Predicting U.S. Drought Monitor States Using Precipitation, Soil Moisture, and Evapotranspiration Anomalies. Part I: Development of a Nondiscrete USDM Index

DAVID J. LORENZ,[a] JASON A. OTKIN,[b] MARK SVOBODA,[c] CHRISTOPHER R. HAIN,[d] MARTHA C. ANDERSON,[e] AND YAFANG ZHONG[b]

[a] *Center for Climatic Research, University of Wisconsin–Madison, Madison, Wisconsin*
[b] *Space Science and Engineering Center, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin–Madison, Madison, Wisconsin*
[c] *National Drought Mitigation Center, University of Nebraska–Lincoln, Lincoln, Nebraska*
[d] *Earth Science Branch, NASA Marshall Space Flight Center, Huntsville, Alabama*
[e] *Hydrology and Remote Sensing Laboratory, Agricultural Research Services, U.S. Department of Agriculture, Beltsville, Maryland*

(Manuscript received 15 March 2016, in final form 3 March 2017)

ABSTRACT

The U.S. Drought Monitor (USDM) classifies drought into five discrete dryness/drought categories based on expert synthesis of numerous data sources. In this study, an empirical methodology is presented for creating a nondiscrete USDM index that simultaneously 1) represents the dryness/wetness value on a continuum and 2) is most consistent with the time scales and processes of the actual USDM. A continuous USDM representation will facilitate USDM forecasting methods, which will benefit from knowledge of where, within a discrete drought class, the current drought state most probably lies. The continuous USDM is developed such that the actual discrete USDM can be reconstructed by discretizing the continuous USDM based on the 30th, 20th, 10th, 5th, and 2nd percentiles—corresponding with USDM definitions for the *D*4–*D*0 drought classes. Anomalies in precipitation, soil moisture, and evapotranspiration over a range of different time scales are used as predictors to estimate the continuous USDM. The methodology is fundamentally probabilistic, meaning that the probability density function (PDF) of the continuous USDM is estimated and therefore the degree of uncertainty in the fit is properly characterized. Goodness-of-fit metrics and direct comparisons between the actual and predicted USDM analyses during different seasons and years indicate that this objective drought classification method is well correlated with the current USDM analyses. In Part II, this continuous USDM index will be used to improve intraseasonal USDM intensification forecasts because it is capable of distinguishing between USDM states that are either far from or near to the next-higher drought category.

1. Introduction

Drought is a naturally recurring climate phenomenon that can reduce agricultural productivity and damage natural ecosystems if it persists for an extended period of time or if it occurs during critical stages of crop development. Extreme drought events have impacted extensive areas of the United States in recent years, most notably the south-central United States in 2010–15, the Corn Belt in 2012, the Pacific Northwest in 2013–15, and California during 2012–16. According to the U.S. Drought Monitor (USDM; Svoboda et al. 2002), extreme *D*3 to exceptional *D*4 drought conditions occurred for at least several months during the growing season during each of these events. These exceptional conditions led to large agricultural losses. For example, the 2011 drought was the most expensive drought in Texas history, with grain and livestock losses estimated at $7.62 billion (Fannin 2012). Drought-related losses were even larger during the 2012 Midwest "flash drought" because of its widespread impact on corn and soybean yields across the Midwest, where county-level crop yields were often less than 75% of normal (Otkin et al. 2016). Indeed, the 2012 flash drought was one of the most expensive natural disasters in U.S. history, with federal crop indemnity payments alone surpassing $17 billion (NRDC 2013) and the total economic cost estimated to be in excess of $35 billion (AON Benfield 2013). More recently, extreme drought conditions occurring over prime agricultural land in California have

Corresponding author: David J. Lorenz, dlorenz@wisc.edu

DOI: 10.1175/JHM-D-16-0066.1

led to estimated agricultural losses in that state exceeding $2.7 billion for 2015 alone (Howitt et al. 2015). Though greater reliance on irrigation has blunted the impact of the drought on California agriculture, losses have been increasing each year as more land is being left fallow because of diminishing irrigation allocations and rapidly falling water tables that make it impossible or too expensive to pump groundwater.

The occurrence of numerous extreme drought events affecting vast areas of the United States in recent years, combined with the high social, economic, and environmental toll exacted by these events, demonstrates our continued vulnerability to drought and the need to develop drought early warning systems (DEWS) that effectively characterize and disseminate value-added information about current and future drought conditions. Vulnerable stakeholders such as farmers and ranchers and governmental organizations tasked with providing aid to those affected by drought can use this information to assist drought mitigation efforts (Otkin et al. 2015). Indeed, a defined goal of the National Integrated Drought Information System (NIDIS), which was enacted by Congress in 2006 and reauthorized in 2013, is to create a DEWS capable of providing probabilistic drought intensity forecasts with sufficient temporal and spatial resolution for stakeholders to make informed management decisions. This task, however, is made more difficult by the diversity of drought types and their differing impacts on various stakeholder groups (Wilhite and Glantz 1985). For example, flash droughts that develop quickly during the growing season (Svoboda et al. 2002; Mozny et al. 2012; Otkin et al. 2013; Hunt et al. 2014) can seriously impact agriculture while leaving other groups sensitive only to long-term hydrological drought unaffected. The extremely rapid drought intensification rates observed during the 2012 flash drought also indicate that a robust drought early warning system should provide subseasonal forecasts updated at weekly intervals in addition to seasonal forecasts that are updated less frequently.

Most drought forecasting systems, however, only provide seasonal forecasts that are updated at monthly intervals, which limits their real-time utility when conditions are rapidly changing. For subseasonal drought prediction (e.g., <3 months), a promising new approach is to use rapid changes in drought indices sensitive to evapotranspiration, precipitation, or soil moisture to identify areas most susceptible to drought development (Otkin et al. 2014, 2015). In this two-part paper, we will expand upon the method described in Otkin et al. (2014, 2015) by using more sophisticated statistical tools to estimate the current drought state and to forecast the probabilistic likelihood of future drought development. In this paper (Part I), the primary focus is not on future prediction but on better characterizing the current drought state. We describe a new method to compute a continuous USDM index (i.e., wetness/dryness is measured on a continuum rather than as six discrete dryness/drought classes) that estimates the likelihood that the current conditions are within each of the six USDM drought categories based on standard precipitation index (SPI), evaporative stress index (ESI), and North American Land Data Assimilation System (NLDAS) anomalies computed over various time periods. The methodology is fundamentally probabilistic, meaning that the probability density function (PDF) of the continuous USDM is estimated and therefore the degree of uncertainty in the fit is properly characterized. The probabilities are objectively derived from the long-term history of the predictor fields and the categorical USDM observations from 2001 to 2014. While the ultimate goal of this research is to predict future changes in the USDM, Part I does not predict the future; instead, we aim to more precisely define the current state of drought as depicted in the USDM. In Lorenz et al. (2017, hereafter Part II), the improved current drought state estimates generated using the method described in this paper are used to compute probabilistic intensification forecasts over subseasonal time scales. The state estimates in this current paper improve future forecasts because they quantify how far the current state is from the next-higher dryness/drought category. For example, the no drought USDM designation can mean near-normal conditions or it can mean extremely wet. Recent dry anomalies are much more likely to lead to an intensification of the drought category in the former case than in the latter case.

There are many potential drought indices/indicators that we could use to characterize drought. We use the USDM because it is the most widely available index that incorporates local drought impacts through its network of observers and contributors around the country. Through expert assessment, the USDM assimilates these contributions with measurements of climatic, hydrologic, and soil conditions into a single drought indicator that attempts to best summarize the drought state across different regions with different drought processes and characteristics. While the impacts-based approach has many advantages, it also introduces some subjectivity into the indicator. We believe a probabilistic approach, such as that used here, is particularly well suited to coping with the subjective aspects of the USDM because it characterizes both the explained and the unexplained portion of the variance. In the USDM's case, the subjective aspects are accounted for in the unexplained portion of the variance. The single composite drought

indicator approach of the USDM also introduces uncertainties because different kinds of droughts (i.e., short vs long term) can have the same USDM designation, and sometimes short- and long-term conditions are on opposite sides of the wet/dry spectrum. Such uncertainties are also are handled well by a probabilistic approach.

In this paper, we begin with a description of the datasets (section 2) and a discussion of the assumptions, fitting, and interpretation of the new methodology (section 3). Results are presented in section 4 and a summary and discussion follow in section 5.

2. Datasets

The purpose of this study is to develop a methodology for weekly predictions of current and future states of the USDM in gridded form; therefore, we focus on predictors that are available in gridded form at weekly time scales. We thus do not use divisional Palmer drought severity index data, streamflow data, crop and range conditions, and other nongridded datasets. In future studies, these additional variables could be added as predictors after gridding. Because the USDM is a relatively low-resolution depiction of drought conditions on the ground, all predictors have been aggregated to a coarse-resolution grid (0.12° × 0.12°) before application of the statistical models (see section 2e).

a. USDM

The USDM is created each week through expert synthesis of numerous data sources, including surface streamflow, soil moisture, rainfall anomalies, temperature anomalies, crop and range conditions, and snow cover. These sources are assimilated in more or less real time into the latest USDM. The above data sources are combined with local impact reports (e.g., crop and range conditions) from over 400 observers across the country. Many times the local impacts are not known until after the USDM is made, and therefore these reports are used to validate and/or adjust the USDM as they become available. Because of the nature of the USDM, there is sometimes a lag between drought-related anomalies on the ground and the USDM (e.g., Otkin et al. 2014, 2015). Nevertheless, for the USDM state estimates in this paper, we find that not using a time lag results in the best skill. In addition, USDM attempts to identify all kinds of drought. Hence, if long-term hydrologic indicators are very dry, then the USDM will show drought even if recent rains have led to good crop and range conditions. The opposite scenario is also possible. This USDM disadvantage is also true of any composite index that condenses multivariate information into a single variable.

The USDM is updated each week based on data obtained through Tuesday morning and is publicly released on Thursday. Using the ray-casting algorithm (Shimrat 1962; Hacker 1962), we assign USDM values to individual grid points in the modeling domain using the USDM shapefile analyses that classify dryness/drought into five categories ranging from abnormal dryness to exceptional drought (Svoboda et al. 2002). We define the drought category for each grid point as the value of the USDM at the center of the grid point. For this study, we designate the six dryness/drought categories as integers from −1 to 4: "no drought" has a value −1, "abnormally dry" = 0, "moderate drought" = 1, "severe drought" = 2, "extreme drought" = 3, and "exceptional drought" = 4.

b. SPI

Precipitation anomalies are computed with the SPI (McKee et al. 1993, 1995), which uses precipitation as its sole input and is widely used to detect meteorological drought conditions. It is standardized so that values less than zero indicate that the observed precipitation was less than the climatological median precipitation over a given time period. The SPI was computed using the 0.25° resolution Climate Prediction Center's gridded analysis of daily precipitation reports from National Weather Service reporting stations and cooperative observers (Higgins et al. 2000). First, the daily precipitation (1948–2014) was interpolated to the 0.12° × 0.12° resolution grid and then the SPI was computed at weekly intervals ending on Tuesday after compositing precipitation over 4-, 8-, 12-, 16-, 20-, 26-, 39-, and 52-week time periods.

c. ESI

Moisture stress can be expressed as the reduction in actual evapotranspiration (ET) from the reference ET (ET_{ref}) expected under non-moisture-limiting conditions. To facilitate identification of anomalous moisture conditions, the ESI was developed to represent standardized anomalies in the ET fraction (ET/ET_{ref}), where the Atmosphere–Land Exchange Inverse model (ALEXI) (Anderson et al. 1997, 2007b) is used to estimate the actual ET and estimates of surface air temperature, surface specific humidity, surface pressure, and wind speed from the Climate Forecast System Reanalysis (CFSR) dataset (Saha et al. 2010) are used to compute ET_{ref} based on the Penman–Monteith formulation (Allen et al. 1998). Normalization by a reference ET is used to minimize the impact of non-moisture-related drivers of ET such as the seasonal cycle in solar radiation.

ALEXI is a two-source energy balance model (Norman et al. 1995) that uses land surface temperatures (LSTs) obtained from satellite thermal infrared imagery

to estimate sensible, latent, and ground heat fluxes for vegetated and bare soil components of the land surface. The partitioning of the energy fluxes into soil and vegetated components is estimated using vegetation cover fraction estimates derived from the MODIS leaf area index product (Myneni et al. 2002). The total surface energy budget is computed using the observed increase in LST from ~1.5 h after local sunrise until 1.5 h before local noon, with the McNaughton and Spriggs (1986) atmospheric boundary layer (ABL) growth model used to provide closure to the energy balance equations. Lower-tropospheric temperature profiles used by the ABL model are obtained from the CFSR dataset (Saha et al. 2010). ALEXI is run each day over the contiguous United States with 4-km horizontal grid spacing using LST retrievals and insolation estimates derived from the Geostationary Operational Environmental Satellite imager. It has been shown to provide reasonable ET estimates for a variety of climate regimes and vegetation types (Anderson et al. 2007a).

Because ALEXI uses the morning rise in LST to estimate ET, incomplete cloud screening can add noise to the ET time series used to compute the ESI. These errors are reduced using a temporal smoothing algorithm that identifies and removes days with ET estimates that differ by more than two standard deviations from surrounding days within a 14-day moving window (Anderson et al. 2013). The remaining clear-sky ET estimates are composited over longer time periods to achieve more complete domain coverage. Standardized ET fraction anomalies, expressed as pseudo z-scores normalized to a mean of 0 and a standard deviation of 1, are computed each week using 4-, 8-, 12-, 16-, 20-, 26-, 39-, and 52-week composite periods ending on Tuesday. For consistency reasons, these are the same composite periods as the SPI. While ESI is not measured during periods of snow cover, the amount of missing data is surprisingly small (92% of the domain has at least 75% temporal coverage for the 4-week composite ESI). The mean ET fraction and standard deviations for each composite period are computed at each grid point using data from 2001 to 2014. Standardized anomalies are computed as

$$\mathrm{ESI}(w,y) = \frac{V(w,y) - \frac{1}{n_y}\sum V(w,y)}{\sigma}, \qquad (1)$$

where $V(w, y)$ is the composite ET fraction for week w and year y at a given grid point, the second term is the mean ET fraction for week w averaged over all years (n_y = number of years), and the denominator is the standard deviation. By standardizing the anomalies, this means that negative (positive) values depict below (above) average ET fluxes, which are typically associated with lower (higher) than average soil moisture content or vegetation health.

d. NLDAS

Soil moisture anomalies in the top 10, 100, and 200 cm of the soil profile were computed for several models in phase 2 of the NLDAS (NLDAS-2; Xia et al. 2012a,b), including Noah (Ek et al. 2003; Barlage et al. 2010; Wei et al. 2013), Mosaic (Koster and Suarez 1994, 1996), and the Variable Infiltration Capacity model (VIC; Liang et al. 1996; Bowling and Lettenmaier 2010). Though each model simulates surface energy, water balance, and soil moisture in multiple layers, their treatment of key processes such as infiltration, drainage, vegetation rooting depth, and canopy uptake differs, which can lead to different soil moisture responses. Given this variability, we use the ensemble average soil moisture (hereafter referred to as NLDAS) because this has been shown to more accurately depict drought conditions (Xia et al. 2014). Daily soil moisture values for each soil layer were averaged each week over 4-, 8-, 12-, 16-, 20-, 26-, 39-, and 52-week time periods and then converted into standardized anomalies using data from 1979 to 2014 in the same way as the ESI.

Given the fact it has been shown that the above drought predictors can lead to USDM changes (e.g., Otkin et al. 2014, 2015), it may seem counterproductive to use such long period composites (i.e., 4, 8, 12, 16, 20, 26, 39, and 52 weeks) and to not use a time lag. We have tried these alternatives but have found that the current approach works best. For Part II, on the other hand, we find that weekly predictors that are lagged in time work best. The predictors are different because the time tendency of any general geophysical time series (i.e., with variations at all time scales) is a high-frequency version of the original time series. Therefore, long-term composites work best for estimating the USDM state and short-term composites work best for predicting USDM changes (Part II).

In summary, we have five fields: SPI, ESI, 0–10 cm soil moisture, 0–100 cm soil moisture, and 0–200 cm soil moisture. Each of these fields is composited over eight different time periods. Hence, we have 40 total predictors. While the different predictor variables have different periods of record, this is of less consequence for this study because we do not consider the raw anomalies in isolation. Instead, the regression techniques used here naturally scale and add an offset that will take into account the different climatologies.

e. Time period and smoothing

The current study focuses on warm season USDM from May through September (technically from the

eighteenth to the thirty-ninth Tuesday of the year; the data cutoff for the weekly USDM is Tuesday) for the years 2001–14. Data from the remainder of the year are included for the calculation of long-term predictor composites. To improve correlation with the low-resolution USDM maps, the predictor variables were spatially smoothed using a two-dimensional Gaussian weighting function:

$$w = \exp\left(-\frac{1}{2}\left\{\left[\frac{\Delta x}{a\cos(40°)}\right]^2 - \left(\frac{\Delta y}{a}\right)^2\right\}\right),$$

where Δx and Δy are the difference in longitude and latitude, respectively, between the central grid point and a surrounding grid point; a is the smoothing radius in latitude; and the cos(40°) factor scales the longitude distance so that the smoothing radius is approximately equal in longitude and latitude.

3. Methodology

Any estimate of the USDM from SPI, ESI, and NLDAS will not have perfect skill, particularly because the USDM includes subjective elements. Therefore, we believe it is most useful to make probabilistic predictions because such predictions characterize the degree of uncertainty and how the estimate deviates from the most likely value. To create an explicit statistical model of the USDM PDF that can be used to make quantifiable predictions, we must first make some explicit assumptions about the USDM.

a. General assumptions

The USDM classifies drought into five discrete dryness/drought categories.[1] The goal of this research is to better characterize current state of USDM so that we can quantify how far the current drought state is from the next-higher or next-lower drought category. In other words, we want to define a "continuous version" of the USDM. We assume that the actual discrete USDM can be reconstructed from the continuous USDM by discretizing based on the 30th, 20th, 10th, 5th, and 2nd percentiles. These are the explicit percentile thresholds that the USDM uses to define the boundaries of the drought categories from wettest to driest. Figure 1 is a schematic showing the PDF of this hypothetical continuous version of the USDM and the five dryness/drought categories. Because only the discretized value of the USDM is known a priori, the most likely value of the continuous version of the USDM must be estimated. Figure 1a shows the climatological PDF at a single grid point over all times in the record. Note that unlike the actual USDM, which only distinguishes between degrees of dryness, the continuous USDM quantifies the degree of wetness and dryness. In the actual USDM, on the other hand, all wet states are aggregated together into a single no drought category.

Figures 1b and 1c are examples of the USDM PDF at two particular times t_1 and t_2 given the values of the SPI, ESI, and NLDAS predictors (i.e., these represent conditional PDFs). It is assumed that SPI, ESI, and NLDAS add information about the USDM state; therefore, the conditional PDFs at times t_1 and t_2 are narrower than the total PDF (Fig. 1a). At time t_2, the predictors suggest conditions are wetter than normal because the PDF is centered to the right of the peak in the total PDF. At time t_1, on the other hand, the predictors suggest that conditions are drier than normal. However, the area of the PDF to the right of the abnormally dry threshold is larger than the area to the left, which implies that conditions are still more likely to be in the no drought category than in the abnormally dry or moderate drought categories. The distributions at times t_1 and t_2 are both cases where the majority of the PDF is in the USDM no drought category; however, the degree of certainty and the closeness to other drought categories is very different in these two cases. In this example, given the predictors alone, the USDM category at time t_1 is much less certain than the USDM category at time t_2. Furthermore, suppose we also know that the actual USDM is in the no drought category at both t_1 and t_2. Then, given an identical forecasted drying trend, one would expect that the USDM is more likely to intensify at time t_1 than at t_2. This information is exploited in Part II in order to predict changes in the USDM.

The above discussion introduces a general framework for generating a nondiscrete representation of the current USDM state. To explicitly define a statistical model that can be fit to the input data, we must specify the functional form of the distribution and how the mean, width, and shape of this distribution vary as a function of the SPI, ESI, and NLDAS predictors. Here we assume that the shape and width are constants independent of SPI, ESI, and NLDAS and that the mean of the distribution is a linear function of SPI, ESI, and NLDAS. The details of the parametric distribution and the fitting of the free constants in the model are described in the next subsection.

b. Specification of the model

Let $f(x)$ and $F(x)$ be the PDF and CDF describing the shape of the continuous USDM distribution given the

[1] Six categories if one includes the "no drought" designation as a category.

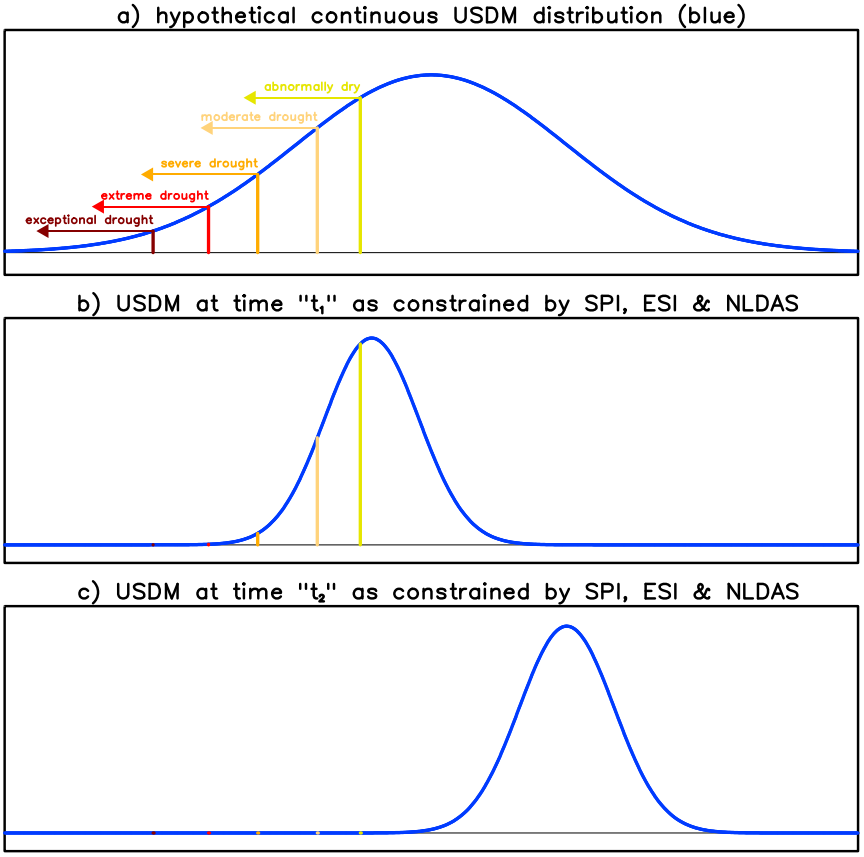

FIG. 1. (a) Schematic of the hypothetical, continuous USDM distribution for all times for a single grid point. (b) Conditional PDF estimate of USDM at time t_1 given the information in the SPI, ESI, and NLDAS. Note the width of the distribution is smaller than in (a) because SPI, ESI, and NLDAS help constrain the range of USDM possibilities. (c) As in (b), but for time t_2.

SPI, ESI, and NLDAS values (Figs. 1b,c). The distribution can be generalized by allowing the width and the position of the distribution to vary:

$$\text{PDF} = sf(s[x-\mu]) \quad \text{and}$$
$$\text{CDF} = F(s[x-\mu]), \tag{2}$$

where s is a constant that controls the width (smaller values of s imply a wider distribution and greater uncertainty) and μ controls the position. If the PDF is the standard normal distribution, then μ is the mean and s is the reciprocal of the standard deviation. As mentioned above, we assume that the mean (i.e., position) is a linear function of the predictors:

$$\mu = a_0 + a_1 p_1 + a_2 p_2 + \cdots, \tag{3}$$

where the a_i are constants that need to be fit and the p_i are the 40 SPI, ESI, and NLDAS predictors. The p_i change in time depending on the state of SPI, ESI, and NLDAS. In addition to the a_i, the constant s also needs to be fit to the data.

For the distribution itself [i.e., $f(x)$], we explored the standard normal distribution as well as the following distributions with heavier tails than the normal distribution: 1) Student's t, 2) logistic, and 3) hyperbolic secant. To quantify the "best" distribution the Brier skill score (BSS) was used, and this led to the Student's t distribution being chosen. The Student's t distribution has a free parameter ν, which corresponds to the degrees of freedom when the distribution is applied to significance testing. Significance testing is not the point here, so instead t distributions with excess kurtosis of 1, 2, and 3 were tried and the distribution with excess kurtosis of 1 was chosen. This corresponds to $\nu = 10$. Note that the above distributions are all symmetric about the mean. We did not use a skewed distribution (i.e., Gamma distribution) because the PDFs of our predictors are nearly symmetric for all but the driest portions of the domain (not shown). The ESI and soil

moisture variables are symmetric because the 4-week (and longer) averaging time scales are long enough for the central limit theorem to apply. The SPI is symmetric by construction.

The thresholds that separate the continuous USDM distribution into discrete categories must also be specified. One choice is to set the thresholds equal to the 30th, 20th, 10th, 5th, and 2nd percentiles of the standard Student's t distribution; however, the distribution of the USDM over all times is not necessarily the same as the conditional distribution but instead depends on the details of the PDFs of the SPI, ESI, and NLDAS predictors that determine the mean at each time. Therefore, alternative forms were also tried, including equally spaced and empirically determined thresholds. Sensitivity tests showed that no single method performed consistently better than the others; therefore, the thresholds were set to the 30th, 20th, 10th, 5th, and 2nd percentiles for the Student's t distribution with $\nu = 10$. In the equations below, the thresholds that bound a drought category are labeled as T_{d-1} and T_d, where d is an integer index for the drought category that increases as drought becomes more intense. For ease of notation, let T_{d-1} for the no drought category be ∞ and T_d for exceptional drought be $-\infty$.

Recently, Hao et al. (2016) developed a similar probabilistic scheme for estimating the USDM from objective drought indicators over the state of Texas. Rather than assuming a continuous USDM that is discretized based on 30th, 20th, 10th, 5th, and 2nd percentiles, they use a general statistical method called ordinal regression. One advantage of our method is that the number of free parameters is less, so the statistical fits are potentially more robust: our method has $n + 2$ free parameters, where n is the number of predictors. The method in Hao et al. (2016) has $n + 5$ free parameters (for the five thresholds separating the categories). Our method also naturally predicts the dependence of skill on USDM category (see section 3c). In future work, we will consider their method and compare it to ours.

c. Fitting of the model

We fit the statistical model described above by first determining the likelihood function associated with our model and then maximizing the likelihood using general optimization software. Suppose the state of the actual, discrete USDM at a certain time is category d, which is bracketed by the continuous USDM thresholds of T_{d-1} and T_d. Meanwhile, the PDF of the USDM at this same time is of the form $sf(s[x - \mu])$, where μ varies with time because it is a linear combination of the predictors [e.g., (3)]. According to this PDF, the likelihood l of the d USDM category is the integral of the PDF from T_{d-1} to T_d:

$$l = \int_{T_{d-1}}^{T_d} sf(s[x-\mu])\,dx = F(s[T_d-\mu]) - F(s[T_{d-1}-\mu]), \quad (4)$$

where the definition of CDF is used to write the PDF integral explicitly in terms of F. Ideally, the likelihood is 1, which means that the statistical model is 100% certain that the USDM is in category d. With only one time (or case) this is trivial: simply choose μ to be between T_{d-1} and T_d and let the width of the PDF be infinitesimally small ($s = \infty$). In reality, we want to explain the USDM variability over many cases, and in this instance the total likelihood over all times L is the product of the likelihoods at each individual time [$= l(t)$]:

$$L = l(t_1)l(t_2)l(t_3)\cdots. \quad (5)$$

With a large collection of times, however, it is not possible to achieve likelihoods of 1 (100% certainty) while simultaneously satisfying the constraints that s is constant and μ is a linear combination of the predictors. Also, to prevent overfitting, we clearly do not want to keep elaborating the model to eliminate these constraints. Instead, we find the parameters $s, a_0, a_1, a_2, \ldots$, which maximize the likelihood function in (5). This is the maximum likelihood solution, which can be thought of as the most probable parameters given the data. The details of the algorithm and numerical scheme to solve this optimization problem are given in the appendix. Once the parameters are found, we have a best guess of the time-varying PDF of the continuous USDM state given the SPI, ESI, and NLDAS predictors.

d. Distilling the PDF into a single value

The time-varying PDF of the USDM state is useful because it quantifies the uncertainty in the prediction of USDM given the SPI, ESI, and NLDAS. Nevertheless, for succinctly comparing the results with the actual USDM, it is helpful to summarize the PDF information into a single value. There are multiple ways of summarizing the PDF depending on the metric one uses to define closeness between the predicted and actual USDM. In this study, we use a metric based on the least squares error, which is perhaps the most widely used measure of error. As mentioned above, the USDM categories are given integers from -1 to 4, representing the no drought, abnormally dry, moderate drought, severe drought, extreme drought, and exceptional drought categories, respectively. Given the PDF of the continuous USDM and the value b for the "best" single number that summarizes the PDF, the expected value of the square error as a function of b is

$$\varepsilon = \int_{-\infty}^{T_{-1}} [b-(-1)]^2 sf(s[x-\mu])\,dx + \int_{T_{-1}}^{T_0} (b-0)^2 sf(s[x-\mu])\,dx + \int_{T_0}^{T_1} (b-1)^2 sf(s[x-\mu])\,dx + \int_{T_1}^{T_2} (b-2)^2 sf(s[x-\mu])\,dx + \int_{T_2}^{T_3} (b-3)^2 sf(s[x-\mu])\,dx + \int_{T_3}^{\infty} (b-4)^2 sf(s[x-\mu])\,dx. \quad (6)$$

Simplifying by taking the derivative with respect to b and setting the derivative to zero, we find the b that minimizes the square error:

$$b = -\int_{-\infty}^{T_{-1}} sf(s[x-\mu])\,dx + \int_{T_0}^{T_1} sf(s[x-\mu])\,dx + 2\int_{T_1}^{T_2} sf(s[x-\mu])\,dx + 3\int_{T_2}^{T_3} sf(s[x-\mu])\,dx + 4\int_{T_3}^{\infty} sf(s[x-\mu])\,dx. \quad (7)$$

This least squares error value b, depicting the most likely category in the continuous USDM distribution, is used to measure the goodness of fit by calculating the correlation between the predicted (i.e., b) and the observed USDM and is also used to compare maps of the observed and predicted USDM values across the United States.

e. Cross validation

When making predictions with multiple input variables, it is essential to avoid overfitting, which can occur when a model has too many predictor variables relative to the number of observations. In this situation, the model may optimize over the random variability in a given sample in addition to any real relationships that might be present. When an overfit model is applied to independent data, it may perform poorly. In this study, cross validation is used to determine if a given predictor adds skill to the statistical model. This process involves 1) removing one year of data from the analysis, 2) fitting the statistical model on all other years, 3) calculating the skill when applying the model to the year that was left out, and 4) repeating until all years have had a chance to be left out. Typically, one prevents overfitting by incrementally adding additional predictors with cross validation until skill decreases on independent data. This process of determining the useful predictors is called forward variable selection. Unfortunately, because of the relatively short period of record, the number of skillful predictors chosen by forward selection is not robust and tends to vary significantly from grid point to grid point. In addition, for more than half the domain only one predictor is chosen, despite the fact that the USDM authors use a wide range of drought indictors. There is also finescale structure in the particular predictors chosen, and typically changes in the dominant predictor occur as sharp transitions over adjacent grid points rather than more gradual transitions where one predictor gains weight while another loses weight.

f. Aggregate variables before model fitting

The issues with robustness suggest that it is helpful to increase the sample size by including data from grid points surrounding a given grid point when fitting the statistical model. Though nearby grid points may not have the same relationship between the various predictors and the USDM, our tests indicate that the errors from the limited sample likely dominate in most locations. While the fitting of the statistical model is performed using nearby grid points, the validation of the model on the left-out year is still done on the central grid point. For fitting, the nearby grid points are not given full weight; instead, the weight w of a grid point in the analysis is given by a Gaussian function:

$$w = \exp\left[-\frac{(\Delta x)^2 + (\Delta y)^2}{(2^\circ)^2}\right], \quad (8)$$

where Δx and Δy are the difference in longitude and latitude, respectively, between the central grid point and a surrounding grid point in degrees. The choice of 2° for the weighting function is a compromise value that was found to perform the best on average. The use of surrounding grid points could potentially be used with the same forward variable selection scheme above to fit the model. Unfortunately, with 40 variables to try, the computational cost of forward selection is simply too much, and therefore alternative methods must be used.

g. Pattern selection using regularized regression

One strategy for addressing the computational cost of forward selection is to aggregate the predictors together before performing the cross validation. For example, empirical orthogonal functions (EOFs) have been used in the past to distill a field of predictors into a few dominant patterns before applying another statistical model. However, the EOFs would optimize the variability

over the 40 predictors instead of optimizing the relationship to USDM. Another strategy is to use the coefficients from linear least squares regression (with USDM as the predictand) as weights to linearly combine predictors. With the limited sample size, however, linear regression will not be robust with 40 predictors, even with the use of surrounding grid points. Because simple regression is prone to overfitting, some form of regularized linear regression to find the weights is more appropriate. Regularized regression methods introduce a penalty for complexity that typically favors models with smaller and/or fewer nonzero regression coefficients. This penalty is typically associated with a parameter that controls the size of the penalty and hence the degree of regularization. Recently, Meinshausen (2013) and Slawski and Hein (2013) have shown that least squares regression with a sign constraint on the regression coefficients can have similar regularization properties without having to determine the ''best'' value of a regularization parameter. Because the sign of the relationship between the drought indicators and the USDM is known a priori (minus random sampling noise), sign-constrained regression is easy to apply. Moreover, the sign constraint is consistent with our understanding of the way the USDM is created. For example, we expect that if the SPI at a certain time scale is used to predict the USDM, its weight is such that a more negative value implies more intense drought. Conversely, if this SPI predictor is not used, then its weight is zero. We do not expect SPI to have the ''wrong'' sign weight, yet this is exactly what standard linear regression will do to some predictors as it tries to reduce the error to the smallest possible value. Sign-constrained regression, on the other hand, respects our expectations about the sign of the predictor weights. In practice, with 40 predictors, this means that only a portion of the coefficients of sign-constrained regression will be nonzero and the rest will be exactly zero. In this respect, it is similar to the standard forward-selection procedure above; however, we find that the number of predictors used is significantly more than those used in forward selection, which is consistent with the wide range of drought indicators considered for the USDM. In addition, we find that the nonnegative least squares (NNLS) method is more robust with higher correlations with the ''observed'' USDM and there are fewer grid points with no skill on independent data.

The details of our calculation of the predictor weights using sign-constrained regression are as follows. First, we specify real numbers for each drought category. Here it is assumed that the continuous USDM follows the standardized normal distribution and that the thresholds separating the six drought categories are the 30th, 20th, 10th, 5th, and 2nd percentiles. The real number for each category is defined as the mean of the standard normal distribution within that category. Using the thresholds above, the calculated means are 0.50, −0.68, −1.04, −1.45, −1.82, and −2.42, for the no drought, abnormally dry, moderate drought, severe drought, extreme drought, and exceptional drought categories, respectively. Let this new observed USDM index with these values for the drought categories be called u. Defined in this way, u is positively correlated with the SPI, ESI, and NLDAS anomalies, and therefore the sign-constrained regression takes the form of least squares regression with nonnegative regression coefficients (except for the intercept). The predictors are normalized before applying the NNLS algorithm. With normalized predictors, the NNLS regression coefficients are determined solely by the correlations among the predictors and between the predictors and u. For efficiency, the correlations are first calculated for each grid point and then later smoothed in space using the weighting scheme in (8). We calculate the NNLS coefficients from the correlations using cyclic coordinate descent (e.g., Franc et al. 2005). The regression coefficients define the weights to linearly combine the 40 variables to create a single ''master index.'' All steps are cross validated by calculating the NNLS coefficients separately for each subperiod. After the master index is created for each subperiod, the standard forward-selection cross validation is then performed using this single ''master predictor'' to determine if it has skill on independent data. Skill is measured with the log likelihood. Unlike the NNLS weighting scheme, which uses surrounding grid points to determine the weights, the final one-predictor model is fit using the master predictor and the USDM at a single grid point.

4. Results

a. Correlation with the USDM

The correlation computed using the above is shown in Fig. 2. All correlations shown here are cross validated, meaning that the statistical model has not ''seen'' the data used to calculate the correlation. There are 22 weeks and 14 years, which implies a sample size of 308 total weeks. Over most of the central and eastern United States correlations are greater than 0.70; however, correlations are lower across large parts of the western United States. In fact, there are a few locations in the western United States where there is no skill (white). In future studies, including runoff from NLDAS may provide skill in these regions. Averaged over the domain, the correlations are 0.79 over the central and eastern United States

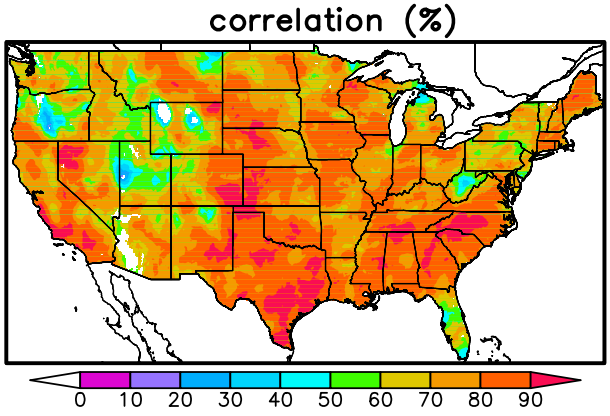

FIG. 2. Cross-validated correlation (%) between USDM and the USDM estimated from the 40 SPI, ESI, and NLDAS predictors using the weighting scheme involving NNLS. The sample size is 308 weeks.

(east of 105°W) and 0.65 over the western United States. To better compare with other studies, the corresponding correlations without cross validation are 0.85 and 0.78.

b. *Chosen predictors*

In this section, we discuss the number and the relative importance of the various predictors. Figure 3a shows the number of predictors with nonzero weight. At least 9 predictors are used throughout most of the central and eastern United States, with some locations using up to 18 predictors. On average, fewer predictors are used in the western United States, but most places still use at least 6 predictors.

Figures 3b–f show the relative weight of the SPI, ESI, and 10-, 100-, and 200-cm NLDAS soil moisture variables, respectively. For this comparison, the contribution of all time periods (4, 8, 12, 16, 20, 26, 39, and 52 weeks) is aggregated for each variable. Because all predictors are standardized and all coefficients are positive, the relative weight for each variable is calculated by summing over the nonzero coefficients from the NNLS. Overall, the SPI has similar weight over most of the United States (Fig. 3b), whereas the other four variables have various regions where they tend to perform well. For example, the ESI has highest weight in Florida where the thermal LST signal is picking up ancillary sources of moisture due to groundwater interactions that are not well represented in NLDAS (Hain et al. 2015). The 100-cm NLDAS soil moisture performs best in a relatively narrow corridor in the central United States and in the Pacific coast states, while soil moisture at 200 cm tends to dominate over the semiarid regions of the central and western United States, where moisture in the soil surface layer is typically quite low. The lower weight assigned to ESI is expected because USDM assessments to date have relied heavily on precipitation and soil moisture indicators. This weighting may evolve in the future as ET becomes more directly incorporated in USDM process to better capture rapid onset (flash) drought events.

There are significant regional differences in the weighting patterns in Fig. 3; however, we believe much of this structure is not based on actual physical differences between regions. The source of the structure is likely due to the fact that the different predictor fields (i.e., SPI and soil moisture) are quite highly correlated, and therefore multicollinearity is an issue. In this case, the individual regression coefficients can change dramatically in response to small changes in the data. Fortunately, while multicollinearity affects the interpretation of the individual predictors, it does not tend to affect the predictive power of the model when regularization (i.e., NNLS) is used.

If the predictors are divided based on time scale, however, then interesting physical structure emerges. For example, when the predictors are divided into short-term (<16 weeks) and long-term (>20 weeks) periods (Figs. 3g and 3h, respectively), a more physically based pattern emerges. For example, the short-time-scale predictors dominate in the central United States where drought conditions can develop very rapidly (e.g., Otkin et al. 2013), whereas the long-time-scale predictors dominate in the western United States where annual-time-scale processes are most important. One should think of the time-scale distinction above as the "average" drought behavior. In reality, short- and long-term droughts can occur anywhere, and our methodology simply determines the typical drought time scale for each region (or combination of drought time scales for cases where both short- and long-term predictors are given significant weight).

c. *Probabilistic skill*

In this subsection, probabilities predicted by the statistical model are evaluated using two verification metrics, including the BSS and reliability diagrams. Note that by "prediction" we mean the estimate of the current state of USDM from SPI, ESI, and NLDAS. In this context, prediction should not be confused with "future forecasts," which will be evaluated in Part II. The BSS measures the difference between the predicted probability of a particular outcome and the actual observed outcome relative to climatology (Wilks 2011). A value of one (zero) indicates perfect (no) skill. The predicted probability of a particular USDM category is given by (4). Cross-validated maps of the BSS for each of the six dryness/drought categories are shown in Fig. 4. Areas in white mean that there are no recorded instances of that

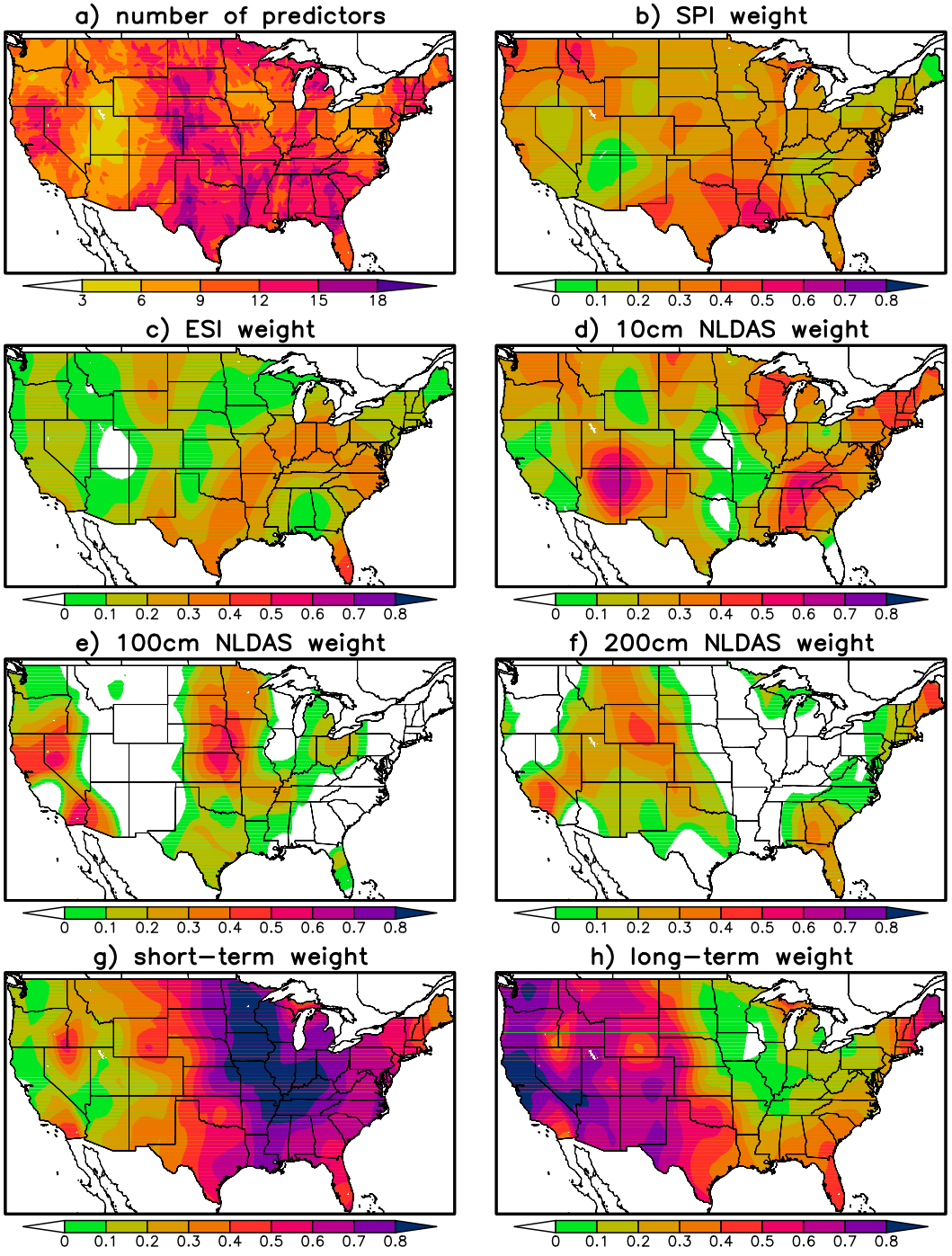

FIG. 3. (a) The number of predictors with nonzero weights using the NNLS scheme. (b) Sum of the weights involving SPI. (c) Sum of the weights involving ESI. (d) Sum of the weights involving the 0–10-cm soil moisture from NLDAS. (e) As in (d), but for the top 0–100 cm. (f) As in (d), but for the top 0–200 cm. (g) Sum of the weights for predictors composited over 4, 8, 12, and 16 weeks. (h) Sum of the weights for predictors composited over 20, 26, 39, and 52 weeks.

drought category during 2001–14. In most places, the no drought category has the best BSS. The lowest skill occurs for the abnormally dry category and then increases for more intense droughts, with the exceptional drought category exhibiting the highest skill of the drought categories. To understand the dependence of the BSS on the drought category, first note that the continuous USDM can potentially take any real number and

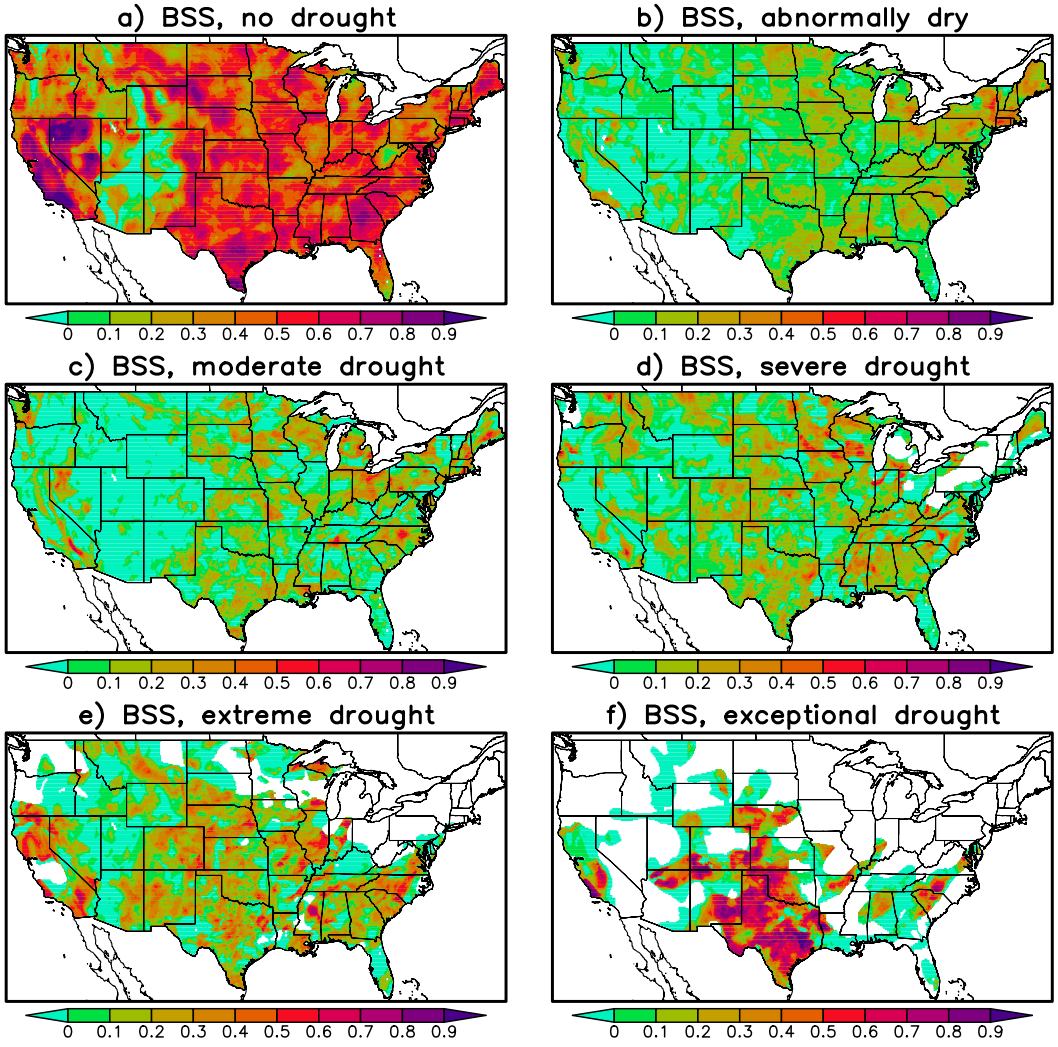

FIG. 4. Cross-validated BSS for the probabilities of (a) no drought, (b) abnormally dry, (c) moderate drought, (d) severe drought, (e) extreme drought, and (f) exceptional drought.

therefore the no drought and exceptional drought categories, which are at the two ends of the USDM distribution, include an infinite portion of the real line. The intermediate drought categories, on the other hand, include a relatively narrow range of values of the continuous USDM. As an example, consider Fig. 1a, which shows the spacing of the USDM thresholds for a continuous USDM that is normally distributed. Because the width of the conditional PDF is independent of the predictors, predictions of the intermediate drought categories are therefore inherently less predictable than the ends of the distribution. Similarly, the no drought category displays the best skill because it includes the widest range of real numbers from normal to extremely wet.

To demonstrate that the above ideas are the source of the BSS differences, assume that our conditional PDF of USDM is in fact the true PDF of the USDM given the predictors. We then randomly sample the PDF to make a synthetic time series of the USDM that is perfectly consistent with our statistical model and then compute the BSS for the synthetic time series. The domain-average BSS for each drought category for the observed USDM and for the synthetic USDM is shown in Fig. 5. As expected, because the synthetic data are perfectly consistent with our predicted PDF, their BSS is elevated. The interesting result is that the observed dependence of the BSS on drought category is very well captured by the synthetic data demonstrating that the low BSS for the intermediate values is not a result of a poorer fit to the data. Instead, the system is inherently less predictable for the intermediate drought categories because the spacing of the USDM thresholds (30th,

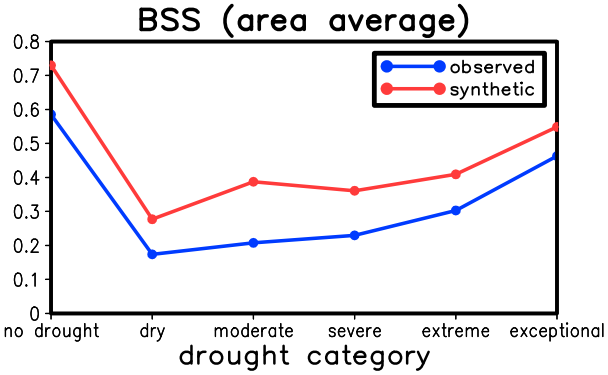

FIG. 5. Domain-averaged BSS as a function of drought category for the cross-validated BSS in Fig. 4 (blue) and the synthetic USDM perfectly consistent with the statistical model (red).

20th, 10th, 5th, and 2nd percentiles) implies that intermediate drought categories are narrower in terms of absolute SPI, ESI, and soil moisture anomalies.

Next, we measure the skill of the estimates using reliability diagrams (Wilks 2011). A "reliable" estimate means that, given a large number of individual cases when the continuous USDM assigns a 60% probability of moderate drought, the actual USDM will be in the moderate drought category in 60% of these cases and will be in some other drought category in 40% of these cases. The reliability diagram is calculated by first ranking the time series of predicted probabilities of drought category d from smallest to largest. Next, the ranked probabilities are divided into n (nearly) equally populated bins. For each of these n bins, the average predicted probability is computed as well as the actual observed fraction of cases in drought category d. Finally, these n pairs of numbers are plotted on a two-dimensional scatter diagram together with the line $y = x$. For a perfectly reliable forecast the predicted and observed probabilities are equal, and therefore the n points on the scatter diagram would lie on the line $y = x$. The different colors denote the reliability diagrams for the six distinct drought categories. For these plots, we choose to divide the data into 20 bins ($n = 20$). To help reduce noise, we average the individual reliability diagrams at each grid point over the northwest, southwest, northeast, and southeast portions of our domain.

Figure 6 shows reliability diagrams for data that have not been cross validated, separated into four regions corresponding to the northwest, southwest, northeast, and southeast parts of the United States. This allows us to examine the maximum reliability of the method because, in this case, the model is trained on the same data used to calculate the reliability diagram. The x axis is the predicted probability and the y axis is the observed probability of occurrence. In general, the reliability diagrams are quite good, with the scatter following the diagonal line for most categories and regions and with the highest reliability in the eastern United States. The reliability for the no drought category is especially good.

The largest deviations from the line $y = x$ are for exceptional drought in the northwest section of the domain (Fig. 6a). In this case, when the prediction says that there is an x percent chance of exceptional drought, the actual chance is significantly smaller than x. The predicted probabilities for moderate drought also tend to be too large in all four regions. Perhaps the thresholds that bound moderate drought (in the statistical model) should be brought closer together so that moderate drought is less likely. Note this threshold bias is also consistent with the fact that the synthetic BSS is especially large compared to the observed BSS for moderate drought (Fig. 5). There is also a smaller tendency for the opposite bias for abnormally dry, particularly in the east.

Figure 7 shows the reliability diagram computed using cross-validated data. Compared to Fig. 6, the most noticeable difference is the downward-curving scatter for some categories as the predicted probabilities increased, particularly for the west, which indicates that the statistical model is overconfident. Note that this bias is only large for the intermediate drought categories that include relatively small portions of the real line (see BSS discussion). More work is necessary to identify the source of this bias. One possibility is the changing character of the USDM analyses during the 2001–14 period. For example, in earlier years, the USDM-depicted drought severity varied slowly in space and typically only showed broad-scale drought features; however, in recent years, it has the tendency to depict much-finer-scale features and to more closely follow fine-scale structures in the precipitation field. Perhaps this nonstationary behavior of the USDM is a source of the biased reliability diagram. Another source of nonstationary statistics is the subjective aspect of the USDM and the fact that different experts take turns producing the maps.

Another potential source of bias is that the USDM distinguishes between short- and long-term drought, which can have different characteristics in space and time. It is possible that including two separate USDM distributions, one with short-term predictors and another with long-term predictors, would improve the reliability of the cross-validated data. In this two-tiered model, the predicted drought state would be the more extreme of the two individual components. These and other strategies to improve the model are the subject of current and future work.

d. Examples

Examples showing comparisons of the observed and predicted USDM analyses are shown in this section. The

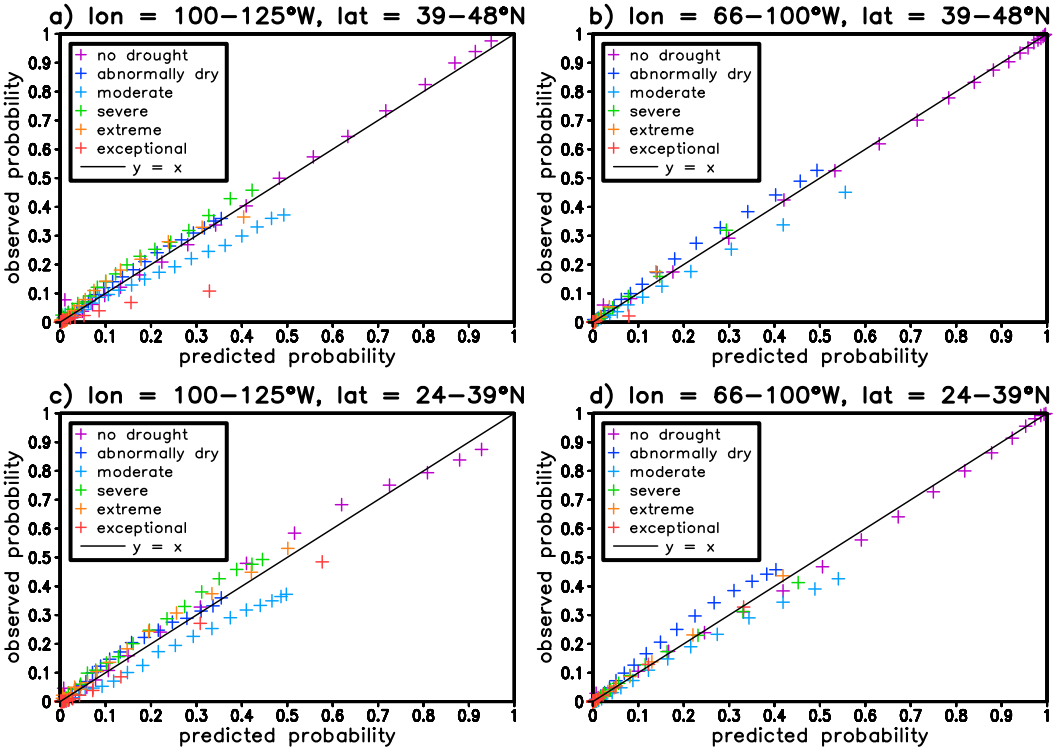

FIG. 6. (a) Reliability diagram for the probabilistic USDM predictions for each drought category (colors). The x axis is the predicted probability and the y axis is the probability from the observed USDM. For an ideal fit the points should lie on the line $y = x$ (black). To reduce noise, the probabilities are averaged over the northwest quarter of the United States. (b) As in (a), but for the northeast. (c) As in (a), but for the southwest. (d) As in (a), but for the southeast.

predicted PDF of the USDM is distilled to a single number using the least squares error method described in section 3d. All results shown here are cross validated, meaning that the statistical model was fit on data from all years besides the one shown. Figures 8–10 show comparisons during the beginning, middle, and end of the growing season for the last 9 years of the period of record (2006–14). The spatial correlations between the observed and predicted USDM states are shown in the title of the prediction panels. Overall, the predictions accurately reproduce the large-scale drought patterns depicted by the USDM; however, the finer details sometimes disagree. In general, the correlations are highest when large portions of the domain are in more intense drought categories. For example, May 2011, May 2013, and May 2014 have the most exceptional drought coverage and also have the highest correlations at 0.93, 0.86, and 0.89, respectively (Fig. 8), with similar behavior evident during mid-July (Fig. 9) and mid-September (Fig. 10). In addition, as the drought in 2012 intensified in July and September, the spatial correlations improved to 0.83 and 0.81, respectively, from 0.70. This fact is also consistent with the idea that spatial correlation improves as the area of intense drought increases. In contrast, the worst spatial correlations (<0.5) occur in July and September 2010, when almost the entire country was drought-free or only in minor drought. This is consistent with the BSS analysis that showed that the intermediate drought categories are the least skillful because of the relatively narrow range.

5. Summary and discussion

In this paper, a statistical method for producing probabilistic estimates of the current drought state as predicted by the USDM was presented. The method uses anomalies in precipitation (SPI), evapotranspiration (ESI), and soil moisture (NLDAS) over different time scales as predictors to estimate the current drought status with high spatial and temporal resolution. The method creates a continuous version of the USDM that is optimally consistent with the actual, discrete USDM state. The discrete USDM can be reasonably reconstructed from the continuous USDM function by discretizing the function into six dryness/drought categories based on five thresholds that define the boundaries of the

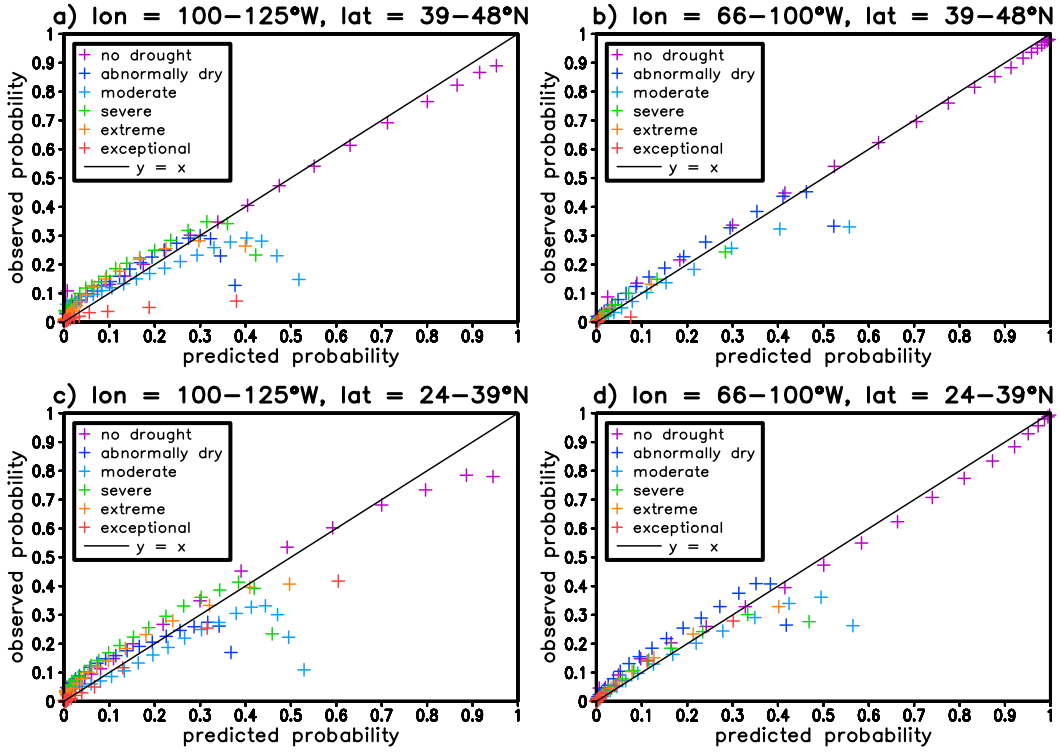

FIG. 7. As in Fig. 6, but for the cross-validated probabilities.

dryness/drought categories. The values of the SPI, ESI, and NLDAS predictors at any given time are used to estimate the PDF of the continuous USDM. To make an explicit statistical model that can be fit to observed data, the following assumptions of the PDF of the USDM conditioned on the SPI, ESI, and NLDAS predictors were made: 1) the mean of the PDF is linearly related to the predictors; 2) the shape and width of the PDF is independent of the predictors; 3) the functional form of the PDF is the Student's t distribution; and 4) the thresholds defining the drought categories are the 30th, 20th, 10th, 5th, and 2nd percentiles of the Student's t distribution. These assumptions are enough to explicitly define the likelihood function for the statistical model. By maximizing the likelihood function, we determine the parameters of the statistical model. Once the conditional PDF of the continuous USDM is found, we can calculate the probabilities for each of the six USDM categories by integrating the PDF between the thresholds that define the drought categories. To more easily compare with the observed USDM, we also developed a scheme to summarize the full PDF with a single least squares best guess of the current USDM.

The cross-validated temporal correlation between the best guess prediction and the observed USDM averages 0.79 over central and eastern United States (east of 105°W) and 0.65 over the western United States. To better compare with other studies, the corresponding temporal correlations without cross validation are 0.85 and 0.78. The cross-validated spatial correlations exceed 0.9 when relatively large regions of the United States are in exceptional drought. Example comparisons between the predictions and the actual USDM show very good agreement over large scales but less agreement on finescale features. For most probabilities, the predictions are reliable, meaning that when the chance of the extreme drought category is 30%, for example, then the actual probability of extreme drought is in fact 30%. The BSS of the probabilistic predictions are best for the no drought and exceptional drought categories and less good for the intermediate drought categories. As described above, our continuous USDM formalism provides a straightforward explanation for the lower skill of the intermediate drought categories.

The method is a very promising objective drought classification technique. Because it is empirically tuned to the USDM, this method can capture regional differences in the processes most relevant for drought impacts. However, because the USDM and this new index are composite indices, they will sometimes identify droughts that do not affect all sectors and/or time scales. Nevertheless, the USDM might be the best available

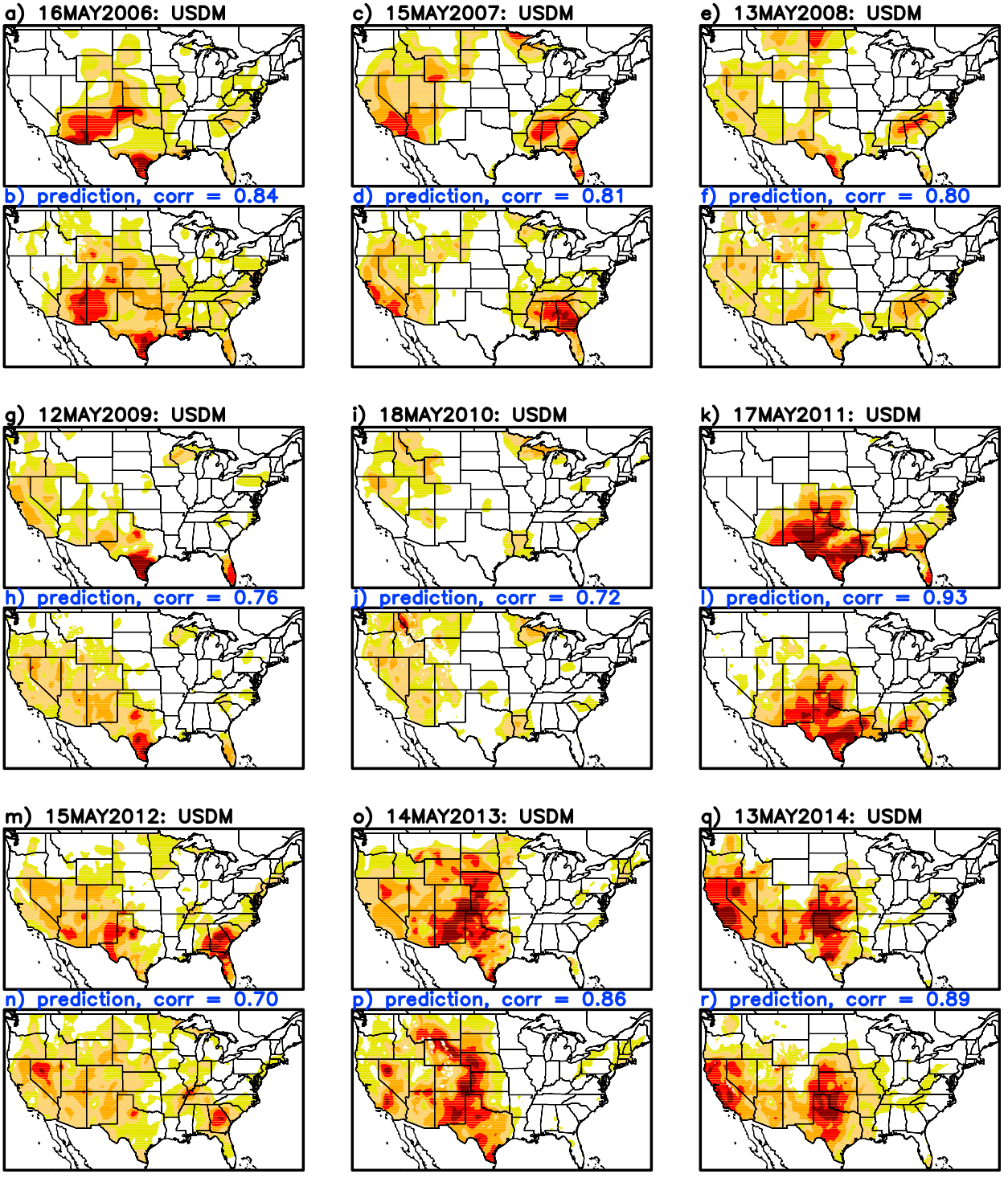

FIG. 8. Comparison of USDM and the cross-validated predictions (i.e., the model has not "seen" the year shown) for mid-May of the latest 9 years: (a) USDM and (b) prediction on 16 May 2006. The spatial correlation between USDM and the prediction is shown in the title of the prediction panel (b). (c),(d) As in (a) and (b), but for 15 May 2007. (e),(f) As in (a) and (b), but for 13 May 2008. (g),(h) As in (a) and (b), but for 12 May 2009. (i),(j) As in (a) and (b), but for 18 May 2010. (k),(l) As in (a) and (b), but for 17 May 2011. (m),(n) As in (a) and (b), but for 15 May 2012. (o),(p) As in (a) and (b), but for 14 May 2013. (q),(r) As in (a) and (b), but for 13 May 2014.

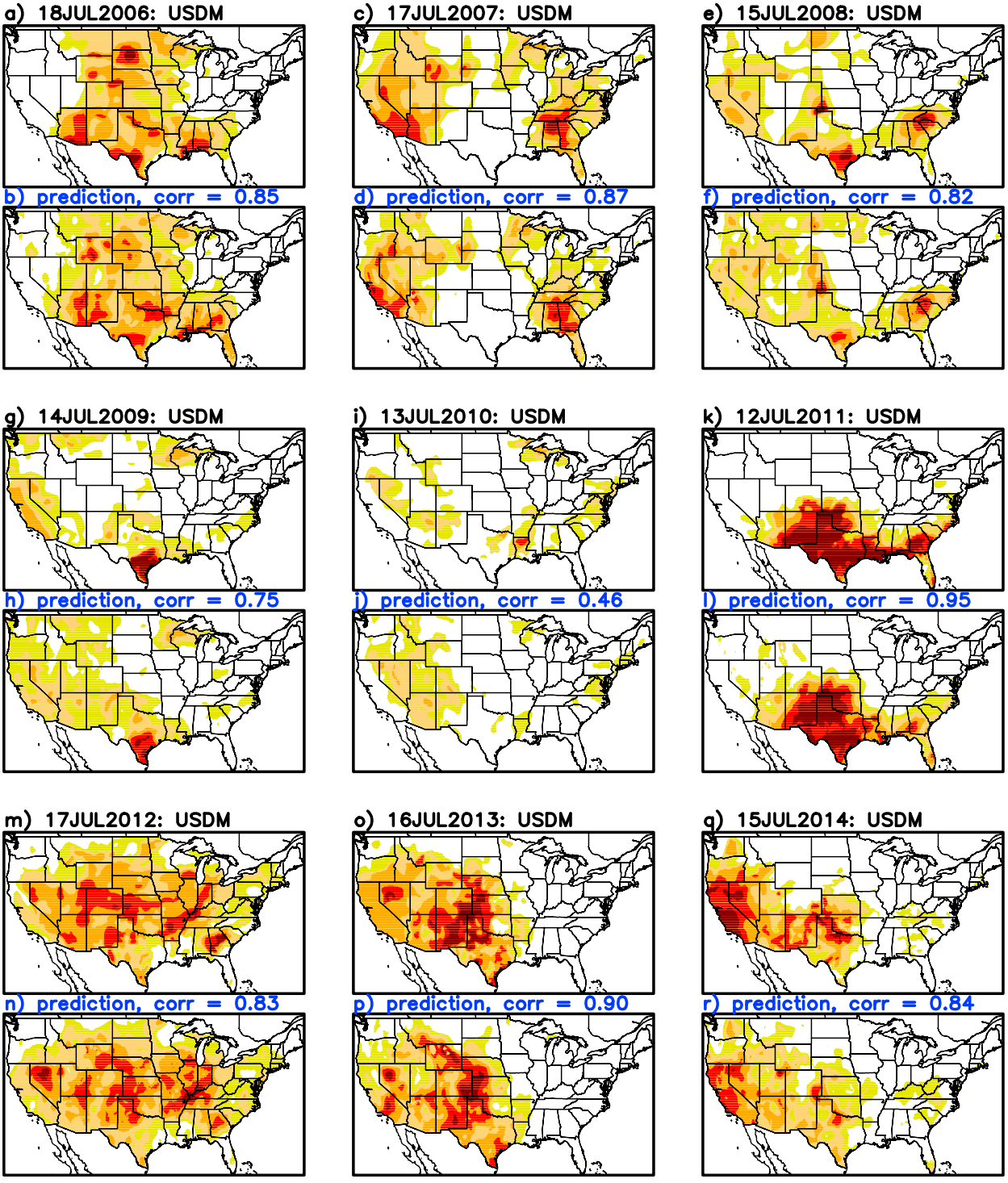

FIG. 9. As in Fig. 8, but for mid-July.

indicator that summarizes the relevant information into a single map. The development and incorporation of gridded fields of other variables such as crop conditions, streamflow, and mountain snowpack would likely lead to further improvements in the method and results. The latter two fields might be particularly beneficial in the western United States, where skill tends to be less than the east using the current predictors.

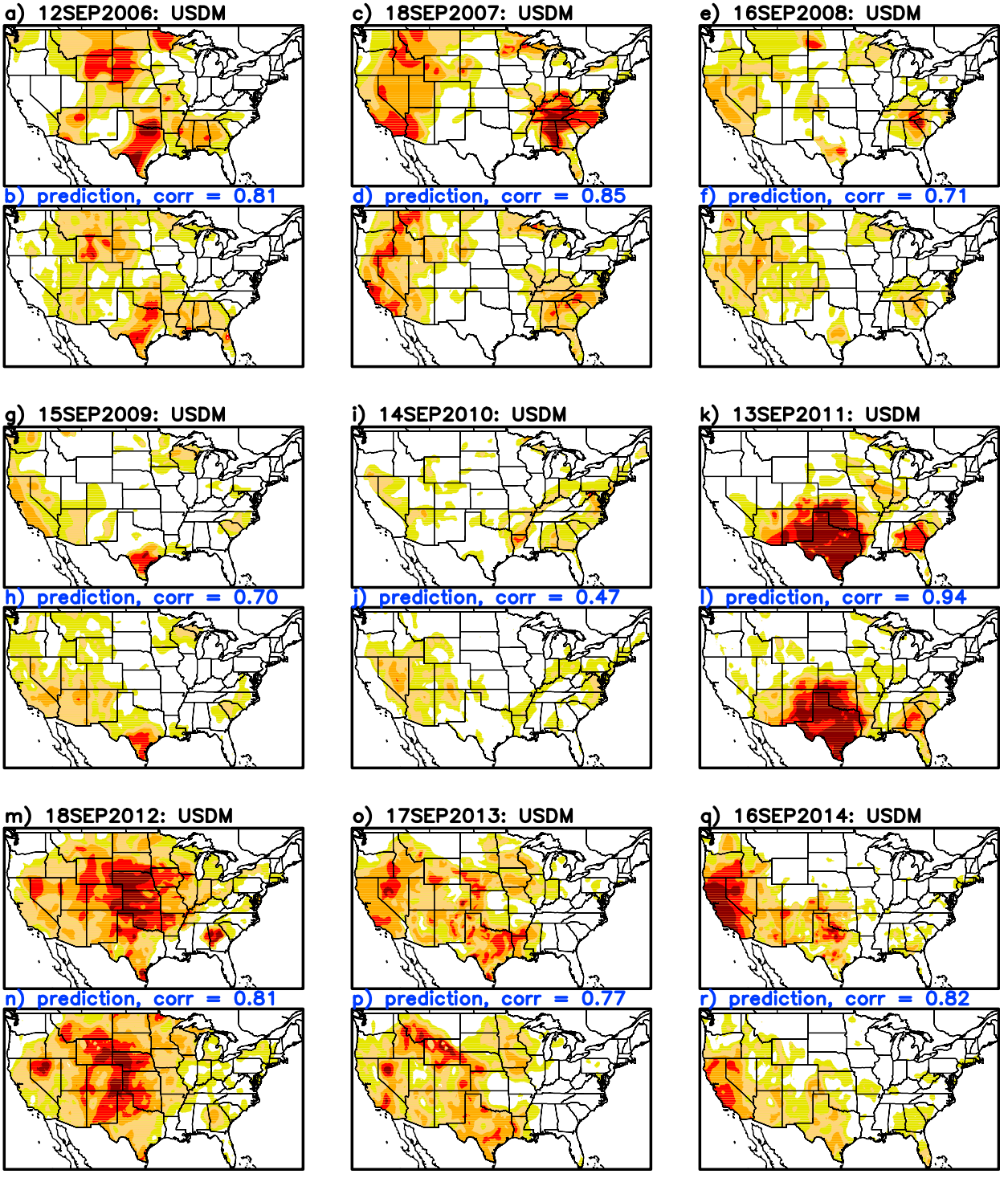

FIG. 10. As in Fig. 8, but for mid-September.

In Part II, the USDM state predictions described in this paper are used in combination with recent precipitation, ESI, and NLDAS soil moisture anomalies to predict future changes in the USDM over intraseasonal time scales (2–8 weeks). The USDM state predictions developed here are useful for predicting future changes in the USDM because 1) they quantify how close or far adjacent drought categories are from the current state

and 2) the USDM is more likely to change categories when conditions are "almost" in another category. In future work, this information will also be combined with data from the North American Multi-Model Ensemble to make future USDM predictions.

Acknowledgments. This work was supported by funds provided by the NOAA Climate Program Office's Modeling, Analysis, Predictions, and Projections (MAPP) program under Grant NA14OAR4310226.

APPENDIX

Maximizing the Likelihood

The values of parameters of the statistical model are those that maximize the likelihood function (5). Provided the likelihood is not zero, the maximum of the logarithm of the likelihood function occurs at the same location as the likelihood. Therefore, as is standard practice, we will maximize the log likelihood:

$$\log(L) = \sum_j \log\{F(s[T_{d_j} - \mu_j]) - F(s[T_{d_j-1} - \mu_j])\}, \tag{A1}$$

where j is an index over time, F is the CDF of the Student's t distribution with 10 degrees of freedom, μ_j depends on time because it is a linear combination of the predictors (see below), d_j is the current USDM category, T are the thresholds that define the boundaries of the drought categories for the continuous "version" of the USDM, and s is inversely proportional to the width of the distribution. The mean μ_j can be written explicitly in terms of the ith predictor p_{ij} as

$$\mu_j = a_0 + \sum_i a_i p_{ij}, \tag{A2}$$

where a_i ($i = 0, 1, 2, 3, \ldots$) are parameters to be fit by maximizing the log likelihood. The only other parameter to fit is s. We maximize the likelihood function using the variable metric algorithm in Press et al. (1992). This algorithm requires the gradient of the function with respect to each of the independent variables. In our case, the function is the log likelihood and the independent variables are the a_i and s. The variables a_i only enter into the likelihood through μ_j, so the chain rule can be used:

$$\frac{d\log(L)}{da_i} = \sum_j \frac{-sf(s[T_{d_j} - \mu_j]) + sf(s[T_{d_j-1} - \mu_j])}{F(s[T_{d_j} - \mu_j]) - F(s[T_{d_j-1} - \mu_j])} \frac{d\mu_j}{da_i}, \tag{A3}$$

where we use the fact that the derivative of the CDF F is the PDF f. The a_i gradient terms are completed by noting that

$$\frac{d\mu_j}{da_i} = \begin{cases} 1, & \text{if } i = 0; \text{ and} \\ p_{ij}, & \text{otherwise} \end{cases}. \tag{A4}$$

Finally, the gradient with respect to s is

$$\frac{d\log(L)}{ds} = \sum_j \frac{(T_{d_j} - \mu_j) f(s[T_{d_j} - \mu_j]) - (T_{d_j-1} - \mu_j) f(s[T_{d_j-1} - \mu_j])}{F(s[T_{d_j} - \mu_j]) - F(s[T_{d_j-1} - \mu_j])}. \tag{A5}$$

REFERENCES

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, 300 pp. [Available online at www.fao.org/docrep/X0490E/X0490E00.htm.]

Anderson, M. C., J. M. Norman, G. R. Diak, W. P. Kustas, and J. R. Mecikalski, 1997: A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.*, **60**, 195–216, doi:10.1016/S0034-4257(96)00215-5.

——, W. P. Kustas, and J. M. Norman, 2007a: Upscaling flux observations from local to continental scales using thermal remote sensing. *Agron. J.*, **99**, 240–254, doi:10.2134/agronj2005.0096S.

——, J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, 2007b: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. *J. Geophys. Res.*, **112**, D10117, doi:10.1029/2006JD007506.

——, C. Hain, J. A. Otkin, X. Zhan, K. Mo, M. Svoboda, B. Wardlow, and A. Pimstein, 2013: An intercomparison of drought indicators based on thermal remote sensing and NLDAS simulations. *J. Hydrometeor.*, **14**, 1035–1056, doi:10.1175/JHM-D-12-0140.1.

AON Benfield, 2013: Annual global climate and catastrophe report: Impact forecasting—2012. Tech Rep., 96 pp. [Available online at http://thoughtleadership.aonbenfield.com/Documents/20130124_if_annual_global_climate_catastrophe_report.pdf.]

Barlage, M., and Coauthors, 2010: Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains. *J. Geophys. Res.*, **115**, D22101, doi:10.1029/2009JD013470.

Bowling, L. C., and D. P. Lettenmaier, 2010: Modeling the effects of lakes and wetlands on the water balance of Arctic environments. *J. Hydrometeor.*, **11**, 276–295, doi:10.1175/2009JHM1084.1.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.

Fannin, B., 2012: Updated 2011 Texas agricultural drought losses total $7.62 billion. Texas A&M Agrilife, accessed 2015. [Available online at http://today.agrilife.org/2012/03/21/updated-2011-texas-agricultural-drought-losses-total-7-62-billion/.]

Franc, V., V. Hlaváč, and M. Navara, 2005: Sequential coordinate-wise algorithm for the non-negative least squares problem. *CAIP 2005: Computer Analysis of Images and Patterns*, Lecture Notes in Computer Science, Vol. 3691, Springer, 407–414, doi:10.1007/11556121_50.

Hacker, R., 1962: Certification of algorithm 112: Position of point relative to polygon. *Commun. ACM*, **5**, 606, doi:10.1145/355580.369118.

Hain, C. R., W. T. Crow, M. C. Anderson, and M. T. Yilmaz, 2015: Diagnosing neglected soil moisture source–sink processes via a thermal infrared–based two-source energy balance model. *J. Hydrometeor.*, **16**, 1070–1086, doi:10.1175/JHM-D-14-0017.1.

Hao, Z., Y. Hong, Y. Xia, V. P. Singh, F. Hao, and H. Cheng, 2016: Probabilistic drought characterization in the categorical form using ordinal regression. *J. Hydrol.*, **535**, 331–339, doi:10.1016/j.jhydrol.2016.01.074.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center Atlas 7, NOAA, 40 pp. [Available online at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/7/toc.html.]

Howitt, R. D. MacEwan, J. Medellín-Azuara, J. Lund, and D. Sumner, 2015: Economic analysis of the 2015 drought for California agriculture. University of California, Davis Tech. Rep., 16 pp. [Available online at https://watershed.ucdavis.edu/files/biblio/Final_Drought%20Report_08182015_Full_Report_WithAppendices.pdf.]

Hunt, E., M. Svoboda, B. Wardlow, K. Hubbard, M. J. Hayes, and T. Arkebauer, 2014: Monitoring the effects of rapid onset of drought on non-irrigated maize with agronomic data and climate-based drought indices. *Agric. For. Meteor.*, **191**, 1–11, doi:10.1016/j.agrformet.2014.02.001.

Koster, R. D., and M. J. Suarez, 1994: The components of a SVAT scheme and their effects on a GCM's hydrological cycle. *Adv. Water Resour.*, **17**, 61–78, doi:10.1016/0309-1708(94)90024-8.

——, and ——, 1996: Energy and water balance calculations in the Mosaic LSM. NASA Tech. Memo. 104606, Vol. 9, 60 pp. [Available online at http://gmao.gsfc.nasa.gov/pubs/docs/Koster130.pdf.]

Liang, X., E. F. Wood, and D. P. Lettenmaier, 1996: Surface and soil moisture parameterization of the VIC-2L model: Evaluation and modifications. *Global Planet. Change*, **13**, 195–206, doi:10.1016/0921-8181(95)00046-1.

Lorenz, D. J., J. A. Otkin, M. Svoboda, C. R. Hain, M. C. Anderson, and Y. Zhong, 2017: Predicting the U.S. Drought Monitor using precipitation, soil moisture, and evapotranspiration anomalies. Part II: Intraseasonal drought intensification forecasts. *J. Hydrometeor.*, **18**, 1963–1982, doi:10.1175/JHM-D-16-0067.1.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scale. Preprints, *Eighth Conf. on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184.

——, ——, and ——, 1995: Drought monitoring with multiple time scales. Preprints, *Ninth Conf. on Applied Climatology*, Dallas, TX, Amer. Meteor. Soc., 233–236.

McNaughton, K. G., and T. W. Spriggs, 1986: A mixed-layer model for regional evaporation. *Bound.-Layer Meteor.*, **34**, 243–262, doi:10.1007/BF00122381.

Meinshausen, N., 2013: Sign-constrained least squares estimation for high-dimensional regression. *Electron. J. Stat.*, **7**, 1607–1631, doi:10.1214/13-EJS818.

Mozny, M., M. Trnka, Z. Zalud, P. Hlavinka, J. Nekovar, V. Potop, and M. Virag, 2012: Use of a soil moisture network for drought monitoring in the Czech Republic. *Theor. Appl. Climatol.*, **107**, 99–111, doi:10.1007/s00704-011-0460-6.

Myneni, R. B., and Coauthors, 2002: Global products of vegetation leaf area and fraction absorbed by PAR from year of MODIS data. *Remote Sens. Environ.*, **83**, 214–231, doi:10.1016/S0034-4257(02)00074-3.

Norman, J. M., W. P. Kustas, and K. S. Humes, 1995: A two-source approach for estimating soil and vegetation energy fluxes from observations of directional radiometric surface temperature. *Agric. For. Meteor.*, **77**, 263–292, doi:10.1016/0168-1923(95)02265-Y.

NRDC 2013: Record-breaking $17.3 billion in crop losses last year; Significant portion potentially avoidable. Natural Resources Defense Council, accessed 25 May 2017. [Available online at https://www.nrdc.org/media/2013/130827.]

Otkin, J. A., M. C. Anderson, C. Hain, I. Mladenova, J. Basara, and M. Svoboda, 2013: Examining flash drought development using the thermal infrared based evaporative stress index. *J. Hydrometeor.*, **14**, 1057–1074, doi:10.1175/JHM-D-12-0144.1.

——, ——, ——, and M. Svoboda, 2014: Examining the relationship between drought development and rapid changes in the evaporative stress index. *J. Hydrometeor.*, **15**, 938–956, doi:10.1175/JHM-D-13-0110.1.

——, ——, ——, and ——, 2015: Using temporal changes in drought indices to generate probabilistic drought intensification forecasts. *J. Hydrometeor.*, **16**, 88–105, doi:10.1175/JHM-D-14-0064.1.

——, and Coauthors, 2016: Assessing the evolution of soil moisture and vegetation conditions during the 2012 United States flash drought. *Agric. For. Meteor.*, **218–219**, 230–242, doi:10.1016/j.agrformet.2015.12.065.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1992: *Numerical Recipes in Fortran 77: The Art of Scientific Computing.* 2nd ed. Cambridge University Press, 933 pp.

Saha, S., and Coauthors, 2010: The NCEP Climate System Forecast Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.

Shimrat, M., 1962: Algorithm 112: Position of point relative to polygon. *Commun. ACM*, **5**, 434, doi:10.1145/368637.368653.

Slawski, M., and M. Hein, 2013: Non-negative least squares for high-dimensional linear models: Consistency and sparse recovery without regularization. *Electron. J. Stat.*, **7**, 3004–3056, doi:10.1214/13-EJS868.

Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181–1190, doi:10.1175/1520-0477(2002)083<1181:TDM>2.3.CO;2.

Wei, H., Y. Xia, K. E. Mitchell, and M. B. Ek, 2013: Improvement of the Noah land surface model for warm season processes: Evaluation of water and energy flux simulation. *Hydrol. Processes*, **27**, 297–303, doi:10.1002/hyp.9214.

Wilhite, D. A., and M. H. Glantz, 1985: Understanding the drought phenomenon: The Role of definitions. *Water Int.*, **10**, 111–120, doi:10.1080/02508068508686328.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences.* 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

Xia, Y., M. B. Ek, H. Wei, and J. Meng, 2012a: Comparative analysis of relationships between NLDAS-2 forcings and model outputs. *Hydrol. Processes*, **26**, 467–474, doi:10.1002/hyp.8240.

——, and Coauthors, 2012b: Continental-scale water and energy flux analysis and validation of the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, doi:10.1029/2011JD016048.

——, J. Sheffield, M. B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E. F. Wood, 2014: Evaluation of multi-model simulated soil moisture in NLDAS-2. *J. Hydrol.*, **512**, 107–125, doi:10.1016/j.jhydrol.2014.02.027.