Linking drought indices to impacts to support drought risk assessment in Liaoning province, China

Yaxu Wang[1,2,3], Juan Lv[1,2], Jamie Hannaford [3,4], Yicheng Wang[1,2], Hongquan Sun[1,2], Lucy J. Barker[3], Miaomiao Ma[1,2], Zhicheng Su[1,2], Michael Eastman[3]

[1]China Institute of Water Resources and Hydropower Research, Beijing 100038, China

[2]Research Center on Flood and Drought Disaster Reduction of the Ministry of Water Resources, Beijing 100038, China

[3]Centre for Ecology & Hydrology, Oxfordshire, OX10 8BB, UK

[4]Irish Climate Analysis and Research UnitS (ICARUS), Maynooth University, Dublin, W23 F2K8, Ireland

Correspondence to： Juan Lv (lujuan@iwhr.com)

Abstract. Drought is a ubiquitous and reoccurring hazard that has wide ranging impacts on society, agriculture and the environment. Drought indices are vital for characterizing the nature and severity of drought hazards, and there have been extensive efforts to identify the most suitable drought indices for drought monitoring and risk assessments. However, to date, little effort has been made to explore which index(s) best represents drought impacts for various sectors in China. This is a critical knowledge gap, as impacts provide important 'ground truth' information. They can be used to demonstrate whether drought indices (used for monitoring or risk assessment) are relevant for identifying impacts, thus highlighting if an area is vulnerable to drought of a given severity. The aim of this study is to explore the link between drought indices and drought impacts, using Liaoning province (northeast China) as a case study due to its history of drought occurrence. To achieve this we use independent, but complementary, methods (correlation and random forest analysis). Using multiple drought indices – Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), Soil Moisture (SoilM) and the Normalized Difference Vegetation Index (NDVI) – and drought impact data (on crop yield, livestock, rural people and the economy) correlation and random forest analysis were used to identify which indices link best to the recorded drought impacts for cities in Liaoning. The results show that the relationship varies between different categories of drought impacts and between cities. SPEI with a 6-month accumulation (SPEI6) had a strong correlation with all categories of drought impacts, while SPI12 had a weak correlation with drought impacts. Of the impact datasets, 'drought suffering area' and 'drought impact area' had a slightly strong relationship with all drought indices in Liaoning province, while 'population and number of livestock with difficulty in accessing drinking water' had weak correlations with the indices. Based on the linkage, drought vulnerability was analyzed using various vulnerability factors. Crop cultivated area was positively correlated to the drought vulnerability for five out of the eight categories of drought impacts, while the total population had a strong negative relationship with drought vulnerability for half the drought impact categories. This study can support drought planning efforts in the region, and provides a methodology for application for other regions of China (and other countries) in the future, as well as providing context for the indices used in drought monitoring applications, so enabling improved preparedness for drought impacts.

1 Introduction

Drought is one of the most pervasive natural hazards with some of the greatest societal impacts (Belal et al., 2014), but is challenging to understand, quantify and manage. These challenges arise from the typically wide spatial extent of droughts, their frequent occurrence and the non-structural, diffuse and delayed nature of drought impacts (Biswas et al., 2013;Mishra and Singh, 2010). China has experienced numerous droughts, which have caused great economic losses since the 1950s, especially in Liaoning province in the dry northeast of the country (Zhang, 2004). From spring 2000 to autumn 2001, Liaoning province experienced a severe drought, which captured a large amount of attention from stakeholders and caused serious impacts on many sectors because of the successive years of drought (Chen et al., 2016).

The costly nature of droughts means it is essential to plan and prepare for droughts proactively. Drought risk assessment is an essential prerequisite of this proactive approach (Wilhite and Buchanan-Smith, 2005;Wilhite et al., 2000), providing methods to predict the potential drought risk to society and the environment. There are numerous approaches to drought risk assessment, and these can be grouped into two broad classes: one based on the definition of drought risk, which combines the frequency of drought and the possible drought impacts. The other is an assessment method for establishing indices to measure the hazard, vulnerability and exposure of drought (Jin et al., 2016). The majority of risk assessment efforts focus primarily on meteorological indices of drought, e.g. assessing the risk of a given severity of meteorological drought using historical precipitation data. However, to adequately define risk it is also necessary to characterize the consequences of drought occurrence, i.e. the impacts of drought on society, the economy and the environment.

A wealth of drought indices have been used in the literature (Lloyd-Hughes, 2014), although predominantly for drought monitoring and early warning (e.g. the review of Bachmair et al. 2016b) rather than risk assessment. The range of drought indices reflects the different types of drought which can be monitored, e.g., meteorological, hydrological and agricultural; the selected index should reflect the type of drought one wishes to monitor and manage. Many indices, such as the Standardized Precipitation Index (SPI), can be calculated over different time scales. This enables deficits to be assessed over different periods, and can help monitor different types of drought. For example, shorter time scales, such as the SPI for three or six months are used for agricultural drought monitoring while SPI values for 12 or 24 months are normally applied to hydrological drought monitoring (Hong et al., 2001;Seiler et al., 2002). The relationship between drought indices and drought impacts, established by a correlation or some other similar analysis (e.g. Bachmair et al. 2016a), can thus be used for drought risk assessment and appraisal of vulnerability. Vulnerability is by its nature difficult to define and measure, but in effect, drought impacts provide a proxy for vulnerability by demonstrating adverse consequences of a given drought severity (Stahl et al., 2016).

There are many different types of drought impact affecting many aspects of society and the environment, but drought impacts are rarely systematically recorded (Bachmair et al., 2016b). Some countries and regions have established drought impact recording systems to analyze historical drought impacts. A leading example of this is the US Drought Impacts Reporter

(Svoboda and Hayes, 2011) which was launched as a web-based system in July 2005. More recently, the European Drought Impact report Inventory (EDII) has been established (Stahl et al., 2016). Such databases are an important step forward, but the information in them is necessarily partial and biased, being effectively crowd-sourced text-based information based on 'reported' impacts a range of sources (the media, grey literature, etc.). In contrast to many other countries, China has a relatively complete and systematically assembled, quantitative drought impact information collection system. Data are collected and checked at the county level by the Drought Resistance Department via a formalized network of reporters, who collect drought impacts statistics in every village. These data then are fed up to the national government and held by the State Flood Control and Drought Relief Headquarters (SFDH). This consistent collection of impact reporting provides a rich resource for drought risk assessment. However, impacts by themselves are not fully informative; to help inform risk assessment there is a need to understand their relationship with quantitative drought indices.

Understanding the relationship between drought indices and drought impacts, and drought vulnerability, is a vital step to improve drought risk management (Hong and Wilhite, 2004). However, whilst there have been many studies developing, applying and validating drought indices, relatively few studies have assessed the link between indices and observed impacts. Bachmair et al. (2016a) noted that this literature tended to be dominated by studies focused on agricultural drought, linking generally indices like the SPI/SPEI and crop yield. Examples appraising multi-sectoral impacts are much sparser – recent studies tend to be in Europe, utilizing the EDII. Bachmair et al. (2014) and Bachmair (2016b) used drought impacts from the EDII, and various time scales of SPI, SPEI and streamflow percentiles. They found that the relationships between indices and impacts varied significantly by region, season, impact types, etc. Whilst Blauhut et al. (2015) developed a quantitative relationship between drought impact occurrence and SPEI using logistic regression in four European regions. They assumed drought impacts were only measured by the drought impact occurrence, meaning that all drought impacts have equal weight without considering the duration, intensity or spatial extent of the impacts; a similar logistic regression approach was also used by Stagge (Stagge et al., 2014). Karavitis et al. (2014) described drought impacts transformed into monetary losses to measure drought impacts in Greece. However, it is challenging to transform all drought impacts into monetary units – especially the indirect impacts of droughts.

In China, previous studies have also focused on agricultural drought. Zhao et al. (2011) established the relationship between drought frequency and simulated crop yield data in Henan Plain. Jia et al. (2011) used the water stress coefficient and duration to establish a drought index. Li et al. (2009) analyzed the links between historical crop yield and meteorological drought and established a meteorological drought risk index by combining the drought frequency, intensity, yield loss and extent of irrigation. The drought index was found to explain 60-75% of the major crop yield reduction.

In summary, previous studies have focused on linking impacts to only one characteristic of drought (such as intensity, duration of occurrence) with most focusing on meteorological drought and agricultural impacts. But with the exception of Blauhut et al. (2015), there is little application of the results to drought vulnerability assessments. Here we link drought indices to drought impacts in 14 cities in Liaoning province, northeast China, showcasing the use of the Chinese drought impact data from the SFDH. Using the drought impact-index linkage, we evaluate the drought vulnerability in Liaoning province and assess what factors affect drought vulnerability. A drought vulnerability evaluation method that can be extended to other areas is then developed. The objectives of this paper are:

1. To identify when and where the most severe droughts occurred between 1990 and 2013 in Liaoning province;
2. To identify which drought indices best link to drought impacts in Liaoning province;
3. To determine which city or area has higher drought vulnerability, based on the correlation analysis from objective 2, in Liaoning province; and,
4. To ascertain which vulnerability factor or set of vulnerability factors have a higher contribution to drought vulnerability, as quantified in objective 3.

2 Materials

2.1 Study area

Located in the northeastern of China, Liaoning province, comprised of 14 cities, has a temperate continental monsoon climate with an annual average precipitation of 686.4mm, which is unevenly distributed both temporally and spatially (Cai et al., 2015). Figure 1 shows the annual average rainfall across Liaoning, the south-east receiving on average more than 1000mm a year, whilst the north-west receives less than 500mm per year.

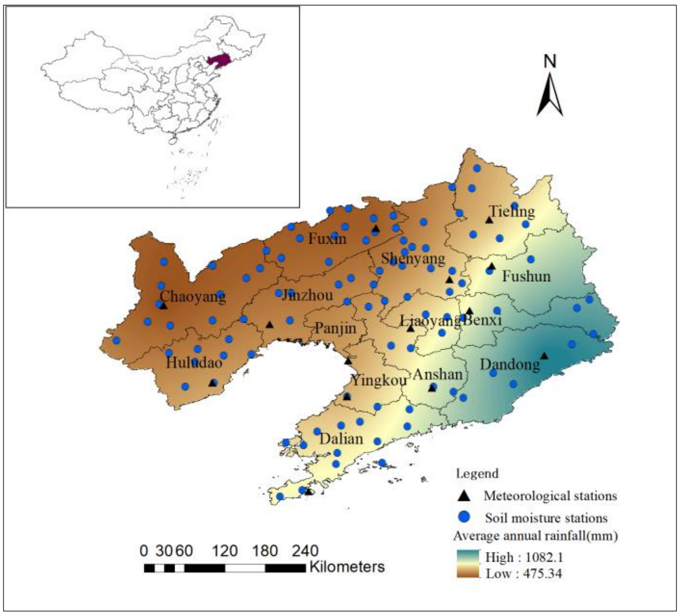

Figure 1: Map showing the 14 cities, the distribution of meteorological and soil moisture stations and the average annual precipitation in Liaoning province.

The annual average volume of water resources is 34.179 billion m^3, and the annual average per capita water resources is 769

m^3 – about one-third of the per capita water resources for the whole China. Thus, Liaoning is one of the severe water-shortage provinces in northern China. Liaoning province is also a highly productive area for agriculture. Spring maize is the dominant crop in agriculture production which makes it an important high-quality maize production area (Liu et al., 2013;Ren and Zhou, 2009). Due to these characteristics, when drought occurs, as has frequently been the case in Liaoning province, it causes a significant reduction in agricultural production (Yan et al., 2012). According to the SFDH, between 2000 and 2016 the average annual yield loss due to drought was 1.89 million tons in Liaoning province, with an average annual direct agricultural economic loss of 1.87 billion yuan.

2.2 Data

1) Meteorological data

Daily precipitation and temperature data for each city in Liaoning province for the period 1990-2013 were obtained from the China Meteorological Administration (http://data.cma.cn/), including daily precipitation and temperature. Although there are 52 meteorological stations in Liaoning province, due to the quality and length of the records, and location of the stations, one representative meteorological site in each city was selected to represent the meteorological condition for the whole city in order to derive drought indices.

2) Soil moisture data

Daily soil moisture data for 96 soil moisture stations in Liaoning province from 1990 to 2006 were obtained from Liaoning Provincial Department of Water Resources. Daily soil moisture was measured at three different depths: 10cm, 20cm and 30cm using frequency domain reflection soil moisture sensors, which are based on the principle of electromagnetic pulse. Soil moisture data were not available between November and February at most stations due to freezing conditions.

3) Normalised Difference Vegetation Index (NDVI) data

Monthly MODIS NDVI data from 2000 to 2013 was collected in Liaoning province from the Geospatial Data Cloud (http://www.gscloud.cn/); the daily maximum data were used to derive the monthly average NDVI.

4) Impact data

In contrast to many other countries, China has a systematic, centralized drought impact information collection system. Drought statistics include drought impacts, drought mitigation actions and benefits of action to agriculture, hydrology and civil affairs. During a drought event, impact statistics are collected from every day to every three weeks, according to the drought warning level (Wang, 2014). When a drought warning is not triggered, drought impact data are collected after an event has ended which could be several months afterwards; and no data are collected when there is no drought event. Statistics for eight drought impact types were collected from the SFDH between 1990 and 2016, and aggregated to annual totals, the impact types used are listed in Table 1.

Table 1: The eight drought impacts used in this study collected by the SFDH for Liaoning province.

Impact	Abbreviation	Description	Unit
Drought suffering area	DSA	The area that was officially declared in drought.	kha
Drought impacted area	DIA	The area that suffered crop yield loss by 10% or more	kha
Disaster area	DA	The area that suffered crop yield loss by 30% or more.	kha
Recessed area	RA	The area that suffered crop yield loss by 80% or more.	kha
Population with difficulty in accessing drinking water	PHD	Rural populations that cannot access normally to drinking water.	10k
Number of livestock with difficulty in accessing drinking water	NLH	Number of livestock that cannot access normally to drinking water.	10k
Yield loss due to drought	YLD	The amount of yield losses due to drought.	10k ton
Direct economic loss in agriculture	DELA	Direct losses of agricultural economy caused by drought.	0.1b yuan

5) Vulnerability factors

Vulnerability factors were collected from the 2017 Liaoning province Statistical Yearbook to explain the drought vulnerability (Liaoning Province Bureau of Statistical, 2017). The drought impacts described above are mainly focused on agriculture, rural populations, agricultural productivity and the agricultural economy; therefore, factors relevant to these sectors were selected. The selected vulnerability factors and data from the 2017 Liaoning Statistical Yearbook are shown in Table 2.

Table 2: Vulnerability factors for Liaoning province collected from the 2017 Liaoning Statistical Yearbook(Liaoning Province Bureau of Statistical, 2017)

City	Per capita gross domestic product(k yuan)	Population (10k)	Crop cultivated area(kha)	Annual per capita water supply(m^3)	Per unit area of Fertilizer application(kg/ha)	Effective irrigation rate (%)	Number of electromechanical wells(k)	Reservoir total storage capacity(m m^3)	Per unit area of major agricultural products(kg/ha)	Livestock production (10k ton)
Shenyang	755.8	733.9	656.0	91.5	1000.4	40.0	27.6	686.6	7090.5	64.5
Dalian	1143.4	595.6	327.0	73.4	1437.2	22.8	19.0	2523.0	4914.3	70.8
Anshan	422.9	345.7	247.7	42.3	1031.8	30.1	4.1	91.9	6641.6	36.7
Fushun	402.7	214.8	116.1	94.7	776.9	37.4	1.8	2575.5	6342.9	10.4
Benxi	511.1	150.0	58.0	167.9	756.3	29.9	0.4	6078.8	6606.3	9.3
Dandong	315.8	237.9	190.4	28.0	1049.7	41.7	1.4	16202.8	6056.9	20.2
Jinzhou	341.8	302.2	457.2	46.6	915.4	41.3	18.7	977.9	6825.7	64.0
Yingkou	496.7	232.8	109.4	42.4	1564.6	67.7	12.3	269.6	7325.0	13.5
Fuxin	215.9	188.9	479.4	39.7	881.9	30.1	26.6	545.0	5243.6	49.6
Liaoyang	373.4	178.6	162.8	42.4	1002.6	44.8	4.0	1418.8	7202.2	11.0
Panjin	778.3	130.1	143.0	70.2	937.0	68.7	1.0	141.5	8918.3	23.8

Tieling	196.5	299.9	548.5	12.2	960.2	32.0	18.1	2174.5	8397.1	46.0
Chaoyang	210.1	341.1	464.5	15.8	874.7	42.0	17.4	2085.6	6292.0	63.6
Huludao	230.8	280.5	249.7	18.7	976.8	28.9	14.0	892.7	4852.3	35.4

2.3 Methods

1) Drought indices

Two meteorological indices were selected, Standardized Precipitation Index (SPI; McKee et al., 1993) and Standardized Precipitation Evapotranspiration Index (SPEI;Vicente-Serrano et al., 2010). These standardized indices are widely used in drought monitoring applications, and the World Meteorological Organization recommend the use of the SPI to monitor meteorological drought (Hayes et al., 2011). This is due to the relatively simple calculation, flexibility of calculation at different time scales, and the fact it can be compared across time and space.

The SPI, in its default formulation, assumes that precipitation obeys the Gamma (Γ) skewed distribution, which is used to transform the precipitation time series into a normal distribution. After normalization, classes of drought can be defined with the cumulative precipitation frequency distribution (Botterill and Hayes, 2012;Hayes et al., 1999). The SPEI is a very similar concept, using the climatic water balance (that is, precipitation minus potential evapotranspiration, PE). Here, PE is calculated by the Thornthwaite method (Thornthwaite, 1948), using observed temperature and sunlight hours (estimated from latitude) as inputs. SPEI are calculated by normalizing the climatic water balance using a log-logistic probability distribution (Yu et al., 2014).

SPI and SPEI are easily calculated and can fit a wide range of time scales (e.g. 1, 3, 12, 24, 72 months) of interest (Edwards, 1997). SPEI has the added advantages of characterizing the effects of temperature and evapotranspiration on drought. In this study, SPI and SPEI were calculated for five accumulation periods (6, 12, 15, 18 and 24-months) from 1990 to 2013 for 14 meteorological stations (i.e. one in each city). In Liaoning province, precipitation is concentrated between April and September; this is also when the growth stage of spring maize occurs. Considering the climatology and crop growth period, SPI6 and SPEI6 ending in September were selected, i.e. calculated using precipitation during April to September. The 12, 15, 18 and 24 months SPI and SPEI in ending December were analyzed with the annual drought impacts.

Using the daily soil moisture of 10 cm, 20 cm and 30 cm depths, the daily average soil moisture for each station was calculated using Eq. (1) and Eq. (2) (Lin et al., 2016).

$$\theta_1=\theta_{10} \qquad \theta_2=\frac{\theta_{10}+\theta_{2}}{2} \qquad \theta_3=\frac{\theta_{20}+\theta_{30}}{2} \tag{1}$$

$$\bar{\theta}=\frac{\sum_{i=1}^{3}(\theta_2\times h_i)}{H} \tag{2}$$

Where θ_i is the soil moisture of the *i*-th layer (i=1, 2, 3). θ_{10}, θ_{20} and θ_{30} are the measured value at different depths (10cm, 20cm and 30cm). $\bar{\theta}$ is the average soil moisture. h_i is the thickness of the *i*-th layer of soil, and H is the total thickness of the measured soil.

Some of the daily soil moisture data were missing, however, this was limited to 17% of total soil moisture data. In some cases there were missing data for one depth of soil moisture measurement. In these cases, the average soil moisture of the other two layers was calculated, and where there was only one layer of soil moisture available it was used to represent the average soil moisture. The annual average soil moisture was calculated based on the available daily soil moisture (March to October) and was analyzed with the annual drought impact data. As each city has more than one station, the annual soil moisture of each station was calculated and then averaged into one value for each city.

The area-averaged NDVI at city unit was calculated based on the monthly NDVI. The critical stages of the spring maize growth in Liaoning is in July, so the area-averaged NDVI in July was selected for the analysis with the annual drought impacts.

2) Correlation analysis

The Pearson correlation method was used to characterize the correlation between indices and various drought impacts (Özger et al., 2009). Due to the limited availability of soil moisture data, correlation analysis of soil moisture and drought impact data was only carried out in 9 cities. The linkage between drought indices and impacts was used to assess the drought vulnerability in Liaoning province. It can be inferred that the greater impact caused by the same severity of drought (as measured by the relevant index e.g. SPI/SPEI), the higher drought vulnerability of the city.

3) Random forest modeling

Random forest (RF) is an algorithm that consists of a series of independent decision trees. RFs can be used for classification and regression (Sethi et al., 2012). Classification RFs aggregate votes from individual trees to estimate the outcome class. In this analysis random forests were built for regression. The results of the leaf nodes at different trees are aggregated for regression (Liaw and Wiener, 2002). The advantages of RF include: its fast training speed, good accuracy and relative efficiency (Mutanga et al., 2012). Additionally, once RF models are established, the values of the predictor that correspond to the first split in the decision tree can be extracted as thresholds corresponding to impact occurrence (Bachmair et al., 2016a).

The R package 'randomForest' was employed to identify the relationship of drought indices to drought impacts in this research (Kursa, 2017;Liaw and Wiener, 2002). There are 5000 decision trees for each RF model. The variance explained was used to determine the goodness of fit of random forest model (Fukuda et al., 2013). The mean squared error (MSE), Eq. (1), was used to evaluate the importance of each index:

$$MSE = \frac{1}{n}\sum_{i=1}^{i=n}(y_i - \hat{y}_i)^2 \tag{3}$$

Where y_i and $\hat{y}_i$ are the each drought impacts and estimated drought impacts of each city, *i*, respectively. *n* is the length of time series.

The percent change of MSE (MSE%) is based on how much the accuracy decreases when the effect of the variable is excluded, as the values are randomly shuffled, the higher the value, the higher the index importance (Carolin et al., 2009). The first splitting values of each decision tree was also extracted. Soil moisture and NDVI were not analyzed using random forest due to missing data and short time series.

4) Standardization of drought impacts and vulnerability factors

To ensure comparability and to facilitate the visualization of the drought impacts and vulnerability factors, they were standardized to a value from 0 to 1 using Eq. (3) and Eq. (4) (Below et al., 2007).

$$SDI_i = \frac{DI_i - \min DI}{\max DI - \min DI} \tag{4}$$

$$SVF_j = \frac{VF_j - \min VF}{\max VF - \min VF} \tag{5}$$

Where SDI_i and DI_i are the Standardized Drought Impacts and drought impacts of year i in Liaoning province, respectively. max DI and min DI is the maximum and minimum values of drought impacts in all year for the given impact type. SVF_j and VF_j is the Standard Vulnerability Factors and vulnerability factors of city j in Liaoning province, and max VF and min VF are the maximum and minimum values of each category of vulnerability factors in all cities.

3. Results

3.1 Drought monitoring and drought impacts

Figure 2 shows high consistency between the drought monitoring indices (in this case the SPI) and the drought impact data.

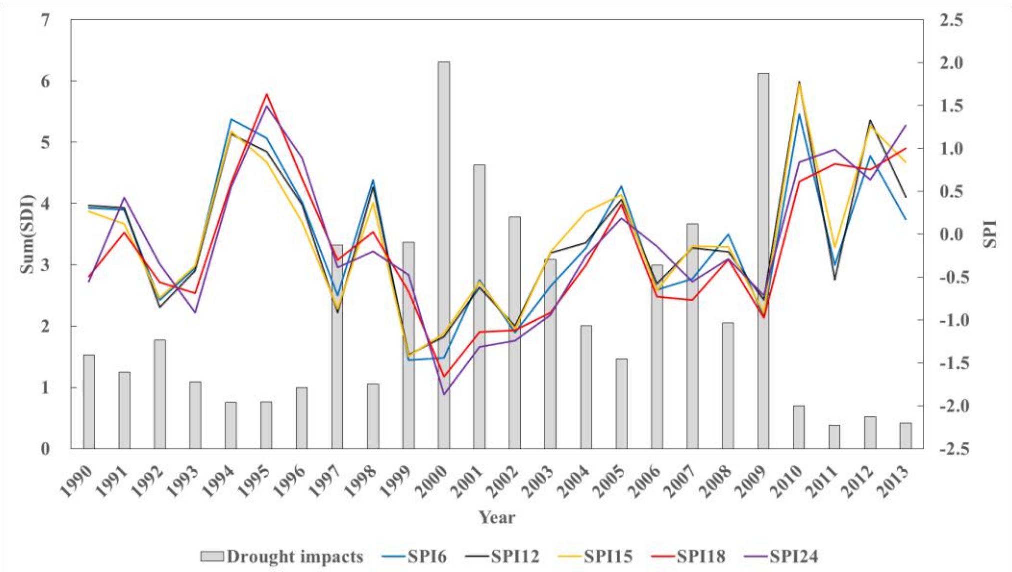

Figure 2: Standardized Precipitation Index (SPI) for 6-, 12-, 15-, 18- and 24-month accumulation periods and the sum of the Standardised Drought Impacts (SDI) for each impact type listed in Table 1 for Liaoning province from 1990 to 2013.

The most severe droughts occurred in 2000, 2001, and 2009, whilst in 1994, 1995, 2012 and 2013 there was above normal precipitation. The largest impacts are generally associated with the lowest index values. This suggests that there is a relationship between the drought indices and drought impacts, and this will be explored quantitatively in the next sections. Figure 3 shows the spatial distribution of the annual average of each drought impact type collected between 1990 and 2016. It shows that for all categories of drought impacts, more drought impacts were recorded in the drier northwestern part of Liaoning province than in eastern parts of the province. The NLH was highest in Dalian, whilst Shenyang had the biggest PHD.

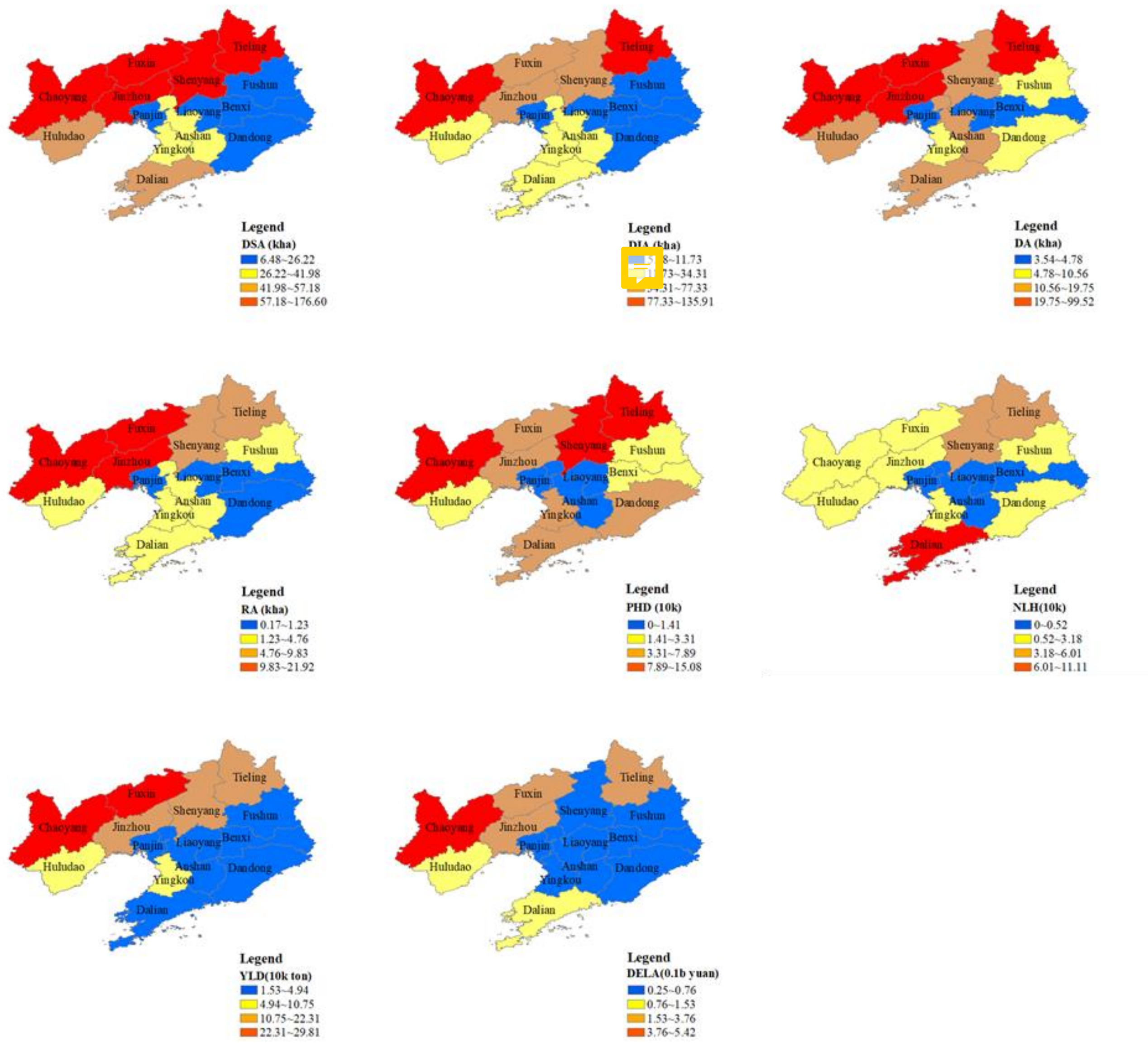

Figure 3: Distribution of average drought impacts (for each impact type, identified by the codes in Table 1) for the period 1990-2013 in Liaoning province.

3.3 Correlation of indices with impacts

The Pearson correlation coefficient (r) for each city and drought impacts is shown in Figure 4. In most cases the drought index is negatively correlated with the drought impacts, suggesting that the lower the drought index, the greater drought impact. However, correlation strength, and direction, varied between the cities and impact types, ranging between -0.890 to 0.621. In most cities of Liaoning province, NDVI and SoilM have a weak correlation with most of types of drought impacts. In Dalian, Chaoyang and Fuxin, all drought indices had a strong correlation with DA, whilst there was a significant correlation for drought impacts area in Jinzhou, Fuxin and Dalian, where most of the correlations were significant ($p < 0.01$). The strongest correlation was found between indices and PHD in Dalian, while it was weakest in Dandong. There is a positive correlation between PHD and NDVI in Fushun, whist NLH has a positive correlation with NDVI in Anshan. Generally, SPEI6 had the strongest correlation with all types of drought impacts, whilst SPI12 had the weakest correlation. SPEI typically exhibited stronger correlations with drought impacts than SPI with the same accumulation period.

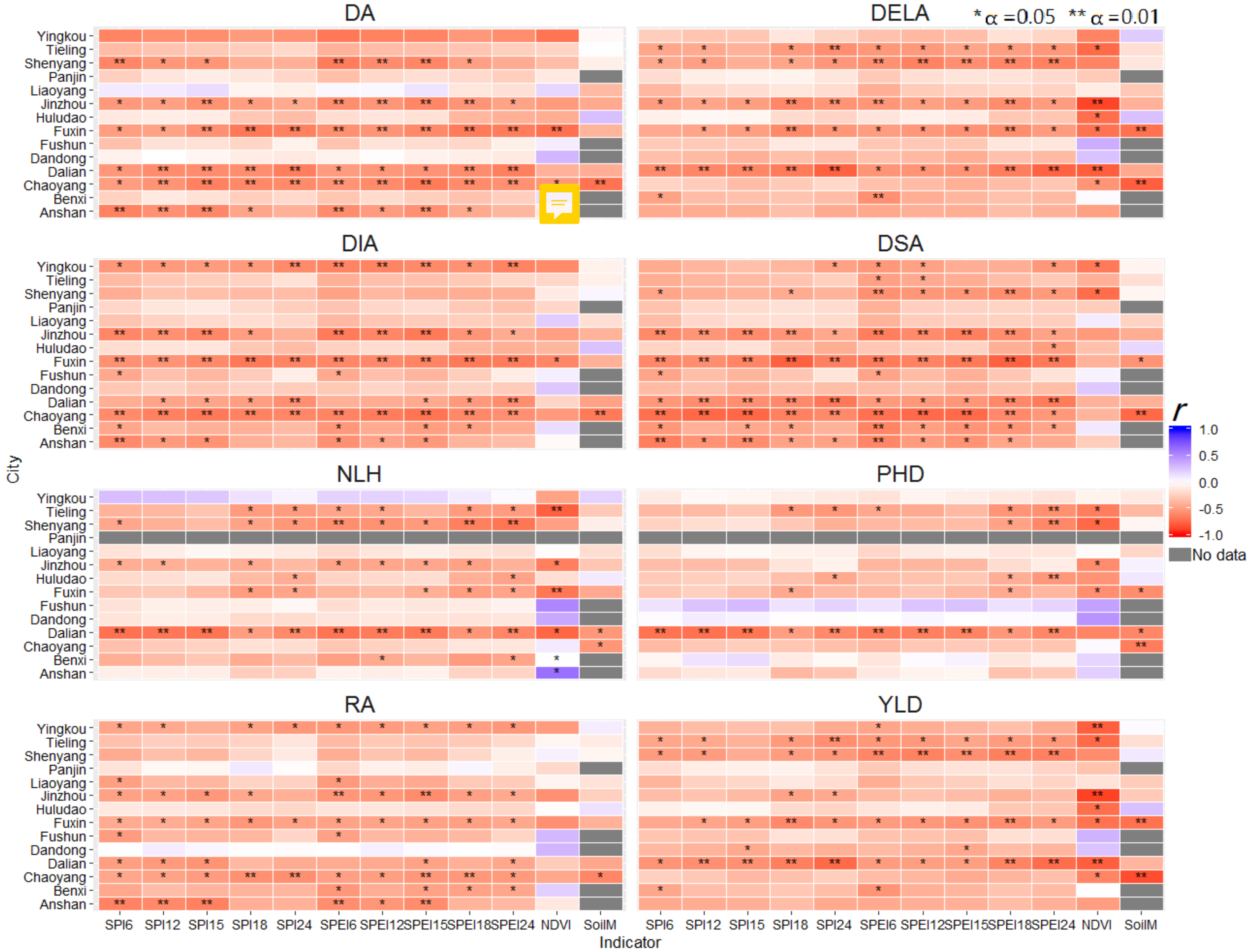

Figure 4: Correlation coefficient (*r*) between drought indices (SPI, SPEI, NDVI and SoilM) and drought impacts for different impact types (identified by the codes in Table 1) in Liaoning province. The significance level of the correlation is indicated using asterisks.

DSA and DIA had a strong correlation with all drought indices in Liaoning province, while PHD and NLH had a weak correlation. The average correlation coefficient across all drought indices and DSA in Liaoning was -0.43, while the average correlation coefficient with PHD and NLH was -0.22 and -0.27, respectively. Drought indices showed a moderate correlation with RA and YLD with average correlation coefficients of -0.32 and -0.37, respectively.

The performance of soil moisture varied significantly between cities and impact types (Figure 4); it had a strong correlation with the impacts in Chaoyang, and a weak correlation in Huludao. In Chaoyang, the correlation between soil moisture and drought impacts was significant (α=0.01), whilst other cities were not significantly correlated.

3.4 Index importance in random forest models

Each drought impact type was selected as the response variable in the random forest. On average the random forests explained 41% of the variance observed within the drought impacts. The MSE% for each city and impact type is shown in Figure 5. The MSE% can be seen to vary between different impact types. DIA and YLD have higher MSE% than other impact types, with average MSE% is 3.02 and 3.01, respectively. The PHD and NLH had lower MSE%, with average of MSE% of 1.58 and 1.39, respectively. DSA and RA had a moderate relationship with drought indices. SPEI performed better than SPI with same durations; SPEI6 had the highest importance with drought impacts. SPI12 was the least important index to drought impacts. Indices had a higher importance with impacts in Anshan and Dalian and lower importance in Yingkou and Dandong.

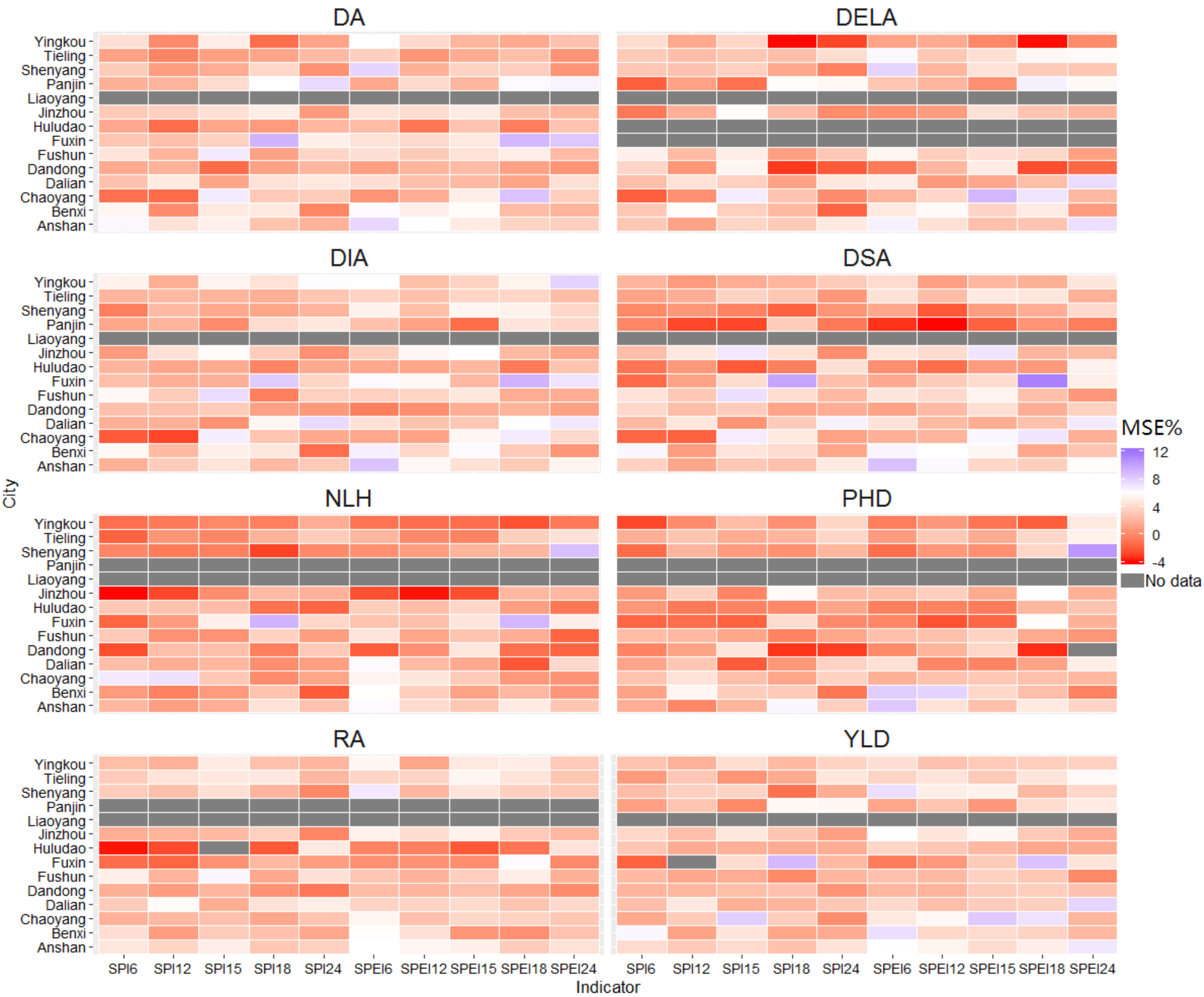

Figure 5: The MSE% of drought indices (SPI and SPEI) with drought impacts (identified by the codes in Table 1) in Liaoning province using random forest.

The variables identified MSE% from the random forest analysis generally match those with strong negative correlations. This supports the statement that indices are negatively related to impacts. The threshold of impact occurrence based on the indices were also identified in the RF analysis using the first splitting value. Figure 6 shows the distribution of first splitting values of each decision tree within the RF. The average first splitting values for SPI18 and SPI24 were higher than those of SPI6, SPI12 and SPI15 (i.e. a more negative index value and more severe meteorological drought state) for all categories of drought impacts. For SPEI, the results were similar (i.e.long-term deficits must be more severe to result in equivalent impacts compared to short-term deficits) but there was more variability between accumulations. When viewed in terms of impact types, DSA had a low threshold, indicating that DSA impacts occur more readily than DA or RA, as may be expected. The impact occurrence of index values increase for DSA, DIS, DA and RA; and YLD and DELA tended to occur for more severe water deficits, with the highest severity threshold being for NLH, indicating that only very severe drought conditions triggered impacts on livestock.

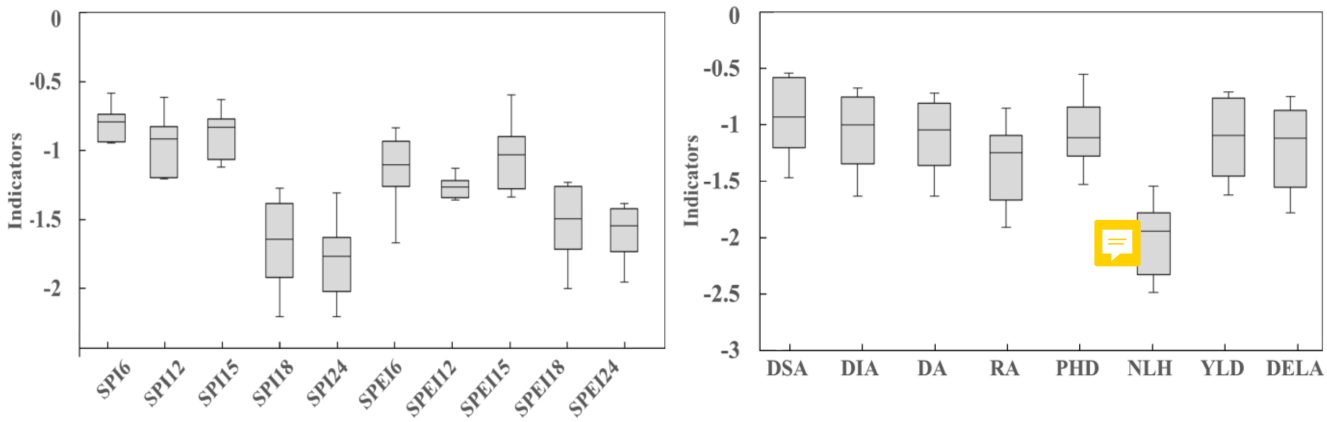

Figure 6: Box plots showing the splitting value (i.e. the thresholds of impacts) in random forest construction across all impact types for each index (left), and across all indices for each impact type (right) in Liaoning province.

3.5 Drought vulnerability evaluation

The results of correlation analysis and random forest show that in most parts of Liaoning province, SPEI at 6-month accumulation period had the strongest correlation with drought impacts. SPEI6 was therefore selected to assess the drought vulnerability of the 14 cities. Regression analysis was performed on the SPEI6 for each category of drought impact, and an example is given in Figure 7 which shows the linear regression of DSA with SPEI6 in the 14 cities. It can be surmised, for practical purposes, that the worse the drought impacts associated with a given drought severity (defined by SPEI6), the higher drought vulnerability of the city to the given impact. Fuxin, Tieling, Chaoyang, Jinzhou and Shenyang have a higher vulnerability to DSA compared to the other cities.

Similar analyses were conducted for all impact types, and Figure 8 summarises which drought impacts each city was the most vulnerable to. It can be seen from Figure 8 that there is little difference between cities in terms of sensitivity to various categories of drought impacts. Considering the various impacts, Chaoyang, Jinzhou, Tieling, Fuxin and Shenyang had the highest drought vulnerability, which are all located in the northwest part of Liaoning province. Dalian was most vulnerable to NLH.

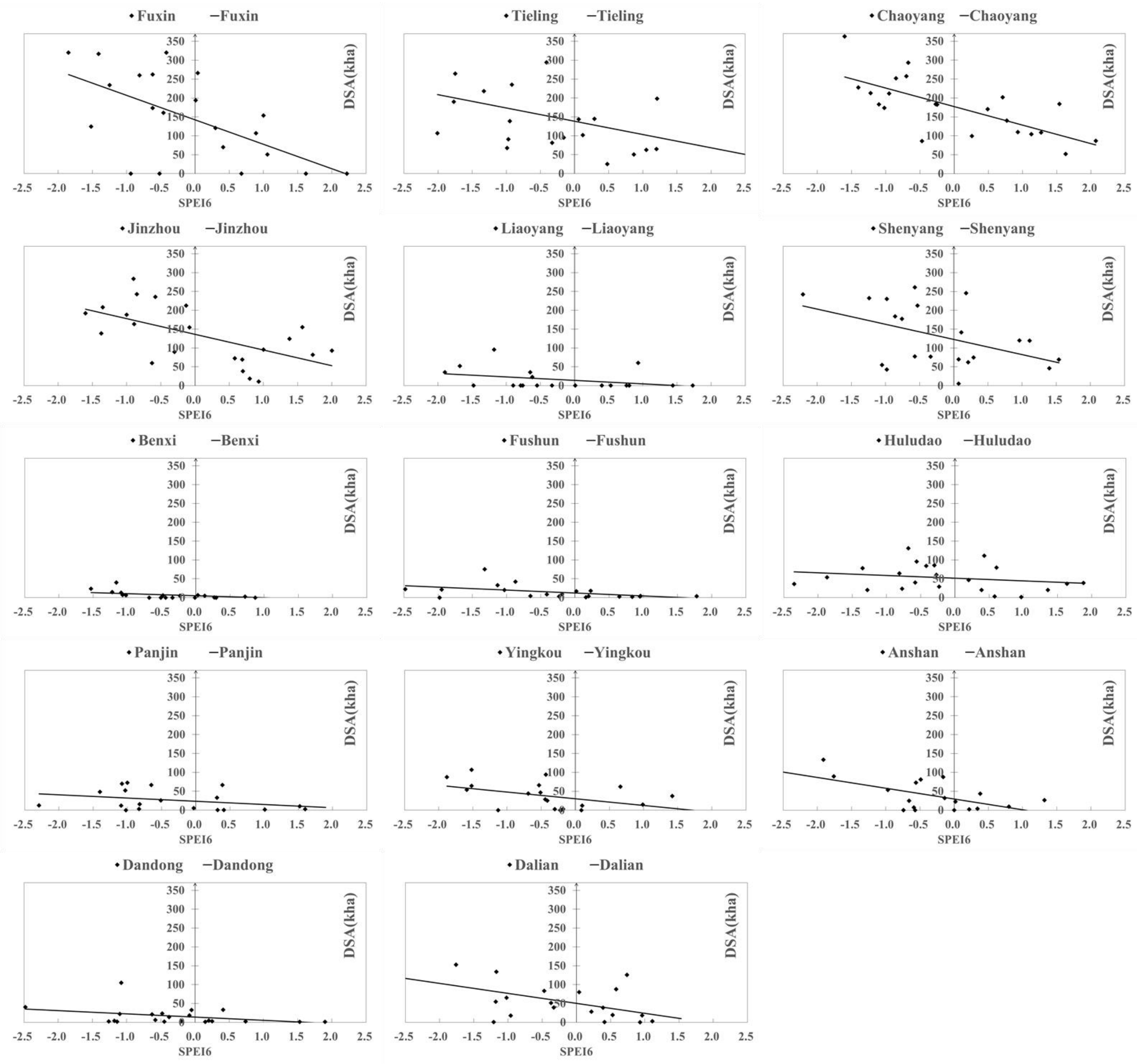

Figure 7: Linear regression results of DSA with SPEI6 in each of the 14 cities in Liaoning Province.

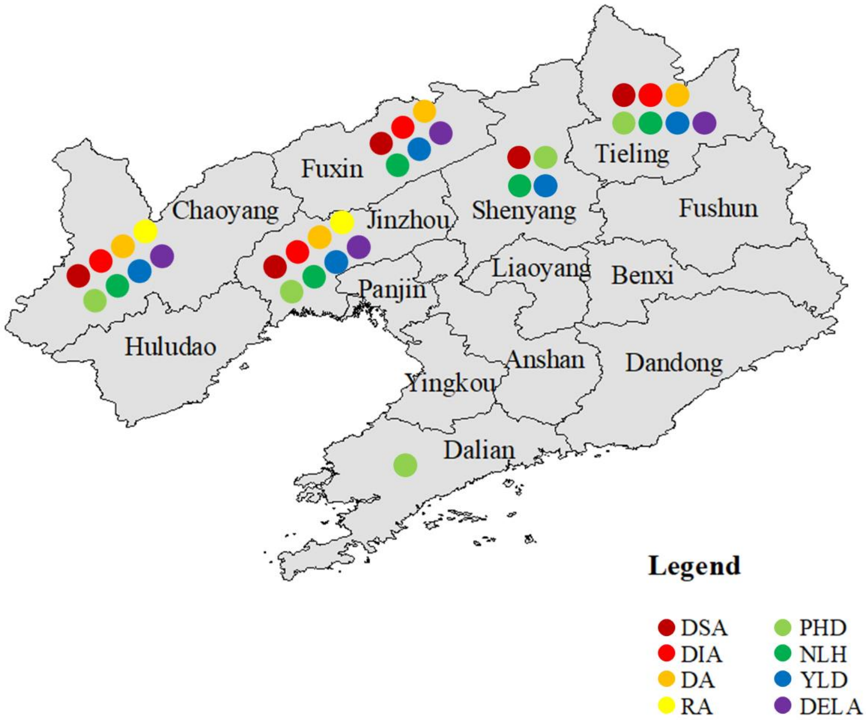

Figure 8: Map showing which drought impacts each city in Liaoning province is most vulnerable to based on the results of the linear regression.

3.6 Vulnerability analysis

A further stepwise regression model was built to explain the variation in each type of Standardised Drought Impact where SPEI6 is equal to -1.5, using vulnerability factors (listed in Table 2) as predictors. Table 3 shows the results of stepwise regression model, demonstrating the contribution of vulnerability factors to each category of drought impact. The results varied for each impact type.

Table 3: The vulnerability factors selected for the stepwise regression model and the R^2 of the resulting model for each impact type (identified by the codes in Table 1).

Drought impact	Predictors (vulnerability factors)	R^2
DSA	Crop cultivated area/Population/Livestock production	0.894
DIA	Crop cultivated area/Population	0.743
DA	Livestock production /Per capita gross domestic product	0.731
RA	Number of electromechanical wells/Per capita gross domestic product	0.541
PHD	Crop cultivated area/Reservoir total storage/Per unit area of Fertilizer application	0.805
NLH	Population	0.474
YLD	Crop cultivated area	0.606
DELA	Crop cultivated area/Population/Livestock production	0.786

Crop cultivated area had a significant relationship with drought vulnerability for DSA, DIA, PHD, YLD and DELA impact types; and population had a significant relationship with DSA, DIA, NLH and DELA. Population was the only significant predictor identified for DELA, with an R^2 of 0.474. Crop cultivated area increases drought vulnerability significantly for 5 out of 8 drought impacts types, while population reduces the drought vulnerability significantly for four drought impact types. With the exception of PHD and NLH, crop cultivated area is directly related to the other drought impact types. Crop cultivated area was the only significant predictor for YLD.

4 Discussion

The methodology in this research has the following characteristics. Firstly, it combines multiple sources of data such as remote sensing data (NDVI data), soil moisture and meteorological data, and takes many drought impacts, across a range of sectors, into consideration. Secondly, the extensive drought impact data was systematically collected from the county level, which is a consistent and reliable data source enabling regional comparisons. The drought impact data used here included impact variables that are rarely available in other studies such as PHD, NLH, YLD and DELA. Thirdly, we not only considered the occurrence of drought events, but also the severity of drought and its spatial extent. Finally, the drought indices-impacts linkage was applied to assess drought vulnerability in Liaoning province.

The biggest challenge of this study was the spatial and temporal matching between the drought impacts and indices. The regularity with which impact data are collected is determined by the drought warning level and as such they are not evenly spaced in time; as a result of this, the data were aggregated to annual totals. It was important to match the accumulation period and timing of the selected drought indices to the timescales critical for the drought impacts; SPEI6 in September covers the critical maize growth period and when the majority of precipitation falls. Soil moisture data are collected at a daily resolution, in order to match up soil moisture and impact data, the March to October average was used in the correlation analysis. However, short term soil moisture deficits can have serious impacts on crops which are sometimes unrecoverable. The average soil moisture may not have captured these short-term deficits, particularly if soil moisture was, in general, sufficient the rest of the year. For this reason, soil moisture data can be used for real-time drought monitoring applications, but may not appropriate to present drought impacts on an annual scale for risk assessment, as applied here. In some cities, the lack of soil moisture data means that the annual average soil moisture does not reflect the occurrence of typical agricultural drought during the year.

NDVI data for the critical growth period of spring maize was used in the analysis with annual drought impacts, but again this does not take all drought events during crop growth period into account. The correlation coefficients characterizing the relationship between NDVI and drought impacts are both positive and negative. This is likely due to the complexity of NDVI drivers (e.g. diversity of land cover

The results from the correlation analysis were consistent with the results from the RF analysis. DSA and DIA had strong correlations with all drought indices in Liaoning province, while PHD and NLH have a weak correlation with indices. This was because DSA and DIA are direct impacts of agricultural drought, whilst PHD and NLH are related to many factors, such as drinking water source location and the amount of water resources available.

The random forest algorithms presented in this paper explained an average of 41% of the variance observed within the drought impact data. This is relatively modest, because of the limitation of the impacts data. Collinearity of the drought indices (e.g. SPI6 is correlated with SPEI6) is a also a potential cause of the low MSE%. The correlation coefficients calculated for drought indices and NLH in Yingkou, and PHD in Fushun were positive. This result is unexpected given the interpretation of these indices as estimations of the drought severity, and the majority of reported correlation coefficients being negative. Therefore, it seems likely this result is not representative of the true relationships between these indices and impacts, and instead an artifact of imperfect data. To explore this the correlation coefficients were estimated with the largest impact years removed. This resulted in a negative correlation coefficient, providing further evidence for the positive correlation coefficients not being representative of the true relationships. The availability of more data would enable a better approximation of the true relationships between indices and impacts.

For all the drought impacts, Dalian and Fuxin had the highest correlation coefficient for all drought impact types and indices. The most vulnerable cities were Fuxin, Tieling, Chaoyang, Jinzhou and Shenyang, which are all located in the northwestern part of Liaoning province indicating there is a high drought vulnerability and drought risk in northwestern Liaoning. This is consistent with existing research by (Yan et al., 2012;Zhang et al., 2012), which established a drought risk assessment index system to assess drought risk in northwestern Liaoning. The number of electromechanical wells is associated with low drought vulnerability – this is the critical water source for irrigation and human drinking.

The first splitting value tended to decrease as the accumulation periods increase, suggesting that higher water deficits are required for the same amount of impact at longer accumulation periods. There is a more severe water deficits of RA occurrence since it caused yield loss by 80% or more, compared to 10% and 30% for DIA and DA, respectively. Livestock drinking water requires lower water quality compare to human and lot of water source available for livestock. For this reason, NLH showed least sensitivity to water deficits.

The relationships analysed in this research support the development of a drought impacts predictor. The drought vulnerability map can be used to support drought risk planning, helping decision makers to inform drought mitigation activities (e.g. sinking more wells to enhance resilience to drought). The impact thresholds identified can also support improved drought warning and planning. The methods used here can be applied in other areas to better understand drought impacts and drought vulnerability. While systematic, statistical archives of drought impact are comparatively rare, globally, there are numerous other potential sources of impact data that could be used (e.g. see Bachmair et al. 2016).

5 Conclusion

This study used correlation analysis and random forest methods to explore the linkage between drought indices and drought impacts. It assessed drought risk in Liaoning province, and proposes a drought vulnerability assessment method which is applied to study the contribution of various socioeconomic factors to drought vulnerability. Here, we return to the original objectives of the study to summarise the key findings.

1: When and where the most severity drought and impact occurred in study area?

Based on the drought monitoring results of SPI, severe drought occurred in 2000, 2001, and 2009. In 2000-2001, drought resulted in many impacts in Liaoning province, particularly in the northwestern part of Liaoning province. The drought monitoring data showed good consistency with the recorded drought impacts.

2：Whether there is an obvious link between drought impact data and drought indices. Which index or set of indices performance best in study area?

The results showed that the indices varied in their capacity to identify the different type of drought and impacts. The strongest correlation was found for SPEI at 6 months, whilst SPI12 had a weak correlation with drought impacts. SPEI was found to better link to drought impacts than SPI of the same accumulation period. NDVI and soil moisture showed some links with impacts in some cities, but the results were generally weaker and less consistent than for either SPI/SPEI – primarily reflecting the limitations in the soil moisture and NDVI datasets

3. Which city or areas have a high drought vulnerability in Liaoning province?

Chaoyang, Jinzhou, Fuxin, Shenyang and Tieling had higher drought vulnerability, all of which are located in the northwestern part of Liaoning province, indicating that drought vulnerability is higher in these regions than in other parts, which is consistent with previous research. However, in contrast with past work, the present research provides a much more comprehensive assessment based on the occurrence of observed impacts data.

4: Which vulnerability factor or set of vulnerability factors contribute most to drought vulnerability?

Population had a strong negative relationship with drought vulnerability, whilst crop cultivated area was positively correlated with drought vulnerability.

The results shown here give a clearer understanding about drought conditions in Liaoning province. The linkage developed can be used to assess drought risk and to map vulnerability. It can also be used to help develop early warning systems and predict drought impacts, which are vital tools for drought management. The results of the vulnerability analysis can guide management measures to mitigate drought impacts – an important step to shift from post-disaster recovery to proactive pre-disaster prevention.

Author Contributions

Yaxu Wang, Juan Lv, Jamie Hannaford, Yicheng Wang and Lucy Barker discussed and developed the aims of the paper. Yaxu Wang was responsible for the data analysis, visualization and prepared the original manuscript, with contributions from

Hongquan Sun, Lucy Barker, Jamie Hannaford, Miaomiao Ma, Zhicheng Su and Michael Eastman.

Competing interests

The authors declare they have no conflict of interest.

Acknowledgements

The authors gratefully acknowledge funding support for these researches provided by the National Key Research and Development Project (No. 2017YFC1502404), and Fund of China Institute of Water Resources and Hydropower Research (JZ0145B592016) and China Scholarship Council. Jamie Hannaford, Lucy Barker and Michael Eastman were supported by the NERC National Capability Official Development Assistance project SUNRISE ("Sustainable Use of Natural Resources to Improve Human Health and Support Economic Development") [NE/R000131/1].

References

Bachmair, S., Kohn, I., and Stahl, K.: Exploring the link between drought indices and impacts, Natural Hazards & Earth System Science, 15, 1381-1397, 2014.

Bachmair, S., Stahl, K., Collins, K., Hannaford, J., Acreman, M., Svoboda, M., Knutson, C., Smith, K. H., Wall, N., and Fuchs, B.: Drought indices revisited: the need for a wider consideration of environment and society, Wiley Interdisciplinary Reviews: Water, 3, 516-536, 2016a.

Bachmair, S., Svensson, C., Hannaford, J., Barker, L. J., and Stahl, K.: A quantitative analysis to objectively appraise drought indices and model drought impacts, Hydrology and Earth System Sciences,20,7(2016-07-04), 12, 9437-9488, 2016b.

Belal, A. A., El-Ramady, H. R., Mohamed, E. S., and Saleh, A. M.: Drought risk assessment using remote sensing and GIS techniques, Arabian Journal of Geosciences, 7, 35-53, 2014.

Below, R., Grover-Kopec, E., and Dilley, M.: Documenting drought-related disasters: A global reassessment, The Journal of Environment & Development, 16, 328-344, 2007.

Biswas, S., Roy, S. S., and Sarkar, S.: Drought Risk Assessment Using GIS and Remote Sensing, 2013.

Blauhut, V., Gudmundsson, L., and Stahl, K.: Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts, Environmental Research Letters, 10, 014008, 2015.

Botterill, L. C., and Hayes, M. J.: Drought triggers and declarations: science and policy considerations for drought risk management, Natural hazards, 64, 139-151, 2012.

Cai, F., Zhang, S. J., Ji, R. P., Mi, N., Wu, J. W., and Zhang, Y. S.: [Spatiotemporal dynamics of maize water suitability and assessment of agricultural drought in Liaoning Province, China from 1981 to 2010], Chinese Journal of Applied Ecology, 26, 233, 2015.

Carolin, S., James, M., and Gerhard, T.: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests, Psychological Methods, 14, 323-348, 2009.

Chen, T., Xia, G., Liu, T., Chen, W., and Chi, D.: Assessment of drought impact on main cereal crops using a standardized precipitation evapotranspiration index in Liaoning Province, China, Sustainability, 8, 1069, 2016.

Edwards, D. C.: Characteristics of 20th century drought in the United States at multiple time scales, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH, 1997.

Fukuda, S., Spreer, W., Yasunaga, E., Yuge, K., Sardsud, V., and Müller, J.: Random Forests modelling for the estimation of mango (Mangifera indica L. cv. Chok Anan) fruit yields under different irrigation regimes, Agricultural Water Management, 116, 142-150, 2013.

Hayes, M., Svoboda, M., Wall, N., and Widhalm, M.: The Lincoln declaration on drought indices: universal meteorological drought index recommended, Bulletin of the American Meteorological Society, 92, 485-488, 2011.

Hayes, M. J., Svoboda, M. D., Wiihite, D. A., and Vanyarkho, O. V.: Monitoring the 1996 drought using the standardized precipitation index, Bulletin of the American meteorological society, 80, 429-438, 1999.

Hong, W., Hayes, M. J., Weiss, A., and Qi, H.: An Evaluation the Standardized Precipitation Index, the China-Z Index and the Statistical Z-Score, International Journal of Climatology, 21, 745-758, 2001.

Hong, W., and Wilhite, D. A.: An Operational Agricultural Drought Risk Assessment Model for Nebraska, USA, Natural Hazards, 33, 1-21, 2004.

Jia, H., Wang, J., Pan, D., and Cao, C.: Maize Drought Disaster Risk Assessment Based on EPIC Model: A Case Study of Maize Region in Northern China, Acta Geographica Sinica, 66, 643-652, 2011.

Jin, J., Song, Z., Cui, Y., Zhou, Y., Jiang, S., and Jun, H. E.: Research progress on the key technologies of drought risk assessment and control, Journal of Hydraulic Engineering, 47, 398-412, 2016.

Karavitis, C. A., Tsesmelis, D. E., Skondras, N. A., Stamatakos, D., Alexandris, S., Fassouli, V., Vasilakou, C. G., Oikonomou, P. D., Gregorič, G., and Grigg, N. S.: Linking drought characteristics to impacts on a spatial and temporal scale, Water Policy, 16, 1172-1197, 2014.

Kursa, M. B.: Efficient All Relevant Feature Selection with Random Ferns, 2017.

Li, Y. P., Wei, Y., Meng, W., and Yan, X. D.: Climate change and drought: a risk assessment of crop-yield impacts, Climate Research, 39, 31-46, 2009.

Liaoning Province Bureau of Statistical: Liaoning Statistical Yearbook 2016, China Statistics Press, 2017.

Liaw, A., and Wiener, M.: Classification and regression by randomForest, R news, 2, 18-22, 2002.

Lin, P., Youhua, M. A., Jiang, Z., Wang, Q., Wang, J., Huang, H., and Jiang, H.: Research Progress of Evaluation Index of Soil Moisture, Agricultural Science & Technology, 2016.

Liu, X., Zhang, J., Ma, D., Bao, Y., Tong, Z., and Liu, X.: Dynamic risk assessment of drought disaster for maize based on integrating multi-sources data in the region of the northwest of Liaoning Province, China, Natural Hazards, 65, 1393-1409, 2013.

Lloyd-Hughes, B.: The impracticality of a universal drought definition, Theoretical and Applied Climatology, 117, 607-611, 2014.

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

Mishra, A. K., and Singh, V. P.: A review of drought concepts, Journal of Hydrology, 391, 202-216, 2010.

Mutanga, Onisimo, ADAM, Elhadi, Cho, A., and Moses: High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm, International Journal of Applied Earth Observations & Geoinformation, 18, 399-406, 2012.

Özger, M., Mishra, A. K., and Singh, V. P.: Low frequency drought variability associated with climate indices, Journal of Hydrology, 364, 152-162, 2009.

Ren, Y. D., and Zhou, J.: Research on the Status of Corn Industry Development in Liaoning Province, Agricultural Economy, 37-38, 2009.

Seiler, R. A., Hayes, M., and Bressan, L.: Using the standardized precipitation index for flood risk monitoring, International Journal of Climatology, 22, 1365-1376, 2002.

Sethi, S. A., Dalton, M., and Hilborn, R.: Quantitative risk measures applied to Alaskan commercial fisheries, Canadian Journal of Fisheries and Aquatic Sciences, 69, 487-498, 2012.

Stagge, J. H., Kohn, I., Tallaksen, L. M., and Stahl, K.: Modeling drought impact occurrence based on climatological drought indices for four European countries, Egu General Assembly Conference, 2014.

Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acacio, V., Dias, S., Bifulco, C., Stagge, J. H., and Tallaksen, L. M.: Impacts of European drought events: insights from an international database of text-based reports, Natural Hazards and Earth System Sciences,16,3(2016-03-21), 3, 5453-5492, 2016.

Svoboda, M. D., and Hayes, M. J.: Enhancing Drought Risk Management: Tools and Services for Decision Support, 2011.

Thornthwaite, C. W.: An approach toward a rational classification of climate, 1, LWW, 1948.

Vicente-Serrano, S. M., Beguería, S., and Lópezmoreno, J. I.: A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index, Journal of Climate, 23, 1696-1718, 2010.

Wang, S. H.: Analysis of Logical Relationship in the Report of National Drought Relief Statistics Management System, Henan Water Resources & South-to-North Water Diversion, 46-47, 2014.

Wilhite, D. A., Hayes, M. J., Knutson, C., and Smith, K. H.: Planning for Drought: Moving From Crisis to Risk Management 1, JAWRA Journal of the American Water Resources Association, 36, 697-710, 2000.

Wilhite, D. A., and Buchanan-Smith, M.: Drought as hazard: understanding the natural and social context, Drought and water crises: science, technology, and management issues, 3-29, 2005.

Yan, L., Zhang, J., Wang, C., Yan, D., Liu, X., and Tong, Z.: Vulnerability evaluation and regionalization of drought disaster risk of maize in Northwestern Liaoning Province, Chinese Journal of Eco-Agriculture, 20, 788-794, 2012.

Yu, C., Li, C., Xin, Q., Han, C., Jie, Z., Feng, Z., Li, X., Clinton, N., Xiao, H., and Yue, Y.: Dynamic assessment of the impact of drought on agricultural yield and scale-dependent return periods over large geographic regions ☆, Environmental Modelling & Software, 62, 454-464, 2014.

Zhang, J.: Risk assessment of drought disaster in the maize-growing region of Songliao Plain, China, Agriculture Ecosystems & Environment, 102, 133-153, 2004.

Zhang, J. Q., Yan, D. H., Wang, C. Y., Liu, X. P., and Tong, Z. J.: A Study on Risk Assessment and Risk Regionalization of Agricultural Drought Disaster in Northwestern Regions of Liaoning Province, Journal of Disaster Prevention & Mitigation Engineering, 2012.

Zhao, H., Gao, G., Yan, X., Zhang, Q., Hou, M., Zhu, Y., and Tian, Z.: Risk assessment of agricultural drought using the CERES-Wheat model: A case study of Henan Plain, China, Climate Research, 50, 247-256, 2011.