A multivariate approach for statistical assessments of compound extremes

Zengchao Hao[1], Fanghua Hao[1], Vijay P. Singh[2], Youlong Xia[3], Chunxiang Shi[4], Xuan Zhang[1]

[1]College of Water Sciences, Beijing Normal University, Beijing, 100875, China

2 Department of Biological and Agricultural Engineering and Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843-2117, USA

3 I.M. System Group at Environmental Modeling Center, National Center for Environmental Prediction, College Park, Maryland, USA

4 National Meteorological Information Center, China Meteorological Administration, Beijing 100081, China

Corresponding Author: Zengchao Hao (Email address: haozc@bnu.edu.cn)

Submit to ***Journal of Hydrology***

Key points:

- Propose a statistical approach for modeling multiple extreme indices
- Estimate the joint severity of compound extreme
- Evaluate the impact of precipitation and temperature on agricultural drought

Abstract

Global warming may affect the regime of hydroclimatic systems and induce more frequent occurrences of extremes, such as drought, heat wave and flood. Apart from the assessment of each extreme, recent decades have witnessed a surge in the study of compound extreme, i.e., the concurrence of multiple extremes. To aid the understanding of compound extremes, a variety of studies has been conducted to assess the dependence among different variables or extremes. As such, it is important to model multiple contributing variables of compound extremes to characterize the associated risk taking into account the dependence. In this study, a multivariate approach based on the meta-Gaussian model is proposed for the statistical analysis of compound extremes in the trivariate case. The application of the proposed approach is illustrated with the compound drought and hot extreme in the U.S. based on monthly precipitation, soil moisture and temperature from the North American Land Data Assimilation System (NLDAS-2). The likelihood of the occurrence of compound drought and hot extreme is assessed based on the joint distribution, which is shown to be higher in regions with significant land atmosphere interactions. The impact of precipitation and temperature on the occurrence of agricultural drought is also assessed based on the conditional distribution. Overall, results show that the proposed method provides a useful tool for statistical assessments of the compound extreme through constructing the joint and conditional distribution.

Keywords: extreme; compound extreme; Meta-Gaussian model; joint distribution

1 Introduction

Extremes, such as drought, heat wave, fluvial (pluvial or coastal) flood, may exert large impacts to the agriculture, energy, and ecosystems. Recent decades have witnessed a large number of occurrence of extremes, such as the 2011 Texas drought in the U.S. or the 2003 Europe heat wave (Coumou and Rahmstorf, 2012). Under the global warming, it is expected that more extremes may be induced imposing great threats to the human society and ecosystems (IPCC, 2012; Rummukainen, 2012). Previous studies have shown the increase of various extremes at different temporal and spatial scales around the globe (Dai, 2011; Heim, 2015; Alexander, 2016).

It has been well recognized that the impact of extremes may be related to multiple variables or processes. Recently, the compound extreme (i.e., the occurrence of concurrent or consecutive events leading to extreme impacts) has attracted much attention due to their even larger impacts on different sectors than that from individual extreme (Seneviratne et al., 2012; Leonard et al., 2014). These examples include a wide ranges of occurrences of multiple extremes at different regions and seasons, such as the drought and hot extreme, storm surge and high rainfall (Kew et al., 2013; van den Hurk et al., 2015; Wahl et al., 2015). The combined drought and hot extreme, which may lead to larger impacts to agriculture and ecosystems than that from either in isolation, has been among the most commonly studied compound extremes (Hao et al., 2013; Mazdiyasni and AghaKouchak, 2015; Cheng et al., 2016; Sharma and Mujumdar, 2017; Zscheischler and Seneviratne, 2017). For example, based on observations from Climatic Research Unit (CRU) (Harris et al., 2013) and the University of Delaware (Willmott and Matsuura, 2001), Hao et al. (2013) showed that the occurrence of warm/dry extremes has increased for the period 1978–2004 relative to 1951–1977 across the globe, including central Africa, eastern Australia, and parts of Russia. Zscheischler and Seneviratne (2017) showed an increase of the occurrence rate of extremely hot and dry warm seasons in the 21st century including the northern extra-tropics, Amazon region and Indonesia based on historical simulations and climate projections in the Coupled Model Intercomparison Project Phase 5 (CMIP5) models.

In defining the compound extreme, it has been highlighted that the dependence generally exists among different contributing variables or processes and is a key factor to characterize the compound extreme (Leonard et al., 2014; Martius et al., 2016; Zscheischler and Seneviratne, 2017). Thus it is important to take into account the dependence among different variables in assessing the likelihood or probability of compound extremes and the potential risk (Risk= probability of events or trends × consequences (Zscheischler et al., 2018)). The joint distribution has been applied for the analysis of multivariate or compound extremes, among which the copula model has been commonly used (Salvadori et al., 2007; Hawkes, 2008; Durante and Salvadori, 2010; Zscheischler and Seneviratne, 2017), mainly in the bivariate case. Due to the multiple components or processes in the occurrence of compound extremes, certain efforts have been devoted to the extreme analysis in higher dimensions (Genest et al., 2007; Zhang and Singh, 2007; Kao and Govindaraju, 2008; Song and Singh, 2010; Wong et al., 2010). However, the commonly used parametric copula family general falls short in modeling dependence for 3 or even higher dimensions (Aas et al., 2009) and thus other models flexible in modeling dependence are desired.

To facilitate the compound extreme modeling in high dimensions, we propose in this study the meta-Gaussian model for the statistical inference of compound extremes by transforming the individual extreme into the standardized extreme index (SEI). The model utilizes the intrinsic nature of SEI and provides an easy way to carry out the statistical modeling of dependence among different extreme indices. The application of the proposed model is illustrated for the compound drought and hot extreme in the U.S. based on monthly precipitation, soil moisture and temperature from the North American Land Data Assimilation System (NLDAS-2)(Xia et al., 2012).

2 Method

The compound extreme of particular interest in this study is the combined condition of low precipitation/soil moisture and high temperature (or the compound drought and hot extreme). This type of compound extreme is closely related to the "precipitation deficit flash droughts" that is also characterized by both high temperature anomalies and soil moisture deficits (Mo and Lettenmaier, 2016). The soil moisture deficit is closely related to the low precipitation and high temperature, which may deplete the soil moisture, leading to the occurrence of agricultural drought. This phenomenon of compound extreme can be assessed based on the conditional property of soil moisture given precipitation and temperature. Apparently, a trivariate distribution is needed to model the three variables to characterize the associated compound extreme. In the following, we introduce the meta-Gaussian model and its application for statistical assessments of compound extremes.

2.1 Meta-Gaussian model

Multivariate distributions are commonly employed to model the dependence between hydrologic variables or properties. The meta-Gaussian model has been proposed to meet the need of representing a full range of association and allowing for flexible forms of marginal distributions for the modeling of bivariate variables in hydrology (Kelly and Krzysztofowicz, 1997). In the context of characterizing compound extremes, the joint modeling of multiple contributing variables in high dimensions is of primary interest. Consider a random vector (X_1, X_2) of two continuous random variables with marginal distribution function $F_1(X_1)$ and $F_2(X_2)$, respectively.

Let N denote the standard normal distribution in the univariate case and N^{-1} its inverse. By applying the normal quantile transformation (NQT)(Herr and Krzysztofowicz, 2005), two standard normal variates can be defined as $Z_1=N^{-1}(F_1(X_1))$ and $Z_2= N^{-1}(F_2(X_2))$. The meta-Gaussian model can be used to construct the bivariate distribution of (X_1, X_2), which can be expressed as (Kelly and Krzysztofowicz, 1997):

$$P(X_1 \le x, X_2 \le x_2) = P(Z_1 \le z_1, Z_2 \le z_2)$$

$$= \int_{-\infty}^{z_2} \int_{-\infty}^{z_1} \frac{1}{(2\pi)\sqrt{(1-\rho^2)}} \exp\left\{ -\frac{t^2 - 2\rho st + s^2}{2(1-\rho^2)} \right\} ds dt \qquad \textbf{(1)}$$

where ρ is the Pearson's correlation coefficient between Z_1 and Z_2; s and t are the integral variables.

The basic idea of the meta-Gaussian model is to transform the variables under investigation into the normal variate based on the NQT and then the multivariate normal distribution can be used to model joint variations of multivariate variables (e.g., Z_1 and Z_2) (Kelly and Krzysztofowicz, 1997; Montanari and Brath, 2004; Wilks, 2011). A suite of standardized extreme indices (SEI), which is advantageous in the consistency and comparability of extremes, has been developed based on the NQT, including the Standardized Precipitation Index (SPI) (McKee et al., 1993), Standardized Soil moisture Index (SSI)(Hao and AghaKouchak, 2013) and Standardized Temperature Index (STI)(Zscheischler et al., 2014). Based on the SEI, the joint and conditional analysis of the extreme indices can be achieved based on meta-Gaussian model. The meta-Gaussian model is closely related to (or sometimes viewed as) the Gaussian copula (Renard and Lang, 2007; Vogl et al., 2012; Ben Alaya et al., 2014; Serinaldi, 2016; Rueda et al., 2016) and has been used for the statistical modeling of hydroclimatic variables (Montanari and Brath, 2004;

Herr and Krzysztofowicz, 2005; Wu et al., 2011). One of our focuses in this study is to derive the explicit form of the conditional distribution and thus we introduce the multivariate modeling of compound extremes based on the framework of the meta-Gaussian model.

2.2 Conditional distribution

Based on the property of the multivariate normal distribution, the conditional distribution of Y (a variable or vector) conditioned on X (a variable or vector) is also normally distributed (Kelly and Krzysztofowicz, 1997; Wilks, 2011). An interesting property from the proposed meta-Gaussian model is that the conditional distribution of a SEI conditioned on other SEIs is normally distributed. We use three variables (Y, X_1, X_2), which represent SSI, SPI and STI, respectively, to illustrate the conditional distribution of Y with respect to $\boldsymbol{X}$=[X_1,X_2]. Specifically, based on the meta-Gaussian model to construct the joint distribution of (Y, X_1, X_2), the explicit form of the conditional distribution of Y conditioned on $\boldsymbol{X}$ can be expressed as (Wilks, 2011):

$$Y \mid X \sim N\left(\mu_{Y|X}, \sum_{Y|X}\right) \quad \textbf{(2)}$$

where $\mu_{Y|X}$ is the conditional mean and $\boldsymbol{\Sigma}_{Y|X}$ is the conditional covariance matrix. These two parameters of the conditional distribution can be expressed as (Wilks, 2011; Hao et al., 2016):

$$\mu_{Y|X} = \mu_y + \sum_{yx} \sum_{xx}^{-1} \left(x - \mu_x\right) \quad \textbf{(3)}$$

$$\boldsymbol{\Sigma}_{Y|X} = \sum_{yy} - \sum_{yx} \sum_{xx}^{-1} \sum_{xy} \quad \textbf{(4)}$$

where $\boldsymbol{\mu}_X$ is the means of the vector $\boldsymbol{X}$ and μ_Y is the mean of the variable Y; $\boldsymbol{\Sigma}_{xx}$, $\boldsymbol{\Sigma}_{xy}$, $\boldsymbol{\Sigma}_{yx}$, and $\boldsymbol{\Sigma}_{yy}$ are covariance matric of the vector $\boldsymbol{X}$ and variable Y. A detailed expression of this equation for 3 variables can be found in Hao et al. (2016).

3 Data and Results

3.1 Data

In this study, the monthly precipitation, 2-m air temperature (meteorological forcing data of NLDAS-2) and root zone soil moisture with a spatial resolution of 0.125 degree for the period from 1979-2014 were obtained from the NLDAS-2 project (Xia et al., 2012). Due to the lack of the large scale and long-term observations of soil moisture for drought prediction(Ford et al., 2015), the simulated soil moisture from land surface model is an useful alternative (Sheffield et al., 2004). In this study, the simulated soil moisture data were obtained from the Noah model, which is used as the land component for several operational model systems of the National Centers for Environmental Prediction (NCEP) (Xia et al., 2014; Sun et al., 2018).

Since the compound drought and hot extreme is of primary interest, we perform the statistical analysis of the compound extreme in the summer season. The meteorological drought is generally characterized by SPI of multiple time scales based on the monthly precipitation. The soil moisture condition responds to the short scale precipitation anomalies while streamflow, groundwater, or reservoir storage responds to the long-term precipitation anomalies (Svoboda et al., 2012; Thilakarathne and Sridhar, 2017). In this study, we use the 3-month SPI as the drought indicator, which has been commonly employed for the assessment of drought (or wetness) and its interaction with hot extreme (Mueller and Seneviratne, 2012; Zscheischler et al., 2014). The 3-month SPI is also effective in capturing available moisture conditions in primary agricultural regions (e.g., a 3-month SPI at the end of August in the U.S. would capture the trend of precipitation during the reproductive and early grain-filling stage of certain crops)(Svoboda et al., 2012). The root zone soil moisture is a governing factor for vegetative growth and can be employed as a direct indicator of agricultural drought (Keyantash and Dracup, 2002; Sheffield et al., 2004). The 1-month soil moisture (e.g., percentile) is commonly used for the monitoring of the agricultural drought (Sheffield et al., 2004; Xia et al., 2014) and is also employed in this study based on the SSI. Following the previous study (Zscheischler et al., 2014), we characterize the hot extreme based on the STI of 1-month time scale, which can be used to capture the instant response of the terrestrial flux related to the plant.

For deriving the SPI (or other standardized extreme indices), a variety of parametric distributions has been proposed and validated (Stagge et al., 2015; Vicente-Serrano and Beguería, 2016). An alternative way to estimate the marginal distribution is the empirical method, which does not rely on a specific form of parametric distributions. In this study, the empirical Gringorten distribution (Gringorten, 1963) is employed to estimate the marginal distribution to derive these standardized extreme indices including SPI, SSI and STI. These indices for August were used for the subsequent analysis to illustrate the application of the proposed model in compound extreme analysis.

3.2 Dependence pattern

The correlation among the three indices during August is shown in Figure 1(a-c) and that significant at a 5% significance level is shown in Figure 1(d-f). It can be seen that there is significant correlation among the SPI, SSI and STI in large regions in the U.S.. The correlation between SPI and SSI during August is positive in all regions. This is intuitive in that the agricultural drought (or soil moisture) is generally dependent on the precipitation (or meteorological drought). The correlation between the SPI and STI is mostly negative during August, especially in the southern and southeastern U.S.. The negative dependence between SPI and STI is mainly due to the land surface interaction (i.e., the soil moisture deficit may induce the decreased evaporative cooling and increased sensible heat flux, leading to the high temperature) (Seneviratne et al., 2010; Hirschi et al., 2011; Mueller and Seneviratne, 2012; Berg et al., 2014; Whan et al., 2015). For example, the negative correlation between the SPI and STI is even lower in the High Plains with significant land atmosphere interactions. The dependence between SSI and STI shows similar pattern to that between SPI and STI. These results highlighted significant dependence among these indices. Thus, it is important to take into account the dependence in the assessment of compound drought and hot extremes.

3.3 Joint probability of compound extreme

In analyzing the compound drought and hot extreme based on the precipitation/soil moisture deficit and high temperature, the joint probability of SPI/SSI lower than certain thresholds and STI higher than certain thresholds is of particular interest. With the three extreme indices SSI, SPI, and STI represented by Y, X_1, and X_2, the joint probability $P(Y<y, X_1<x_1 \text{ and } X_2>x_2)$ of the compound extreme based on the meta-Gaussian model can be obtained as:

$$\begin{aligned} & P(Y < y, X_1 < x_1, X_2 > x_2) \\ & = P(Y < y, X_1 < x_1) - P(Y < y, X_1 < x_1, X_2 < x_2) \\ & = \Phi_2(y, x_1) - \Phi_3(y, x_1, x_2) \end{aligned} \quad \mathbf{(5)}$$

where the function Φ_2 (Φ_3) is the bivariate (trivariate) standard normal distribution function. The joint percentile (or probability) has been commonly used for the characterization of the overall condition of multiple hydrologic variables/extremes (Beersma and Buishand, 2004; Kao and Govindaraju, 2010; Chebana and Ouarda, 2011). In the context of multiple extreme indices, the joint percentile in equation **(5)** summarizes the joint condition from extreme indices and can be used to measure the severity of the compound drought and hot extremes. Specifically, the low probability indicates severe conditions of the compound extreme.

The joint probability estimated from equation (5) for the period August 2011 in U.S. is shown in Figure 2, along with the other three indices SPI, SSI, and STI. The SPI and SSI show severe drought conditions (with index values lower than -1) in large regions in southern U.S.. The STI shows anomaly high temperature (with index value higher than 1) in these regions. The extremely low SPI (and SSI) and high STI imply compound drought and hot extreme during this period. The joint percentile (or probability) for the period is extremely low (<0.05) in the southern U.S., particularly in Texas, indicating the severe condition of the compound extreme during 2011.

For the statistical assessment of compound extremes, it is also of interest to assess the probability of specified thresholds of the individual component. In this section, we focus on the concurrent extreme of lower SPI/SSI and higher STI. Here we specify the threshold value for SSI/SPI as 0, -0.5, -0.8 and -1.2 and that for STI as 0, 0.5, 0.8 and 1.5, which corresponds to 50^{th}, 30^{th}, 20^{th} and 10^{th} percentile for SPI/SSI (or 50^{th}, 70^{th}, 80^{th} and 90^{th} for STI). The threshold values -0.5, -0.8 and -1.2 correspond to the abnormally dry, moderate drought, and severe drought, as defined in the U.S. Drought Monitor (USDM)(Svoboda et al., 2002). The joint probability of this type of compound extreme in the U.S. is shown in Figure 3. We take the compound extreme with SPI/SSI<-0.8 and STI>0.8 as an example. It can be seen that the probability of compound drought and hot extreme is relatively high along the regions with significant negative correlations between SPI (or SSI) and STI. The is likely due to the interaction between the moisture deficit and high temperature in this region, which may induce the concurrent dry and hot extremes, leading to the high likelihood of compound extremes. Similar patterns of the low joint probability for other thresholds were also revealed from Figure 3. The difference is that with thresholds of the SPI and SSI get lower (and that of the STI gets higher), the joint probability of the compound extreme become even lower.

3.4 Conditional probability of compound extremes

The conditional probability enables the assessment of the impact of temperature and precipitation on the occurrence of agricultural drought. The conditional probability of soil moisture lower than 30^{th} percentile (or SSI<-0.5) given different precipitation deficit (SPI=-0.5, -1.2) and high temperature condition (STI=0.5,1.2) can be obtained from equation (2) and is shown in Figure 4(a-d). Comparing Figure 4(a) and (b) shows that the precipitation deficit significantly affects the agricultural drought in most regions in the U.S., which is easy to understand since meteorological drought is generally the prerequisite of the agricultural drought. Comparing Figure 4 (a) and (c) shows the impact of the temperature on the agricultural drought. It can be seen that significant changes of the probability are mainly in the High Plains region and southeastern regions, where the soil moisture-temperature interactions are most profound (Koster et al., 2009). The main reason is that, as stated before, the soil moisture deficit induces warm temperature due to the soil moisture-temperature interaction during summer. Meanwhile, the warm temperature and associated high evaporation may lead to dry soil moisture and further exacerbate the occurrence of drought, leading to the co-occurrence of drought and hot extreme. In the western region (with low correlation between SSI and STI), the changes of probability from Figure 4(a) and (c) is relatively small, which implies that the temperature does not play an important role (compared with precipitation) in affecting the agricultural drought in this region. Comparing Figure 4 (a) and (d) shows the combined impact from the both precipitation and temperature, implying that given the low precipitation and high temperature, the condition of the agricultural drought is expected to be even severe.

3.5 Conditional return period

A commonly used way to assess the likelihood of extremes is through the return period (RP) with respect to an exceedace or nonexceedance probability (P) of interest (RP=1/P). We use the 2011 Texas drought and hot extreme as an example to estimate the conditional return period. We first obtain the monthly precipitation, soil moisture and temperature of Texas by taking the average of all grids for the whole state. The three indices are then computed based on the statewide average of each variable. The scatter plot of these indices are show in Figure 5. A noticeable pattern is the negative (positive) dependence between SSI and STI (SPI), which implies the importance of analyzing the agricultural drought from a multivariate approach. The meta-Gaussian model is then constructed to estimate the conditional distribution and return period.

To assess the performance of the mega-Gaussian model in modeling these three indices, realizations of the random vector (Z_1, Z_2, Z_3) are generated and compared with the observations. The statistical simulation is achieved based on the conditional distribution method (Johnson, 1987; Aas et al., 2009). First, generate three uniform random values w_1, w_2 and w_3 between [0, 1], which can be achieved with the use of the random number generator function *rand* in MATLAB. The random variate z_1 can be generated by solving $z_1=Q^{-1}(w_1)$. Based on equation (2), one can derive the conditional distribution of Z_2 condition on Z_1, i.e., $F(Z_2|Z_1)$. The random variate z_2 can be generated by solving $w_2= F^{-1}(z_2|z_1)$. Similarly, the conditional distribution of Z_3 condition on Z_1 and Z_2, i.e., $G(Z_3|Z_1,Z_2)$, can be derived from equation (2). The random variate z_3 can then be generated by solving $w_3= G^{-1}(z_3|z_1,z_2)$. A sequence of simulations with the sample size n=150 from this procedure is shown in Figure 5(a-c). It can be seen that this model performs relatively well in modeling the dependence among different indices.

The conditional distribution of the SSI conditioned on the SPI and STI for August 2011 is estimated from equation (2), which is normal distribution and is shown in Figure 5(d). The conditional return period can be employed to assess the likelihood of the agricultural drought given meteorological drought and hot extreme. We compare the univariate return period of the agricultural drought and the conditional return period given the meteorological drought and hot extreme during this period. In the univariate case, the empirical return period of the soil moisture (or agricultural drought) for August 2011 is estimated as 64.5 years based on the Gringorten plotting position formula (Gringorten, 1963). The conditional return period of the agricultural drought with respect to the SPI and STI for August 2011 is estimated as 3.0 year. One can see that the conditional return period is much shorter than the univariate period. The reason is that given the low precipitation and high temperature, it is expected that the occurrence of the agricultural drought will be more frequently. The proposed framework thus provides a useful tool in estimating the likelihood of the compound extreme in this regard.

4 Conclusions

Due to the intrinsic nature of dependence among multiple contributing variables or processes of the compound extreme, statistical inference of the relationship among different extremes requires suitable dependence modeling. In this study, the meta-Gaussian model is proposed for the modeling of multiple extremes for the statistical analysis of the compound extreme in the U.S. based on monthly precipitation, soil moisture and temperature from NLDAS-2. The application of the model is illustrated for statistical assessments of the compound drought (with SPI and SSI as the indicator) and hot extreme (with STI as the indicator). The likelihood of the occurrence of compound drought and hot extreme is higher in regions, such as High Plains, with significant land atmosphere interactions. The compound drought and hot extreme during August 2011 in

Texas was used as a case study for the multivariate analysis. Results show the conditional return period of SSI with respect to the SPI and ST is around 3.0 years, which is much shorter than the univariate empirical return period 64.5 years.

The use of Meta-Gaussian method bears the advantage that the multivariate distribution in the high dimension can be derived straightforward with an explicit form. Though extreme indices may not strictly follow the multivariate normal assumption, the meta-Gaussian model is still a useful approximation to the parent distribution of the underlying data and many properties are robust to the departure from normality (Rencher and Christensen, 2012). This study mainly focused on the compound extreme defined by the precipitation deficit, soil moisture deficit and high temperature. It can be essentially applied to other types of compound extremes, such as the heat wave flash droughts characterized by the low soil moisture (SM) induced by high temperature and evapotranspiration (Mo and Lettenmaier, 2015; Otkin et al., 2018).

In this study, the SPI, SSI and STI based on the monthly precipitation, soil moisture and temperature are employed for characterizing the compound drought and hot extremes. Certain extremes, such as the flash drought and heat wave, may occur in a much shorter time scale. Thus, extreme indices of shorter time scales (e.g., daily, 5-day) may be used instead for the assessment of the compound extremes or impacts. In addition, we use the SPI, SSI and STI of the same month to define the compound extreme in this study. In reality, there may be some time lags between the meteorological drought and hydrological drought, between the agricultural drought and high temperature, or between extremes drivers and impacts (Mueller and Seneviratne, 2012; Zscheischler et al., 2014; Van Loon, 2015; Nicolai-Shaw et al., 2017). Thus, the time lag between the extremes (or impacts) should also be taken into account in statistical characterizations of the compound extreme and its impact based on the proposed model.

Moreover, the selection of time scales of SPI, SSI and STI was mainly based on the common practices in previous studies for illustration purposes in this study. These indices have been used to study drought or hot extremes and the relationship may vary among the indices of different time scales (or regions, layers of soil moisture) (Mo, 2008; Mishra and Singh, 2010; Sehgal et al., 2017). Thus careful assessments are needed in selecting the time scale of different indices to assess compound extremes of interest. Due to the limited availability of the soil moisture observations, the soil moisture simulated from land surface models was used in the statistical assessments of the compound extreme. This may induce uncertainties in the results (e.g., correlation pattern) and thus findings from this study should be interpreted with caveats.

The potential limitation of the proposed model is that it may not be quite accurate in characterizing the dependence at the extremal level (e.g., 99 percentile) (Wang et al., 2014). For the complicated dependence modeling (e.g., asymmetric dependence) in high dimensions, the vine copula or the pair-copula construction (PCC) is a promising alternative, which has been developed through decomposing the dependence structure into bivariate dependence that can be modeled with bivariate copulas (Aas et al., 2009; Brechmann and Schepsmeier, 2013; Liu et al., 2015; Liu et al., 2016). For example, the compound flood event resulting from the joint storm surge and high river runoff in Ravenna, Italy has been studied using the vine copula to construct the high-dimensional distribution (Bevacqua et al., 2017). We stress that the proposed method provides an alternative way to other tools (e.g., vine copula) for compound extreme analysis but does not mean to replace the currently used models. Statistical assessments of the multivariate behavior of different contributing variables of the compound extreme in this study may provide useful insights into the likelihood of compound extremes and aid the mitigation efforts under climate change.

5 Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC, No. 41601014). We thank the Environmental Modeling Center (EMC) of the National Centers for Environmental Prediction (NCEP), National Oceanic and Atmospheric Administration (NOAA) for providing NLDAS-2 datasets.

6 Reference

Aas, K., C. Czado, A. Frigessi, et al. (2009). Pair-copula constructions of multiple dependence. *Insur. Math. Econ.*, 44(2): 182-198.

Alexander, L. V. (2016). Global observed long-term changes in temperature and precipitation extremes: A review of progress and limitations in IPCC assessments and beyond. *Weather Clim. Extrem.*, 11: 4-16.

Beersma, J. J. and T. A. Buishand (2004). Joint probability of precipitation and discharge deficits in the Netherlands. *Water Resources Research*, 40, W12508.

Ben Alaya, M. A., F. Chebana and T. B. M. J. Ouarda (2014). Probabilistic Gaussian Copula Regression Model for Multisite and Multivariable Downscaling. *J. Climate*, 27(9): 3331-3347.

Berg, A., B. R. Lintner, K. L. Findell, et al. (2014). Impact of Soil Moisture–Atmosphere Interactions on Surface Temperature Distribution. *J. Climate*, 27(21): 7976-7993.

Bevacqua, E., D. Maraun, I. Hobæk Haff, et al. (2017). Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). *Hydrol. Earth Syst. Sci.*, 21(6): 2701-2723.

Brechmann, E. C. and U. Schepsmeier (2013). Modeling dependence with C-and D-vine copulas: The R-package CDVine. *Journal of Statistical Software*, 52(3): 1-27.

Chebana, F. and T. Ouarda (2011). Multivariate quantiles in hydrological frequency analysis. *Environmetrics*, 22(1): 63-78.

Cheng, L., M. Hoerling, A. AghaKouchak, et al. (2016). How Has Human-induced Climate Change Affected California Drought Risk? *J. Climate*, 29(1): 111-120.

Coumou, D. and S. Rahmstorf (2012). A decade of weather extremes. *Nat. Clim. Change.*, 2(7): 491-496.

Dai, A. (2011). Drought under global warming: a review. *Wiley Interdiscip. Rev.: Clim. Change*, 2(1): 45-65.

Durante, F. and G. Salvadori (2010). On the construction of multivariate extreme value models via copulas. *Environmetrics*, 21(2): 143-161.

Ford, T. W., D. B. McRoberts, S. M. Quiring, et al. (2015). On the utility of in situ soil moisture observations for flash drought early warning in Oklahoma, USA. *Geophys. Res. Lett.*, 42(22): 9790-9798.

Genest, C., A. C. Favre, J. Béliveau, et al. (2007). Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resour. Res*, 43: W09401.

Gringorten, I. I. (1963). A Plotting Rule for Extreme Probability Paper. *Journal of Geophysical Research*, 68(3): 813-814.

Hao, Z. and A. AghaKouchak (2013). Multivariate Standardized Drought Index: a parametric approach for drought analysis. *Advances in Water Resources*, 57: 12–18.

Hao, Z., A. AghaKouchak and T. J. Phillips (2013). Changes in concurrent monthly precipitation and temperature extremes. *Environ Res Lett*, 8(3): 034014.

Hao, Z., F. Hao, V. P. Singh, et al. (2016). Probabilistic prediction of hydrologic drought using a conditional probability approach based on the meta-Gaussian model. *J. Hydrol.*, 542: 772-780.

Harris, I., P. D. Jones, T. J. Osborn, et al. (2013). Updated high - resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.*, 34(3): 623-642.

Hawkes, P. J. (2008). Joint probability analysis for estimation of extremes. *Journal of Hydraulic Research*, 46(S2): 246-256.

Heim, R. R. (2015). An overview of weather and climate extremes – Products and trends. *Weather Clim. Extrem.*, 10: 1-9.

Herr, H. D. and R. Krzysztofowicz (2005). Generic probability distribution of rainfall in space: The bivariate model. *J. Hydrol.*, 306(1): 234-263.

Hirschi, M., S. I. Seneviratne, V. Alexandrov, et al. (2011). Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.*, 4(1): 17-21.

IPCC (2012). Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. C. B. Field, V. Barros, T.F. Stocker,D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen,M. Tignor, and P.M. Midgley. Cambridge, UK, and New York, NY, USA, Cambridge University Press**:** 582.

Johnson, M. (1987). Multivariate Statistical Simulation. John Wiley & Sons, New York.

Kao, S. C. and R. S. Govindaraju (2008). Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. *Water Resour. Res.*, 44(2): W02415

Kao, S. C. and R. S. Govindaraju (2010). A copula-based joint deficit index for droughts. *J. Hydrol.*, 380(1-2): 121-134.

Kelly, K. and R. Krzysztofowicz (1997). A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics*, 11(1): 17-31.

Kew, S. F., F. M. Selten, G. Lenderink, et al. (2013). The simultaneous occurrence of surge and discharge extremes for the Rhine delta. *Nat. Hazards Earth Syst. Sci.*, 13(8): 2017-2029.

Keyantash, J. and J. A. Dracup (2002). The quantification of drought: An evaluation of drought indices. *Bull. Amer. Meteor. Soc*, 83: 1167-1180.

Koster, R., S. Schubert and M. Suarez (2009). Analyzing the concurrence of meteorological droughts and warm periods, with implications for the determination of evaporative regime. *J. Climate*, 22(12): 3331-3341.

Leonard, M., S. Westra, A. Phatak, et al. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdiscip. Rev.: Clim. Change*, 5(1): 113–128.

Liu, Z., T. Törnros and L. Menzel (2016). A probabilistic prediction network for hydrological drought identification and environmental flow assessment. *Water Resour. Res.*, 52(8): 6243-6262.

Liu, Z., P. Zhou, X. Chen, et al. (2015). A multivariate conditional model for streamflow prediction and spatial precipitation refinement. *J. Geophys. Res-Atmos.*, 120(19): 10116-10129.

Martius, O., S. Pfahl and C. Chevalier (2016). A global quantification of compound precipitation and wind extremes. *Geophys. Res. Lett.*, 43(14): 7709-7717.

Mazdiyasni, O. and A. AghaKouchak (2015). Substantial increase in concurrent droughts and heatwaves in the United States. *Proc. Natl. Acad. Sci.*, 112(37): 11484-11489.

McKee, T. B., N. J. Doesken and J. Kleist (1993). The relationship of drought frequency and duration to time scales. Eighth Conference on Applied Climatology, Am. Meteorol. Soc., Anaheim, CA.

Mishra, A. K. and V. P. Singh (2010). A review of drought concepts. *J. Hydrol.*, 391(1-2): 202-216.

Mo, K. C. (2008). Model-based drought indices over the United States. *J. Hydrometeorol.*, 9(6): 1212-1230.

Mo, K. C. and D. P. Lettenmaier (2015). Heat wave flash droughts in decline. *Geophys. Res. Lett.*, 42: 2823–2829.

Mo, K. C. and D. P. Lettenmaier (2016). Precipitation deficit flash droughts over the United States. *J. Hydrometeorol.*, 17(4): 1169-1184.

Montanari, A. and A. Brath (2004). A stochastic approach for assessing the uncertainty of rainfall - runoff simulations. *Water Resour. Res.*, 40(1): W01106.

Mueller, B. and S. I. Seneviratne (2012). Hot days induced by precipitation deficits at the global scale. *Proc. Natl. Acad. Sci.*, 109(31): 12398-12403.

Nicolai-Shaw, N., J. Zscheischler, M. Hirschi, et al. (2017). A drought event composite analysis using satellite remote-sensing based soil moisture. *Remote Sensing of Environment*, 203: 216-225.

Otkin, J. A., M. Svoboda, E. D. Hunt, et al. (2018). Flash Droughts: A Review and Assessment of the Challenges Imposed by Rapid-Onset Droughts in the United States. *Bull. Amer. Meteor. Soc*, 99(5): 911-919.

Renard, B. and M. Lang (2007). Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources*, 30(4): 897-912.

Rencher, A. C. and W. F. Christensen (2012). Methods of multivariate analysis. John Wiley & Sons.

Rueda, A., P. Camus, A. Tomás, et al. (2016). A multivariate extreme wave and storm surge climate emulator based on weather patterns. *Ocean Modelling*, 104: 242-251.

Rummukainen, M. (2012). Changes in climate and weather extremes in the 21st century. *Wiley Interdiscip. Rev.: Clim. Change*, 3(2): 115–129.

Salvadori, G., C. de Michele, N. Kottegoda, et al. (2007). Extremes in Nature: An Approach Using Copulas. Springer, New York.

Sehgal, V., V. Sridhar and A. Tyagi (2017). Stratified drought analysis using a stochastic ensemble of simulated and in-situ soil moisture observations. *J. Hydrol.*, 545: 226-250.

Seneviratne, S. I., T. Corti, E. L. Davin, et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-sci. Rev.*, 99(3-4): 125-161.

Seneviratne, S. I., N. Nicholls, D. Easterling, et al. (2012). Changes in climate extremes and their impacts on the natural physical environment. In: Field CB, Barros V, Stocker TF, Qin D, Dokken

D, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, et al, eds. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC) (Eds.). Cambridge University Press. Cambridge, UK, pp**.** 109–230.

Serinaldi, F. (2016). Can we tell more than we can know? The limits of bivariate drought analyses in the United States. *Stoch. Env. Res. Risk. A.,* 30(6): 1691-1714.

Sharma, S. and P. Mujumdar (2017). Increasing frequency and spatial extent of concurrent meteorological droughts and heatwaves in India. *Sci. Rep.*, 7(1): 15582.

Sheffield, J., G. Goteti, F. Wen, et al. (2004). A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res-Atmos.*, 109: D24108.

Song, S. and V. P. Singh (2010). Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. *Stoch. Env. Res. Risk. A.*, 24(3): 425-444.

Stagge, J. H., L. M. Tallaksen, L. Gudmundsson, et al. (2015). Candidate distributions for climatological drought indices (SPI and SPEI). *Int. J. Climatol.*, 35(13): 4027-4040.

Sun, A. Y., Y. Xia, T. G. Caldwell, et al. (2018). Patterns of precipitation and soil moisture extremes in Texas, US: A complex network analysis. *Advances in Water Resources*, 112: 203-213.

Svoboda, M., M. Hayes and D. Wood (2012). Standardized precipitation index user guide. World Meteorological Organization Geneva, Switzerland.

Svoboda, M., D. LeComte, M. Hayes, et al. (2002). The drought monitor. *Bull. Amer. Meteor. Soc*, 83(8): 1181-1190.

Thilakarathne, M. and V. Sridhar (2017). Characterization of future drought conditions in the Lower Mekong River Basin. *Weather Clim. Extrem.*, 17: 47-58.

van den Hurk, B., E. van Meijgaard, P. de Valk, et al. (2015). Analysis of a compounding surge and precipitation event in the Netherlands. *Environ Res Lett*, 10(3): 035001.

Van Loon, A. F. (2015). Hydrological drought explained. *Wiley Interdiscip. Rev.: Water*, 2: 359–392.

Vicente-Serrano, S. M. and S. Beguería (2016). Comment on ‘Candidate distributions for climatological drought indices (SPI and SPEI)’ by James H. Stagge et al. *Int. J. Climatol.*, 36(4): 2120-2131.

Vogl, S., P. Laux, W. Qiu, et al. (2012). Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. *Hydrol. Earth Syst. Sci*, 16(7): 2311-2328.

Wahl, T., S. Jain, J. Bender, et al. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nat. Clim. Change.*, 5(12): 1093-1097.

Wang, Z., J. Yan and X. Zhang (2014). Incorporating spatial dependence in regional frequency analysis. *Water Resour. Res.*, 50(12): 9570-9585.

Whan, K., J. Zscheischler, R. Orth, et al. (2015). Impact of soil moisture on extreme maximum temperatures in Europe. *Weather Clim. Extrem.*, 9: 57-67.

Wilks, D. S. (2011). Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego, CA.

Willmott, C. J. and K. Matsuura (2001). Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999),http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html.

Wong, G., M. F. Lambert, M. Leonard, et al. (2010). Drought analysis using trivariate copulas conditional on climatic states. *Journal of Hydrologic Engineering*, 15(2): 129-141.

Wu, L., D.-J. Seo, J. Demargne, et al. (2011). Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. *J. Hydrol.*, 399(3): 281-298.

Xia, Y., M. Ek, H. Wei, et al. (2012). Comparative analysis of relationships between NLDAS - 2 forcings and model outputs. *Hydrol.Proccess*, 26: 467-474.

Xia, Y., M. B. Ek, C. D. Peters-Lidard, et al. (2014). Application of USDM Statistics in NLDAS-2: Optimal Blended NLDAS Drought Index Over the Continental United States. *J. Geophys. Res-Atmos.*, 119(6): 2947–2965.

Xia, Y., K. Mitchell, M. Ek, et al. (2012). Continental - scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS - 2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, 117(D3).

Xia, Y., J. Sheffield, M. B. Ek, et al. (2014). Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.*, 512: 107-125.

Zhang, L. and V. P. Singh (2007). Gumbel–Hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering*, 12(4): 409-419.

Zscheischler, J., A. M. Michalak, C. Schwalm, et al. (2014). Impact of large - scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data. *Global Biogeochem. Cy*, 28(6): 585-600.

Zscheischler, J. and S. I. Seneviratne (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, 3(6): e1700263.

Zscheischler, J., S. Westra, B. J. J. M. van den Hurk, et al. (2018). Future climate risk from compound events. *Nat. Clim. Change.*, 8(6): 469-477.

7 Figure

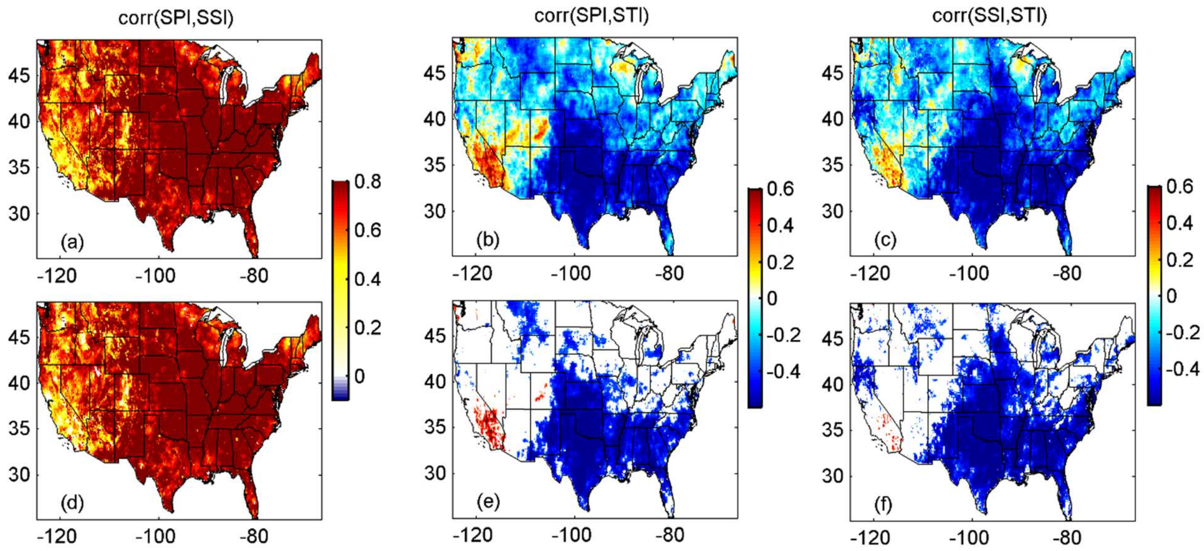

Figure 1 Dependence pattern among SPI, SSI and STI of August for the period from 1979-2014 in the U.S. . (a)-(c) Correlation coefficients among three indices. (d)-(f) Significant correlation coefficients at a 5% significance level among three indices.

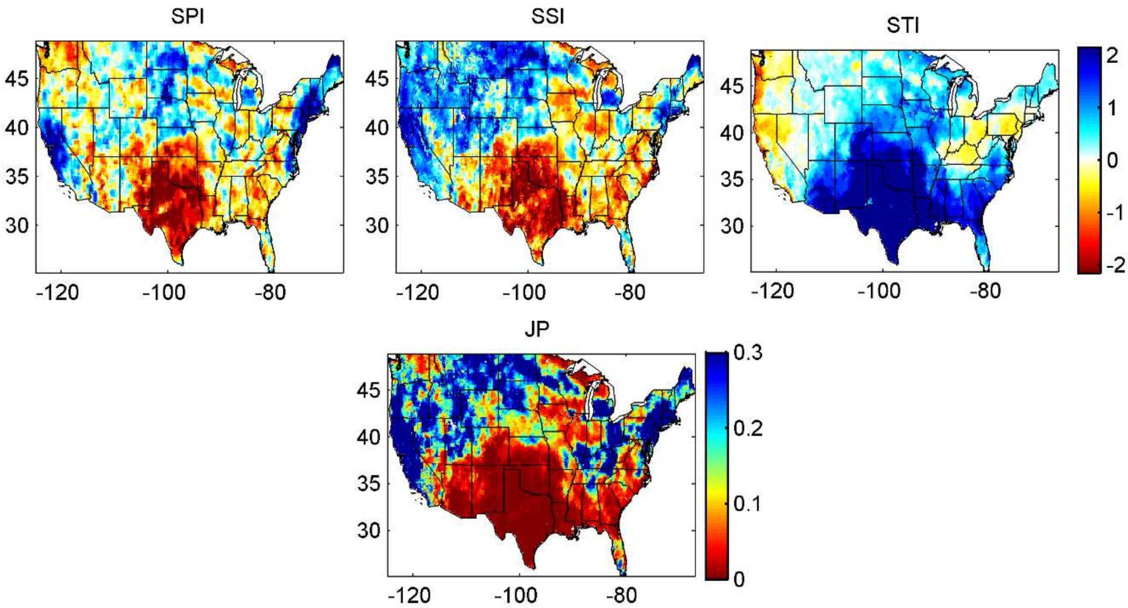

Figure 2 The SPI, SSI and STI values and the joint probability (JP) for August 2011 in the U.S.. The joint probability (JP) of the compound drought and hot extreme is defined as P(SSI $<y$, SPI $<x_1$ and STI $>x_2$) , where y, x_1 and x_2 are the SSI, SPI and STI value of August 2011.

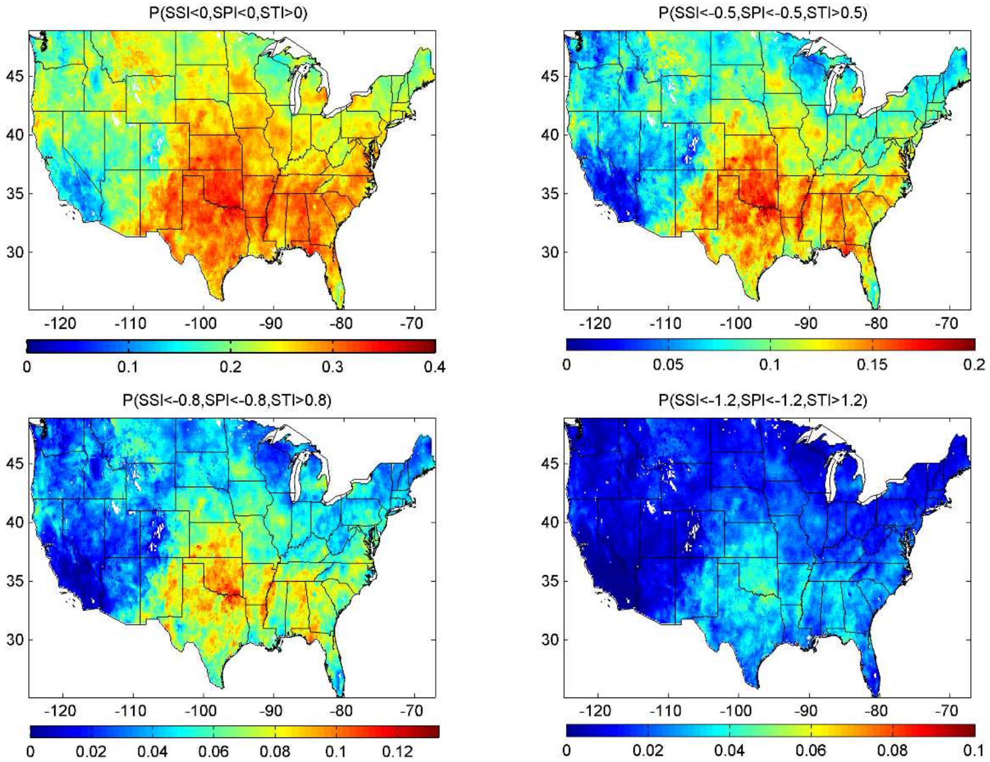

Figure 3 The probability of the compound extreme of low SPI/SSI and high STI with threshold values 0, -0.5, -0.8 and -1.2 for SPI/SSI and 0, 0.5, 0.8 for STI.

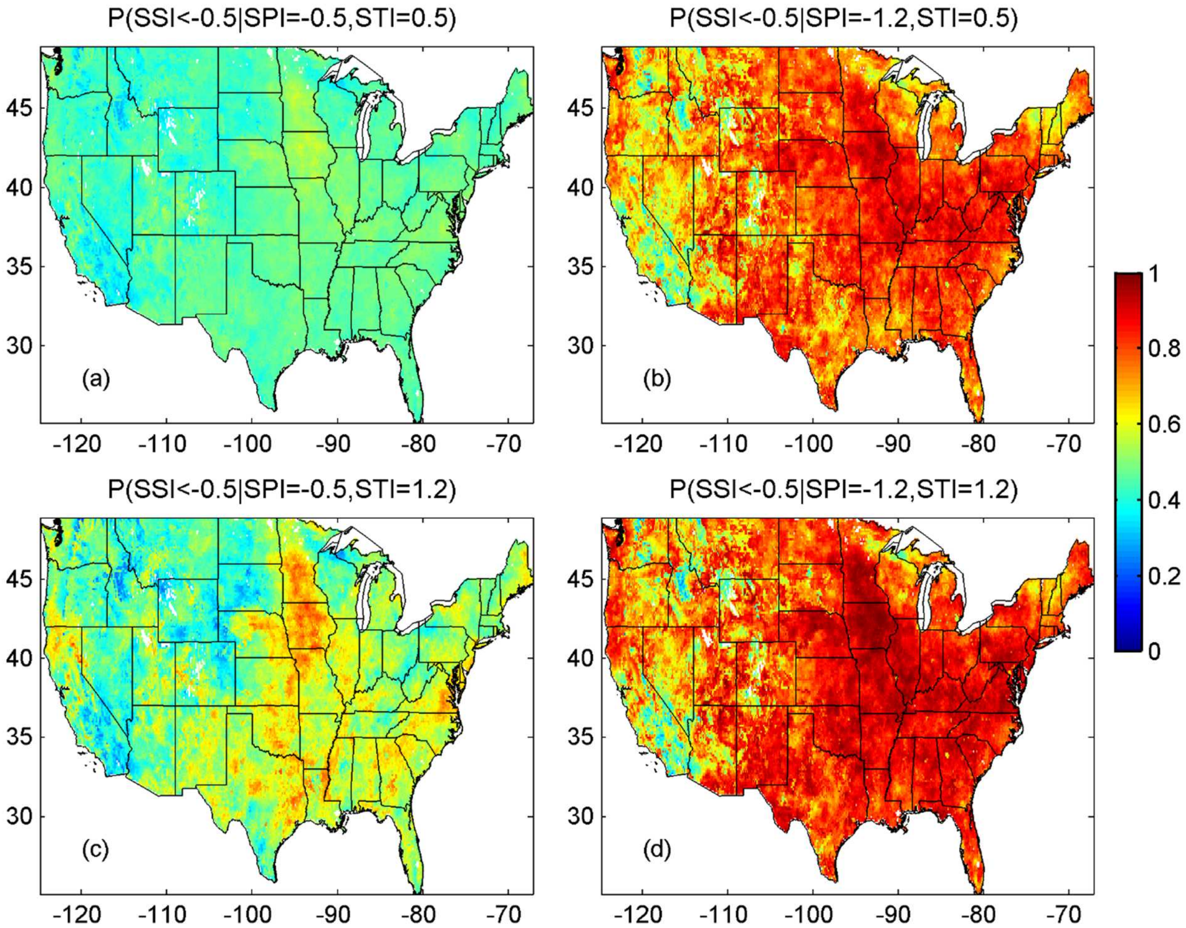

Figure 4 The conditional probability of SSI<-0.5 given different values of SPI and STI. (a) *P*(SSI<-0.5 |SPI=-0.5, STI=0.5); (b) *P* (SSI<-0.5|SPI=-1.2, STI=0.5); (c) *P* (SSI<-0.5 |SPI=-0.5, STI=1.2) and (d) *P* (SSI<-0.5 |SPI=-1.2, STI=1.2).

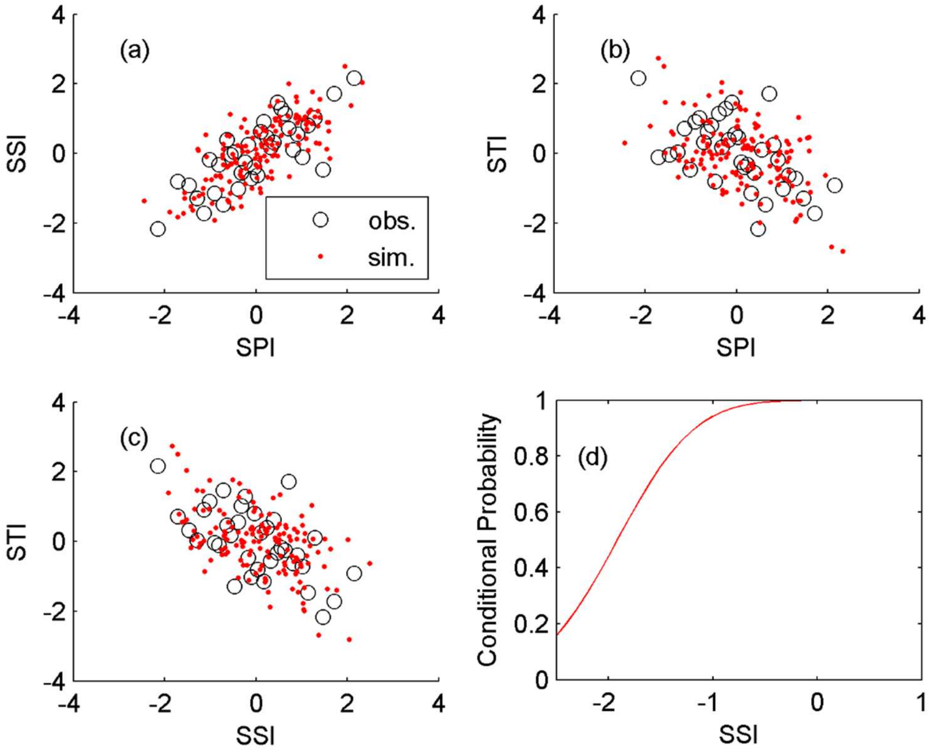

Figure 5 Assessments of the agricultural drought conditioned on SPI and STI for the period August 2011 in Texas, U.S.. (a-c) Scatterplots of observed and simulated pairs of SPI, SSI and STI. (d) Conditional probability of SSI conditioned on SPI and STI for August 2011 in Texas.