Cuadernos de Investigación Geográfica	2016	Nº 42 (1)	pp. 167-183	ISSN 0211-6820 EISSN 1697-9540

DOI: 10.18172/cig.2916

SPATIAL AND TEMPORAL VARIATIONS IN THE ACCURACY OF METEOROLOGICAL DROUGHT MAPS

S. YUAN*, S.M. QUIRING, S. PATIL

Climate Science Lab, Department of Geography, Texas A&M University, USA.

ABSTRACT. *Meteorological drought indices are commonly calculated using data from weather stations and then interpolated to create a map of moisture conditions. These maps are used to communicate drought information to decision makers and the general public. This study analyzes five of the factors (drought index, interpolation method, seasonality, climate region, and station density) that influence the accuracy of these maps. This study compared the Standardized Precipitation Index (SPI) and Standardized Precipitation Evapotranspiration Index (SPEI) using data from the Cooperative Observer Network (COOP) and United States Historical Climatology Network (USHCN). The accuracy of the drought maps varied significantly over time and space. The most significant factor affecting the accuracy of the meteorological drought maps was seasonality. Errors were higher in regions (e.g., southeastern U.S.), and months (e.g., summer), dominated by convective precipitation. The choice of interpolation method also had an influence. We found that Ordinary Kriging (OK) performed better than Inverse Distance Weighting (IDW) in all cases and therefore it was recommended for interpolating drought indices. Not surprisingly, maps that were created using more stations (COOP) were more accurate. The normalized errors of SPI and SPEI were very similar and so the choice of drought index had little impact on the accuracy of the drought maps.*

Variaciones espaciales y temporales en la precisión de los mapas de sequías meteorológicas

RESUMEN. *Los índices de sequía meteorológica se calculan normalmente utilizando datos de estaciones meteorológicas que luego son interpolados para crear un mapa de las condiciones de humedad. Este estudio analiza cinco de los factores (índice de sequía, método de interpolación, estacionalidad, región climática y densidad de estaciones) que influyen en la precisión de estos mapas. Este estudio compara el Índice de Precipitación Estandarizada (SPI) y el Índice de Precipitación Evapotranspiración estandarizada (SPEI) utilizando datos de la Cooperative Observer Network (COOP) y la Red Climatológica Histórica de los Estados Unidos (USHCN). La precisión de los mapas de sequía varió significativamente en el tiempo y en el espacio. El factor que afectó de manera* *más significativa a los mapas de sequía meteorológica fue la estacionalidad. Los errores fueron mayores en las regiones (por ej., sureste de los Estados Unidos) y meses (por ej., el verano) dominados por precipitación convectiva. La elección del método de interpolación también tuvo influencia, de manera que el Kriging Ordinario (OK) ofrecía mejores resultados que el Peso de la Distancia Inversa (IDW) en todos los casos, y por ello se recomienda para la interpolación de índices de sequía. Como cabría esperar, los mapas que fueron creados utilizando más estaciones (COOP) fueron más precisos. Los errores normalizados de SPI y SPEI fueron muy similares, de manera que la elección del índice de sequía tuvo poco impacto en la precisión de los mapas de sequía.*

Key words: mapping drought indices, spatial and temporal variation, accuracy.

Palabras clave: cartografía de índices de sequía, variación especial y temporal, precisión.

Received 2 December 2015
Accepted 5 February 2016

*Corresponding author: Shanshui Yuan. Climate Science Lab, Department of Geography, 814 O&M Building, 3147 Texas A&M University, College Station, Texas, 77843-3147, USA. E-mail: yuanshanshui@tamu.edu

1. Introduction

Drought is a naturally recurring feature of the climate system that is characterized by a prolonged deficiency of precipitation (Dai, 2011). The social and economic costs of drought can be enormous and therefore decision makers seek to develop better mitigation and adaptation strategies (Zarafshani *et al.*, 2012). Quantitative information on the duration, severity and spatial extent of drought events are used to help monitor conditions and make decisions (Rhee *et al.*, 2008). Accurate information is important for making good decisions. However, lots of factors can affect the accuracy of drought information, such as the choice of drought indices (Quiring, 2009), the source of the meteorological data (Zhou *et al.*, 2011; Naumann *et al.*, 2014), the qualitative and quantitative methods that are used to combine drought indices (Yuan and Quiring, 2014; Steinemann *et al.*, 2015), and the methods used to interpolate drought indices to a continuous grid (Akhtari *et al.*, 2009).

There is no uniform method to characterize drought conditions and many different drought indices have been used to monitor meteorological drought (Heim, 2002; Quiring, 2009). Some of the drought indices that are commonly used to monitor drought conditions include: Palmer Drought Severity Index (PDSI) introduced by Palmer (1965), Standardized Precipitation Index (SPI) introduced by McKee *et al.* (1993), Standardized Precipitation Evapotranspiration Index (SPEI) introduced by Vicente-Serrano *et al.* (2009), and Effective Drought Index (EDI) introduced by Byun and Wilhite (1999). Each drought index has advantages and disadvantages and it may not accurately represent drought conditions at every location (Vicente-Serrano *et al.*, 2010). For example, SPI does not account for the influence of potential evapotranspiration and so it neglects the effects of temperature on drought conditions.

Previous studies have evaluated the performance of drought indices in different regions. For example, Keyantash and Dracup (2002) assessed seven meteorological drought indices and they found that the SPI ranked highly in terms of robustness, sophistication and extendibility. Morid *et al.* (2006) compared seven drought indices in Iran and found that there was significant variability in terms of their ability to accurately detect drought onset and to represent the spatial and temporal patterns of drought. Jain *et al.* (2015) evaluated six drought indices in India and found that the time scale and location of interest had a significant influence on the performance of the drought indices and that there was not a single best index for all locations and time scales. McEvoy *et al.* (2012) evaluated two multiscalar drought indices, SPI and SPEI, in Nevada and California at time scales ranging from 1 to 72 months. They found that the SPEI performed slightly better than SPI when compared to summer stream flow. Vicente-Serrano *et al.* (2010) also compared a multiscalar drought index (SPEI) with two different versions of the PDSI and found that the PDSI provides information on medium-term or long-term drought conditions in most regions.

In addition to the choice of drought indices, the method used for drought index interpolation also has an impact on the accuracy of spatial depictions of drought conditions. Previous studies have evaluated different interpolation methods that are commonly used in drought monitoring. Carbone *et al.* (2008) assessed the suitability of Inverse Distance Weighting (IDW), Thin Plate Splines (TPS), Kriging and Thiessen Polygons (TP) using a cross-validation analysis for both PDSI and SPI based on 316 stations in North and South Carolina (~12 stations per 10,000 km^2). They concluded that IDW and Kriging had similar accuracy and both outperformed TPS and TP by a significant margin. Rhee *et al.* (2008) compared two interpolation methods used for mapping drought indices, simple unweighted average and spatial interpolation (IDW) plus aggregation, to examine the effects on droughts across space. They found that spatial interpolation plus aggregation is a superior method. In general, applying interpolation methods before calculating drought indices may increase error relative to performing the interpolation after calculating the drought index. Akhtari *et al.* (2009) compared the accuracy of drought index interpolations using IDW, Ordinary Kriging (OK) and TPS based on 43 stations in Iran (~22 stations per 10,000 km^2). They observed that IDW and OK outperformed TPS. Ali *et al.* (2011) carried out a similar study using 27 climatic stations in the Boushehr province of Iran (~12 stations per 10,000 km^2) and their results were generally in agreement with Akhtari *et al.* (2009).

Other interpolation methods such as regression, Bayesian approaches (Li and Heap, 2011) and the reduced optimal interpolation (ROI) method (Kaplan *et al.*, 2000) have also been analyzed. For example, Yuan and Quiring (under review) applied ROI to interpolate soil moisture in Oklahoma. This method uses a secondary dataset to capture the spatial patterns and improve the interpolation accuracy. However, these methods have not been frequently applied for interpolating drought indices. In addition, most of the previous studies that have evaluated interpolation of drought indices have focused on a regional scale (e.g., a relatively small number of stations in a single state or province). There have been relatively few studies comparing interpolation methods for drought indices over a large spatial extent (e.g., national to continental) that spans diverse topography and climate regions. This paper will address this gap by evaluating the spatial and temporal patterns of interpolation accuracy over the United States.

In addition to the influence of selecting a drought index and interpolation method, the accuracy of spatial depictions of drought are also influenced by the climatic datasets that are used (Sheffield *et al.*, 2012). Tucker (1989) compared two satellite-derived data sets for drought monitoring in sub-Saharan Africa and northeastern Brazil, one based on the Scanning Multi-channel Microwave Radiometer (SMMR) and one based on Advanced Very High Resolution Radiometer (AVHRR). Significant differences between the two were found during certain time periods even though the two satellite-derived datasets are highly correlated. Sheffield *et al.* (2012) calculated the PDSI using observation (HADCRU), reanalysis (e.g., ECMWF ERA-40, ERA-Interim and NCEP/NCAR) and combined datasets (Sheff2006) to evaluate global drought trends. They found that the choice of datasets had an impact on the results and the largest drought trends were found when using the NCEP/NCAR reanalysis dataset. Mo *et al.* (2010) compared SPI calculated based on the Climate Forecast System Reanalysis (CFSR), North American Land Data Assimilation System (NLDAS) and the North American Regional Reanalysis (NARR). They found significant regional and seasonal differences in the SPI between the three datasets. Naumann *et al.* (2014) compared SPI and SPEI in Africa derived from five datasets: ECMWF ERA-Interim reanalysis, Tropical Rainfall Measuring Mission (TRMM) monthly rainfall (3B43), Global Precipitation Climatology Centre (GPCC) precipitation, Global Precipitation Climatology Project (GPCP) precipitation, and the Climate Prediction Center (CPC) merged analysis of precipitation. They found that use of different datasets had an influence on the magnitude of the wet seasons and extremes and the areal extent of drought events. The differences were especially large in regions with few precipitation gauges.

The goal of this study is to quantify how five factors influence the accuracy of spatial depictions of meteorological drought conditions. Specifically, this paper will evaluate how the choice of drought indices, interpolation methods and weather data influences the accuracy of drought maps. In addition, it will examine how the accuracy of these maps varies over time (seasonality) and space (climate region). This study has significant practical implications since maps of drought conditions are commonly used to develop products like the United States Drought Monitor (Lawrimore *et al.*, 2002) and to make operational decisions such as drought and disaster declarations (Quiring, 2009). A detailed description of the data and methods used in this study are presented in section 2. This is followed by presentation of the results and discussion in section 3 and the conclusions are summarized in section 4.

2. Data and Methods

2.1. Data

Precipitation and temperature data from 2001 to 2010 were obtained from the National Weather Service Cooperative Observation Network (COOP) and the United States

Historical Climatology Network (USHCN). The COOP network consists of volunteer observers that span the continental United States (CONUS) with over 11,000 observers taking measurements for daily variables. Monthly data from COOP are available from the National Centers for Environmental Information. The majority of COOP stations were not included in this study because of missing data. There are 3680 stations that are used to calculate drought indices (Figure 1a). The average station density of the COOP stations, based on the entire CONUS study region, is ~5.80 stations per 10,000 km^2.

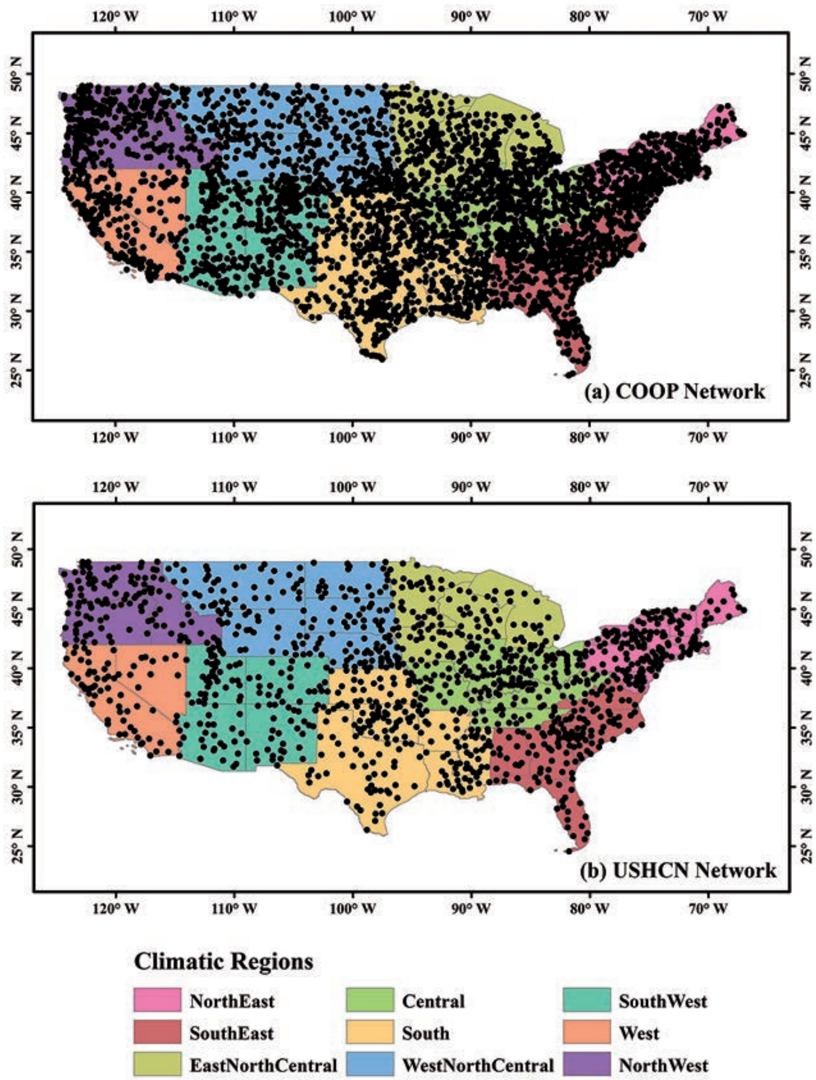

Figure 1. Spatial distribution of: (a) COOP stations and, (b) USHCN stations across the 9 climate regions.

USHCN version 2 is a dataset of 1147 stations (Figure 1b) across the 48 contiguous states that provide longer daily and monthly records of basic meteorological variables. The USHCN sites are a subset of the COOP network and they were selected due to their spatial coverage, record length, data completeness, and historical stability (Menne *et al*., 2009). Most of these stations are in rural locations, while some are National Weather Service First-Order stations that are located at airports. The average station density of the USHCN stations, based on the entire CONUS study region, is 1.5 stations per 10,000 km^2. Relative to COOP, USHCN has a longer period of record, but lower station density.

2.2. Methods

2.2.1. Drought Indices

Two commonly used drought indices are analyzed in this study, the Standardized Precipitation Index (SPI) and the Standardized Precipitation and Evapotranspiration Index (SPEI). The SPEI is an improved version of the SPI that includes a temperature-based estimate of potential evapotranspiration (Vicente-Serrano *et al*., 2009).

The SPI was developed by McKee *et al*. (1993) to provide a moisture supply index that performed better than the PDSI. The SPI is produced by standardizing the probability of observed precipitation for any duration. For example, durations of weeks or months can be used to apply this index for agricultural or meteorological purposes, and longer durations of years can be used to apply this index to water supply and water management purposes (Guttman, 1999). The SPI can be calculated for any location that has a long-term precipitation record. The precipitation record is fit with a probability density function and subsequently transformed using an inverse normal (Gaussian) function (Guttman, 1999). This insures that the mean SPI value for any given location (and duration) is zero and the variance is one.

There are different probability distributions (e.g., Pearson Type III or Gamma) that are commonly used to calculate the SPI. This is important to note because using a different probability distribution will produce different SPI values, even with the same input data. Guttman (1999) experimented with different probability distributions and concluded that the Pearson Type III distribution provides the best model for calculating the SPI. However, this remains a matter for debate since other studies have identified different probability distributions as being the most appropriate for evaluating monthly precipitation probabilities (Legates, 1991, Husak *et al*., 2007). For example, the National Drought Mitigation Center (NDMC, drought.unl.edu) uses the 2-parameter gamma PDF to fit the frequency distribution of precipitation and calculate the SPI (Wu *et al*., 2007). In this study, we will use the Gamma distribution to calculate the SPI.

Vicente-Serrano *et al*. (2009) introduced the Standardized Precipitation Evapotranspiration Index (SPEI) which is calculated using both precipitation and potential evapotranspiration (PET). PET is estimated using Thornthwaite equation (Thornthwaite, 1948) which is based on monthly air temperature, latitude, and month. Thornthwaite equation has some limitations because it is an empirical function solely depends on the temperature. Relative to more physical based equation, such as Penman-Monteith equation (Allen *et al*., 1998) and two-source PET model (Shuttleworth and

Wallace, 1985), Thornthwaite equation overestimates PET under global warming. This leads to further overestimation on drought conditions. Detailed evaluation on the impacts of different PET calculations on droughts was described by Yuan and Quiring (2014). The inclusion of PET means that this index represents the difference between water supply from precipitation (P) and atmospheric demand for water through PET. The SPEI is based on this difference (P_i – PET_i) for any time period of interest. It is necessary to use a 3-parameter model to fit the frequency distribution of the P-PET differences since, unlike when only precipitation is considered, negative values are possible. This study uses a three-parameter Gamma distribution to calculate the SPEI.

2.2.2. Interpolation Methods

Two interpolation methods are evaluated in this study: Inverse Distance Weighting (IDW) and Ordinary Kriging (OK). These two were selected because they are the most commonly used interpolation methods in the environmental sciences (Li and Heap, 2011).

IDW is a weighting algorithm, so the value at the target location is most strongly influenced by the nearest stations. IDW is calculated as follows:

$$Z(S_0) = \sum_{i=1}^{n} \lambda_i Z(S_i) \quad \text{(eq. 1)}$$

$$\lambda_i = \frac{d_i^{-p}}{\sum_{i=1}^{n} d_i^{-p}} \quad \text{(eq. 2)}$$

where, $Z(S_i)$ is the measured value at the i_{th} station; λ_i is the weight for the i_{th} station; $Z(S_0)$ is the interpolated value at the target location, n is the number of stations, d_i is the distance from the i_{th} station to the target location; p is power parameter.

We evaluated the performance of IDW using a variety of power parameters and found that parameters >2.5 or <2.0 led to a degradation in interpolation accuracy (results not shown). Therefore, in this study IDW is employed using power parameters of 2 and 2.5 (hereafter referred to as IDW 2 and IDW 2.5).

OK is based on a semi-variogram, which is used for estimating the dissimilarity between observations. The semi-variogram is calculated using:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n} \left[Z(x_i) - Z(x_i + h)\right]^2 \quad \text{(eq. 3)}$$

where n is the number of pairs of stations which are separated by distance h. $Z(x_i)$ and $Z(x_i + h)$ are the measured value at x_i and $x_i + h$ locations.

In this study, the nearest 500 COOP (200 USHCN) stations are used for generating fitting functions. The theoretical semi-variogram is fit using a least-squares approach based on a Gaussian distribution.

2.2.3. Performance Evaluation

This study uses leave-out-one cross-validation (Isaaks and Srivastava, 1989) to examine the spatial accuracy (i.e., performance) of different drought indices, interpolation methods and datasets. In this approach, one station is removed and the value at that location is interpolated using the remaining stations. This process is then repeated until all stations have been held out. The difference between the measured value at each station and the interpolated value is the residual error and its absolute value is called the absolute error. The absolute error and normalized absolute error (calculated over a group of absolute errors, e.g., all stations within a climatic region) are used to quantify the accuracy for each interpolation method, drought index and dataset. The results are summarized on a regional basis to illustrate the spatial variations in performance.

2.3. Climatic Regions

Figure 1 shows the 9 climatic regions within CONUS that were identified by Karl (1983). In this paper, these regions are referred to as: NorthEast, SouthEast, Central, EastNorthCentral, South, SouthWest, WestNorthCentral, West and NorthWest. These climatic regions are used to aggregate the results of the performance evaluation and to compare how they vary across the United States.

3. Results and Discussion

3.1. Comparison of Interpolation Methods

Figure 2 compares the mean absolute error (MAE) for OK, IDW 2 and IDW 2.5 for 1-month SPI and 1-month SPEI based on COOP stations and USHCN stations for CONUS. Overall, the OK method had the lowest MAE over the entire CONUS, followed by IDW 2.5 and IDW 2. The differences between interpolation methods are small, but statistically significant between OK and IDW based on a paired t-test (90% confidence level).

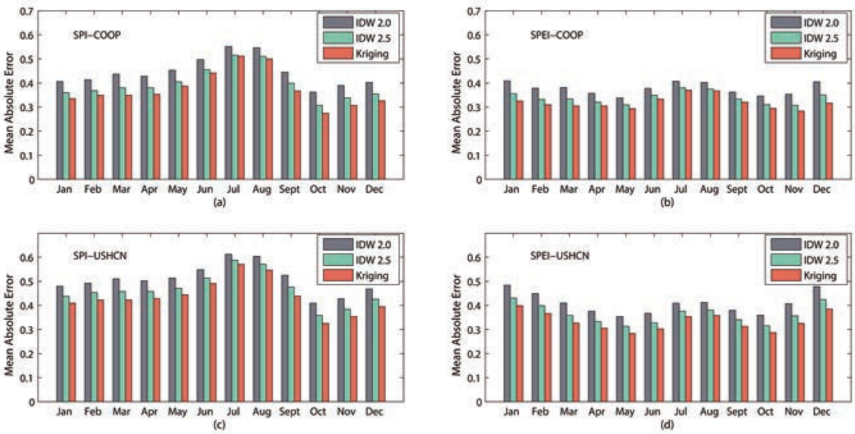

Figure 2. Monthly variations in mean absolute error for the three interpolation methods (IDW 2.0, IDW 2.5 and OK) and the two drought indices (1-month SPI and 1-month SPEI) based on the COOP data (upper) and USHCN data (lower) over CONUS.

Given that OK consistently had the lowest error and IDW 2 had highest error regardless of the drought index and dataset that was used, the results are only shown for one representative drought index/dataset combination (Table 1). Table 1 shows the mean MAE for the three interpolation methods in the 9 climate regions based on the 1-month SPI. IDW 2 has the highest error, while OK has the lowest error in all the regions. All the three interpolation methods have the smallest MAE in the West region and the highest MAE in the WestNorthCentral region. MAE is also consistently high in the SouthEast and SouthWest. Our results show that although the choice of interpolation methods has a relatively small impact on the depiction of drought conditions, OK is slightly more accurate.

Table 1. Mean MAE of 1-month SPI for each interpolation method (IDW 2, IDW 2.5 and OK) in each of the 9 climate regions.

	Mean MAE		
	OK	IDW 2.5	IDW 2
NorthEast	0.35	0.37	0.39
SouthEast	*0.39	0.40	0.44
Central	0.34	0.36	0.41
EastNorthCentral	0.35	0.38	0.42
South	0.36	0.38	0.43
SouthWest	*0.39	*0.41	0.44
WestNorthCentral	*0.39	*0.41	*0.45
West	0.30	0.31	0.33
Northwest	0.34	0.35	0.37

Note: *corresponds to largest mean monthly MAE across the 9 climate regions and corresponds to smallest mean monthly MAE across the 9 climate regions.

3.2. Comparison of Drought Indices

This section compares two commonly used drought indices, SPI and SPEI, to determine whether there are significant differences in their ability to accurately represent spatial variations in drought conditions. SPI depends solely on precipitation while SPEI uses the same algorithm but also accounts for PET. Normalized errors were computed (using the mean and standard deviation of each index) and they are used to compare the pair of drought indices (SPI vs. SPEI) over the 9 climatic regions using USHCN datasets, as shown in Figure 3. A paired t-test is used to compare the drought index pairs in 27 combinations (9 climate regions and 3 selected months: January, July and October). January, July and October were selected because they represent climatic conditions during different seasons. Specifically, precipitation patterns and precipitation mechanisms differ from season to season in CONUS. Obviously our analyses could be repeated for other months, but we believe that these three months are relatively representative. Drought maps are usually created on a weekly or monthly timescale and therefore it is not appropriate to perform this analysis using seasonal precipitation (e.g., total December-January-February precipitation).

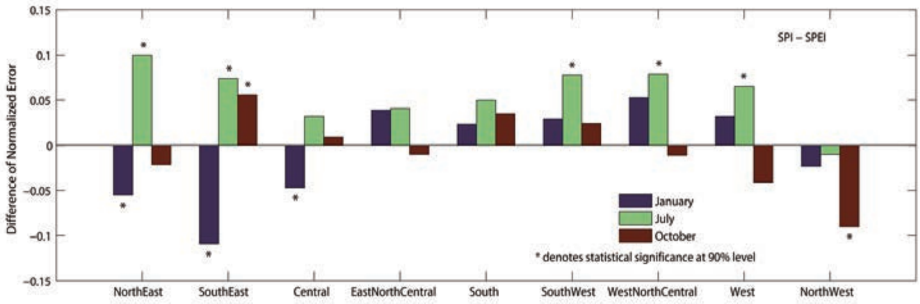

Figure 3. Differences in normalized error between SPI and SPEI over the 9 climatic regions in the United States in January, July and October using USHCN data.

SPI shows slightly lower accuracy than SPEI (Figure 3). 4 out of 27 combinations show that the normalized errors for SPI are significantly lower (90% confidence level) than SPEI and 6 out of 27 combinations show the opposite results. Therefore, maps of drought conditions based on the SPI tend to have higher errors than SPEI, but the results are not statistically significant in all regions and months. A more detailed discussion of the reasons for this is provided in the following section.

3.3. Seasonal Variations

Figure 4a shows that there is substantial seasonal variability in the accuracy of 1-month SPI maps. In most regions, except the West and NorthWest, the largest MAE is in July or August and the smallest MAE is in October or November. This intra-annual variation coincides with the precipitation climatology since SPI is a function of precipitation (not shown). Therefore months with more spatially variable precipitation patterns (e.g., due to convection) have higher interpolation errors than months with more spatially homogeneous precipitation patterns (e.g., due to mid-latitude cyclones). The higher errors in August and July are associated with mesoscale convective systems (Velasco and Fritsch, 1987; Ashley *et al.*, 2003; Murray and Colle, 2010), tropical cyclone systems (Larson *et al.*, 2005) and the North American Monsoon (Adams and Comrie, 1997).

In the West, the largest errors occur in July and the lowest errors occur in December, so clearly the amount of precipitation is not directly driving the seasonal variation in errors since July is one of the driest months in this region and December is one of the wettest months (not shown). In the NorthWest, the errors are relatively high from January through July and then they decrease, with minimum error in October. This region has the lowest seasonal variation in drought index interpolation error. Precipitation patterns in the West and NorthWest are quite different from other regions because both of these two regions have a winter precipitation maximum.

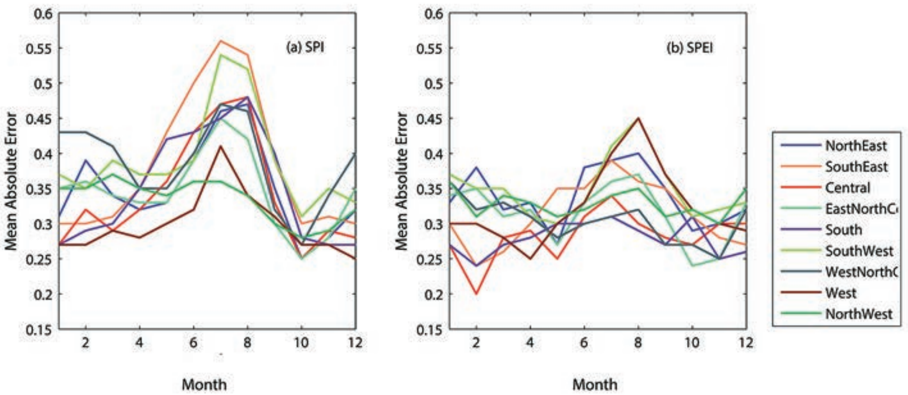

Figure 4. Seasonal variations in mean absolute error for (a) 1-month SPI and (b) 1 month SPEI.

SPEI is dependent on both precipitation and temperature and therefore the seasonal variations in error are driven by temperature and precipitation patterns in each region. It is interesting to note that there is much less seasonal variation in error for the SPEI as compared to the SPI in most of the climatic regions (Figure 4b). This suggests that the higher summer temperatures (and therefore higher PET) reduce the impact of the spatial variability in precipitation. The differences in the seasonal variations in error for the SPI and SPEI are particularly notable in the SouthEast, South, Central, and WestNorthCentral regions. The maximum error tends to occur in August (about 1 month later than the maximum error for SPI). The minimum error tends to occur in spring (April or May) and autumn (October and November).

Paired t-tests (90% confidence level) are used to compare the relative accuracy of the interpolated drought indices in January versus July and July versus October (Table 2). The interpolation error across the 9 climatic regions is lower in January than July (21 out of 27 combinations have lower errors in January based on COOP stations and 19 out of 27 combinations show lower error based on USHCN stations). Similarly, the interpolation error in October is also lower than in July in 22 out of 27 cases for the COOP network and 19 out of 27 cases for the USHCN network. This demonstrates that there are significant seasonal variations in the accuracy of drought index maps. In most locations within the United States, the depictions of drought conditions are more accurate during the winter (January) and fall (October), than they are in summer (July). This presents a challenge given that drought impacts are often most pronounced during the summer.

Table 2. Paired t-test (90% confidence level) for January versus July and July versus October (n = 27; 3 drought indices and 9 climatic regions). The results show, for example, the number of times that the interpolation in January is more accurate than July.

Comparison	Results	Number of Occurrences	
		COOP	USHCN
January vs. July	Interpolation in January performs better than in July	**21**	**19**
	Interpolation in January performs worse than in July	0	0
	No statistically significant difference	6	8
July vs. October	Interpolation in July performs better than in October	2	3
	Interpolation in July performs worse than in October	**22**	**19**
	No statistically significant difference	3	5

3.4. Spatial Variations

Figure 5 shows the spatial distribution of error over CONUS based 1-month SPI and 1-month SPEI in January and July (using USHCN data). Figure 5a shows that in January the majority of CONUS has relatively low errors that range from 0.2 to 0.4 for the 1-month SPI. The highest errors tend to occur in the WestNorthCentral region. This region has complex terrain (from Rocky Mountains to Great Plains) and it has a relatively low station density (1.38 stations per 10,000 km^2). Together, these two factors account for the higher errors in this region.

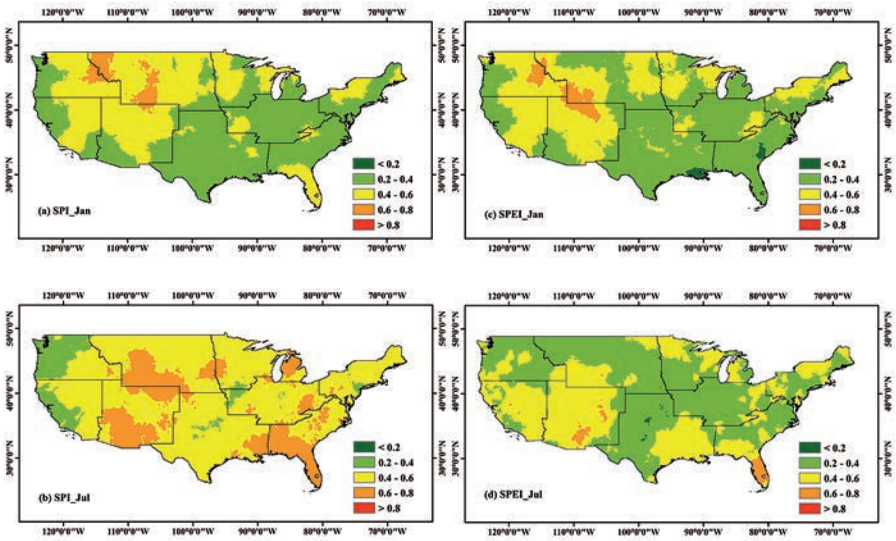

Figure 5. Spatial patterns of mean absolute error for 1-month SPI and 1-month SPEI in January (top row) and July (bottom row) based on USHCN data.

Figure 5b shows that the errors for the 1-month SPI across most of the U.S. are higher in July (0.4 to 0.6) than in January. The highest errors are found in the SouthEast, SouthWest and WestNorthCentral regions in July. These regions are influenced by tropical storms, North American monsoon and summer convective storms in July and therefore have more spatially heterogeneous precipitation patterns that other regions of the U.S. (or other months of the year). In contrast, the errors in the 1-month SPI are relatively consistent from January to July in the west coast of U.S. because these regions receive relatively little precipitation during the summer.

The spatial distribution of error for 1-month SPEI in January and July is shown in Figure 5c and 5d, respectively. In January, the 1-month SPEI (Figure 5c) is very similar to the 1-month SPI (Figure 5a), which confirms the results described in Section 3.2. However, in July it is apparent that the 1-month SPEI (Figure 5d) has lower MAE than the 1-month SPI (Figure 5b). This is because SPEI considers the difference between precipitation and potential evapotranspiration and, as noted above, the higher summer temperatures (and therefore higher PET) reduce the impact of the spatial variability in precipitation. The spatial patterns of error in the SPEI are relatively consistent between January and July, although in July there are higher errors along the Gulf Coast and in Florida due to the heterogeneity of precipitation in these regions.

3.5. Comparison of Station Density

Two networks with different station densities (COOP has a higher station density than USHCN) were compared to evaluate how station density influences interpolation accuracy. We found that using the COOP network results in a small improvement in interpolation accuracy over CONUS (Figure 2). Table 3 shows the monthly averaged MAE from COOP and USHCN for the two drought indices and 9 climatic regions. The 1-month SPI has lower error (10% to 23%) when it is based on the COOP in all 9 climatic regions. The MAE shows that the largest difference is in the West region (23%). This makes sense because the mean station density for the COOP network is ~5.80 stations per 10,000 km^2 versus ~1.50 stations per 10,000 km^2 for USHCN. Therefore, the COOP network is able to more accurately resolve precipitation heterogeneity.

The COOP network is also more accurate than USHCN for the 1-month SPEI. However, the differences in MAE are much smaller, ranging from a 0% to 15% reduction in MAE. The largest difference in MAE for SPEI occurs in the EastNorthCentral region. The results suggests that station density is not as important for SPEI as SPI because the SPEI includes the influence of PET which is a more spatially homogenous field than precipitation and therefore it is not as sensitive to station density.

Table 3. MAE for 1-month SPI and 1-month SPEI based on COOP and USHCN networks.

	Number of Stations		MAE of 1-month SPI			MAE of 1-month SPEI		
	COOP	USHCN	COOP	USHCN	Δ	COOP	USHCN	Δ
NorthEast	261	135	0.37	0.41	-0.04	0.36	0.37	-0.01
SouthEast	382	121	0.41	0.46	-0.05	0.34	0.34	0.00
Central	512	164	0.37	0.42	-0.05	0.32	0.33	-0.01
EastNorthCentral	385	96	0.38	0.46	-0.08	0.35	0.41	**-0.06**
South	658	184	0.39	0.45	-0.06	0.31	0.31	0.00
SouthWest	415	118	0.41	0.45	-0.04	0.37	0.39	-0.02
WestNorthCentral	543	166	0.42	0.48	-0.06	0.33	0.36	-0.03
West	225	60	0.31	0.40	**-0.09**	0.33	0.36	-0.03
Northwest	299	103	0.35	0.40	-0.05	0.34	0.36	-0.02

4. Conclusions

The objective of this study was to quantify how five factors (drought index, interpolation method, seasonality, climate region, and station density) influence the spatial and temporal variability in the accuracy of interpolated maps of meteorological drought indices. Seasonality was found to have the greatest impact on accuracy followed by climate region. Both of these factors are important because they are associated with temporal and spatial variations in meteorological conditions. Time periods and locations with more spatially heterogeneous precipitation are associated with larger errors. The highest errors were consistently observed for 1-month SPI and 1-month SPEI in months with high precipitation (generally summer) in climatic regions dominated by convective precipitation, tropical storms or monsoon regimes. Even with the use of the best interpolation method and highest station density, relatively large interpolation errors were found during these months.

The accuracy of maps of meteorological drought indices varies significantly from region to region and from season to season. Errors are consistently higher in July than January. Regions such as the WestNorthCentral have high errors because of the significant topographical variation, while other regions such as the SouthEast have higher errors in the summer and fall due to sea-breeze and tropical storm activity. The lowest errors are generally found in the Great Plains, Midwest and Northwest U.S. The spatial and temporal variations in the accuracy of drought maps poses a significant challenge for communicating drought information to decision makers and the general public because they are unlikely to be familiar with the precipitation climatology and most operational drought products do not provide estimates of uncertainty/error. Therefore, there is a need to develop better methods for representing and visualizing the inherent uncertainty in maps of drought conditions. This is a fertile area for future research.

The choice of drought indices also has an influence on the accuracy of drought maps. This study compared the 1-month SPI and 1-month SPEI and found that the SPEI has slightly lower interpolation errors than the SPI. We conclude that the accuracy of maps based on drought indices that only use precipitation (e.g., SPI) are more variable than those based on drought indices that use both temperature and precipitation (e.g., SPEI). Therefore, the selection of a drought index can have an important influence on the accuracy of drought maps. Based on the results of this study, the SPEI is more suitable for mapping meteorological drought conditions than the SPI.

Station density has less influence than season, climate region or drought index, but it can be important in regions where station density is very low. For example, our results showed that using the denser COOP network in the West and EastNorthCentral regions lead to a significant reduction in interpolation error as compared to the less dense USHCN network. However, in most locations, using a different station network had a relatively small impact on the interpolation error for the SPEI. Also, mapping drought conditions in regions with complex topography requires a higher density of stations. We found that the relatively low density of stations and complex topography in the WestNorthCentral region lead to high interpolation errors.

In this study, the selection of interpolation method had the least influence on the accuracy of the drought index maps. Although the differences were relatively small, OK consistently performed better than IDW across all season, regions, drought indices and station densities. Therefore, OK is the recommended method for interpolating drought indices. IDW is a reasonable choice in most situations and using a weighting parameter of 2.5 was consistently better than a weighting parameter of 2.0. The interpolation methods applied in this study are all geospatial interpolation methods. Developing and evaluating more advanced interpolation methods may improve the accuracy of drought mapping.

Future research will focus on evaluating how accuracy of drought maps produced using other indices such as the Palmer Drought Severity Index (PDSI). The PDSI is a widely used drought index that integrates weather conditions over many months. The characteristic timescale of the PDSI is on the order of 9 to 12 months (Dai and National Center for Atmospheric Research Staff, 2015). Therefore, evaluating indices such as the PDSI will be useful for examining how the accuracy of maps of longer-term drought conditions (i.e., 6-month, 9-month and 12-month). In addition, this study only used MAE to evaluate the accuracy of drought maps. Future research should utilize a more comprehensive set of performance evaluation statistics (e.g., correlation, coefficient of efficiency and skill scores).

References

Adams, D.K., Comrie, A.C. 1997. The North American Monsoon. *Bulletin of the American Meteorological Society* 78 (10), 2197-2213.

Akhtari, R., Morid, S., Mahdian, M.H., Smakhtin, V. 2009. Assessment of areal interpolation methods for spatial analysis of SPI and EDI drought indices. *International Journal of Climatology* 29 (1), 135-145.

Ali, M.G., Younes, K., Esmaeil, A., Fatemeh, T. 2011. Assessment of geostatistical methods for spatial analysis of SPI and EDI drought indices. *World Applied Sciences Journal* 15 (4), 474-482.

Ashley, W.S., Mote, T.L., Dixon, P.G., Trotter, S.L., Powell, E.J., Durkee, J.D., Grundstein, A.J. 2003. Distribution of Mesoscale Convective Complex Rainfall in the United States. *Monthly Weather Review* 131 (12), 3003-3017.

Byun, H.R., Wilhite, D.A. 1999. Objective Quantification of Drought Severity and Duration. *Journal of Climate* 12 (9), 2747-2756.

Carbone, G.J., Rhee, J., Mizzell, H.P., Boyles, R. 2008. Decision support: A Regional-Scale Drought Monitoring Tool for the Carolinas. *Bulletin of the American Meteorological Society* 89 (1), 20-28.

Dai, A. 2011. Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change* 2 (1), 45-65.

Guttman, N.B. 1999. Accepting the Standardized Precipitation Index: A calculation algorithm. *JAWRA Journal of the American Water Resources Association* 35 (2), 311-322.

Heim, R.R. 2002. A Review of Twentieth-Century Drought Indices Used in the United States. *Bulletin of the American Meteorological Society* 83 (8), 1149-1165.

Husak, G.J., Michaelsen, J., Funk, C. 2007. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *International Journal of Climatology* 27 (7), 935-944.

Isaaks, E.H., Srivastava, R.M. 1989. *An introduction to applied geostatistics*. Oxford University Press, Oxford, New York, 400 pp.

Jain, V.K., Pandey, R.P., Jain, M.K., Byun, H.R. 2015. Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin. *Weather and Climate Extremes* 8, 1-11.

Kaplan, A., Kushnir, Y., Cane, M.A. 2000. Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854-1992*. *Journal of Climate* 13 (16), 2987-3002.

Karl, T.R. 1983. Some Spatial Characteristics of Drought Duration in the United States. *Journal of Climate and Applied Meteorology* 22 (8), 1356-1366.

Keyantash, J., Dracup, J.A. 2002. The Quantification of Drought: An Evaluation of Drought Indices. *Bulletin of the American Meteorological Society* 83 (8), 1167-1180.

Larson, J., Zhou, Y., Higgins, R.W. 2005. Characteristics of Landfalling Tropical Cyclones in the United States and Mexico: Climatology and Interannual Variability. *Journal of Climate* 18 (8), 1247-1262.

Lawrimore, J., Heim, R.R., Svoboda, M., Swail, V., Englehart, P.J. 2002. Beginning a New Era of Drought Monitoring Across North America. *Bulletin of the American Meteorological Society* 83 (8), 1191-1192.

Legates, D.R. 1991. An evaluation of procedures to estimate monthly precipitation probabilities. *Journal of Hydrology* 122 (1), 129-140.

Li, J., Heap, A.D. 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics* 6 (3-4), 228-241.

McEvoy, D.J., Huntington, J.L., Abatzoglou, J.T., Edwards, L.M. 2012. An Evaluation of Multiscalar Drought Indices in Nevada and Eastern California. *Earth Interactions* 16 (18), 1-18.

Menne, M.J., Williams, C.N., Vose, R.S. 2009. The U.S. Historical Climatology Network Monthly Temperature Data, Version 2. *Bulletin of the American Meteorological Society* 90 (7), 993-1007.

Mo, K.C., Long, L.N., Xia, Y., Yang, S.K., Schemm, J.E., Ek, M. 2010. Drought Indices Based on the Climate Forecast System Reanalysis and Ensemble NLDAS. *Journal of Hydrometeorology* 12 (2), 181-205.

Morid, S., Smakhtin, V., Moghaddasi, M. 2006. Comparison of seven meteorological indices for drought monitoring in Iran. *International Journal of Climatology* 26 (7), 971-985.

Murray, J.C., Colle, B.A. 2010. The Spatial and Temporal Variability of Convective Storms over the Northeast United States during the Warm Season. *Monthly Weather Review* 139 (3), 992-1012.

Naumann, G., Dutra, E., Barbosa, P., Pappenberger, F., Wetterhall, F., Vogt, J.V. 2014. Comparison of drought indicators derived from multiple data sets over Africa. *Hydrology and Earth System Sciences* 18 (5), 1625-1640.

Palmer, W.C. 1965. *Meteorological Drought*. U.S. Weather Bureau, Washington, D.C., 58 pp.

Quiring, S.M. 2009. Monitoring Drought: An Evaluation of Meteorological Drought Indices. *Geography Compass* 3 (1), 64-88.

Rhee, J., Carbone, G.J., Hussey, J. 2008. Drought Index Mapping at Different Spatial Units. *Journal of Hydrometeorology* 9 (6), 1523-1534.

Sheffield, J., Wood, E.F., Roderick, M.L. 2012. Little change in global drought over the past 60 years. *Nature* 491 (7424), 435-438.

Shuttleworth, W.J., Wallace, J.S. 1985. Evaporation from sparse crops-an energy combination theory. *Quarterly Journal of the Royal Meteorological Society* 111 (469), 839-855.

Steinemann, A., Iacobellis, S.F., Cayan, D.R. 2015. Developing and Evaluating Drought Indicators for Decision-Making. *Journal of Hydrometeorology* 16 (4), 1793-1803.

Thornthwaite, C.W. 1948. An Approach toward a Rational Classification of Climate. *Geographical Review* 38 (1), 55-94.

Tucker, C.J. 1989. Comparing SMMR and AVHRR data for drought monitoring. *International Journal of Remote Sensing* 10 (10), 1663-1672.

Velasco, I., Fritsch, J.M. 1987. Mesoscale convective complexes in the Americas. *Journal of Geophysical Research: Atmospheres* 92 (D8), 9591-9613.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I. 2009. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate* 23 (7), 1696-1718.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., Angulo, M., El Kenawy, A. 2010. A New Global 0.5° Gridded Dataset (1901-2006) of a Multiscalar Drought Index: Comparison with Current Drought Index Datasets Based on the Palmer Drought Severity Index. *Journal of Hydrometeorology* 11 (4), 1033-1043.

Wu, H., Svoboda, M.D., Hayes, M.J., Wilhite, D.A., Wen, F. 2007. Appropriate application of the standardized precipitation index in arid locations and dry seasons. *International Journal of Climatology* 27 (1), 65-79.

Yuan, S., Quiring, S.M. 2014. Drought in the U.S. Great Plains (1980-2012): A sensitivity study using different methods for estimating potential evapotranspiration in the Palmer Drought Severity Index. *Journal of Geophysical Research: Atmospheres* 119 (19), 10996-11010.

Zarafshani, K., Sharafi, L., Azadi, H., Hosseininia, G., De Maeyer, P., Witlox, F. 2012. Drought vulnerability assessment: The case of wheat farmers in Western Iran. *Global and Planetary Change* 98-99, 122-130.

Zhou, L., Zhang, J., Wu, J., Zhao, L., Liu, M., Lü, A., Wu, Z. 2011. Comparison of remotely sensed and meteorological data-derived drought indices in mid-eastern China. *International Journal of Remote Sensing* 33 (6), 1755-1779.