https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Three-Dimensional Identification of Compound Heat–Drought Events in Maize in Northeast China Based on the Process-based Cumulative Heat–Drought Index (PCHDI)

Yashu Gao[1,2,3,4], Cha Ersi[1,2,3,4], Dan Chen[1,2,3,4], Sicheng Wei[1,2,3,4], Ying Guo[1,2,3,4*], Ziyuan Zhou[1,2,3,4], Zhijun Tong[1,2,3,4], Xingpeng Liu[1,2,3,4], Jiquan Zhang[1,2,3,4**], Chunli Zhao[5]

1 School of Environment, Northeast Normal University, Changchun 130024, China

2 State Environmental Protection Key Laboratory of Wetland Ecology and Vegetation Restoration, Changchun 130024, China

3 Jilin Province Science and Technology Innovation Center of Agro-meteorological Disaster Risk Assessment and Prevention, Changchun 130024, China

4 Key Laboratory for Vegetation Ecology, Ministry of Education, Changchun 130024, China

5 College of Forestry and Grassland, Jilin Agricultural University, Changchun 130024, China

*Correspondence: guoy471@nenu.edu.cn, Tel.: +86-175-4372-8607

**Correspondence: zhangjq022@nenu.edu.cn, Tel.: +86-135-9608-6467

Email address: zhangjq022@nenu.edu.cn (J. Zhang)

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Abstract

Under global warming, compound heat–drought events pose a serious threat to maize production in Northeast China. Existing indices are insufficient to capture the synergistic interactions among multiple stresses and their cumulative effects during the growing season. In this study, a novel Process-based Cumulative Heat–Drought Index (PCHDI) was developed based on soil–plant–atmosphere continuum (SPAC) theory by integrating the Vegetation Condition Index (VCI), Soil Temperature Index (STI), and Standardized Soil Moisture Index (SSMI).The response thresholds of maize to compound stress were optimized using the XGBoost–SHAP framework, and the intensity and synergistic effects of stresses were quantified through three-dimensional vector analysis. The spatiotemporal evolution of compound events from 1990 to 2023 was further analyzed using the 3D DBSCAN algorithm. The results show that high-intensity events are concentrated in the central and western agricultural-pastoral transition zones and the western Songnen Plain, exhibiting a "central-western concentration, peripheral diffusion" pattern. The event centroid follows a "southwest-northeast-southwest" oscillation, with the migration step length reaching a peak of nearly 900 km from 2015 to 2022. The intensification of stress propagates from the later growth stages to the earlier stages: the earliest mutation occurred in the flowering stage (2008), followed by the ear stage (2010) and the seedling stage (2015), with no significant mutation observed during the emergence stage. The annual pixel mutation rate for the flowering stage peaked between 2010 and 2018.Validation results demonstrate that the PCHDI has a stronger correlation with maize yield reduction than traditional indices, with improvements of 62% compared to empirical threshold-based indices and 33% compared to existing datasets. The PCHDI also shows high consistency with historical disaster records. This study provides a new methodological

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

framework for assessing compound climate disasters and supporting agricultural adaptation.

Keywords: Maize; Compound Heat-Drought Events; 3D DBSCAN; XGBoost

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

1. Introduction

Under global warming, the frequency, duration, and intensity of extreme climate events have shown a significant increasing trend. Among these, compound extreme events—resulting from the interaction of multiple hazards—have emerged as a key frontier in climate change research (Zscheischler et al., 2018). Compared with single extreme events, compound events exhibit more complex formation mechanisms and stronger amplification effects, leading to more pronounced impacts on ecosystem stability and socioeconomic systems (Zscheischler et al., 2020; Lesk et al., 2016). According to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), human activities have caused the global mean surface temperature to increase by approximately 1.07 °C over the past two decades, and the frequency and intensity of extreme heat and drought events are projected to further increase (Ipcc, 2023). In Northeast China, maize is predominantly cultivated under rainfed conditions, making it particularly vulnerable to climate change (Disasa et al., 2025). It has been reported that for every 1 °C increase in mean temperature during the growing season, maize yield decreases by 5.8% (Hou et al., 2021). In recent years, increasingly frequent compound heat–drought events have become the dominant abiotic stress limiting stable and high maize yields (Wang et al., 2018; Yu et al., 2023).

Regarding the impact mechanisms of compound heat–drought events on crop production, existing studies consistently indicate that the effects of heat and drought are not simply additive but involve complex synergistic interactions that jointly regulate crop physiological processes. While drought acts as a direct environmental stress affecting crop water status, its severity is strongly modulated by temperature (Olufayo et al., 1993). In other words, crop responses to one stress are influenced by the presence and intensity of other stresses. For instance, high temperatures can

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

intensify water stress by increasing transpiration and evapotranspiration, thereby elevating water consumption and either triggering drought conditions or exacerbating existing drought stress (Hao et al., 2022; Blum, 1990). Both controlled experiments and observational studies have confirmed the synergistic amplification effects of heat and drought on maize growth (Zhang et al., 2021; Shan), In addition, maize exhibits pronounced cumulative responses under such compound stresses (Guo et al., 2025).

In terms of quantification and identification methods for compound heat–drought events, research has progressed from single indices to composite indices. Early studies primarily used the Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), or Standardized Temperature Index (STI) to characterize drought and heat separately (Zhang et al., 2022; Hu et al., 2024). To capture the compound nature of both, researchers have developed composite indices based on threshold overlap, Euclidean distance, Copula function joint probability models, and other methods (Ni et al., 2024; Hosseinzadehtalaei et al., 2024; Lu et al., 2026). For example, Zhang et al. (2022) used the Copula model to identify compound events in North China (Zhang et al., 2022), Guo et al. (2023) developed a drought-heatwave compound magnitude index for the maize growing season in the Songliao Plain (Guo et al., 2023), and Lu et al. (2026) proposed a composite index based on Euclidean distance to evaluate compound heat–drought events in maize. In recent research on quantification indices, there has been a growing focus on the cumulative effects of stresses. (Lu et al., 2026). For instance, Zhang et al. (2017) introduced the Process-based Accumulated Drought Index (PADI), which integrates the evolution of precipitation, soil moisture, and vegetation conditions to cumulatively quantify the sustained impact of drought on crops (Zhang et al., 2017). Similarly, Guo et al. (2023) also incorporated cumulative effects when studying

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

compound heat–drought events (Guo et al., 2023).

In terms of event identification, traditional Run Theory defines events by identifying periods of time that continuously exceed a threshold. However, its intensity calculation is generally a static summary conducted after the event (Xu et al., 2023; Zhou et al., 2024), which limits its ability to dynamically characterize events.In recent years, hydrometeorological studies have introduced event detection methods based on point clustering, most of which have been applied to drought identification (Andreadis, 2005; Sheffield et al., 2009; Herrera - Estrada et al., 2017; Zhou et al., 2020; Zuo et al., 2022) as well as heatwaves (Fang and Lu, 2020; Sánchez-Benítez et al., 2020). A significant advantage of these methods is their ability to detect climate extremes that occur, intensify, and diminish at any location and during any season, rather than relying on area-averaged time series from specific seasons or fixed regions. This demonstrates their strong capability in capturing spatiotemporal dynamics.

Despite significant progress in the studies mentioned above, several deficiencies remain in the existing methodological framework. First, the most critical gap is the lack of a quantitative characterization of the synergistic degree of compound heat–drought events and the development of an integrated index. Current compound indices, whether based on probabilistic coupling or simple addition, are essentially separate approaches. That is, heat and drought events are first identified independently, and then their temporal overlap is determined before conducting compound analysis (Bogenreuther et al., 2025; Liu et al., 2025), These methods fail to separate and quantify the inherent synergistic amplification effect between heat and drought, and cannot distinguish between events with different levels of synergy in terms of disaster impacts. Second, existing indices generally overlook the cumulative effect of stress. Classic indices such as the SPI, SPEI, and STI, as well as

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

most compound indices, are based on daily instantaneous values and fail to incorporate the duration of stress and its nonlinear cumulative impact into their core framework (Wang et al., 2024; Guo et al., 2025). Third, the research framework remains fragmented, and a full-chain process space has yet to be established. While some studies have integrated remote sensing data, such as the Normalized Difference Vegetation Index (NDVI), it is typically used as a validation indicator rather than embedded as a core component of index construction. As a result, an integrated process expression of "heat stress–drought stress–vegetation physiological response" has not been established (Kang et al., 2022; Yin et al., 2024). Fourth, traditional event identification methods face dimensional limitations. These methods struggle to dynamically couple intensity with spatiotemporal dimensions, while 3D DBSCAN, though promising, has only been applied to clustering single meteorological variables (Cammalleri and Toreti, 2023), Liu et al. (2023) used this method to identify major heatwave events in the Yangtze River Basin in 2022 (Liu, 2023), but it has not yet incorporated the synergistic degree and cumulative intensity of compound stresses as core attributes in the three-dimensional (longitude-latitude-time) clustering framework. Finally, the adaptability of threshold definition methods is insufficient. Empirical thresholds or percentile-based thresholds have not sufficiently accounted for sensitivity differences across maize growth stages or the shifts in extreme values under climate change (Liu et al., 2026a).

This study focuses on the main maize-producing area in Northeast China and presents the following contributions: (1) Constructing a Process-based Compound Heat–Drought Index (PCHDI), innovatively proposing a three-dimensional process state space of "heat stress–drought stress–vegetation response," while incorporating stress duration, multi-day cumulative intensity, and vegetation physiological response into the core of the index. This results in an integrated, full-chain

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

representation of compound stress. The index considers stress thresholds and uses machine learning methods to select recognition thresholds for compound heat–drought events at different growth stages, based on the nonlinear relationship between climate factors and maize yield loss. (2) Establishing a 3D DBSCAN framework for event identification, incorporating synergistic and cumulative attributes to dynamically and accurately identify compound event clusters in the longitude-latitude-time three-dimensional space. (3) Combining the Mann–Kendall trend test with Sen's slope estimation to systematically reveal the spatiotemporal evolution characteristics of compound heat–drought events during maize growth stages in Northeast China. This study provides new methodological tools for the fine quantification and identification of compound heat–drought events in Northeast China's maize-producing areas.

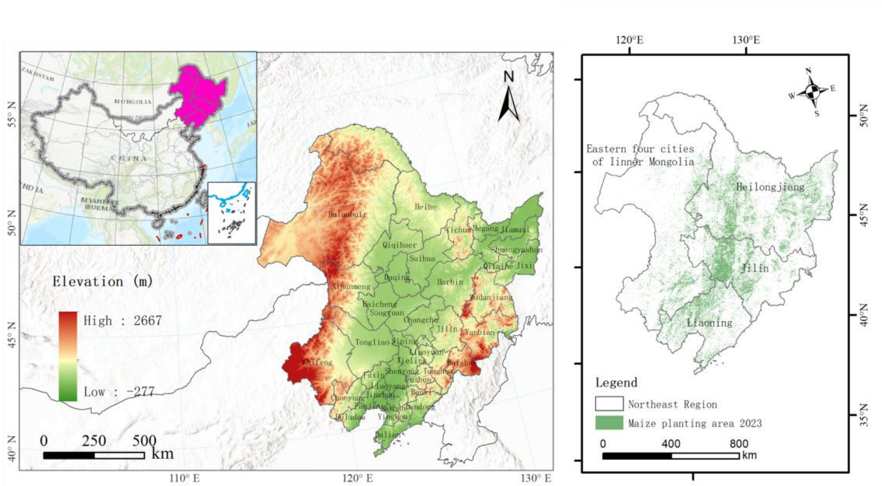
2 Study Area and Data Sources

2.1 Study Area

The study area is located in Northeast China (Figure 1), encompassing Heilongjiang, Jilin, Liaoning, and the eastern four leagues of Inner Mongolia (from north to south: Hulunbuir City, Hinggan League, Tongliao City, and Chifeng City). The longitude range is 115.51–135.78°E, and the latitude range is 38.72–53.56°N, covering an area of approximately 1.24 million square kilometers. As one of the four major black soil regions in the world (Yu et al., 2025), Northeast China's plains are flat, and the soil is fertile. This region is a critical strategic reserve for China's food security and the country's primary maize-producing area. In 2019, the total cultivated land area was 35.9 million hectares, with maize cultivation accounting for 54.4% of the total crop area. From 1980 to 2019, the proportion of maize cultivation steadily increased from 35%, making it the core of the regional planting structure (Li, 2021). However, the region frequently experiences compound

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

heat–drought events, which lead to maize yield reductions.



(a) (b)

Fig 1 Location of Northeast China and its (a) digital elevation model (DEM), (b) Maize planting area in Northeast China.

2.2 Data Sources

The meteorological data used in this study are daily data from 1990 to 2023 with a spatial resolution of 0.1°×0.1°, obtained from ERA5-Land. The daily NDVI data, covering 1981–2023, were sourced from the China seamless daily NDVI dataset released by the Original Work Team of Northwestern Polytechnical University. This dataset, based on a fusion of AVHRR and MODIS data, has a spatial resolution of 0.05° and has undergone cloud removal and missing value interpolation. The data are publicly available on the Figshare platform (Li et al., 2024). The maize yield data were obtained from the National Ecological Science Data Center and the National Earth Observation Science Data Center, supplemented by interpolation calculations based on statistical yearbook data. All raster data involved in the calculations were resampled to the same daily spatiotemporal resolution of 0.1° using ArcGIS 10.8 software. The disaster data were obtained from the Chinese

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Provincial Meteorological Disaster Catalog, which includes relevant information on disaster occurrence times, locations, and damages. DEM (Digital Elevation Model) data were retrieved from the Geospatial Data Cloud and processed based on the digital elevation model by the Chinese Academy of Sciences. Maize planting area data are based on the 30-meter resolution maize distribution map of China, derived from Landsat and Sentinel satellite imagery. The data used in this study are shown in Table 1 below:

Table1 A brief description of the data used in this research

Data type	Indicator	Time scale	Spatial scale	Source
Meteorological Data	Maximum temperature (Tmax); Root-zone soil moisture (Soil Moisture)	1990~2023, Daily	0.1°	https://cds.climate.copernicus.eu/datasets/derived-era5-land-daily-statistics?tab=download
Remote Sensing Data	Normalized Difference Vegetation Index (NDVI)	1990~2023, Daily	0.1°	https://figshare.com/collections/a_daily_gap-free_normalized_difference_vegetation_index_product_from_1981_to_2023_in_china/7002225
Maize Yield Data	GlobalCropYield5min	1990~2015, Yearly	0.1°	https://www.nesdc.org.cn
	Global Gridded Crop Production	2000~2022, Yearly	0.05°	https://noda.ac.cn/datasharing/datasetDetails/67578f822011d517fa

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

	Dataset			e6f6de
	Statistical Yearbook	2023,Yearly	-	https://www.stats.gov.cn/sj/ndsj/
Disaster Data	Time, location, and damage caused by meteorological disasters	1990~2023, Yearly	-	https://www.agridata.cn/
Maize Planting Area	The distribution maps of maize in China	2016-2023,Yearly	30 m	https://doi.org/10.6084/m9.figshare.17091653
DEM	SRTM（Shuttle Radar Topography Mission）	-	30 m	https://www.gscloud.cn/

The maize phenological data were obtained from the China Agricultural Crop Growth and Development Dataset provided by the China Meteorological Data Network (https://data.cma.cn/). The maize growing season in Northeast China is divided into four growth stages (Huang and Liu,

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

2024), as shown in Table 2, with each stage encompassing different developmental periods.

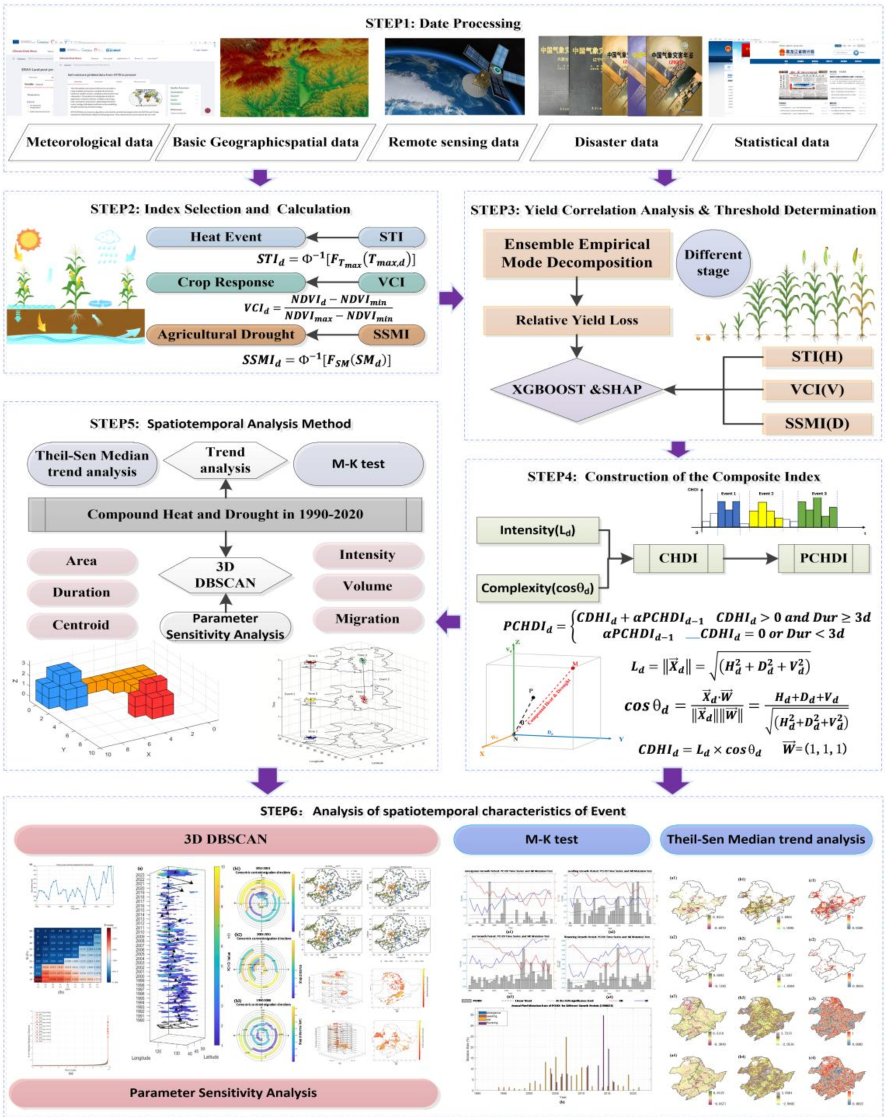
Table 2: Maize Growth Stage Division

Growth Stage	Specific Stage	Time
emergence stage	Sowing to Germination	April 20 – May 10
seedling stage	Germination to Jointing	May 11 – June 17
ear stage	Jointing to Tasseling	June 18 – July 25
flowering and maturity stage	Tasseling to Maturity	July 26 – September 10

3 Research Methods

The overall flowchart of this study is shown in Figure 2. First, multi-source data were collected and processed, followed by the selection and calculation of indicators for characterizing the process, followed by the optimization of thresholds for different growth stages. The basic indicators are then processed based on these thresholds. Next, a new index is constructed, and finally, spatiotemporal characteristics of compound heat–drought events in Northeast China from 1990 to 2023 are analyzed using the new index, DBSCAN, the Mann–Kendall trend test, and Sen's slope estimation

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

183 

184 **Fig 2** Technical Flow Chart

185 3.1 Standardized index calculation

186 The Soil–Plant–Atmosphere Continuum (SPAC) is a unified, dynamic, and continuous system (as
187 shown in Figure 3), where feedback interactions occur (Wang et al., 2019). Based on SPAC theory,
188 the daily-scale Standardized Temperature Index (STI) is selected to represent heat stress, the

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Standardized Soil Moisture Index (SSMI) is used to represent root-zone water stress, and the Vegetation Condition Index (VCI) is used to represent crop physiological status.

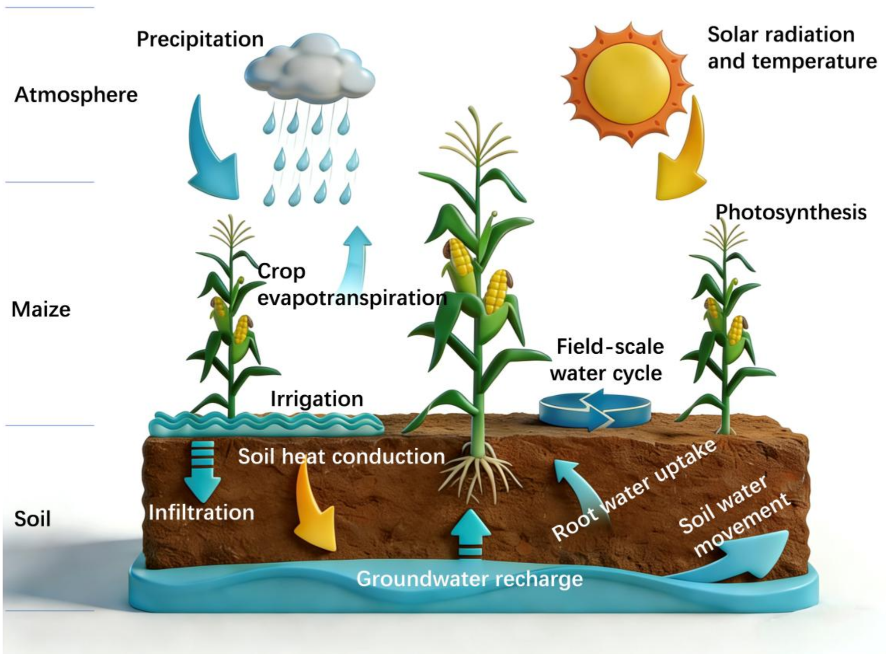


Fig 3 Schematic Diagram of the Soil–Plant–Atmosphere Continuum

（1）STI (Standardized Temperature Index)

The construction of the Standardized Temperature Index (STI) is based on the method of the Standardized Precipitation Index (SPI) proposed by McKee et al. (Mckee, 1993), which has been widely used to identify and assess climate extreme events. In this study, daily maximum temperature is used to identify temperature extreme events. The calculation formula is as follows:

$$STI_d = \Phi^{-1}[F_{T_{max}}(T_{max,d})] \quad (1)$$

Where, $T_{\max,d}$ represents the daily maximum temperature on day d, and $F_{T_{max,}}$ represents the empirical cumulative distribution function (CDF) of the daily maximum temperature during the same growth stage in the reference period (1990–2023). This function indicates the cumulative

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

probability that the daily maximum temperature during the same growth stage in the reference period is less than or equal to $T_{\max,d}$ followed by normalization. The Standardized Temperature Index, STI_d then lies within the range of $(-\infty, +\infty)$.

（2）VCI (Vegetation Condition Index)

The core of the Vegetation Condition Index (VCI) is the normalization of the current NDVI with the historical extreme values for the same period. The lower the value, the greater the vegetation stress. It is commonly used for agricultural drought and vegetation stress monitoring (Quiring and Ganesh, 2010) , The following are the calculation method and steps:

$$VCI_d = \frac{NDVI_d - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \tag{2}$$

$NDVI_{min}$ and $NDVI_{max}$ represent the minimum and maximum values during the reference period, respectively.

（3）SSMI(Standardized Soil Moisture Index)

The main characteristic of agricultural drought is the insufficient soil moisture for crop root absorption and normal transpiration, leading to reduced crop and pasture yields (Xu et al., 2020). The Standardized Soil Moisture Status Index (SSMI) uses soil moisture to describe drought (Lu et al., 2026), The calculation formula is as follows:

$$SSMI_d = \Phi^{-1}[F_{SM}(SM_d)] \tag{3}$$

Where, SM_d is the soil moisture on day d; $SSMI_d \in (-\infty, +\infty)$. The closer the value is to 1, the wetter the soil; the closer the value is to 0, the drier the soil.

3.2 Yield Decomposition and Threshold Determination

3.2.1 (EEMD) and Yield Reduction Rate

To eliminate the long-term yield growth trend caused by non-climatic factors such as

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

technological advances, agricultural management improvements, and variety improvements, and to extract the impact of climate change on the interannual fluctuations of maize yield, the method from related studies (Ceglar et al., 2016). was referenced. The maize yield time series for Northeast China was detrended. Considering the requirements of the Ensemble Empirical Mode Decomposition (EEMD) for time series length, a complete annual series from 1990 to 2023 was selected for analysis. In this study, EEMD decomposition was applied to each maize planting grid at the raster scale (0.1° × 0.1°), For the *i-th* grid, its original maize yield time series $Y_i(t)$ can be expressed as the sum of several Intrinsic Mode Functions (IMFs) and a residual term (Wu, 2005).

$$Y_i(t) = \sum_{j=1}^{n} IMF_{i,j}(t) + r_i(t) \qquad (4)$$

Where, $IMF_{i,j}(t)$ is the *j-th* Intrinsic Mode Function (IMF), which mainly reflects interannual to decadal fluctuations at different time scales; $r(t)$ is the residual term, representing the long-term trend, typically associated with non-climatic factors such as technological advancements, improved management measures, and variety updates.

In this study, the focus is on the interannual yield fluctuations driven by climatic factors, rather than long-term trend changes. Based on the dominant periodic characteristics of the IMFs, the IMF components with a main period greater than 20 years and the residual term are combined into the trend component, while the sum of the remaining IMF components is considered as the climate-driven interannual fluctuation component. The specific definition is as follows:

$$Y_{i,trend}(t) = \sum_{j\in L} IMF_{i,j}(t) + r_i(t) \qquad (5)$$

$$Y_{i,c;lim}(t) = Y_i(t) - Y_{i,trend}(t) \qquad (6)$$

Where, $Y_{i,clim}(t)$ represents the climate-driven interannual yield fluctuation component, which is represented by the high-frequency IMF components extracted by EEMD, reflecting the impact of

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

climate variability on maize yield; $Y_{i,trend}(t)$ represents the long-term trend yield for the *i-th* grid, composed of low-frequency IMF components and the residual term, reflecting the combined impact of non-climatic factors such as technological advancements and management improvements.

To quantitatively assess the contribution of climatic events to maize yield loss, the yield reduction rate (Relative Yield Loss, $RY_i(t)$) is further defined as the ratio of climate-driven yield fluctuations to the trend yield:

$$RY_i(t) = \frac{Y_{i,clim}(t)}{Y_{i,trend}(t)} \times 100\% \tag{7}$$

When $RY_i(t)<0$, it indicates that the year is a yield loss year, and its absolute value $|RY_i(t)|$represents the magnitude of the yield loss. This index effectively eliminates the influence of temporal, spatial, and agricultural technology level differences, making the yield loss comparable across different regions and years, and provides a quantitative basis for subsequent analysis of the impacts of extreme heat, drought, and their compound events on maize yield changes.

3.2.2 XGBOOST&SHAP

XGBoost (eXtreme Gradient Boosting) has recently been developed as a supervised machine learning method (Chen, 2016). The XGBoost algorithm has achieved excellent results in many machine learning challenges; it is characterized by being 10 times faster than existing popular solutions, and it can handle sparse datasets as well as process hundreds of millions of examples. SHapley Additive exPlanations (SHAP) (Lundberg and Lee, 2017) is a unified approach used to explain the output of any machine learning model, and it can visualize and describe the complex causal relationships between driving factors and the predicted outcomes. To determine the thresholds for drought and heat events in each growth stage, we trained the XGBoost model and used SHAP analysis to extract the contribution of each feature. The study used VCI, STI, and SSMI

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

as the main climate factors, combined with SHAP values to analyze different growth stages, and derived the climate index thresholds corresponding to each growth stage. SHAP values help us quantify the role of each climate factor in predicting crop yield loss or stress. Through SHAP analysis, we identified the best relationship between the climate factors and crop stress, and determined the corresponding thresholds.

In this study, the model's identification performance was evaluated using the coefficient of determination (R^2), and the expression for the validation index is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \qquad (8)$$

Where, y_i represents the observed value, $\hat{y}_i$ represents the predicted value from the model, and n is the sample size. $\bar{y}_i$ represents the mean of the observed values. The R^2 value ranges from 0 to 1, with values closer to 1 indicating a better model fit.

3.3 Construction of the Composite Index: PCHDI

Based on thresholds, we first screened the values of the three climate factors—VCI, STI, and SSMI: for VCI and SSMI, pixels greater than the set threshold were excluded. For STI, pixels smaller than the set threshold were excluded. Subsequently, we performed normalization on the screened data and flipped the VCI and SSMI data, so that higher values indicate more severe heat-drought stress. Since lower values of VCI and SSMI indicate more severe drought stress, we flipped VCI and SSMI using the formula 1-current value to ensure consistency in the calculation results of all indicators, meaning "the larger the value, the more severe the heat-drought stress." These indices are used to represent atmospheric heat stress, root-zone soil moisture status, and vegetation physiological response, and are then uniformly mapped to the three-dimensional process state space. As shown in the figure 4:

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

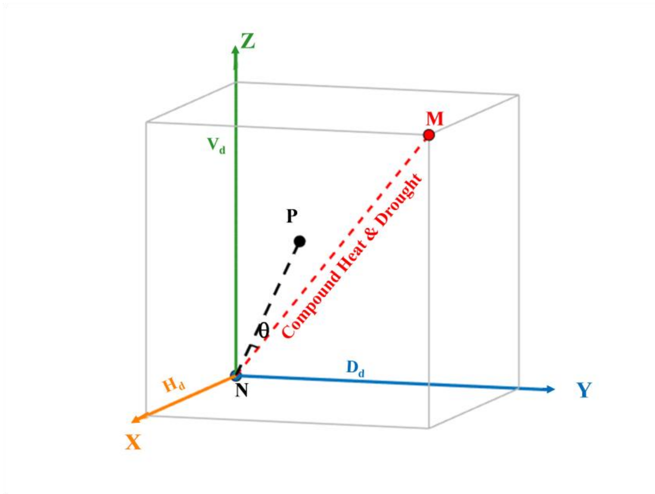


Fig 4 Schematic Diagram of Composite Index Construction

The definition of heat, vegetation, stress and drought magnitude are are defined as:

$$H_d = STI_d \tag{9}$$

$$V_d = 1 - VCI_d \tag{10}$$

$$D_d = 1 - SSMI_d \tag{11}$$

Where, $H_d \in [0,1]$, the larger the value, the more extreme the heat stress. $D_d \in [0,1]$, the larger the value, the more extreme the drought stress. $V_d \in [0,1]$, the poorer the crop growth, the larger the V_d.

Thus, the position of each day in the three-dimensional state space can be represented as a vector:

$$\overrightarrow{X_d} = (H_d, D_d, V_d) \tag{12}$$

Where $Ht, Dt, Vt \in [0,1]$. The origin $N(0,0,0)$ represents no significant heat-drought stress, and the point $M(1,1,1)$ represents the extreme values of heat, drought, and vegetation compound stress, corresponding to the line between point N and point M in the diagram (Compound Heat & Drought direction).

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

In the three-dimensional state space, the total intensity of compound heat-drought stress can be characterized by the magnitude of the vector $\vec{X}_d$:

$$L_d = \|\vec{X}_d\| = \sqrt{(H_d^2 + D_d^2 + V_d^2)} \tag{13}$$

Where $\|\vec{X}_d\|$ is the Euclidean norm. Since $H_d, D_d, V_d \in [0,1]$, $Ld \in [0, \sqrt{3}]$. When L_d approaches 0, it indicates that the three types of stress are weak or only one mild stress exists; when L_d approaches $\sqrt{3}$, the compound stress is close to extreme, In the diagram, L_d represents the length of vector.

The severity of compound events not only depends on the magnitude of individual stresses but also on whether the different stresses occur simultaneously over time. When all three stresses are simultaneously strong, the resulting damage to crops is often greater than the sum of their individual impacts. To measure this "synergistic effect," this study defines the synergy index X_d as the sine of the angle between the unit direction vector $\vec{W}$=(1,1,1) and the vector:

$$\cos\theta_d = \frac{\vec{X}_d \cdot \vec{W}}{\|\vec{X}_d\|\|\vec{W}\|} = \frac{H_d + D_d + V_d}{\sqrt{(H_d^2 + D_d^2 + V_d^2)}} \tag{14}$$

When $\vec{X}_d$=（0,0,0,）,we define $\cos\theta_d$=0. The value of $\cos\theta_d$ ranges from [0,1]. When $\cos\theta_d \approx 1$, he magnitudes of the three stress components are similar, and the compound heat-drought event exhibits a highly synchronized characteristic. When $\cos\theta_d$ is small (close to 0), it indicates that the stress is mainly dominated by a single factor. For example, there may only be heat stress, while soil moisture and vegetation conditions are close to normal, and at this point, the compound effect is relatively weak. The angle between the vector from N to P and the vector from N to M in the diagram represents $\cos\theta_d$.

Considering both the overall intensity of compound stress and the degree of multi-factor synergy, the daily-scale Compound Drought and Heat Index (CDHI) is constructed as follows:

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

$$CDHI_d = L_d \times \cos\theta_d \tag{15}$$

Since $L_d \in [0, \sqrt{3}]$, and $\cos\theta_d$ ranges from [0,1], it follows that $CDHI_d \in [0, \sqrt{3}]$, The larger the value, the more synchronized the three types of stress are, and the greater the intensity.

To represent the process accumulation effect of compound heat-drought events during maize growth stages, this study constructs the Process-based Cumulative Heat-Drought Index (PCHDI) based on the daily-scale instantaneous compound stress index $CDHI_d$. According to Run Theory, continuous sequences of CDHI>0 lasting no less than 3 days（Dur≥3d）are defined as the effective occurrence period of a compound heat-drought process (Zhou et al., 2024; Guo et al., 2023). Only when the process meets this continuity constraint is the daily $CDHI_d$ allowed to make a "new contribution" to the cumulative process. Considering the lag in crop system recovery from stress, the cumulative effect does not instantly disappear after the stress ends. This study uses an exponential decay form to describe the process tail and gradual recovery. The specific recursive form is:

$$PCHDI_d = \begin{cases} CDHI_d + \alpha PCHDI_{d-1} & CDHI_d > 0 \; and \; Dur \geq 3d \\ \alpha PCHDI_{d-1} & CDHI_d = 0 \; or \; Dur < 3d \end{cases} \tag{16}$$

Where $\alpha \in (0,1)$ is the memory coefficient. The above equation indicates that during the effective occurrence period, the current stress has the greatest contribution weight to the cumulative process, while the influence of historical stresses decreases exponentially. When the stress temporarily alleviates or the process does not meet the continuity constraint, $PCHDI_d$ no longer accumulates new $CDHI_d$, and instead only decays exponentially by α\alphaα, to represent the gradual recovery of the crop system from past stress effects. To avoid interference from very weak tails in event identification, when $PCHDI_d$ decays to a very small threshold（e.g., $PCHDI_d < \varepsilon$）, it can be set to zero, indicating that the cumulative effect has largely dissipated. Currently, ε=0. 05

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

The effective memory length of the exponential decay model is approximately:

$$\gamma = \frac{1}{1-\alpha} \tag{17}$$

When α=0.95,$\gamma \approx 20days$, This timescale is on the same order of magnitude as key growth window periods for maize, such as tasseling—flowering—early grain filling (approximately 15–25 days), and is capable of covering the cumulative and lag effects of compound stress during the critical stages of yield formation. It should be noted that, γ represents the effective decay scale of historical influences, and does not mean that the effects completely disappear after this period. Rather, it indicates that their weight has significantly decreased, reflecting the gradual recovery characteristics of the crop system from early stresses.

3.3 Identification and Spatiotemporal Feature Analysis of Compound Heat-Drought Events

3.3.1 3D DBSCAN

DBSCAN is an unsupervised machine learning clustering algorithm that relies on density for clustering. It identifies points exceeding a specified density threshold by calculating the number of points within a specified radius and then classifies them into clusters (Birant and Kut, 2007; Sun, 2021; Cammalleri and Toreti, 2023). DBSCAN stands out among clustering algorithms because it can recognize clusters of arbitrary shapes, distinguish noise, and perform clustering using both spatial and non-spatial data attributes without the need to predefine the number of clusters (Sander, 1998). In this study, a three-dimensional spatiotemporal extended DBSCAN algorithm is used to identify the spatiotemporal continuous clustering of drought events. First, each pixel-date combination is abstracted as a three-dimensional spatiotemporal point(x_i,y_i,t), where, x_i and y_i represent longitude and latitude, and t represents time, as shown in Figure 5. To achieve

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

spatiotemporal integrated clustering, a time scaling factor, TIME_SCALE, is introduced to convert time distance into a spatially equivalent distance, thus providing a unified metric for spatiotemporal distance. Next, based on the K-distance method and sensitivity heatmap, the core parameters of DBSCAN are calibrated: neighborhood radius ε and minimum sample number for core points MinPts. Finally, clustering is performed on the preprocessed daily PCHDI grid data, and after filtering noise points, effective compound heat-drought event clusters are obtained. The spatiotemporal distribution characteristics of the events are intuitively presented through three-dimensional voxel visualization (with cubes representing voxels, different colors distinguishing clusters, and each cluster representing an event).

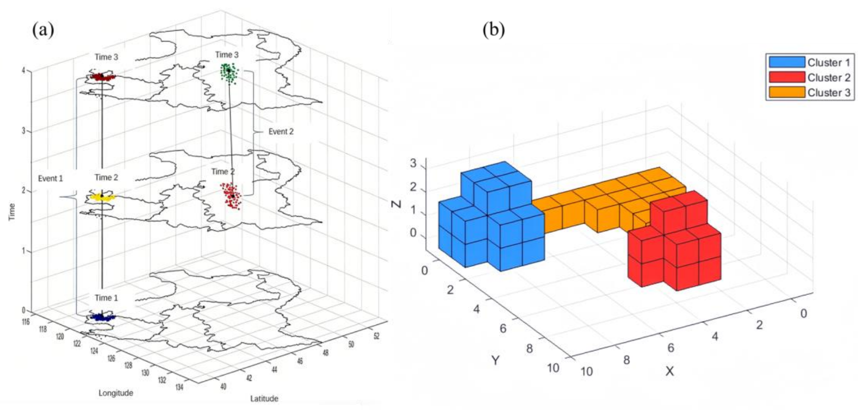


Fig 5 DBSCAN Schematic Diagram, (a) Macroscopic Identification, (b) Microscopic Identification

For each DBSCAN cluster, that is, the characteristics of the event group, the following evolution features can be calculated::

DU refers to the time period over which a compound heat-drought event cluster spans from start to end. It is formally defined as the difference between the maximum end time (T_e) and the minimum start time (T_s) of a single compound heat-drought event cluster within the same label. The

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

mathematical expression is shown in the following formula:

$$DU = T_e - T_s + 1 \tag{18}$$

A_{max} refers to the maximum impact range of the event during its development. N_t is the number of grids belonging to the cluster at time t, and A_{cell} is the area of a single grid. The formula is shown as follows:

$$A_{max} = \max(N_t \times A_{cell}) \tag{19}$$

I is used to measure the overall intensity of the compound heat-drought event. D_i is the drought index value of the i-th grid, n is the number of voxels included in the event.

$$I = \sum_{i=1}^{n} D_i \tag{20}$$

V represents the overall scale of the compound heat-drought event in the spatiotemporal-intensity structure. $D_{i,t}$ is the drought index of the i-th voxel at time t, DU is the duration, and N_t is the number of grids at time t.

$$V = \sum_{t=1}^{DU} \sum_{i=1}^{N_t} D_{i,t} \tag{21}$$

The centroid(X_i,Y_j) takes into account both the spatial extent and severity（Andreadis et al., 2005; Herrera-Estrada et al., 2017）(Andreadis, 2005; Herrera‐Estrada et al., 2017), where the non-negative value of the index, after filtering low values, is used as the weight, as shown in the formula. In the formula, Wi represents the weight factor of the i-th event（i = 1,2,3...... n）, x_i and y_i represent the longitude and latitude of the i-th voxel, D_i is the drought index value, and q_{low} is the low-value quantile threshold. This definition emphasizes the contribution of high-intensity regions in the centroid calculation.

$$\mathrm{W}_i = max(D_i - q_{low}, 0) \tag{22}$$

$$X_i = \frac{\sum_{i=1}^{n} x_i W_i}{\sum_{i=1}^{n} W_i} \tag{23}$$

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

$$Y_j = \frac{\sum_{i=1}^{n} y_i W_i}{\sum_{i=1}^{n} W_i} \quad (24)$$

d is used to calculate the spherical distance between the positions of two adjacent centroids.

$$d = 2Rarcsin\left(\sqrt{\alpha}\right) \quad (25)$$

Where, $\alpha = sin^2\left(\frac{\Delta\emptyset}{2}\right) + cos\emptyset_1 cos\emptyset_2 sin^2\left(\frac{\Delta\lambda}{2}\right)$, This formula is the Haversine formula (Hansen, 2020). θ is used to represent the migration direction of the drought structure centroid, and is further divided into eight directions: N, NE, E, SE, S, SW, W, and NW.

$$\theta = mod(atan2(y, x), 360°) \quad (26)$$

Where, $y = \sin(\Delta\lambda)\cos\emptyset_2$, $x = cos\emptyset_1 sin\emptyset_2 - sin\emptyset_1 cos\emptyset_2 \cos(\Delta\lambda)$.

3.3.2 Theil - Sen median trend analysis and the Mann - Kendall test

The Theil-Sen median method is a robust trend calculation method suitable for non-parametric statistics. The Mann-Kendall test is a non-parametric statistical test method used to determine the visibility of trends. The common features of these two methods are as follows: (1) the data does not need to follow a specific distribution; (2) they are highly resistant to outliers in the data; (3) they have high computational efficiency and are not sensitive to measurement errors and outliers. (Dinpashoh et al., 2014; Jiang et al., 2015; Serrano-Notivoli, 2018).

$$\beta = mean\left(\frac{x_j - x_i}{j - i}\right) \quad (27)$$

Where x_i and x_j represent the time series data of the grid units, respectively. If $\beta > 0$, it indicates that the time series exhibits an increasing trend; if $\beta < 0$, it indicates that the time series exhibits a decreasing trend.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign\left(x_j - x_i\right) \quad (28)$$

$$\theta = \left(x_j - x_i\right) \quad (29)$$

$$sign(\theta) = mean\left(\frac{x_j - x_i}{j - i}\right) \quad (30)$$

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

$$sign(\theta) = \begin{cases} 1 & \theta > 0 \\ 0 & \theta = 0 \\ -1 & \theta < 0 \end{cases} \quad (31)$$

$$E(s) = 0 \quad (32)$$

$$Var(s) = \left(\frac{n(n+1)(2n+5)}{18}\right) \quad (33)$$

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & S < 0 \end{cases} \quad (34)$$

Where n is the number of days. When the absolute value of Z is greater than 1.65, 1.96, and 2.58, the displayed trend passes the significance test, with reliability levels of 90%, 95%, and 99%, respectively (Table 3).

Table 3 Theil - Sen Median trend analysis and the Mann Kendall test trend categories

β	Z	Trend type	Trend features
β>0	2.58<Z	4	Very significant increase
	1.96<Z≤2.58	3	Significant increase
	1.65<Z≤1.96	2	Slight significant increase
	Z≤1.65	1	Non-significant increase
β=0	Z	0	Unchanged
β<0	Z≤1.65	-1	Non-significant decrease
	1.65<Z≤1.96	-2	Slight significant decrease
	1.96<Z≤2.58	-3	Significant decrease
	2.58<Z	-4	Very significant decrease

4 Results

4.1 Threshold Determination Results for Different Maize Growth Stages

The SHAP dependence plot generated through machine learning and the fitted curve used to

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

find the intersection (Figure 6) shows that the XGBoost model achieves fitting for each growth stage, with the highest explanatory power during the tasseling stage（R^2=0.170）, indicating that this growth stage is a key period for the impact of compound heat-drought disasters on yield. The R^2 values for each indicator during the emergence stage are 0.153743, 0.096322 during the seedling stage, 0.170213 during the jointing stage, and 0.103488 during the flowering and maturity stage. Although the overall R^2 is low, the red fitted curve (Red line) in the SHAP dependence plot clearly presents a very distinct nonlinear trend. Although the model cannot "precisely predict," it can "qualitatively identify" — meaning it captures the critical impact of indicator mutations on yield reduction. The results are shown in the figure:

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

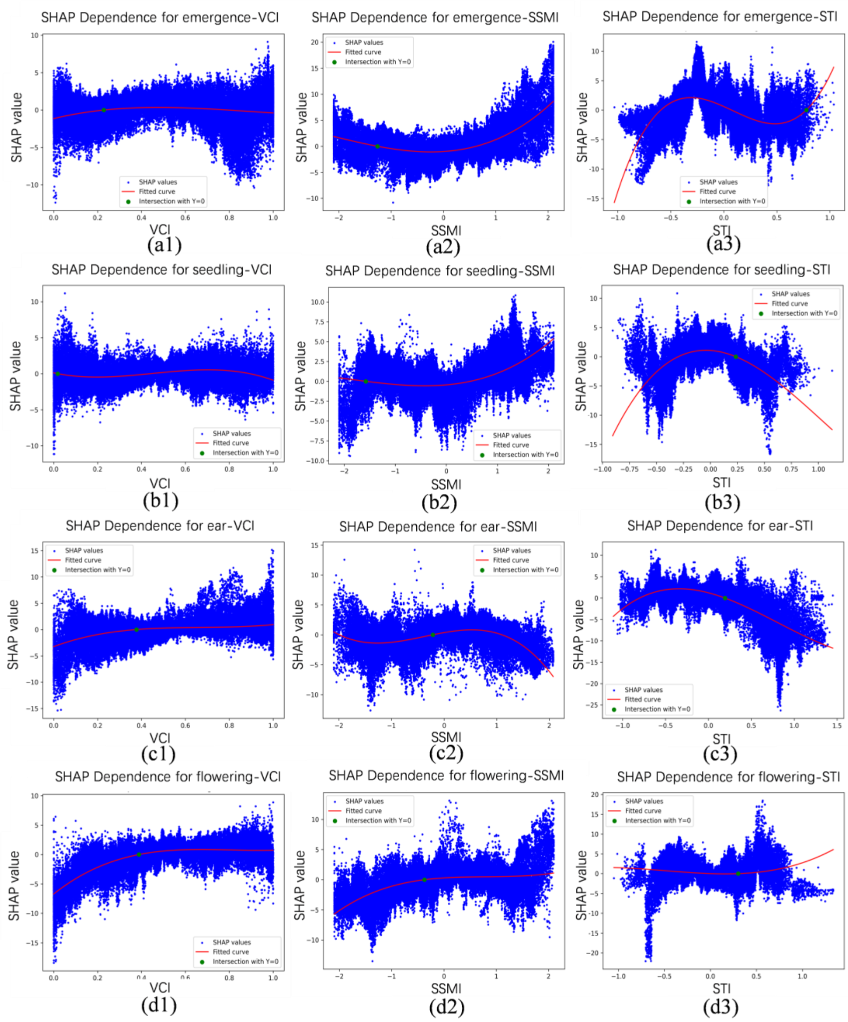


Fig 6 SHAP dependence plots for yield reduction in different growth stages and each individual index (VCI, STI, SSMI). (a1-a3) Emergence stage; (b1-b3) Seedling stage; (c1-c3) Jointing stage (Dheed); (d1-d3) Flowering and maturity stage.

Based on the intersection points identified after curve fitting, the optimal threshold results for different basic indicators during the four growth stages in Northeast China are shown in the table below:

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Table 4 Growth Stage Threshold Table

Growth Stage	VCI	STI	SSMI
emergence stage	0.228	0.778	-1.262
seedling stage	0.018	0.235	-1.579
ear stage	0.377	0.193	-0.215
Flowering and maturity stage	0.389	0.301	-0.374

4.2 Parameter Sensitivity Analysis

To achieve robust identification of compound meteorological disaster events from 1990 to 2023 in the voxel spatiotemporal domain, this study adopts a two-step method to systematically select the core parameters of 3D DBSCAN. First, the K-distance method (Figure 7a&Figure 8) is used for parameter estimation. The raw features (longitude, latitude, PCHDI index) are standardized using Z-scores. Then, the K-th nearest neighbor distance for each data point is calculated, and the K-distance curve is obtained by sorting the distances in ascending order. The distance value corresponding to the slope change (inflection point) can be used as a candidate value for ϵ.(eps) In this study, K values of 10, 15, ..., 50 were tested, totaling 9 values. Figure 8 shows the K-distance curves for different K values. All curves exhibit a typical shape of being flat at first and steep later: the flat section corresponds to dense areas, and the steep section corresponds to sparse areas. The inflection points of all curves are located at about 87.5% on the horizontal axis (point index about 437,560), indicating that approximately 87.5% of the data points are located in the dense background area, while only 12.5% of the points may correspond to sparse compound heat-drought events. This distribution aligns with the spatial characteristics of the drought index—most areas are in a non-compound heat-drought state, with compound heat-drought events being scattered. Thus,

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

the reasonable range for the neighborhood radius ϵ is preliminarily set as 2.0–4.0. Then, sensitivity heatmaps are constructed in the range of eps ∈ [2.1, 3.7] with a step of 0.2 and MinPts ∈ [10, 50] (Figure 7b), using event count as the core indicator to evaluate clustering performance. The results show that when ϵ is too small(⩽2.3), many voxels are classified as noise, leading to event fragmentation. When ϵ is too large(⩾3.6), independent events are easily merged incorrectly, underestimating disaster frequency. When ϵ is in the range of 2.3–3.3, the number of events stabilizes (1300–2000), effectively identifying statistically significant disaster events. After comprehensive analysis, this study ultimately selects eps = 2.5 和 MinPts = 20 as the optimal parameter combination. This configuration ensures effective spatiotemporal connectivity of voxels while filtering background noise through strict density constraints. A total of 1947 events are identified for the period from 1990 to 2023, avoiding both excessive clustering and event disappearance, providing a reliable data foundation for subsequent disaster pattern analysis and risk assessment.

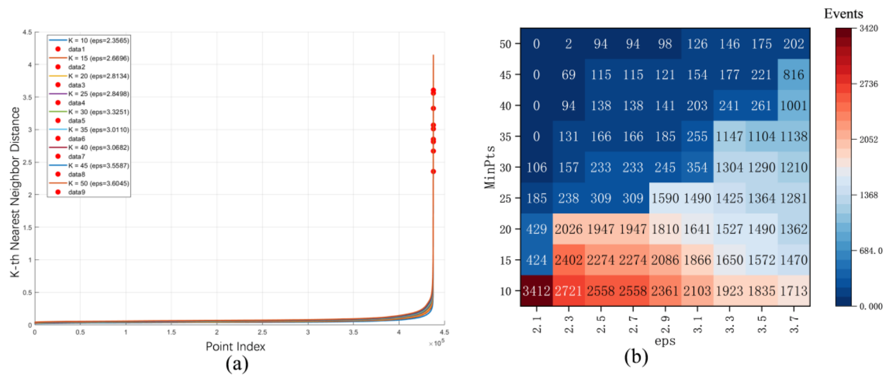


Fig 7 K-Distance Plot Collection for DBSCAN with Different Parameters (a) and Sensitivity Analysis of Event Clusters for Different Parameter Combinations (b)

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

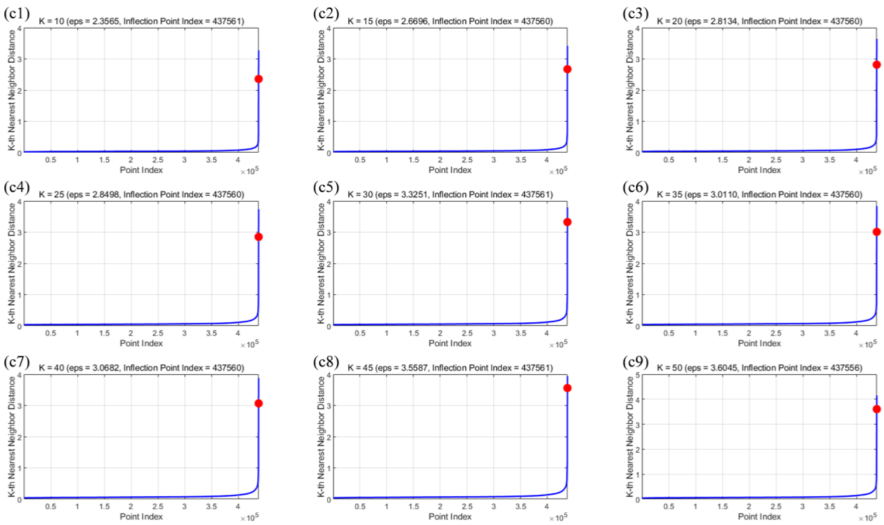


Fig 8 Scatter Subplots of K-Distance Collection with Different Parameters

4.3 Spatiotemporal Distribution and Evolution Characteristic Analysis

4.3.1 Spatial characteristics of clusters

The structural characteristics of the compound heat-drought event clusters identified using the DBSCAN method are shown in Figure 9. The four structural indicators (maximum impact area, cumulative intensity, duration, and volume) exhibit significant regional differences in space.

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

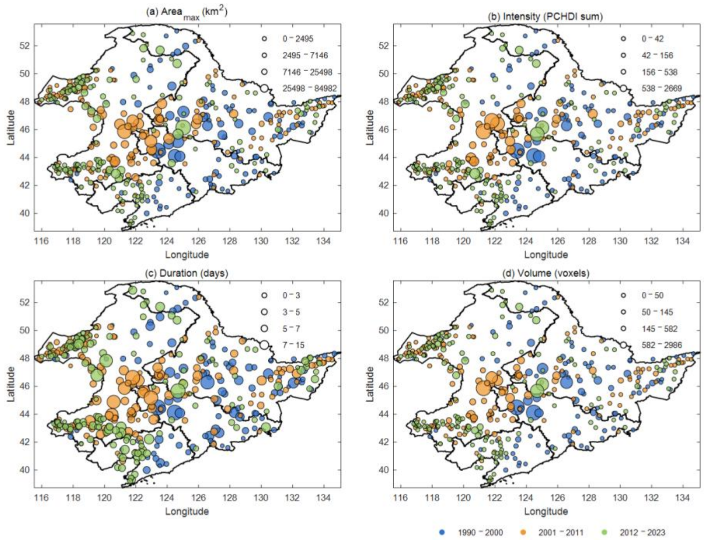


Fig 9 The maximum spatial area (Area$_{max}$), event intensity (Intensity), duration (Duration), and event volume (Volume) of compound heat-drought events in Northeast China from 1990 to 2023.

The spatial distribution of compound heat-drought events in Northeast China exhibits a distinct "western-central concentration, peripheral diffusion" pattern. The eastern agricultural-pastoral transition zone of Inner Mongolia (including Hulunbuir, Tongliao, and Chifeng) and the western Songnen Plain (around Baicheng and Daqing) are the core high-frequency areas for these events. High-value clusters of maximum impact area (Area_max), event volume (Volume), and cumulative intensity (Intensity) are concentrated in these regions, indicating that they are the key zones most prone to large-scale, long-duration compound disasters in Northeast China. In contrast, the Sanjiang Plain (northeastern Heilongjiang), southern Liaoning, and peripheral regions are dominated by small-area, short-duration localized events.The duration (Duration) of the events exhibits a clear spatial gradient: the core regions are characterized by events lasting longer than 7 days, where atmospheric circulation anomalies and positive farmland-surface feedbacks jointly sustain the compound stress over time. In peripheral areas, events tend to be short and abrupt, typically driven

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

by short-term heatwaves lasting less than 3 days.

From an interdecadal perspective, the point density of high-intensity events in the core region from 2012 to 2023 is significantly higher than that in 1990–2000, visually reflecting the continuous intensification of compound heat-drought disaster risks in Northeast China over the past decade, particularly under the backdrop of climate warming.

4.3.2 Migration characteristics of clusters

To reveal the spatiotemporal evolution of compound heat-drought events, this study analyzes the centroid migration characteristics of PCHDI events in Northeast China from 1990 to 2023 (Figure 10). The three-dimensional spatiotemporal trajectory shown in panel (a) indicates that the high-value areas of PCHDI were primarily concentrated between 1996 and 2010. The event centroid exhibited a significant "southwest-northeast-southwest" periodic oscillation pattern: from 1990 to 1997, the centroid migrated southwest, gathering in the western Songnen Plain; from 1998 to 2010, it expanded northeast, moving northward to Heilongjiang and eastern Inner Mongolia; after 2011, it again reversed and predominantly retreated southwest.The polar coordinate plots for different time periods (b1–b3) precisely validate the directional shifts: from 1990 to 2000, the migration predominantly followed the southwest–southeast direction, with step lengths fluctuating between 200–600 km, and the diffusion path was relatively dispersed; from 2001 to 2011, the migration direction significantly shifted to the northeast–northwest, with step lengths becoming more concentrated, indicating enhanced northward movement of the event core; from 2012 to 2023, the migration direction reverted to the southwest–southeast dominant direction, reflecting the centroid's return oscillation.The year-to-year step length series (Figure c) further reveals the changes in spatial instability: from 1990 to 2005, the step length fluctuation was relatively moderate, primarily within

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

the 300–600 km range; after 2006, the step length fluctuation significantly intensified, with multiple instances between 700–900 km during 2015–2020 (peaking near 900 km), far exceeding earlier levels. This indicates that since the 21st century, particularly in the past decade, the spatial volatility of compound heat-drought events has significantly increased, with the migration amplitude and instability of disaster impact areas continuing to rise.

Overall, the periodic oscillation of the event centroid intuitively reflects the spatiotemporal instability of the compound disaster's impact range, which is likely closely related to the interdecadal fluctuations of the East Asian summer monsoon, the subtropical high-pressure system, and the adjustment of extreme event-driving mechanisms under the background of regional climate warming..

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

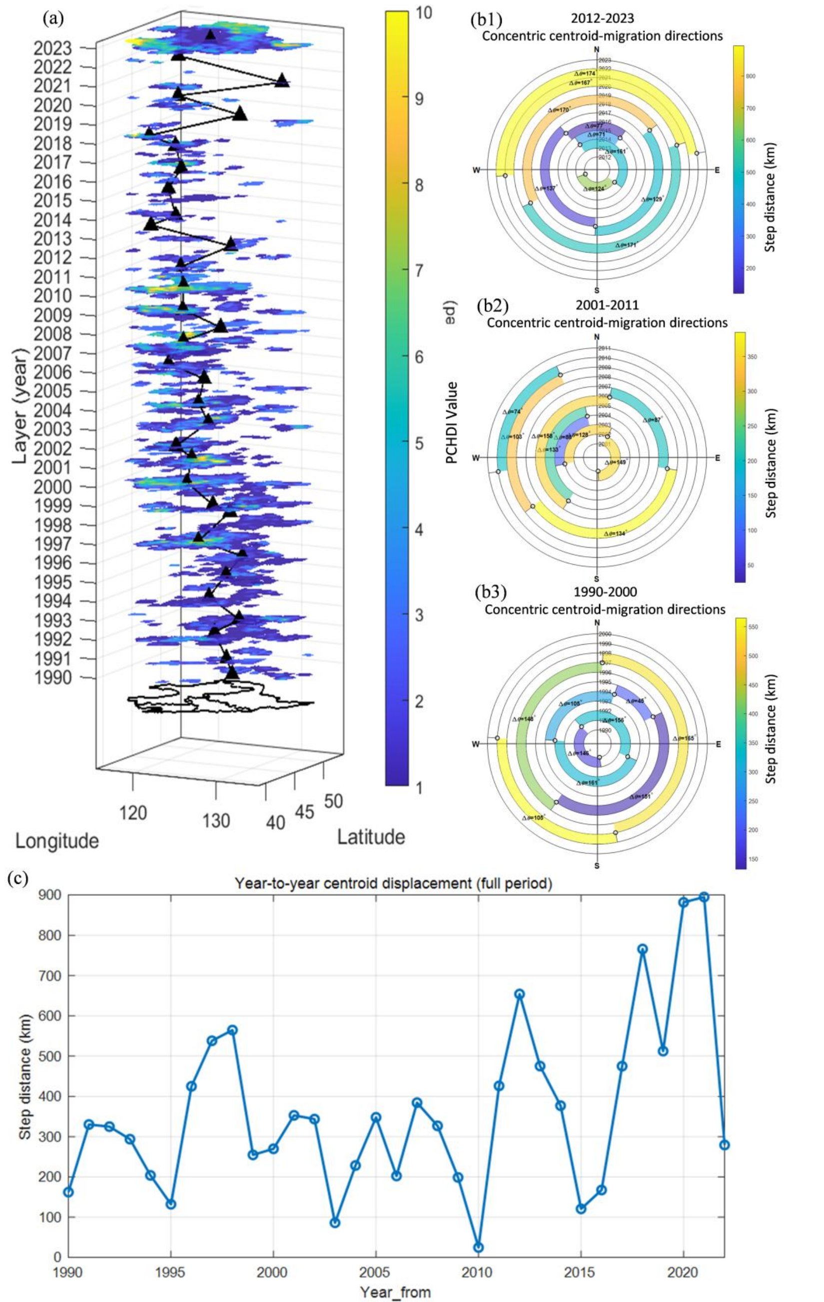


Fig 10 Centroid Migration Trajectory Plot Group.(a) The interannual centroid migration in three dimensions for compound heat-drought events in Northeast China, with the circles representing the centroid positions when the compound heat-drought event reaches its maximum impact range.(b1-b3)

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Interannual centroid migration angle and distance circular plots.(c) Interannual centroid migration distance line chart.

4.3.3 Long-term trends of PCHDI

The spatial distribution of the Sen trend slope and Z statistic for PCHDI during different maize growth stages in Northeast China from 1990 to 2023 is shown in Figure 11. The results indicate significant growth stage differences in the trend characteristics of compound heat-drought stress.

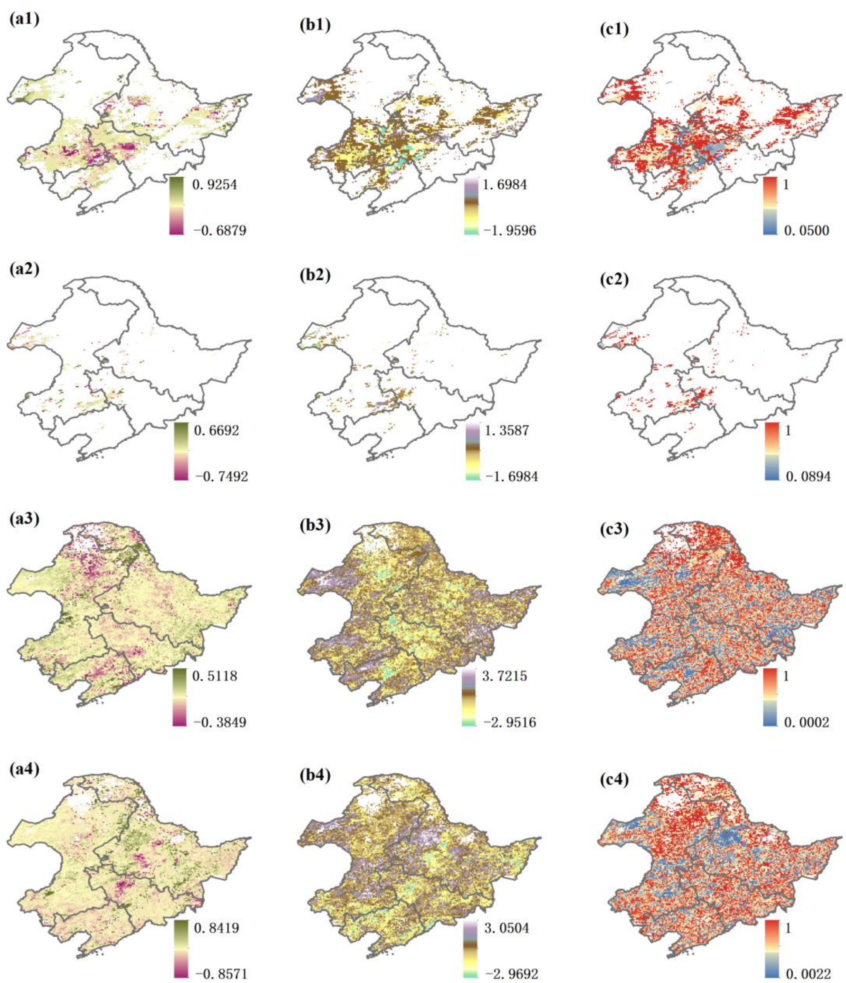


Fig 11 Sen Trend Slopes for Different Growth Stages (a1-a4), Z Statistics (b1-b4), and p

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Significance Values (c1-c4) Plot Group

During the emergence stage, the Sen slope predominantly shows weak trends, with significantly increasing areas ($p < 0.05$) scattered in the central, eastern coastal, and western edge regions. The Z statistic mostly falls between ±1.96, indicating smooth stress changes. The seedling stage has the smallest range of significant trends, with positive slope areas only scattered in the western and central regions, and no area reaching the 95% confidence significant level, making it the most stable stage of stress. During the ear stage, positive slope areas cover the western, southern, and eastern coastal regions, with the Z statistic generally exceeding 1.96. A significant decline area is present in the northern part of the central region, forming a pattern of "overall intensification, local alleviation." In the flowering and maturity stage, the positive slope area nearly covers the entire region, with the Z statistic consistently exceeding 1.96. The significantly increasing area is the largest among the four stages, representing the most significant intensification of stress. Overall, the significantly increasing areas of PCHDI gradually expanded from the early stage of scattered local distribution to contiguous distribution in the ear stage, and eventually covered the entire region in the flowering and maturity stage, showing an evolution characteristic of "gradually expanding range and continuously increasing intensity."

4.3.4 Abrupt changes of PCHDI

The Mann-Kendall (M-K) mutation detection results (Figure 12) show significant heterogeneity in the mutation time and trend intensity of PCHDI for the four maize growth stages in Northeast China from 1990 to 2023. For the emergence stage (a1), the UF curve consistently fluctuates within the ±1.96 significance threshold, with no obvious crossover between UF and UB, and the PCHDI linear trend exhibits a gradual decline, indicating that stress changes in this stage

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

are mild, with no significant mutation points, and the intensification is both the latest and weakest. For the seedling stage (a2), the UF curve rapidly rises and surpasses 1.96 after 2015, with the UF and UB curves crossing in 2015. After the mutation, the PCHDI significantly increases, and the linear trend rises slowly, reflecting the significant intensification of the compound stress response in the seedling stage after 2015. In the ear stage (a3), the UF curve remains above 1.96 after 2010, with the UF and UB curves crossing in 2010. After the mutation, the PCHDI gradually increases, and the linear trend steadily rises, indicating that stress in the ear stage has entered a phase of significant intensification since 2010. The flowering stage (a4) is the earliest to experience mutation, with the UF curve remaining above 1.96 after 2008. The UF and UB curves cross in 2008, and after the mutation, the PCHDI continuously and steadily increases, making it the stage with the earliest mutation, reflecting that the intensification of stress began the earliest and has the most sustained trend among the four growth stages.

Overall, the PCHDI mutation time follows a "later growth stages → earlier growth stages" transmission pattern: the earliest mutation occurred in the flowering stage (2008), followed by the ear stage (2010) and seedling stage (2015), with no significant mutation in the emergence stage. This indicates that the intensification of compound heat-drought stress spreads from the critical late growth stages of maize yield formation to the earlier stages. The annual pixel mutation rate (Figure b) further confirms this pattern: from 1990 to 2023, PCHDI mutation events were primarily concentrated between 2000 and 2018, highly coinciding with the historical high-frequency disaster periods in the region. Among them, the peak mutation rate for the flowering stage was the highest (close to 35%), with high values concentrated from 2010 to 2018, making it the most sensitive growth stage to stress changes. The peak mutation rate for the ear stage was second, with a broader

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

distribution between 2005 and 2015 and a more gradual evolution. The mutation rates for the emergence and seedling stages were generally low (peak < 15%), with scattered mutation years and a moderate response. This result corroborates the M-K time series, confirming that compound stress mutations are more likely to occur in the later growth stages and that the intensification is more significant.

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

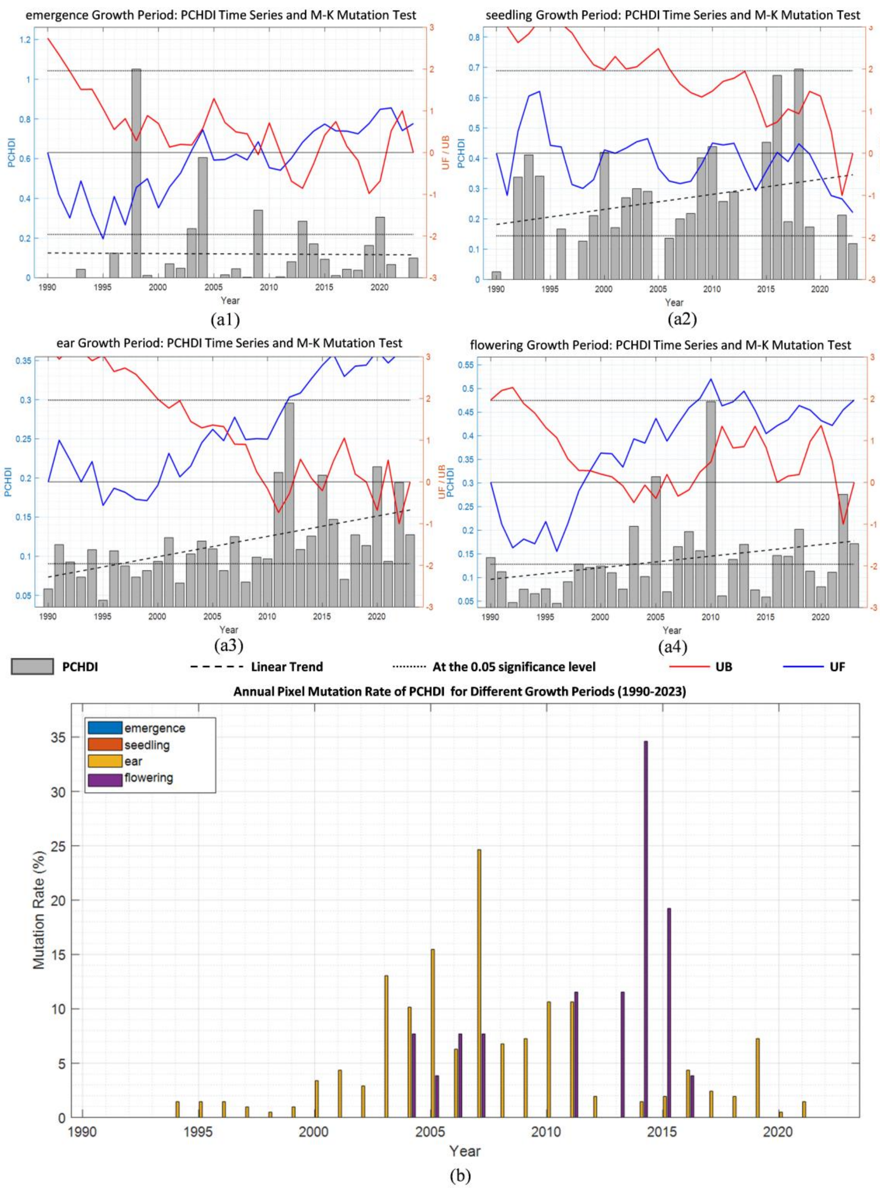


Fig 12 (a1-a4) PCHDI and M-K detection curves for different maize growth stages in Northeast China (b) Annual PCHDI mutation rates for different maize growth stages

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

4.3.5 Daily-Scale Event Visualization

To intuitively identify the dynamic evolution of compound heat-drought events in the time dimension and their spatial impact range, based on the 3D DBSCAN density clustering algorithm and the index constructed in this study, the recognition results of five typical compound heat-drought events in Northeast China (1992, 2000, 2003, 2010, 2020) were selected for three-dimensional visualization and superimposed top-down view analysis. This analysis reveals the spatiotemporal evolution and intensity distribution patterns.

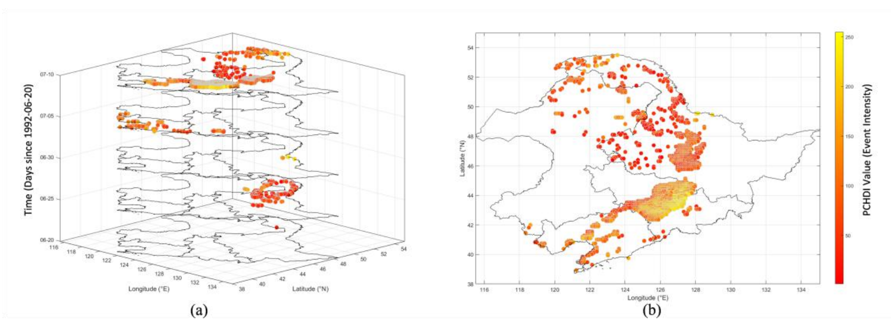


Fig 13 June 20, 1992 – July 10, 1992 Identification Visualization: (a) Three-Dimensional Stereoscopic View (b) Superimposed Top-Down View

The identification results from June 20, 1992 – July 10, 1992 (Figure 13): During this period, the compound heat-drought event was centered around the Songnen Plain. Early clustering was concentrated in the Baicheng area of Jilin and the Tongliao area of Inner Mongolia, while in the middle period, it gradually expanded to Daqing, Harbin, and Changchun in Jilin, forming a continuous spatiotemporal cluster covering the central part of the Songnen Plain. The top-down view shows that high-intensity areas are concentrated in the western part of Jilin and the southern part of Heilongjiang (around Baicheng, Daqing, and Changchun), exhibiting a clear "core-edge" gradient decrease, with a significant impact on the dryland farming areas in western Jilin and the

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

central grain production base in Heilongjiang.

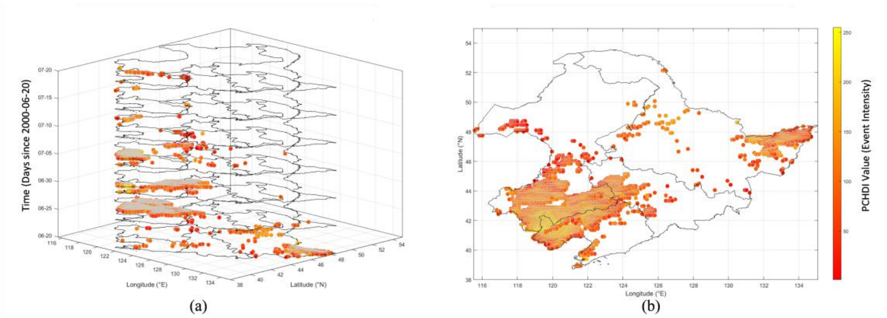


Fig 14 June 20, 2000 – July 20, 2000 Identification Visualization: (a) Three-Dimensional Stereoscopic View (b) Superimposed Top-Down View

The identification results for June 20, 2000 – July 20, 2000 (Figure 14) show a cross-regional, contiguous spread pattern. Clustering began in areas such as Fuxin and Jinzhou in Liaoning, and Chifeng in Inner Mongolia, gradually expanding northeastward to Siping and Changchun in Jilin, and eventually extending to Qiqihar and Daqing in Heilongjiang. In the top-down view, the high-intensity areas are distributed in a strip shape along the Liaohe Plain and Songnen Plain, covering central and western Liaoning, central and western Jilin, and southwestern Heilongjiang, causing sustained stress on the ecologically vulnerable areas in western Liaoning and the main grain-producing regions in central Jilin.

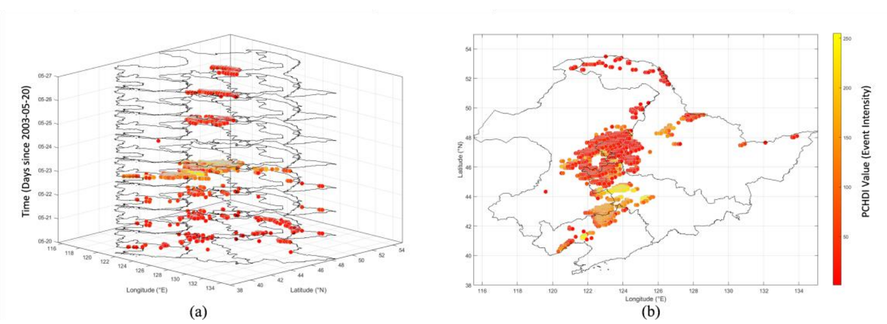


Fig 15 May 20, 2003 – May 27, 1992 Identification Visualization: (a) Three-Dimensional

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Stereoscopic View (b) Superimposed Top-Down View

The identification results for May 20 to May 27, 2003 (Figure 15) show that this event was a short-duration, localized extreme event. The clustering was highly concentrated in central Jilin and southern Heilongjiang (around Changchun, Jilin City, and Harbin), with no significant cross-regional expansion. The three-dimensional voxel plot shows that high-value voxels rapidly formed within a short period of time. In the top-down view, the PCHDI peak areas are concentrated around Changchun and Jilin City, reflecting the sudden and localized nature of compound drought in central Jilin's agricultural areas under the background of spring warming.

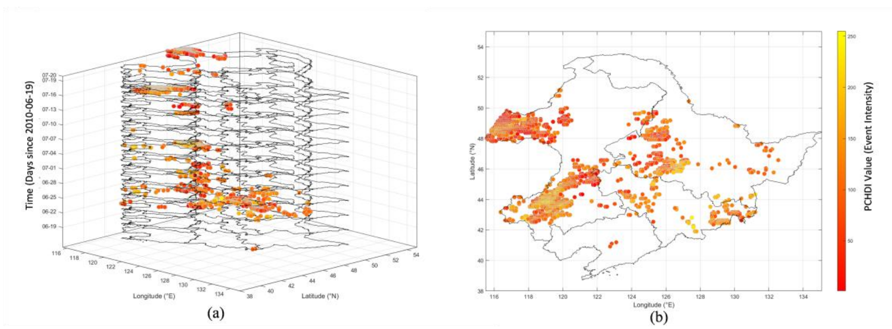


Fig 16 June 20, 2010 – July 20, 2010 Identification Visualization: (a) Three-Dimensional Stereoscopic View (b) Superimposed Top-Down View

The identification results for June 10 – July 20, 2010 (Figure 16) show a multi-core initiation and contiguous expansion pattern. Early clustering was scattered in areas such as Shenyang and Tieling in Liaoning, Siping and Songyuan in Jilin, and Daqing and Suihua in Heilongjiang. In the middle period, the clusters gradually connected to form a continuous spatiotemporal cluster covering the Songnen Plain and Liaohe Plain. In the top-down view, the high-intensity areas are scattered in the northern part of Liaoning, central and western Jilin, and southwestern Heilongjiang (around Shenyang, Siping, Songyuan, and Daqing). The multi-core distribution pattern reflects the

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

complex propagation paths of compound events in Northeast China's major agricultural areas under atmospheric circulation fluctuations.

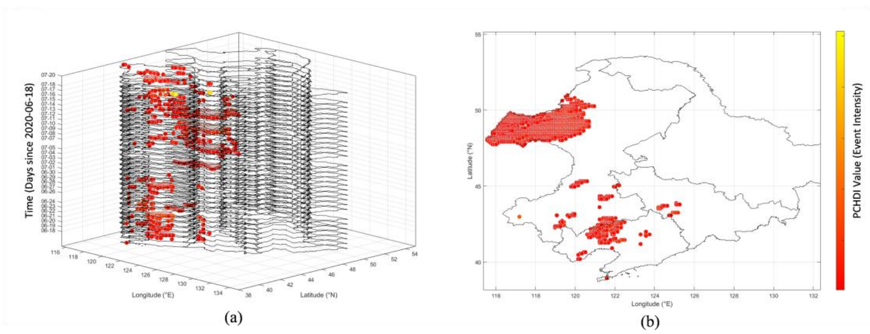


Fig 17 June 10, 2020 – July 20, 2020 Identification Visualization: (a) Three-Dimensional Stereoscopic View (b) Superimposed Top-Down View

The identification results for June 10 – July 20, 2020 (Figure 17) show a widespread high-intensity extreme pattern in the western region. The clustering is concentrated in eastern Inner Mongolia, western Liaoning, and western Jilin (around Hulunbuir, Chifeng, Fuxin, and Baicheng), and rapidly expands eastward to Qiqihar and Daqing in Heilongjiang. Figure 17 shows that the high-value clusters almost span the entire study period. In the top-down view, the PCHDI peak areas are concentrated in eastern Inner Mongolia and western Liaoning, reflecting the trend of extreme and widespread compound extreme events in the drought-sensitive areas of western Northeast China under the background of climate warming.

5 Discussion

5.1 Comparison and Historical Case Validation of PCHDI Based on Different Threshold Calculation Methods with Existing Indices

The average Pearson correlation coefficient between the PCHDI index constructed in this study and maize yield reduction is 0.2186. Based on the high-temperature and drought empirical

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

thresholds for Northeast maize, PCHDI was also calculated (PCHDI traditional , PCHDItra) with STI set to 1 as the critical value(Zhou et al., 2024), SSMI set to -0.5(Lu et al., 2026), and VCI set to 0.4(Pius et al., 2025; Jiang et al., 2021). For better comparison, this study employed the XGBOOST model to optimize the 9-year moving threshold, with the results presented in Supplementary Table S1. Subsequently, the PCHDI (PCHDI dynamic, abbreviated as PCHDIdyn) was calculated. The PCHDItra index's correlation coefficient is 0.2046, significantly higher than the traditional datasets Dheed (0.0679)(Weynants et al., 2025) , PCHDIdyn（0.1256）, SSMI (-0.0679), VCI (-0.1052), and STI (0.1240) indices. This indicates that the overall correlation strength between PCHDI and yield reduction is the best among all the indices. In terms of result stability, the standard deviation of the correlation coefficient for PCHDI is 0.1861, lower than PCHDItra(0.2034) and other traditional indices, suggesting that its correlation with yield reduction is more stable in spatial distribution. Furthermore, the proportion of significant pixels ($p < 0.05$) for PCHDI is 25.6%, and for PCHDItra it is 25.8%, both significantly higher than Dheed (10.5%), SSMI (9.0%), VCI (7.3%), STI (14.3%) and PCHDIdyn（5.6%）. Combining the average correlation coefficient and standard deviation, it can be concluded that the results from PCHDI have better statistical credibility. In terms of the significant spatial distribution (Figure 18), the significant correlation areas between each index and maize yield reduction show clear differences: PCHDI index (Figure a): The area with a significant positive correlation ($p < 0.05$, $r > 0$) has a wider spatial range, forming a continuous high-value region in the central and eastern parts of the study area. This indicates that its significant correlation with yield reduction is more universally applicable and stable in space. PCHDIdyn index (Figure b): Although the proportion of pixels with significant correlations relative to PCHDI is not high, the standard deviation of the correlation coefficient is

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

relatively small. PCHDItra index (Figure c): The proportion of significant pixels is slightly higher, but the standard deviation of the correlation coefficient is larger, and the spatial distribution is relatively scattered, making it less stable than PCHDI. Dheed index (Figure d): The proportion of significant pixels (10.5%) is significantly lower than PCHDI and PCHDItra, with significant positive correlation mainly concentrated in the center, and the spatial distribution is fragmented, indicating weaker stability in the correlation. SSMI index (Figure e): The average correlation coefficient is -0.0679, showing a negative correlation. Although it can explain yield reduction, its explanatory power is lower than PCHDI due to the negative nature of the index. VCI index (Figure f): The average correlation coefficient is -0.1052, with a more pronounced negative correlation. Although it can explain yield reduction, its explanatory power is still weaker than PCHDI. STI index (Figure g): The proportion of significant pixels (14.3%) is the highest among the traditional indices. It performs well in some areas but has a weaker average correlation coefficient (0.1240) and result stability (standard deviation 0.1911) compared to PCHDI.

Overall, the PCHDI index outperforms PCHDItra , PCDHIdyn and other traditional indices in three core dimensions: overall correlation strength, spatial significance, and result stability. It more effectively represents the driving effect of compound heat-drought stress on maize yield reduction.

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

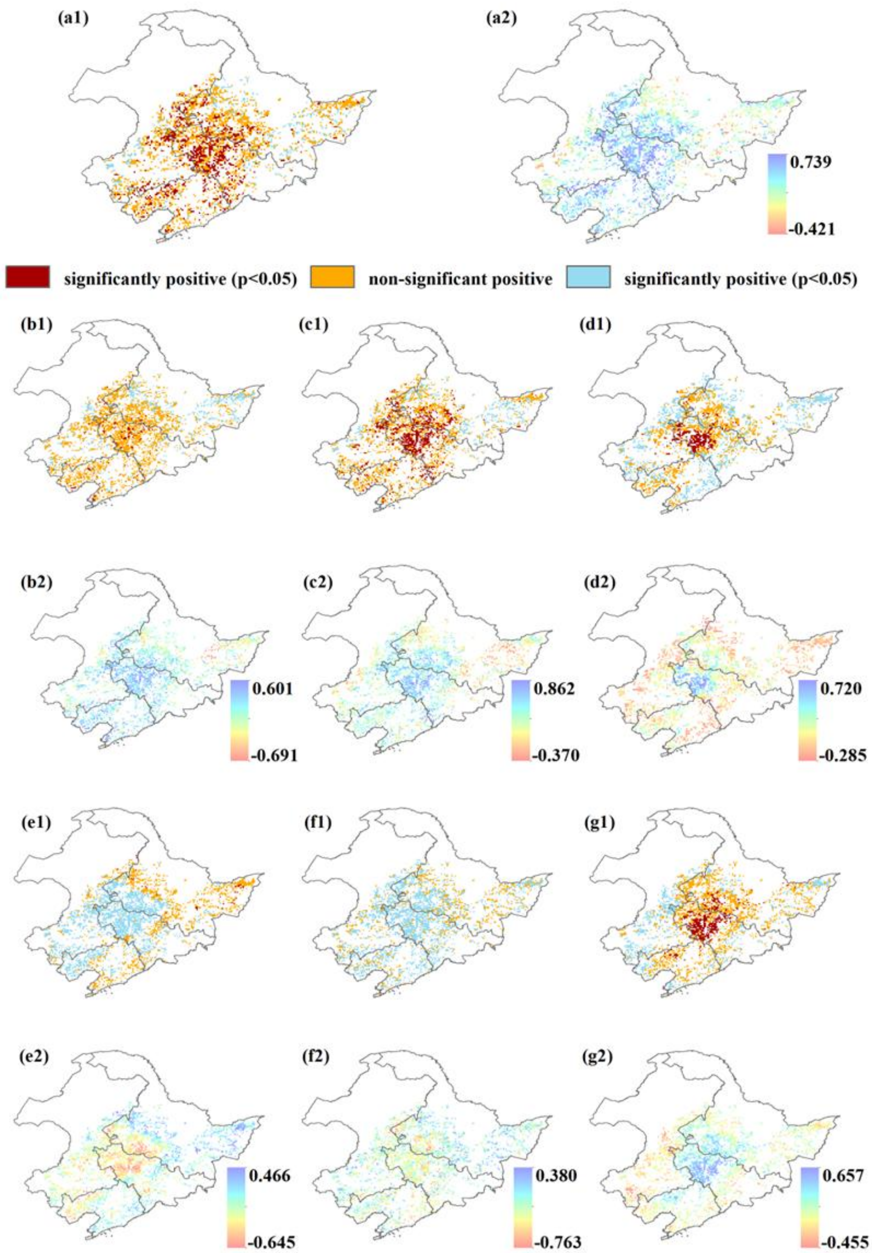


Fig 18 Comparison of correlations between yield reduction and traditional Compound Drought and Heat Event (CDHE) indices, as well as each individual index.(a1–a2) Process-based Cumulative

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Heat-Drought Index based on machine learning thresholds (PCHDI);(b1–b2) Process-based Cumulative Heat-Drought Index based on dynamic thresholds (PCHDIdyn);(c1–c2) Process-based Cumulative Heat-Drought Index based on traditional thresholds (PCHDItra);(d1–d2) Global Database for Heat-Drought Compound Extreme Events (Dheed);(e1–e2) Standardized Soil Moisture Index (SSMI);(f1–f2) Vegetation Condition Index (VCI);(g1–g2) Standardized Temperature Index (STI).

In addition, we also calculated the correlation at the site scale with the yield loss magnitude, as shown in Figure 19. PCHDI demonstrates significant advantages in both the average correlation coefficient and the proportion of significant sites. PCHDI shows a 62% higher average yield correlation than the PCHDItra based on traditional thresholds, 85% higher than the sliding threshold-based PCHDIdyn, and approximately 33% higher yield correlation than the existing Dheed dataset. These results suggest that integrating a threshold optimized over a longer time span and accounting for the cumulative process of crop damage effectively enhances the index's ability to capture compound heat-drought stress relevant to crop impacts. The stronger statistical relationship further validates the reliability and applicability of PCHDI in agricultural disaster assessment.

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

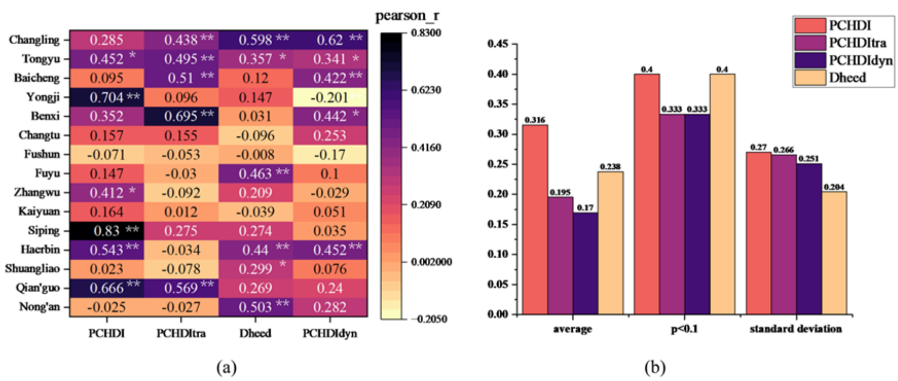


Fig 19 Comparison of PCHDI constructed in this study based on machine learning thresholds, PCHDItra based on traditional thresholds, PCHDIdyn based on dynamic thresholds, and the traditional compound index DHEED at the site level.(a) Pearson correlation coefficient heatmap between different indices and yield reduction rates at various sites.(b) Average Pearson correlation coefficients, proportions of significant sites, and standard deviations of different indices across sites.

Referring to the visualization results in Figures 13-16 of section 4.3.5, we validated the results using historical disaster statistical yearbooks. The validation results are very promising, with almost no misjudgment of severity or disaster occurrence areas. The occurrence times and regions of multiple historical compound heat-drought events coincide with the recorded events in the historical records (Table 5), further proving the superiority of the index constructed in this study.

Table 5 Comparison of Historical Classic Cases

Year	Event
1992	In June and July, compound heat-drought disasters led to a yield loss of 1.212 $\times 10^6$ hm^2 in Liaoning(Dictionary, 2005)； while parts of Hulunbeier and Xinanmeng in Inner Mongolia experienced drought (Dictionary, 2008).
2000	After summer began, Liaoning faced rare high temperatures and low rainfall, with

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

	agricultural losses reaching 2.7862×10^6 hm^2 (Dictionary, 2005), From June 25 to July 25, Chifeng City in Inner Mongolia had 320 consecutive days of drought, affecting over 86 kha of farmland (Dictionary, 2008).
2003	Rainfall from January to May was the third-lowest since 1961, with above-average temperatures leading to severe droughts in western Northeast China and eastern Inner Mongolia. Some areas experienced five years of consecutive drought since 1999, hindering crop planting(Administration, 2017).
2010	In late June, Heilongjiang's Greater Khingan Mountains and Mudanjiang experienced severe drought, while Liaoning had minimal drought. Jilin lost 349,000 hectares to compound heat-drought disasters, and 12 cities in Inner Mongolia experienced summer droughts(Administration, 2012).
2020	From June to July, the southwest of Northeast China received 30-80% less rainfall than usual. Liaoning had the lowest rainfall since 1981, and the drought intensity was the second-highest in the past decade. Summer droughts impacted spring maize in western and central-northern Liaoning, while Hulunbeier areas faced continuous drought (Administration, 2022).

5.2 Explanatory Advantage of the Process-based Cumulative Index for Yield Variability and Its Physiological Basis

Based on the theoretical framework of the Soil-Plant-Atmosphere Continuum (SPAC), after crops experience stress, even if environmental factors return to normal levels, the recovery of their physiological functions is not immediate. Studies have shown that stomatal conductance exhibits a lagged response to atmospheric changes (Bhatt et al., 2025), and the repair of xylem embolism takes

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

time (Galviz et al., 2025). More importantly, crop hormones and irreversible morphological changes triggered by stress act as 'memory factors,' leading to 'historical dependency' in their responses (Aswathi et al., 2025; Alongi et al., 2025). These mechanisms together result in a gradual recovery of stomatal conductance, photosynthetic rate, and root water uptake after stress, and this lag effect has a significant cumulative contribution to yield formation. Traditional analysis methods based on daily-scale intensity often overlook the temporal continuity of stress and its cumulative effects, making it difficult to fully explain the mechanisms behind final yield formation. Therefore, this study constructs the Process-based Cumulative Heat-Drought Index (PCHDI) based on the daily-scale compound index CDHI, aiming to capture the dynamic cumulative characteristics of compound stress during key growth stages of maize. The following explains why the cumulative process has a stronger explanatory power for yield from three perspectives:

First, the application of Run Theory effectively filters out noise interference and highlights the contribution of the effective stress processes. As an annual crop, maize yield formation depends on the accumulation of photosynthetic products during specific growth stages (e.g., tasseling to grain filling). If only daily-scale intensity analysis is used, short-term temperature fluctuations or slight declines in soil moisture (e.g., CDHI > 0 but lasting only 1-2 days) may be misjudged as stress events. However, these short disturbances are often insufficient to trigger irreversible physiological responses in crops. PCHDI addresses this by setting a duration threshold (Dur $\geqslant$ 3 days), allowing the daily CDHI to contribute to the cumulative process only when the continuous stress satisfies the effective process criteria of Run Theory. This design effectively filters out high-frequency noise and focuses the index on persistent stress processes that can truly impact yield formation.

Second, the exponential decay form of the index accurately captures the lagged recovery

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

characteristics of the crop system. After experiencing stress, even when environmental factors return to normal levels, stomatal conductance, photosynthetic rate, and root water uptake still need time to gradually recover. This lag effect has an undeniable cumulative contribution to yield formation. PCHDI introduces a memory coefficient α (set at 0.95) to model the exponential decay of historical stresses through a recursive form: during the effective occurrence period, the current stress contributes the most, while historical stress decays exponentially. When stress alleviates or the process no longer meets the continuity constraint, PCHDI stops accumulating new values and only decays exponentially by α, simulating the gradual recovery process of the crop system. This mechanism allows PCHDI to reflect the complete dynamic trajectory of "stress occurrence – sustained impact – lagged recovery," rather than just recording the instantaneous state of stress.

Additionally, the alignment of memory length with the maize critical growth window enhances the physiological significance of the cumulative process. The effective memory length γ derived from the memory coefficient is approximately 20 days (with α =0.95), which is of the same order of magnitude as the key yield formation window (approximately 15-25 days) during tasseling, flowering, and early grain filling. This means that PCHDI can effectively accumulate the contribution of compound stresses during this critical period, while also considering the lag effect of historical stresses. For example, brief heat-drought stress during tasseling may still affect yield during early grain filling by influencing ear grain number or grain weight. The exponential decay structure of PCHDI captures this cross-stage accumulation process, while daily-scale CDHI or simple accumulation indices struggle to reflect this nonlinear time-dependent relationship.

In summary, PCHDI effectively characterizes the dynamic cumulative characteristics of compound heat-drought stress by filtering out effective stress processes using Run Theory,

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

simulating crop lagged recovery with exponential decay, and matching the memory length to critical growth stages. Compared to indices that only focus on instantaneous intensity or simple accumulation, PCHDI more closely aligns with the real laws of stress signal transmission and crop response in the SPAC system, thus providing a stronger explanatory power for yield variation. This finding also suggests that when assessing the impacts of climate change on agriculture, the cumulative effect of stress processes should be emphasized rather than isolated events, in order to more accurately reveal the nonlinear relationship between extreme climate and crop yield.

5.3 Limitations and Uncertainty

This study finds that there is a clear nonlinear dependence between the thresholds of compound heat-drought events and key climate factors (as shown in Figure 6). Although the overall explanatory power of the model is only $R^2 \approx 0.2$, it is noteworthy that whether at the pixel level or the site level, comparisons with indices based on empirical thresholds or existing datasets show that the thresholds derived in this study better explain yield reduction, which sufficiently demonstrates the significance of these thresholds..

So, why does the clear nonlinear relationship have only $R^2 \approx 0.2$ explanatory power? Liu et al. (2026), when analyzing the relationship between climate and crops in their study, faced the same contradiction and provided a profound explanation: In highly heterogeneous drought and semi-arid ecosystems, even with real and clear bioclimate coupling processes, the quantitative explanatory power(R^2) of simple models relying on a few core climate variables is inherently limited due to spatial heterogeneity, unobserved covariates (such as soil, topography, and disturbance history), and the superposition of multiple process interactions. In other words, a low R^2 quantifies the gap between the 'core processes captured' and the 'unobserved system complexities,' rather than negating

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

the existence of the core processes themselves (Liu et al., 2026b).

The nonlinear dependence trend we identified represents the dominant mechanism in the formation of compound event thresholds. However, the $R^2 \approx 0.2$ suggests that beyond this dominant mechanism, there are many local modulating factors (such as prior soil moisture, vegetation status, and atmospheric circulation anomalies) that have not yet been incorporated into the model. This is a typical characteristic of complex system modeling—the core laws are clear, but quantitative prediction is limited by the completeness of the information.

In future research, more climate factors (such as soil temperature, atmospheric humidity, sunshine duration, radiation intensity, and wind speed) should be considered, along with the integration of higher-resolution ground observation data, to improve the spatial resolution and accuracy of compound heat-drought event identification. Additionally, a more comprehensive deepening of pre-impact factors could be incorporated, and finally, better phenological monitoring technologies could be combined to optimize threshold determination and extend its applicability to other crops and regions, thereby improving the practicality and generality of compound event recognition.

6 Conclusion

This study focuses on maize in Northeast China and constructs a three-dimensional process-based cumulative heat-drought index (PCHDI) based on the SPAC theory. It systematically analyzes the spatiotemporal evolution characteristics of compound heat-drought events from 1990 to 2023 and their yield effects. The main conclusions are as follows:

（1）A composite stress evaluation framework that integrating intensity, synergy, and cumulative process was constructed. VCI, STI, and SSMI were mapped to a three-dimensional state

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

space, with vector magnitude representing stress intensity and the angle between the vector and (1,1,1) direction representing stress synergy. Response thresholds for compound heat-drought events were optimized using XGBoost-SHAP.

（2）This study elucidates the spatial clustering characteristics and dynamic migration patterns of compound events. The spatial distribution of events exhibits a "western-central concentration, peripheral diffusion" pattern: high-value areas of maximum impact area, cumulative intensity, duration, and volume are concentrated in the eastern agricultural-pastoral transition zone of Inner Mongolia and the western Songnen Plain, forming a large-scale, high-intensity, and long-duration core high-frequency area. In contrast, the northeastern Sanjiang Plain and the southern edges are dominated by small-area, short-duration events. From 2012 to 2023, the density of high-intensity events in the core region significantly increased, further confirming the intensification of compound disaster risks over the past decade. The event centroid follows a "southwest-northeast-southwest" oscillation pattern: migrating southwest from 1990 to 2000, northeast from 2001 to 2011, and southwest again after 2012. After 2006, the migration step length fluctuated significantly, with multiple instances reaching high values between 700 and 900 km from 2015 to 2023, indicating a significant increase in spatial pattern instability and migration amplitude since the 21st century.

（3）The study also clarifies the growth stage differences and mutation transmission patterns of compound stress. The trend intensity and significance of compound heat-drought stress exhibit a "stronger in later stages, weaker in earlier stages" pattern: the positive slope regions for flowering and ear stages almost cover the entire region, with Z-statistics generally exceeding 1.96 ($p<0.05$), showing a significant intensification of stress. In contrast, only localized weak trends are present during the emergence and seedling stages, with no significant changes over large areas. The

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

significant upward regions gradually expanded from scattered distributions in the early stages to full coverage in the later stages, reflecting the evolution of stress from "expanding range and increasing intensity". The M-K mutation detection reveals that the mutation time of intensified stress follows a "later growth stages → earlier growth stages" transmission pattern: the earliest mutation occurred during the flowering stage (2008), followed by the ear stage (2010), and the seedling stage (2015), with no significant mutation observed during the emergence stage. After the 21st century, mutation events concentrated, with the peak annual pixel mutation rate occurring in the flowering stage (close to 35%), indicating that the later growth stages are the most sensitive to stress changes and are the critical windows for the impact of compound disasters on maize yield.

（4）The superiority of PCHDI in yield correlation and historical event inversion was validated. The pixel-level correlation and stability of PCHDI with maize yield reduction were generally superior to traditional indices and indices based on empirical thresholds. At the site level, PCHDI showed a 62% higher average yield correlation than the PCHDItra based on traditional thresholds and approximately 33% higher correlation than the existing Dheed dataset. The spatiotemporal inversion of historical disaster cases highly matches the recorded events, proving that this index can effectively capture the actual impact of compound stress on crop yield.

In conclusion, the three-dimensional process-based cumulative framework constructed in this study provides a new method for the quantitative identification of compound extreme events. It can more accurately and effectively identify compound heat-drought events, offering greater support for global climate change-driven food security risk assessments.

CRediT authorship contribution statement

Yashu Gao: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Cha Ersi:** Conceptualization, Software, Validation, Writing – review & editing, **Dan Chen:** Conceptualization, Supervision. **Sicheng Wei:** Supervision. **Ying Guo:** Data Management, Formal Analysis, Supervision. **Ziyuan Zhou:** Conceptualization, supervision. **Zhijun Tong:** Conceptualization, data curation. **Xingpeng Liu:** Conceptualization. **Jiquan Zhang :** Supervision, Project management, Funding acquisition. **Chunli Zhao :** Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

This study was supported by the National K&D Program of China (2023YFD2301701), Jilin Province's Special Project (Topic) for Focused Efforts and Breakthroughs (2025JL0013GX).

Science and Technology Development Plan Project of Jilin Province, China （20260203067SF）

References

Administration, C. M.: Yearbook of Meteorological Disasters in China (2011), 1st Edition, China Meteorological Press, Beijing2012.

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

Administration, C. M.: Yearbook of Meteorological Disasters in China (2004), 3rd Edition, China Meteorological Press, Beijing2017.

Administration, C. M.: Yearbook of Meteorological Disasters in China (2021), 1st Edition, China Meteorological Press, Beijing2022.

Alongi, F., Petek-Petrik, A., Mukarram, M., Torun, H., Schuldt, B., and Petrík, P.: Somatic drought stress memory affects leaf morpho-physiological traits of plants via epigenetic mechanisms and phytohormonal signalling, Plant Gene, 42, https://doi.org/10.1016/j.plgene.2025.100509, 2025.

Andreadis, K. M. C., E. A.; Wood, A. W.; Hamlet, A. F.; Lettenmaier, D. P.: Twentieth-Century Drought in the Conterminous United States, Journal of Hydrometeorology, 6, 985–1000, https://doi.org/10.1175/JHM450.1, 2005.

Aswathi, K. P. R., Ul‐Allah, S., Puthur, J. T., Siddique, K. H. M., Frei, M., and Farooq, M.: The Plant Mind: Unraveling Abiotic Stress Priming, Memory, and Adaptation, Physiol. Plant., 177, https://doi.org/10.1111/ppl.70372, 2025.

Bhatt, T., Rathore, N., Semerád, J., Cajthaml, T., Thakur, D., and Münzbergová, Z.: Transgenerational Drought and Methyl Jasmonate Memory Interactively Shape Metabolome and Physiology in Clonal Grass, Physiol. Plant., 178, https://doi.org/10.1111/ppl.70720, 2025.

Birant, D. and Kut, A.: ST-DBSCAN: An algorithm for clustering spatial–temporal data, Data & Knowledge Engineering, 60, 208-221, https://doi.org/10.1016/j.datak.2006.01.013, 2007.

Blum, A. R., S.; Kanemasu, E. T.; Paulsen, G. M.: The physiology of heterosis in sorghum with respect to environmental stress, Annals of Botany, 65, 149–158, https://doi.org/10.1093/oxfordjournals.aob.a087919, 1990.

Bogenreuther, J., Bogner, C., Siebert, S., and Koellner, T.: Heat, drought, and compound events: Thresholds and impacts on crop yield variability, AGRICULTURAL AND FOREST METEOROLOGY, 375, https://doi.org/10.1016/j.agrformet.2025.110836, 2025.

Cammalleri, C. and Toreti, A.: A Generalized Density-Based Algorithm for the Spatiotemporal Tracking of Drought Events, Journal of Hydrometeorology, 24, 537-548, https://doi.org/10.1175/jhm-d-22-0115.1, 2023.

Ceglar, A., Toreti, A., Lecerf, R., Van der Velde, M., and Dentener, F.: Impact of meteorological drivers on regional inter-annual crop yield variability in France, Agricultural and Forest Meteorology, 216, 58-67, https://10.1016/j.agrformet.2015.10.004, 2016.

Chen, T. G., Carlos; Sander, Jörg: XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, https://doi.org/10.1145/2939672.2939785, 2016.

Dictionary, E. B. o. C. M. D.: China Meteorological Disaster Dictionary: Liaoning Volume, 1st Edition, China Meteorological Press, Beijing2005.

Dictionary, E. B. o. C. M. D.: China Meteorological Disaster Dictionary: Inner Mongolia Volume, 1st Edition, China Meteorological Press, Beijing2008.

Dinpashoh, Y., Mirabbasi, R., Jhajharia, D., Abianeh, H. Z., and Mostafaeipour, A.: Effect of Short-Term and Long-Term Persistence on Identification of Temporal Trends, Journal of Hydrologic Engineering, 19, 617-625, https://doi.org/10.1061/(asce)he.1943-5584.0000819, 2014.

Disasa, K. N., Yan, H., Bao, R., Zhang, J., Zhang, C., Wang, B., and Wang, G.: Modelling the impacts of future climate variable changes on rainfed maize production, evapotranspiration, and adaptation measures in Northeast China, Journal of Hydrology, 662, https://doi.org/10.1016/j.jhydrol.2025.134051, 2025.

Fang, B. and Lu, M.: Heatwave and Blocking in the Northeastern Asia: Occurrence, Variability, and Association, Journal of Geophysical Research: Atmospheres, 125, https://doi.org/10.1029/2019jd031627, 2020.

Galviz, Y. C., Bortolin, G. S., and Souza, G. M.: Trade-off between water deficit memory and growth under limiting conditions in rice, Plant Physiol. Biochem., 226, https://doi.org/10.1016/j.plaphy.2025.110048, 2025.

Guo, Y., Zhang, J. Q., Li, K. W., Aru, H., Feng, Z., Liu, X. P., and Tong, Z. J.: Quantifying hazard of drought and heat compound extreme events during maize (Zea mays L.) growing season using Magnitude Index and Copula, Weather and Climate Extremes, 40, 12, https://doi.org/10.1016/j.wace.2023.100566, 2023.

Guo, Y. L., Si, R. R., and Shang, D. Y.: Research on the Compound Drought and High-Temperature Events During the Flowering Period of Summer Maize in the Huang-Huai-Hai Region of China, Advances in Meteorology, 2025, 13, https://doi.org/10.1155/adme/1581745, 2025.

Hansen, G. J., Emily; Yan, Jason: Analysis of Climate Change and Water Access on the Production of Leading Agricultural Commodities in California, USA, 1–32, 2020.

Hao, Z. C., Hao, F. H., Xia, Y. L., Feng, S. F., Sun, C., Zhang, X., Fu, Y. S., Hao, Y., Zhang, Y., and Meng, Y.: Compound droughts and hot extremes: Characteristics, drivers, changes, and impacts, EARTH-SCIENCE REVIEWS, 235, https://doi.org/10.1016/j.earscirev.2022.104241, 2022.

Herrera - Estrada, J. E., Satoh, Y., and Sheffield, J.: Spatiotemporal dynamics of global drought, Geophysical Research Letters, 44, 2254-2263, https://doi.org/10.1002/2016gl071768, 2017.

Hosseinzadehtalaei, P., Termonia, P., and Tabari, H.: Projected changes in compound hot-dry events depend on the dry indicator considered, Communications Earth & Environment, 5, https://doi.org/10.1038/s43247-024-01352-4, 2024.

Hou, P., Liu, Y., Liu, W., Yang, H., Xie, R., Wang, K., Ming, B., Liu, G., Xue, J., Wang, Y., Zhao, R., Zhang, W., Wang, Y., Bian, S., Ren, H., Zhao, X., Liu, P., Chang, J., Zhang, G., Liu, J., Yuan, L., Zhao, H., Shi, L., Zhang, L., Yu, L., Gao, J., Yu, X., Wang, Z., Shen, L., Ji, P., Yang, S., Zhang, Z., Xue, J., Ma, X., Wang, X., Lu, T., Dong, B., Li, G., Ma, B., Li, J., Deng, X., Liu, Y., Yang, Q., Jia, C., Chen, X., Fu, H., and Li, S.: Quantifying maize grain yield losses caused by climate change based on extensive field data across China, Resources, Conservation and Recycling, 174, https://doi.org/10.1016/j.resconrec.2021.105811, 2021.

Hu, Y., Wang, W., Wang, P., Teuling, A. J., and Zhu, Y.: Spatial-temporal variations and drivers of the compound dry-hot event in China, Atmospheric Research, 299, https://doi.org/10.1016/j.atmosres.2023.107160, 2024.

Huang, Y. and Liu, Z.: Improving Northeast China's soybean and maize planting structure through subsidy optimization considering climate change and comparative economic benefit, Land Use Policy, 146, https://doi.org/10.1016/j.landusepol.2024.107319, 2024.

IPCC: Summary for Policymakers, Cambridge, UK; New York, NY, USA9781009157896, 3-32, https://doi.org/10.1017/9781009157896.001, 2023.

Jiang, R., Liang, J., Zhao, Y., Wang, H., Xie, J., Lu, X., and Li, F.: Assessment of vegetation growth and drought conditions using satellite-based vegetation health indices in Jing-Jin-Ji region of China, Scientific Reports, 11, https://doi.org/10.1038/s41598-021-93328-z, 2021.

Jiang, W., Yuan, L., Wang, W., Cao, R., Zhang, Y., and Shen, W.: Spatio-temporal analysis of vegetation variation in the Yellow River Basin, Ecological Indicators, 51, 117-126, https://doi.org/10.1016/j.ecolind.2014.07.031, 2015.

Kang, Y., Guo, E., Wang, Y., Bao, Y., Zhao, S., and A, R.: Spatiotemporal Variation in Compound Dry and Hot Events and Its Effects on NDVI in Inner Mongolia, China, Remote Sensing, 14, https://doi.org/10.3390/rs14163977, 2022.

Lesk, C., Rowhani, P., and Ramankutty, N.: Influence of extreme weather disasters on global crop production, NATURE, 529, 84-+, https://doi.org/10.1038/nature16467, 2016.

Li, B. L., Zhong; Huang, Feng; Yang, Xiaoguang; Liu, Zhijuan; Wan, Wei; Wang, Jingkuan; Xu, Yingde; Li, Zizhong; Ren, Tusheng: Ensuring National Food Security by Strengthening High-productivity Black Soil Granary in Northeast China, Bulletin of Chinese Academy of Sciences, 36, 1184–1193, https://doi.org/10.16418/j.issn.1000-3045.20210706003, 2021.

Li, H., Cao, Y., Xiao, J., Yuan, Z., Hao, Z., Bai, X., Wu, Y., and Liu, Y.: A daily gap-free normalized difference vegetation index dataset from 1981 to 2023 in China, Scientific Data, 11, https://doi.org/10.1038/s41597-024-03364-3, 2024.

Liu, J., Li, M., Pham, Q. B., Liu, T., Ren, Y., Shalamzari, M. J., and He, P.: From historical to future scenarios: A deep learning and dynamical detection framework for compound drought and heatwave events, Expert Systems with Applications, 307, https://doi.org/10.1016/j.eswa.2026.131105, 2026a.

Liu, T., Zhu, X., Sun, H., and Tang, M.: The spatiotemporal characteristics of extreme drought events in China from 1961 to 2022 via a copula function, Weather and Climate Extremes, 50, https://doi.org/10.1016/j.wace.2025.100815, 2025.

Liu, Z., Feng, Y., Cao, X., Peñuelas, J., Descals, A., Yang, D., Cheung, K. K., Liu, S., Lin, Z., Liu, X., Wu, X., Ma, H., Li, W., Xia, H., Jin, S., Liang, S., Liu, L., Liu, L., Su, Y., Chen, J., and Wu, J.: Integrating very-high-resolution imagery, Sentinel-2 time-series data, and machine learning to map shrub fractional abundance across arid and semi-arid ecosystems in China, Remote Sensing of Environment, 335, https://doi.org/10.1016/j.rse.2026.115300, 2026b.

Liu, Z. Z., Wen; Yuan, Y.: 3D DBSCAN detection and parameter sensitivity of the 2022 Yangtze river summertime heatwave and drought, Atmospheric and Oceanic Science Letters, 16, 100324, https://doi.org/10.1016/j.aosl.2022.100324, 2023.

Lu, X., Wei, S., Guo, Y., Ersi, C., Chen, D., Zhou, Z., Li, Y., Dong, Y., Tong, Z., Liu, X., Zhang, J., Zhao, C., and Zhao, Y.: Dynamic hazard assessment of compound drought and heat events in maize based on a soil-temperature-integrated index in Songliao Plain, European Journal of Agronomy, 175, https://doi.org/10.1016/j.eja.2026.127984, 2026.

Lundberg, S. M. and Lee, S. I.: A Unified Approach to Interpreting Model Predictions, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09, WOS:000452649404081, https://doi.org/10.48550/arXiv.1705.07874, 2017.

McKee, T. B. D., Nolan J.; Kleist, John: The relationship of drought frequency and duration to time scales, Eighth Conference on Applied Climatology, Anaheim, California, 17–22 January 19931993.

Ni, Y., Qiu, B., Miao, X., Li, L., Chen, J., Tian, X., Zhao, S., and Guo, W.: Shift of soil moisture-temperature coupling exacerbated 2022 compound hot-dry event in eastern China, Environmental Research Letters, 19, https://doi.org/10.1088/1748-9326/ad178c, 2024.

Olufayo, A., Baldy, C., Ruelle, P., and Konate, J.: DIURNAL COURSE OF CANOPY TEMPERATURE AND LEAF WATER POTENTIAL OF SORGHUM (SORGHUM-BICOLOR L MOENCH) UNDER A MEDITERRANEAN CLIMATE, Agricultural and Forest Meteorology, 64, 223-236, https://doi.org/10.1016/0168-1923(93)90030-l, 1993.

Pius, M. N., Yelwa, S. A., and Adewuyi, T. O.: From Deficit to Dysfunction: Rethinking Drought through a Multi-Index Functional Framework for Ecologically Vulnerable Systems, International Journal

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

of Research in Environmental Science, 11, 1-15, https://doi.org/10.20431/2454-9444.1103001, 2025.

Quiring, S. M. and Ganesh, S.: Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas, Agricultural and Forest Meteorology, 150, 330-339, https://doi.org/10.1016/j.agrformet.2009.11.015, 2010.

Sánchez-Benítez, A., Barriopedro, D., and García-Herrera, R.: Tracking Iberian heatwaves from a new perspective, Weather and Climate Extremes, 28, https://doi.org/10.1016/j.wace.2019.100238, 2020.

Sander, J. E., Martin; Kriegel, Hans-Peter; Xu, Xiaowei: Density-based clustering in spatial databases the algorithm GDBSCAN and its applications, Data Mining and Knowledge Discovery, 2, 169–194, https://doi.org/10.1023/A:1009745219129, 1998.

Serrano-Notivoli, R. B., Santiago; Saz, Miguel Ángel; de Luis, Martín: Recent trends reveal decreasing intensity of daily precipitation in Spain, International Journal of Climatology, 38, 4211–4224, https://doi.org/10.1002/joc.5562, 2018.

Shan, B. D. B., Bernard; Verhoest, Niko: Compound drought and heatwave identification: daily-scale independent extreme events based on 120-year observations, EGU General Assembly 2022, Vienna, Austria & Online, 2022, EGU22-3877, https://doi.org/10.5194/egusphere-egu22-3877,

Sheffield, J., Andreadis, K. M., Wood, E. F., and Lettenmaier, D. P.: Global and Continental Drought in the Second Half of the Twentieth Century: Severity–Area–Duration Analysis and Temporal Variability of Large-Scale Events, Journal of Climate, 22, 1962-1981, https://doi.org/10.1175/2008jcli2722.1, 2009.

Sun, L. J., Lin; Wei, Zhonghua; Li, Junfeng: Demand Forecasting of Taxi Travel Based on GPS Data, Journal of Transport Information and Safety, 39, 128-136, https://doi.org/10.3963/j.issn.1674-4861.2021.05.016, 2021.

Wang, L. J., Liao, S. H., Huang, S. B., Ming, B., Meng, Q. F., and Wang, P.: Increasing concurrent drought and heat during the summer maize season in Huang-Huai-Hai Plain, China, International Journal of Climatology, 38, 3177-3190, https://doi.org/10.1002/joc.5492, 2018.

Wang, R., Zhang, J., Wang, C., and Guo, E.: Characteristic Analysis of Droughts and Waterlogging Events for Maize Based on a New Comprehensive Index through Coupling of Multisource Data in Midwestern Jilin Province, China, Remote Sensing, 12, https://doi.org/10.3390/rs12010060, 2019.

Wang, R., Zhang, X. X., Cong, L. P., Wang, Y. L., and Bai, X. T.: Risk Assessment of Compound Dry-Hot Events for Maize in Liaoning Province, Atmosphere, 15, 19, https://doi.org/10.3390/atmos15070834, 2024.

Weynants, M., Ji, C., Linscheid, N., Weber, U., Mahecha, M. D., and Gans, F.: Dheed: an ERA5 based global database of compound dry and hot extreme events from 1950 to 2023, Earth System Science Data, 17, 6621-6645, https://doi.org/10.5194/essd-17-6621-2025, 2025.

Wu, Z. H., Norden E.: Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method, Advances in Adaptive Data Analysis, 1, 1–41, https://doi.org/10.1142/S1793536909000047, 2005.

Xu, L., Abbaszadeh, P., Moradkhani, H., Chen, N., and Zhang, X.: Continental drought monitoring using satellite soil moisture, data assimilation and an integrated drought index, Remote Sensing of Environment, 250, https://doi.org/10.1016/j.rse.2020.112028, 2020.

Xu, P., Wang, D., Wang, Y., Singh, V. P., Qiu, J., Wu, J., Zhang, A., and Ju, X.: +Dynamic identification and risk analysis of compound dry-hot events considering nonstationarity, Journal of Hydrology, 616, https://doi.org/10.1016/j.jhydrol.2022.128852, 2023.

Yin, H., Bai, F. Q., Wu, H. M., Yan, M., and Zhou, S.: Assessment of Vegetation Vulnerability in the Haihe River Basin Under Compound Heat and Drought Stress, Sustainability, 16, 19,

https://doi.org/10.5194/egusphere-2026-3379
Preprint. Discussion started: 29 July 2026

https://doi.org/10.3390/su162310489, 2024.

Yu, W. Y., Ji, R. P., Wu, J. W., Feng, R., Mi, N., and Chen, N. N.: Combined Effects of Heat and Drought Stress on the Growth Process and Yield of Maize (Zea mays L.) in Liaoning Province, China, Atmosphere, 14, 37, https://doi.org/10.3390/atmos14091397, 2023.

Yu, Y., Xu, X., Yan, Q., Lu, J., and Yu, Y.: Characteristics of Regional Hourly Extreme Precipitation With Different Durations Over the Northeast Plain, China During Summer, Earth and Space Science, 12, https://doi.org/10.1029/2024ea003973, 2025.

Zhang, Q., Yu, X., Qiu, R., Liu, Z., and Yang, Z.: Evolution, severity, and spatial extent of compound drought and heat events in north China based on copula model, Agricultural Water Management, 273, https://doi.org/10.1016/j.agwat.2022.107918, 2022.

Zhang, W., Luo, M., Gao, S., Chen, W., Hari, V., and Khouakhi, A.: Compound Hydrometeorological Extremes: Drivers, Mechanisms and Methods, Frontiers in Earth Science, 9, https://doi.org/10.3389/feart.2021.673495, 2021.

Zhang, X., Chen, N., Li, J., Chen, Z., and Niyogi, D.: Multi-sensor integrated framework and index for agricultural drought monitoring, Remote Sensing of Environment, 188, 141-163, https://doi.org/10.1016/j.rse.2016.10.045, 2017.

Zhou, H., Zhou, W., Liu, Y., Yuan, Y., Huang, J., and Liu, Y.: Identifying spatial extent of meteorological droughts: An examination over a humid region, Journal of Hydrology, 591, https://doi.org/10.1016/j.jhydrol.2020.125505, 2020.

Zhou, Z. Y., Guo, Y., Chen, D., Li, K. W., Wang, R., Wei, X., Zhang, J. Q., Zhao, C. L., Tong, Z. J., and Liu, X. P.: Identification and Spatiotemporal Characteristic Analysis of Compound Weather and Climate Extremes for Maize in Different Climate Zones of the Songliao Plain, Int. J. Disaster Risk Sci., 15, 831-851, https://doi.org/10.1007/s13753-024-00585-3, 2024.

Zscheischler, J., Westra, S., van den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, 8, 469-477, https://doi.org/10.1038/s41558-018-0156-3, 2018.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., Maraun, D., Ramos, A. M., Ridder, N. N., Thiery, W., and Vignotto, E.: A typology of compound weather and climate events, Nature Reviews Earth & Environment, 1, 333-347, https://doi.org/10.1038/s43017-020-0060-z, 2020.

Zuo, Y., Hu, Z., Yuan, S., Zheng, J., Yin, X., and Li, B.: Identification of Convective and Stratiform Clouds Based on the Improved DBSCAN Clustering Algorithm, Advances in Atmospheric Sciences, 39, 2203-2212, https://doi.org/10.1007/s00376-021-1223-7, 2022.