.

1. First of all, the manuscript needs significant improvements in terms of language, there are so many grammatical errors, long sentences that make it hard to follow the text, and so many abbreviations even in the Abstract that confuse the reader.

Response: Thank you for the comments. We apologize for the poor language of our manuscript. We have corrected the grammar errors, revised the long sentences into short sentences, and made efforts to reduce the abbreviations in the Abstract. We have now worked on both language and readability and have also involved native English speakers for language corrections. We really hope that the flow and language level have been substantially improved.



AJE

Editing Certificate

This document certifies that the manuscript

Flash drought characteristics based on three identification methods in the North China Plain, China

prepared by the authors

Siyao Zhang, Jianzhu Li, Ting Zhang, Ping Feng

was edited for proper English language, grammar, punctuation, spelling, and overall style by one or more of the highly qualified English speaking editors at AJE.

This certificate was issued on **March 5, 2025** and may be verified on the AJE website using the verification code **1878-66CB-E67D-B347-0785**.

Neither the research content nor the authors' intentions were altered in any way during the editing process. Documents receiving this certification should be English-ready for publication; however, the author has the ability to accept or reject our suggestions and changes. To verify the final AJE edited version, please visit our verification page at aje.com/certificate. If you have any questions or concerns about this edited document, please contact AJE at support@aje.com.

AJE provides a range of editing, translation, and manuscript services for researchers and publishers around the world. For more information about our company, services, and partner discounts, please visit aje.com.

2. In the Introduction, the literature review merely mentions other studies using different drought indicators (~3 paragraphs and introducing a new abbreviation in every line) without highlighting the knowledge gap and the challenges that need to be addressed.

Response: Thank you for the comments. The literature review in the Introduction introduces many drought indicators but lacks focus and knowledge gaps. We have revised the writing logic of the Introduction. Firstly, it pointed out the serious harm of FD and the widespread global concern. Then, the application of FD identification indicators, which are developed by conventional drought indicators, soil moisture, and atmospheric evaporation demand, was introduced. The advantages and disadvantages of these FD identification indicators were pointed out, emphasizing two widely applied

indicators, RZSM and SESR. After that, previous applications of RZSM and SESR for FD identification were discussed, especially their findings on FD characteristics. It was also pointed out that FDs based on RZSM and SESR are from agricultural and meteorological perspectives, respectively. The significance of exploring FD from different perspectives was highlighted. Finally, limitations of FD identification by SESR were displayed. In the last paragraph of Introduction, improvement methods to address those limitations were proposed, and the objectives of this paper were stated (see Introduction in lines 25-109).

*"**1 Introduction***

The terrestrial water cycle has accelerated and droughts have become more frequent under global warming. Persistent and slow drought disasters, caused by long-term inadequate precipitation, can severely harm ecology, agriculture, and economy (Li et al., 2020; Limones, 2021). In addition to these long-term and slow droughts, some droughts characterized by rapid occurrence and intensification, known as flash droughts (FDs), also occur frequently (Deng et al., 2022; Yuan et al., 2023). FDs are driven by both water and energy limitations, and their rapid onset poses significant challenges for drought monitoring and forecasting (Yuan et al., 2023; Zhang et al., 2022a). A severe FD occurred in the United States during the summer of 2012. The conditions rapidly shifted from no drought to extreme drought within less than a month. Its rapid intensification made prediction difficult, resulting in economic losses exceeding $30 billion (Yuan et al., 2023). Other notable FD events included those in western Russia in 2010, southern Great Plains in 2015, and southern China in 2019 (Edris et al., 2023; Hunt et al., 2021; Wang and Yuan, 2021). Because of the severe societal and environmental impacts of FD, its occurrence and development have received significant attention worldwide (Deng et al., 2022; Li et al., 2020).

FDs frequently occur alongside increasing temperature and decreasing precipitation, resulting in a rapid increase in evapotranspiration and a reduction in soil moisture (Tyagi et al., 2022). Consequently, some hydrological and meteorological variables have been applied as key indicators to identify FD in previous studies. Currently, FD identification methods can be categorized into three main classifications on the basis of conventional drought indicators, soil moisture, and atmospheric evaporation demand.

The conventional drought indicators employed for identifying FDs include United States Drought Monitor (USDM) (Chen et al., 2019; Edris et al., 2023; Otkin et al., 2019; Pendergrass et al., 2020), standardized precipitation evapotranspiration index (SPEI) (Fu and Wang, 2022; Noguera et al., 2020; Noguera et al., 2021), and standardized precipitation index (SPI) (Noguera et al., 2021; Parker et al., 2021). However, these conventional drought indices have drawbacks. For example, USDM provides weekly updates on drought location and intensity but is applicable only within the United States. SPI focuses only on the precipitation deficit and neglects the influences of temperature and evaporation. SPEI integrates both insufficient precipitation and increased evapotranspiration, yet its response to climatic changes is slow (Deng et al., 2022). Consequently, several specific indicators for FD identification that respond quickly to climate change and are applicable in broader regions have also been developed using soil moisture and climate factors.

Soil moisture reflects the combined influence of precipitation and evapotranspiration and is directly related to agricultural production. Therefore, soil moisture-based indicators have been widely employed in agricultural flash drought identification. These indices include root zone soil moisture (RZSM, Yuan et al., 2019), soil moisture

index (SMI, Hunt et al., 2009, 2021), soil moisture volatility index (SMVI, Osman et al., 2021; Osman et al., 2022), and flash drought stress index (FDSI, Sehgal et al., 2021). SMI uses wilting and field capacity soil metrics to analyze the changes in soil moisture stress. However, these metrics are influenced by soil types and can be challenging to estimate (Garg et al., 2017; Otkin et al., 2018; Rab et al., 2011; Sehgal et al., 2021). SMVI must be reset by precipitation and is sensitive to interruptions associated with drought onset (Osman et al., 2021). Due to the limited research and application of SMVI, its effectiveness and universality need further verification. In addition, FDSI combines soil moisture and climate factors, which can comprehensively reflect FD stress. However, the diverse required data and complex calculations limit its widespread application (Sehgal et al., 2021). RZSM can directly monitor the rapid decreases in soil moisture over a short period and is closely related to soil moisture stress (Yuan et al., 2019). Owing to its clear physical mechanism, simple data requirements, and ease of calculation, RZSM is widely used worldwide.

In addition to soil moisture, various indicators have been employed to identify FDs on the basis of the atmospheric evaporation demand. The evaporative stress index (ESI) is calculated as 1 minus the ratio of actual evapotranspiration (ET) to potential evapotranspiration (PET) (Ahmad et al., 2022; Anderson et al., 2007; Hunt et al., 2021; Otkin et al.,2019). It is sensitive to energy- and water-limited situations, although the spatial resolution is typically 5 ~ 10 km (Anderson et al., 2007). Standardized evapotranspiration deficit index (SEDI; Vicente-Serrano et al., 2010) evaluates both heatwave- and water deficit-driven FDs by measuring the difference between ET and PET. However, SEDI exhibits significant uncertainty in humid and sub-humid regions (Li et al., 2020; Rakkasagi et al., 2023). The evaporative demand drought index (EDDI; Hobbins et al., 2016) detects FDs through the PET response to surface dryness anomalies. While it provides weekly water stress, it can be utilized only in the contiguous United States (Rakkasagi et al., 2023). Standardized evaporative stress ratio (SESR; Christian et al., 2019) incorporates near-surface state variables and plant health, making it sensitive to rapid changes in evaporative stress and capable of capturing early FD signals (Gou et al., 2022). Identifying FDs by SESR is based on the objective percentiles on a grid, making it globally applicable (Christian et al., 2019).

Therefore, RZSM and SESR are commonly utilized as indicators for identifying agricultural FDs and meteorological FDs in regional studies. For example, Yuan et al. (2019) characterized FD frequency and duration using RZSM and projected FD severity and risk exposure. They reported that southern China experiences more frequent, severe, and higher-risk FDs. Zhang et al. (2022a) assessed FDs using RZSM and simulated FD intensity by three machine learning models, revealing greater FD intensity in southeastern China. Zhong et al. (2022) evaluated FD spatial and temporal characteristics, such as duration, severity, and intensity, via SESR and identified increasing FD risk from west to east in the Pearl River Basin, China. Gou et al. (2022) illustrated the spatial distribution of FD frequency and duration using SESR and tracked FD spatial paths in the Huaibei Plain of China from 2001 to 2019. Deng et al. (2022) used SESR to determine global FD spatiotemporal characteristics and found a significant decrease in FD coverage from 1981 to 2020. Mukherjee and Mishra (2022b) investigated global FD frequency and intensity using SESR and RZSM, highlighting the influence of climate factors and background aridity on FD onset and evolution. Therefore, RZSM and SESR are widely applied to identify regional FD (i.e., FD_{RZSM} and FD_{SESR}). FD_{SESR}, identified by anomalous evaporative stress, is classified as a meteorological drought. However, FD_{RZSM}, determined by soil moisture, is categorized as an agricultural drought. Although they are interrelated, they identify FDs from

different perspectives. Exploring FD characteristics from multiple perspectives plays an important role in the comprehensive understanding and effective response to FD events. Therefore, this study employs RZSM and SESR to further analyze regional FD characteristics and investigate the effects of climate factors and background aridity.

However, the SESR application has several problems. When a FD identified via SESR, both SESR and the change in SESR (ΔSESR) are assumed to follow normal distributions. Therefore, the 50th percentiles of ΔSESR ($\Delta SESR_{50th}$) equal 0, indicating no change in SESR (ΔSESR = 0). A ΔSESR below the 40th percentile of ΔSESR ($\Delta SESR_{40th}$) denotes a decrease in the SESR. However, whether SESR and ΔSESR follow normal distributions remains uncertain. If they do not, $\Delta SESR_{40th}$ may not be less than 0. In the study by Gou et al. (2022), the 36th percentile of ΔSESR ($\Delta SESR_{36th}$) corresponded to an increase in SESR, where $\Delta SESR_{36th}$ was greater than 0. This phenomenon might be because SESR gradually decreases during dry periods and increases during precipitation process. Consequently, ΔSESR can be less than 0 during dry periods and greater than 0 during precipitation periods, potentially leading to underestimation or overestimation of FDs. Furthermore, Christian et al. (2019) did not provide a criterion for SESR at FD onset, only specifying a criterion for the minimum SESR at the FD onset stage greater than the 20th percentile ($SESR_{20th}$). This leads to the possibility that SESR at FD onset is below $SESR_{20th}$. Therefore, it is necessary to add a criterion for SESR at FD onset.

In this study, a new method based on the multiples of the mean evaporative stress ratio (MESR) for FD identification was developed to address the aforementioned limitations. Unlike SESR, MESR does not rely on the assumption of a specific probability density function (PDF) for evaporative stress ratio (ESR). Instead, MESR and ΔMESR are fitted by multiple PDFs and converted into percentiles. Variable thresholds are employed for identifying FDs to guarantee that the thresholds of MESR and ΔMESR are less than 0, reflecting a lower level of MESR and a decrease in MESR. The objectives of this study are to: (1) propose an improved FD identification method called MESR based on SESR; (2) characterize FDs in the North China Plain (NCP) from 1981 to 2022 using RZSM, SESR, and MESR and investigate FD temporal and spatial trends; and (3) identify FD hotspots in the NCP and evaluate the impact of thresholds on FD identification."

3. The same happens in the Discussion where a lot of other studies are mentioned without explaining what has been done differently or what is different in their results from those other studies.

Response: Thank you for the comments. We have revised Section 4.1, which compared FD characteristics of this study with those of previous studies. The revised manuscript has emphasized the consistency between the findings of this study and those of others and explained the reason for the inconsistency with other studies. In addition, the impact of FD identification indicators on FD characteristics has also been discussed (see Section 4.1 in lines 554-606). Furthermore, the innovation of this study, namely what has been done differently, has also been added in Section 4.4 (see lines 673-685).

lines 554-606

*"**4.1 FD characteristics compared with previous studies***

A new FD identification method called MESR, which is an improved version of SESR, was developed in this study. On the basis of RZSM, SESR, and MESR, FD events in the NCP from 1981 to 2022 were identified from both agricultural and meteorological

drought perspectives. The spatial distributions of FD characteristics, such as frequency, duration, severity, and intensity, were analyzed. These findings suggest that FD frequency is high in the north-central NCP and low in the southern NCP. Frequency of agricultural FD (FD_{RZSM}) ranges mostly 0 ~ 0.5 times/year, whereas FD_{RZSM} IR ranges primarily from 10 to 50 %. Previous studies have also reported FD_{RZSM} frequency in the NCP. Yuan et al. (2019) estimated FD_{RZSM} frequency in the NCP at 0 ~ 0.6 times/year using the fifth Coupled Model Intercomparison Project (CMIP5). Mukherjee and Mishra (2022a, 2022b) identified FD_{RZSM} events from 1980 to 2018 using ERA5, reporting a FD_{RZSM} frequency of 0.08 ~ 0.38 times/year and a FD_{RZSM} IR of 15 ~ 30 % in the NCP. Wang and Yuan (2022) found a FD_{RZSM} frequency of primarily 0.1 ~ 0.4 times/year in the NCP by ERA5, and Mahto and Mishra (2023) showed a FD_{RZSM} IR of 11 ~ 15 % in the NCP using ERA5. These studies illustrate that frequency and IR determined by the ERA5 dataset are lower than those found in this study, which might be related to the different data sources used.

Moreover, frequency and IR of meteorological FDs (FD_{SESR} and FD_{MESR}) identified in this study are mostly 0.4 ~ 0.8 times/year and 30 ~ 60%, respectively. Gou et al. (2022) identified FD_{SESR} events in the Huaibei Plain (south of the NCP) from 2001 to 2019 using MODIS Global Evapotranspiration Project (MOD16) data with a spatial resolution of 500 m. They discovered a FD_{SESR} frequency of 0.79 ~ 1.05 times/year, which was higher than that in this study. The Huaibei Plain is located in the sub-humid and humid regions with adequate precipitation, abundant heat, and developed agriculture, contributing to its higher FD_{SESR} frequency than the NCP. Deng et al. (2022) found a FD_{SESR} IR of 15 ~ 40 % from 1981 to 2020 in the NCP using ERA5-Land, whereas Christian et al. (2023) obtained a mean FD_{SESR} IR of 25 ~ 35% in the NCP using four reanalysis datasets from 1980 to 2014. These variations highlight the influences of data sources and their spatial resolution on the FD identification. Overall, the FD frequency in this study is within a reasonable range compared with those in previous studies.

Agricultural FDs (FD_{RZSM}) are less frequent than meteorological FDs (FD_{SESR} and FD_{MESR}). The FD_{RZSM} identification is based on soil moisture from an agricultural perspective, whereas FD_{SESR} and FD_{MESR} focus on atmospheric evaporation demand from a meteorological perspective. FD_{RZSM} and FD_{SESR}/FD_{MESR} are independent, whereas FD_{MESR} is an advanced version of FD_{SESR}. The higher FD_{SESR} and FD_{MESR} frequencies than FD_{RZSM} frequency is related to the fact that meteorological FDs are directly determined by atmospheric conditions and is more sensitive to FD occurrence and development. In contrast, agricultural FDs are dependent on soil moisture and influenced by crop regulatory effects and human activities (Meng et al., 2024).

The onset and recovery stages of FD events were distinguished in this study. FD_{RZSM} events have $duration_{Onset}$ of 2 ~ 4 pentads, $duration_{Recovery}$ of 3 ~ 6 pentads, and $duration_{Total}$ of 5 ~ 9 pentads. FD_{SESR} and FD_{MESR} have similar $duration_{Onset}$ (2 ~ 4 pentads), $duration_{Recovery}$ (4 ~ 6 pentads), and $duration_{Total}$ (7 ~ 9 pentads). The findings of this study align with those of other studies on the FD_{RZSM} duration of the NCP. Yuan et al. (2019) determined FD_{RZSM} $duration_{Total}$ to be 4 ~ 8 pentads in the NCP using CMIP5. They also found $duration_{Total}$ to be 6 ~ 8 pentads from 1951 to 2014 in the NCP with three reanalysis datasets with a 1° resolution (Yuan et al., 2023). Zhang et al. (2020) identified FD using RZSM using ERA5 with a 0.25° resolution from 2003 to 2018 and found FD_{RZSM} $duration_{Onset}$ of 3 ~ 4 pentads, $duration_{Recovery}$ of 1 ~ 6 pentads, and $duration_{Total}$ of 5 ~ 10 pentads in the NCP. These findings show that different datasets within various spatial resolutions have a smaller effect on FD duration. Some studies have also identified FD_{SESR} events. Deng et al. (2022) revealed FD_{SESR}

duration$_{Total}$ to be 6 ~ 9 pentads in the NCP by three reanalysis datasets. In the Huaibei Plain, the adjacent region of the NCP, Gou et al. (2022) found FD$_{SESR}$ duration$_{Onset}$ to be 3 ~ 4 pentads and duration$_{Recovery}$ to be 2 ~ 4 pentads. The FD$_{SESR}$ duration$_{Onset}$ in the NCP is high in the south but small in the north, but duration$_{Recovery}$ is high in the north, as shown in Figure S2. The Huaibei Plain is located in the south of the NCP, therefore duration$_{Onset}$ in the Huaibei Plain is greater, but duration$_{Recovery}$ is smaller than in the NCP. The findings of this study demonstrate significant rationality compared with previous findings.

The spatiotemporal variabilities in FD characteristics are explored in this study, revealing a decreasing trend in the NCP. The spatial distribution patterns of duration$_{Total}$ and severity trends are comparable. Similar trends have been shown in other studies. Chen et al. (2019) reported a decreasing FD affected area in the United States from 2000 to 2017 using USDM. Zhang et al. (2021) determined a decreasing tendency in FD$_{RZSM}$ frequency, duration, severity, and intensity in the Gan River Basin from 1961 to 2018 using the variable infiltration capacity (VIC) model driven by meteorological data. Liu et al. (2020) and Deng et al. (2022) also reported that there was a notable decline in FD$_{RZSM}$ and FD$_{SESR}$ affected area in the NCP. Besides, Christian et al. (2023) revealed a noteworthy reduction in FD$_{SESR}$ frequency in the NCP from 1950 to 2015 using four reanalysis datasets. However, other studies have indicated different trends that cannot be disregarded (Yuan et al., 2019, 2023; Zhang et al., 2022a; Zhang et al., 2022c), which concern study areas, data sources, and study periods (Liu et al., 2020a). In addition, FD tendency might be affected by FD identification methods as well. For example, Noguera et al. (2021) revealed that FD frequency identified by the SPI decreased in Spain, whereas those identified by the EDDI and SPEI increased slowly."

lines 673-685

*"**4.4 Innovation and prospects***

This study introduced a new FD indicator, MESR, as an improvement over SESR for regional FD identification. FD identified by MESR addressed the limitations of SESR, and the detailed improvement was explained in Section 2.3.3. Multiple indicators (i.e., RZSM, SESR, and MESR) were utilized to identify and characterize FD events from both agricultural and meteorological perspectives in this study, providing a comprehensive investigation of FD events in the NCP. The differences in FD characteristics are related to background aridity and climatic conditions. Besides, a FD hotspot identification method that combines agricultural and meteorological FDs and considers FD frequency, severity, and intensity was proposed. Percentile was also introduced, and two hotspots that were prone to frequent, severe, and intense FD were identified. Furthermore, the impacts of different thresholds in FD identification methods on FD frequency were quantified. Theoretical and statistical analyses were conducted to demonstrate the improvement in MESR over SESR. Furthermore, control variables were used to investigate the rationality of the spatial heterogeneity in FD$_{MESR}$ and FD$_{SESR}$ frequencies. Typical historical FD events were analyzed to demonstrate the applicability of FD indicators in the NCP as well. However, it is still worth exploring in depth how to quantitatively evaluate whether SESR or MESR is the optimal FD indicator."

4. The novelty of this work is questionable. They have used three indicators to identify flash drought frequency, intensity, severity, and duration of flash droughts. Two of these

indicators have been used before, and it is not well described why the third indicator might have outperformed the others. There are differences between the values derived from the indicators although the spatial pattern is similar. My concern is how one could say which indicator is reporting flash drought characteristics closest to reality. In lines 245-255, the authors have tried to justify the preference of MESR over SESR by showing their distribution, but this does not convince the reader, how about RZSM? It is hard to understand what makes this study different from other studies on flash drought identification in China!

Response: Thank you for the comments. We have emphasized the limitations of FD_{SESR} in lines 91-98 and added the detailed reasons for using MESR instead of SESR to identify FD in lines 204-214. FD_{MESR} is an improvement method proposed based on the limitations of FD_{SESR}. The differences between FD_{MESR} and FD_{SESR} and detailed improvement have also been applied in lines 102-106 and lines 623-625. The distributions of SESR and MESR have been illustrated in Section 3.1, emphasizing the necessity of improving FD_{SESR} method and the improvement effect of FD_{MESR} method (see Section 3.1 in lines 294-317). Furthermore, the limitations of FD_{SESR} and the rationality of the spatially heterogeneous frequencies between FD_{SESR} and FD_{MESR} have been displayed in Section 4.2 (see Section 4.2 in lines 607-653).

Two typical FD events occurring in 2017 and 2019 were utilized to validate the applicability of three FD identification methods in Section 3.4 (see Section 3.4 in lines 477-499). Furthermore, the performance of MESR has been evaluated in Text S2 (see Text S2 in lines 23-55 of Supplementary Materials). Especially in Fig.S18 (c), MESR has slightly higher explanatory ability for RZSM than SESR (see lines 42-44 of Supplementary Materials). Based on the maximal information coefficient (MIC), the correlation between pentad SPI series and MESR and SESR percentile series from 1981 to 2022 in the NCP have also been explored. A stronger correlation between SPI and MESR than between SPI and SESR has been shown in Fig.S19, highlighting a better performance of MESR than SESR (see lines 56-62 of Supplementary Materials). Besides, the uncertainties from the different reanalysis datasets have been evaluated in the Text S1 (see Text S1 in 2-22 of Supplementary Materials). It demonstrates the reliability of the findings using ERA5-Land in this study.

FD_{RZSM} identification, based on soil moisture, is categorized as an agricultural drought. FD_{SESR} is identified by atmospheric evaporation demand and is classified as a meteorological drought. FD_{MESR} identification is an improvement on the FD_{SESR} identification method. It can be thought that FD_{RZSM} and FD_{SESR}/FD_{MESR} are independent, while FD_{MESR} is an advanced version of FD_{SESR} (see lines 85-90 and lines 576-581). The spatial heterogeneity of FD characteristics based on the three indicators has been analyzed, combined with the land use types, Aridity Index (AI), and the ratio of mean annual ET and PET in the revised manuscript (see lines 326-339, lines 363-368, lines 380-388, lines 397-405). Furthermore, the innovation of this study has also been supplemented in the revised manuscript (see Section 4.4 in lines 673-685).

lines 85-90

"FD_{SESR}, identified by anomalous evaporative stress, is classified as a meteorological drought. However, FD_{RZSM}, determined by soil moisture, is categorized as an agricultural drought. Although they are interrelated, they identify FDs from different perspectives. Exploring FD characteristics from multiple perspectives plays an important role in the comprehensive understanding and effective response to FD events. Therefore, this study employs RZSM and SESR to further analyze regional FD characteristics and investigate the effects of climate factors and background aridity."

lines 91-98

"However, the SESR application has several problems. When a FD identified via SESR, both SESR and the change in SESR (ΔSESR) are assumed to follow normal distributions. Therefore, the 50th percentiles of ΔSESR ($\Delta SESR_{50th}$) equal 0, indicating no change in SESR (ΔSESR = 0). A ΔSESR below the 40th percentile of ΔSESR ($\Delta SESR_{40th}$) denotes a decrease in the SESR. However, whether SESR and ΔSESR follow normal distributions remains uncertain. If they do not, $\Delta SESR_{40th}$ may not be less than 0. In the study by Gou et al. (2022), the 36th percentile of ΔSESR ($\Delta SESR_{36th}$) corresponded to an increase in SESR, where $\Delta SESR_{36th}$ was greater than 0. This phenomenon might be because SESR gradually decreases during dry periods and increases during precipitation process. Consequently, ΔSESR can be less than 0 during dry periods and greater than 0 during precipitation periods, potentially leading to underestimation or overestimation of FDs."

lines 102-106

"In this study, a new method based on the multiples of the mean evaporative stress ratio (MESR) for FD identification was developed to address the aforementioned limitations. Unlike SESR, MESR does not rely on the assumption of a specific probability density function (PDF) for evaporative stress ratio (ESR). Instead, MESR and ΔMESR are fitted by multiple PDFs and converted into percentiles. Variable thresholds are employed for identifying FDs to guarantee that the thresholds of MESR and ΔMESR are less than 0, reflecting a lower level of MESR and a decrease in MESR."

lines 204-214

"There are three main differences between FD identification via MESR and SESR. First, the ESR value is divided by its mean to construct the MESR series instead of normalizing ESR to create SESR. Regardless of whether ESR is standardized as SESR or MESR, their percentile values are unaffected by linear transformations based on ESR. However, whether ESR follows a normal distribution requires further determination, which makes the rationality of normalizing ESR into SESR debatable. Therefore, dividing ESR by its mean and transforming it into MESR can avoid considering the PDF that ESR follows. Second, the percentiles of MESR and ΔMESR for each pentad are fitted using an optimal PDF rather than an EDF for conversion into percentiles. Because the distributions of MESR and ΔMESR are yet unknown, several PDFs are fitted to select the optimal PDF, ensuring more precise percentiles. Finally, the variable thresholds of MESR and ΔMESR are employed in the process of FD identification. Due to uncertainty regarding whether SESR and ΔSESR follow a normal distribution, and $SESR_{40th}$ and $\Delta SESR_{40th}$ might be greater than 0, variable percentiles of MESR and ΔMESR are used as thresholds for FD identification. The variable thresholds ensures that the threshold is less than 0."

lines 294-317

*"**3.1 Shortcomings of FD_{SESR} and rationality of FD_{MESR}***

To graphically illustrate the limitations of the FD_{SESR} method, the distributions of $SESR_{50th}$ and $\Delta SESR_{50th}$ are shown in Fig.3 (a) (b). SESR and ΔSESR percentiles are calculated on a grid by pentads, meaning that $SESR_{50th}$ and $\Delta SESR_{50th}$ vary across grids and pentads, and are temporally and spatially affected. Theoretically, SESR and ΔSESR follow a standard normal distribution, where $SESR_{50th} = 0$ and $\Delta SESR_{50th} = 0$. This implies that $SESR_{40th}$ should be below the average level and that $\Delta SESR_{40th}$ should be

less than 0, indicating a decreasing SESR. However, the skewed distributions of $SESR_{50th}$ and $\Delta SESR_{50th}$ in the NCP demonstrate that $SESR_{50th}$ and $\Delta SESR_{50th}$ are generally not 0. When $SESR_{50th} > 0$, $SESR_{40th}$ may also exceed 0, indicating that ESR_{40th} exceeds $\overline{ESR}$ and that $SESR_{40th}$ cannot reliably indicate a low evaporative stress value. When $\Delta SESR_{50th} > 0$, $\Delta SESR_{40th}$ may also exceed 0, reflecting an increasing SESR. This could result in underestimation of the evaporative stress value and inaccurate capture of FD events that do not actually occur. When $SESR_{50th} < 0$, ESR_{40th} may be significantly lower than $\overline{ESR}$, indicating a lower evaporative stress value. When $\Delta SESR_{50th} < 0$, $\Delta SESR_{40th}$ is also significantly less than 0, indicating a severely decreasing SESR. This could lead to overestimation of the evaporative stress value and failure to detect FD events that have actually occurred.

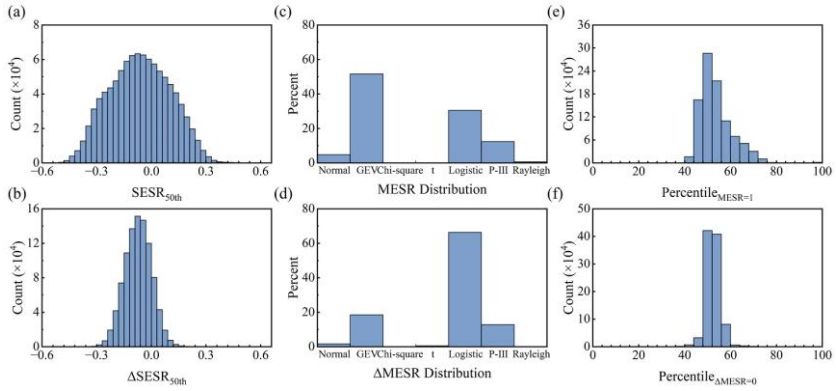


Figure 3 The distribution of (a) $SESR_{50th}$, (b) $\Delta SESR_{50th}$, the optimal PDF that (c) $MESR_{50th}$ and (d) $\Delta MESR_{50th}$ follow, and distribution of (e) Pr_1 and (f) Pr_2.

Because both MESR and SESR are obtained by the linear transformations of ESR, it can be assumed that MESR also follows a skewed distribution, corresponding to Fig.3 (a) (b). Figure 3 (c) (d) shows the optimal PDFs for MESR and $\Delta MESR$ (f_1 and f_2). The typical distributions for MESR (f_1) and $\Delta MESR$ (f_2) are the GEV and logistic distributions. Pr_1 and Pr_2 are shown in Fig.3 (e) and (f). While some Pr_1 and Pr_2 values may still deviate from the 50th percentile, there is a notable improvement over the SESR and $\Delta SESR$ distributions in Fig.3 (a) (b). In addition, the variable thresholds used in FD_{MESR} identification ensure that the thresholds related to Pr_1 and Pr_2 remain less than 0. Therefore, the FD_{MESR} method, where MESR is fitted by several PDFs for FD identification, theoretically offers a more accurate method for identifying FD"

lines 326-339

"The frequency of FD_{RZSM} is high in the central and northeastern NCP and low in the southern NCP, with two high-frequency regions in the central and northeastern NCP. The AI values in the central and northeastern NCP are 0.2 ~ 0.3 and less than 0.2, respectively (Fig.S1 (b)), indicating relatively dry conditions. The lack of precipitation and increased evapotranspiration accelerate the decrease in soil moisture, increasing the likelihood of FDs in the central NCP compared with the southern NCP (Gou et al., 2022; Yuan et al., 2023). In addition, the northeastern NCP is predominantly woodland with high ET (Guo et al., 2007) and experiences frequent FD_{RZSM} due to high water demand. In contrast, FD_{SESR} and FD_{MESR} frequencies are both high in the north-central NCP and low in the northeastern and southern NCP. It is inversely correlated with the ratio of annual ET to PET (Fig.S1 (c)), indicating that a region with greater evaporative stress would experience more FD_{SESR} and FD_{MESR}. The north-central NCP is primarily cultivated land and is influenced by irrigation. This leads to the significant fluctuations in evaporative stress and results in higher FD_{SESR} and FD_{MESR} frequencies.

In contrast, the northeastern NCP, dominated by woodland, is usually not irrigated artificially. Evaporative stress is influenced mostly by climate conditions. Therefore, the northeastern NCP experiences fewer evaporative stress fluctuations and lower FD_{SESR} and FD_{MESR} frequencies (Guo et al., 2007). The southern NCP is characterized by higher temperatures, greater evapotranspiration, and more abundant precipitation as latitude decreases. The balanced hydrothermal conditions in the southern NCP leads to a greater AI and lower FD frequency."

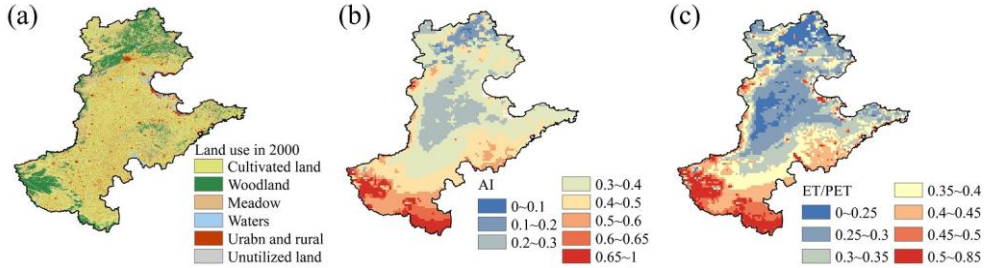


Figure S1 Spatial distribution of (a) the land use in 2010, (b) AI, and (c) the ratio of annual ET and PET in NCP.

lines 363-368

"Warmer temperatures may result in longer FD durations (Zhang et al., 2022c). Woodlands take longer to recover from drought than cultivated lands do (Wu et al., 2024). Additionally, human activities might influence FD $duration_{Total}$. Irrigation could significantly alleviate FDs in cultivated land, whereas woodlands in the northeastern NCP are less impacted by human activity and might have a longer FD $duration_{Total}$. The density distributions of $duration_{Total}$ for FD_{SESR} and FD_{MESR} are similar, predominantly ranging between 7 and 9 pentads, whereas that of FD_{RZSM} is lower, ranging between 5 and 9 pentads. This is connected to their identification methods."

lines 380-388

"Overall, the severities of FD_{RZSM} and FD_{SESR} exhibit spatial distributions similar to their $duration_{Total}$ distributions. FD_{RZSM} severity is high in the southern, eastern, and northeastern NCP and low in the central NCP, whereas FD_{SESR} severity is greater in the southern NCP and lower in the northern NCP. Zhang et al. (2021) reported that the spatial distribution patterns of severity and $duration_{Total}$ were similar in the Gan River Basin. Warming exacerbates drought severity by increasing surface evapotranspiration losses and decreasing soil moisture (Yuan et al., 2019; Zhang et al., 2021). High temperatures also the duration of drought in the southern NCP, which might also result in high severity. In contrast, FD_{MESR} severity is high in the northern and central NCP and low in the southern NCP. It is inversely correlated with the ratio of annual ET to PET (Fig.S1 (c)). These findings indicate that FD_{MESR} $duration_{Total}$ and severity are more strongly correlated with evaporative stress than FD_{SESR} is."

lines 397-405

"Lower intensity values indicate more intense FD events. FD_{RZSM} intensity values are low in the southern, northern, and eastern NCP and high in the central NCP. However, FD_{SESR} and FD_{MESR} intensity values are low in the northern NCP and high in the southern NCP. Intensity reflects the rate of decrease in RZSM, SESR, and MESR. For FD_{RZSM}, RZSM percentiles decrease more slowly in the west-central NCP. The west-central NCP is located within an arid state with an AI ranging 0.2 ~ 0.4. But it is cultivated land, and irrigation has a significant impact on soil moisture, effectively mitigating the decline in RZSM. Although the AI in the northern NCP is less than 0.4,

the woodlands in this region are less impacted by human activities, such as irrigation, which causes RZSM to rapidly decrease. Additionally, high temperatures and great ET in the southern NCP accelerated the RZSM decline. For FD_{SESR} and FD_{MESR}, high ET in the southern NCP might not result in a low ESR value or high evaporative stress (Fig.S1). Abundant precipitation and low evaporative stress slow the decline in SESR and MESR."

lines 477-499

*"**3.4 Typical historical events***

To evaluate the applicability of the three FD identification methods, two typical drought events occurring in 2017 and 2019 were analyzed (Chen et al., 2024; Xue, 2023; Yao et al., 2022). Xue (2023) identified FD events in the NCP between 1978 and 2020 using soil moisture and reported that a FD_{RZSM} event began in late July 2017, peaked in early August, and terminated by mid-August. Figure 9 (a) shows the spatial evolution of FD_{RZSM}, FD_{SESR}, and FD_{MESR} from July to August 2017. FD_{SESR} and FD_{MESR} began in the southwestern NCP on July 5th, were alleviated by August 4th, and were followed by sporadic FDs. A FD_{RZSM} event started on July 15th and eased until August 9. After that, the affected area rapidly shrank and ended on August 29th. In late July, the affected area of FD_{RZSM}, FD_{SESR}, and FD_{MESR} were all large, indicating a severe FD. Furthermore, FD_{SESR} and FD_{MESR} started and developed before FD_{RZSM}, indicating that they may precede and influence FD_{RZSM}. Therefore, the FD_{RZSM}, FD_{SESR}, and FD_{MESR} occurred in 2017 in this study align with the findings of Xue (2023).

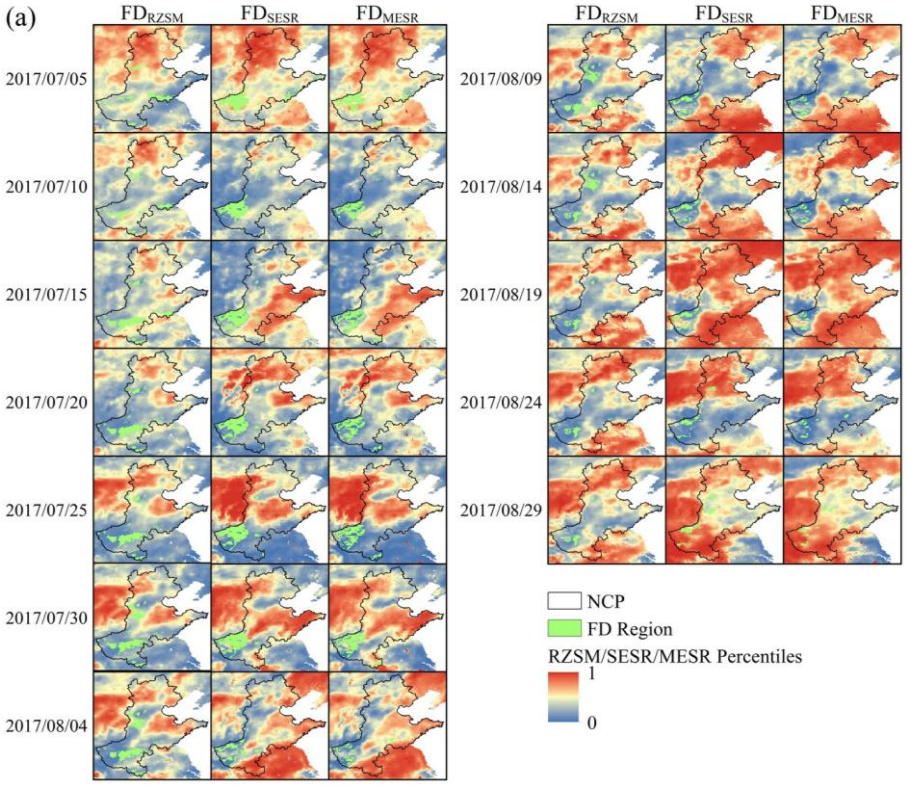


Figure 9 The spatiotemporal evolution process of FD events in (a) 2017 and (b) 2019.

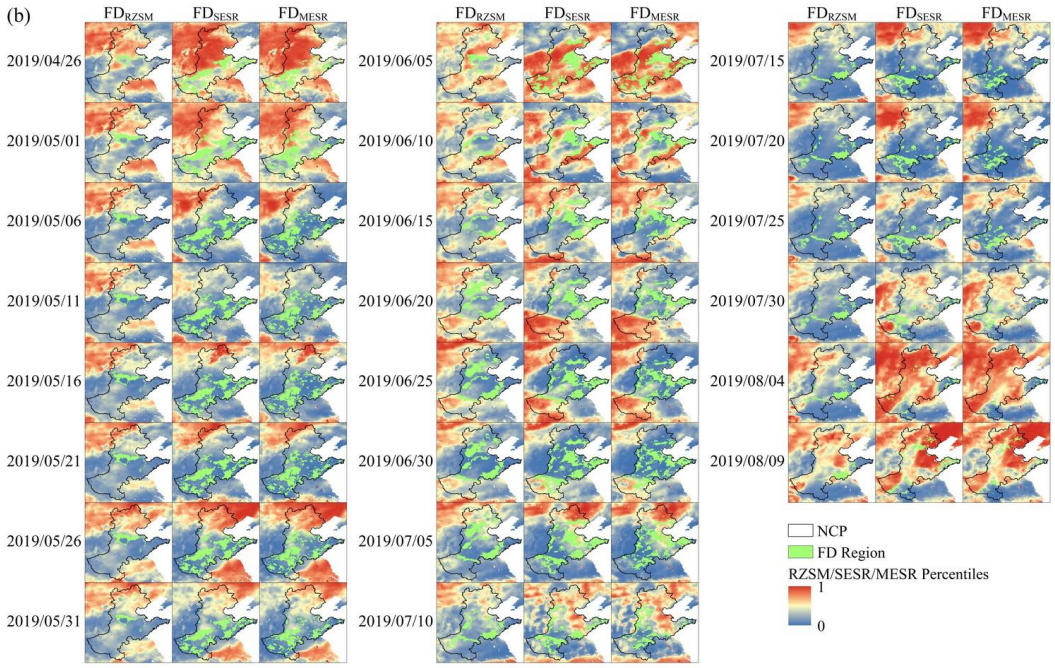


Figure 9 (Continued).

Yao et al. (2022) analyzed the 2019 FD event and discovered that FD_{RZSM} developed rapidly from April 30th to June 9th, during which the RZSM percentiles decreased sharply from 86% to 25%. Afterward, the RZSM percentile decreased once again, FD_{RZSM} severity peaked in July, and recovered in August. Figure 9 (b) shows that FD_{RZSM} started on April 26th and briefly recovered on June 5th. The FD_{RZSM} worsened again on June 20th, reached its peak affected area from late June to early July, and gradually recovered and terminated in August. FD_{SESR} and FD_{MESR} exhibited evolution similar to that of FD_{RZSM}. It began on April 26th, recovered on May 31st, worsened again on June 15th, eased on July 10th, and ended on July 30th. The evolution is consistent with that reported by Yao et al. (2022). Therefore, FD identification via RZSM, SESR, and MESR in this study might be in good agreement with the actual FD events, and the applicability of FD identification methods used in this study is validated."

lines 576-581

"Agricultural FDs (FD_{RZSM}) are less frequent than meteorological FDs (FD_{SESR} and FD_{MESR}). The FD_{RZSM} identification is based on soil moisture from an agricultural perspective, whereas FD_{SESR} and FD_{MESR} focus on atmospheric evaporation demand from a meteorological perspective. FD_{RZSM} and FD_{SESR}/FD_{MESR} are independent, whereas FD_{MESR} is an advanced version of FD_{SESR}. The higher FD_{SESR} and FD_{MESR} frequencies than FD_{RZSM} frequency is related to the fact that meteorological FDs are directly determined by atmospheric conditions and is more sensitive to FD occurrence and development. In contrast, agricultural FDs are dependent on soil moisture and influenced by crop regulatory effects and human activities (Meng et al., 2024)."

lines 607-653

*"**4.2 Attribution analysis of frequency differences between FD_{SESR} and FD_{MESR}***

Compared with FD_{SESR}, FD_{MESR} frequency shows spatial heterogeneity, which is related to the frequency distributions of $SESR_{50th}$ and $\Delta SESR_{50th}$. Figure S11 illustrates the spatial distributions of the mean $SESR_{50th}$ and $\Delta SESR_{50th}$ during FD_{SESR} pentads. Except for the southern NCP, where $SESR_{50th}$ is less than 0, $SESR_{50th}$ in the northern and central NCP is predominantly greater than 0. This may result in $SESR_{40th}$ failing

to represent the real low evaporative stress value and an underestimation of the evaporative stress value. In addition, $\Delta SESR_{50th}$ is greater than 0 in the NCP, which may be accompanied by increasing SESR and inaccurate capture of FD_{SESR} events that do not occur, potentially causing overestimation in the NCP (see Sect.3.2). Theoretically, $SESR_{40th}$ and $\Delta SESR_{40th}$ in the NCP should be less than 0, representing a low evaporative stress value and a decreasing ESR. The distributions of average $SESR_{40th}$ and $\Delta SESR_{40th}$ in Fig.S11 (c) (d) indicate that both are less than 0, except in the northern NCP, where $SESR_{40th}$ exceeds 0. This further shows that unobserved FD_{SESR} events might be captured in the northern NCP. The histograms of $SESR_{50th}$, $\Delta SESR_{50th}$, $SESR_{40th}$, and $\Delta SESR_{40th}$ during FD_{SESR} events are shown in Fig.S12. $SESR_{50th}$ and $\Delta SESR_{50th}$ values are mostly greater than 0, particularly $\Delta SESR_{50th}$, corresponding to Fig.S11 (a) (b). However, most $SESR_{40th}$ and $\Delta SESR_{40th}$ are less than 0. Notably, approximately 35% of $SESR_{40th}$ and $\Delta SESR_{40th}$ are greater than 0, supporting the likelihood of FD_{SESR} overestimation in the NCP. Notably, the difference between Fig.3 (a) (b) and Fig.S12 (a) (b) is that Fig.S12 (a) (b) shows the $SESR_{50th}$ and $\Delta SESR_{50th}$ values during FD_{SESR} occurrence, whereas Fig.3 (a) (b) shows those for all pentads.

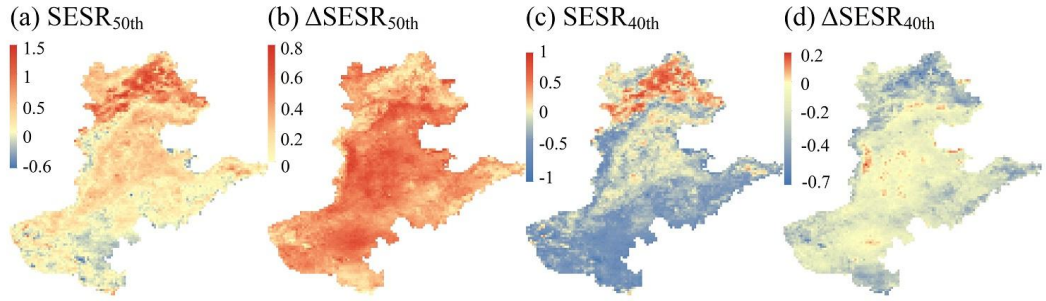


Figure S11 Distribution of the mean (a) $SESR_{50th}$, (b) $\Delta SESR_{50th}$, (c) $SESR_{40th}$, and (d) $\Delta SESR_{40th}$ on FD_{SESR} pentads.

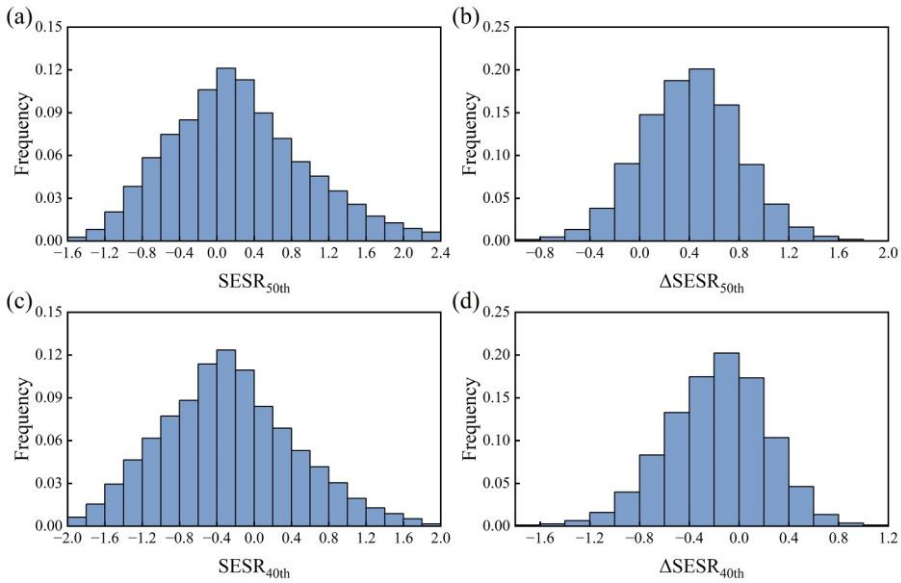


Figure S12 Histogram of (a) $SESR_{50th}$, (b) $\Delta SESR_{50th}$, (c) $SESR_{40th}$, and (d) $\Delta SESR_{40th}$ on FD_{SESR} pentads.

The FD identification based on SESR and MESR differs in three aspects. First, ESR is divided by its mean to construct MESR rather than being normalized into SESR. Second, MESR is fitted using various PDFs instead of EDF. Third, variable thresholds are utilized in FD_{MESR} identification. Both MESR and SESR are linearly transformed from ESR to facilitate a comparison of FD identification results between different regions.

Converting ESR to MESR rather than SESR enhances the rationality of the ESR standardization process. However, the linear transformation of ESR into SESR or MESR does not affect its corresponding percentiles. Therefore, the differences between FD_{SESR} and FD_{MESR} identification results can be attributed to two factors: PDF fitting and variable thresholds.

To quantify the contributions of PDF fitting and variable thresholds to the differences between FD_{SESR} and FD_{MESR}, a modified version of FD_{MESR}, termed $FD_{MESR\text{-}invariable}$ was created. The fixed thresholds from the FD_{SESR} method instead of variable thresholds were applied, i.e., MESRonset1 = 40, MESRonset2 = 20, ΔMESRonset = 40, and MaxMESRchange = 25. Figure S13 (a) ~ (c) displays the frequency of $FD_{MESR\text{-}invariable}$ and the differences between $FD_{MESR\text{-}invariable}$ frequency and FD_{SESR} and FD_{MESR} frequencies. Figure S13 (d) (e) shows the contributions of PDF fitting and variable thresholds to the differences between FD_{SESR} and FD_{MESR} frequencies, and Fig.S13 (f) illustrates their relative contributions.

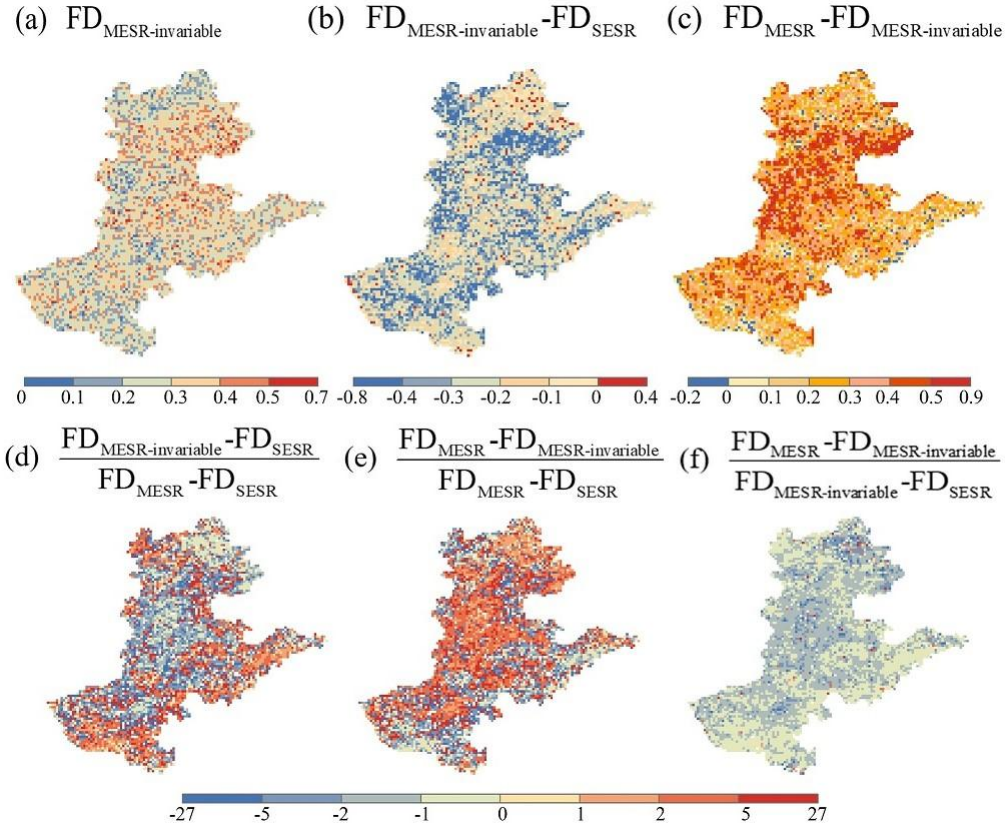


Figure S13 Frequency of $FD_{MESR\text{-}invariable}$ and its difference with FD_{MESR} and FD_{SESR}. *$FD_{MESR\text{-}invariable}$ and FD_{MESR} frequencies exhibit similar spatial distributions, with higher frequencies in the north-central NCP and lower frequencies in the southern NCP. In addition, $FD_{MESR\text{-}invariable}$ has a lower frequency than both FD_{SESR} and FD_{MESR} do, indicating that PDF fitting decreases FD frequency, whereas variable thresholds increase it. According to Fig.S13 (d) (e), PDF fitting has a negative effect on the difference between FD_{SESR} and FD_{MESR} frequencies in the northeastern and west-central NCP, whereas variable thresholds have a positive effect. However, it is opposite in the other NCP regions. As shown in Fig.S13 (f), the relative contribution of variable thresholds to PDF fitting mostly ranges between -2 and -1 in the northeastern and west-central NCP but between -1 and 0 in other regions. Therefore, the absolute contribution of PDFs fitting is greater than 1 in the northeastern and west-central NCP but less than 1 in other regions. Considering the negative contribution of PDF fitting in the NCP, it*

can be assumed that the contribution of PDF fitting is less than that of variable thresholds in the northeastern and west-central NCP but greater in other regions. Therefore, the frequency difference between FD_{SESR} and FD_{MESR} frequencies is driven mainly by variable thresholds in the northeastern and west-central NCP, whereas PDF fitting plays a greater role in other regions.

FD_{SESR} frequency is lower than FD_{MESR} frequency in the northeastern and west-central NCP but higher in other regions. Although FD_{SESR} frequency is overestimated in the NCP, particularly in the northern NCP, PDF fitting decreases FD frequency in the FD_{MESR}. However, in the northeastern and west-central NCP, the increase in FD frequency due to variable thresholds outweighs the decreases due to PDF fitting, resulting in a higher FD_{MESR} frequency than FD_{SESR} frequency. In other regions, the decrease in FD frequency from PDF fitting surpasses the increase from variable thresholds, leading to a lower FD_{MESR} frequency than FD_{SESR} frequency."

lines 2-22 of Supplementary Materials

*"**Text S1 Uncertainties from the reanalysis datasets***

To assess data-related uncertainties, the soil moisture, ET, and PET obtained from two additional reanalysis datasets, the GLEAM and GLDAS 2 datasets, were utilized to identify FD_{RZSM}, FD_{SESR}, and FD_{MESR}. Because the RZSM, SESR, and MESR are the basis for FD identification, their pentad percentiles were determined. Figure S14 shows the Taylor diagrams comparing the pentad RZSM, SESR, and MESR percentile series from 1981 to 2022 for the ERA5-Land, GLEAM, and GLDAS 2 datasets, with the ERA5-Land dataset as the reference. The pentad RZSM, SESR, and MESR percentiles from the GLEAM and GLDAS 2 datasets are highly consistent. Their correlation coefficients are approximately 0.7, centered root mean square differences are approximately 0.8, and standard deviations are close to 1. Therefore, the pentad percentiles of RZSM, SESR, and MESR of the ERA5-Land dataset are consistent with those of the GLEAM and GLDAS 2 datasets. Figure S15 displays the spatial distributions of correlations for the pentad RZSM, SESR, and MESR percentiles between the ERA5-Land dataset and the GLEAM and GLDAS 2 datasets. As shown in Fig.S15 (a) and (d), the RZSM percentile correlation between the ERA5-Land and the GLEAM is comparable to that between the ERA5-Land and the GLDAS 2, primarily exceeding 0.6. SESR percentile correlation between the ERA5-Land and the GLEAM is similar to MESR percentile correlation, which mostly exceeds 0.5. The correlations of SESR and MESR percentiles between the ERA5-Land and the GLDAS 2 are primarily between 0.4 and 0.7, with a comparable spatial distribution pattern. The greater correlation between the ERA5-Land and the GLEAM than between the ERA5-Land and the GLDAS 2 might be due to the coarse spatial resolution of the GLDAS 2.

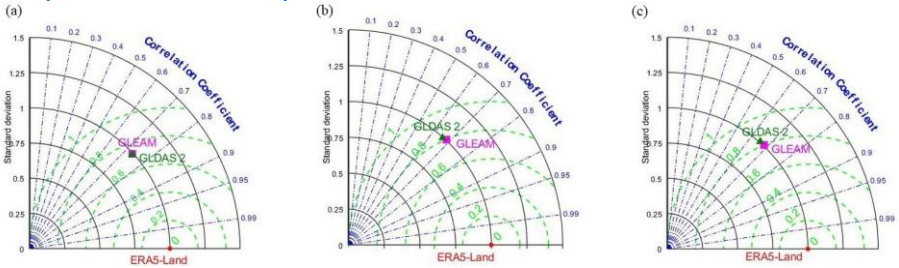


Figure S14 Taylor diagram for the pentad (a) RZSM, (b) SESR, and (c) MESR percentiles based on ERA5-Land, GLEAM, and GLDAS 2 datasets.

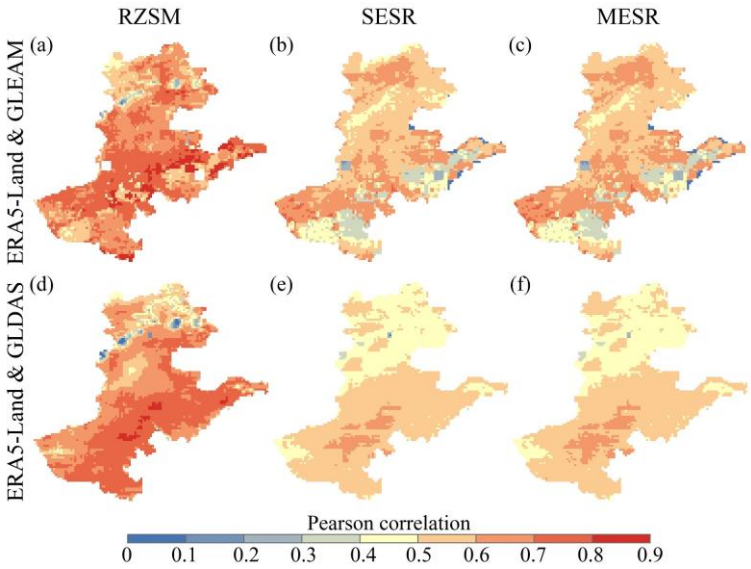


Figure S15 Spatial distribution of the Pearson correlation of the pentad RZSM, SESR, and MESR percentiles between ERA5-Land and (a) ~ (c) GLEAM and (d) ~ (f) GLDAS 2 datasets.

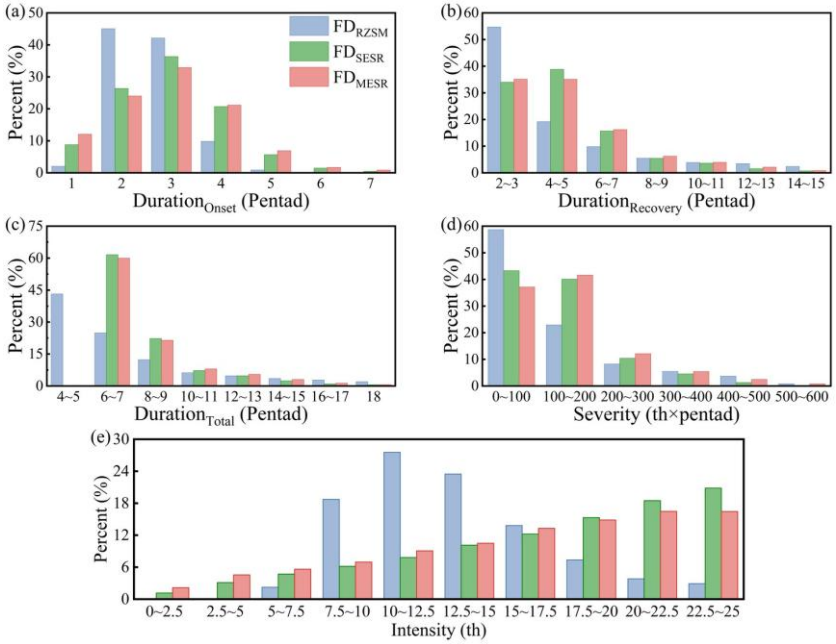


Figure S5 Histogram of FD characteristics identified by RZSM, SESR, and MESR based on ERA5-Land.

In addition, FD characteristics identified using the ERA5-Land, GLEAM, and GLDAS 2 datasets are displayed in Figs.S5, S16 and S17. The distributions of FD_{RZSM}, FD_{SESR}, and FD_{MESR} characteristics are consistent across the datasets, except for FD_{MESR} intensity based on the GLDAS 2. Moreover, the proportions of various FD grades determined by intensity from diverse datasets also demonstrate strong agreement, as shown in Fig.S6. The similarities in the pentad RZSM, SESR, and MESR percentiles from various datasets and FD characteristics across datasets effectively demonstrate the reliability of our findings."

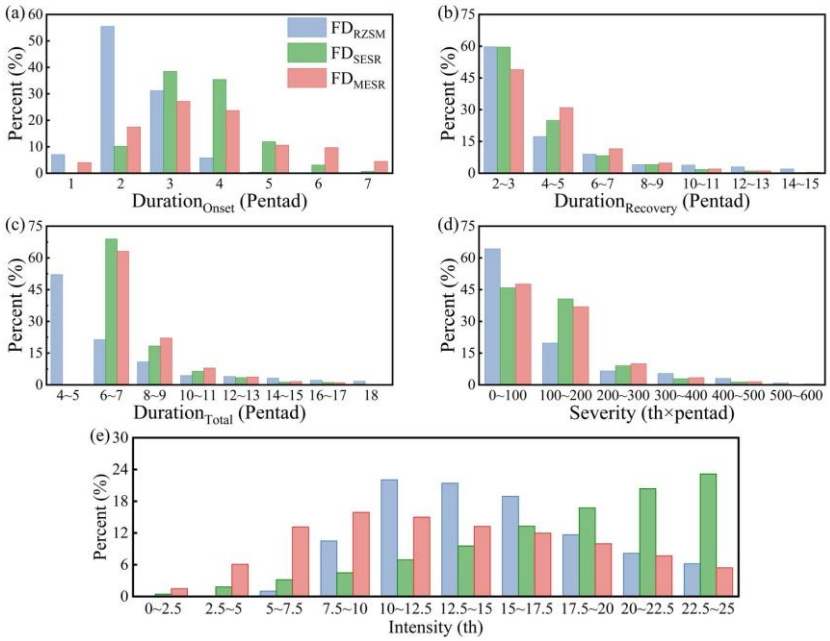


Figure S16 Same as Figure S5, but based on GLEAM.

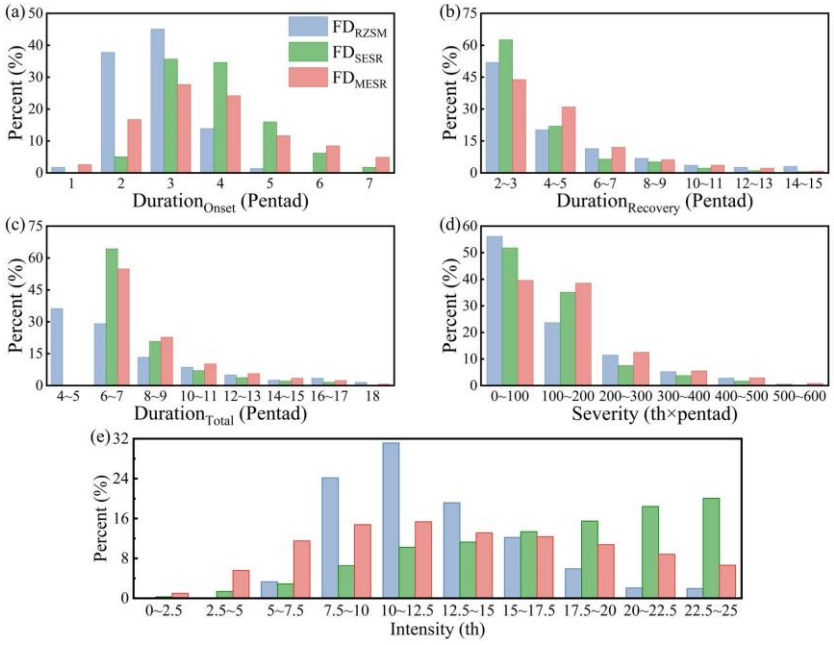


Figure S17 Same as Figure S5, but based on GLDAS 2.

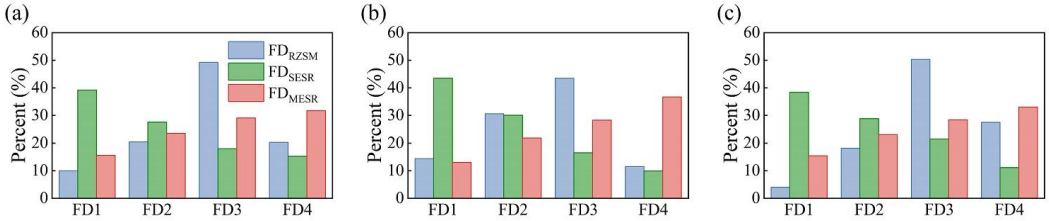


Figure S6 Proportion of FD1 ~ 4 for FD_{RZSM}, FD_{SESR}, and FD_{MESR} based on (a) ERA5-Land, (b) GLEAM, and (c) GLDAS 2 datasets.

lines 23-55 of Supplementary Materials

*"**Text S2 Explanatory ability between different FD types***

Given the influence of climate control on FD occurrence (Mukherjee and Mishra, 2022), there might be relationships between different FD types. Therefore, the coefficient of

determination (R^2) was used to quantify the relationships among FD_{RZSM}, FD_{SESR}, and FD_{MESR}. R^2 represents the capacity of linear regression to explain the variance in the dependent variable based on the independent variables (Mukherjee and Mishra, 2022). In particular, the relationship between FD_{RZSM} and FD_{SESR}, denoted as "RZSM ~ SESR", is represented by the R^2 derived from the linear regression between RZSM percentile (dependent variable) and SESR percentile (independent variable) during FD_{RZSM} pentads. Moreover, other relationships, such as "RZSM ~ MESR", "SESR ~ RZSM", "MESR ~ RZSM", "SESR ~ MESR", and "MESR ~ SESR", were determined, as shown in the first two columns of Fig.S18.

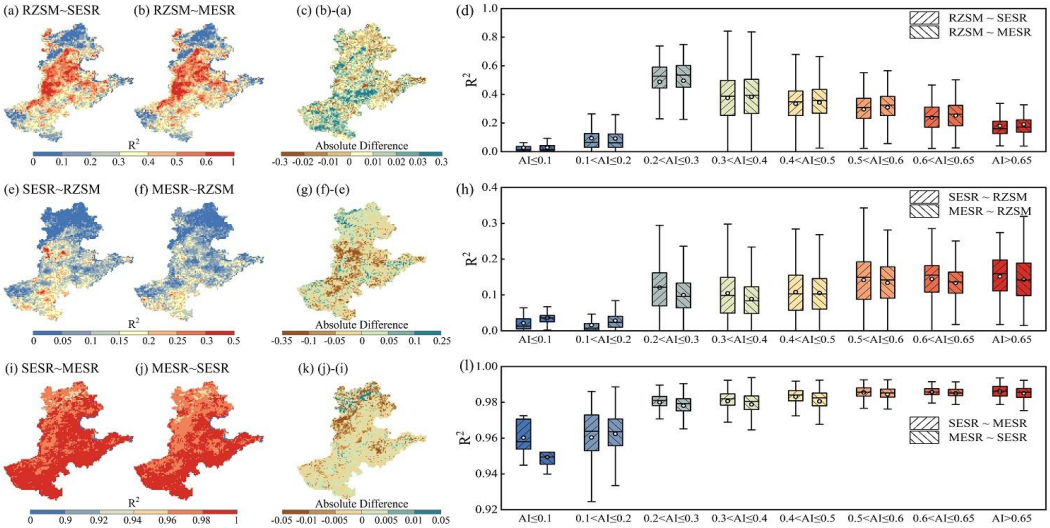


Figure S18 Spatial distribution of the R^2 determined by (a) "RZSM ~ SESR", (b) "RZSM ~ MESR", (e) "SESR ~ RZSM", (f) "MESR ~ RZSM", (i) "SESR ~ MESR", and (j) "MESR ~ SESR", as well as the differences of R^2 (c) between "RZSM ~ SESR" and "RZSM ~ MESR", (g) between "SESR ~ RZSM" and "MESR ~ RZSM", and (k) between "SESR ~ MESR" and "MESR ~ SESR". The boxplots in the (d), (h), and (l) illustrate the R^2 in the "RZSM ~ SESR" and "RZSM ~ MESR", "SESR ~ RZSM" and "MESR ~ RZSM", and "SESR ~ MESR" and "MESR ~ SESR" over different AI values.

In Fig.S18 (a) and (b), both "RZSM ~ SESR" and "RZSM ~ MESR" explain more than 40% of the variance in RZSM percentile in the central NCP but less than 30% in other regions. In Fig.S18 (e) and (f), the southern NCP shows higher R^2 values for "SESR ~ RZSM" and "MESR ~ RZSM" (mostly approximately 15% ~ 25%) than the northern NCP does (less than 15%). However, "SESR ~ MESR" explains more than 90% of the variance in SESR, and "MESR ~ SESR" explains more than 90% of the variance in MESR, as shown in Fig.S18 (i) and (j). Overall, the explanatory ability of these relationships, ranked from highest to lowest, is "SESR ~ MESR" and "MESR ~ SESR" > "RZSM ~ SESR" and "RZSM ~ MESR" > "SESR ~ RZSM" and "MESR ~ RZSM". Because SESR and MESR are both derived from the linear transformation of ESR, they exhibit strong mutual explanatory ability. The relationships between MESR and RZSM ("RZSM ~ MESR" and "MESR ~ RZSM") are quite comparable to those between SESR and RZSM ("RZSM ~ SESR" and "SESR ~ RZSM"), highlighting the reliability of FD identification using MESR. The differences in Fig.S18 (c), (g), and (k) further demonstrate the similarities between SESR and MESR. However, considering the propagation from meteorological to agricultural drought, "RZSM ~ MESR" has a slightly greater explanatory ability for RZSM than "RZSM ~ SESR" in Fig.S18 (c), illustrating that MESR performs better than SESR in explaining RZSM.

The spatial distributions of R^2 also reveal sensitivity to the AI, as shown in Fig.S18 (d),

(h), and (l). For "RZSM ~ SESR" and "RZSM ~ MESR", the explanatory ability increases with increasing AI in the region where AI < 0.3 but decreases in the region where AI > 0.3. The region with AIs between 0.2 and 0.3 has the highest explanatory ability for RZSM percentile (approximately 60%). Overall, the RZSM percentiles could be better explained by the SESR and MESR percentiles in drier regions, except where AI < 0.2, which might be related to RZSM being greater initially in wetter regions with extended memory (Mukherjee and Mishra, 2022). For "SESR ~ RZSM" and "MESR ~ RZSM", the explanatory ability is less than 20%, and is obviously greater in the region with AI > 0.2 than in the region with AI < 0.2. SESR and MESR percentiles could be better explained by RZSM percentile in wetter regions with less evaporative stress and greater evaporation. Meteorological drought (FD_{SESR} and FD_{MESR}) might lead to agricultural drought (FD_{RZSM}), resulting in lower R^2 for "SESR ~ RZSM" and "MESR ~ RZSM" compared to "RZSM ~ SESR" and "RZSM ~ MESR". For "SESR ~ MESR" and "MESR ~ SESR", the explanatory ability exceeds 90% and generally increases with increasing AI overall."

lines 56-62 of Supplementary Materials

"To quantify which indicator, SESR or MESR, can better identify FDs in the NCP, the relationships between the two indicators and the pentad SPI series were measured. The maximal information coefficient (MIC) can measure both linear and nonlinear relationships between two series (Cao et al., 2021). Therefore, the MICs between the pentad SPI series and SESR or MESR percentile series (denoted as $MIC_{SPI\&SESR}$ and $MIC_{SPI\&MESR}$) were calculated grid by grid, and their differences are shown in Fig.S19. The difference between $MIC_{SPI\&SESR}$ and $MIC_{SPI\&MESR}$ is predominantly greater than 0, demonstrating a stronger correlation between SPI and MESR than between SPI and SESR. This suggests that the performance of MESR is better than that of SESR in the meteorological drought identification in the NCP."

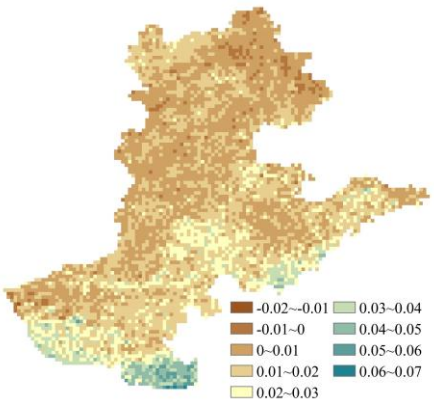


Figure S19 Difference between $MIC_{SPI\&SESR}$ and $MIC_{SPI\&MESR}$.