Response letter to reviewer's comments

Please find the responses to the reviewer's comments in the blue font and the revised manuscript.

Anonymous Referee # 2

The authors use two global precipitation and temperature datasets to calculate the SPEI drought metric using various choices of the calibration period. By quantifying drought severity, duration and extent in maps and by aggregating SPEI values over four regions, conclusions are reached on 'best practice' in setting the calibration period. A very nice touch is that the authors have looked at specific record-dry years for the regions under consideration, and visualised the effects of the choice of calibration period on the drought estimate.

The SPEI, like many other drought metrics, is a standardised metric making its estimates for dry or wet conditions comparable over diverse climatological areas. The issue what calibration period to use, and the effects of not using the full-length of the period for which data is available as calibration, has been debated in the literature. This makes the study a very welcome contribution to the discussion.

However, my view on the manuscript that is currently submitted is that it may raise more questions than answers. There are quite a few things unclear and difficult to believe. Some of the choices made are unlucky, like the regions over which the SPEI is averaged. Some of these analyses need to be looked at again. Nevertheless, the analysis of the very dry years in the final parts of the paper show that the authors are capable of making some fine analysis - it is just a pity that they fail to observed some of the interesting aspects of their results.

The two main concerns relate 1) to the quality and validity of figs. 6, 7 and 8. I have trouble understanding what they mean and (especially fig. 7) can't be correct. 2) The analysis of figs. 9, 10, 11, 12, which touches at the essence of what the authors aim to investigate, is incomplete.

There are many other less serious concerns.

Main concerns

1.

(a) fig. 6: I simply do not understand the quantity that is on the y-axis. The text (page 5, line 9-11) says: "the drought frequency as the ratio between the total number of drought events (...) relative to the total effective grid points." Are you calculating the number of grid points with SPEI12 ≤ -1, and then divide this number by the total number of grid points? Fig. 6 gives me ratios well above 1, so this can't be the case. There is also 'duration' on the x-axis. This is not the length of the time window over which the SPEI is calculated, is it? It seems to be the period for which a grid point stays at or below the SPEI12=-1 threshold, right?

>> First of all, the frequency can be well above 1. The multiple drought events can occur in one grid cell as the monthly SPEI-12 time series over the periods of 1901-195 and 1958-2014 are examined. Second, the duration is how long the SPEI-12 stays

at or below -1, not the length of the time window over the SPEI-12 is calculated. To clarify our approach to calculate the duration-frequency relation for Fig. 7 in the revised manuscript (Fig. 6 in the original manuscript), we have re-written the explanation about it.

Page 7: *"As explained above, a drought event is counted when the monthly SPEI-12 is estimated to be at or below -1.0 for the drought duration-frequency relation. For each drought event of grid cell, the duration is how long the SPEI-12 stays at or below -1. The frequency is the ratio between the total number of drought events and the effective grid points in each region (Fig. 7)."*

(b) Also fig. 7. This can't be true. There is a continuous upward line for Europe (and less so for the US) from 1901-1957 to 1958-2014. This would mean that in 1901 the moving average of SPEI was lowest for the coming century. I do not see dry years in this series like 1921, 1976 or 2003. All lines (for each period and region) have upward slopes. I think why this is (it is because of the use of Thornthwaite) but this is not discussed anywhere. It is strange that fig. 7 has upward trends for EU and US, while none of this is seen in fig. 4

>> Yes, we agree that the continuous upward line from P1 to P2 for EU and US seems strange. Therefore we have performed the analysis again and found the same results as in the original manuscript. To better understand our results in Fig. 9 of the revised manuscript (Fig. 7 of the original manuscript), we have performed an additional analysis as presented in Fig. 4 of the revised manuscript, presenting the averaged temporal variations of annual precipitation, PET and water surplus or deficit D (=P-PET in Eq.1). We have checked those variables as the PET increased significantly from P1 to P2 as the air temperature, which increasing significantly from P1 to P2, is a key controlling factor for the PET, as estimated based on the Thornthwaite approach in this study.

We therefore argue that such continuous upward line from P1 to P2 could be obtained because of the following reasons. First of all, we examine the regionally averaged indices, which do not necessarily capture the local severe drought events. Second, we find that it is consistent with Fig. 4. In US, the increase in precipitation is higher than that in PET, which leads the increase in D. In EU, the increase in PET is higher than that in precipitation, and thus the decrease in D is found in terms of average but the slight increase in the lower extreme of D is found. Therefore the severest drought events present less significant in P2 compared to those in P1. This point has been added in the manuscript as it follows.

Page 7: *"Such findings are seemingly inconsistent with the recently observed severe drought events in US and EU, but it is possible since we examine the regionally averaged indices, not the local extremes of SPEIs. Also it is consistent with Fig. 4. In US, the increase in precipitation is higher than that in PET, which leads the increase* *in D (=P-PET in Eq.1). In EU, the increase in PET is higher than that in precipitation, and thus the decrease in D is found in terms of average but the slight increase in the lower extreme of D is found. Therefore the severest drought events present less severe in P2 compared to those in P1."*

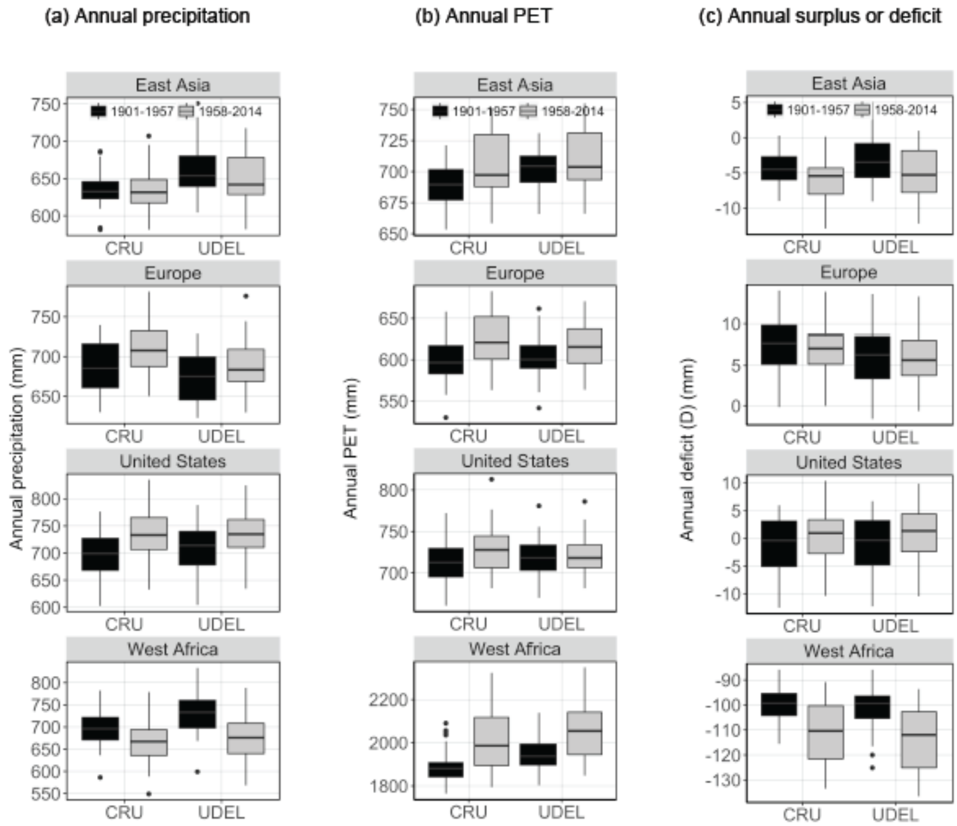

Figure 4. Temporal variations of annual precipitation, PET and surplus or deficit (D=P-PET) depending on two datasets (CRU and UDEL) and periods (1901-1957 and 1958-2014). In the box plot, the center line represents the median value; the top and bottom of box represent the 25th and 75th percentile of the data, respectively; the dot represents the outlier.

(c) fig. 8 is not understandable. The x-axis says 'ascending order', claiming that for 'ascending order' ~ 60 for CRU 1958-2014 and West Africa, 100% has SPEI ≤ -1. What does this mean?

>> As per reviewer's suggestion, we have added the detailed explanation about how we derive Figure 9 in the revised manuscript (Fig. 8 in the original manuscript) as it follows.

Page 8: *"We count the numbers of grid points with the SPEI-12 less than -1.0 for each period (i.e., P1 and P2) and divide them with the effective grid numbers in the region to derive the spatial extent, i.e., the grid percentage of droughts. Then the annual time series of spatial extent are sorted in ascending order, from the smallest to the largest."*

>> In addition, the spatial extents in WA are as large as 89.6% with CRU and 87.7 % with UDEL for 1958-2014. It means that at least one year showed the drought events (with the SPEI-12 less than -1) occur over the 89.6% (87.7%) of the study area in WA with CRU (UDEL).

2. The analysis of figs. 9-12 is a good idea, but there are a few things the authors need to explain and re-do.

(a) it is not clear which SPEI12 value is taken. For instance, the 2003 heat wave in Europe, which coincided with a dry period, the height of the heat wave was early August 2003. The Spring was rather dry and the heat wave stopped when Autumn was a bit wetter than usual for France and Spain and in December, northern Europe received more rain than usual. The question is now: which month provides the SPEI12 value which is characteristic of the 2003 drought in Europe? SPEI12 is based on 12-month accumulated precipitation, so do you take the December value (so that the whole of 2003 is captured)? Or do you take the annual averaged SPEI12, which then has a small influence of January 2002 as well in it..... A pragmatic approach is to make time series of monthly SPEI12 values and then take the month in 2003 with the lowest value, but over which area to average? The whole of Europe is nonsense, since the heatwave was much more local than that.

>> We have examined the annual SPEI-12 based on the data from January to December in each year. Therefore our results show the average stage of the European heat wave in 2003. It is possible to take the suggested approach by the reviewer (i.e., taking the lowest value of monthly SPEI-12 for each grid cell), but it is not relevant to the focus of this study. Therefore it could be performed in the future study. In the revised manuscript, we have clarified our approach to estimate the SPEI-12 for Figs. 11-14 (Figs. 9-12 in the original manuscript) as well as possibly different results according to the local area as it follows.

Page 9: *"Here the annual SPEI-12s with the monthly climate data from January to December in each year are first constructed and then the SPEI-12s for a chosen year are examined in detail."*

Page 10: *"Although this study with historical data may shows the different results depending on the selected local area, a similar study with historical data or climate change scenarios in different regions would undoubtedly strengthen our findings."*

(b) All figures 9-12 show that for Ref3, the SPEI values are off the scale for some areas. I think that this is the main issue with the Ref3 approach. The SPEI (and SPI) are more-or-less normally distributed. By using the calibration from one period (like 1901-1957), you run the risk that the metric 'explodes' beyond the range in which the SPEI/SPI lives when droughts occur in the 1958-2014 period which are (much) more severe than anything seen in the calibration period. Essentially, using Ref3, you are not only assuming stationarity of the climate but also that the the whole probability distribution of droughts (and pluvials) is sampled in this period.

>> Yes, we agree. As per reviewer's suggestion, we have added a relevant discussion in the revised manuscript as it follows.

Page 9: *"In particular, the several extremes (i.e., out of the scale ranges in Figs. 11-14) of SPEI-12 in Ref3 cases highlight the importance of the reference period. By using the reference period of the certain past time (P1 in this study, i.e., Ref3), the drought events in the estimation period could be beyond the range in which the distribution is calibrated for the index. Essentially, using Ref3, it is assumed that not only the stationarity of the climate but also that the whole probability distribution of droughts is sampled in this period."*

Other issues the authors may want to look into

1. The CRU dataset (and presumably the UDEL dataset as well) relax values to climatological values when data is insufficiently available. For Europe and North America, this will not be a big problem I think, but for West Africa and South Asia the number of records going back to 1901 are few and far in between. This means that the early period in these regions sees much less month-to-month variability as the more recent periods. Discuss the implications of this on your results.

>> We have added the suggested discussion about the lack of ground-based observations in the early period in the revised manuscript as it follows:

Page 6: *"Furthermore, the variances of SPEI are relatively small for P1 compared to those for P2 in EA and WA while no noticeable differences in the variances are captured in EU and US. It may attribute to the lack of ground-based observations before 1950 (i.e., the most of P1) (Becker et al., 2013; Vittal et al., 2013; Nasrollahi et al., 2015) and such limit in data availability seems play a role in reducing the variance of SPEI for P1 in EA and WA."*

2. The issue of the sensitivity has been raised earlier by Van der Schrier et al. 2013 (doi:10.1002-jgrd.50355) and Trenberth et al. 2014.

>> The past studies about the sensitivity in regard to the reference period have been included as it follows:

Page 2: *"It has already been pointed out for the self-calibrated PDSI that trends towards more extreme conditions are amplified when the calibration period does not include the recent part of data, including the recent effects of climate change (Van der Schrier et al., 2013; Trenberth et al., 2013)."*

3. The Thornthwaite (1948) parameterization is directly related to temperature and has a huge trend. Even without a trend in precipitation amounts, the difference between the two will have a drying trend. This should be noted in the ms. and observed in the figures.

>> We have pointed out the use of Thornthwaite and its influences throughout the revised manuscript as it follows.

Page 5: "Precipitation, air temperature and PET are investigated because they are used to estimate the SPEIs (Figs. 2 and 3 and Table 3). As noted already, the SPEIs are estimated based on the distribution of D (=P-PET in Eq. 1) and here the air temperature is directly related to PET because we use the Thornthwaite approach to estimate PET."

>> Furthermore, we have added the analyses of PET and D(=P-PET) in addition to the analyses of precipitation and air temperature in Figs. 3, 4 and 7 as well as Table 3 in the revised manuscript.

4. Sect. 2.2 It would be helpful for many readers what the descriptions are associated with the various SPI/SPEI values and the chance that 'severe' or 'extreme' drought is likely to occur. These are available in the McKee article or in: edo.jrc.ec.europa.eu - documents - factsheets - factsheet_spi.pdf

>> As per reviewer's suggestion, we have added Table 1 to show the categories of drought.

Table 1. Classification of dry status in this study (Mckee et al., 1993).

Category	Description	SPEI
D1	Moderate dry	≤ -1.0
D2	Extreme dry	≤ -2.0
D3	Very extreme dry	≤ -3.0

5. page 5, line 13-14. It is a good idea to see how large the region is with SPEI ≤ -1. However, simply counting grid squares does not work. You need to calculate area, where the grid areas are weighted with the cosine of the latitude.

>> We agree that the area of each grid cell varies significantly according to the latitude. Therefore we have revised the relevant text to point out that for the spatial extent of drought, we calculate the number of drought grid points relative to the total effective grid points in each region, not the drought area relative to the total area in each region. Consequently, we revised the definition word from "area" to "spatial extent" not to misunderstand and added the detail meaning in the manuscripts as it follows:

Page 8: *"The spatial extents of droughts for the annual SPEI-12 ≤ -1.0 are examined by sorting the results in ascending order (Fig. 10). We count the numbers of grid points with the SPEI-12 less than -1.0 for each period (i.e., P1 and P2) and divide them with the effective grid numbers in the region to derive the spatial extent, i.e., the grid percentage of droughts. Then the annual time series of spatial extent are sorted in ascending order, from the smallest to the largest."*

6. page 6, lines 1-10. Interesting analysis, but the areas defined are more-or-less arbitrary. In Europe, for instance, there is a wetting trend in northern Europe and a drying trend in southern Europe. It makes more sense to separate these two. Take a selection of the Giorgi regions: www.ipcc.ch - ipccreports - tar - wg1 - images - fig10-1.gif

>> Yes, our selection of study region is more-or-less arbitrary. Each region could be divided into the sub-regions based on the climate classification as the reviewer suggested. While different studies use different climate sub-regions in their analyses, we more focus on the detailed changes within the large region such as EA, WA, US and EU. In this study, Fig. 4 and Table 3 do not show the detailed change in precipitation, air temperature and PET with the areal averages, but Figs 2 and 3 show the spatially distributed maps of averages and temporal changes. With putting Figs 2, 3 and 4 and Table 3 together, we were able to present different trends within the region in this study. Indeed, Fig. 3 presents a wetting trend in the northern Europe and a drying trend in the southern Europe. Therefore we would perform such sub-regional analyses based on the climate classification in the future study as it is pointed out in the revised manuscript.

Page 3: *"We perform the analyses based on the spatially distributed patterns over those regions as well as their averages, but without distinguishing the sub-regions based on the climate characteristics."*

Page 10: *"Although this study with historical data may shows the different results depending on the selected local area, a similar study with historical data or climate change scenarios in different regions would undoubtedly strengthen our findings."*

7. page 6, Instead of calculating the trends in temperature, it makes more sense to calculate the Thornthwaite PET value. This makes this analysis directly comparable to the trends in precipitation.

>> As per reviewer's suggestion, we have added the analyses of PET throughout the manuscript. Figs. 3, 4 and 7, and Table 3 are revised to include PET in addition to the precipitation and air temperature.

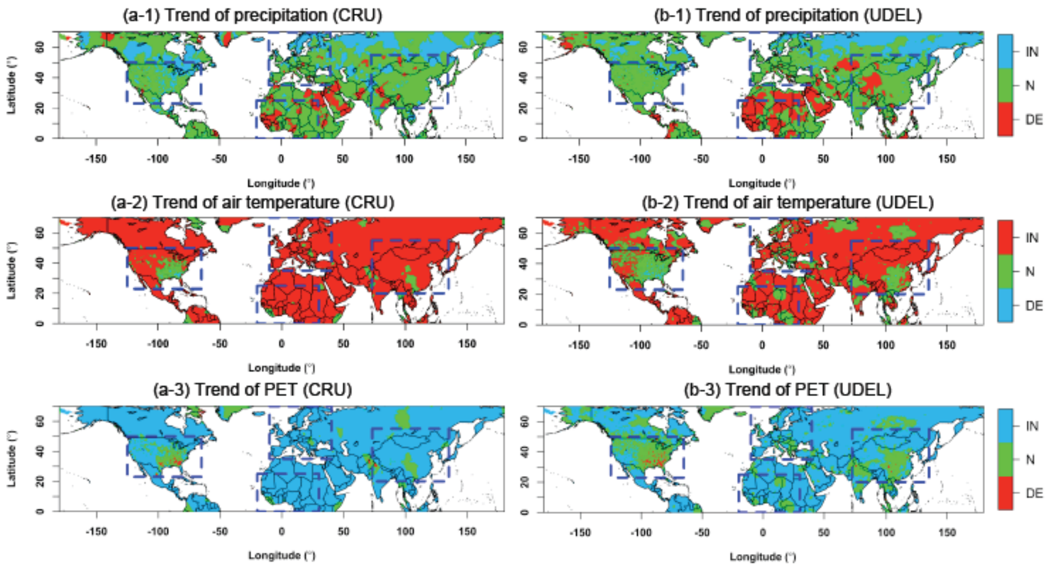

Figure 3. Trends of annual precipitation, annual averaged temperature and annual PET for the CRU and UDEL datasets. PR and TA denote precipitation and temperature, respectively, and IN, N and DE indicate increasing, no trend and decreasing, respectively.

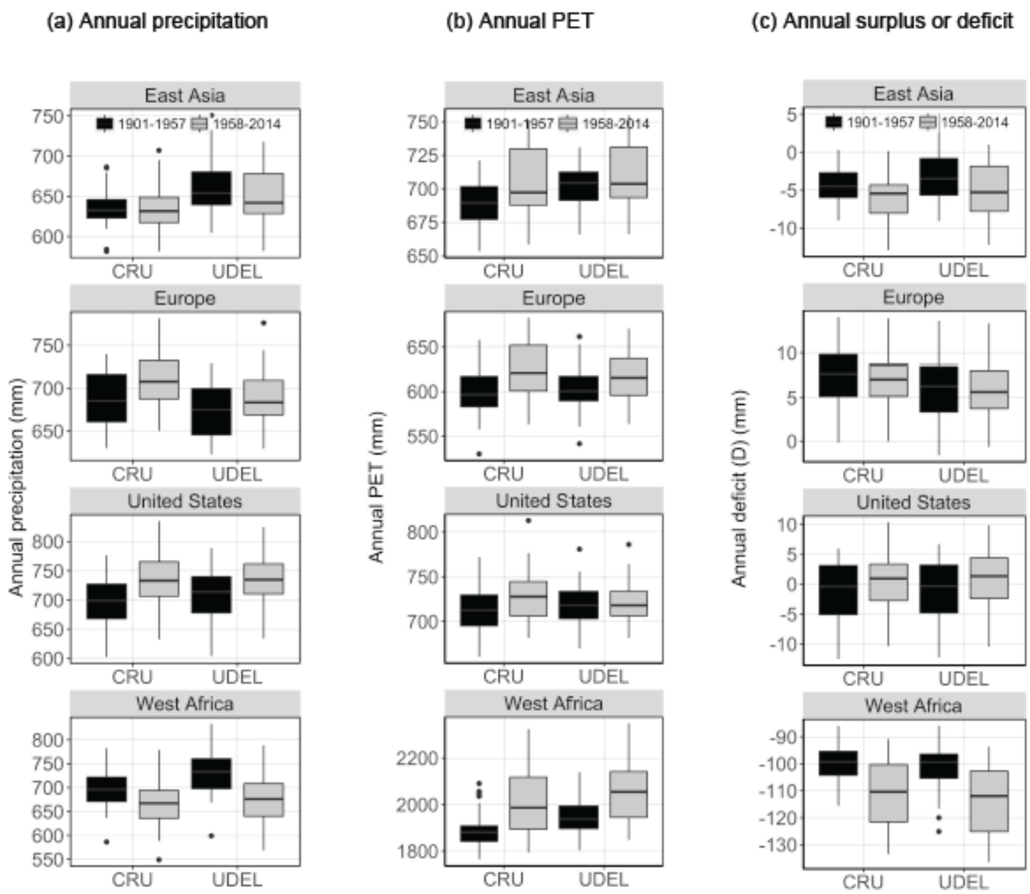

Figure 4. Temporal variations of annual precipitation, PET and surplus or deficit (D=P-PET) depending on two datasets (CRU and UDEL) and periods (1901-1957 and 1958-2014). In the box plot, the center line represents the median value; the top and bottom of box represent the 25th and 75th percentile of the data, respectively; the dot represents the outlier.

8. fig. 2. Perhaps show the CRU climatology and the difference between UDEL and CRU? The pictures are very similar now.

>> As we focus on the difference between P1 and P2, we have added the difference figure between P1 and P2 in Figure 2.

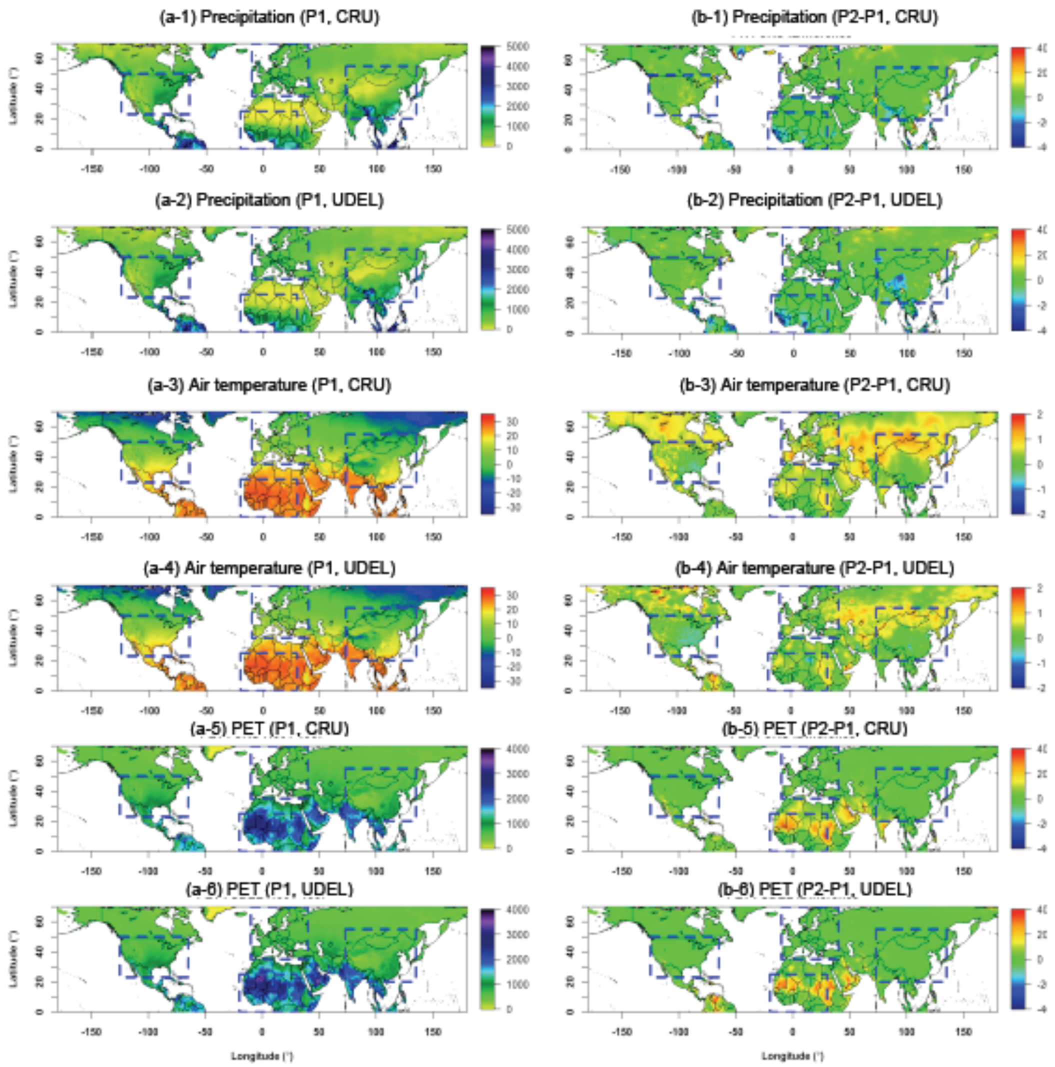

Figure 2. Annual precipitation (mm), annual averaged temperature (°C) and annual PET (mm) for the CRU and UDEL datasets for P1 and their difference between P1 and P2.

9. fig. 4. What are we seeing? Is that the median, the 25th and 75th percentiles and min & max values? This is nowhere in the text or caption.

>> As per reviewer's suggestion, we have added the detailed explanation for the box plot.

Figure 6 in the revised manuscript: *"In the box plot, the center line represents the median value; the top and bottom of box represent the 25th and 75th percentile of the data, respectively; the dot represents the outlier."*

10. table 2. I see values of 26.1 degrees for area NA. This can't be North America (which is US in the text).

>> We have corrected North America to US in abstract and Table 3.

Page 1: *"Focusing on East Asia, Europe, United States and West Africa"*

Effects of different reference periods on drought index estimations for 1901-2014

Myoung-Jin Um[1], Yeonjoo Kim[1,*], Daeryong Park[2], Jeongbin Kim[1]

[1]Department of Civil and Environmental Engineering, Yonsei University, Seoul, 03722, Republic of Korea
[2]Department of Civil, Environmental and Plant Engineering, Konkuk University, Seoul 05029, Republic of Korea

Correspondence to: Yeonjoo Kim (yeonjoo.kim@yonsei.ac.kr)

Abstract. This study aims to understand how different reference periods (i.e., calibration periods) of climate data for estimating the drought index influence regional drought assessments. Specifically, we investigate the influence of different reference periods on historical drought characteristics such as trends, frequency, intensity and spatial extents using the standard precipitation evapotranspiration index with a 12-month lag (SPEI-12) estimated from the datasets of the climate research unit (CRU) and the University of Delaware (UDEL). For the 1901–1957 (P1) and 1958–2014 (P2) estimation periods, three different types of reference periods are used: P1 and P2 together, P1 and P2 separately and P1 only. Focusing on East Asia, Europe, United States and West Africa, we find the influence of the reference periods to be significant in East Asia and West Africa, with dominant drying trends from P1 to P2. The reference periods influence the assessment of drought characteristics, particularly for severity and spatial extent, whereas their influence on the frequency is relatively small. Finally, self-calibration, which is the most common practice with an index such as SPEI, tends to underestimate the drought severity and spatial extent relative to the other approaches used in this study. Although the conclusions drawn in this study are limited to two global datasets, they nevertheless highlight the need for the reference period to be clarified in drought assessments to better understand regional drought characteristics and their temporal changes, particularly under climate change scenarios.

1 Introduction

Drought is a complex, slow-onset natural phenomenon affecting more people than any other hazards and seriously influencing water resources, agriculture, society and ecosystems (Hagman, 1984; Wilhite, 2002; Ionita et al., 2015). As drought impacts are largely nonstructural and spread over a relatively large region, the onset and end of a drought as well as its severity are often difficult to determine (Wilhite, 2002). Furthermore, based on recent changes in the 21st century and projected climate warming, such drought phenomena will likely worsen (Sheffield and Wood, 2008; Dai, 2010). Sheffield et al. (2012) stated that the severe and prolonged drought events have been witnessed since the 1970s and their changes are related to higher temperatures and lower precipitation.

Drought can be defined and explained using absolute or relative terminology, allowing these terms or measures to be compared to each other (Dai, 2011; Trenberth et al., 2014). For the absolute term, the amount of precipitation, the amount of soil moisture and other metrics can be used. The relative measures include the Palmer drought severity index (PDSI), the standardized precipitation index (SPI), the standardized precipitation and evapotranspiration index (SPEI) and others. Vicente-Serrano et al. (2010) and Vicente-Serrano and Beguería-Portugués (2003) suggested that drought indices were not as useful because they were based on standardized or normalized shortages relative to average conditions in a given station and period. Nevertheless, various drought indices have been widely used in many drought studies.

Dracup et al. (1980) suggested three components of drought: duration, magnitude (average water deficiency) and severity (cumulative water deficiency). Such concepts have been applied to various drought indices to analyze historical characteristics. Wang et al. (2011) defined the intensity-duration-frequency of droughts with the SPI, standardized runoff index (SRI), standardized soil water index (SSWI) derived from observations and future regional climate change projections in central Illinois. To evaluate how well the global climate models simulate observed drying or wetting trends, Nasrollahi et al. (2015) applied the Mann-Kendall trend test to the SPIs derived from global observational climate data, that is, the dataset from the climate research unit (CRU), and 41 predictions with global climate models (GCMs) from the Coupled Model Intercomparison Project Phase 5 (CMIP5). Similarly, Tan et al. (2015) utilized the climate data from 22 meteorological stations in Ningxia, a well-known food production area in Northwest China, and performed the Mann-Kendall trend tests with SPI and SPEI. The degreesn of increasing drought frequency and intensity varied with the stations in the study region. Furthermore, Touma et al. (2015) used data from 15 GCMs in the CMIP5 and assessed the likelihood of changes in the spatial extent, duration and number of occurrences of four drought indices, including SPI, SPEI, and others.

Estimating the drought index requires a calibration step. Specifically, historical data such as precipitation should be fitted to a specific probability distribution function (PDF) to be used for estimating drought indices. A few previous studies addressed the issue of data periods for the calibration step (e.g., Karl et al., 1996; Dubrovsky et al., 2009). While it is common to use self-calibrated indices (i.e., using the same dataset for calibration and index estimation), some studies suggest calibrating with the reference climate data to allow an inter-comparison of the index among stations or different periods (Dubrovsky et al., 2009). Such the reference period (i.e., calibration period) of climate data would be particularly important in climate change studies. It has already been pointed out for the self-calibrated PDSI that trends towards more extreme conditions are amplified when the calibration period does not include the recent part of data, including the recent effects of climate change (Van der Schrier et al., 2013; Trenberth et al., 2013). Still, only limited studies clarify their approaches to calibration.

For this reason, we aim to understand how a different reference period (i.e., calibration period) of climate data influences the regional drought assessment. Specifically, we investigate the influence of different reference periods on historical drought characteristics such as trends, frequency, intensity and spatial extents with the SPEI estimated with two historical global climate datasets from the CRU and the University of Delaware (UDEL). This study shows that the reference periods influence the assessment of drought characteristics, particularly for severity and spatial extent, while its influence on the frequency is relatively small. These influences are especially significant in regions with dominant drying trends such as East Asia and West Africa. These findings suggest that the reference period should be clarified in drought assessments for a better understanding of regional drought characteristics and their temporal changes.

2 Materials and methods

2.1 Study area and climate data

We investigate the drought characteristics over the Northern Hemisphere with a focus on four different regions, including East Asia (EA), Europe (EU), the United States (US) and West Africa (WA) (Fig. 1). We perform the analyses based on the spatially distributed patterns over those regions as well as their averages, but without distinguishing the sub-regions based on the climate characteristics. Two widely used global observational datasets from the CRU and UDEL are utilized in this study. From these two datasets, monthly precipitation and temperature data are used with a spatial resolution of 0.5° from 1901 to 2014.

This study uses the latest CRU dataset (CRU TS3.10) as described in Harris et al. (2014). The principal sources of the CRU are the World Meteorological Organization (WMO) in collaboration with the US National Oceanographic and Atmospheric Administration (NOAA). Covering all land area between 60°S and 80°N at a spatial resolution of 0.5°, the dataset includes global monthly climate data on ten variables: precipitation, mean temperature, diurnal temperature range, minimum and maximum temperature, vapor pressure, cloud cover, rain days, frost days and potential evapotranspiration. The dataset is derived from archives of climate station records with extensive manual and semi-automated quality control measures.

The UDEL dataset (V 4.01, Willmott and Matsuura, 2001) is also used in this study. The dataset includes gridded monthly precipitation and temperature data at a spatial resolution of 0.5° across the land over the globe. The dataset was compiled from sources including the Global Historical Climatology Network (GHCN) and the Global Surface Summary of Day (GSOD). To interpolate the station values to the grid, climatologically aided interpolation (CAI) and traditional interpolation were used for precipitation and digital-elevation-model (DEM)-assisted interpolation, traditional interpolation and CAI for temperature. In this work, traditional interpolation is a spherical version of Shepard's algorithm, which employs an enhanced distance-weighting method (Shepard, 1968; Willmott et al., 1985).

2.2 Meteorological drought index

Various drought indices have been used to understand different types of droughts, including meteorological drought, agricultural drought and hydrological drought (Heim, 2002). For meteorological droughts, the indices include the PDSI (Palmer, 1965), the SPI (McKee et al., 1993) and the SPEI (Vicente-Serrano et al., 2010). As different studies used different meteorological drought indices (Seneviratne, 2012; Sheffield et al., 2012; Trenberth et al., 2014; Nasrollahi et al., 2015; Touma et al., 2015), this study focuses on the SPEI. Devised by Vicente-Serrano et al. (2010), the SPEI has the advantage of being able to consider the effects of temperature variability for the drought relative to the SPI (Naumann et al., 2014) because the potential evapotranspiration (PET) can be calculated with air temperature based on Thornthwaite (1948). The SPEI uses the amount of precipitation minus PET and fits the data to the log-logistic probability distribution function. Here, we summarize the steps to estimate SPEI based on monthly precipitation and temperature. The detailed procedure for estimating the SPEI is well presented in Vicente-Serrano et al. (2010).

Step 1: Estimate the water surplus or deficit in month j (D_j) using the difference between precipitation (P_j) and potential evapotranspiration (PET_j):

$$D_j = P_j - PET_j \quad (1)$$

Here, the potential evapotranspiration is estimated based on the method of Thornthwaite (1948), which requires the monthly temperature, latitude, day and month.

Step 2: Estimate the accumulated difference ($X_{i,j}^k$) over the time scale k in a given month j and year i. For example, the accumulated difference for a month in a particular year with a 12-month time scale is calculated as follows:

$$X_{i,j}^k = \sum_{l=13-k+j}^{12} D_{i-1,l} + \sum_{l=1}^{j} D_{i,j}\,, \qquad if\ j < k \quad (2)$$

$$X_{i,j}^k = \sum_{l=j-k+1}^{j} D_{i,l}\,, \qquad if\ j \geq k \quad (3)$$

Step 3: Fit the accumulated difference to a log-logistic distribution as follows:

$$F(X) = \left[1 + \left(\frac{\alpha}{x-\gamma}\right)^{\beta}\right]^{-1} \quad (4)$$

where $F(X)$ is the cumulative probability function of a three-parameter log-logistic distribution with α, β and γ representing the scale, shape and origin parameters, respectively. For the model fitting, the L-moment procedure (Hosking, 1990) is employed as it is one of the most robust and easy-to-use approaches.

Step 4: Estimate the SPEI based on the estimated $F(X)$. The SPEI can be derived from the standardized values of $F(X)$ and the classical approximation of Abramowitz and Stegun (1965) following Vicente-Serrano et al. (2010). The estimated drought index is classified as in Table 1 for moderate, extreme and very extreme dry cases. In this study, we focus on the SEPI with the 12-month lag (SPEI-12). SPEI can be estimated for different lag times such as 1, 3, 6, 9, 12 and 24 months.

2.3 Temporal trends and statistical characteristics

This study investigates various measures of historical droughts, including trend, frequency, severity and spatial extent (Lloyd-Hughes and Saunders, 2002; Wang et al., 2011; Hoerling et al., 2012; Seneviratne, 2012; Trenberth et al., 2014; Touma et al., 2015).

The temporal trend is investigated with a nonparametric and monotonic trend test with the S-statistic (Mann, 1945; and Kendall, 1976) as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i) \quad (5)$$

where $sgn(x_j - x_i) = \begin{cases} +1, & (x_j - x_i) > 0 \\ 0, & (x_j - x_i) = 0 \\ -1, & (x_j - x_i) < 0 \end{cases}$ (6)

where sgn is the sign function and n is the sample size. The statistical significance of the trend can be predicted by a Z test as follows:

$$Z = \begin{cases} (S-1)/\sigma_s\,, & if\ S > 0 \\ 0, & if\ S = 0 \\ (S+1)/\sigma_s\,, & if\ S < 0 \end{cases} \quad (7)$$

$$\sigma_s = \sqrt{\left(n(n-1)(2n+5) - \sum_{j=1}^{q} t_j(t_j-1)(2t_j+5)\right)/18} \quad (8)$$

where σ_s is the square root of S in the case that the x values are possible tie situations, q is the number of ties in the dataset and t_j is the number of data in the jth tie group. The trend in the data does not exist for $Z < Z_{\alpha/2}$ at the significance level α.

For the frequency, severity and spatial extent of drought, different measures have been defined and used in past studies (e.g., Wang et al., 2011; Touma et al., 2015; Um et al., 2016) because it is not straightforward to define these quantities in practice. For example, Touma et al. (2015) defined the duration, occurrence and spatial extent of drought to investigate the drought changes with 15 CMIP5 models throughout the world for the 21st century: the duration of drought is defined as the consecutive period below a certain drought status, the occurrence of droughts is defined as the total number of droughts in the period of interest, and the spatial extent of droughts is defined as the percentage of grid points below the given drought level, in which the corresponding drought index is less than the given drought category for each month.

In this study, we define three measures of droughts with the SPEI-12: (1) the drought frequency as the ratio between the total number of drought events, which is defined as the SPEI-12 ≤ -1, and the total effective grid points; (2) the severity as the lowest estimates among the regional monthly average SPEI-12 with moving windows with periods of 1 to 12 months; here, the regional averages are estimated for the four study regions depicted in Fig. 1; and (3) the spatial extent as the number of grids with the annual SPEI-12 ≤ -1.0 relative to the total grids.

2.4 Design of data analysis

To understand the influence of the reference period (i.e., calibration period) on the drought index, three different types of reference periods are used to estimate the SPEI-12 with the CRU and UDEL. To analyze separately the drought characteristics for the estimation periods of 1901–1957 (P1) and 1958–2014 (P2), different sets of reference periods are used (Table 2). Here, we assume that the mean climates of P1 and P2 are different to some extent because of global climate and environmental changes, which will be discussed further in Section 3. For the first type of reference period (Ref1), we calibrate the distribution of a specific PDF (Step 3 in Section 2.2) using the data from 1901 to 2014, which is used for estimating the SPE12 for the P1 and P2 estimation periods. For the second type of reference period (Ref2), calibrations are performed separately for P1 and P2, and thus so-called self-calibrated indices are derived. For the third type (Ref 3), we calibrate the distribution using the data from P1 (i.e., 1910–1957) and then use this distribution for both estimation periods.

3 Results and discussion

3.1 Spatial and temporal patterns of climate variables

Precipitation, air temperature and PET are investigated because they are used to estimate the SPEIs (Figs. 2 and 3 and Table 3). As noted already, the SPEIs are estimated based on the distribution of D (=P-PET in Eq. 1) and here the air temperature is directly related to PET because we use the Thornthwaite approach to estimate PET. The selected regions show different climate features (Fig. 2), and EA and WA include the regions with a relatively wide range of mean precipitation from almost zero to more than 2000 mm per year. In terms of mean air temperature, it is clear that WA is generally quite warmer than other regions. Thus the relatively high PET in WA is clearly presented. Furthermore, the mean precipitation, air temperature and PET are quite similar between the CRU and UDEL.

To investigate the temporal changes of precipitation, air temperature and PET, we compared the means and the standard deviations between two periods (i.e., P1 and P2) in Table 3 and performed the Mann-Kendall trend test (Fig. 3). Table 3 presents clearly different temporal patterns for precipitation depending on the regions and all increasing temporal patterns for air temperature. Additionally, the annual precipitation in EA slightly decreased from 637.19 mm to 635.52 mm in the CRU (-0.2%) and from 659.67 mm to 649.21 mm in the UDEL (-1.6%). Moreover, in WA, the annual precipitation decreased substantially from 698.49 mm to 666.59 mm in the CRU (-4.6%) and from 734.84 mm to 676.11 mm in the UDEL (-8.0%). However, the annual precipitation increased in EU (25.17 mm (5.4%) in the CRU and 14.14 (3.5%) mm in the UDEL) and US (37.78 mm (3.7%) in the CRU and 24.92 mm (2.1%) in the UDEL).

For annual averaged air temperature, the averaged changes of air temperature between P1 and P2 in CRU (UDEL), which were 0.59 (0.37)°C in EA, 0.50 (0.27)°C in EU, 0.32 (0.05)°C in US and 0.35 (0.26)°C in WA, were generally greater than the differences between the CRU and UDEL. The annual averaged temperature became higher from the EA (6.16°C) to EU (6.99°C) to US (10.59°C) to WA (10.52°C) for P1. Consequently, the increasing ratios of annual averaged temperature in CRU (UDEL) were 9.70 (5.92)%, 7.18 (3.85)%, 3.06 (0.47)% and 1.33 (0.98)% in the EA, EU, US and WA, respectively. Such changes in air temperature directly influence the changes in PET, as we used the Thornthwaite approach for estimating PET.

For annual PET, the average growth amounts in the CRU and UDEL in P2 were higher than those in P2, in which increases are 17.09 mm and 9.08 mm in CRU and UDEL in EA, 26.42 mm and 14.20 mm in CRU and UDEL in EU, 17.11 mm and 2.42 mm in CRU and UDEL in US, 111.80 mm and 95.37 mm in CRU and UDEL in WA, because the air temperature is main factor to estimate the PET in this study. Consequently, the increasing ratios of annual PET in CRU (UDEL) were 2.5 (1.3)%, 4.4 (2.4)%, 2.4 (0.3)% and 5.9 (4.9)% in the EA, EU, US and WA, respectively.

The Mann-Kendall trend tests for annual precipitation, annual averaged temperature and annual PET were also performed, as shown in Fig. 3. The data reflect whether these variables showed statistically increasing, decreasing or no trends. For annual precipitation in EA, the areal extent with increasing trend was almost twice than that with a decreasing trend in the CRU, but the areal extent with a decreasing trend in the UDEL was broader than that with increasing area. In EU and US, the areal extent with an increasing trend was clearly greater than that with decreasing area in both the CRU and UDEL. However, in WA, the areal extent with a decreasing trend was larger than that with an increasing trend in both the CRU and UDEL. These patterns were usually more severe in the CRU than those in the UDEL. For annual averaged air temperature and PET, the CRU showed an increasing trend over most of the regions. Similar patterns were found in the UDEL, but the areal extent of the decreasing trend was slightly larger than that in the CRU.

3.2 Temporal patterns of drought index

The drought index (i.e., SPEI-12) is estimated for two periods of P1 and P2 with three different reference periods (Table 2) as described in Section 2.4. Fig. 5 shows the temporal variations of SPEI-12 depending on the reference periods (Ref1, Ref2 and Ref3) and datasets (CRU and UDEL) for the two periods. For US and EU, the SPEI-12 averages are very similar for the two periods: 0.005 (P1) and 0.118 (P2) in the US and -0.011 (P1) to -0.001 (P2) in EU. In EA, the SPEI-12 averages with the three different reference periods slightly decrease from P1 to P2, whereas the deviations of SPEI-12 increase markedly. In WA, the averages and deviations of SPEI-12 significantly decrease and increase, respectively, from P1 to P2. Furthermore, the variances of SPEI are relatively small for P1 compared to those for P2 in EA and WA while no noticeable differences in the variances are captured in EU and US. It may attribute to the lack of ground-based observations before 1950 (i.e., the most of P1) (Becker et al., 2013; Vittal et al., 2013; Nasrollahi et al., 2015) and such limit in data availability seems play a role in reducing the variance of SPEI for P1 in EA and WA. With regional averages, the role of the reference period is not clear and thus we investigate the spatial patterns of SPEI-12 hereafter.

Based on the Mann-Kendall trend test with annual SPEI-12 from 1901 to 2014, we present the increasing (i.e., wetting), decreasing (i.e., drying) or no trend over the regions (Fig. 6). First, the spatial distribution of SPEI-12 trends is identical between Ref1 and Ref3 and that in Ref2 is different. Ref1 and Ref2 use different calibration datasets but are similar in using one dataset for the two estimation periods; Ref2 uses a different calibration dataset for different estimation periods (Table 4). Therefore, SPEI-12 with Ref2 shows relatively less area with wetting and drying trends for the first and second periods relative to Ref1 and Ref2.

Regarding the temporal characteristics over different regions, the following are our findings based on Ref1 and Ref3: In WA, the drying trends are clearly dominant. In EU, the drying trends are scattered over the domain. In US, the wetting trends are scattered in the eastern region and the drying trends in the southwestern region. In EA, the drying trends are clearly in the western region.

Based on the grid-level trend analyses of precipitation, air temperature, PET and SPEI-12, we categorize each grid cell based on increasing, decreasing or neutral trends for each variable (i.e., precipitation, air temperature, PET and SPEI-12). For SPEI-12, increasing and decreasing trends represent wetting and drying trends. We present the ratio of each case relative to the total number of cases (i.e., total number of effective grid cells in all four regions), as shown in Fig. 7. First, the SPEI-12 trends are the same between Ref1 and Ref3, as the estimation periods share the one reference period in both Ref1 and Ref3 while each estimation period uses its own reference period in Ref2. Thus, the values of SPEI-12 are different in both cases, but the trends (i.e., relative values) are the same. Second, precipitation and air temperature exhibit neutral (or no) trends (in the center panel; presumably stationary climate), and the grid percentages of different trends in SPEI-12 vary between Ref1/Ref3 and Ref2. However, the ratio is relatively small, as most grid cells display increasing temperature and PET trends. Finally, based on neutral precipitation and increasing air temperature and PET trends in most grid cells, the numbers of cells with neutral and drying SPEI-12 trends are notably different between Ref1/Ref3 and Ref2. An increasing temperature and PET trend can be observed in most regions; thus, it is important to consider its impact on SPEI. It is particularly true as we use the Thornthwaite approach using the air temperature as a significant control variable of PET.

3.3 Frequency, severity and spatial extent of drought

In this section, we examine how the reference periods play a role in assessing the frequency, severity and spatial extent of drought using SPEI-12. The definitions of frequency, severity and spatial extent of drought used in this study are clarified in Section 2.3, and they may differ in different studies.

As explained above, a drought event is counted when the monthly SPEI-12 is estimated to be at or below -1.0 for the drought duration-frequency relation. For each drought event of grid cell, the duration is how long the SPEI-12 stays at or below -1. The frequency is the ratio between the total number of drought events and the effective grid points in each region (Fig. 8). We first find that the drought events with longer durations (prolonged right tails in the plot) occur more frequently in P2 than in P1 in all regions. However, we do not find any particular differences among the three different reference periods except in WA. The drought frequencies differ among the three reference periods. The frequencies with Ref2 and Ref3 are higher than those with Ref1 for P1, and slight differences in the frequency among the three reference periods are found around the 12-month duration for P2.

We examine how the severity of drought varies with the moving window sizes for the averaged monthly SPEI-12. Fig. 9 shows the severest SPEI-12 estimates, defined as the lowest value among the regional monthly average of SPEI-12 for the moving windows from 1 month to 12 months. In EU and US, we find no large differences among the SPEI-12s with Ref1, Ref2 and Ref3 for the same period. In these regions, the severest SPEI-12s for P1 are higher than those for P2. Such findings are seemingly inconsistent with the recently observed severe drought events in US and EU, but it is possible since we examine the regionally averaged indices, not the local extremes of SPEIs. Also it is consistent with Fig. 4. In US, the increase in precipitation is higher than that in PET, which leads the increase in D (=P-PET in Eq.1). In EU, the increase in PET is higher than that in precipitation, and thus the decrease in D is found in terms of average but the slight increase in the lower extreme of D is found. Therefore the severest drought events present less severe in P2 compared to those in P1. Nonetheless such changes in SPEI-12s according to the relative changes between P and PET suggests the important role of air temperature in drought severity in particular because the Thornwaite approach, using the monthly temperature as a major control variable for PET, is used to estimate the SPEI in this study.

In EA and WA, there exist different patterns in the severest SPEI-12s. The annual precipitation and air temperature (and thus PET) exhibit regionally scattered decreases and widespread increases, respectively (Fig. 3). Consequently, the droughts in 1958–2014 are more severe than those in P1. Furthermore, the severities vary significantly with the calibration periods in EA and WA, where the changes in precipitation and air temperature between two periods are marked.

The spatial extents of droughts for the annual SPEI-12 ≤ -1.0 are examined by sorting the results in ascending order (Fig. 10). We count the numbers of grid points with the SPEI-12 less than -1.0 for each period (i.e., P1 and P2) and divide them with the effective grid numbers in the region to derive the spatial extent, i.e., the grid percentage of droughts. Then the annual time series of spatial extent are sorted in ascending order, from the smallest to the largest. No specific patterns are evident for EU and US. In EA and WA, the spatial extents are generally broader in P2 than in P1. In particular, the spatial extents in 1958–2014 clearly diverge among the different calibration periods, suggesting the importance of the calibration (i.e., reference periods in assessing the droughts in a region).

To understand how the drought characteristics would change if the reference period is dry or wet, we compare the drought spatial extent (%) for dry and wet cases in EA, EU, US and WA with defining dry and wet cases as below. We define the drought and wet cases with using a water surplus or deficit D (=P-PET in Eq. 1). We compare Ds between the reference period and estimation period. A value of D in the estimation period less than that in the reference period represents the dry case, i.e., the estimation period is drier than the reference period. We perform such analyses only in Ref1 for estimation periods of 1901-1957 (P1) and 1958-2014 (P2) and a reference/calibration period from 1901-2014 (P1+P2). For dry and wet cases, we quantify the spatial extent (%) according to the three different drought levels (D1, D2 and D3, which denote the cases of SPEI < -1.0, SPEI < -2.0 and SPEI < -3.0, respectively) in the four regions.

As presented in Table 5, the average D in P1 or P2 (estimation period) is smaller than that in P1+P2 (reference period), and it is considered to be the dry case. For example, in EA, the Ds values in P2 and P1+P2 are -4.89 mm/month and -5.07 mm/month, respectively; thus, it is a dry case. Then, for each case, the drought spatial extents, the number of drought grid cells relative to the total number of effective grid cell, are analyzed as shown in Fig. 11. The drought spatial extent tends to be larger in the dry case than that in the wet case in most regions, particularly in West Africa. However, we also note there are a few exceptions, which may be attributed to the fact that we use the regional average Ds. Thus, we cannot consider the grid-level variability in Ds.

3.4 Case studies with historical drought events

SPEI-12s with different reference periods are evaluated for historical drought events selected in each region to investigate how different reference periods influence the drought assessments of historical events. One drought event is chosen for each region as follows: 1) For East Asia, droughts that occurred in northern China in 2001 are chosen. These events caused economic losses of USD 1.52 billion (Zhang and Zhou, 2015). 2) For EU, we chose a 2003 drought that was caused by the European heat wave and spread over the majority of Europe (Stagge et al., 2013; Spinoni et al., 2015). 3) For US, we chose 2012 as the period of study as drought in that year was the most extensive drought over half of the US since the 1930s and caused economic losses of USD 31.2 billion (Smith and Katz, 2013; National Climate Data Center, 2015). 4) For West Africa, the drought in 1984 is chosen because it is one of severest droughts that has occurred over most Sahel countries (Gommes and Petrassi, 1994; Rojas et al., 2011; Masih et al., 2014).

By estimating SPEI-12 for a chosen year in each region, we compare the magnitudes of SPEIs (Figs. 12, 13, 14 and 15). Here the annual SPEI-12s with the monthly climate data from January to December in each year are first constructed and then the SPEI-12s for a chosen year are examined in detail. All SPEI-12s with the different reference periods present the drought status because we chose specific years with drought events. In general, all cases reveal that the SPEI-12 estimates in Ref2 are relatively high (i.e., wet) and those in Ref3 are relatively low (i.e., dry) for EA and WA, where drying temporal trends are clear. In particular, the several extremes (i.e., out of the scale ranges in Figs. 12-15) of SPEI-12 in Ref3 cases highlight the importance of the reference period. By using the reference period of the certain past time (P1 in this study, i.e., Ref3), the drought events in the estimation period could be beyond the range in which the distribution is calibrated for the index. Essentially, using Ref3, it is assumed that not only the stationarity of the climate but also that the whole probability distribution of droughts is sampled in this period.

Furthermore, the percentages of drought spatial extent, i.e., the number of drought grid points relative to the total grid points, are assessed with different drought thresholds (Table 6). In most cases, the spatial extents of drought with the SPEI less than certain thresholds, such as -1, -2 or -3 (i.e., D1, D2 and D3 as in Table 1) are the greatest in Ref3 among the three cases with different reference periods. These results with the spatial extent are consistent with the results with the SPEI-12 estimates above. In addition, for the severe droughts with the drought events, defined with low thresholds such as SPEI-12 less than -2 or -3, greater percentages in Ref3 than in Ref1 and Ref2 are consistently obtained without exception in all regions of EA, EU, US and WA.

4 Conclusions

This study seeks to understand how a different reference period (i.e., calibration period) of climate data for estimating the drought index would influence the regional drought assessment. Specifically, we investigate the influence of different reference periods on historical drought characteristics such as trends, frequency, intensity and spatial extents using SPEI-12 from the CRU and UDEL datasets. For the 1901–1957 (P1) and 1958–2014 (P2) estimation periods, three different types of reference periods are used. For the first case, the data from 1901 to 2014 (P1+P2) are used for both estimation periods; for the second case, the data from P1 and P2 are used separately for the estimation periods of P1 and P2, respectively (self-calibrated); and for the final case, the data from P1 (1910–1957) are used for both estimation periods.

Focusing on the EA, EU, US and WA regions, we find the influence of the reference periods is significant in the regions with dominant drying trends from P1 to P2, such as EA and WA. It is also suggested that it is necessary to quantify the trends of climate variables such as precipitation and air temperature as the first step in selecting a reference period. We find that the reference periods influence the assessment of drought characteristics, particularly for severity and spatial extent, based on two datasets; however, their influence on the frequency is relatively small. Finally, self-calibration, the most common practice with an index such as SPEI, tends to underestimate the drought severity and spatial extent relative to the other approaches examined in this study.

This study highlights the need for the reference period to be clarified in drought assessments for a better understanding of regional drought characteristics and their temporal changes, particularly under climate change scenarios. Although this study with historical data may shows the different results depending on the selected local area, a similar study with historical data or climate change scenarios in different regions would undoubtedly strengthen our findings. We note that this study focuses on the temporal aspects of calibration data (i.e., calibration period). As briefly mentioned in the Introduction, using data from a particular station or grid, the averaged data for calibration would permit a meaningful comparison of the drought index among different locations. In conjunction with temporal considerations, such spatial issues could readily be addressed in future studies.

Acknowledgements

This study was supported by the Korea Meteorological Administration R&D Program under Grant KMIPA 2015-6180 and by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning (2015R1C1A2A01054800).

References

Abramowitz, M. and Stegun, I. A.: Handbook of Mathematical Functions. National Bureau of Standards, 1966.

Dai, A.: Drought under global warming: a review, Wiley Interdisciplinary Reviews: Climate Change, 2, 45–65, 2011.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth Syst. Sci. Data, 5, 71–99, 2013.

Dai, A.: Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008, J. Geophys. Res., 116, D12115, 2011.

Dracup, J. A., Lee, K. S., and Paulson Jr, E. G.: On the definition of droughts, Water Resour. Res., 16, 297-302, 1980.

Dubrovsky, M., Svoboda, M. D., Trnka, M., Hayes, M. J., Wilhite, D. A., Zalud, Z., and Hlavinka, P.: Application of relative drought indices in assessing climate-change impacts on drought conditions in Czechia, Theor. Appl. Climatol., 96, 155–171, 2009.

Gommes, R. and Petrassi, F., 1994: Rainfall Variability and Drought in Sub-Saharan Africa sicn 1960. FAO Agrometeorology: Series Working Papers 9: 100pp.

Hagman, G.: Prevention better than cure: report on human and natural disasters in the third world, Swedish Red Cross, Stockholm, 1984.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset, Int. J. Climatol., 34, 623–642, 2014.

Heim, R. R.: A review of twentieth-century drought indices used in the United States, Bull. Amer. Meteor. Soc., 83, 1149–1165, 2002.

Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T., and Pegion, P.: On the increased frequency of Mediterranean drought, J. Climate, 25, 2146-2161, 2012.

Hosking, J. R. M.: L-Moments: analysis and estimation of Distributions using linear combinations of order statistics, J. R. Stat. Soc. B, 52, 105-124, 1990.

Ionita, M., Scholz, P., and Chelcea, S.: Spatio-temporal variability of dryness/wetness in the Danube river basin, Hydrological Processes, 29, 4483-4497, 2015.

Karl, T. R., Knight, R. W., Easterling, D. R., and Quayle, R. G.: Indices of climate change for the United States. Bull. Amer. Meteor. Soc.77, 279-292, 1996.

Kendall, M. G.: Rank Correlation Methods. Griffin, London, 1957.

Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, Int. J. Climatol., 22, 1571-1592, 2002.

Mann, H. B.: Nonparametric tests against trend, Econometrica, 13, 245-259, 1945.

Masih, I., Maskey, S., Mussá, F. E. F., and Trambauer, P.: A review of droughts on the African continent: A geospatial and long-term perspective, Hydrol. Earth Syst. Sci., 18, 3635-3649, 2014.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to Time scales: Eighth Conference on Applied Climatology, in: January 17-22. Anaheim, CA: 179–184, 1993.

Nasrollahi, N., AghaKouchak, A., Cheng, L., Damberg, L., Phillips, T. J., Miao, C., Hsu, K., and Sorooshian, S.: How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts?. Water Resour. Res., 51, 2847-2864, 2015.

Naumann, G., Dutra, E., Barbosa, P., Pappenberger, F., Wetterhall, F., and Vogt, J. V.: Comparison of drought indicators derived from multiple data sets over Africa, Hydrol. Earth Syst. Sci., 18, 1625-1640, 2014.

National Climate Data Center, 2014: Billion-dollar weather and climate disasters. http://www.ncdc.noaa.gov/billions.

Palmer, W.: Meteorological drought, Res. Pap., 45, 1965, U.S. Department of Commerce Weather Bureau: 58 p. Available online by the NOAA National Climatic Data Center at http://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf.

Rojas, O., Vrieling, A., and Rembold, F.: Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery, Remote Sens. Environ., 115, 343-352, 2011.

Seneviratne, S. I.: Historical drought trends revisited, Nature, 491, 339, 2012.

Sheffield, J. and Wood, E. F.: Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations, Clim. Dyn., 31, 79–105, 2008.

Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years, Nature, 491, 435-440, 2012.

Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data: Proceedings, 1968 ACM National Conference, in: 517-523, 1968.

Smith, A. B. and Katz, R. W.: US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases, Nat. Hazards, 67, 387–410, 2013.

Spinoni, J., Naumann, G., Vogt, J. V., and Barbosa, P.: The biggest drought events in Europe from 1950-2012, J. Hydrol.: Reg. Stud. 3, 509-524, 2015.

Stagge, J. H., Tallaksen, L. M., Kohn, I., Stahl, K., and van Loon, A. F.: A European drought reference (EDR) database: design and online implementation. Drought-R&SPI deliverable, D1.1. Technical report no. 12, 2013.

Tan, C., Yang, J., and Li, M.: Temporal-spatial variation of drought indicated by SPI and SPEI in Ningxia Hui Autonomous Region, China, Atmosphere, 6, 1399-1421, 2015.

Thornthwaite, C. W.: An approach toward a rational classification of climate, Geogr. Rev., 38, 55-94, 1948.

Touma, D., Ashfaq, M., Nayak, M. A., Kao, S., and Diffenbaugh, N. S.: A multi-model and multi-index evaluation of drought characteristics in the 21st century, J. Hydrol., 526, 196-207, 2015.

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J.: Global warming and changes in drought, Nat. Clim. Change, 4, 17-22, 2014.

Van der Schrier, G., Barichivich, J., Briffa, K. R., Jones, P. D.: A scPDSI-based global data set of dry and wet spells for 1901-2009, J. Geophys. Res., 118(10), 4025-4048, 2013.

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

Vicente-Serrano, S. M. and Beguería-Portugués, S.: Estimating extreme dry-spell risk in the middle Ebro valley (northeastern Spain): a comparative analysis of partial duration series with a general Pareto distribution and annual maxima series with a Gumbel distribution, Int. J. Climatol., 23, 1103–1118, 2003.

Vittal, H., Karmakar, S., and Ghosh, S.: Diametric changes in trends and patterns of extreme rainfall over India from pre-1950 to post-1950, Geophys. Res. Lett., 40, 3253–3258, 2013, doi:10.1002/grl.50631.

Wang, D., Hejazi, M., Cai, X., and Valocchi, A. J.: Climate change impact on meteorological, agricultural, and hydrological drought in central Illinois, Water Resour. Res., 47, 2011. doi: 10.1029/2010WR009845.

Wilhite, D. A.: Reducing drought vulnerability through mitigation and preparedness, report to the inter-Agency Task Force For Disaster Reduction Sixth Meeting, Geneva. National Drought Mitigation Center, Lincoln, 2002.

Willmott, C. J. and Matsuura, K., 2001: Terrestrial Air temperature and precipitation: monthly and annual Time series (1950-1999). http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html.

Willmott, C. J., Rowe, C. M., and Philpot, W. D.: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring, Am. Cartographer, 12, 5-16, 1985.

Zhang, L. and Zhou, T.: Drought over East Asia: a review, J. Climate, 28, 3375-3399, 2015.

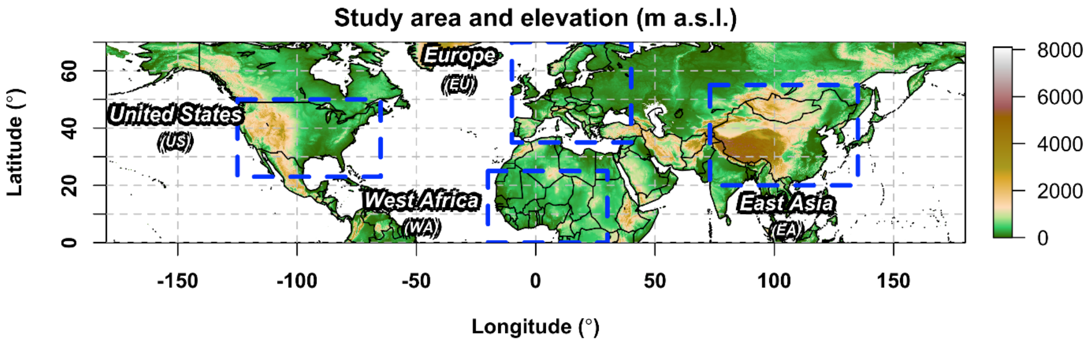

Figure 1. Study area and elevation investigated in this work.

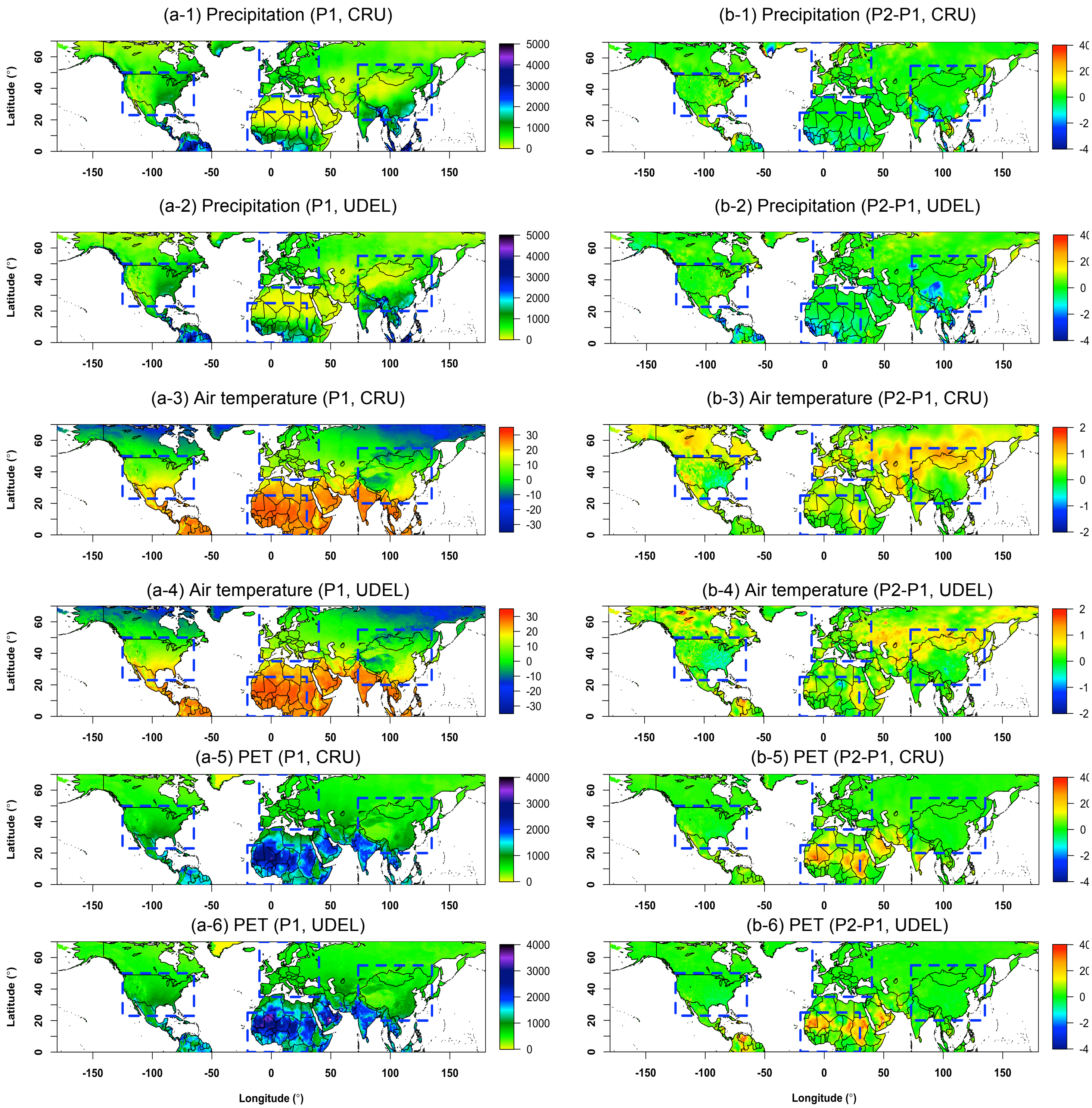

Figure 2. Annual precipitation (mm), annual averaged temperature (°C) and annual PET (mm) for the CRU and UDEL datasets for P1 and their difference between P1 and P2.

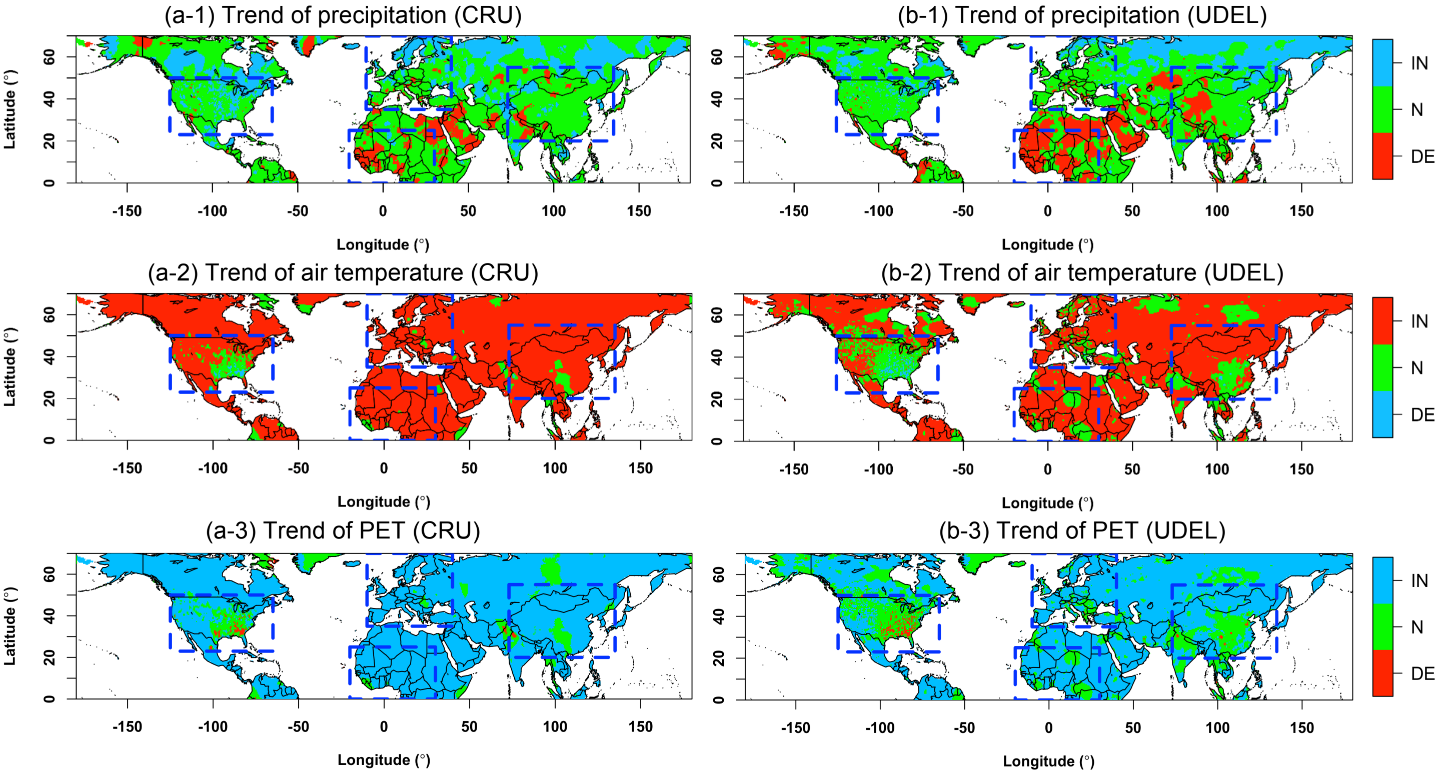

Figure 3. Trends of annual precipitation, annual averaged temperature and annual PET for the CRU and UDEL datasets. PR and TA denote precipitation and temperature, respectively, and IN, N and DE indicate increasing, no trend and decreasing, respectively.

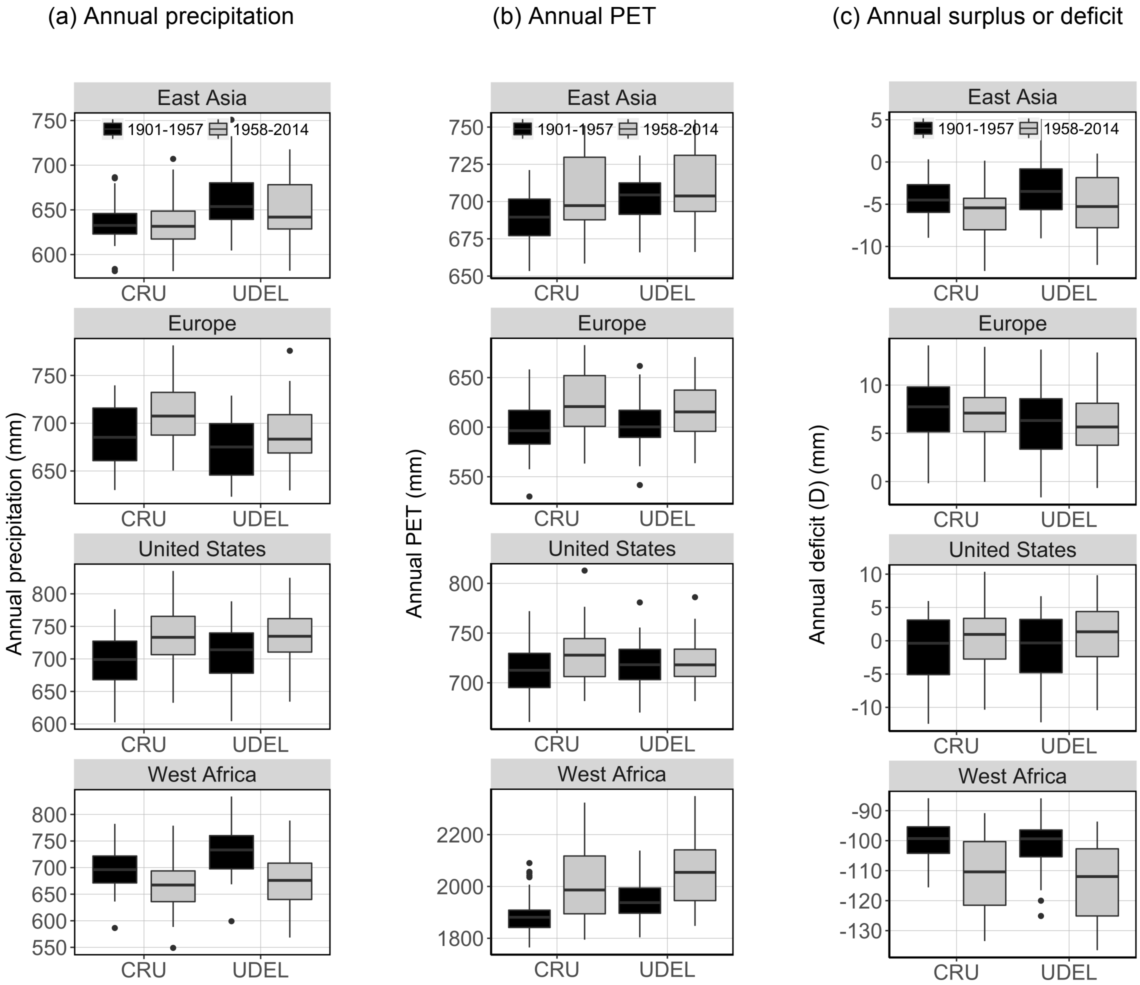

Figure 4. Temporal variations of annual precipitation, PET and deficit (D) depending on two datasets (CRU and UDEL) and periods (1901-1957 and 1958-2014). In the box plot, the center line represents the median value; the top and bottom of box represent the 25th and 75th percentile of the data, respectively; the dot represents the outlier.

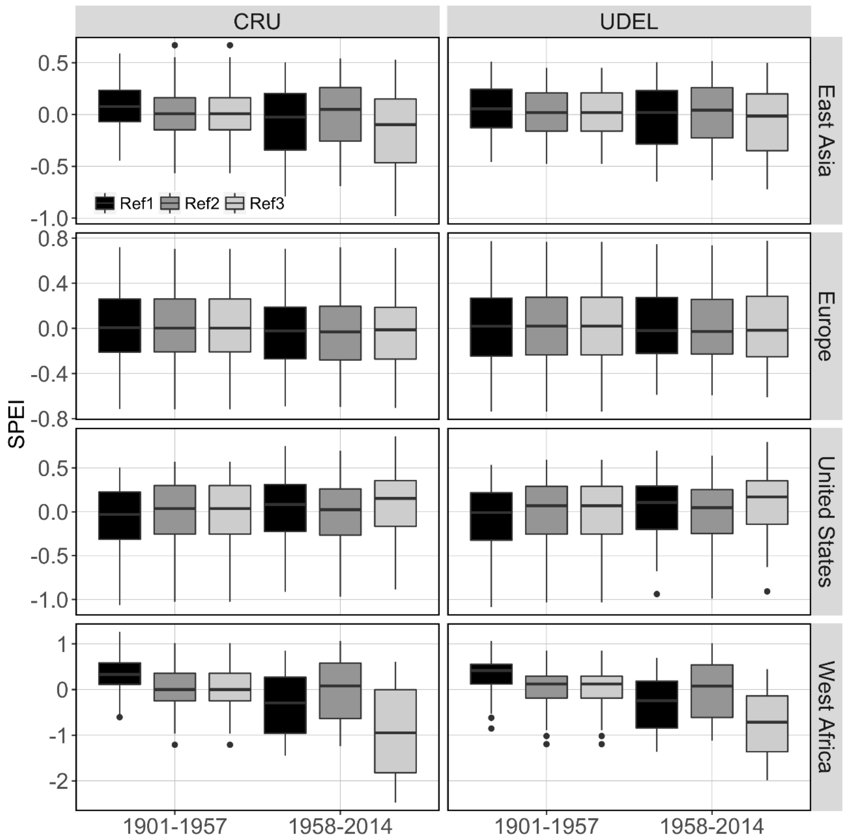

Figure 5. Temporal variations of SPEI with 12-month lag for three different reference periods (Ref1, Ref2 and Ref3) for the CRU and UDEL datasets and the periods 1901–1957 and 1958–2014. In the box plot, the center line represents the median value; the top and bottom of box represent the 25th and 75th percentile of the data, respectively; the dot represents the outlier.

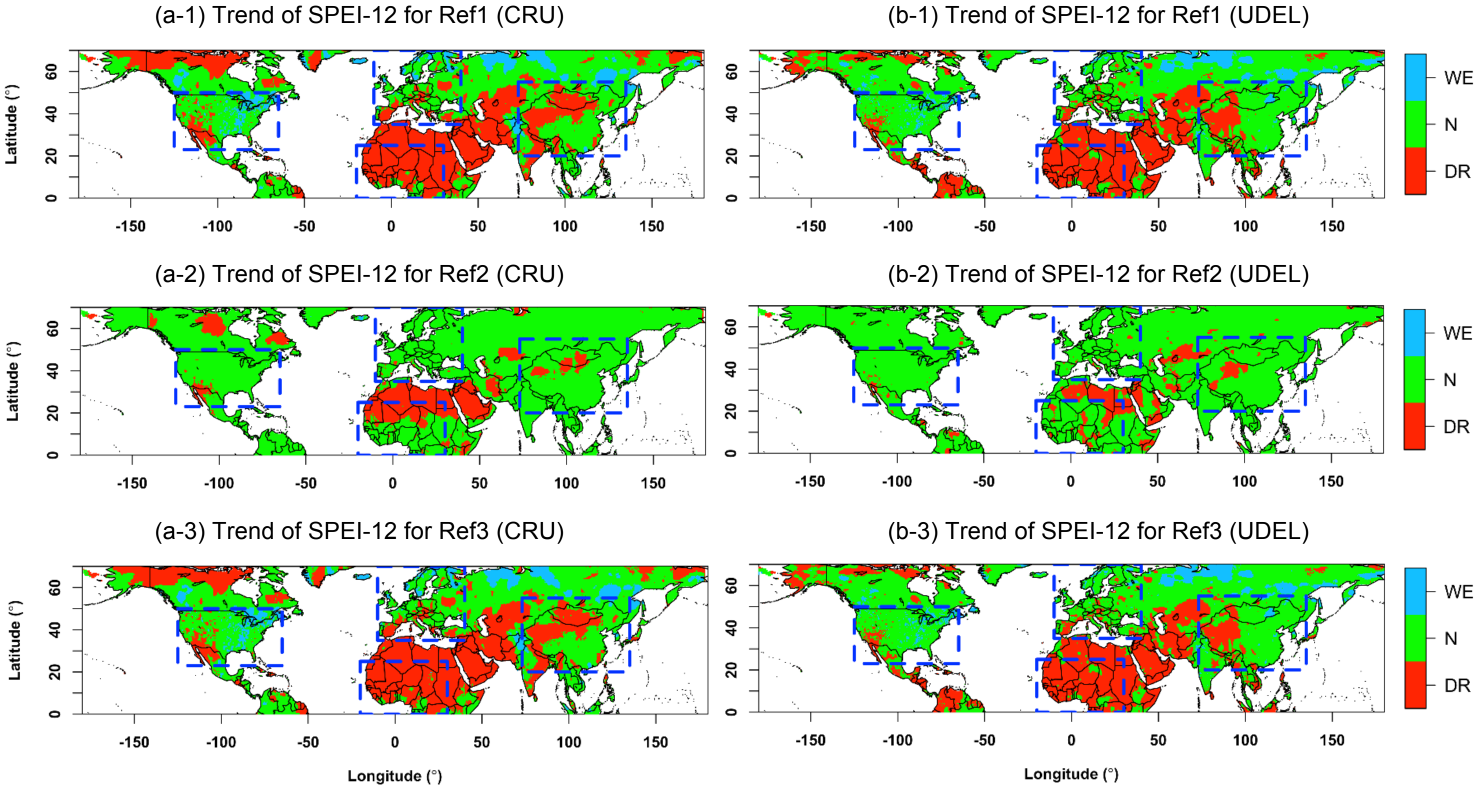

Figure 6. Trend of SPEI with 12-month lags (SPEI-12) for three different reference periods (Ref1, Ref2 and Ref3) for the (a) CRU and (b) UDEL datasets. WE, N and DR denote wetting, no trend and drying, respectively.

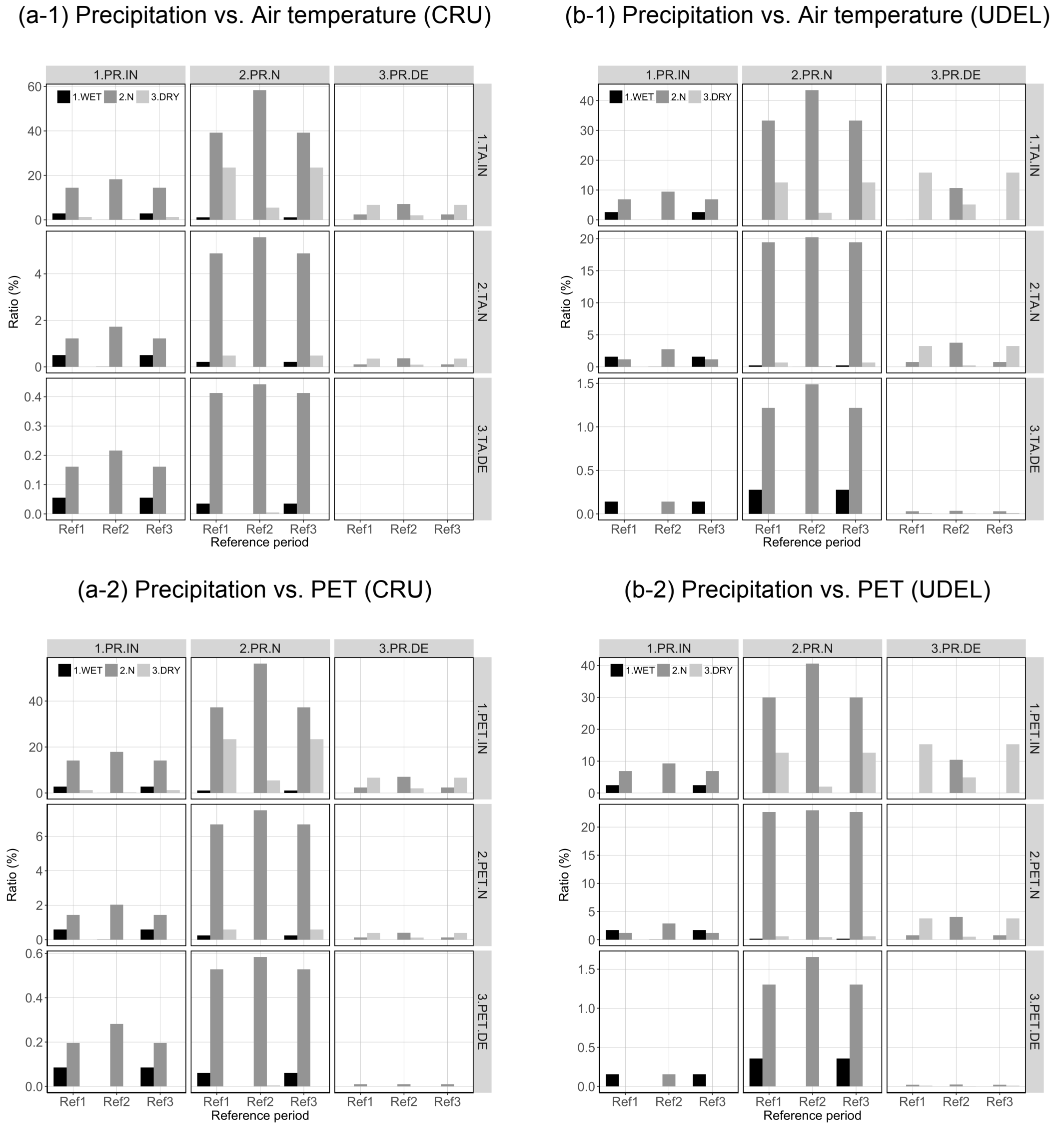

Figure 7. The SPEI trends with 12-month lag (SPEI12) for three different reference periods (Ref1 to Ref3) for the CRU and UDEL datasets based on the trends of monthly precipitation and temperature (or PET) in the four zones

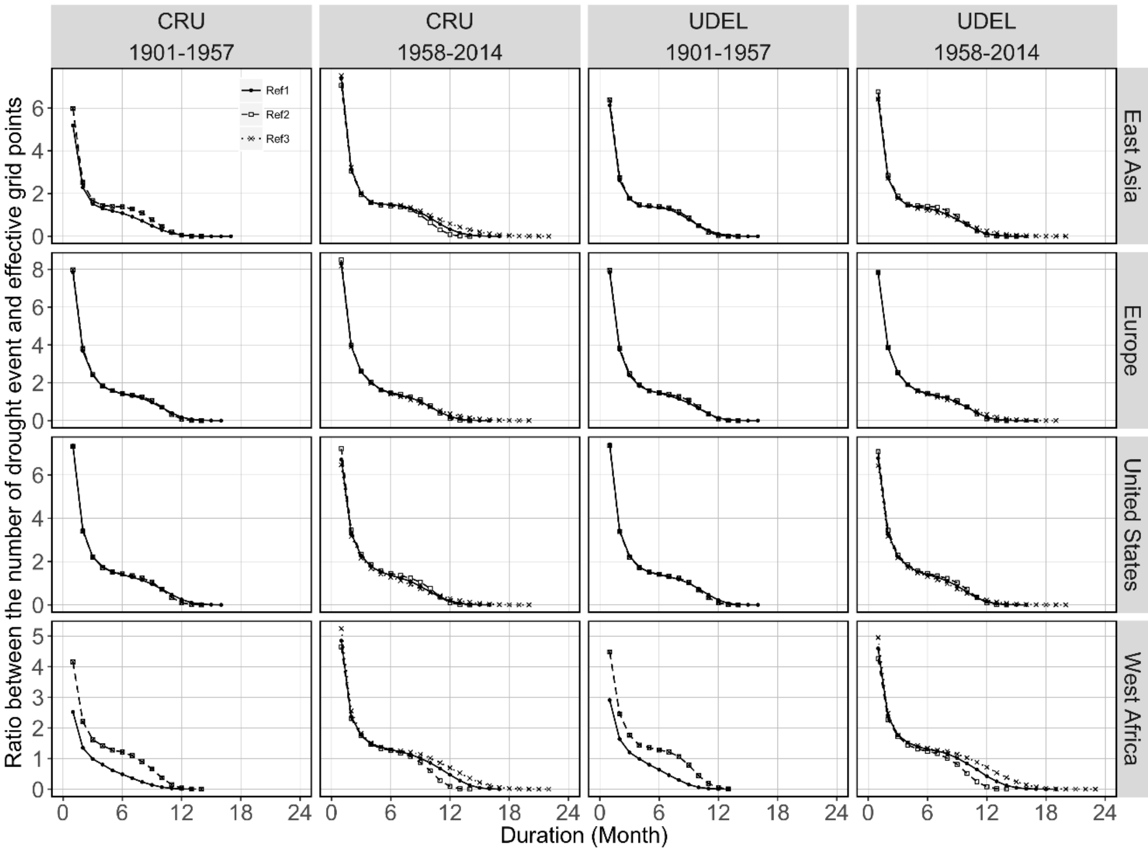

Figure 8. Ratio of the number of drought events and effective data grid points for the CRU and UDEL datasets and the periods 1901–1957 and 1958–2014.

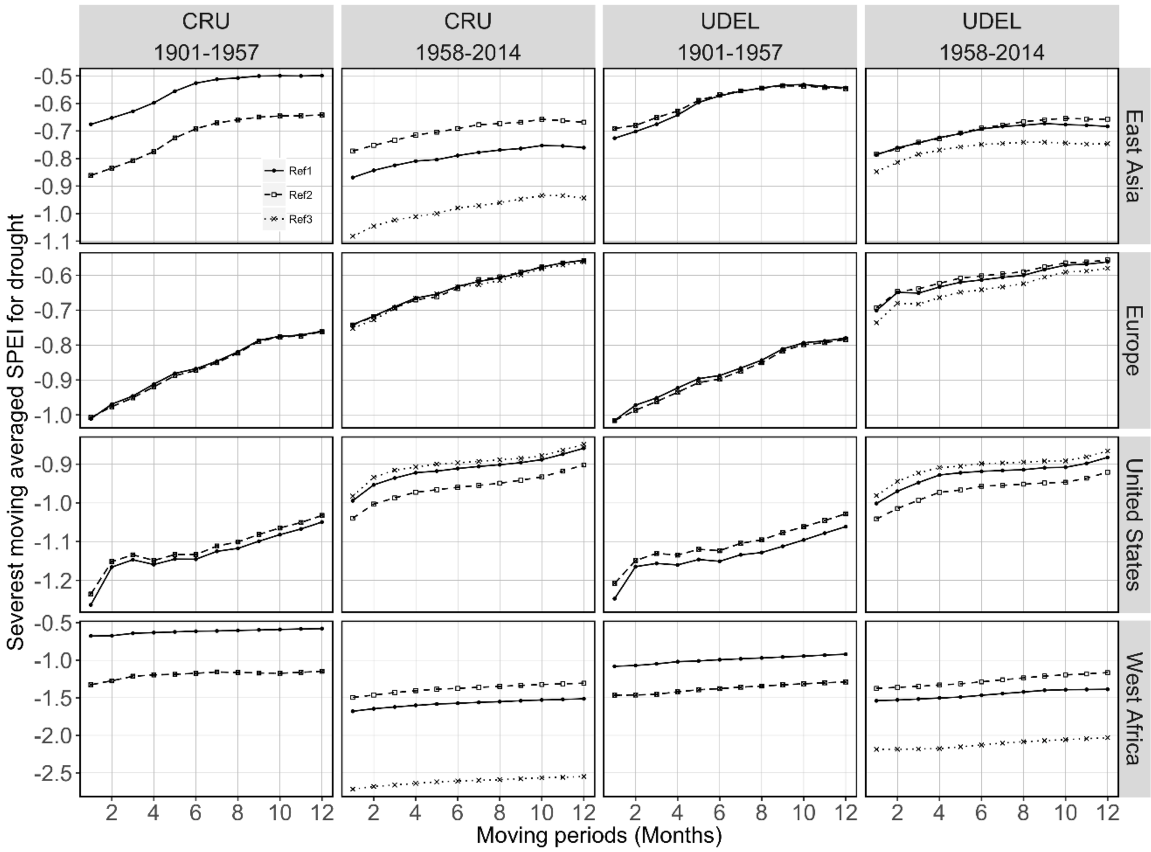

Figure 9. Severest moving average of regional average SPEI for 1–12 months for three different reference periods (Ref1, Ref2 and Ref3) for the CRU and UDEL datasets and the periods 1901–1957 and 1958–2014.

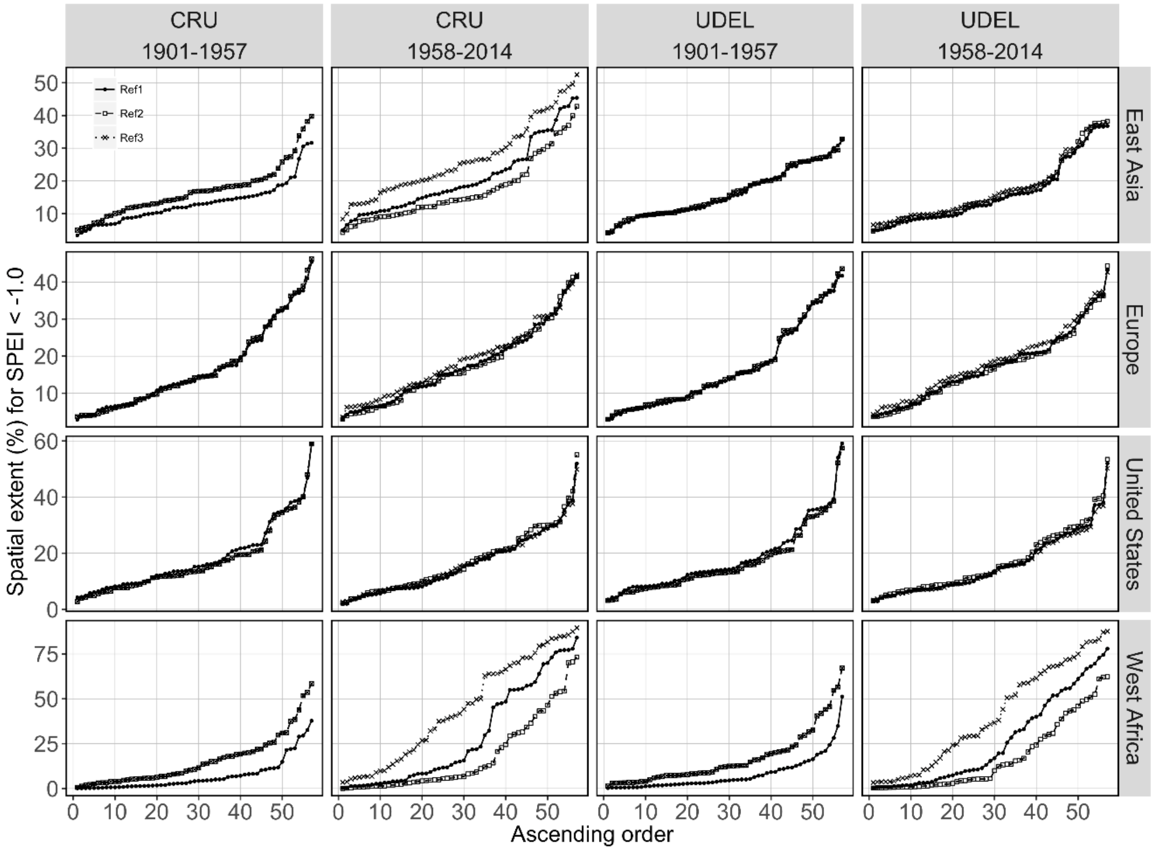

Figure 10. Spatial extent (%) for SPEI with 12-month lag < -1.0 for three different reference periods (Ref1, Ref2 and Ref3) for the CRU and UDEL datasets and the periods 1901–1957 and 1958–2014.

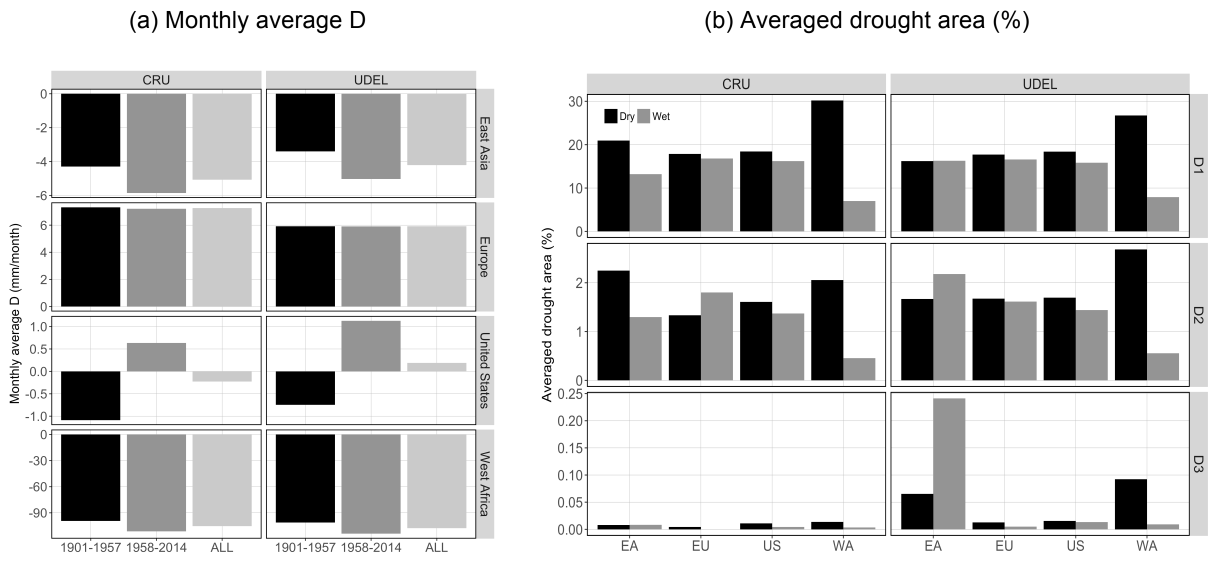

Figure 11. Monthly average D in Eq. (1) and averaged drought area depending on two datasets (CRU and UDEL) and four zones (EA, EU, US and WA) with the Ref1 condition (In Fig. 10 (a), all denotes the period for 1901-2014. In Fig. 10 (b), the dry status means that monthly average D in assessment period is less than those in reference period and the wet status denotes that monthly average D in assessment period is greater than those in reference period in the Ref1 condition.

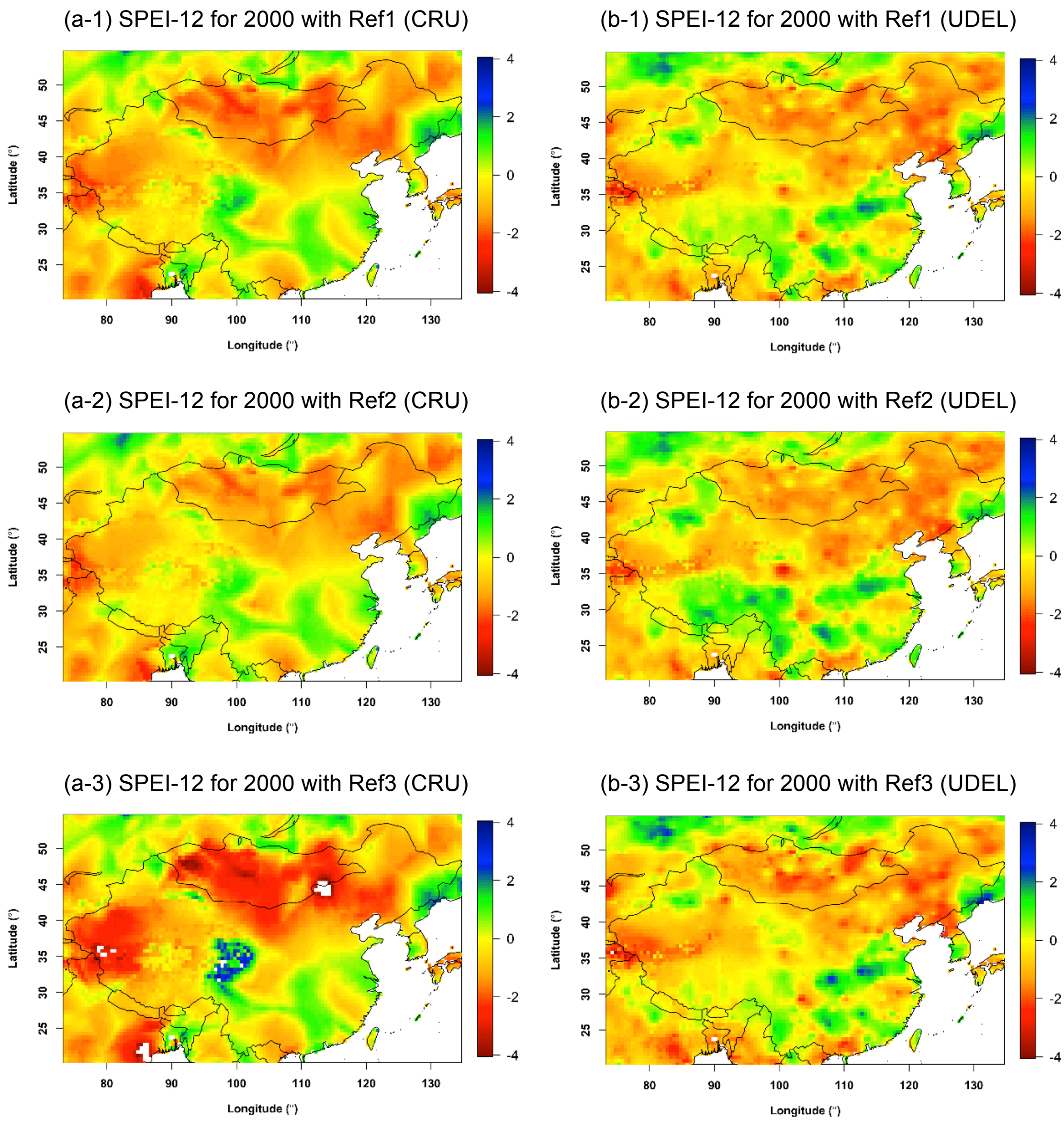

Figure 12. SPEI with 12-month lag (SPEI12) for three different reference periods (Ref1, Ref2 and Ref3) for the (a) CRU and (b) UDEL datasets in East Asia in 2000.

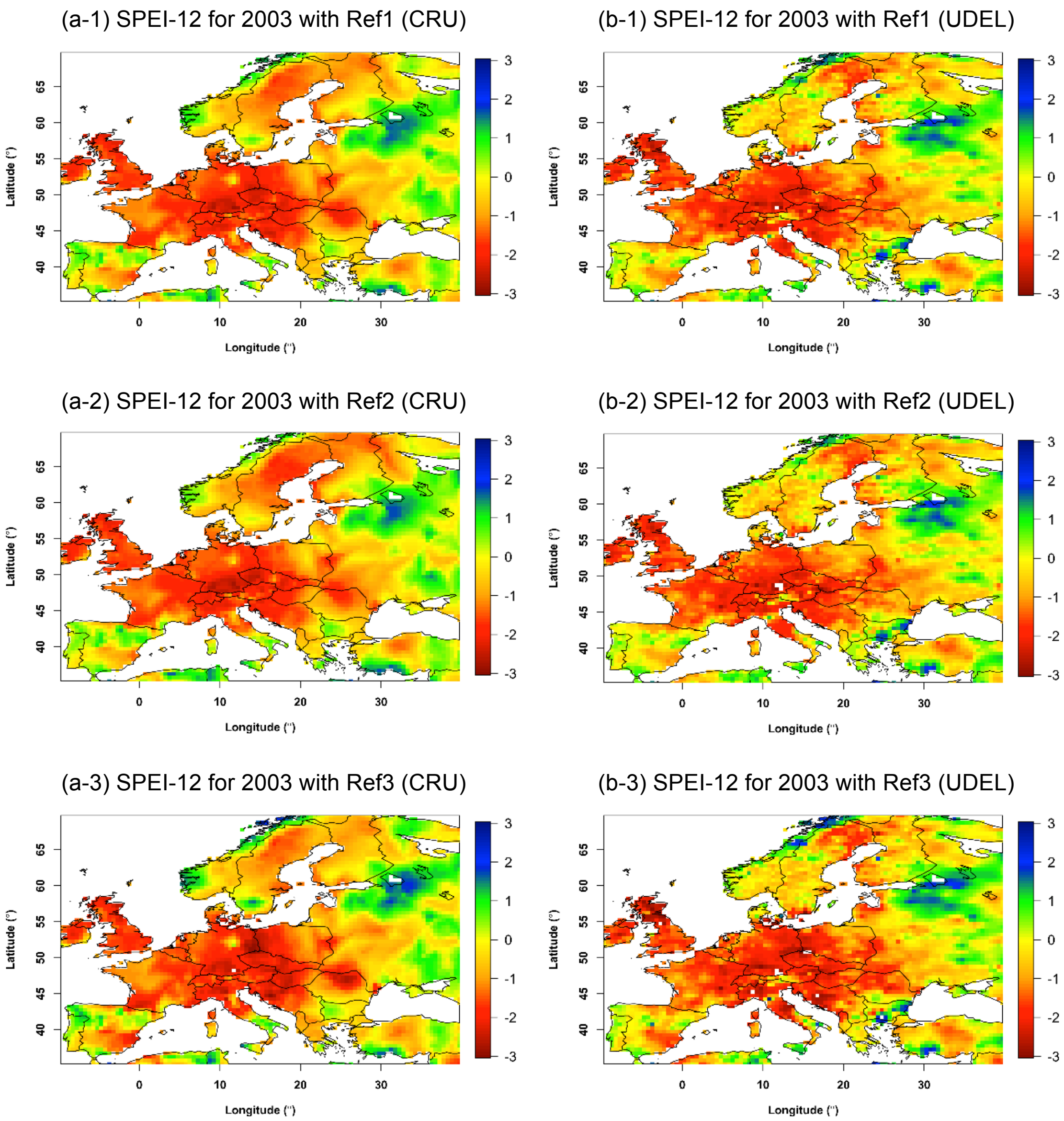

Figure 13. SPEI with 12-month lag (SPEI12) for three different reference periods (Ref1, Ref2 and Ref3) for the (a) CRU and (b) UDEL datasets in Europe in 2003.

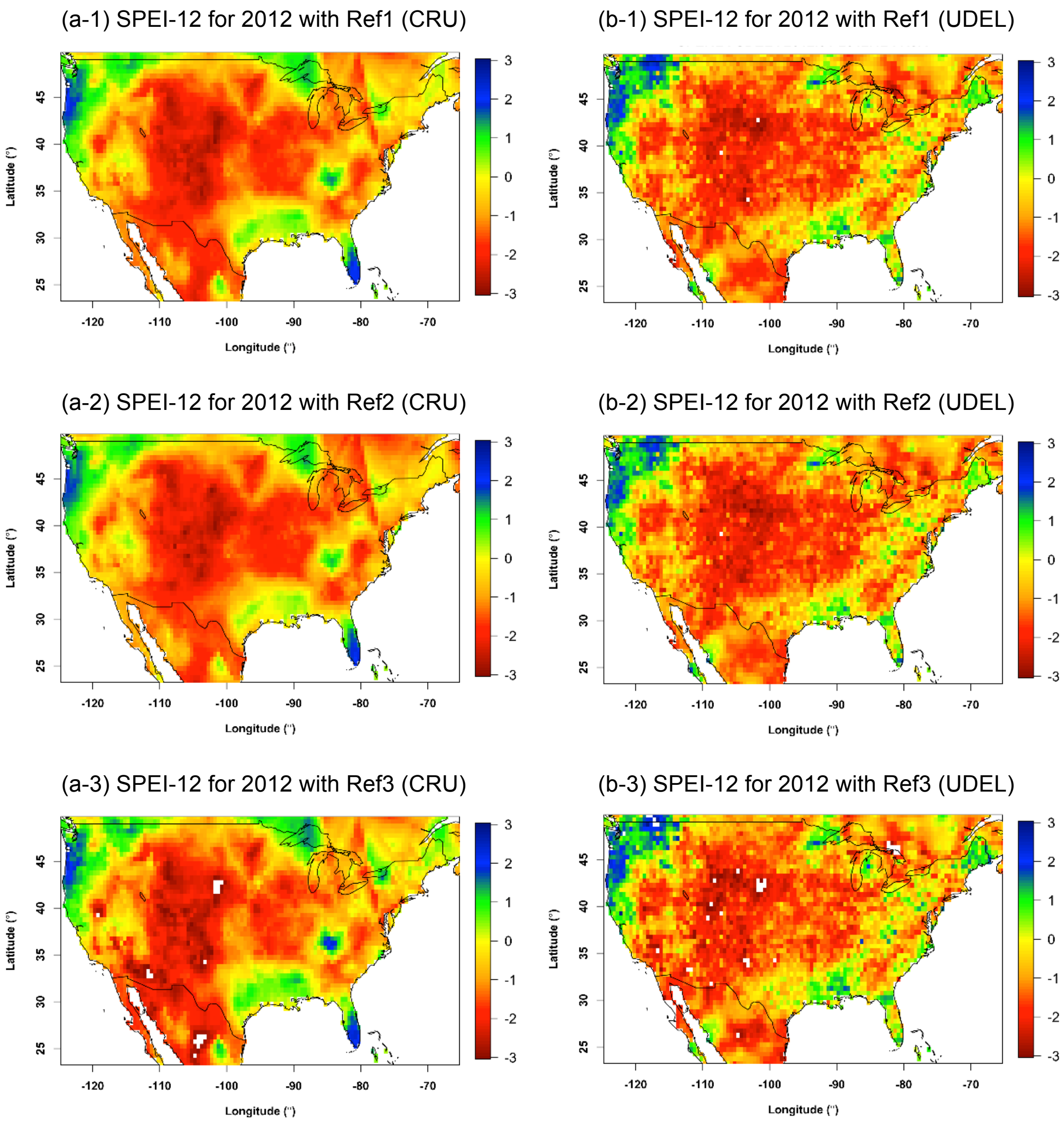

Figure 14. SPEI with 12-month lag (SPEI12) for three different reference periods (Ref1, Ref2 and Ref3) for the (a) CRU and (b) UDEL datasets in the United States in 2012.

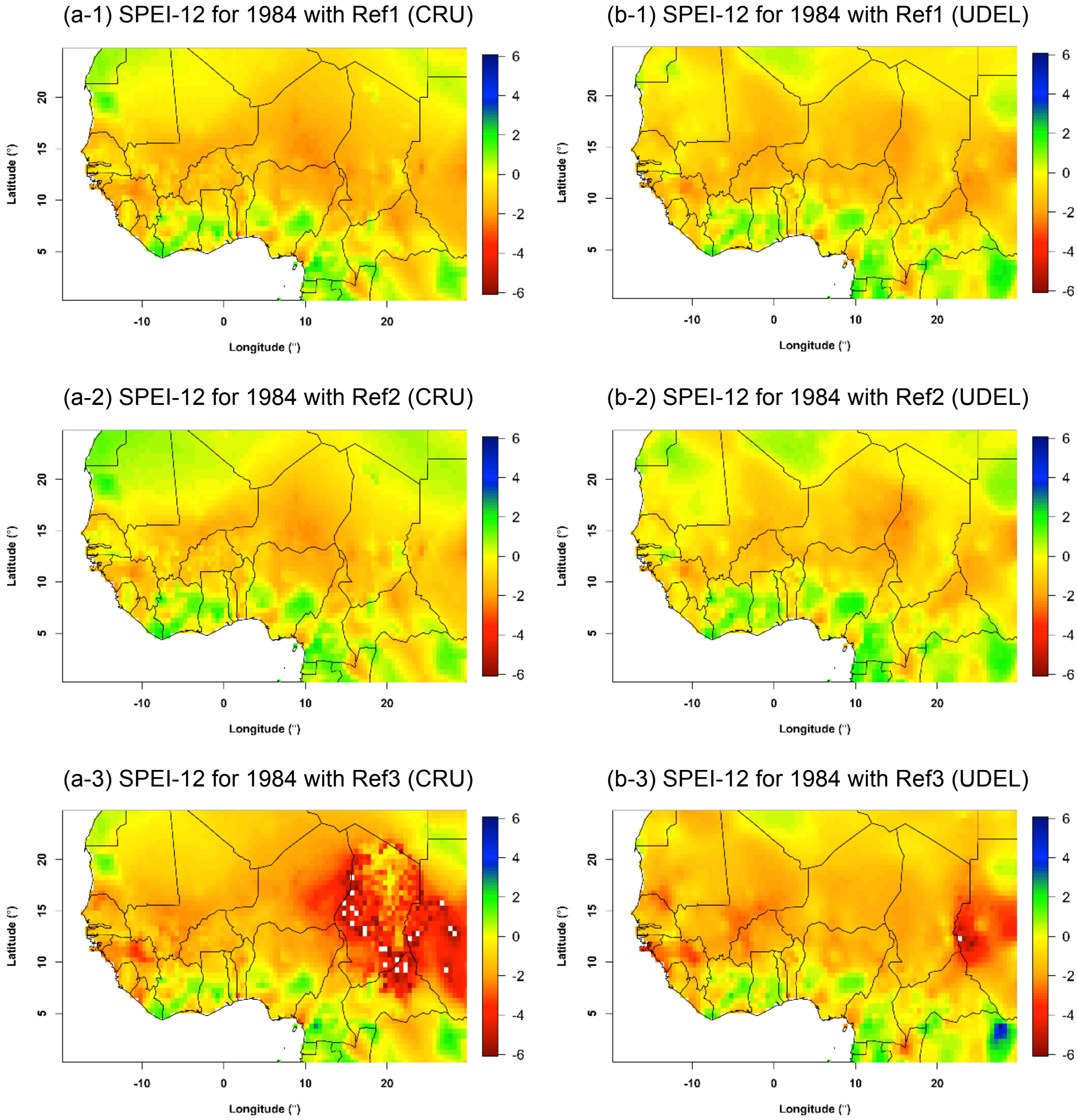

Figure 15. SPEI with 12-month lag (SPEI12) for three different reference periods (Ref1, Ref2 and Ref3) for the (a) CRU and (b) UDEL datasets in West Africa in 1984.

Table 1. Classification of dry status in this study (Mckee et al., 1993).

Category	Description	SPEI
D1	Moderate dry	≤ -1.0
D2	Extreme dry	≤ -2.0
D3	Very extreme dry	≤ -3.0

Table 2. Climate variables and conditions for SPEI.

Type	Estimation Period	Calibration Period
Ref1	1901–1957	1901–2014
	1958–2014	
Ref2	1901–1957	1901–1957
	1958–2014	1958–2014
Ref3	1901–1957	1901–1957
	1958–2014	

Table 3. Mean and standard deviation (STD) of precipitation and air temperature over the regions.

			CRU		UDEL	
			1901–1957	1958–2014	1901–1957	1958–2014
Annual Precipitation (mm)	EA	Mean	637.19	635.52	659.67	649.21
		STD	22.36	30.05	30.67	31.76
	EU	Mean	685.86	711.03	674.17	688.31
		STD	31.08	32.43	30.97	31.16
	US	Mean	698.44	736.22	709.50	734.42
		STD	43.31	41.48	44.06	41.55
	WA	Mean	698.49	666.59	734.84	676.11
		STD	36.87	43.84	44.89	48.00
Air Temperature (°C)	EA	Mean	6.08	6.67	6.25	6.62
		STD	0.28	0.52	0.31	0.48
	EU	Mean	6.96	7.46	7.02	7.29
		STD	0.56	0.68	0.55	0.64
	US	Mean	10.46	10.78	10.59	10.64
		STD	0.45	0.50	0.43	0.43
	WA	Mean	26.27	26.62	26.40	26.66
		STD	0.25	0.48	0.26	0.41
Annual potential evapotranspiration (mm)	EA	Mean	688.69	705.78	700.44	709.52
		STD	16.06	23.92	15.65	22.73
	EU	Mean	598.06	624.48	603.15	617.35
		STD	24.86	31.66	23.46	28.54
	US	Mean	711.49	728.60	718.48	720.90
		STD	23.84	26.28	22.59	21.28
	WA	Mean	1889.57	2001.37	1948.91	2044.28
		STD	72.61	136.96	78.17	129.94

Table 4. Spatial extent (%), the number of grid points relative to the total effective grid point in each region for different trends with the different reference periods of SPEI-12.

Zone	CRU						UDEL					
	Ref1		Ref2		Ref3		Ref1		Ref2		Ref3	
	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry
EA	2.5	36.3	0.0	8.0	2.5	36.5	3.4	23.2	0.0	7.7	3.4	23.2
EU	10.4	24.9	0.0	1.7	10.4	24.9	5.3	15.8	0.0	2.2	5.3	15.8
US	18.6	16.2	0.0	67	18.6	16.2	11.3	9.7	0.1	3.1	11.3	9.7
WA	0.0	90.2	0.0	40.4	0.1	89.8	0.0	90.9	0.0	19.5	0.0	90.9

Table 5. Monthly average D (mm/month) in four study regions.

	CRU			UDEL		
	P1	P2	P1+P2	P1	P2	P1+P2
EA	-4.29	-5.85	-5.07	-3.40	-5.03	-4.21
EU	7.32	7.21	7.26	5.92	5.91	5.92
US	-1.09	0.64	-0.23	-0.75	1.13	0.19
WA	-99.26	-111.23	-105.24	-101.17	-114.01	-107.59

Table 6. Spatial extent (%) (the number of grid points belonging to each drought category, relative to the total grid point) for the major drought events.

Zone	Period	Type	CRU			UDEL		
			Ref1	Ref2	Ref3	Ref1	Ref2	Ref3
EA	2000	D1	32.63	27.48	38.80	26.81	27.39	29.62
		D2	2.45	0.75	14.64	0.92	0.73	2.64
		D3	0.05	0.00	1.83	0.04	0.01	0.07
EU	2003	D1	37.58	39.10	36.68	35.30	34.67	36.61
		D2	5.33	3.97	7.68	5.93	4.82	8.50
		D3	0.00	0.00	0.02	0.02	0.10	0.22
US	2012	D1	52.16	55.01	50.02	54.69	56.32	52.92
		D2	11.97	11.90	15.74	10.36	11.76	11.63
		D3	0.02	0.00	0.53	0.09	0.05	0.87
WA	1984	D1	44.06	31.04	62.18	37.13	27.15	57.78
		D2	3.42	1.87	28.62	2.07	1.72	13.80
		D3	0.00	0.00	14.30	0.00	0.00	2.99