A baseline probabilistic drought forecasting framework using
standardized soil moisture index: application to the 2012 United
States drought

A. AghaKouchak
Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering,
University of California, Irvine, CA, USA

Correspondence to: A. AghaKouchak (amir.a@uci.edu)

Received: 28 January 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 11 February 2014
Revised: 9 May 2014 – Accepted: 21 May 2014 – Published: 4 July 2014

Abstract. The 2012 drought was one of the most extensive
drought events in half a century, resulting in over USD 12
billion in economic loss in the United States and substan-
tial indirect impacts on global food security and commod-
ity prices. An important feature of the 2012 drought was
rapid development and intensification in late spring/early
summer, a critical time for crop development and invest-
ment planning. Drought prediction remains a major chal-
lenge because dynamical precipitation forecasts are highly
uncertain, and their prediction skill is low. Using a prob-
abilistic framework for drought forecasting based on the
persistence property of accumulated soil moisture, this pa-
per shows that the US drought of summer 2012 was pre-
dictable several months in advance. The presented drought
forecasting framework provides the probability occurrence
of drought based on climatology and near-past observations
of soil moisture. The results indicate that soil moisture ex-
hibits higher persistence than precipitation, and hence im-
proves drought predictability.

1 Introduction

According to United States Department of Agriculture
(USDA) estimates, about 80 percent of US agricultural land
experienced drought in 2012, which made the event more ex-
tensive than any since 1950 (USDA, 2012). A striking as-
pect of the 2012 drought was rapid increase in severity in
early July during a critical time of crop development (USDA,
2012). The quick onset of the drought in the central plains
during late spring led to a so-called “flash drought” (Ho-
erling et al., 2013, 2014). A drought early warning system
with seasonal predictions of drought onset, severity, persist-
tence, and spatial extent in a timely manner would provide
invaluable information to decision-makers and stakeholders.
There are a number of research and operational drought (or
hydrologic) prediction systems (Pozzi et al., 2013; Mishra
and Singh, 2010; AghaKouchak and Nakhjiri, 2012), includ-
ing the Climate Prediction Center Seasonal Drought Out-
look (Steinemann, 2006), the University of Washington’s
Surface Water Monitor (Wood and Lettenmaier, 2006; Wood,
2008), Princeton University’s drought forecast system (Luo
and Wood, 2007; Li et al., 2008; Sheffield et al., 2008), US–
Mexico Drought Prediction Tool (Lyon et al., 2012), and the
Global Integrated Drought Monitoring and Prediction Sys-
tem (GIDMaPS; Hao et al., 2014). Despite all these efforts,
a community white paper by the World Climate Research
Program identified sub-seasonal to seasonal drought predic-
tion as one of the major research gaps in hydroclimatology
(WCRP, 2010).

Drought forecasting is generally based on drought indi-
cators computed using dynamic or statistical model simul-
ations of drought-related variables (e.g., Mishra et al., 2009;
Madadgar, and Moradkhani, 2013). Droughts are classified
as agricultural (soil moisture deficit), meteorological (precip-
itation deficit), and hydrological (streamflow/groundwater
deficit), and various drought indicators based on soil mois-
ture, precipitation and runoff have been developed to de-
scribe different aspects of droughts (Heim, 2002; Wood et
al., 2002; Wood and Lettenmaier, 2006; Mo, 2008; Shukla
and Lettenmaier, 2011; Hao and AghaKouchak, 2013). Most
drought prediction studies are based on the standardized pre-
cipitation index (SPI; McKee et al., 1993) with the input pre-
cipitation derived from dynamical weather/climate models.
A baseline probabilistic drought forecasting framework using SSI

2 Data

The data sets used in this study include the monthly precipitation and soil moisture from the NASA Modern-Era Retrospective analysis for Research and Applications (MERRA-Land), available on a 2/3° × 1/2° grid from 1 January 1980 onwards (Reichle et al., 2011; Rienecker et al., 2011). MERRA data sets have been used in numerous studies in different climatic regions (Boslavich et al., 2011; Golian et al., 2014; Wong et al., 2011). Uncertainties in MERRA data sets have been evaluated against different observations (e.g., Yi et al., 2011; Kennedy et al., 2011). The results show that MERRA provides valuable information consistent with observations especially in the midlatitudes, while uncertainties in high latitudes are often large (Yi et al., 2011; Reichle et al., 2011).

3 Methodology

The standardized soil moisture index (SSI; Hao and AghaKouchak, 2014) can be defined in a similar way to the commonly used standardized precipitation index (SPI; McKee et al., 1993) that has been used in a wide variety of studies (Dutra et al., 2013; Damberg and AghaKouchak, 2013). Here, the SSI is estimated using a nonparametric approach in which the empirical probability \( p \) of the historical soil moisture data is derived using the empirical Gringorten plotting position (Gringorten, 1963). In other words, instead of fitting a distribution function to soil moisture data, the probabilities \( p \) are obtained empirically using the empirical Gringorten approach: 

\[
(i - 0.44)/(n + 0.12),
\]

where \( n \) denotes the sample size and \( i \) refers to the rank of soil moisture data from the smallest to the largest.

The empirical probabilities, derived from the Gringorten plotting position, are then transformed into the standard normal distribution function: 

\[
SSI = \Phi^{-1}(p),
\]

where \( \Phi \) is the standard normal distribution function. In this approach, one can avoid making a decision about the parametric distribution function of accumulated soil moisture at different timescales. Assume that soil moisture for the month \( i \) is \( S_i \). Then the 6-month accumulation of the soil moisture \( A_i \) for the month \( i \) can be expressed as (Hao et al., 2014)

\[
A_i = S_{i-5} + S_{i-4} + S_{i-3} + S_{i-2} + S_{i-1} + S_i. \tag{1}
\]

In this study, the ensemble streamflow prediction (ESP) method (Twedt et al., 1977; Day, 1985) is used for resampling from historical records of soil moisture to obtain monthly moisture at the target season with the 6-month SSI as the drought indicator. Assume the \( l \)-month lead forecasting is needed based on the monthly soil moisture observations with forecast initialization at month \( i \). Then the \( l \) month

\[
A_{i+l} = S_{i+l-5} + S_{i+l-4} + S_{i+l-3} + S_{i+l-2} + S_{i+l-1} + S_{i+l}. \tag{2}
\]
Assume that 1-month lead forecasting (i.e., \( l = 1 \)) based on the 6-month SSI is needed. The unknown \( S_{i+1} \) is predicted by resampling the soil moisture from the historical record of the target month (i.e., \( i + 1 \)). As a result, an ensemble of \( m \) (i.e., the length of observation in the historical record) sequence of the monthly soil moisture in the target season can be obtained from the observed monthly soil moisture. In this manner, \( m \) sequences of accumulated 6-month soil moisture for the 1-month lead time can be generated by blending the observed and predicted monthly soil moisture. For example, for \( l = 1 \), the blended sequences of accumulated 6-month soil moisture can be expressed as (Hao et al., 2014)

\[
A_{i+1}^{(1)} = S_{i-4} + S_{i-3} + S_{i-2} + S_{i-1} + S_i + S_{i+1},
\]

\[
A_{i+1}^{(2)} = S_{i-4} + S_{i-3} + S_{i-2} + S_{i-1} + S_i + S_{i+1},
\]

\[
\vdots
\]

\[
A_{i+1}^{(m)} = S_{i-4} + S_{i-3} + S_{i-2} + S_{i-1} + S_i + S_{i+1},
\]

(3)–(5)

where \( S_{i-4}, \ldots, S_i \) are the observed soil moisture prior to the target month in the 6-month window, while \( S_{i+1}^{(1)}, \ldots, S_{i+1}^{(m)} \) are the sequences of sampled monthly soil moisture from the observations in the historical record for the target month (here, \( S_{i+1}^{(1)} \)). For any timescale (sc) and lead time (\( l \)), Eq. (3) can be generalized as

\[
A_{i+l}^{(1,\ldots,m)} = \sum_{j=0}^{sc-l-1} S_{i-j}^{(1)} + \sum_{k=1}^{l} S_{i+k}^{(1,\ldots,m)}.
\]

(6)

Note that the lead time (\( l \)) should be less than the timescale (sc) – here, 6 months. Each sequence of the blended 6-month soil moisture \( A^{(j)} \), \( j = 1, 2, \ldots, m \) in Eq. (3) can be combined with the observed 6-month accumulated soil moisture in the past years to derive the corresponding SSI\(^{(j)}\). Here, the probability of drought is defined as the probability that a future drought condition (SSI) is lower than an alarm threshold (e.g., SSI < −0.8 corresponding to ~20th percentile). The empirical probability is estimated by dividing the number of the forecasted values below the threshold (e.g., −0.8) by the number of the ensemble members.

4 Results

First, it is shown that the accumulated soil moisture typically exhibits much higher persistence compared to precipitation, and hence can be used for drought forecasting with up to several months lead time. Then, the 2012 summer drought conditions are predicted using the SSI with different lead times. The SSI is obtained using predicted soil moisture information using the ESP concept based on long-term climatology and near-past observations (see Sect. 3). The study focuses on the drought prediction for May–August, which is an important period for agricultural decision-making.

Understanding the persistence property of soil moisture is fundamental to drought forecasting. It is hypothesized that using accumulated soil moisture would improve persistence-based drought forecasting relative to using accumulated precipitation. First, the persistence property of accumulated soil moisture is evaluated against the accumulated precipitation that has been used for meteorological drought prediction (Lyon et al., 2012; Quan et al., 2012; Hao et al., 2014; Yoon et al., 2012). The monthly precipitation and soil moisture data from MERRA-Land (Reichle et al., 2011; Rienecker et al., 2011) in California and Texas are used to examine the persistence of accumulated soil moisture relative to precipitation. Both states are among the most important producers of agricultural products, and have experienced severe/extreme drought conditions in the past decade. The autocorrelations of accumulated 6-month precipitation and soil moisture for 1–6-month time lags and four different initial conditions (March, April, May and June) for summer drought prediction are provided in Fig. 1. In the figure, the term initial is defined as similar to initial conditions in the Methodology section. For example, March corresponds to precipitation and soil moisture from October 2011 through March 2012. The box plots present the median, 25th, 75th percentiles, and whiskers of the autocorrelations. Lyon et al. (2012) showed that variance of the accumulated precipitation can enhance or diminish the persistence of the SPI at different start times, mainly due to seasonality of precipitation. As shown, the autocorrelation of the accumulated soil moisture (or SSI) is generally higher than that of accumulated precipitation (or SPI) for the four different initial conditions. The figure shows that the autocorrelations of the accumulated 6-month soil moisture decay at a slower rate than the accumulated 6-month precipitation in both California (Fig. 1a) and Texas (Fig. 1b). For example, in California and for the initial condition in April, the medians of the autocorrelation coefficients are higher than 0.6 even at a 5-month lag. However, the medians of the autocorrelations of the 6-month SPI drop below 0.6 after a 4-month lag. The higher persistence of the SSI relative to SPI implies that a persistence-based model based on SSI would lead to better predictions as compared to a similar model based on SPI (see also Changnon Jr., 1987).

The 6-month SSI is used as the drought indicator to monitor and predict the 2012 (May–August) US drought. Figure 2a shows observed drought conditions from May to August 2012. As shown, the drought develops and intensifies quickly, affecting most of the continental US including the Great Plains, the Midwest, and west and southeast. By August, a large portion of the country experienced severe, extreme, or exceptional drought conditions. In operational drought early warning, the severe drought condition is of critical concern. In this paper, the proposed methodology is tested for predicting the moderate and severe drought conditions in summer 2012. Following the US Drought Monitor (USDM), D-scale, the moderate drought is defined as SSI below −0.8 (corresponding to nonexceedance probability of ~0.2), whereas the severe drought is defined as SSI below −1.3 (or nonexceedance probability of ~0.1) (Svoboda et
Figure 1. Box plots of autocorrelation coefficients (up to 6 months) of accumulated 6-month precipitation (blue) and soil moisture (red) from MERRA-Land for different initial month for (a) California and (b) Texas. The box plots show the median (center), 25th (lower) and 75th (upper) percentile edges.

Figure 2. (a) Observed 6-month SSI for May–August 2012; (b) observed 6-month SSI with severe drought condition (SSI < −1.3) for May–August 2012.

The observed drought conditions below the severe level (D2) for May–August are shown in Fig. 2b. The 1- and 2-month lead drought (SSI < −0.8) forecasts for May–August 2012 are presented in Fig. 3a and b, respectively. The 1-month lead forecasted SSI maps for different initializations resemble the observed SSI well in terms of the spatial extent (compare Fig. 3a with Fig. 2a). As shown, the regions with high probability of drought (e.g., above approx. 90 %) are in very good agreement with the observations. For example, the outlined methodology predicts high probability of drought over the western US and high plains in August, which is consistent with observations. Furthermore, as the 2012 drought intensifies, the area with high probability of drought (Fig. 3a) increases in a similar manner to the observations (Fig. 2a). A visual comparison of the 2-month lead drought forecasts (Fig. 3b) and observations (Fig. 2a) reveals that the predicted drought conditions are in very good agreement with probabilities higher than 0.8 in most regions. The 1-month and 2-month lead severe drought (SSI < −1.3) forecasts for May–August 2012 are presented in Fig. 4a and b. The 1-month lead forecasts of May–August severe drought conditions are in very good agreement with observations. As shown, the severe drought conditions from May–August in northern Texas and the western US are captured in the predictions. Figure 4b highlights that, even at a 2-month lead, the proposed model predicts the 2012 summer drought reasonably well.

The predicted drought probability maps for July and August 2012 for 3-month and 4-month lead time are presented in Fig. 5a (SSI < −0.8) and b (SSI < −1.3). One can see that the 3- and 4-month lead forecasts capture the observed drought conditions with probabilities ranging from 0.1 to 0.8.
The prediction skill of the model is higher in the western US where drought conditions are predicted at higher probabilities relative to the Midwest. A review of Figs. 3 and 4 indicates that the predicted probabilities in longer leads (i.e., 3 and 4 months) are typically lower than those of shorter (1 and 2 months) lead forecasts. Basically, in persistence-based models, as the lead month increases, one expects the forecast probabilities to decrease as well. This can be partly explained from the autocorrelations of accumulated soil moisture presented in Fig. 1. As shown, in the western US, the 4-month lead forecasted drought probabilities for July and August 2012 are relatively high and in fairly good agreement with observations. In the Midwest and eastern US, the proposed model indicates relatively low probabilities of drought for 3- and 4-month lead forecasts. While the forecasted drought probabilities are lower at a 4-month lead, still they provide valuable information by showing the drought signal. While the 3- and 4-month lead forecasted probabilities of severe droughts are substantially less compared to the 2-month lead forecasts, the drought signal in the western US is still strong (see Fig. 5b).

It should be noted that the seasonal climate predictions based on weather/climate models initialized in April and May 2012 revealed limited drought information for May–July and June–August 2012 (Hoerling et al., 2013). This highlights that improvements in just two 2-month lead forecasts could be very important for risk assessment and decision-making. The presented persistence-based model with the SSI as the drought indicator provides the potential capability to predict droughts that would be of great value to agricultural planning.

The quality and the latency of predictions rely on the quality and availability of input data sets. Currently, limited observations of soil moisture are available across the globe, and soil moisture estimation relies on model simulations. The soil moisture ocean salinity (SMOS; Kerr et al., 2001) and...
the upcoming soil moisture active and passive (SMAP; En-
tekhabi et al., 2010) mission may provide the opportunity to
integrate near real-time satellite data with long-term climate
data records such as MERRA to improve drought monitoring
and prediction.

5 Conclusions

Using the standardized soil moisture index (SSI) as the
drought indicator, a persistence-based drought prediction
method is presented and used for predicting the 2012 United
States drought. It is shown that because of high persistence
property of soil moisture, the SSI can be used for seasonal
drought forecasting. The presented statistical approach pre-
dicted the May–August drought conditions relatively well,
especially for 1- and 2-month lead forecasts. The 3- and 4-
month lead forecasts of the western US were in good agree-
ment with observations. However, the drought prediction sig-
nal in the eastern US was not as strong at 3- and 4-month lead
time. Given the persistence-based nature of the methodology,
uncertainties of predictions increase with lead time. Similar
behavior has been observed in persistence-based drought re-
covery assessment (Pan et al., 2013). However, even 1- and
2-month lead information is valuable to some end users in-
cluding farmers and commodity investors.

It is acknowledged that, similar to other methods, both the
presented modeling framework and input data sets are sub-
ject to uncertainties (e.g., see Quan et al., 2012). The pre-
sented model is based on near-past soil moisture conditions
and long-term climatology. Soil moisture responds to precip-
itation with some delay, and for this reason the methodol-
ogy may not capture rapid developments. Furthermore, this
methodology relies on historical observations; because of
limited samples of extreme conditions in historical records,
it should not be used for predicting extreme droughts.

It is stressed that the proposed approach is not meant to
replace the currently available dynamic drought forecasting
models. Rather, the persistence-based predictions should be
used as additional information that can potentially improve
drought predictability. Finally, it should be pointed out that
SSI is not suggested as an alternative to using SPI (or other
indicators) for seasonal drought prediction. The best choice
of index or the best set of indicators depends on the problem
at hand and the climate of the study area. It is our view that
drought monitoring and prediction should be based on mul-
tiple sources of information (data and indicator) as well as
models (e.g., dynamic, statistical).

Acknowledgements. The author appreciates the constructive
comments and suggestions from three anonymous reviewers that
led to substantial improvement in the current version. This study is
partially supported by the NOAA Modeling, Analysis, Prediction,
and Projections (MAPP) program (award no. NA14OAR4310222),
and the Hellman Foundation.

Edited by: F. Pappenberger

References

AghaKouchak, A. and Nakhjiri, N.: A Near Real-Time Satellite-
Based Global Drought Climate Data Record, Environ. Res. Lett.,
Bosilovich, M. G., Robertson, F. R., and Chen, J.: Global energy
and water budgets in MERRA, J. Climate, 24(22), 5721–5739,
2011.
Changnon Jr., S. A.: Detecting drought conditions in Illinois, Illi-
nois State Water Survey, Champaign, IL, USA, ISWS C-169,
36 pp., 1987.
Damberg, L. and AghaKouchak, A.: Global Trends and Pat-
terns of Droughts from Space, Theor. Appl. Climatol.,
A. AghaKouchak: A baseline probabilistic drought forecasting framework using SSI 2491


