Predicting the U.S. Drought Monitor Using Precipitation, Soil Moisture, and Evapotranspiration Anomalies. Part II: Intraseasonal Drought Intensification Forecasts

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

Probabilistic forecasts of U.S. Drought Monitor (USDM) intensification over 2-, 4-, and 8-week time periods are developed based on recent anomalies in precipitation, evapotranspiration, and soil moisture. These statistical forecasts are computed using logistic regression with cross validation. While recent precipitation, evapotranspiration, and soil moisture do provide skillful forecasts, it is found that additional information on the current state of the USDM adds significant skill to the forecasts. The USDM state information takes the form of a metric that quantifies the ''distance'' from the next-higher drought category using a nondiscrete estimate of the current USDM state. This adds skill because USDM states that are close to the next-higher drought category are more likely to intensify than states that are farther from this threshold. The method shows skill over most of the United States but is most skillful over the north-central United States, where the cross-validated Brier skill score averages 0.20 for both 2- and 4-week forecasts. The 8-week forecasts are less skillful in most locations. The 2- and 4-week probabilities have very good reliability. The 8-week probabilities, on the other hand, are noticeably overconfident. For individual drought events, the method shows the most skill when forecasting high-amplitude flash droughts and when large regions of the United States are experiencing intensifying drought.

1. Introduction

Drought can impact the health and diversity of natural ecosystems and severely reduce agricultural output in regions that depend on rain-fed crops and forage. Because drought can take many forms through variations in intensity and longevity (Wilhite and Glantz 1985), its impact can also vary significantly for different socioeconomic groups. For example, severe drought that is short in duration but occurs during critical crop yield development stages can have a large impact on farm production, yet have minimal impact on other stakeholder groups that are less sensitive to short-term drought. Even within the agricultural sector, impacts can differ greatly because of varying exposure to drought due to differences in soil type, crops, and agricultural practices, among other factors. Extreme flash drought events that rapidly develop over short time periods (e.g., Svoboda et al. 2002; Mozny et al. 2012; Otkin et al. 2013; Hunt et al. 2014), such as those that occurred over large areas of the central United States in 2011 and 2012, can have an especially large impact because there is less time to prepare for its adverse effects. Robust drought early warning systems capable of supporting flash drought mitigation efforts necessitate the development of methods that can produce reliable subseasonal forecasts with frequent updates (Otkin et al. 2015b).

Most drought forecasting systems, however, typically provide seasonal forecasts that are updated each month and thus do not provide sufficient temporal resolution to capture the onset and intensification of flash drought.

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

DOI: 10.1175/JHM-D-16-0067.1

This limits their utility for drought mitigation because it can diminish the ability of vulnerable groups to implement proactive measures in a timely manner to lessen the detrimental impacts of drought (e.g., Otkin et al. 2015b). Drought forecasts produced using output from coupled ocean–land–atmosphere general circulation models (e.g., Wood et al. 2002; Luo et al. 2007; Quan et al. 2012; Yuan et al. 2013a,b; Bell et al. 2013; Pan et al. 2013; Dutra et al. 2014; McEvoy et al. 2016) provide valuable information. For example, outputs from the North American Multi-Model Ensemble (Kirtman et al. 2014) have been used to forecast future drought (Yuan and Wood 2013; Mo and Lyon 2015; Thober et al. 2015). However, this guidance has typically been for seasonal time scales. To help address short-range forecasting needs, the Climate Prediction Center (CPC) issues a Monthly Drought Outlook product, but it is only updated at monthly intervals. Statistical methods have also been developed to predict future drought conditions over different time scales (e.g., Steinemann 2003; Kim et al. 2003; Mishra and Desai 2005, 2006; Sen and Boken 2005; Barros and Bowden 2008; Hwang and Carbone 2009; Özger et al. 2012).

A promising approach to produce subseasonal drought predictions (e.g., from weeks to several months) is to use temporal tendencies in drought indices sensitive to precipitation, temperature, soil moisture, or evapotranspiration to identify regions with an increased risk for drought development. Otkin et al. (2014, 2015a) showed that drought was more likely to develop over subseasonal time scales when a rapid change index (RCI), designed to capture the rate of moisture stress change over short time periods, indicated that moisture stress was rapidly increasing. Case study and climatological analyses demonstrated that this relationship was strongest across the central and eastern United States in regions where flash droughts occur more frequently. For example, during the 2012 flash drought, the RCI became strongly negative across portions of the central United States more than a month before the U.S. Drought Monitor (USDM; Svoboda et al. 2002; Otkin et al. 2014) indicating that drought was rapidly intensifying. This event resulted primarily from natural variations in the weather rather than external forcing due to sea surface temperatures, and little to no warning of its rapid onset was evident in traditional drought metrics, drought outlooks, or climate model simulations (Kumar et al. 2013; Hoerling et al. 2014). These results suggest that new forecasting methods based on temporal tendencies of drought indices could augment existing drought monitoring and forecasting systems based on prognostic models.

In this two-part paper, we expand upon the method described in Otkin et al. (2014, 2015a) by using more sophisticated statistical tools to depict the current drought state and to forecast the probabilistic likelihood of future drought development in order to improve the early warning of drought. In this paper (Part II), we develop probabilistic forecasts of drought intensification at 2-, 4-, and 8-week time periods using current trends in precipitation, evapotranspiration, and soil moisture anomalies. A key result of this paper is that incorporation of a continuous depiction (i.e., wetness/dryness is measured on a continuum rather than as six discrete dryness/drought classes) of the current USDM state described in Lorenz et al. (2017, hereafter Part I) greatly increases the forecast skill. We begin with a description of the datasets used in the forecast system (section 2). In section 3, we introduce the statistical methodology, discuss how USDM state information is incorporated as a predictor of the future state, and demonstrate the impact of the USDM state predictor. Results are presented in section 4 and conclusions in section 5.

2. Datasets

a. *Summary of standard predictors*

The purpose of this study is to develop a methodology for forecasting a gridded depiction of the USDM; therefore, we focus on predictors that are available in gridded form. The datasets are the same as those in Part I, and the reader should look there for details. Briefly, the following datasets are used to predict USDM: 1) the CPC gridded analysis of daily precipitation (Higgins et al. 2000); 2) the evaporative stress index (ESI; Anderson et al. 1997, 2007); and 3) model average soil moisture anomalies in the top 10, 100, and 200 cm of the soil profile from phase 2 of the North American Land Data Assimilation System (NLDAS-2; Xia et al. 2012a,b). The data smoothing and the time of year (May–September) are also like that in Part I. Unlike Part I, the "raw" precipitation anomalies are used rather than the standardized precipitation index (SPI) because this provides a slight increase in skill. In addition, because we are predicting changes in the USDM rather than the USDM state as in Part I, weekly anomalies are used here rather than long-term composites ranging from 4 to 52 weeks as in Part I. This is because the time tendency (or change) of a time series with variations at a range of frequencies is necessarily of higher frequency than 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. The one exception to weekly anomalies is ESI, where 4-week composites are used to substantially reduce the number of missing data.

b. USDM state PDFs

In Part I, we developed a probabilistic estimate of the current state of the USDM using the SPI and anomalies in ESI and the NLDAS soil moisture averaged to the 10-, 100-, and 200-cm layers. Each of these five fields was composited over 4-, 8-, 12-, 16-, 20-, 26-, 39-, and 52-week periods before being used as predictors of the current USDM state. The probabilities from this method will be used to forecast the changes in the USDM as described in section 3d. All USDM state probabilities used in this study are cross validated, meaning that the statistical models below do not use USDM state probabilities that were fit on the period used for validation.

3. Methodology

a. Logistic regression

For this study, we develop probabilistic predictions of drought intensification. Because the USDM is discrete, predicting whether USDM intensifies or not over a certain time period involves predicting a discrete yes/no variable. For this study, we only predict whether the USDM intensifies or not—we do not predict the magnitude of the intensification (i.e., the USDM increases by two drought categories). Also, we only predict intensification of drought and not the amelioration of drought. The method in this paper can easily be applied to these alternative "events" by changing the definition of the discrete yes/no variable (i.e., yes when drought intensifies by two or more drought categories, otherwise no). Standard methods such as multilinear regression are not appropriate for predicting yes/no variables because linear regression assumes the errors follow a normal distribution and linear regression does not exclude the possibility of negative probabilities of intensification or probabilities greater than 100%. Instead, the natural statistical model for a discrete yes/no variable is logistic regression. Unlike linear regression, which minimizes the squared error between the predictand y and the weighted sum of the predictors $a_0 + a_1x_1 + a_2x_2 + a_3x_3$ (where x_j are the predictors and a_j are the regression coefficients), logistic regression inserts the weighted sum of predictors into the logistic function to predict the probability p that drought intensification occurs:

$$p = L(a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \cdots), \tag{1}$$

where $L(x)$ is the logistic function:

$$L(x) = \frac{1}{1 + \exp(-x)}. \tag{2}$$

The logistic function goes from 0 to 1 as x goes from $-\infty$ to ∞ and therefore ensures that the probability is between 0 and 1. The parameters $(a_0, a_1, a_2, a_3, \ldots)$ are fit by maximum likelihood using iteratively weighted least squares (Dobson and Barnett 2008).

b. Predictors and predictand

In this study, we predict the probability that the USDM will intensify over 2-, 4-, and 8-week time periods after the current week's USDM value.[1] Consider the 4-week predictions: if the USDM drought category exceeds the current drought category during any point in the next 4 weeks, then the USDM has intensified; otherwise, it has not. The same rule applies to the 2- and 8-week predictions. For the logistic regression, weeks that intensify are given the value 1, and weeks that do not are given the value 0. Therefore, the probabilities from (1) represent the probability of intensification. For all statistics and analysis in this study, we do not use weeks where the USDM is in the most intense "exceptional drought" category $D4$. Instead, we treat such weeks as "missing data." The reason is simply that the prediction is trivial in this case because, by definition, the USDM cannot be more intense.

The predictors include weekly precipitation anomalies from the mean seasonal cycle, the ESI time tendency, and the soil moisture tendencies for three different soil layers: 0–10, 0–100, and 0–200 cm. The tendencies are computed from the difference in adjacent weeks. The ESI and soil moisture tendencies are used because we are forecasting the change in the USDM. Like the USDM, the ESI and soil moisture are variables that measure the state of the land surface; therefore, the change in USDM is analogous to a change in ESI and soil moisture. Precipitation, on other hand, is not a state variable but is a "forcing" variable that directly imposes a change on the state variables. Therefore, the precipitation anomalies are used directly. We also explored treating precipitation as a state variable and ESI and soil moisture as forcing variables, and the skill was degraded, as expected. The RCI of Otkin et al. (2014, 2015a) was also considered, but the straightforward and more parsimonious ESI tendency performed slightly better, so it was chosen.

Note that, unlike the USDM state estimate, which uses wide range of composite periods (see section 2e), weekly composites are used in this current study because the focus is on USDM changes over relatively short time scales.[2]

[1] In other words, our data cutoff is Tuesday and the USDM is released on Thursday. We predict intensification in the following week (and beyond) compared to Thursday's USDM.

[2] To reduce missing data periods, ESI is composited over 4 weeks.

To extend the time scale of the predictors, each of the above fields is used at multiple time lags starting at the present time and including the values for the past 1, 2, and 3 weeks. In other words, there are four precipitation predictors that are staggered in time: the weekly mean precipitation for the most recent time and the weekly mean for each of the preceding weeks. This allows the data to decide the best weighting of the different time lags. For example, the method can empirically determine whether the 1-week average is the best predictor (all regression coefficients are zero except for lag 0), whether a 4-week average is best (the coefficients are equal for all lags), or whether any other combination of weighting is best. Time-lagged predictors are also used for each of the following four fields: 1) the standardized ESI tendency and the NLDAS soil moisture tendencies averaged from the surface to 2) 10, 3) 100, and 4) 200 cm.

There are two additional predictors: one based on the USDM state estimate from Part I and another based on the climatological seasonal cycle of drought intensification. The first additional predictor is described in detail in section 3b. The second, the climatological intensification, is used because some locations have a preferred time of year when intensification is most likely to occur. The calculation of the climatological intensification is performed multiple times so that when we validate the model for 2012, for example, the climatological intensification is calculated only using the other years. To reduce noise, the climatological intensification probability is smoothed in time by a Gaussian kernel of the form: $\exp[-0.5(t/s)^2]$, where t is time and the constant s is 5 weeks. The 5 weeks was chosen because it was the smallest value that appeared to remove most of the noise in the seasonal cycle (estimated visually).

All predictors are standardized prior to the logistic regression so that the size of the regression coefficients can be used to gauge the impact of each predictor. With five fields at four different time lags and two additional predictors, the total number of predictors is 22.

c. Predictor selection

With multiple predictors, avoiding overfitting is essential. The standard approach for determining the important predictors is to incrementally add individual predictors with cross validation until skill decreases on independent data. In Part I, we found that aggregating predictors together into a single "master index" improves skill over the individual predictor approach. In that study, the predictors were linearly combined using weights that were empirically calculated using nonnegative least squares (NNLS) regression. Unlike standard linear regression, the signs of the regression coefficients are constrained to be greater than or equal to zero. The advantage of the NNLS regression is its regularization properties (Meinshausen 2013; Slawski and Hein 2013), which penalize excessive complexity, and therefore NNLS almost always shows better skill on independent data. Moreover, NNLS was easy to apply in Part I because the sign of the true, physically based coefficients are known a priori (if the physically based coefficient is negative, simply multiply the predictor by −1). Note that with many predictors, NNLS regression will result in nonzero weights for a subset of the predictors, and the rest of the weights will be exactly zero. In this respect, it is similar to the standard individual predictor approach described above. In this paper, the NNLS approach of Part I is modified slightly: the nonnegativity constraint is imposed on the coefficients of logistic regression ($a_1, a_2, a_3, \ldots$) rather than on the coefficients of linear regression.[3] Also, in order to avoid artificially inflating forecast skill, the coefficients are calculated with cross validation. First, one year is removed from the data. Second, all model coefficients are fit using data from the remaining years. Next, skill is calculated when applying the model to the year that was left out. Finally, the process is repeated until all years have a chance to be left out.

Part I also found that, given the relatively short USDM record, aggregating surrounding points together when fitting the statistical models improves the fit. While it is true that nearby grid points may not have the same relationship between the various predictors and USDM intensification, the errors from the limited sample (14 years of data) apparently dominate. The validation of the model on the left-out year is still done on the central grid point—it is only when fitting the model that surrounding grid points are used. The surrounding grid points are not given full weight; instead, the weight w of a grid point in the analysis is given by a Gaussian function:

$$w = \exp\left[-\frac{1}{2}\frac{(\Delta x)^2 + (\Delta y)^2}{(2^\circ)^2}\right],$$

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 tends to perform the best on average.

[3] Note that the nonnegativity constraint is not applied to the intercept term a_0 because the physically based sign is not known a priori. Roughly, the intercept "scales" the probability so that the time-mean probability is the right magnitude. We want freedom to fit any mean probability; therefore, a_0 has no constraints.

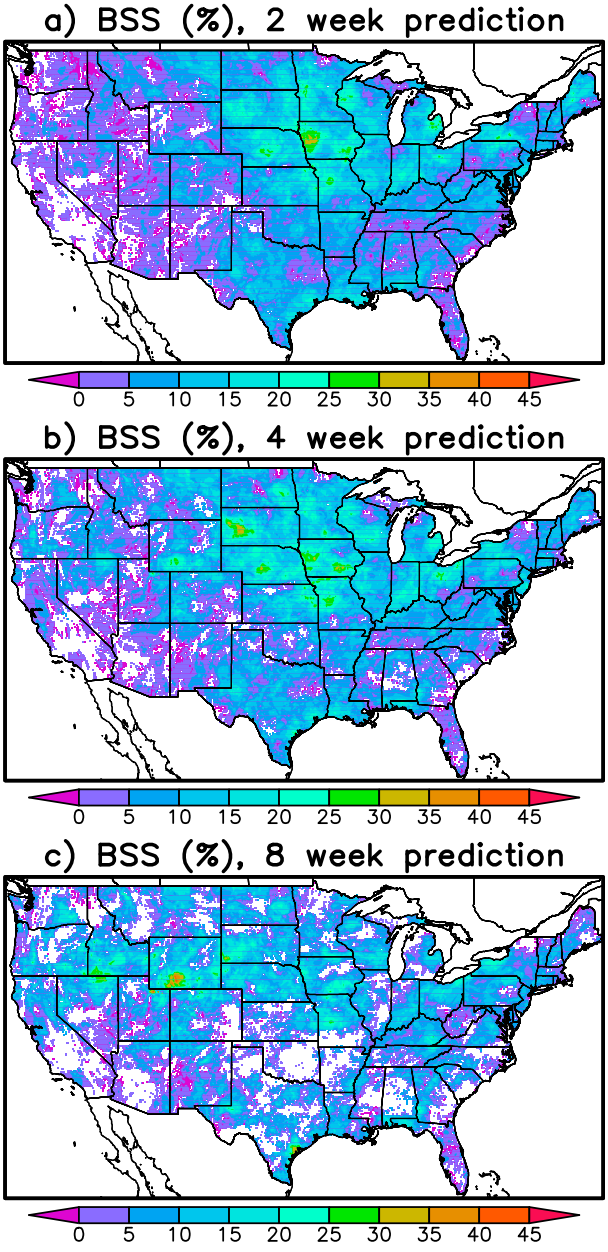

FIG. 1. Cross-validated BSS for logistic regression without USDM state information: (a) 2-, (b) 4-, and (c) 8-week forecast.

In summary, we use standard logistic regression with the following modifications: 1) a sign constraint on the coefficients and 2) surrounding grid points are used for model fitting (but not for model validation) in order to increase sample size. In the next subsection, we explore the best way to incorporate USDM state information as an additional predictor. To motivate the USDM state predictor methodology, we first show the preliminary results of applying the logistic regression with all predictors except the state predictor. In Fig. 1, the cross-validated Brier skill score (BSS) for the 2, 4, and 8-week USDM intensification forecasts are shown. The cross validation is done by removing one year from the record, training the model on the remaining years, and then validating on the left-out year. For the validation, goodness of fit is measured using the likelihood. This process is repeated until all years have a turn to be removed. 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 2- and 4-week predictions have the most skill in the north-central United States and the least skill in the west and parts of the southeast. The 8-week predictions have the most areas with no skill on independent data (white) and the regions of largest skill are in the Intermountain West.

d. USDM state predictor

An important result of this paper is that predictions of future USDM are significantly improved given information about the current state of the USDM. This additional information is more than simply the value of the current drought category, but involves an estimate of the "distance" to the next-higher drought category. For example, suppose the USDM is in the "no drought" category. We show that it is useful to know whether conditions are normal or extremely wet because recent dry anomalies are much more likely to tip the former state into drought than the latter. The development of a methodology to more precisely characterize the state of the USDM was described in Part I. Briefly, Part I develops an empirical methodology for making a non-discrete USDM index that is most consistent with the time scales and processes of the actual USDM presented. In other words, Part I developed a USDM that classifies wetness/dryness on a continuum rather than as five discrete drought/dryness categories together with the no drought designation. Anomalies in precipitation, soil moisture, and evapotranspiration over a range of different time scales are used as predictors to estimate this continuous USDM. 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 2 is a schematic showing the PDF of this hypothetical continuous version of the USDM and the five dryness/drought categories. Figure 2a shows the "climatological" distribution of USDM at a single grid point over all times in the record. Part I then estimated the PDF of the continuous USDM conditioned on SPI, ESI, and

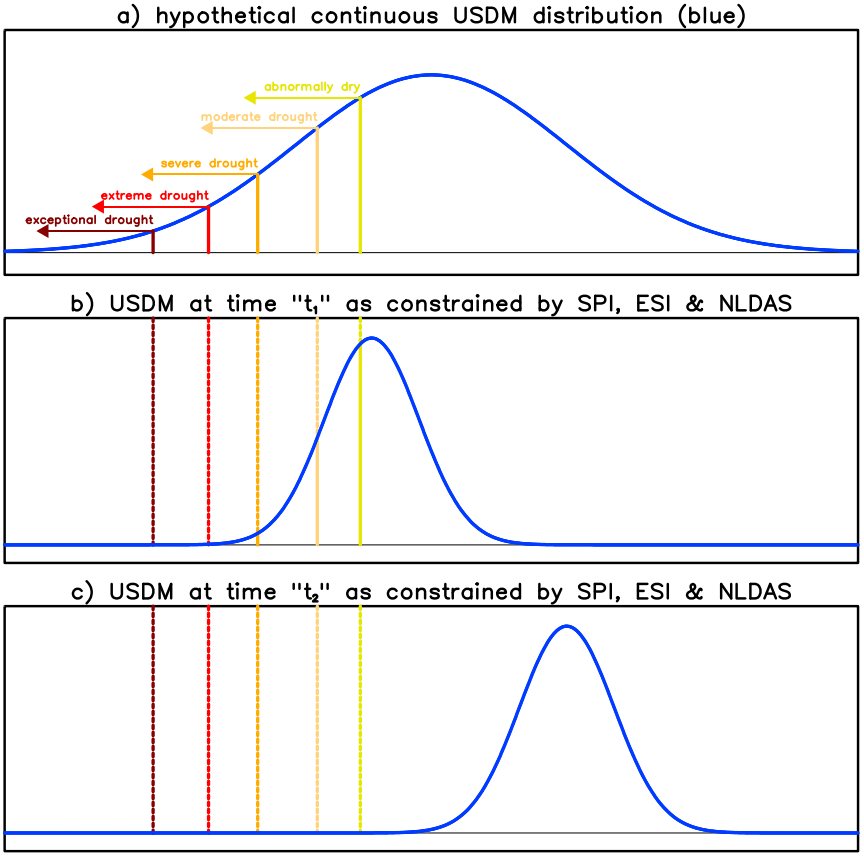

FIG. 2. (a) Schematic of the hypothetical, continuous USDM distribution (for all times). (b) PDF estimate of USDM at time t_1 given the information in the SPI, ESI, and NLDAS. Note the width of the distribution is smaller than in (a) because SPI, ESI, and NLDAS help constrain the range of USDM possibilities. (c) As in (b), but for time t_2.

NLDAS anomalies. Figures 2b and 2c show these conditional distributions at two example times: t_1 and t_2. Notice that the widths of the distributions at times t_1 and t_2 are less than the climatological distribution because the SPI, ESI, and NLDAS predictors add information about the current state that constrains the range of possible USDM drought categories. On the other hand, also note that the PDFs do not precisely pinpoint the true USDM category; instead, they allow for the inherent uncertainty in estimating USDM. 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 much different in these two cases. For example, at time t_2 the integral of the PDF to the left of the "abnormally dry" category threshold is nearly the same as the integral over the entire domain. Therefore, the values of the SPI, ESI, and NLDAS predictors imply that the USDM is almost certainly not in drought. For t_1, on the other hand, there is roughly a 50% chance that the USDM is not in drought. 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 observed USDM is more likely to intensify at time t_1 than at time t_2. This type information will be exploited below to enhance USDM predictability.

To exploit the information in the USDM state PDF, we must determine a good way to quantify the "distance" to the next-higher drought category. One potential distance metric is the probability h that the USDM should be in a higher drought category based on the continuous USDM PDF (i.e., the integral of the PDF below the current drought category). For example, consider the situations in Figs. 2b and 2c and suppose the current USDM category is no drought. The value of h is the integral of the blue PDF to the left of the yellow line. The state in Fig. 2b has a much higher value for h than Fig. 2c, and therefore one expects it to be more likely to intensify. Note that h depends on both the empirical

PDF from Part I and the actual USDM state. So if instead the current USDM category is "moderate drought," then the probabilities (i.e., h) in Fig. 2 would be the integral to the left of the dark orange line, which is less than the h value if USDM was in the no drought category. This conforms to our expectations that, given the same SPI, ESI, and NLDAS anomalies, a moderate drought category is much less likely to intensify than no drought category.

We now determine whether the state predictor h is in a form that is consistent with logistic regression. First, we calculate the two-dimensional histogram of all observations based on the "total" logistic regression predictor from the "no USDM state" model (section 3a) on the x axis and the value of h on the y axis. We define the total logistic regression predictor as the weighted sum of all predictors inside the logistic function: $a_0 + a_1x_1 + a_2x_2 + a_3x_3\cdots$ [(1)]. For this analysis, the 4-week predictions are used and all grid points are aggregated together for more robust statistics. The gray shading in Fig. 3a shows the total number of observations, or cases, in each two-dimensional bin. The fact that the density of observations is largest near the x axis means that most of the time the probability that the USDM should be in a higher drought category is relatively small. This is consistent with the high skill scores in Part I: the PDF is usually centered on the actual USDM category, so h is "small." To gauge the consistency with logistic regression, we next compute the empirical probability of intensification in each bin (i.e., p) by calculating the ratio of the cases that intensify to the total number of cases. Because the probability is nonlinearly related to the predictors in (1), the probability is transformed by the inverse of the logistic function (called the logit function):

$$q = L^{-1}(p) = \text{logit}(p) = \log\left(\frac{p}{1-p}\right), \qquad (3)$$

where q is the transformed probability. The value of q is contoured in Fig. 3a (to focus on the robust results, q is only contoured where there are at least 1000 observations). If the predictors are consistent with logistic regression, then q should be a linear function of the predictors or, in other words, the contours of q should be straight lines that are equidistant from each other. Unfortunately, the q contours are significantly curved, and moreover this curvature occurs near the highest density of observations (i.e., the darkest gray). However, notice that the probability of intensification (contours) does depend on h, suggesting that there is skill and that by transforming h with some nonlinear function, one may make the q contours straight lines. After some

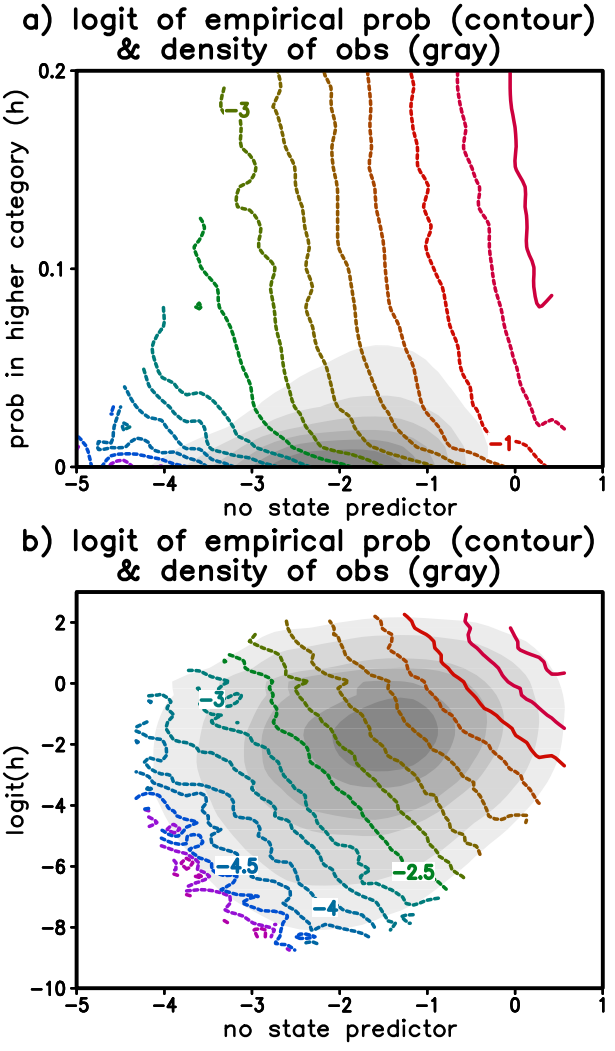

FIG. 3. (a) Two-dimensional histogram of density of observations as a function of the total predictor for the no-state logistic regression (x axis) and the probability that the USDM should be in higher drought category (i.e., h) (y axis; gray shading). Empirical probability of intensification transformed by the logit function (contours). (b) As in (a), but for the logit of h on the y axis.

experimentation, we determined that the logit of h is a predictor consistent with the form of logistic regression:

$$\text{USDM state predictor} = \log\left(\frac{h}{1-h}\right). \qquad (4)$$

For example, Fig. 3b shows the empirical probabilities as a function of the "no state" predictor and the logit of h. The contours of q are equally spaced straight lines in most regions, confirming that q is a linear function of the logit of h. Also, it turns out that the logit function is an increasing function of h, so all of the qualitative interpretation and reasoning regarding the h predictor

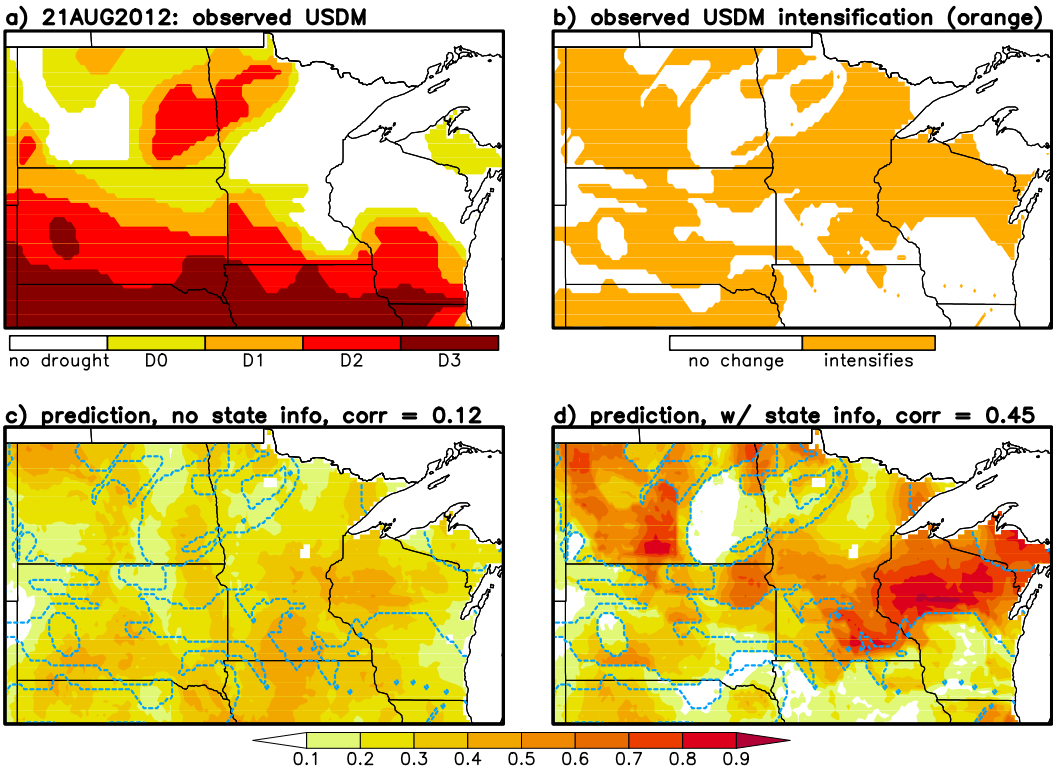

FIG. 4. (a) USDM on 21 Aug 2012. (b) Regions where USDM intensifies over the 4 weeks following 21 Aug 2012 are orange. (c) Probability of intensification from logistic regression with no USDM state information. The blue dashed lines are the boundaries between the orange and white in (b). (d) As in (c), but for the logistic regression with USDM state information.

above also applies to the logit of h. We should also note that the USDM state PDF used to calculate h in this study is cross validated.

4. Results

a. Example showing effect of USDM state predictor

Here we describe an example (Fig. 4) in the north-central United States from the 2012 drought that shows the advantages of using the USDM state predictor. The USDM on 21 August 2012 shows extreme drought in the southern portion of the domain with no drought over portions of Minnesota, northern Wisconsin, and western North Dakota (Fig. 4a). Regions that experienced drought intensification over the next 4 weeks are shown in orange in Fig. 4b. The probability of drought intensification from the logistic regression without and with the USDM state information is shown in Figs. 4c and 4d, respectively. For reference, the boundaries of the intensification regions are shown by the blue dashed lines, and the spatial correlation between the probabilities and the actual intensification pattern (i.e., Fig. 4b) are shown in the top-right corner of Figs. 4c and 4d. For these spatial correlations, the observed intensification map is 1 where the USDM intensifies and 0 otherwise. Also, the raw continuous probabilities are used rather than values discretized by rounding to 0 or 1. In this way, the degree of certainty in the forecasts is taken into account.

The regression with state information performs significantly better, particularly in delineating the boundaries between regions that intensify and those that do not. For example, Fig. 4d shows an abrupt increase in intensification probability as one travels from south to north across Wisconsin. This abrupt transition coincides with the boundary between the orange and white in Fig. 4b. The regression with state information also captures these boundaries in northeastern South Dakota and in North Dakota. Note that these abrupt transitions occur at the boundaries of the USDM categories in Fig. 4a. The reason is as follows: the USDM is a discrete variable quantifying a continuous distribution, and therefore, changes tend to occur via the expansion and contraction of the discrete USDM contours surrounding existing drought regions. Hence, changes in USDM

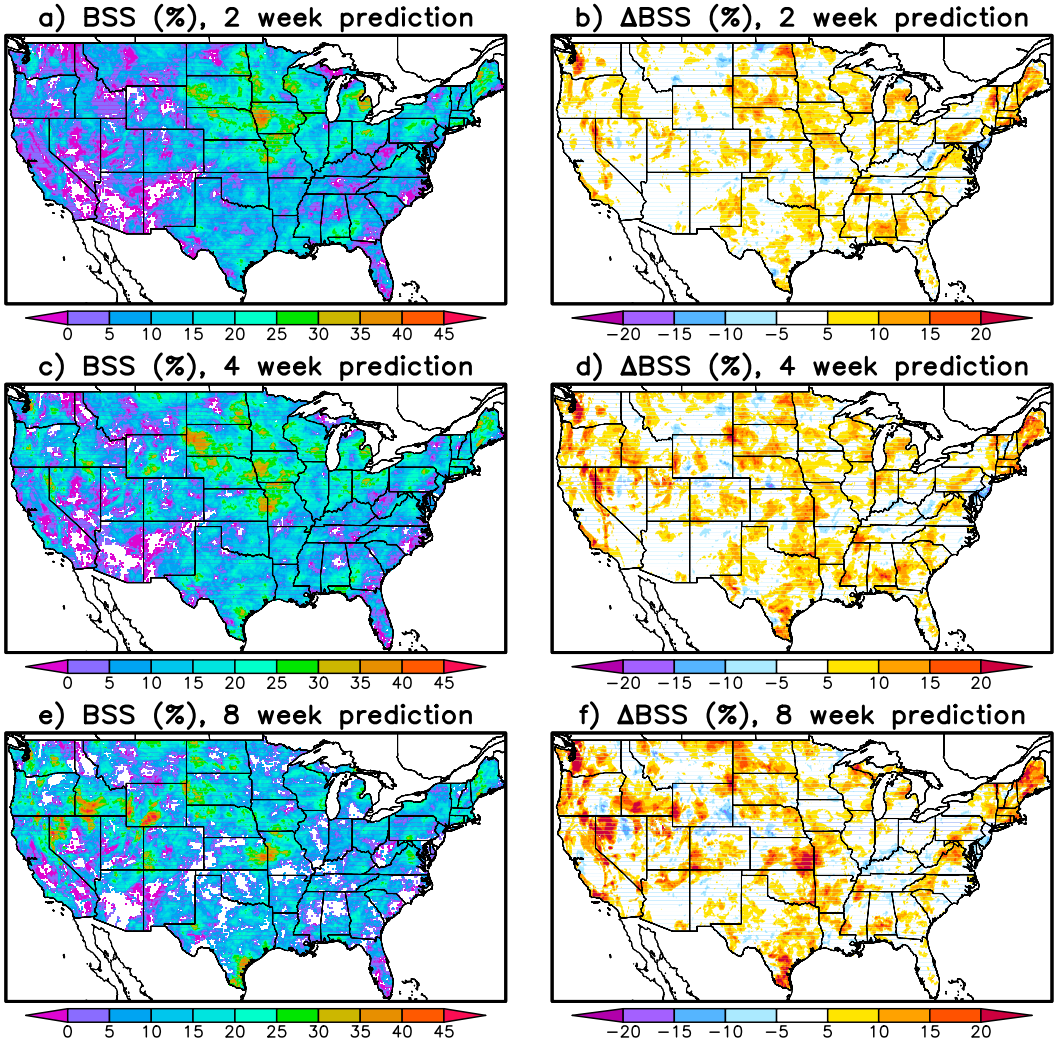

FIG. 5. Cross-validated BSS for logistic regression with USDM state information and change in BSS compared with no USDM state information: (a) 2-week forecast, (b) BSS with state information minus BSS with no state information for 2-week forecast, (c) 4-week forecast, (d) BSS with state information minus BSS with no state information for 4-week forecast, (e) 8-week forecast, and (f) BSS with state information minus BSS with no state information for 8-week forecast.

drought severity are most likely to occur close to the current USDM contours. The USDM state variable used here captures this effect because the probability that the USDM should be in a higher drought category (i.e., h) also tends to change abruptly at the USDM contours because the limit of the integration of the USDM PDF (Fig. 2) changes abruptly.

b. *Probabilistic skill*

The cross-validated BSSs of logistic regression with USDM state information included as a predictor are shown in Fig. 5 together with the change in BSSs relative to the no-state predictor regression. The spatial patterns of skill are similar to Fig. 1, with better skill in the north-central United States for the 2- and 4-week forecasts, for example. While the patterns are similar, the skill scores are improved at nearly all locations. Averaged over the north-central United States where skill is highest (38°–46°N, 87°–104°W), the BSSs for both the 2- and 4-week forecasts improve from 0.15 to 0.20. Unlike the rest of the domain, the skill in the western United States is largest for the 8-week forecasts. This is likely due to the longer time scales of USDM variations in the west. The standard BSS compares model skill to the "baseline" skill of the climatology. We also tried persistence as a baseline and found that the "skill" scores increase everywhere significantly. These inflated skill scores are not shown because climatology is a stricter metric in this

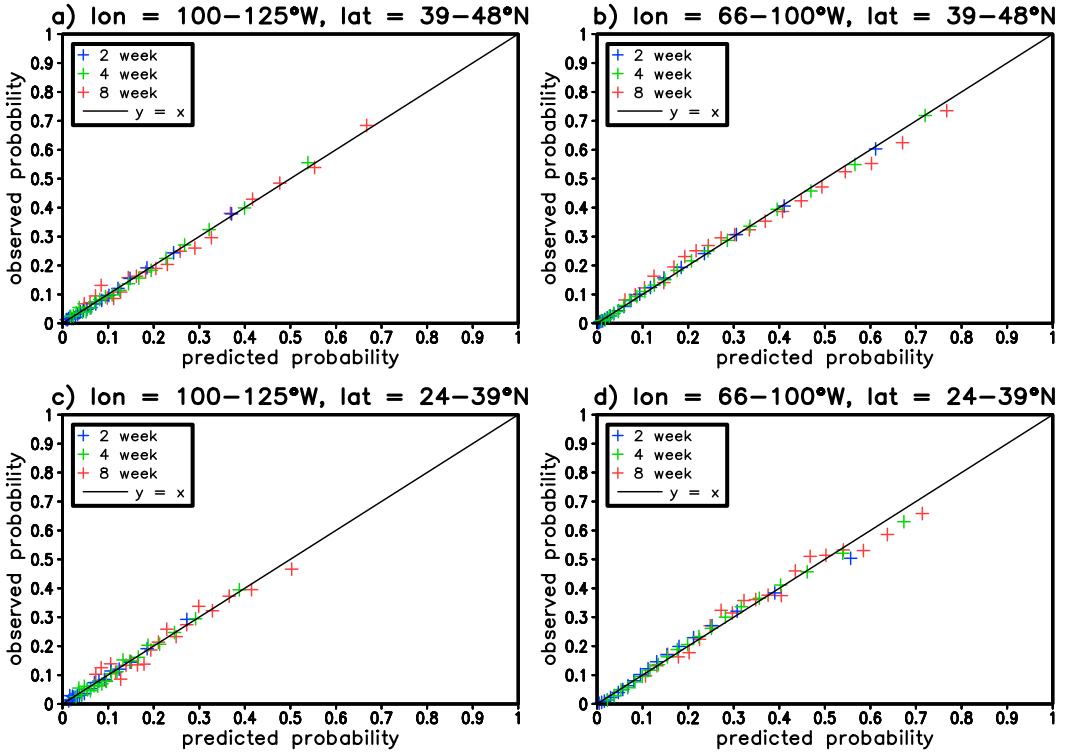

FIG. 6. (a) Reliability diagram for the probabilistic USDM forecasts for each forecast period (colors). The x axis is the predicted probability and the y axis is the probability from the observed USDM record. For an ideal fit, the points should lie on the $y = x$ line (black). To reduce noise, the probabilities are averaged over the northwestern quarter of the United States. (b) As in (a), but for the northeastern quarter of the United States. (c) As in (a), but for the southwestern quarter of the United States. (d) As in (a), but for the southeastern quarter of the United States.

case. The standard BSS compares model skill to the baseline skill of the climatology. We also tried persistence as a baseline and found that the skill scores increase everywhere significantly. These inflated skill scores are not shown because climatology is a stricter metric in this case.

The above results suggest that the USDM state information adds skill to the predictions; however, it is possible that the no-state regression is not the most skillful (and therefore meaningful) null hypothesis. For example, perhaps the USDM state regression has more skill because climatologically the no drought category is broad while the other USDM categories are relatively narrow. USDM categories that are narrow have less distance to adjacent categories on average, so the probability of increase is larger. Perhaps this is the effect that the USDM forecasts are using to increase skill. To test this hypothesis, we use the width of the current USDM category (in percentiles) as a predictor. The BSS for this new predictor, however, is very much like the no-state regression, which suggests that the no-state regression is an appropriate null hypothesis. We also use a climatological USDM state PDF (i.e., one that is independent of time) to form the state predictor by the method described in section 3d. This null hypothesis is also no better than the no-state regression.

We have also calculated the BSS only during times when the USDM is in the abnormally dry category or more intense. The BSS in this restricted case is comparable to the full BSS (not shown), which says that the USDM state information is useful even when the USDM is not in the broad no drought category.

Next, the skill of the predictions is assessed using reliability diagrams (Wilks 2011). A "reliable" forecast means that, given a large number of individual cases when we predict a 60% chance of drought intensification, the actual USDM drought depiction will intensify 60% of the time and will not intensify during the remaining 40%. In a reliability diagram (e.g., Fig. 6), the observed probability (y axis) is plotted as a function of the predicted probability (x axis). Thus, for a perfectly reliable forecast, the points on the diagram would lie along the $y = x$ line. For these reliability scatterplots, the data are divided into 20 bins (see Part I for more details). To help eliminate noise, the individual reliability diagrams at each grid point are averaged over the

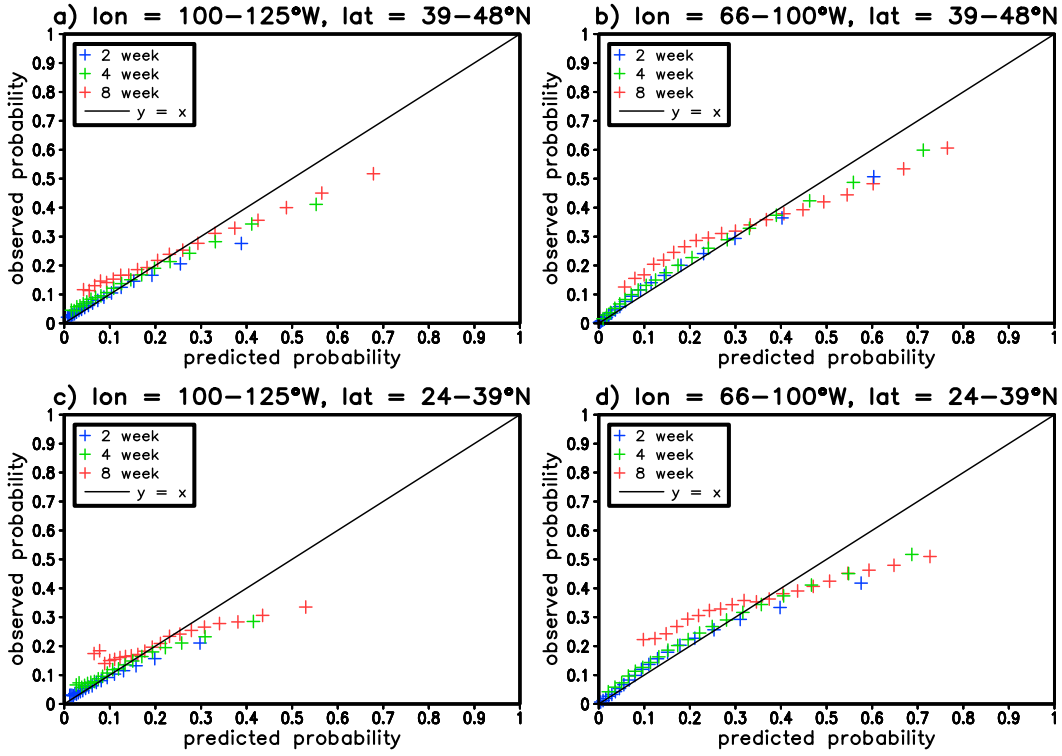

FIG. 7. As in Fig. 6, but for the cross-validated probabilities.

northwestern, southwestern, northeastern, and southeastern portions of the contiguous U.S. domain (see figure titles for the boundaries of the domains). The different colors denote the reliability diagrams for the 2-, 4-, and 8-week forecast time periods.

Figure 6 shows reliability diagrams for data that have not been cross validated. 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 reliability diagrams are very good, with points falling very close to the $y = x$ line, demonstrating that the logistic regression with the canonical logistic "link function" captures the dependence of the probability on the predictors. The corresponding cross-validated reliability diagrams are shown in Fig. 7. The 2- and 4-week forecasts are still quite good, although the forecasts tend to be slightly overconfident: the predicted probability is larger than the observed probability for high-probability events and smaller than the observed probability for low-probability events (i.e., the scatter intersects the $y = x$ line but the slope is less than one). The 8-week forecasts, on the other hand, noticeably deteriorate when applied to independent data. It appears that the 8-week time scale is approaching the limit of USDM predictability given current and past conditions of the land state. In future work, these forecasts will incorporate additional predictors from actual climate model forecasts. While we expect this information to improve these forecasts at all lead times, perhaps the biggest improvement will occur for longer forecast lead times when the current state has less of an impact.

c. *Predictors chosen*

In this section we discuss the number and relative importance of the various predictors used to predict changes in the USDM. Figure 8a shows the number of predictors with nonzero coefficients or weights for the 4-week forecasts, which are quite similar to the results for 2- and 8-week forecasts (not shown). Throughout much of the central and northeastern United States, at least 12 predictors are chosen. In areas of the southwestern and the southeastern United States, as few as 6 predictors are chosen.

Figures 8b–h show the relative weight of the precipitation; ESI; 10-, 100-, and 200-cm NLDAS soil moisture variables; the USDM state predictors; and the mean seasonal cycle, respectively. For this comparison, the contribution of all time lags (0, 1, 2, and 3 weeks) is aggregated for each variable that uses multiple time lags. Because all predictors are standardized and all coefficients are positive (by construction), the relative

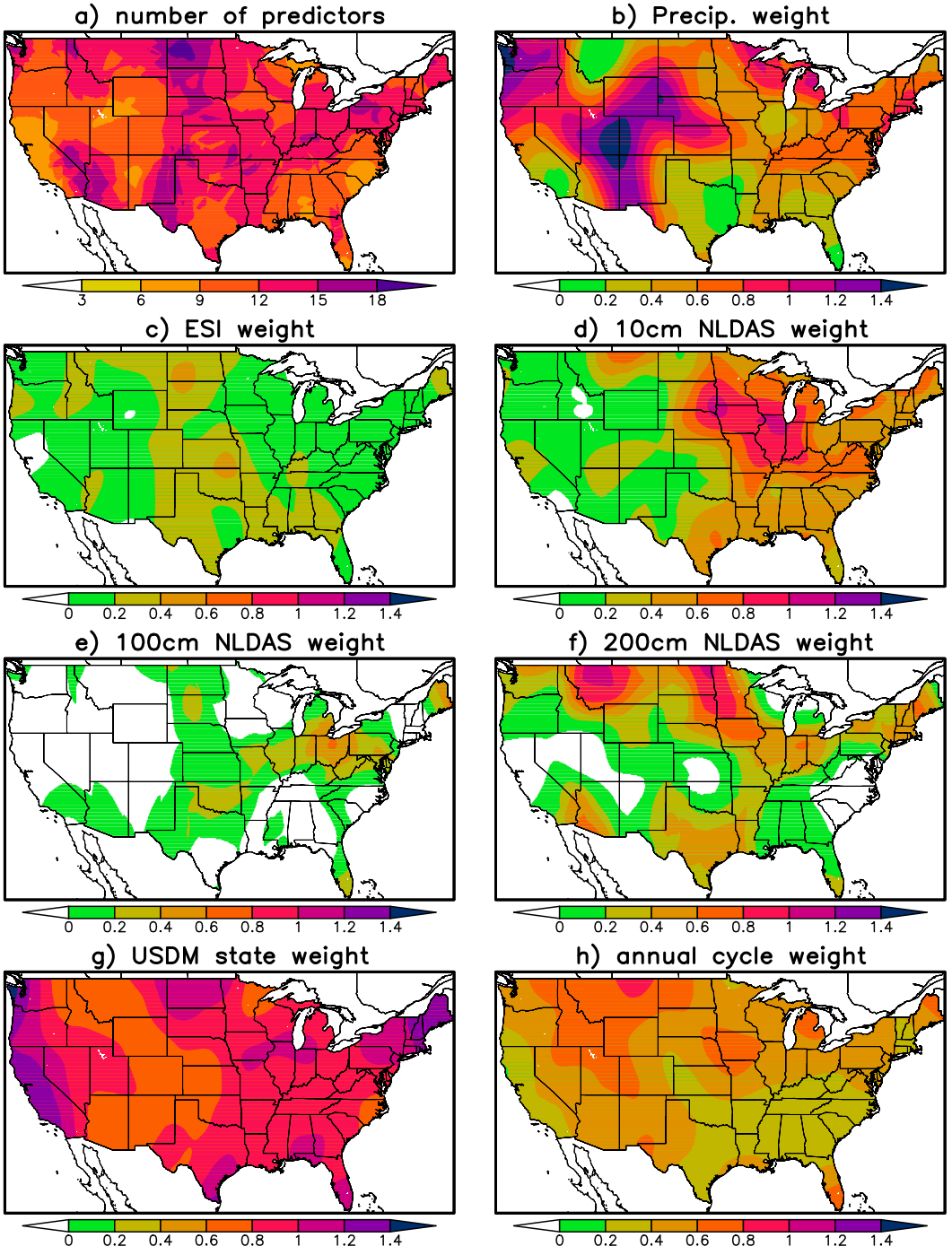

FIG. 8. (a) The number of predictors with nonzero weights for nonnegative logistic regression. (b) Sum of the weights involving precipitation. (c) Sum of the weights involving ESI tendency. (d) Sum of the weights involving the 0–10-cm soil moisture tendency from NLDAS. (e) As in (d), but for the top 0–100 cm. (f) As in (d), but for the top 0–200 cm. (g) Sum of the weights for the USDM state predictor. (h) Sum of the weights for the climatological annual cycle predictor.

weight for each variable is calculated by simply summing over the coefficients from the nonnegative logistic regression. Precipitation has the most weight in the Pacific Northwest and in the central and southern Rocky Mountains (Fig. 8b). It also has relatively large weighting in the northeastern United States and the upper Midwest. The ESI tendency has significantly less weight than the precipitation (Fig. 8c), which is consistent with

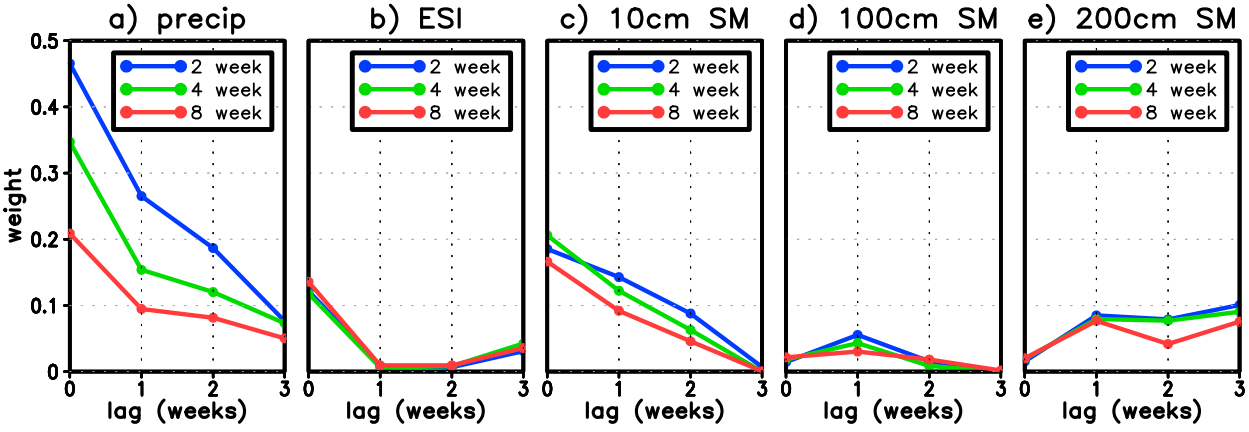

FIG. 9. Predictor weights as a function of time lag. Weights are averaged in space over the entire domain. (a) precipitation, (b) ESI tendency, (c) 0–10-cm soil moisture tendency, (d) 0–100-cm soil moisture tendency, (e) 0–200-cm soil moisture tendency.

the fact that ESI has only recently been used by the USDM to estimate drought intensity. The ESI tendency is used mostly in the central United States, perhaps because the USDM is focused on agricultural drought in this region. ESI is important because, unlike precipitation, it takes into account the effect of warm temperatures, high winds, and low relative humidity on drought development. The shallow soil moisture tendency (Fig. 8d) is most important in the eastern United States, where it tends to be the most important of the soil moisture predictors in regions that are susceptible to flash droughts (e.g., Otkin et al. 2014). Between the two deeper soil moisture level tendencies, however, the 200-cm soil moisture tendency has more weight (Figs. 8e,f). Perhaps this is because the 100-cm layer is more closely related to the 10-cm layer, which is already included as a predictor. Unlike the other variables, the USDM state predictor has weights that are both large and relatively uniform across the country (Fig. 8g), demonstrating that the USDM state information makes important contributions everywhere. For the contributions of the precipitation, ESI, and soil moisture to the USDM state predictor, see Fig. 4 of Part I. The mean seasonal cycle also has relatively uniform but moderate weight (Fig. 8h).

Figure 9 shows the domain-averaged weights as a function of time lag for the five fields that depend on time lag: precipitation, ESI, and NLDAS soil moisture integrated to depths of 10, 100, and 200 cm. The short-term (e.g., 2 and 4 weeks) variables for precipitation and 10-cm soil moisture tendency have similar, nearly linear drop-offs from lag 0 to 3 weeks. Physically, the small lags have more weight because they are closer to the verification time. Because ESI is the tendency of a 4-week composite, on the other hand, the lag 1 and 2 week values are closely related to the lag 0 week value, and therefore the weights are reduced for lag 1 and 2 weeks. The deeper soil moisture variables are different in that the weights do not peak at lag 0 weeks. The weights for all forecast time periods (2, 4, and 8 weeks) are remarkably similar, with the largest differences occurring for precipitation. The differences for precipitation are consistent with the difference in forecast lead time: for example, the 2-week forecasts have more weight for the most recent (lag 0 weeks) anomalies than the 4- and 8-week predictions.

d. Examples

In this subsection, the probabilities of intensification from the logistic regression method are compared to the observed occurrence of USDM intensification. Figures 10–12 show comparisons during the beginning, middle, and end of the growing season for the last 9 years of the period of record (2006–14) for the 4-week forecasts, with all results being cross validated. The panels come in pairs: Figs. 10–12 (top) show the observed intensification and Figs. 10–12 (bottom) show the probability of intensification from the logistic regression. The observed plots are orange if the USDM intensifies over the specified time period and white otherwise. The spatial correlations between the observed occurrence and the predicted probabilities are shown in the title of the forecast panels (panels with a blue title caption). These correlations are calculated in the same way as Fig. 4. In some years the forecasts are quite good, for example, the years 2007, 2008, and 2012 in the mid-May examples (Fig. 10). In general, it appears that the forecasts are best when a relatively large portion of the domain experiences drought intensification. Conversely, when the USDM does not depict widespread drought intensification, the spatial correlations tend to be weaker, for example, the years 2013

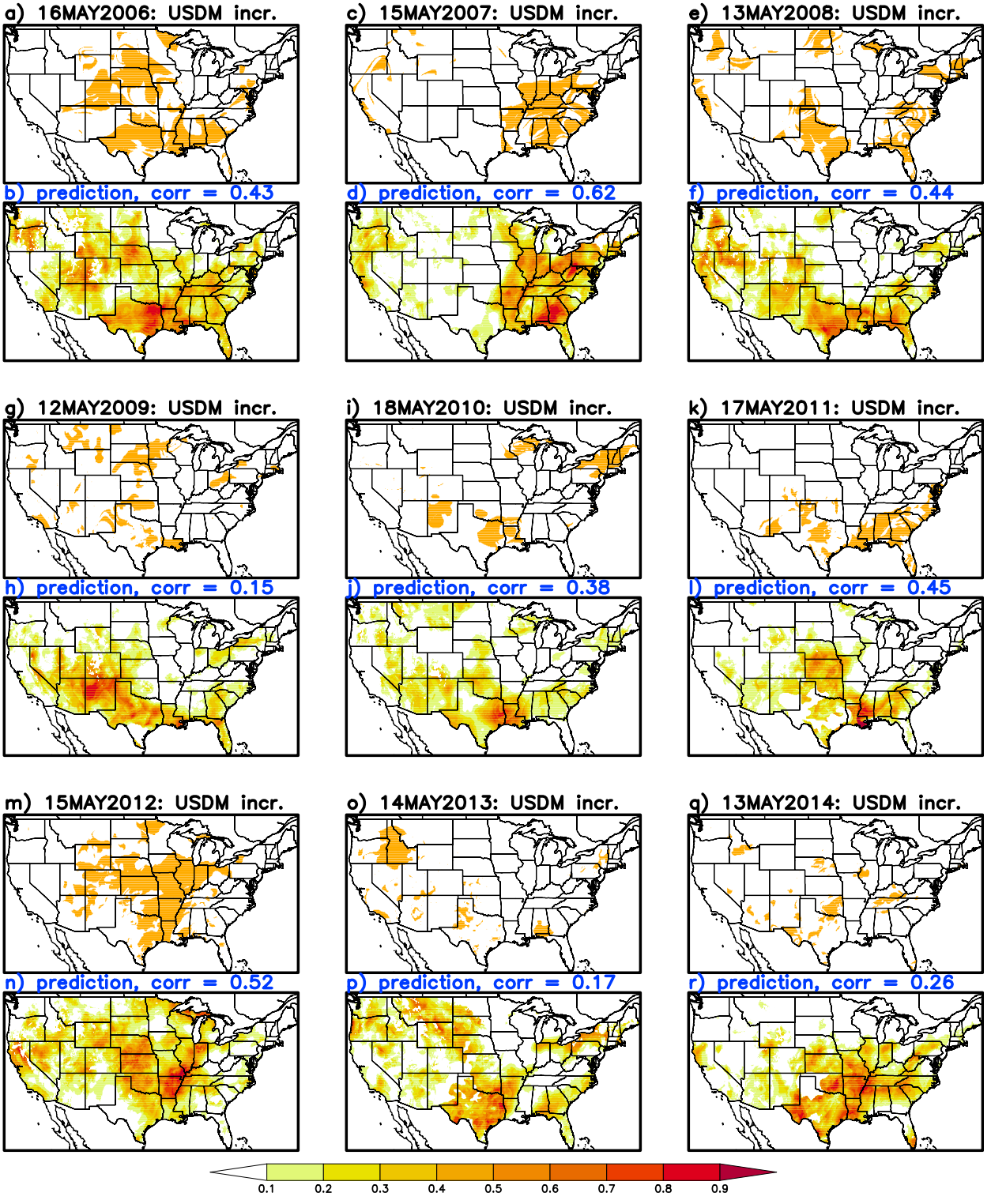

FIG. 10. Comparison of USDM intensification and the cross-validated probabilities (i.e., the model has not been trained on the year shown) for mid-May of the latest 9 years: (a) USDM intensification on 16 May 2006. Locations that intensify over 4 weeks are shown in orange. (b) Probability of intensification over 4 weeks on 16 May 2006. The spatial correlation between USDM intensification and the probabilities is shown. (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.

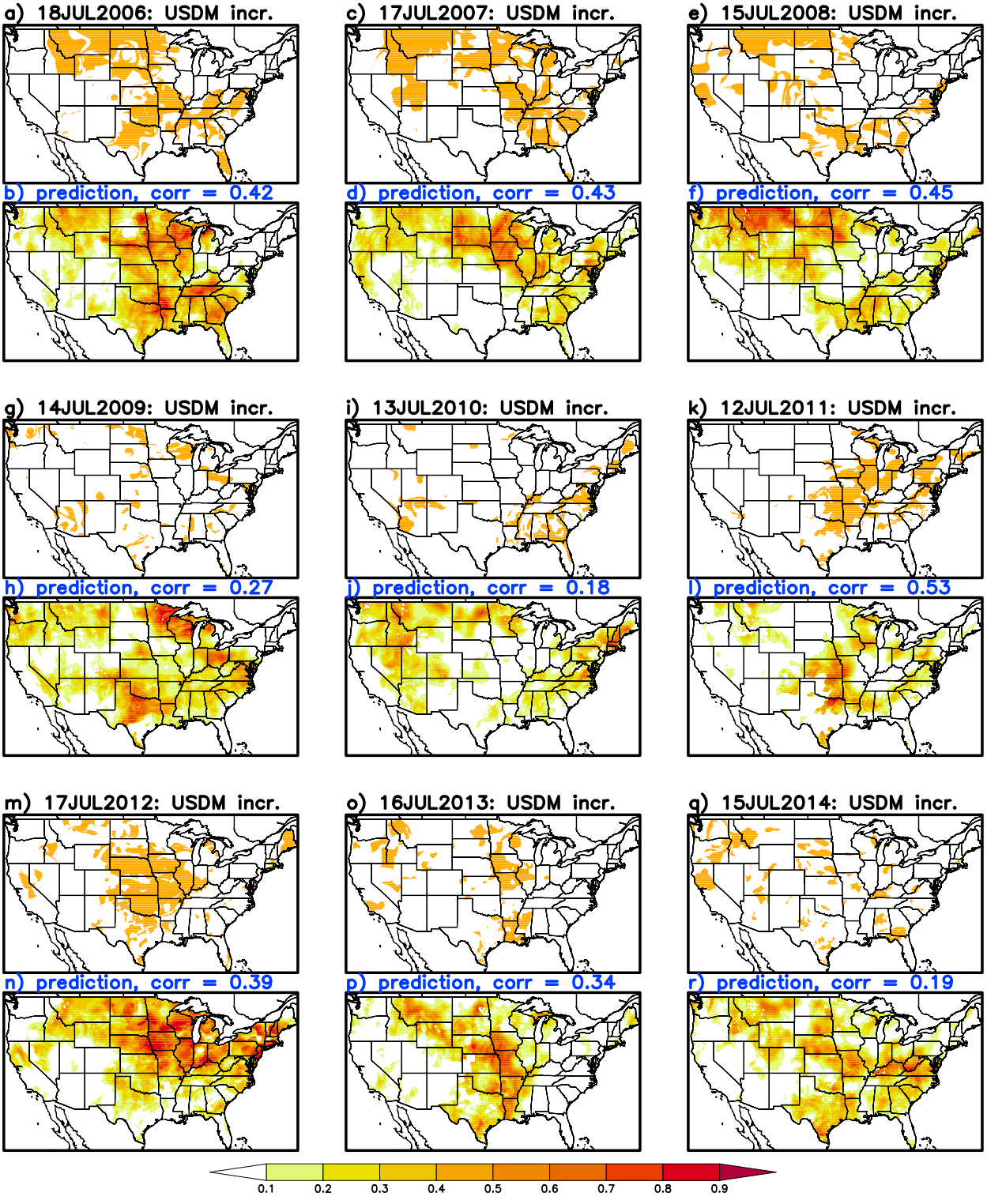

FIG. 11. As in Fig. 10, but for mid-July.

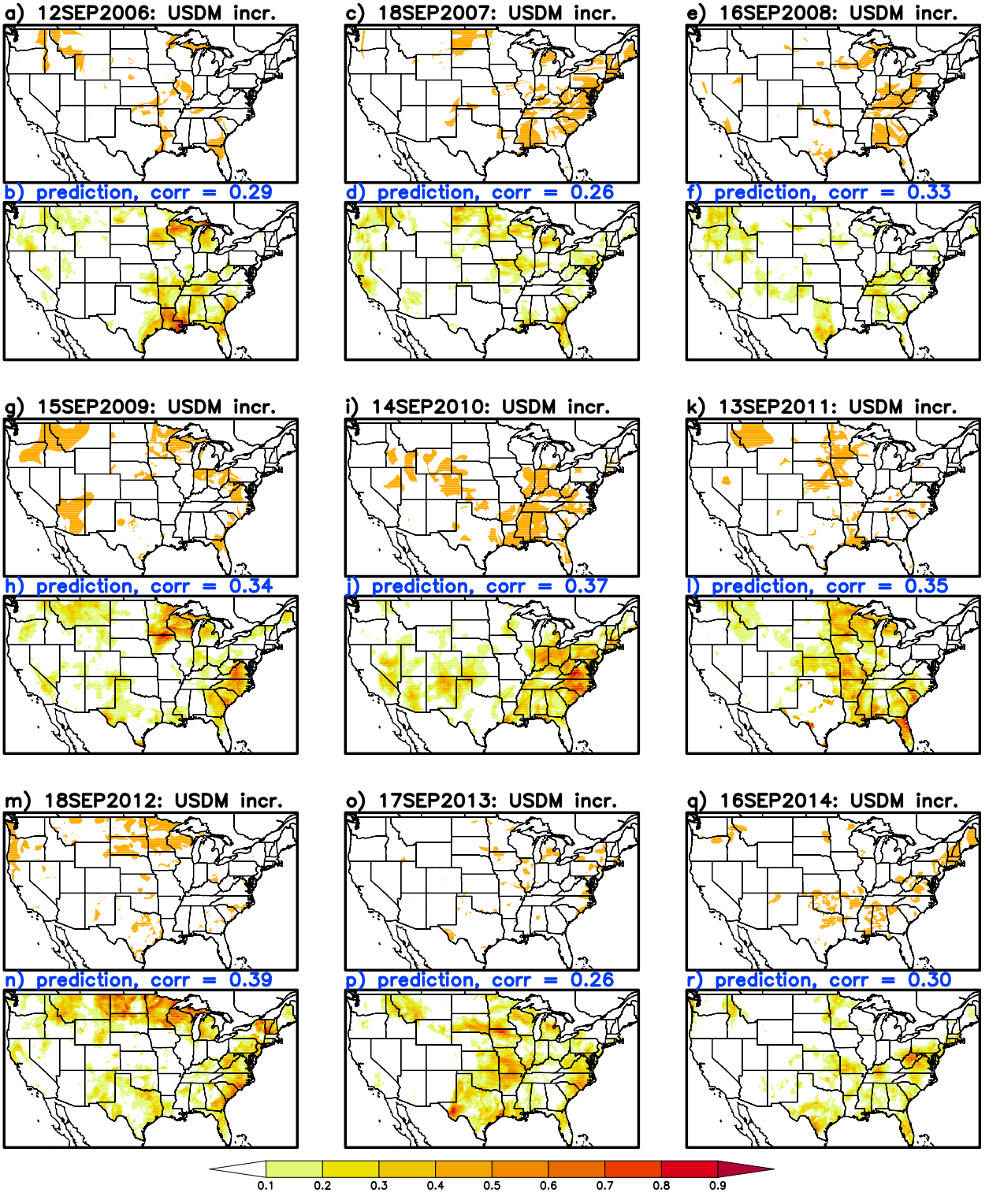

FIG. 12. As in Fig. 10, but for mid-September.

TABLE 1. The largest flash droughts in the record for the months May–September for the years 2001–14. The amplitude of a flash drought is quantified by the largest increase in the USDM averaged over a climate division over a 2-, 4-, or 8-week time period. The change in the average USDM, the year, the state, and the climate division are shown. Events in boldface are shown in more detail in Fig. 13.

	2 weeks				4 weeks				8 weeks			
Rank	ΔUSDM	Year	State	Division	ΔUSDM	Year	State	Division	ΔUSDM	Year	State	Division
1	**2.51**	**2005**	**Louisiana**	**7**	**3.78**	**2012**	**Wisconsin**	**8**	**4.15**	**2012**	**Arkansas**	**5**
2	**2.12**	**2006**	**Minnesota**	**2**	3.32	2012	Wisconsin	9	**4.12**	**2011**	**Oklahoma**	**9**
3	**2.11**	**2012**	**Wisconsin**	**8**	3.18	2012	Wisconsin	7	**4.12**	**2012**	**Nebraska**	**5**
4	2.10	2013	Arkansas	1	3.15	2012	Nebraska	3	4.05	2012	Oklahoma	2
5	2.07	2006	Alabama	8	3.00	2007	Missouri	6	4.05	2012	Oklahoma	3

and 2014 in Fig. 10. Similar results hold for July (Fig. 11): the years 2006–08 and 2011–12 have relatively large areas with observed drought intensification, and the spatial correlations are also relatively good. Conversely, the years 2009–10 and 2014 have a relatively small amount of intensification and the spatial correlations are weak. For the September results (Fig. 12), the most notable result is that the forecasts seem less skillful compared to May and July. This observation holds up under closer scrutiny: first the warm season (weeks 18–39) is divided into three nearly equal-length subseasons. The cross-validated BSSs are then calculated separately for each subseason. Averaged over the north-central United States where skill is highest (38°–46°N, 87°–104°W), the BSS is 0.24, 0.20, and 0.17 for the early, middle, and late parts of the warm season, respectively. Similar results are obtained for the domain-averaged BSS and for the 2-week predictions. The 8-week predictions, on the other hand, have similar skill for each subseason. The reasons for the seasonality in skill are the subject of future work.

We also show examples of the time progression of the probabilities of intensification together with the observed USDM. First, the largest ''flash drought'' events during the analysis period (May–September) are selected based on the areas with the largest increase in the USDM over the 2-, 4-, or 8-week time periods. The analysis is performed over the U.S. climate divisions, and thus the USDM values are averaged over the climate divisions first before selecting the largest increase. The top five events are shown in Table 1. The ΔUSDM column shows the number of categories of the increase for each time period and can be a fractional value due to the average over the climate division. The year, state, and climate division of each event are also shown. Some events appear in multiple time periods, such as the flash drought in Wisconsin in 2012. Six events that are well separated in space and time are chosen for the time progression plots in Fig. 13 (boldface in Table 1). Figures 13a–c document the expansion from south to north of the drought of 2012. The drought began in Arkansas near the end of May and reached peak intensity by August (orange bars in Fig. 13a). In Nebraska, the drought did not achieve the moderate drought category until the end of June, and in Wisconsin this threshold was not reached until July. The probabilities of intensification for the 2-, 4-, and 8-week periods are shown by the green, blue, and purple lines, respectively. In general, the forecasts anticipate the drought development several weeks before the USDM depicted drought intensification, particularly for Wisconsin, where probabilities began rising in May and eventually reached about 90% two weeks before any significant drought intensification occurred.

In the 2011 Oklahoma case, the forecasts do very well anticipating the drought with very high probability a month in advance. The Minnesota drought of 2006, on the other hand, did not achieve the same level of certainty even though the timing of the probabilities is still good. The Louisiana flash drought of September 2005 also did not achieve the same level of certainty. Moreover, the predicted probability of intensification nearly reached the same magnitude earlier in the year when little or no drought intensification occurred. When looking at other weaker flash droughts (not shown), it becomes evident that the forecasts tend to be better when the events have greater amplitude and vice versa. This is consistent with the spatial correlations in Figs. 10–12, which were better when more of the country was experiencing intensifying drought.

5. Conclusions

The development of probabilistic USDM drought intensification forecasts over 2-, 4-, and 8-week time periods are described. In this paper, the forecasts only involve current and past conditions, but the method can easily be generalized to include weather forecast model output. Logistic regression is used to estimate the probabilities. The predictors include the four most recent weeks of precipitation, evapotranspiration (ESI) time tendencies, and soil moisture time tendencies at

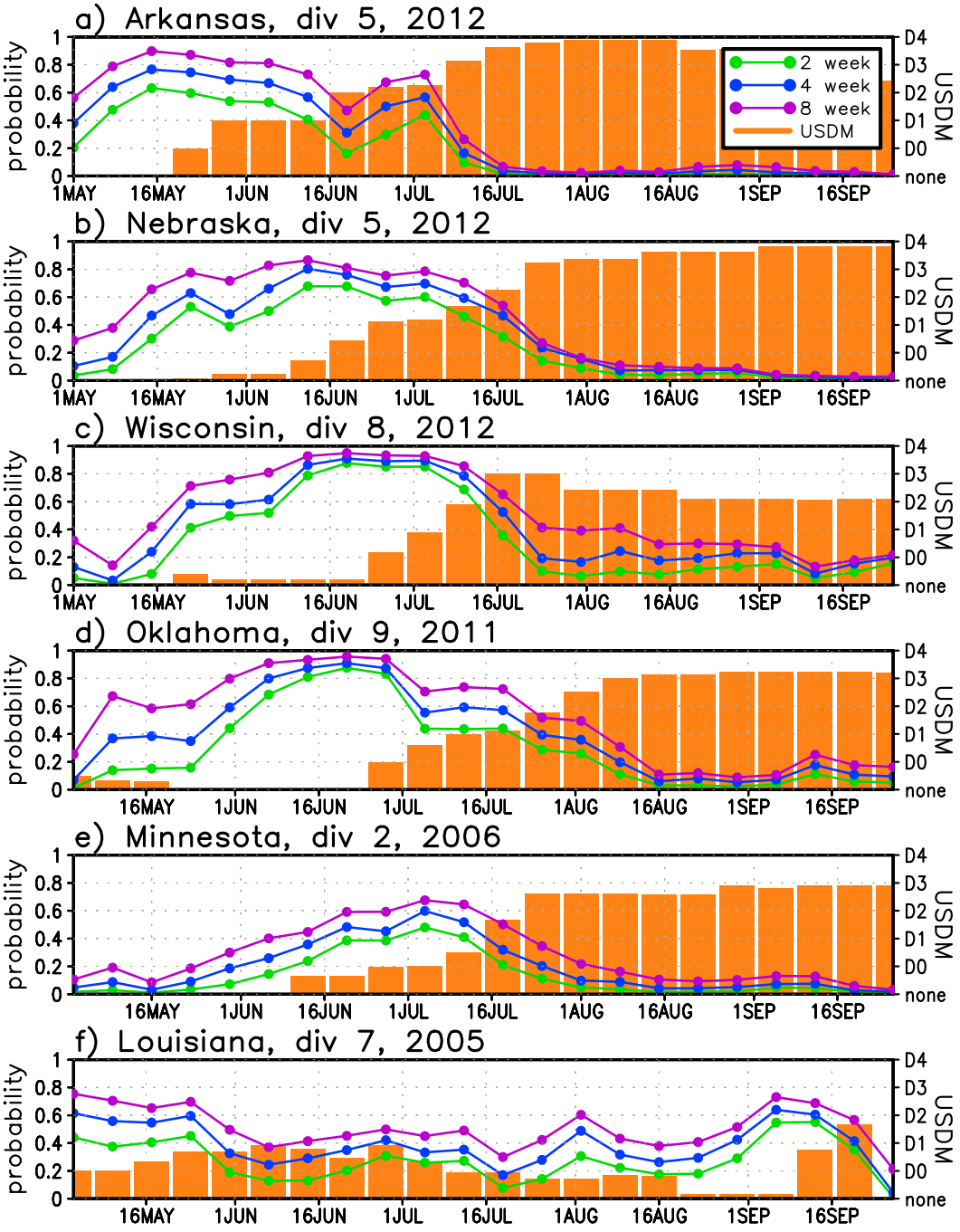

FIG. 13. Time series of probability of intensification for the 2- (green), 4- (blue), and 8-week (purple) forecasts and the observed USDM (orange bars) averaged over climate division. The probability axis runs from 0 to 1 (left *y* axis) and the USDM axis runs from no drought (none), abnormally dry *D*0, moderate drought *D*1, severe drought *D*2, extreme drought *D*3, and exceptional drought *D*4 (right *y* axis). (a) Division 5 in Arkansas in 2012, (b) division 5 in Nebraska in 2012, (c) division 8 in Wisconsin in 2012, (d) division 9 in Oklahoma in 2011, (e) division 2 in Minnesota in 2006, and (f) division 7 in Louisiana in 2005.

several depths. In addition, it is found that detailed information about the current state of the USDM adds skill to the forecasts. The information on the state of the USDM is derived from the probabilistic estimates of the USDM category based on SPI, ESI, and soil moisture anomalies as described in Part I. This additional USDM state predictor quantifies the "distance" to the next-higher drought category. This adds skill because

locations that are close to the next USDM drought category are more likely to intensify than states that are far from the next drought category. Finally, a predictor based on the climatological seasonal cycle of USDM intensification is used.

The method shows skill over most of the United States: cross-validated BSSs are greater than zero for 90%, 93%, and 96% of the domain for the 2-, 4-, and 8-week predictions, respectively. The method is most skillful over the north-central United States, where the cross-validated BSS average is 0.20 for both 2- and 4-week forecasts. The 8-week forecasts are less skillful in most locations. The 2- and 4-week probabilities have very good reliability. The 8-week probabilities, on the other hand, are noticeably overconfident.

Examples of the probabilistic forecasts are shown for the beginning, middle, and end of the growing season for 2006–14. The spatial correlations between the probabilities and the observed occurrence of intensification show that the method tends to be most skillful when large areas of the United States are experiencing intensifying drought. When few regions are intensifying, on the other hand, forecast skill is smaller. Example time series from the most intense flash droughts in the record are also shown. Consistent with Otkin et al. (2014, 2015a), the method typically does very well anticipating these high-amplitude events a month or more in advance.

The results described here define a baseline skill level that can be improved upon using climate model forecast output. Thus, future work will develop new methods to combine predictions from the method presented in this paper with model forecast data from the North American Multi-Model Ensemble to further increase the accuracy of the USDM drought intensification forecasts. We will also explore additional ways to improve the predictions based on recent anomalies. For example, the USDM categorizes droughts as short- and/or long-term droughts. If the method in Part I can be extended to separate these droughts, then skill might improve because short-term droughts are more likely to intensify over subseasonal times scales than long-term droughts.

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.

REFERENCES

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.

——, ——, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, 2007: 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.

Barros, A. P., and G. J. Bowden, 2008: Toward long-lead operational forecasts of drought: An experimental study in the Murray–Darling River basin. *J. Hydrol.*, **357**, 349–367, doi:10.1016/j.jhydrol.2008.05.026.

Bell, V. A., H. N. Davies, A. L. Kay, T. J. Marsh, A. Brookshaw, and A. Jenkins, 2013: Developing a large-scale water-balance approach to seasonal forecasting: Application to the 2012 drought in Britain. *Hydrol. Processes*, **27**, 3003–3012, doi:10.1002/hyp.9863.

Dobson, A. J., and A. Barnett, 2008: *An Introduction to Generalized Linear Models.* CRC Press, 320 pp.

Dutra, E., and Coauthors, 2014: Global meteorological drought—Part 2: Seasonal forecasts. *Hydrol. Earth Syst. Sci.*, **18**, 2669–2678, doi:10.5194/hess-18-2669-2014.

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.]

Hoerling, M., J. Eischeid, A. Kumar, R. Leung, A. Mariotti, K. Mo, S. Schubert, and R. Seager, 2014: Causes and predictability of the 2012 Great Plains drought. *Bull. Amer. Meteor. Soc.*, **95**, 269–282, doi:10.1175/BAMS-D-13-00055.1.

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.

Hwang, Y., and G. Carbone, 2009: Ensemble forecasts of drought indices using a conditional resampling technique. *J. Appl. Meteor.*, **48**, 1289–1301, doi:10.1175/2009JAMC2071.1.

Kim, T. W., J. B. Valdes, and C. Yoo, 2003: Nonparametric approach for estimating return periods of droughts in arid regions. *J. Hydrol. Eng.*, **8**, 237–246, doi:10.1061/(ASCE)1084-0699(2003)8:5(237).

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.

Kumar, A., M. Chen, M. Hoerling, and J. Eischeid, 2013: Do extreme climate events require extreme forcings? *Geophys. Res. Lett.*, **40**, 3440–3445, doi:10.1002/grl.50657.

Lorenz, D. J., J. A. Otkin, M. Svoboda, C. R. Hain, M. C. Anderson, and Y. Zhong, 2017: Predicting U.S. Drought Monitor states using precipitation, soil moisture, and evapotranspiration anomalies. Part I: Development of a nondiscrete USDM index. *J. Hydrometeor.*, **18**, 1943–1962, doi:10.1175/JHM-D-16-0066.1.

Luo, L., E. F. Wood, and M. Pan, 2007: Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *J. Geophys. Res.*, **112**, D10102, doi:10.1029/2006JD007655.

McEvoy, D. J., J. L. Huntington, J. F. Mejia, and M. T. Hobbins, 2016: Improved seasonal drought forecasts using reference evapotranspiration anomalies. *Geophys. Res. Lett.*, **43**, 377–385, doi:10.1002/2015GL067009.

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

Mishra, V. R., and A. K. Desai, 2005: Drought forecasting using stochastic models. *Stochastic Environ. Res. Risk Assess.*, **19**, 326–339, doi:10.1007/s00477-005-0238-4.

——, and ——, 2006: Drought forecasting using feed-forward recursive neural network. *Ecol. Modell.*, **198**, 127–138, doi:10.1016/j.ecolmodel.2006.04.017.

Mo, K. C., and B. Lyon, 2015: Global meteorological drought prediction using the North American Multi-Model Ensemble. *J. Hydrometeor.*, **16**, 1409–1424, doi:10.1175/JHM-D-14-0192.1.

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.

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 ——, 2015a: 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, 2015b: Facilitating the use of drought early warning information through interactions with agricultural stakeholders. *Bull. Amer. Meteor. Soc.*, **96**, 1073–1078, doi:10.1175/BAMS-D-14-00219.1.

Özger, M., A. K. Mishra, and V. P. Singh, 2012: Long lead time drought forecasting using a wavelet and fuzzy logic combination model: A case study in Texas. *J. Hydrometeor.*, **13**, 284–297, doi:10.1175/JHM-D-10-05007.1.

Pan, M., X. Yuan, and E. F. Wood, 2013: A probabilistic framework for assessing drought recovery. *Geophys. Res. Lett.*, **40**, 3637–3642, doi:10.1002/grl.50728.

Quan, X.-W., M. P. Hoerling, B. Lyon, A. Kumar, M. A. Bell, M. K. Tippett, and H. Wang, 2012: Prospects for dynamical prediction of meteorological drought. *J. Appl. Meteor. Climatol.*, **51**, 1238–1252, doi:10.1175/JAMC-D-11-0194.1.

Sen, Z., and V. K. Boken, 2005: Techniques to predict agricultural droughts. *Monitoring and Predicting Agricultural Drought*, V. K. Boken, A. P. Cracknell, and R. L. Heathcote, Eds., Oxford University Press, 40–54.

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.

Steinemann, A., 2003: Drought indicators and triggers: A stochastic approach to evaluation. *J. Amer. Water Resour. Assoc.*, **39**, 1217–1233, doi:10.1111/j.1752-1688.2003.tb03704.x.

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.

Thober, S., R. Kumar, J. Sheffield, J. Mai, D. Schäfer, and L. Samaniego, 2015: Seasonal soil moisture drought prediction over Europe using the North American Multi-Model Ensemble (NMME). *J. Hydrometeor.*, **16**, 2329–2344, doi:10.1175/JHM-D-15-0053.1.

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.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.*, **107**, 4429, doi:10.1029/2001JD000659.

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.

Yuan, X., and E. F. Wood, 2013: Multimodel seasonal forecasting of global drought onset. *Geophys. Res. Lett.*, **40**, 4900–4905, doi:10.1002/grl.50949.

——, ——, N. W. Chaney, J. Sheffield, J. Kam, M. Liang, and K. Guan, 2013a: Probabilistic seasonal forecasting of African drought by dynamic models. *J. Hydrometeor.*, **14**, 1706–1720, doi:10.1175/JHM-D-13-054.1.

——, ——, J. K. Roundy, and M. Pan, 2013b: CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. *J. Climate*, **26**, 4828–4847, doi:10.1175/JCLI-D-12-00683.1.