Precipitation and vegetation transpiration variations dominate the dynamics of agricultural drought characteristics in China

Wenwen Guo [a], Shengzhi Huang [a*], Qiang Huang [a], Dunxian She [b], Haiyun Shi [c], Guoyong Leng [d], Ji Li [d], Liwen Cheng [a], Yuejiao Gao [a], Jian Peng [e,f]

a. State Key Laboratory of Eco-Hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, China

b. State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

c. Shenzhen Key Laboratory of Precision Measurement and Early Warning Technology for Urban Environmental Health Risks, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

d. Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

e. Department of Remote Sensing, Helmholtz Centre for Environmental Research-UFZ, Permoserstrasse 15, 04318, Leipzig, Germany

f. Remote Sensing Centre for Earth System Research, Leipzig University, Talstr. 35, 04103, Leipzig, Germany

---

[*] Corresponding author at: State Key Laboratory of Eco-hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, China. Tel.: +86 29 82312801; fax: +86 29 82312797. E-mail Address: huangshengzhi7788@126.com.

Abstract: Agricultural drought posing a significant threat to agricultural production is subject to the complex influence of ocean, terrestrial and meteorological multi-factors. Nevertheless, which factor dominating the dynamics of agricultural drought characteristics and their dynamic impact remain equivocal. To address this knowledge gap, we used ERA5 soil moisture to calculate the standardized soil moisture index (SSI) to characterize agricultural drought. The extreme gradient boosting model was then adopted to fully examine the influence of ocean, terrestrial and meteorological multi-factors on agricultural drought characteristics and their dynamics in China. Meanwhile, the shapley additive explanation values were introduced to quantify the contribution of multiple drivers to drought characteristics. Our analysis reveals that the drought frequency, severity and duration in China ranged from 5-70, 2.15-35.02 and 1.76-31.20, respectively. Drought duration is increasing and drought intensity is intensifying in southeast, north and northwest China. In addition, potential evapotranspiration is the most significant driver of drought characteristics at the basin scale. Regarding the dynamic evolution of drought characteristics, the percentages of raster points for drought duration and severity with evapotranspiration as the dominant factor are 30.7% and 32.7%, and the percentages with precipitation are 35.3% and 35.0%, respectively. Precipitation in northern regions has a positive effect on decreasing drought characteristics, whilst in southern regions, evapotranspiration dominates the dynamics in drought characteristics due to increasing vegetation transpiration. Moreover, the drought severity is exacerbated by the Atlantic Multidecadal Oscillation in the Yangtze and Pearl River basins, while the contribution of the North Atlantic Oscillation to the drought duration evolution is increasing in the Yangtze River basin. Generally, this study sheds new insights into

agricultural drought evolution and driving mechanism, which are beneficial for agricultural drought early warning and mitigation.

Keywords: Agricultural drought; Drought events; Drought dynamics; Driving Factors; China

1. Introduction

Drought, a natural disaster caused by water scarcity, is becoming more frequent, severe, and prolonged in many parts of the world due to climate change (Felsche and Ludwig, 2021; Li et al., 2022c; Wu et al., 2021a; Xu et al., 2015). Agricultural drought is a phenomenon in which insufficient rainfall or surface water supply leads to continuous decline in soil moisture, inhibits crop production and reduces grain yield (Crow et al., 2012; Li et al., 2022b). As an important component of the regional water cycle, soil moisture is the main source of water for vegetation growth, and it is also a sensitive indicator for evaluating the development of agricultural drought, which is of great significance to the monitoring of agricultural drought (Somorowska, 2022; Wu et al., 2021b). In historical period, frequent droughts severely constrain socio-economic development, and threaten agricultural production and ecosystem security (Feng et al., 2021; Guo et al., 2023; Liu et al., 2020a; Ma et al., 2015). On the one hand, persistent soil moisture deficit can affect crop growth and lead to lower crop yields (Wei et al., 2019). For example, the annual average drought-related grain loss in China from 2000 to 2020 reached 25.719 billion kg (Ministry of Water Resources, 2023), and the cumulative global grain production losses from 1983 to 2009 amounted to 166 billion U.S. dollars (Kim et al., 2019). On the other hand, soil moisture deficit can prolong drought recovery time and exacerbate the impact of drought on terrestrial ecosystems (Yao et al., 2023). In particular, from July to August in 2013, a two-month drought reduced the carbon sequestration in southern China by 101.54 Tg C, accounting for 39-53% of the annual net carbon sink of China's land ecosystem (Yuan et al., 2016). In 2010, southwest China experienced a severe spring drought, which reduced regional annual GPP by about 65 Tg C (Li et al., 2019; Zhang et al., 2012). Moreover, drought-induced water stress was the main cause for the reduction of terrestrial carbon sinks in northern China, and the study showed that the maximum reduction of GPP in the region was 0.09 Pg C yr^{-1} in 1999-2011 compared to 1982-1998 (Yuan et al., 2014). Therefore, exploring the dynamic evolution of agricultural drought characteristics and the driving patterns is crucial for determining future development directions and adopting drought mitigation measures to cope with climate change.

In recent years, many studies have investigated the drivers of drought and their impacts (Deng et al., 2021; Ma et al., 2020; Qiu et al., 2017). For example, Deng et al. (2021) used a stepwise regression approach to identify the drivers of drought and showed that precipitation causes an extremely severe deficit in terrestrial water storage in the Huang-Huai-Hai Plain. In addition, precipitation deficits have been proven to be a major cause of multi-year agricultural droughts in California, and warming will also exacerbate the likelihood of extreme droughts (Luo et al., 2017; Williams et al., 2015).The global sensitivity of SOBOL was used to assess the sensitivity of precipitation and potential evapotranspiration to the frequency of drought events, and it was found that the drought events frequency dominated by potential evapotranspiration decreases from southeast to northwest in China (Ma et al., 2020). In addition, Zhang et al. (2018) evaluated the effects of climate change and human activities on hydrological drought events based on different hydrological models. The results found that the dominant factor of hydrological drought severity was precipitation, followed by potential evapotranspiration and human activities in the middle reaches of the Yangtze River. Moreover, it is worth noting that climate extremes have been shown to be related to circulation factors in China and globally. It was found that agricultural drought represented by soil moisture was influenced by the El Niño-Southern Oscillation (ENSO) (Zhang et al., 2021b), and the influence of the North Atlantic Oscillation (NAO) on dry-heat complex events was mainly concentrated in northwest, northeast and east China (Wu et al., 2021c). These studies lay the groundwork for understanding the driving mechanisms of agricultural drought. However, previous studies focused on the relationship between agricultural drought and meteorological factors or circulation factors, while the analysis of the drivers of agricultural drought characteristics considering multiple factors (marine, terrestrial and meteorology) is quite limited. Due to the uncertainty of climate change caused by global warming, which factor dominates the agricultural drought dynamics and the dynamic impact of driving factors on drought characteristics are still open questions.

Recently, the application of machine learning methods to drought monitoring, prediction, and attribution has gained increasing recognition. Current studies using machine learning methods to assess drought in China focused on developing a comprehensive agricultural drought index for agricultural drought monitoring (Cheng et al., 2023; Liu et al., 2020b), building disaster vulnerability models to assess the potential impact of crop disaster risk (Li et al., 2021), constructing drought prediction models (Felsche and Ludwig, 2021; Li et al., 2020), and analyzing the drivers of agricultural drought-affected area and drought-suffering area (Deng et al., 2022). In general, previous studies have explored the application of machine learning to agricultural drought and provided valuable insights for drought impact. However, few studies have applied machine learning to the identification of the drivers of the agricultural drought characteristics dynamics in raster data from drought event perspective. In the context of global warming, the dynamic response of agricultural drought characteristics caused by soil moisture stress to multiple factors is not revealed. Furthermore, how different factors affect the spatial pattern of agricultural drought characteristics dynamics is also a question that needs to be explored. Studying the changes in drivers affecting agricultural drought characteristics can help policy makers to adopt drought mitigation measures to reduce the adverse effects of drought. Recent studies found that tree-based machine learning models, such as extreme gradient boosting (XGB), are popular non-parametric models for attribution analysis (Ebrahimi-Khusfi et al., 2022; Felsche and Ludwig, 2021; Li et al., 2022a; Lundberg et al., 2020). To better explore the mechanisms of agricultural drought evolution and investigate the influencing factors and their relative contributions of drought, we used the XGB algorithm to identify the response of different factors to drought characteristics. On this basis, the contribution of individual factors to drought characteristics was determined by calculating shapley additive explanation (SHAP) values (Lundberg and Lee, 2017), which can improve the interpretability of the XGB model and increase our knowledge of the contribution of variables. Therefore, we applied an interpretable machine learning framework to identify the potential mechanisms affecting the dynamics of drought characteristics. In summary, the main objectives of this study are: 1) to identify regional drought characteristics in China; 2) to assess the main factors of basin-scale drought characteristics; 3) to determine the contribution of influencing factors to the dynamic evolution of drought characteristics and their dynamic impacts.

2. Study area and data

2.1. Study area

China has a vast territory and many rivers. The terrain is high in the west and low in the east, with a terraced topographic distribution. In addition, the spatiotemporal distribution of precipitation is uneven. Precipitation is mostly concentrated in the summer and autumn, while precipitation decreases from the southeast coast to the northwest inland spatially. According to the Institute of Geographical Sciences and Resources of the Chinese Academy of Sciences, the basins are divided into the following nine basin areas (Fig.1), including Songhua and Liaohe River Basin (R1), Haihe River Basin (R2), Huaihe River Basin (R3), Yellow River Basin (R4), Yangtze River Basin (R5), Pearl River Basin (R6), Southeast Basin (R7), Southwest Basin (R8) and Continental Basin (R9). The hydrological characteristics of the basins are shown in Table 1.

Place Figure 1 here.

Place Table 1 here.

2.2. Data

Monthly soil moisture (SM) data for 1980-2021 is obtained from the ERA5 reanalysis dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=form). Monthly precipitation (P), temperature (T), and vapor pressure (VAP) are obtained from the Climatic Research Unit (CRU) at CRU TS v.4.06

(https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.06/), and VPD is calculated from T and VAP. Monthly actual evapotranspiration (E), potential evapotranspiration (Ep) and transpiration (Et) are available at the Global Land Evaporation Amsterdam Model (GLEAM v3.6a, https://www.gleam.eu/). LAI data is obtained from ERA5. The DEM data and soil texture data from the Resource and Environment Science and Data Center are used in the study (https://www.resdc.cn/Default.aspx). The Circulation factor used in this study come from the Physical Sciences Laboratory (https://www.psl.noaa.gov/data/climateindices/list/). Finally, the SM, E, Ep and Et datasets are resampled to 0.5◦ × 0.5◦ spatial resolution using the mean aggregation method.

3. Methods

3.1. Drought index calculation and drought event identification

3.1.1. Standardized soil moisture index

Based on the ERA5 soil moisture data from 1980-2021, the standardized soil moisture index (SSI) was calculated. The SSI was calculated with reference to the calculation of the standardized precipitation index (SPI) (McKee et al., 1993). First, appropriate distribution functions were selected to fit the soil moisture sequence of each raster in China. Six commonly used probability density functions were selected to fit the soil moisture series, which were gamma distribution, exponential distribution, weibull distribution, generalized extreme value distribution, log-normal distribution and normal distribution. By Kolmogorov-Smirnov test (K-S test) and root mean square error (RMSE), the distribution function that best conforms to the empirical cumulative distribution probability function curve was selected as the optimal distribution function for each raster. Then, the SSI was obtained by normalizing the cumulative probability of the optimal distribution function. However, at some raster points, the above parametric distribution functions may not be suitable. Therefore, for these raster points, the Gringorten plotting position algorithm was used to calculate the marginal probability of soil moisture to obtain a nonparametric normalized index instead of the empirical probability distribution (Farahmand and AghaKouchak, 2015; Gringorten, 1963).

3.1.2. Drought event identification

The run theory is a method to extract drought events by setting relevant thresholds based on the characteristics of drought index on the time sequence (Yevjevich, 1967). Since a large number of mild droughts in the sample may have an impact on the statistical features (Fleig et al., 2006). Based on this, for the calculated SSI series, we used three-threshold optimized run theory to identify agricultural drought events in China, and then drought events were eliminated and merged to obtain drought characteristics (drought frequency, drought duration and drought severity) (He et al., 2016; Shen et al., 2016; Shi et al., 2023; Wang et al., 2019). The specific process of drought event identification using the threshold method is as follows:

1) It is initially identified as a drought event when SSI is less than -0.5 (blue area in Fig. 2), as shown in Fig. 2 there are five droughts (a-e).

2) On the basis of 1), small drought events are eliminated, i.e. for drought events with drought duration of only 1 month, if SSI > -1.0, it is classified as no drought occurred in this month (Fig. 2a), otherwise it is considered that an independent drought occurred (Fig. 2b).

3) For two adjacent drought events with an interval of 1 month (Fig. 2d and 2e), if the interval month SSI < 0, the two adjacent drought events are merged into one drought event,

otherwise, they are two independent drought events. The drought duration is the sum of the two drought duration plus 1, and the drought severity is the sum of the severity of the two drought events.

Place Figure 2 here.

3.1.3. Trend analysis

The Mann-Kendall (MK) trend test is a nonparametric test that distinguishes trends in time sequences (Mann, 1945). It has the advantage that the sample series do not need to follow a specific distribution and is often used to test the trend of variable time series (Guo et al., 2021; Yue et al., 2018). For time series x_i, the specific principle of MK trend test is as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}\left(x_j - x_i\right) \tag{1}$$

$$\mathrm{sgn}(\theta) = \begin{cases} 1, & \theta > 0 \\ 0, & \theta = 0 \\ -1, & \theta < 0 \end{cases} \tag{2}$$

where n is the data length; sgn is the sign function.

Then the variance is:

$$var(S) = \frac{n(n-1)(2n+5)}{18} \tag{3}$$

The standard normalization statistic Z could be expressed as:

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{var(S)}} & S > 0 \\ 0 & S = 0 \\ \dfrac{S+1}{\sqrt{var(S)}} & S < 0 \end{cases} \tag{4}$$

when $Z > 0$, the sequence has an upward trend, otherwise it has a downward trend. The significance level is set at 0.05. When $|Z| \geq 1.96$, it represents that the trend of the series passes the 95% significance test and the trend is significant. Conversely, the trend of the series is not significant.

3.2. Vapor pressure deficit

VPD is the difference between the saturation vapor pressure and the actual vapor pressure. In this study, the saturation vapor pressure (SVP) is first calculated using the Goff-Gratch formula, and then the actual vapor pressure (VAP) is subtracted to obtain the VPD. The Goff-Gratch formula is the saturation vapor pressure calculation formula recommended by the World Meteorological Organization in 1966. The VPD is calculated as:

$$\mathrm{VPD} = \mathrm{SVP} - \mathrm{VAP} \tag{5}$$

$$\lg(\mathrm{SVP}) = c_1(1 - 273..16/T) + c_2 \lg(T/273..16) + c_3\left[1 - 10^{c_4(T/273.16-1)}\right] + c_5\left[10^{c_6(1-273.16/T)} - 1\right] + c_7 \tag{6}$$

where c_1=10.79574, c_2=-5.02800, c_3=1.50475 × 10^{-4}, c_4=-8.29690, c_5=0.42873 × 10^{-3}, c_6=4.76955, c_7=0.78614, and T=273.15+t, t is the Celsius temperature (℃).

3.3. Drivers of the drought characteristics dynamics

Based on the identification of drought characteristics, the XGB algorithm and SHAP values were combined to quantify the effects of marine, terrestrial and meteorological drivers on drought duration and drought severity (factors in Table 2).

Place Table 2 here.

The XGB uses a gradient boosting structure and has the advantage of parallel tree boosting (Chen and Guestrin, 2016). It integrates weak classifiers into a strong classifier to obtain a better regression performance than a single model. By introducing regular items to control the complexity of the model, it can prevent model overfitting and improve modeling performance (Fan et al., 2018; Shin et al., 2019). Moreover, due to the lack of transparency and interpretability of traditional machine learning methods, the visibility of feature importance is poor. Therefore, determining the contribution of influencing factors to target variables changes and improving the interpretability of models are important issues in the modeling process of machine learning algorithms (Gilpin et al., 2018). Recently, the emergence of interpretable methods has improved the understanding of learning model or predictions (Deng et al., 2022; Wang et al., 2022b). The SHAP value is one such interpretable approach that quantifies feature importance, determines the contribution of drivers and elucidates the dependencies between input features and output targets (Lundberg and Lee, 2017). Therefore, in this study, the XGB algorithm was used to construct regression relationships between drought characteristics and factors, and Grid Search method was used to determine the optimal combination of parameters. Then a model based on the optimized parameters was built to identify the response of factors to drought characteristics, and finally SHAP value was used to quantify the magnitude of the effect of each factor on drought characteristics.

4. Results

4.1. Spatial and temporal evolution of agricultural drought in China

The time series curves of monthly SSI for nine basins in China from 1980 to 2021 are shown in Fig. 3. It can be seen that the regional average SSI in China ranges from -0.65 to 0.80, with the smallest fluctuation range of -1.03 - 0.71 in the continental basin and the largest fluctuation range of -2.16 - 2.23 in the Huaihe River basin among the nine major basins. The statistical monthly SSI trends from 1980 to 2021 are shown in Fig. 4. Most of the northwest regions, especially the Qinghai-Tibet Plateau region, show an increasing trend of SSI, accompanied by weakening trend of agricultural drought. In contrast, SSI tends to decrease in northeastern China, Yellow and Huaihua regions and southeastern China. Due to the influence of summer precipitation, there is a trend of becoming wet in parts of the middle and lower reaches of the Yangtze River from May to July.

Place Figure 3 here.

Place Figure 4 here.

4.2. Agricultural drought characteristics and their variations

For each raster, drought events were extracted according to the run theory, and then the mean values of drought characteristics under all drought events were calculated. As shown in Fig. 5, the frequency of drought in China ranges from 5 - 70, and the regions with higher frequency are mainly concentrated in the southern region (R5-R8). The drought severity and drought duration range from 2.15 - 35.02 and 1.76 - 31.20, respectively. Spatially, drought severity is greater and has a longer duration in the northwest (R9) and the western northeast (R1). In summary, the drought duration and drought severity in northern China are higher than those in other regions, but southern China has a high frequency of droughts and will face a higher risk of agricultural drought (Fig. 5g, h).

Place Figure 5 here.

4.3. Driving factors of drought characteristics

4.3.1. Dominant factors of basin-scale drought characteristics

We used the XGB model to establish a model between drought characteristics and driving factors, and then used the Grid Search method to find the optimal parameters of the XGB model for each watershed (Table 3). Based on optimized parameters, a model was established to identify the response of different drivers to drought characteristics at the watershed scale, and then the feature importance was calculated based on the model and quantified as the mean absolute SHAP value of each factor. Fig. 6 shows the influence of the drivers on drought severity and drought duration in nine basins of China. The results indicate that the dominant drivers of drought characteristics of agricultural drought events in China vary among basins, with potential evapotranspiration (Ep) dominating in the majority of regions. For the basin scale, the influence of meteorological factors plays the largest role, followed by the impact of vegetation on the drought duration and severity. Compared to the influence of meteorological and vegetation factors on drought, the influence of circulation factors is weak, but cannot be ignored (Forootan et al., 2019; Wang et al., 2022a). As can be seen from Fig. 6, it was found that the closer the watershed to the ocean, the more prominent the influence of circulation factors, such as the R6 and R7 watersheds.

Place Table 3 here.

Place Figure 6 here.

From the importance scores of drivers on drought severity (Table 4), it can be seen that the SHAP value of Ep spans 1.7, from the lowest value in the R6 (0.13) to the highest value in the R9 (1.83). The influence of vegetation on drought severity is also significant in R4 and R8, where droughts are frequent and vegetation vulnerability is high. Soil moisture is a direct source of water available to vegetation. The promotion of ecological projects such as reforestation increases the amount of water dissipated by vegetation, which will further exacerbate soil water deficit and thus affect the severity of drought. As can be seen from the impact of the driving factors on the drought duration (Table 5), the SHAP values of Ep range from 0.13 to 1.68. In addition to Ep, precipitation (P) is the second most important factor affecting the drought duration in the R1 and R2, while VPD is the secondary factor affecting duration in R9. In the R1, R2, R5 and R6 basins, Pacific Decadal Oscillation (PDO) has a greater impact on drought duration than other circulation factors, while Nino3.4 plays a greater role in the R4 and the R7.

Place Table 4 here.

Place Table 5 here.

4.3.2. Grid-based dominant factors on drought characteristic evolution and their dynamics

To investigate the causes of the dynamics of drought characteristics, variable importance was calculated for each raster based on the XGB model and SHAP values to assess the influence of marine, terrestrial, and meteorological factors on drought characteristics, where the maximum score is identified as the dominant factor in the drought characteristics dynamics (Fig. 7). Fig. 7 shows that the dynamics of drought characteristics in southern China is mainly attributed to the actual evapotranspiration (E), while the dynamics of drought characteristics in the north is mainly dominated by P. Among them, the percentages of drought duration and severity with E as the dominant factor are 30.7% and 32.7%, respectively, while the proportion of P as the dominant factor in drought duration and severity are 35.3% and 35.0%, respectively. In addition, for the drought duration dynamics, the raster points with VPD as the dominant factor are mainly concentrated in the northern part of R1 and the western part of R4 and R5. The area of VPD influence on drought severity is mainly distributed in the western part of R4 and R5.

Place Figure 7 here.

SHAP values provide both global and local interpretability for machine learning models by providing feature importance values. Therefore, for each raster point, the SHAP values of individual factors in each drought event were calculated (Fig. 8). The temporal variation of SHAP values was used to evaluate the changing influence of each factor on drought characteristics. An increasing trend in SHAP values indicates that the contribution of the factor to drought characteristics is increasing, while a decreasing trend suggests that the factor is weakening the drought characteristics. As can be seen from Fig. 8 and Fig. 9, the decreased precipitation has an increasing influence on the dynamic evolution of drought severity at the majority of raster points in southern China basins, especially in the Pearl River basin (R6). While the contribution of water demand, such as potential evapotranspiration, VPD and vegetation transpiration, to drought dynamics is gradually increasing. Consequently, the decrease in water supply and the increase in water demand depletion further exacerbate the increased drought risk in southern China. Moreover, in the Yangtze and Pearl River basins

(R5 and R6), Atlantic Multidecadal Oscillation (AMO) exacerbates the severity of drought. And the rising trend of North Atlantic Oscillation (NAO) is also much more concentrated in the south, leading to a larger dynamic contribution to the drought duration. Therefore, future research should focus on the evolution of droughts in the southern basins of China.

Place Figure 8 here.

Place Figure 9 here.

----

5. Discussion

In the context of climate change, our findings show that agricultural drought events in northern China are low in frequency but high in severity and long in duration. In contrast, the southern region experiences high frequency, short duration and weak severity of droughts. Also, there is a trend of increasing duration and severity in this region, indicating an increased agricultural risk due to drought in this area. These results are consistent with previous studies that have found that droughts become more frequent, with a progressively larger impact area and extremely prominent extreme weather events under a warming climate (Ayantobo et al., 2017; He et al., 2016; Ma et al., 2020; Zhou et al., 2021).

We also investigated the spatial distribution and dynamic evolution of agricultural drought characteristics in China, and then used the XGB model and SHAP values to assess the contributions of marine, terrestrial, and meteorological factors to the duration and severity of agricultural drought. Regarding the dynamic evolution of drought characteristics, there are regional differences in the drivers of the duration and severity of agricultural droughts in

China. In the same basin, droughts would exhibit different conditions depending on meteorological factors, groundwater storage and underlying surface conditions (Han et al., 2020; Yang et al., 2022). This is because the frequency of droughts occurring in each raster is different, and the corresponding drought duration and severity are different, so there is a slight difference in the relative importance derived from the modeling of each raster. However, the dominant factor still has a certain spatial distribution pattern in the watershed. It is found that drought characteristics are mainly influenced by P in the northern China, while E is the main driver for the evolution of drought characteristics in the southern region. To further investigate the influence of dominant factors on drought characteristics, we analyzed the inter-annual trend of annual precipitation. As seen in Fig. 9 and Fig. 5(g, h), the annual precipitation at the raster points where drought severity and drought duration decreased in the northern China showed an increasing trend. This indicates that the increased precipitation in the northern China alleviates the drought and mitigates the risk of agricultural drought. It is similar to the finding of Huang et al. (2015) that the frequency of extreme droughts decreases with increasing precipitation in northwest China.

As known in previous studies, the frequency of drought events dominated by evapotranspiration decreases from southeast to northwest in China, and the frequency of drought events by Ep in southeast China is greater than that of drought events dominated by precipitation deficit (Ma et al., 2020). In terms of dominant factors, the dynamics of drought characteristics in the southern region are mostly influenced by E. To clarify the mechanisms by which drought characteristics are influenced by E, we explored the relationship between E and vegetation transpiration (Et). It is found that Et is closely related to E with high correlation (Fig. 10a), and vegetation transpiration significantly increases in most of the regions (Fig. 10b), suggesting that vegetation transpiration has an important role in the intensification of agricultural drought in the southern China. Similar findings of vegetation significantly increasing evaporative water consumption and exacerbating the risk of agricultural drought are also confirmed in the Loess Plateau region (Han et al., 2021; Shao et al., 2019). Moreover, the second and third factors in the ranking of importance are mostly Ep and VPD in southern China (Fig. 7), indicating that the atmospheric evaporation demand due to temperature rise also has a greater impact on drought in this region. The significant increase of Ep and VPD in Fig.10 (c-d) also confirms the finding that the atmospheric evaporation demand increases in this region. Wang et al. (2022c) found that the contribution of increased terrestrial evapotranspiration is greater in humid areas, mainly because humid areas could provide sufficient water supply to meet atmospheric evaporation demands and vegetation physiological activities. Increased evapotranspiration indicates more surface water loss and less soil moisture, which could exacerbate drought stress in terrestrial ecosystems, affecting water resources, climate, and agriculture (Wang et al., 2022c; Zhang et al., 2021a). In summary, water shortage caused by increased atmospheric evaporative demand and water depletion caused by vegetation transpiration jointly contribute to the exacerbation of agricultural drought in the southern region. Our study improves the understanding of the response of the evolution of agricultural drought characteristics to the driving factors and provides a scientific basis for drought adaptation strategies.

Place Figure 10 here.

In this study, machine learning with interpretable methods is used to obtain the drivers of the evolution of agricultural drought characteristics in China. Machine learning has the characteristics of nonlinearity, high estimation accuracy, and strong generalization ability, which can effectively process large amounts of data. However, it should be noted that machine learning methods cannot directly quantify the internal mechanisms of model behavior. It is a data-driven model subject to data and algorithm constraints, which may introduce some uncertainty to the quantification of the contribution of drivers. In addition, the method of parameter optimization in model construction is also one of the sources of uncertainty. Another deficiency of the attribution analysis is the insufficient consideration of human activities (e.g., CO_2 emissions, irrigation, land use change, etc.), which needs to be enhanced in subsequent studies. Despite uncertainties and limitations, the XGB model can still estimate the impact of drivers on drought characteristics. It improves the understanding of machine learning models by combining with interpretability methods (SHAP), making it easier to quantify feature importance and clarify dependencies between input features and output targets. Therefore, to reduce uncertainty, future studies also need to evaluate the effects of variables on changes in agricultural drought characteristics under multiple models such as RNNs and LSTM models to obtain more reliable attribution results.

6. Conclusion

In this study, we applied an interpretable machine learning framework to identify the potential mechanisms affecting the dynamics of drought characteristics. Our findings showed that agricultural drought events in northern China have low frequency, high severity and long duration in the context of climate change, while those have the opposite characteristics in southern China. In addition, there is an increasing trend in drought duration and severity in southeast, north, and northwest China. At the basin-scale, evapotranspiration is the most influential driver of drought characteristics. Moreover, we identified regional differences in the drivers of drought dynamics. The percentages of drought duration and severity with E as the dominant factor were 30.7% and 32.7%, and 35.3% and 35.0% with P as the dominant factor, respectively. Precipitation in northern China positively contributes to reducing drought duration and intensity, while water scarcity caused by increased atmospheric evaporation demand and water depletion due to vegetation transpiration led to the intensification of agricultural drought in southern China. Furthermore, the contributions of AMO and NAO to drought characteristics are gradually increasing. Our study explores the dynamics and driving patterns of agricultural drought characteristics and improves the understanding of the evolution of agricultural drought characteristics in response to drivers, which are important for developing effective drought mitigation measures and adapting to climate change.

Acknowledgements

This research was jointly supported by the National Natural Science Foundation of China (grant number 52279026), the National Key R&D Program of China (grant number 2021YFC3000203), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant number XDA28060100), and the Natural Science Foundation of Inner Mongolia Autonomous Region (grant number 2021ZD12).

References

Ayantobo, O.O., Li, Y., Song, S. et al., 2017. Spatial comparability of drought characteristics and related return periods in mainland China over 1961-2013. Journal of Hydrology, 550: 549-567. https://doi.org/10.1016/j.jhydrol.2017.05.019

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing

Machinery, San Francisco, California, USA, pp. 785–794. http://doi.org/10.1145/2939672.2939785

Cheng, M., Zhong, L., Ma, Y. et al., 2023. A new drought monitoring index on the Tibetan Plateau based on multisource Data and machine learning methods. Remote Sensing, 15(2). https://doi.org/10.3390/rs15020512

Crow, W.T., Kumar, S.V., Bolten, J.D., 2012. On the utility of land surface models for agricultural drought monitoring. Hydrology and Earth System Sciences, 16(9): 3451-3460. https://doi.org/10.5194/hess-16-3451-2012

Deng, S., Liu, S., Mo, X., 2021. Assessment and attribution of China's droughts using an integrated drought index derived from GRACE and GRACE-FO data. Journal of Hydrology, 603. https://doi.org/10.1016/j.jhydrol.2021.127170

Deng, X., Wang, G., Yan, H. et al., 2022. Spatial-temporal pattern and influencing factors of drought impacts on agriculture in China. Frontiers in Environmental Science, 10: 820615. https://doi.org/10.3389/fenvs.2022.820615

Ebrahimi-Khusfi, Z., Dargahian, F., Nafarzadegan, A.R., 2022. Predicting the dust events frequency around a degraded ecosystem and determining the contribution of their controlling factors using gradient boosting-based approaches and game theory. Environmental Science and Pollution Research, 29(24): 36655-36673. https://doi.org/10.1007/s11356-021-17265-0

Fan, J., Wang, X., Wu, L. et al., 2018. Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. Energy Conversion and Management, 164: 102-111. https://doi.org/10.1016/j.enconman.2018.02.087

Farahmand, A., AghaKouchak, A., 2015. A generalized framework for deriving nonparametric standardized drought indicators. Advances in Water Resources, 76: 140-145. https://doi.org/10.1016/j.advwatres.2014.11.012

Felsche, E., Ludwig, R., 2021. Applying machine learning for drought prediction in a perfect model framework using data from a large ensemble of climate simulations. Natural Hazards and Earth System Sciences, 21(12): 3679-3691. https://doi.org/10.5194/nhess-21-3679-2021

Feng, S., Hao, Z., Zhang, X. et al., 2021. Changes in climate-crop yield relationships affect risks of crop yield reduction. Agricultural and Forest Meteorology, 304. https://doi.org/10.1016/j.agrformet.2021.108401

Fleig, A.K., Tallaksen, L.M., Hisdal, H. et al., 2006. A global evaluation of streamflow drought characteristics. Hydrology and Earth System Sciences, 10: 535–552. https://doi.org/10.5194/hess-10-535-2006

Forootan, E., Khaki, M., Schumacher, M. et al., 2019. Understanding the global hydrological droughts of 2003-2016 and their relationships with teleconnections. Science of the Total Environment, 650: 2587-2604. https://doi.org/10.1016/j.scitotenv.2018.09.231

Gilpin, L.H., Bau, D., Yuan, B.Z. et al., 2018. Explaining explanations: an overview of interpretability of machine learning, 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), pp. 80-89. http://doi.org/10.1109/DSAA.2018.00018

Gringorten, I.I., 1963. A plotting rule for extreme probability paper Journal of Geophysical Research, 68(3): 813-814. https://doi.org/10.1029/jz068i003p00813

Guo, W., Huang, S., Huang, Q. et al., 2023. Drought trigger thresholds for different levels of vegetation loss in China and their dynamics. Agricultural and Forest Meteorology, 331: 109349. https://doi.org/10.1016/j.agrformet.2023.109349

Guo, Y., Huang, Q., Huang, S. et al., 2021. Elucidating the effects of mega reservoir on watershed drought tolerance based on a drought propagation analytical method. Journal of Hydrology, 598.

https://doi.org/10.1016/j.jhydrol.2020.125738

Han, Z., Huang, Q., Huang, S. et al., 2021. Spatial-temporal dynamics of agricultural drought in the Loess Plateau under a changing environment: Characteristics and potential influencing factors. Agricultural Water Management, 244. https://doi.org/10.1016/j.agwat.2020.106540

Han, Z., Huang, S., Huang, Q. et al., 2020. Effects of vegetation restoration on groundwater drought in the Loess Plateau, China. Journal of Hydrology, 591: 125566. https://doi.org/10.1016/j.jhydrol.2020.125566

He, J., Yang, X., Li, Z. et al., 2016. Spatiotemporal variations of meteorological droughts in China during 1961–2014: An investigation based on multi-threshold identification. International Journal of Disaster Risk Science, 7: 63-76. https://doi.org/10.1007/s13753-016-0083-8

Huang, X., Li, Y., Feng, J. et al., 2015. Climate characteristics of precipitation and extreme drought events in Northwest China. Acta Ecologica Sinica, 35(5): 1359-1370.(in Chinese). https://doi.org/10.5846/stxb201305101013

Kim, W., Iizumi, T., Nishimori, M., 2019. Global patterns of crop production losses associated with droughts from 1983 to 2009. Journal of Applied Meteorology and Climatology, 58(6): 1233-1244. https://doi.org/10.1175/jamc-d-18-0174.1

Li, L., Zeng, Z., Zhang, G. et al., 2022a. Exploring the individualized effect of climatic drivers on MODIS net primary productivity through an explainable machine learning framework. Remote Sensing, 14(17). https://doi.org/10.3390/rs14174401

Li, P., Huang, Q., Huang, S. et al., 2022b. Various maize yield losses and their dynamics triggered by drought thresholds based on Copula-Bayesian conditional probabilities. Agricultural Water Management, 261. https://doi.org/10.1016/j.agwat.2021.107391

Li, X., Li, Y., Chen, A. et al., 2019. The impact of the 2009/2010 drought on vegetation growth and terrestrial carbon balance in Southwest China. Agricultural and Forest Meteorology, 269-270: 239-248. https://doi.org/10.1016/j.agrformet.2019.01.036

Li, Y., Huang, S., Wang, H. et al., 2022c. High-resolution propagation time from meteorological to agricultural drought at multiple levels and spatiotemporal scales. Agricultural Water Management, 262. https://doi.org/10.1016/j.agwat.2021.107428

Li, Z., Chen, T., Wu, Q. et al., 2020. Application of penalized linear regression and ensemble methods for drought forecasting in Northeast China. Meteorology and Atmospheric Physics, 132(1): 113-130. https://doi.org/10.1007/s00703-019-00675-8

Li, Z., Zhang, Z., Zhang, L., 2021. Improving regional wheat drought risk assessment for insurance application by integrating scenario-driven crop model, machine learning, and satellite data. Agricultural Systems, 191. https://doi.org/10.1016/j.agsy.2021.103141

Liu, L., Gudmundsson, L., Hauser, M. et al., 2020a. Soil moisture dominates dryness stress on ecosystem production globally. Nature Communications, 11(1): 4892. https://doi.org/10.1038/s41467-020-18631-1

Liu, X., Zhu, X., Zhang, Q. et al., 2020b. A remote sensing and artificial neural network-based integrated agricultural drought index: Index development and applications. Catena, 186. https://doi.org/10.1016/j.catena.2019.104394

Lundberg, S.M., Erion, G., Chen, H. et al., 2020. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1): 56-67. https://doi.org/10.1038/s42256-019-0138-9

Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Advances in neural information process, 30: 4768–4777.

Luo, L., Apps, D., Arcand, S. et al., 2017. Contribution of temperature and precipitation anomalies to the

California drought during 2012-2015. Geophysical Research Letters, 44(7): 3184-3192. https://doi.org/10.1002/2016gl072027

Ma, B., Zhang, B., Jia, L. et al., 2020. Conditional distribution selection for SPEI-daily and its revealed meteorological drought characteristics in China from 1961 to 2017. Atmospheric Research, 246. https://doi.org/10.1016/j.atmosres.2020.105108

Ma, M., Ren, L., Singh, V.P. et al., 2015. Evaluation and application of the SPDI-JDI for droughts in Texas, USA. Journal of Hydrology, 521: 34-45. https://doi.org/10.1016/j.jhydrol.2014.11.074

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica: Journal of the econometric society, 13(3): 245-259. https://doi.org/10.2307/1907187

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales, Proceedings of the 8th Conference on Applied Climatology. Boston, pp. 179-183.

Ministry of Water Resources Resources., 2023. China flood and drought disaster bulletin. China Water Resources and Hydropower Press, Beijing.

Qiu, J., Wang, Y., Xiao, J., 2017. Spatiotemporal distribution of droughts in the Xijiang River Basin, China and its responses to global climatic events. Water, 9(4). https://doi.org/10.3390/w9040265

Shao, R., Zhang, B., Su, T. et al., 2019. Estimating the increase in regional evaporative water consumption as a result of vegetation restoration over the Loess Plateau, China. Journal of Geophysical Research: Atmospheres, 124(22): 11783-11802. https://doi.org/10.1029/2019jd031295

Shen, D., Shang, C., Fang, X. et al., 2016. Spatio-temporal characteristics of drought duration and drought severity in Guizhou. Journal of Arid Land Resources and Environment, 30(07): 138-143.(in Chinese). https://doi.org/10.13448/j.cnki.jalre.2016.229

Shi, P., Tang, H., Qu, S. et al., 2023. Characteristics of propagation from meteorological drought to hydrological drought in Southwest China. Water Resources Protection, 39(01): 49-56.(in Chinese).

Shin, J.-y., Ro, Y., Cha, J.-W. et al., 2019. Assessing the applicability of random forest, stochastic gradient boosted model, and extreme learning machine methods to the quantitative precipitation estimation of the radar data: a case study to Gwangdeoksan Radar, South Korea, in 2018. Advances in Meteorology, 2019. https://doi.org/10.1155/2019/6542410

Somorowska, U., 2022. Amplified signals of soil moisture and evaporative stresses across Poland in the twenty-first century. Science of the Total Environment, 812. https://doi.org/10.1016/j.scitotenv.2021.151465

Wang, F., Lai, H., Li, Y. et al., 2022a. Identifying the status of groundwater drought from a GRACE mascon model perspective across China during 2003-2018. Agricultural Water Management, 260. https://doi.org/10.1016/j.agwat.2021.107251

Wang, H., Yan, S., Ciais, P. et al., 2022b. Exploring complex water stress-gross primary production relationships: Impact of climatic drivers, main effects, and interactive effects. Global Change Biology, 28(13): 4110-4123. https://doi.org/10.1111/gcb.16201

Wang, R., Li, L., Gentine, P. et al., 2022c. Recent increase in the observation-derived land evapotranspiration due to global warming. Environmental Research Letters, 17(2). https://doi.org/10.1088/1748-9326/ac4291

Wang, R., Zhang, J., Guo, E. et al., 2019. Integrated drought risk assessment of multi-hazard-affected bodies based on copulas in the Taoerhe Basin, China. Theoretical and Applied Climatology, 135(1-2): 577-592. https://doi.org/10.1007/s00704-018-2374-z

Wei, Y., Jin, J., Jiang, S. et al., 2019. Simulated assessment of summer maize drought loss sensitivity in Huaibei Plain, China. Agronomy-Basel, 9(2). https://doi.org/10.3390/agronomy9020078

Williams, A.P., Seager, R., Abatzoglou, J.T. et al., 2015. Contribution of anthropogenic warming to California drought during 2012-2014. Geophysical Research Letters, 42(16): 6819-6828. https://doi.org/10.1002/2015gl064924

Wu, H., Su, X., Singh, V.P., 2021a. Blended dry and hot events index for monitoring dry-hot events over global land areas. Geophysical Research Letters, 48(24). https://doi.org/10.1029/2021gl096181

Wu, H., Su, X., Singh, V.P. et al., 2021b. Agricultural drought prediction based on conditional distributions of vine copulas. Water Resources Research, 57(8). https://doi.org/10.1029/2021wr029562

Wu, X., Hao, Z., Hao, F. et al., 2021c. Influence of large-scale circulation patterns on compound dry and hot events in China. Journal of Geophysical Research-Atmospheres, 126(4). https://doi.org/10.1029/2020jd033918

Xu, K., Yang, D., Yang, H. et al., 2015. Spatio-temporal variation of drought in China during 1961–2012: A climatic perspective. Journal of Hydrology, 526: 253-264. https://doi.org/10.1016/j.jhydrol.2014.09.047

Yang, F., Duan, X., Guo, Q. et al., 2022. The spatiotemporal variations and propagation of droughts in Plateau Mountains of China. Science of The Total Environment, 805: 150257. https://doi.org/10.1016/j.scitotenv.2021.150257

Yao, Y., Liu, Y., Zhou, S. et al., 2023. Soil moisture determines the recovery time of ecosystems from drought. Global change biology. https://doi.org/10.1111/gcb.16620

Yevjevich, V., 1967. An objective approach to definitions and investigations of continental hydrologic droughts. Journal of Hydrology, 7: 353-356.

Yuan, W., Cai, W., Chen, Y. et al., 2016. Severe summer heatwave and drought strongly reduced carbon uptake in Southern China. Scientific Reports, 6(1): 1-12 https://doi.org/10.1038/srep18813

Yuan, W., Liu, D., Dong, W. et al., 2014. Multiyear precipitation reduction strongly decreases carbon uptake over northern China. J. Geophys. Res.-Biogeosci., 119(5): 881-896. https://doi.org/10.1002/2014jg002608

Yue, Y., Shen, S.-h., Wang, Q., 2018. Trend and variability in droughts in northeast China based on the reconnaissance drought index. Water, 10(3). https://doi.org/10.3390/w10030318

Zhang, D., Zhang, Q., Qiu, J. et al., 2018. Intensification of hydrological drought due to human activity in the middle reaches of the Yangtze River, China. Science of the Total Environment, 637: 1432-1442. https://doi.org/10.1016/j.scitotenv.2018.05.121

Zhang, L., Xiao, J., Li, J. et al., 2012. The 2010 spring drought reduced primary productivity in southwestern China. Environmental Research Letters, 7(4): 045706. https://doi.org/10.1088/1748-9326/7/4/045706

Zhang, W., Li, Y., Wu, X. et al., 2021a. Divergent response of vegetation growth to soil water availability in dry and wet periods over Central Asia. Journal of Geophysical Research-Biogeosciences, 126(6). https://doi.org/10.1029/2020jg005912

Zhang, Y., Hao, Z., Feng, S. et al., 2021b. Agricultural drought prediction in China based on drought propagation and large-scale drivers. Agricultural Water Management, 255. https://doi.org/10.1016/j.agwat.2021.107028

Zhou, S., Wang, Y., Li, Z. et al., 2021. Characterizing spatio-temporal patterns of multi-scalar drought risk in mainland China. Ecological Indicators, 131: 108189. https://doi.org/https://doi.org/10.1016/j.ecolind.2021.108189

Table 1. Nine basins of China

Codes	Name	Areas ($10^4 km^2$)	Average annual soil moisture ($m^3 \cdot m^{-3}$)	Average annual precipitation (mm)
R1	Songhua and Liaohe River Basin	124	0.938	538
R2	Haihe River Basin	32	0.777	520
R3	Huaihe River Basin	32	0.910	874
R4	Yellow River Basin	80	0.897	479
R5	Yangtze River Basin	180	1.178	1060
R6	Pearl River Basin	57	1.172	1539
R7	Southeast Basin	24	1.229	1736
R8	Southwest Basin	85	1.078	747
R9	Continental Basin	334	0.548	166

Table 2. Drivers in the study

Categories	Drivers	Symbols
Meteorological factors	Potential evapotranspiration	Ep
	Actual evapotranspiration	E
	Precipitation	P
	Vapour pressure deficit	VPD
	Aridity index	AI
Terrestrial factors	Vegetation leaf area index	Laihv, Lailv
	Digital elevation model	dem
	Clay content	clay
	Sand content	sand
	Silt content	silt
Circulation factors	Correlation between precipitation and circulation factors	P_AMO, P_AO, P_NAO, P_Nino, P_PDO, P_SOI
	Correlation between temperature and circulation factors	T_AMO, T_AO, T_NAO, T_Nino, T_PDO, T_SOI

Table 3. Optimal parameters of the XGB model

		R1	R2	R3	R4	R5	R6	R7	R8	R9
colsample_bytree	Duration	0.7	0.6	0.8	0.7	0.6	0.8	0.7	0.9	0.7
	Severity	0.6	0.8	0.7	0.7	0.9	0.8	0.7	0.9	0.9
learning_rate	Duration	0.10	0.10	0.15	0.15	0.10	0.05	0.15	0.05	0.10
	Severity	0.05	0.10	0.05	0.10	0.15	0.10	0.05	0.05	0.15
max_depth	Duration	3	3	4	3	4	3	3	3	3
	Severity	4	7	3	3	3	3	3	3	4
n_estimators	Duration	500	500	500	500	500	500	300	500	500
	Severity	500	200	500	400	500	400	500	500	500
subsample	Duration	0.6	0.8	0.6	0.6	0.6	0.6	0.7	0.6	0.9
	Severity	0.6	0.7	0.8	0.6	0.6	0.6	0.6	0.6	0.8
R^2	Duration	0.95	0.83	0.79	0.93	0.91	0.82	0.90	0.93	0.92
	Severity	0.93	0.83	0.88	0.89	0.92	0.78	0.88	0.91	0.89
RMSE	Duration	0.37	0.52	0.31	0.31	0.29	0.19	0.14	0.48	0.78
	Severity	0.49	0.55	0.28	0.40	0.30	0.23	0.17	0.62	1.04

Table 4. Importance of driving factors on drought severity

R1		R2		R3		R4		R5		R6		R7		R8		R9	
factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value
Ep	0.858	Ep	0.636	Ep	0.354	Ep	0.806	Ep	0.227	Ep	0.132	Ep	0.132	AI	0.621	Ep	1.834
E	0.324	dem	0.256	P_Nino	0.189	Lailv	0.287	E	0.194	E	0.129	E	0.104	Laihv	0.401	VPD	0.401
P	0.222	P	0.164	P_AO	0.146	P	0.111	Lailv	0.140	P	0.099	P	0.060	Ep	0.283	P	0.395
AI	0.128	VPD	0.075	AI	0.060	dem	0.087	P	0.126	P_PDO	0.058	P_Nino	0.045	T_NAO	0.139	Lailv	0.352
P_AMO	0.106	Lailv	0.055	P	0.057	E	0.078	Laihv	0.126	P_NAO	0.051	T_SOI	0.023	Lailv	0.121	T_PDO	0.285
P_AO	0.093	P_PDO	0.041	VPD	0.046	AI	0.074	P_SOI	0.125	dem	0.040	T_AO	0.020	VPD	0.095	T_AMO	0.237
T_NAO	0.067	E	0.028	dem	0.036	T_AO	0.058	T_PDO	0.112	T_AMO	0.035	P_AMO	0.019	dem	0.093	T_NAO	0.234
P_NAO	0.066	clay	0.022	T_NAO	0.028	P_PDO	0.045	P_Nino	0.102	T_AO	0.029	Lailv	0.019	P	0.080	T_AO	0.231
T_PDO	0.063	AI	0.021	E	0.025	P_NAO	0.043	VPD	0.066	P_AO	0.028	Laihv	0.017	T_Nino	0.071	dem	0.149
VPD	0.056	Laihv	0.020	T_AO	0.021	Laihv	0.041	T_AO	0.046	P_Nino	0.026	T_NAO	0.016	E	0.067	T_Nino	0.141
T_Nino	0.048	P_Nino	0.017	Lailv	0.020	P_Nino	0.039	P_PDO	0.044	Laihv	0.021	slit	0.013	T_SOI	0.061	P_Nino	0.107
T_AO	0.045	P_AO	0.016	slit	0.020	T_AMO	0.037	T_AMO	0.038	Lailv	0.020	P_SOI	0.011	T_AMO	0.057	E	0.093
Laihv	0.043	P_AMO	0.013	T_AMO	0.018	P_AO	0.032	T_Nino	0.032	T_NAO	0.018	dem	0.011	P_AMO	0.057	P_SOI	0.090
T_AMO	0.042	P_NAO	0.012	T_Nino	0.015	T_Nino	0.028	P_AO	0.025	T_SOI	0.017	P_PDO	0.010	P_NAO	0.055	T_SOI	0.081
Lailv	0.041	slit	0.010	T_PDO	0.015	T_NAO	0.025	dem	0.024	T_Nino	0.016	P_AO	0.010	P_PDO	0.051	sand	0.077
dem	0.033	P_SOI	0.009	P_SOI	0.013	clay	0.022	AI	0.024	P_SOI	0.015	T_PDO	0.010	P_AO	0.039	slit	0.076
P_SOI	0.032	T_Nino	0.009	clay	0.012	P_AMO	0.021	T_SOI	0.021	AI	0.013	VPD	0.010	T_AO	0.036	P_AMO	0.074
P_Nino	0.026	T_PDO	0.007	P_PDO	0.010	T_SOI	0.018	P_AMO	0.021	VPD	0.012	clay	0.008	P_Nino	0.025	P_NAO	0.069
T_SOI	0.025	T_AO	0.007	P_NAO	0.010	T_PDO	0.017	T_NAO	0.021	P_AMO	0.010	T_Nino	0.006	T_PDO	0.023	P_AO	0.059
P_PDO	0.023	T_SOI	0.006	Laihv	0.010	P_SOI	0.015	P_NAO	0.021	T_PDO	0.009	P_NAO	0.006	P_SOI	0.016	P_PDO	0.052
slit	0.014	T_NAO	0.005	sand	0.008	VPD	0.015	slit	0.019	slit	0.009	T_AMO	0.005	slit	0.010	AI	0.049
sand	0.012	sand	0.003	P_AMO	0.008	sand	0.010	clay	0.007	clay	0.006	AI	0.005	clay	0.010	clay	0.030
clay	0.011	T_AMO	0.003	T_SOI	0.004	slit	0.007	sand	0.006	sand	0.004	sand	0.001	sand	0.009	Laihv	0.008

Table 5. Importance of driving factors on drought duration

R1		R2		R3		R4		R5		R6		R7		R8		R9	
factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value	factor	shap value
Ep	1.000	Ep	0.593	Ep	0.398	Ep	0.630	Ep	0.222	Ep	0.127	Ep	0.139	AI	0.590	Ep	1.680
P	0.292	P	0.147	P_Nino	0.138	Lailv	0.229	E	0.156	E	0.111	E	0.089	Laihv	0.365	VPD	0.386
AI	0.254	dem	0.146	P_AO	0.109	AI	0.217	P	0.142	P	0.108	P	0.081	Ep	0.275	P	0.327
T_PDO	0.142	VPD	0.086	VPD	0.058	E	0.133	T_PDO	0.136	P_PDO	0.050	P_Nino	0.057	T_NAO	0.152	Lailv	0.293
VPD	0.110	P_PDO	0.070	P	0.046	P	0.127	Lailv	0.125	T_AO	0.046	P_AMO	0.029	VPD	0.100	T_PDO	0.251
E	0.094	Lailv	0.061	Lailv	0.031	dem	0.086	Laihv	0.101	T_NAO	0.027	P_PDO	0.023	P	0.099	T_AMO	0.216
P_AMO	0.077	E	0.034	AI	0.026	P_Nino	0.066	P_Nino	0.079	P_Nino	0.026	P_SOI	0.019	Lailv	0.097	T_NAO	0.203
Laihv	0.068	Laihv	0.033	P_AMO	0.025	T_Nino	0.047	VPD	0.062	P_AO	0.025	T_AMO	0.017	T_Nino	0.084	T_AO	0.139
P_AO	0.051	AI	0.031	dem	0.023	T_SOI	0.043	P_SOI	0.056	P_NAO	0.024	T_SOI	0.015	P_AMO	0.071	dem	0.122
Lailv	0.046	P_AMO	0.029	E	0.023	VPD	0.042	T_AO	0.051	T_AMO	0.023	Lailv	0.012	dem	0.068	T_Nino	0.109
T_AMO	0.045	slit	0.025	clay	0.020	Laihv	0.042	P_PDO	0.045	dem	0.021	T_PDO	0.012	P_PDO	0.067	P_SOI	0.099
T_SOI	0.042	P_Nino	0.022	T_AMO	0.017	T_NAO	0.035	AI	0.031	Laihv	0.014	dem	0.011	E	0.054	slit	0.081
T_Nino	0.041	P_AO	0.018	sand	0.017	T_AMO	0.034	P_AO	0.026	T_Nino	0.013	P_AO	0.011	T_SOI	0.048	T_SOI	0.077
T_NAO	0.040	T_AO	0.017	T_Nino	0.014	P_AMO	0.032	T_AMO	0.025	AI	0.012	VPD	0.009	P_NAO	0.045	E	0.074
dem	0.035	T_Nino	0.017	slit	0.014	P_NAO	0.029	T_Nino	0.024	Lailv	0.011	AI	0.008	P_AO	0.039	P_AMO	0.072
P_NAO	0.028	P_NAO	0.015	T_PDO	0.012	P_PDO	0.028	dem	0.024	VPD	0.011	T_AO	0.007	T_AMO	0.038	P_AO	0.063
P_SOI	0.021	sand	0.013	P_PDO	0.010	P_AO	0.026	T_SOI	0.023	P_SOI	0.010	T_NAO	0.006	T_AO	0.035	P_NAO	0.056
T_AO	0.020	P_SOI	0.013	Laihv	0.010	P_SOI	0.022	P_AMO	0.020	T_PDO	0.008	Laihv	0.006	T_PDO	0.034	sand	0.053
P_PDO	0.017	T_AMO	0.012	P_SOI	0.010	T_PDO	0.022	P_NAO	0.020	P_AMO	0.007	T_Nino	0.006	P_Nino	0.017	AI	0.052
clay	0.017	clay	0.008	P_NAO	0.010	slit	0.020	slit	0.016	clay	0.006	clay	0.005	P_SOI	0.014	P_Nino	0.051
P_Nino	0.015	T_PDO	0.008	T_AO	0.007	T_AO	0.019	T_NAO	0.016	T_SOI	0.006	P_NAO	0.004	slit	0.013	P_PDO	0.051
slit	0.013	T_SOI	0.007	T_NAO	0.006	clay	0.013	clay	0.010	slit	0.004	slit	0.003	sand	0.008	clay	0.025
sand	0.007	T_NAO	0.007	T_SOI	0.003	sand	0.013	sand	0.006	sand	0.003	sand	0.001	clay	0.008	Laihv	0.007

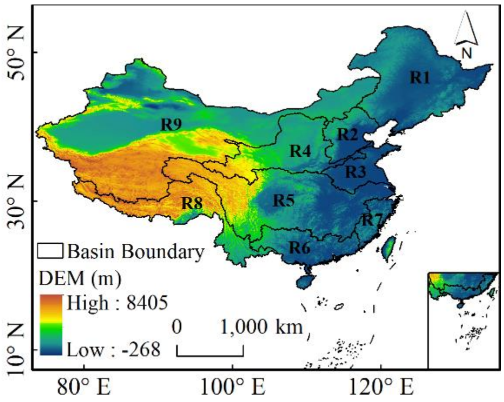

Fig.1 Basin division in China (Songhua and Liaohe River Basin (R1), Haihe River Basin (R2), Huaihe River Basin (R3), Yellow River Basin (R4), Yangtze River Basin (R5), Pearl River Basin (R6), Southeast Basin (R7), Southwest Basin (R8) and Continental Basin (R9)).

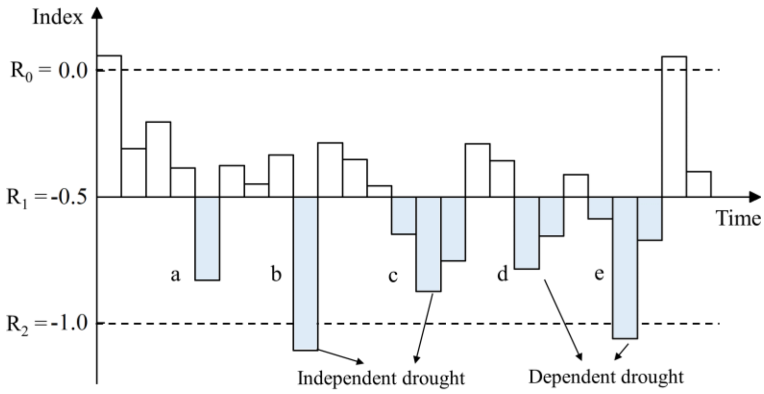

Fig. 2. Schematic diagram of drought event identification, elimination and fusion process

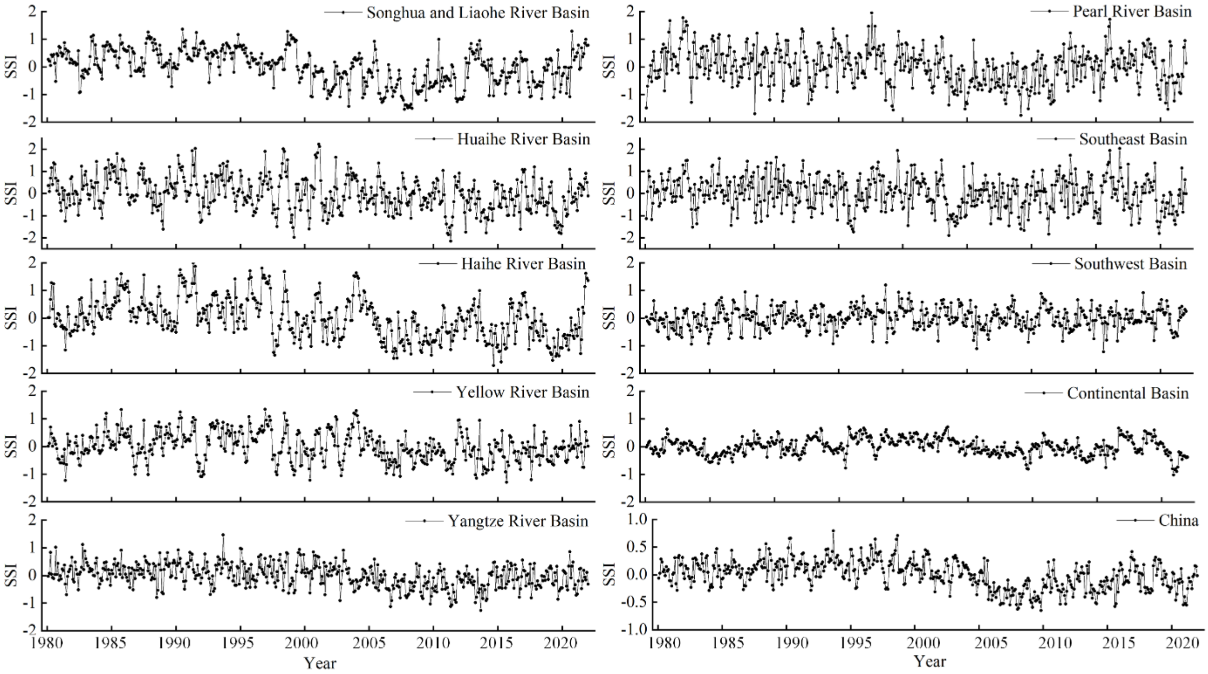

Fig.3 Temporal variation of SSI on the scale of nine basins and China from 1980 to 2021

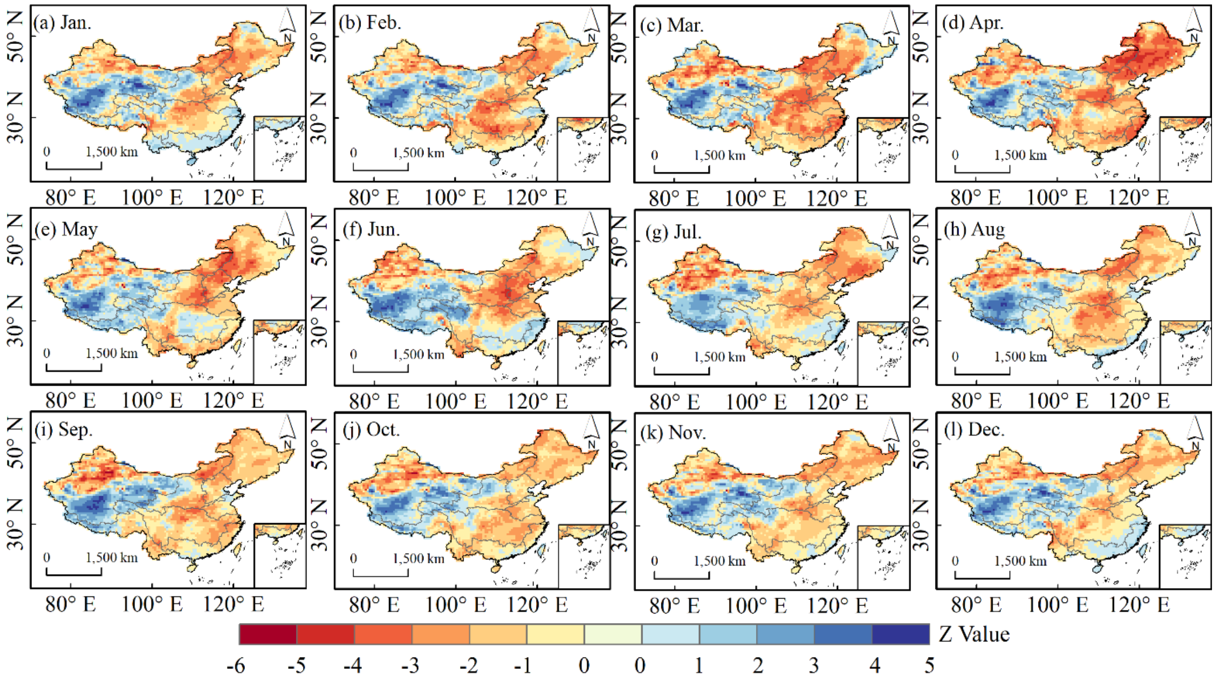

Fig.4 Spatial variation trend of monthly SSI in China from 1980 to 2021

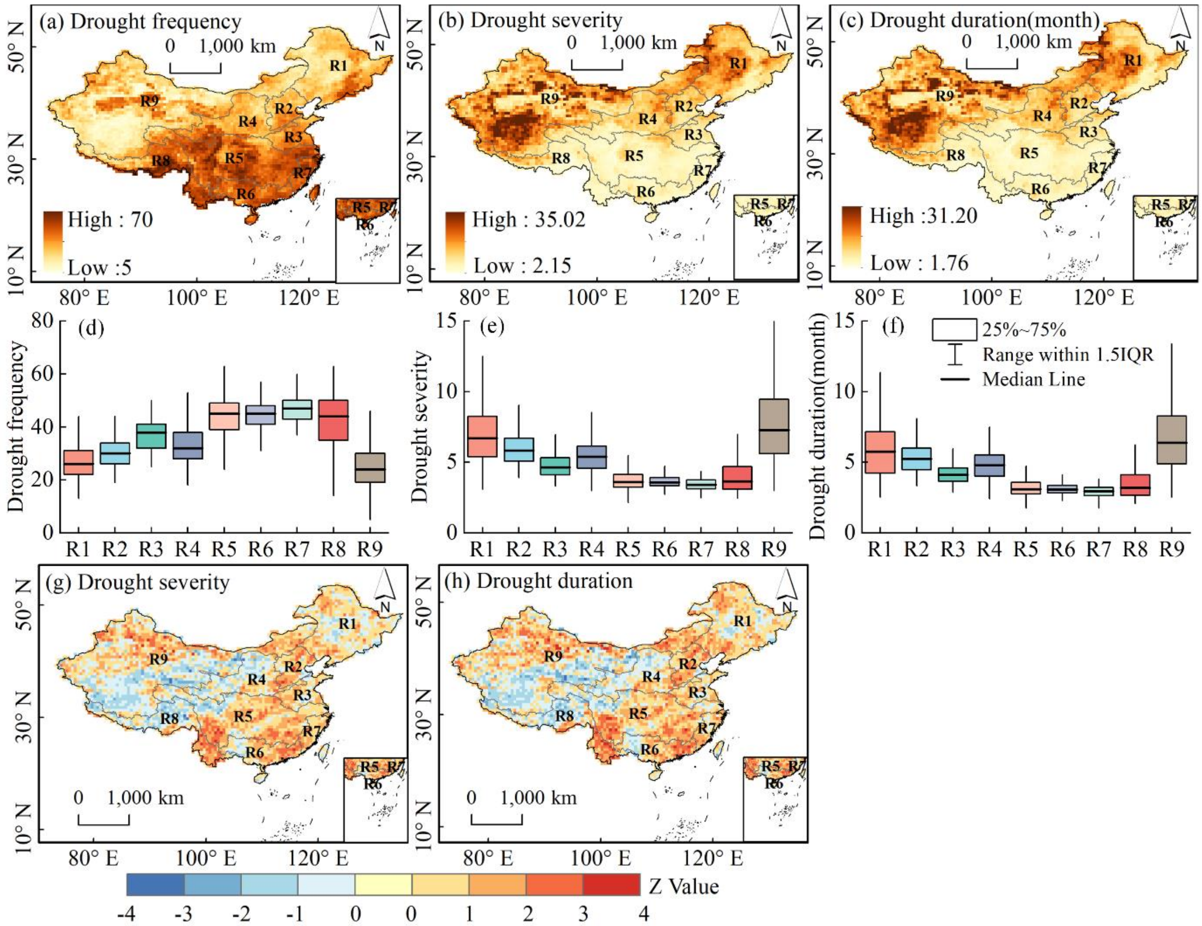

Fig. 5 Spatial distribution of agricultural drought characteristics and their changes in China (a-c) drought frequency, severity, and drought duration, (d-f) drought characteristics of the basin, (g-h) drought characteristics trend.

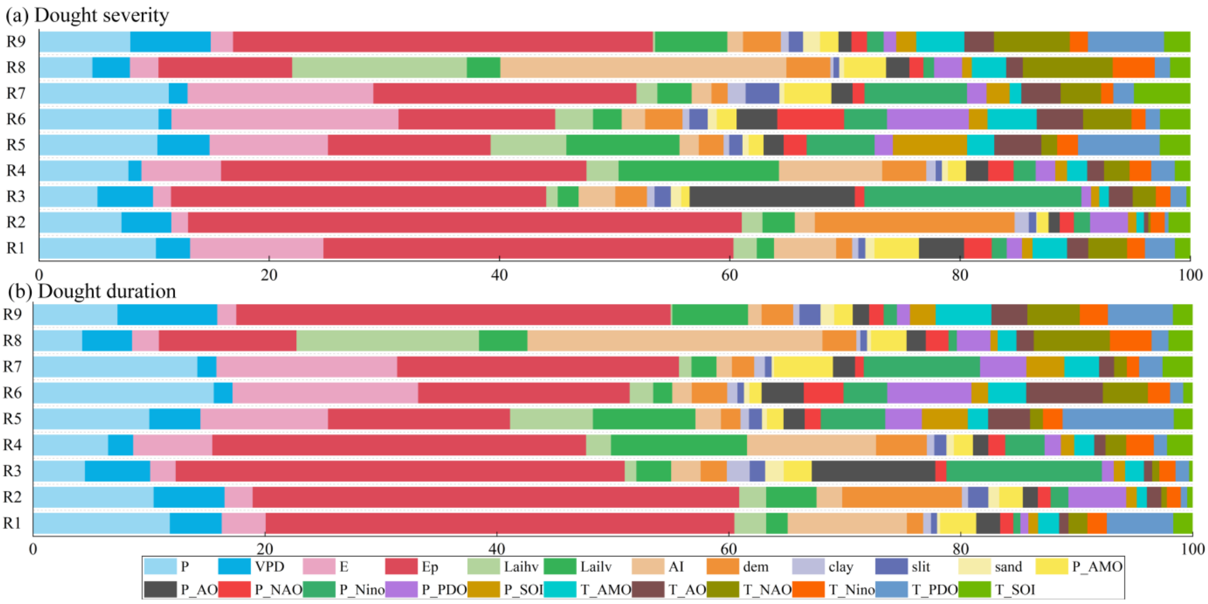

Fig. 6 Importance of drivers at the basin scale.

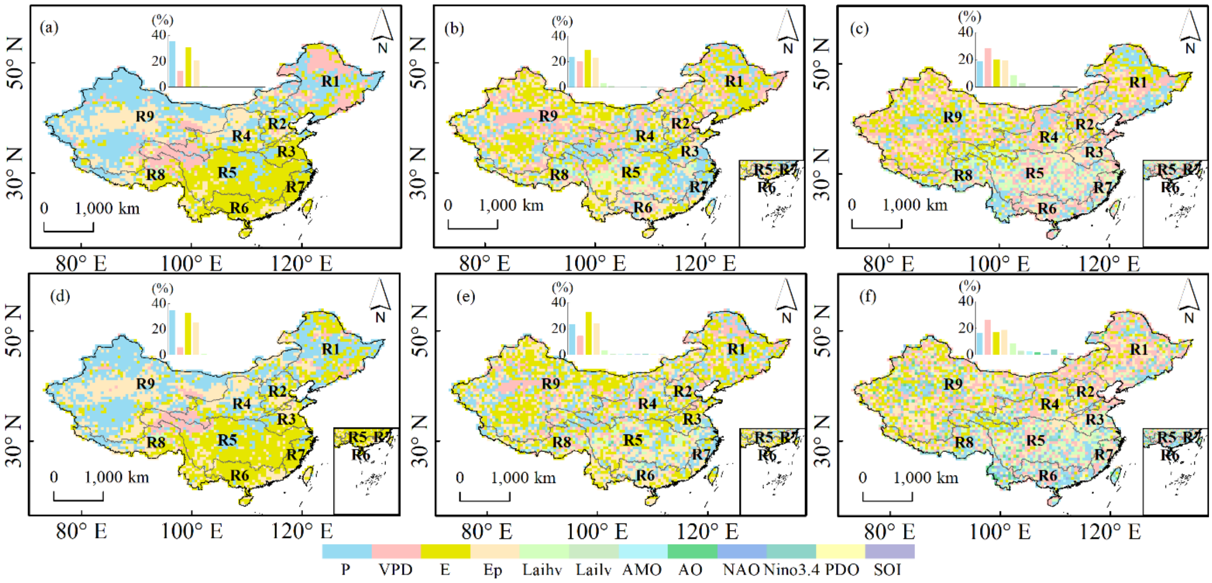

Fig. 7 Driving factors based on SHAP values. (a-c) the most, the second most, and third most important factor for drought duration, (d-e) the most, the second most, and third most important factor for drought severity.

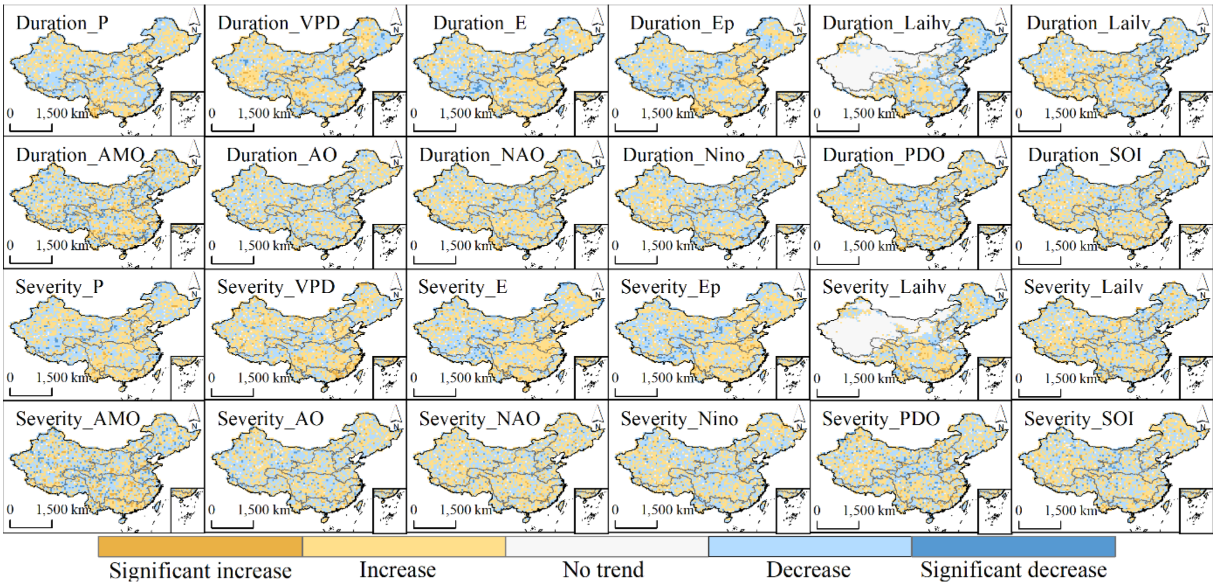

Fig.8 Trend of SHAP value of each factor

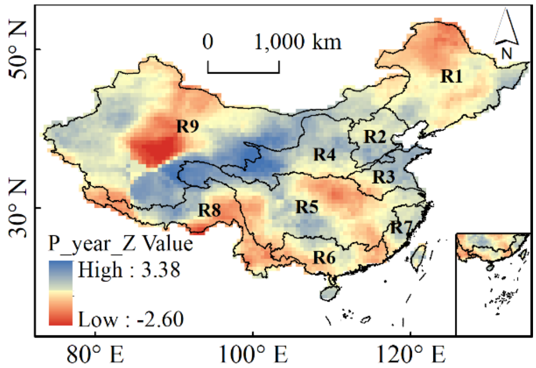

Fig.9 Spatial distribution of regional annual precipitation trends in China

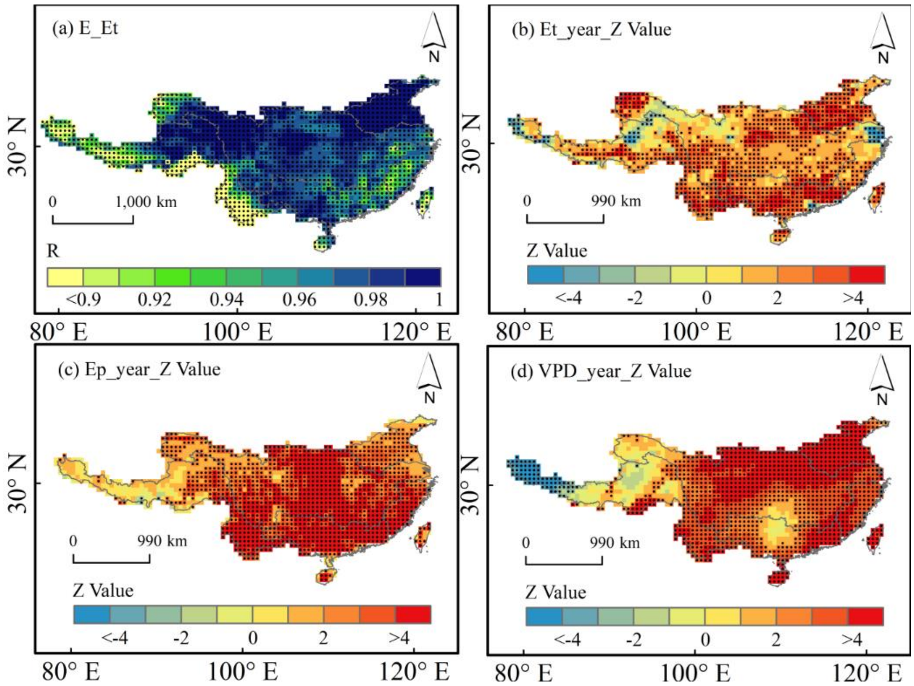

Fig. 10 Spatial distribution of factors in southern basins of China (a) correlation analysis of vegetation transpiration (Et) and actual evapotranspiration (E), (b-d) inter-annual variation of Et, E and VPD. Black dots indicate passing the significance test of P<0.05