



## Technical Note 15-02

### Early warning of summer drought over Texas and the south central United States: *spring conditions as a harbinger of summer drought*

D. Nelun Fernando<sup>1</sup>, Ph.D., Rong Fu<sup>2</sup>, Ph.D., Ruben S. Solis<sup>1</sup>, Ph.D., P.E., Robert E. Mace<sup>1</sup>, Ph.D., P.G., Ying Sun<sup>2</sup>, Ph.D., Binyan Yang<sup>2</sup>, and Bing Pu<sup>2</sup>, Ph.D.

<sup>1</sup> Water Science & Conservation, Texas Water Development Board

<sup>2</sup> Department of Geological Sciences, The University of Texas at Austin

January 2015

## Table of Contents

Acronyms .....	3
List of Figures .....	4
Executive Summary.....	6
1. Introduction .....	9
2. Datasets and methods .....	13
2.1: Datasets .....	13
2.2: Statistical prediction tool .....	15
3. Results .....	19
3.1: Predictor season .....	19
3.2: Predictand variation.....	19
3.3: Statistical prediction skill versus baseline predictability and dynamical predictions .....	22
3.4: Lead times for forecast .....	24
3.5: Indicator sensitivity to data input .....	27
3.6: Comparison of the 2014 indicator and official NOAA forecasts from the NMME.....	28
4. Conclusions and recommendations for further study .....	30
Acknowledgments.....	32
References .....	33
Appendix 1: Table with variance explained by the EOF modes .....	38
Appendix 2: Canonical Correlation Analysis input window in the Climate Predictability Tool .....	39
Appendix 3: Sample Rotated EOF input data for Canonical Correlation Analysis using the Climate Predictability Tool .....	40

## Acronyms

2AFC	2 Alternate Forced Choice
CAMS_OPI	Climate Anomaly Monitoring System (CAMS) and Out-going longwave radiation Precipitation Index (OPI)
CCA	Canonical Correlation Analysis
CFSR	Climate Forecast System Reanalysis
CFSv2	Climate Forecast System version 2
CIN	Convective Inhibition
CPC	Climate Prediction Center
CPT	Climate Predictability Tool
ECV-SM	Essential Climate Variable-Soil Moisture
ENSO	El Nino Southern Oscillation
EOF	Empirical Orthogonal Function
ERA-Interim	Reanalysis of the global atmosphere produced by the European Center for Medium-Range Weather Forecasts
JJA	June through August
MAI3CPASM	3-hourly instantaneous meteorology product
MAIMNPANA	Monthly means of analyzed state meteorology product
MAM	March through May
MATMNXLND	Monthly mean land surface diagnostics product
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MJJ	May through July
NASA	National Aeronautics and Space Administration
NCDC	National Climate Data Center
NCEP	National Centers for Environmental Prediction
NMME	North American Multi-model Ensemble
NOAA	National Oceanographic and Atmospheric Administration
REOF	Rotated Empirical Orthogonal Function
ROC	Relative Operating Characteristics
SC US	South Central United States
SPI6	Six-monthly Standardized Precipitation Index
SPI12	Twelve-monthly Standardized Precipitation Index
SVD	Singular Value Decomposition

## List of Figures

- Figure 1:** Time series of average anomalous monthly (a) rainfall, (b) maximum temperature at the surface and relative humidity, (c) temperature at 700 hectopascals and convective inhibition, (d) streamfunction at 200, 500 and 850 hectopascals, and (e) relative vorticity at 200, 500 and 850 hectopascals for the 7 strongest droughts over the southern Great Plains from 1948-2013. 11
- Figure 2:** Skill level as depicted by the Spearman's Correlation (top), Relative Operating Characteristics Area (below-normal) (middle), and two-alternative Forced Choice (forecast categories) (bottom), using April (left), March (center) and March through May seasonal average (right) initial conditions for the predictor fields. 21
- Figure 3:** Skill level as depicted using the Spearman's Correlation (top), Relative Operating Characteristics Area (below-normal) (middle), and two-alternative Forced Choice (forecast categories) (bottom), using May through July seasonal rainfall, July through August seasonal rainfall (center), and July SPI6 (right) as the predictand. 21
- Figure 4:** Time series plot of observed (blue bar) six-monthly Standardized Precipitation Index for July (July SPI6) and indicator-predicted July SPI6 at the grid point 33°N and 99°W for the period 1982-2014. The correlation coefficient between the two time series is depicted as  $r$ . 22
- Figure 5:** Comparison of Spearman correlation between predicted and observed six-monthly Standardized Precipitation Index (SPI6) for July (left), baseline predictability represented as the autocorrelation function (ACF) of April SPI6 with July SPI6 (center), and the difference between prediction skill and baseline predictability. 23
- Figure 6:** Forecast skill depicted using Spearman's correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and two-Alternate Forced Choice (forecast categories) (bottom), for the indicator forecast of the six-monthly Standardized Precipitation Index for July (July SPI6) using April initial conditions (top), and the North American Multi-model Ensemble mean three-month lead prediction of July SPI6 (bottom). 23
- Figure 7:** Prediction skill for forecasts using April forecast fields initialized with January, February and March observations, and for forecasts made using observed April fields. Prediction skill is expressed as the percentage of grid points with the 2-Alternate Forced Choice (forecast categories) exceeding 50 percent (blue line), percentage of grid points with Relative Operating Characteristics Area (below-normal) exceeding 0.5 (green line), and number of grid points with Spearman's correlation exceeding 0.6 (red line). 25

- Figure 8:** Forecast skill depicted using Spearman’s correlation (left), 2-Alternate Forced Choice (forecast categories) (center), and Relative Operating Characteristics Area (below-normal) (right), using three-month (six-month lead forecast), two-month (five-month lead forecast), one-month (four-month lead forecast) April forecast fields and realtime April observations (three-month lead forecast) as initial conditions. 26
- Figure 9:** Six- to three-month lead categorical forecasts of the 6-monthly Standardized Precipitation Index for July (SPI6) in 2011 (top, from left to right), and observed July SPI6 in 2011 (bottom). 27
- Figure 10:** Forecast skill depicted using Spearman’s correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and 2-Alternate Forced Choice (forecast categories) (bottom), using April predictor fields from the Climate Forecast Reanalysis (left) and from the Modern-Era Retrospective Analysis for Research and Applications (right). 29
- Figure 11:** Forecast skill depicted Spearman’s correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and 2-Alternate Forced Choice (forecast categories) (bottom), where the soil moisture field is varied using Climate Forecast System Reanalysis volumetric soil moisture (left) and Essential Climate Variable-Soil Moisture data (right). 29
- Figure 12:** Statistical prediction (left), official forecast using the North American Multi-model Ensemble, and observed precipitation anomalies for the 2014 May through July season. 30

## Executive Summary

Strong summer droughts over the Southern Great Plains region are often characterized by rapid intensification in the late-spring and early-summer. The decreased rainfall in these drought years are coupled with strong increases in summertime temperature extremes, as for example, was the case with the 2011 drought over Texas and the Great Plains drought in 2012. Dynamic climate models failed to predict these summer droughts. This is largely due to model weaknesses in predicting summer rainfall, underestimating summer rainfall variance, and weaknesses in capturing soil moisture feedbacks. By contrast, climate models are more reliable in capturing the variability in large-scale circulation features and temperature during winter and spring.

Observations show that severe-to-extreme summer drought events over Texas are preceded by dry springs. Over the period 1895-2014, there were 13 severe-to-exceptional droughts (defined as the 12-monthly Standardized Precipitation Index for August  $\leq -1.2$ ) over Texas. Ninety two percent of these drought events were characterized by anomalously low rainfall in the spring (March through May). Dry springs cause anomalous high pressure systems and anticyclonic (clockwise) flow in the prevailing wind system, which lead to subsidence (sinking motion) in the atmosphere. Such subsidence persists through much of the summer and inhibits rainfall from convective (rising motion) processes. This work explored the potential predictability of strong summer droughts, and the feasibility of using a process-based empirical model to predict summer droughts, over the Southern Great Plains based on such persistent drought-inducing atmospheric circulation patterns and surface dryness in spring.

We developed a process-based statistical model to provide an early warning indicator of summer (meteorological) drought based on the anomalous large-scale middle tropospheric (that is, 500 hectopascals level, or approximately 5,500 meters above sea level) circulation, convective inhibition energy (a numerical measure in meteorology indicating the negative energy available in the environment to prevent development of convective weather systems), and land surface moisture

conditions in spring (March to May). We used the three aforementioned conditions in spring (March, April, and seasonal mean March through May conditions) as inputs to our statistical model to predict cumulative rainfall deficits or surplus (referred to as the cumulative rainfall anomalies hereinafter) during May-July and the six-month standardized precipitation index (a rainfall based drought index) for July.

Comparison of the hindcasts made using the statistical model with the observations shows that the model can predict summer droughts over Texas and the southern Great Plains region in spring with skill levels acceptable decision makers (~60 percent or higher) — particularly those tasked with drought emergency management. The drought indicator shows higher success rates in correctly predicting the occurrence of dry or wet summers than the baseline drought predictability (that is, autocorrelation of rainfall anomalies) and forecasts from dynamical models over south central, central, northern and eastern Texas, western Louisiana, most of Oklahoma and southern Kansas at three to six months lead time. In all realizations of the model, we find that the grid points with the highest skill scores lie within Texas. The best skill is achieved when using April initial conditions of the three predictor variables. As a rule of thumb, if there is a high pressure system at 500 hectopascals over the Southern Great Plains, very high values of convective inhibition, and dry land-surface conditions in the region in April, there is a strong probability of an impending intense summer drought over this region.

We made a first forecast for the summer of 2014 using observed April fields. The forecast showed abnormally wet conditions, which better matched observed conditions than the dynamical model forecast from the National Oceanic and Atmospheric Administration's Climate Prediction Center for this region.

We developed a combined dynamic-statistical prediction approach to assess the feasibility of providing an early warning of summer drought at the four- to six-month lead time. This approach uses the ensemble mean dynamic prediction for April conditions, initialized by observed conditions in

January, February and March, respectively, as input to the statistical model to predict May through July rainfall anomalies, or the six-monthly Standardized Precipitation for July, at the 6-month, 5-month and 4-month lead time. The categorical seasonal forecasts (that is, probabilistic estimates of whether a season will be below-, near- or above-normal) from the indicator provide added information on drought susceptibility for upto six-month lead time with the skill levels acceptable to decision makers. Of notable interest is the ability of this combined dynamic and statistical approach to hindcast the 2011 summer drought in January with upto six month lead time. This implies that the 2011 summer drought over Texas could have been predicted in January 2011.

Given the performance of the drought early warning indicator over Texas, we will explore the possibility for providing summer drought forecasts from January onwards to the Texas Drought Preparedness Council, state emergency management initiatives, and water planners. Such forecasts would ideally be made available to the public through the TWDB's drought web page and the Water Data for Texas web site.



## 1. Introduction

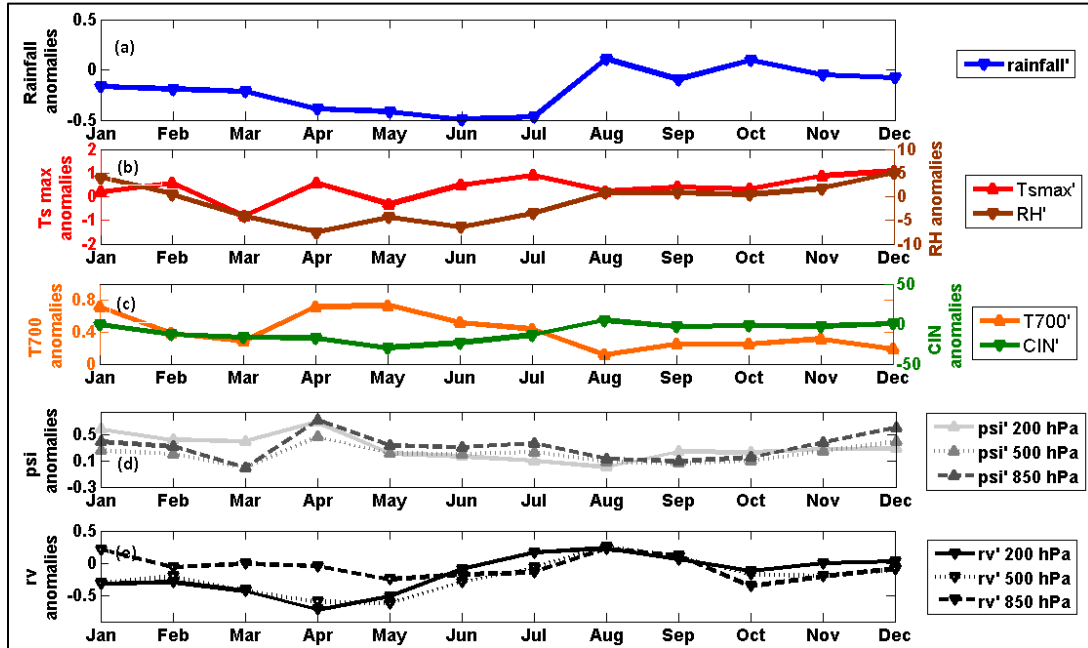
Strong summer droughts over the U.S. Southern Great Plains region (110°W-92°W and 24°N-40°N) are often characterized by rapid intensification in the late-spring and early-summer. The decreased rainfall in these drought years are coupled with strong increases in summertime temperature extremes, as shown, for example, by the 2011 drought over Texas and the 2010 Great Plains drought. Dynamic climate models did not predict these summer droughts (Seager and others, 2013; Hoerling and others, 2013; Kumar and others, 2013). They are also unable to provide more skill than that provided by the autocorrelation of rainfall anomalies permits, particularly during summer over the U.S. Great Plains (Quan and others, 2012). This is in part due to model limitations in representing summer thunderstorms and land-surface feedbacks, which occur at scales much smaller than the current spatial resolution of climate models, and due to models underestimating rainfall variance (Kam and others, 2014). By contrast, climate models are more reliable in capturing the variability in large-scale circulation features and temperature during winter and spring.

What causes drought onset over the US Southern Great Plains? Our prior drought research based on observational data indicates that severe-to-extreme summer drought events over Texas are preceded by dry springs (Fernando, 2014). The phenomenon of dry springs preceding strong summer droughts was also documented earlier by Namias (1982) for the U.S. Great Plains. The strongest drought events that occurred over the southern Great Plains since 1948 show some salient characteristics. A marked decrease in rainfall from April through July is led by sharp increases in temperature at the surface and at 700 hectopascals, and a decrease in relative humidity in April. These changes lead to strong convective inhibition energy (CIN) — a numerical measure in meteorology indicating the negative energy available in the environment to prevent development of convective weather systems — in April and May. It is also accompanied by an increased geopotential height and persistent negative vorticity (indicating subsidence) at mid-to higher levels of the atmosphere from April

to June (Figure 1). The persistence of rainfall anomalies from winter through early-spring leads to significant cumulative soil moisture deficits, a reduction in evapotranspiration, and an increase in sensible heating to balance the decrease in evapotranspiration. This leads to increased temperature at the surface. Drought years with persistent rainfall anomalies from winter through spring are characterized by strong westerly winds in April, particularly at 850 hectopascals (approximately 1 kilometer above the surface). These winds advect warm dry air eastwards over Texas from the Mexican plateau. Dry air advection, and the associated cap inversion, cause an increase in temperature at 700 hectopascals, and sharp increases in convective inhibition in April and May. These anomalous conditions are also the main causes of drought in the summer of Texas (Myoung and Nielsen-Gammon, 2010).

Droughts in Texas are generally associated with colder than normal sea surface temperatures in the equatorial Pacific corresponding to La Niña events (Ropelewski and Halpert 1986; 1987; Schubert and others, 2009). La Niña-induced cooler SST anomalies are usually established in the fall, which contribute to winter drought over Texas. This is because a La Niña event induces the poleward displacement of the sub-tropical jet stream, which deflects winter storm tracks to the north of their climatological location and causes a reduction of precipitation over the U.S. Southern Great Plains (Eichler and Higgins, 2006; Kousky and Ropelewski, 1989). The reduction in winter precipitation leads to soil moisture deficits and increased surface temperature. The increase in surface temperature in turn leads to 1000-500 hectopascals geopotential height thickness over the southern United States, which could explain the presence of a mid-tropospheric high in the spring over the region.

However, La Niñas do not always cause summer droughts. Historically, about 11 percent of La Niña-induced winter droughts over Texas ended in spring. Those that persisted through spring tended to develop into severe [ $-1.2 \leq$  Standardized Precipitation Index  $\leq -1.5$ ] to extreme droughts (Standardized Precipitation Index  $\leq -1.5$ ) in summer (Fernando and others, in preparation).



**Figure 1:** Time series of average anomalous monthly (a) rainfall, (b) maximum temperature at the surface and relative humidity, (c) temperature at 700 hectopascals and convective inhibition, (d) streamfunction at 200, 500 and 850 hectopascals, and (e) relative vorticity at 200, 500 and 850 hectopascals for the 7 strongest droughts over the southern Great Plains from 1948–2013.

What causes the persistence of a summer drought, once it is established? Many studies have emphasized the importance of regional land-atmosphere feedbacks in affecting drought persistence (Lyon and Dole, 1995; Hong and Kalnay, 2002). Dry soils represent the cumulative impact of deficit precipitation and provide a positive feedback, further enhancing precipitation deficits (Mueller and Seneviratne, 2012). Modeling analyses identified the Central United States, including much of Texas, as an area of strong coupling between soil moisture and precipitation (Koster and others, 2004). Fernando and others (in-prep), in investigating whether drought memory is due to the persistence of remote forcing or due to land surface feedbacks, found that dry soil moisture anomalies over the South Central United States might have a stronger influence on positive 500 hectopascals height anomalies 2 to 3 weeks later than that of remote forcing in the late-spring/early-summer.

Building on the above understanding of the causes that initiate and maintain persistent drought, we use the following circulation and land surface fields in spring (April, March, and average March through May) to build a statistical model to predict cumulative summer rainfall (e.g. May through July seasonal rainfall or July SPI6):

- a) Geopotential height at 500 hectopascals
- b) Difference in temperature at 700 hectopascals and surface dewpoint (proxy for convective inhibition, Myoung and Nielsen-Gammon, 2010)
- c) Soil moisture

We address the following questions: Can summer drought over the southern Great Plains be predicted empirically? How sensitive is the statistical prediction tool to predict season? Can such a prediction scheme outperform baseline drought predictability and drought predictions from dynamical models? What lead times are skillful for issuing the spring forecast of summer drought? How sensitive is the statistical prediction tool to data input?

## 2. Datasets and methods

### 2.1: Datasets

We used monthly 500 hectopascals geopotential height, temperature at 700 hectopascals, 2-meter dewpoint temperature and 0–10 cm depth liquid volumetric soil moisture (non-frozen) for April, March, and March through May (MAM) from the National Centers for Environmental Prediction Climate Forecast System Reanalysis (CFSR, Saha and others, 2010) for 1982–2010 and derive monthly values of the same fields from 6-hourly Climate Forecast System version 2 (CFSv2) realtime data for the period 2011–2013. We accessed the CFSR and CFSv2 data through the Data Library of the International Research Institute for Climate and Society (<http://iridl.ldeo.columbia.edu>). We used the Climate Forecast System Reforecasts (Saha and others, 2014) for hindcasts at the three-month (3.5), two-month (2.5), and one-month (1.5) lead times of April 500 hectopascals geopotential height, temperature at 700 hectopascals, 2-meter dewpoint temperature and 0–10cm depth liquid volumetric soil moisture for the time periods 1982–2013. The data from 1982–2010 are from the Climate Forecast System Reanalysis and Reforecasts archive and the data from 2011–2014 were from the 6-hourly CFSv2 realtime data. The latitudinal means were removed from the 500 hectopascals geopotential height data.

We used 500 hectopascals geopotential height and temperature at 700 hectopascals from monthly means of analyzed state meteorology product (MAIMNPANA); relative humidity and air temperature at 925 hectopascals from the 3-hourly instantaneous meteorology product (MAI3CPASM); and root zone soil wetness data from the monthly mean land surface diagnostics product (MATMNXLND) for April 1982–2013 from the National Aeronautics and Space Administration (NASA) Modern Era Retrospective-analysis for Research and Applications (MERRA, Rienecker and others, 2011). Monthly dewpoint is derived using the 3-hourly relative humidity and air temperature fields. We access MERRA data using OPeNDAP/DODS access to the NASA GES DISC GrADS data server

(<http://goldsmr3.sci.gsfc.nasa.gov/dods>). The latitudinal means are removed from the 500 hectopascals geopotential height data.

We used daily soil moisture, aggregated to monthly values, from the merged active and passive microwave retrievals product of the European Space Agency Climate Change Initiative Essential Climate Variable Soil Moisture (ECV-SM) dataset version 1 (<http://www.esa-soilmoisture-cci.org>; Liu and others, 2012; Liu and others, 2011; Wagner and others, 2012) for the period 1982–2010. These data are available at 25 km resolution.

We obtained the ensemble mean of the three-month lead time May through July (MJJ) rainfall, for the period 1982–2010, using all the ensemble members (listed in parentheses) of seven models participating in the North American Multi-model Ensemble Project (NMME, Kirtman and others, 2013). These models are the CMC1-CanCM3 (10), CMC2-CanCM4 (10), COLA-RSMAS-CCSM3 (6), GFDL-CM2p1-aer04 (10), GFDL-CM2p5-FLOR-A06 (12), NASA-GMAO-062012 (12), and CFSv2 (28). We also obtained ensemble mean monthly rainfall for 1982–2010 from the NMME to derive hindcasts of the Standardized Precipitation Index (SPI6).

We obtained observed values of July SPI6, for the period 1982–2014, from SPI-CAMS\_OPI\_six-month dataset ([http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/.SPI-CAMS\\_OPI\\_six-month/](http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/.SPI-CAMS_OPI_six-month/)) made available through the International Research Institute for Climate and Society. The SPI analysis uses the CAMS\_OPI [Climate Anomaly Monitoring System (CAMS) and Out-going longwave radiation Precipitation Index (OPI), Janowiak and Xie, 1999] satellite and raingauge merged data product from the Climate Prediction Center (CPC, [http://www.cpc.ncep.noaa.gov/products/global\\_precip/html/wpage.cams\\_opi.html](http://www.cpc.ncep.noaa.gov/products/global_precip/html/wpage.cams_opi.html)).

We used monthly  $1^{\circ} \times 1^{\circ}$  rainfall from the CPC global land precipitation dataset (Chen and others, 2002) to derive seasonal rainfall for MJJ and June through August (JJA) for period 1982-2014.

All data fields cover the domain 24N-40N and 110W to 92W. We refer to this domain as the South Central United States. All atmospheric and soil moisture fields were aggregated to the monthly

time step when such aggregated was needed prior to extracting the fields for April. All data fields were regridded to 1° horizontal resolution. At this resolution, there are 323 grid cells over the study domain.

## 2.2: Statistical prediction tool

All fields were converted to standardized anomalies, using the base period 1982–2012, prior to analysis. Original predictor inputs could have multicollinearity, noise, and variance irrelevant to the drought prediction. Therefore, we first applied Multivariate Empirical Orthogonal Function Analysis (EOF)<sup>1</sup> to the three predictor fields as a way to filter out noise and highlight the most coherent spatial and temporal variances. We retained the first two EOF modes, accounting for at least 70 percent of the variance in the predictor fields, to minimize the potential multicollinearity of the original predictor fields. The two retained EOF modes are linear combinations of the three predictor variables and represent the highest fraction of co-variability in the original predictor fields.

The data vector we use has the three predictor input datasets. The original 3-dimensional input data of 17 latitude points, 19 longitude points and 33 years (1982–2014) are transformed into a 10659 x 1 matrix prior to multivariate EOF analysis using the Singular Value Decomposition (SVD) algorithm. Singular Value Decomposition is a method used to compute EOFs by factorizing the data matrix. The two EOF modes we retained were subjected to Rotated Empirical Orthogonal Function Analysis (REOF, Richman 1986). Details on the variance explained by each EOF mode for every iteration of the model run used in this study are given in Annexure 1. The technique is designed to extract components from unrotated EOFs as individual portions of variance belonging to empirically distinct inter-correlational

---

<sup>1</sup> Empirical Orthogonal Function analysis is a data compression technique such that a dataset containing a large number of samples is reduced to a dataset that captures the dominant modes (or correlated variance), which explain a large fraction of the squared total co-variance among these fields. The new variables are linear combinations of the original variables and represent the highest possible proportion of co-variability found in the original dataset (Wilks, 2006, pg. 463).

complexes in the dataset (Barnston and Livezey, 1987). In rotation, principal component weights are designed to maximize the variance in the selected modes. We used the Varimax (Richman, 1986 and references therein) to rotate the EOFs is an orthogonal rotation (Richman, 1986). Orthogonal rotations must satisfy the equation:

$$B = AT^{-1}$$

Where,

B –  $k \times m$  rotated EOF loadings matrix

A –  $k \times m$  initial unrotated loading matrix

T –  $m \times m$  orthonormal transformation matrix

As T is an orthonormal matrix,

$$TT^T = T^T T = I$$

Varimax attempts to simplify the EOF modes of B to achieve simple structure (Richman, 1986) such that:

$$V^* = \frac{[k \sum_{i=1}^k (b_{ij}^2)^2 - (\sum_{i=1}^k b_{ij}^2)^2]}{k^2}$$

Where,  $j = 1 \dots V^*$

$V^*$  - simplicity of an EOF mode

$m$  – number of modes

$b$  – EOF loadings

We input the spatial loadings of the rotated EOF modes as predictor variables to a Canonical Correlation Analysis (CCA) model. The predictand variable is summer rainfall (i.e. either July SPI6 or seasonal



average MJJ rainfall). CCA identifies a sequence of pairs of patterns in two multivariate datasets. Linear combinations of the original data are produced by projecting the original data onto the identified patterns. New variables — known as “canonical variates” — that maximize interrelationships between the two data sets are then identified [Wilks, 2006, pg. 509]. CCA can be used as a statistical forecasting technique if one of the input data fields is observed prior to the other field (e.g. the predictand or the ‘y’ field) (von Storch and Zwiers, 2002). Such application of CCA has been undertaken in forecasting SSTs (Landman and Mason, 2001), in predicting seasonal temperatures over land (Shabbar and Barnston, 1996), and in predicting ENSO episodes (Barnston and van den Dool, 1993). When using CCA for forecasting purposes, a simple linear regression model is constructed that relates the predictand canonical variates  $w_m$  to the predictor canonical variates  $v_m$  (Wilks, 2006, pg. 509).

$$w_m = \hat{\beta}_{0,m} + \hat{\beta}_{1,m}v_m$$

where  $m = 1, 2, \dots, M$  ( $M$  – the number of canonical pairs)

We used the Climate Predictability Tool (CPT, <http://iri.columbia.edu/our-expertise/climate/tools/cpt/>) to run the CCA model. Forecast skill assessment was undertaken using cross-validation (Michaelson, 1987; Barnston and Ropelewski, 1992) over a 24 year training period from 1982 through 2005. We use a cross validation window of three years. The strength of the predictor fields was assessed based on the goodness-of-fit between the cross-validated forecasts and the observation time series. Goodness-of-fit is a measure of the average correlation between the cross-validated forecasts and the observation time series. The measure is reported for every possible combination of predictor and predictand modes. When the goodness-of-fit value is closer to 1 the correlation between the predictor and predictand fields is stronger.

Model skill was assessed in two steps. First, the skill of deterministic forecasts of seasonal cumulative rainfall or the six-monthly Standardized Precipitation Index was assessed using Spearman’s Ranked Correlation between observations and predictions at each grid point. Second, the skill of

probabilistic forecasts was assessed using the Two Alternate Forced Choice score for forecast categories (2AFC; Mason and Weigel, 2009), and the area under the Relative Operating Characteristics curve for the below-normal category (ROC; Mason and Graham, 2002) at each grid point. The 2AFC score ranges from 0 to 100 percent. If the 2AFC score is greater than 50 percent, it means that the forecast is able to discriminate beyond random guessing. The value of 0.5 indicates that the forecast cannot do better than a forecast made using purely climatology. The ROC score also ranges from 0 to 1 and only a score greater than 0.5 indicates that the forecast is able to discriminate beyond random guessing.

To address the question of whether the statistical prediction improves on baseline drought predictability, we compared the Spearman's Ranked Correlation for observed versus predicted July SPI6 with persistence (measured as the autocorrelation function) of April SPI6. To assess if the statistical prediction improves on dynamical prediction skill, we also compared the 2AFC and ROC skill from three-month lead NMME forecasts of MJJ precipitation with the same skill scores for forecast MJJ precipitation from the indicator.

To address the question of what season and lead times are feasible for issuing the spring prediction of summer drought conditions, we ran the model using observed predictor fields for March, April and MAM and with 3-, 2-, and one-month forecasts and real-time observations of the predictor fields. For example, if the optimum lead-time for the predictor is April, we used January (three-month), February (two-month), and March (one-month) forecasts and observed values of April mean fields as the predictors versus either July SPI6 or MJJ mean rainfall as the predictand. The best lead times were assessed based on the three skill metrics listed above.

## 3. Results

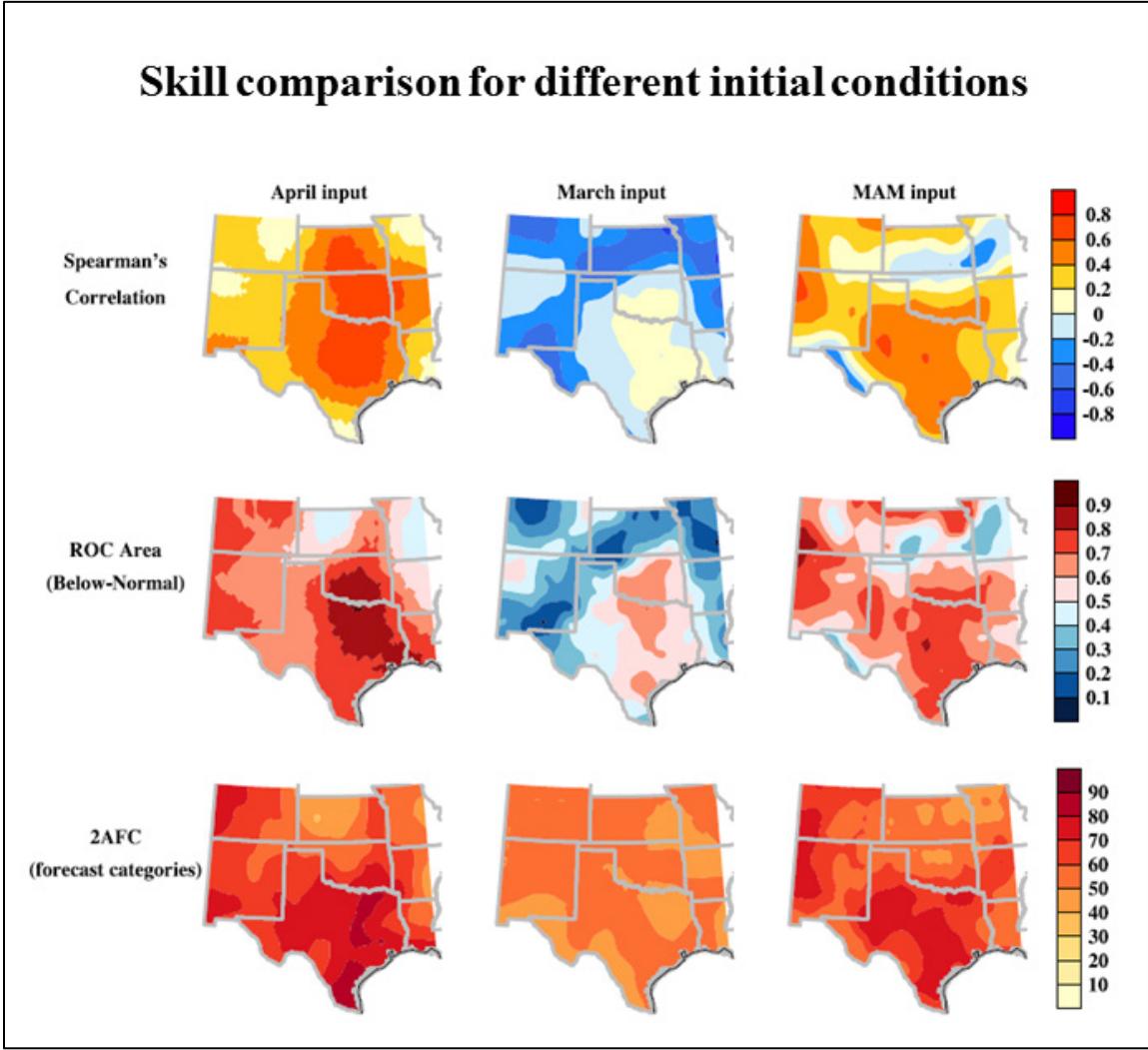
### 3.1: Predictor season

The best skill is achieved when using April initial conditions of the three predictor variables — that is, 500 hectopascals geopotential height, difference in temperature between 700 hectopascals and dewpoint at the surface, and soil moisture — to predict July SPI6 (Figure 1, first column). March initial conditions yields poor skill. Using March through May (MAM) average initial conditions improves the prediction skill compared to using only March initial conditions. However, it still shows a lower skill compared to that obtained using April initial conditions.

### 3.2: Predictand variation

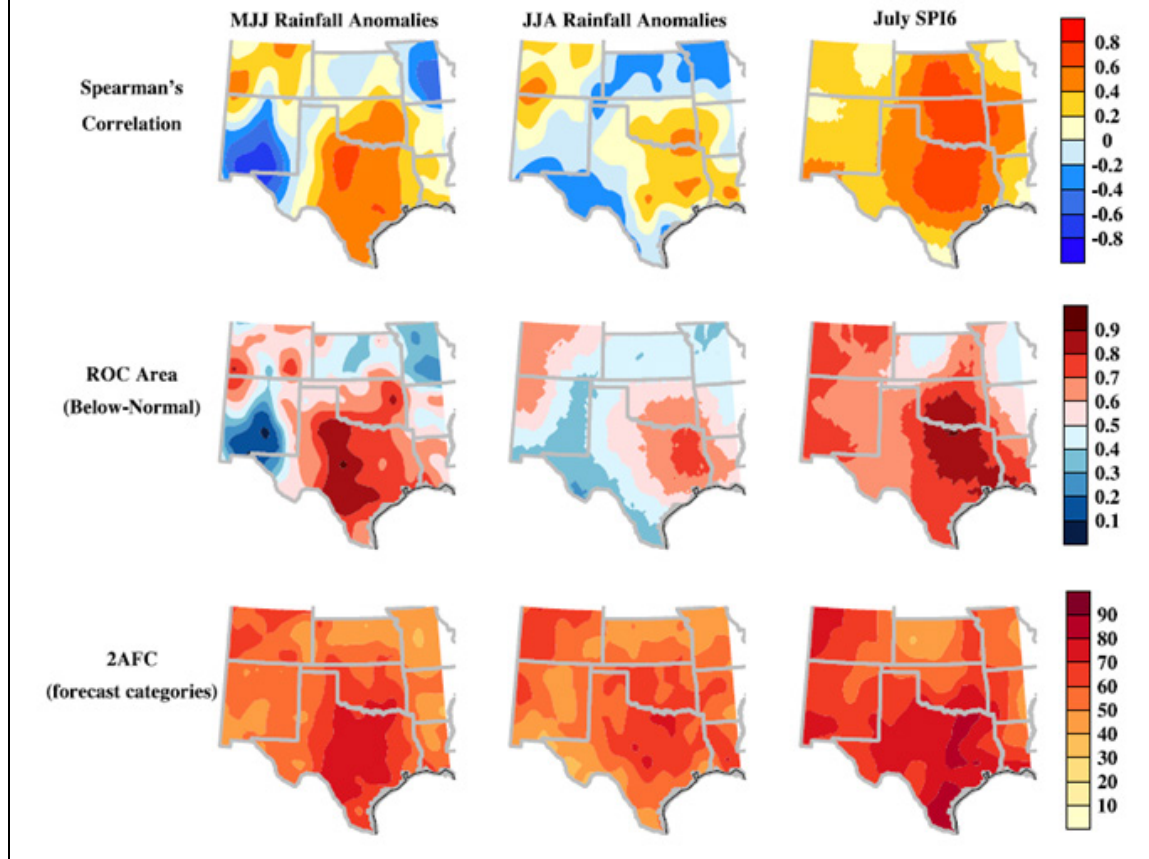
As a first step, we use both July SPI6 and MJJ seasonal rainfall as the predictand (that is, the variable that is predicted). The former is a commonly used drought indicator, and includes comprises a measure of rainfall departures from the mean over the preceding 6 month period. The latter, being a measure of the cumulative rainfall over MJJ, permits a better assessment of the prediction skill of our empirical model. For both predictands, our model shows better skill than using JJA seasonal rainfall (Figure 2). The skill is higher with July SPI6, particularly over the domain  $100^{\circ}\text{W}$ - $96^{\circ}\text{W}$  and  $32^{\circ}\text{N}$ - $38^{\circ}\text{N}$ . The reasons for poor skill when using JJA as the predictand could be due to the influence of tropical synoptic systems in August.

A time series plot at a grid point falling within the region with high skill shows that the indicator captures very well the observed sign (red line with circle) of the July SPI6 (i.e. whether wet or dry) for the period of analysis (Figure 4), although the severity of the events is not captured as well. This is an area that needs further study to improve skill at the local level in capturing the magnitudes of the observed drought events.

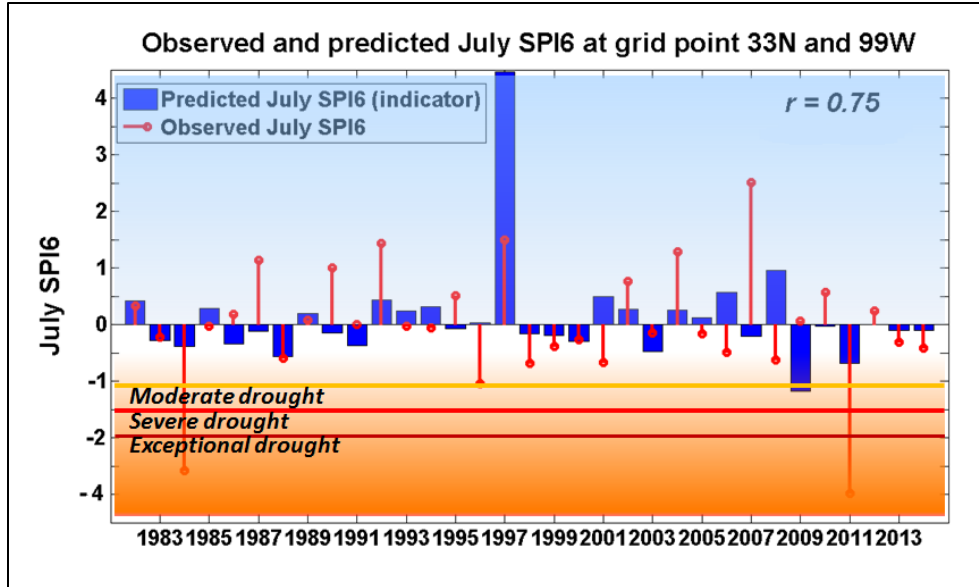


**Figure 2:** Skill level as depicted by the Spearman's Correlation (top), Relative Operating Characteristics Area (below-normal) (middle), and two-alternative Forced Choice (forecast categories) (bottom), using April (left), March (center) and March through May seasonal average (right) initial conditions for the predictor fields.

## Skill comparison for different predictands



**Figure 3:** Skill level as depicted using the Spearman's Correlation (top), Relative Operating Characteristics Area (below-normal) (middle), and two-alternative Forced Choice (forecast categories) (bottom), using May through July seasonal rainfall, July through August seasonal rainfall (center), and July SPI6 (right) as the predictand.

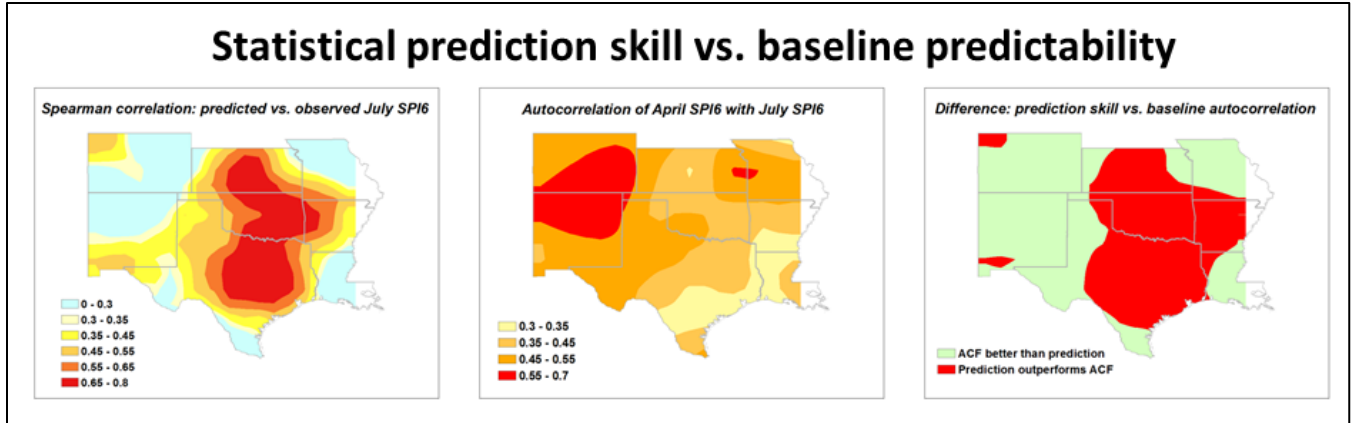


**Figure 4:** Time series plot of observed (blue bar) six-monthly Standardized Precipitation Index for July (July SPI6) and indicator-predicted July SPI6 at the grid point 33°N and 99°W for the period 1982-2014. The correlation coefficient between the two time series is depicted as  $r$ .

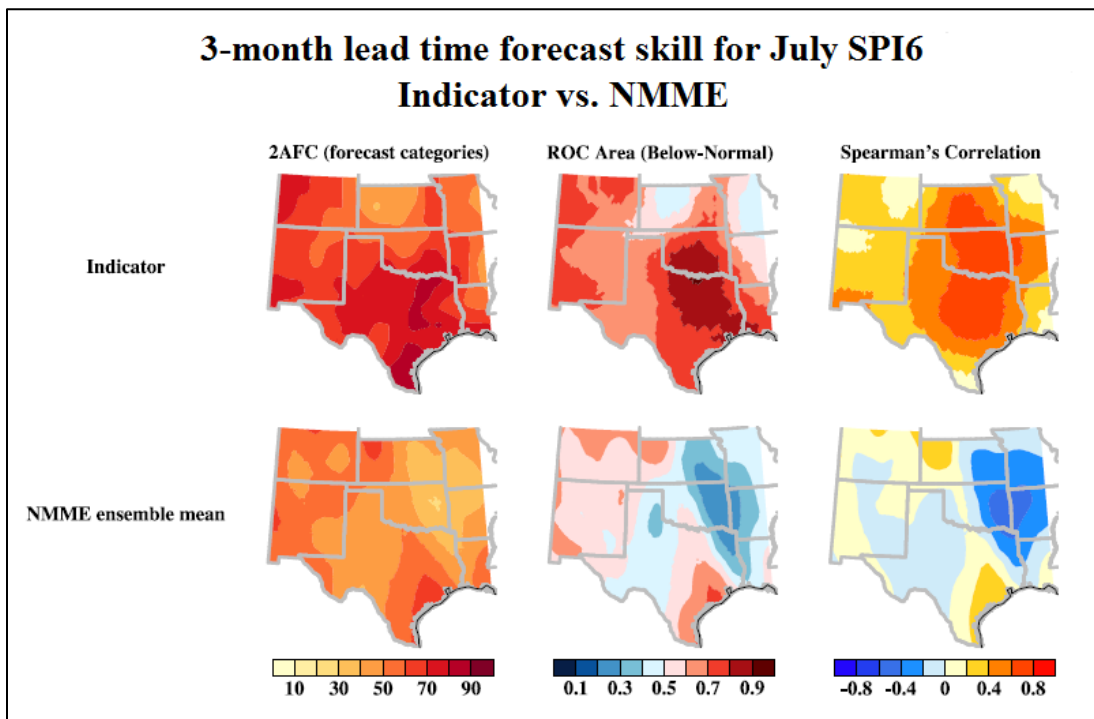
### 3.3: Statistical prediction skill versus baseline predictability and dynamical predictions

The skill of the indicator beats baseline predictability (red, right column) over south central, central northern and eastern Texas, western Louisiana, most of Oklahoma (except for the western-most edge of the panhandle) and southern Kansas. However, over west Texas, New Mexico and southeastern Colorado, the indicator does not provide more skill than that which can be achieved from the pure persistence of April SPI6.

In terms of deterministic forecast skill, the indicator outperforms the dynamical predictions of ensemble mean MJJ rainfall from the NMME over Texas much of the central and eastern regions of the study domain (Figure 6, right panel). There is small region in the north of the domain where the dynamical forecast outperforms the statistical forecast. However, In terms of the categorical forecast skill, the indicator clearly outperforms the dynamical prediction from the 2AFC (forecast categories) and ROC (below-normal) skill scores, particularly over Texas (Figure 6, left and center panels).



**Figure 5:** Comparison of Spearman correlation between predicted and observed six-monthly Standardized Precipitation Index (SPI6) for July (left), baseline predictability represented as the autocorrelation function (ACF) of April SPI6 with July SPI6 (center), and the difference between prediction skill and baseline predictability.



**Figure 6:** Forecast skill depicted using Spearman's correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and two-Alternate Forced Choice (forecast categories) (bottom), for the indicator forecast of the six-monthly Standardized Precipitation Index for July (July SPI6) using April initial conditions (top), and the North American Multi-model Ensemble mean three-month lead prediction of July SPI6 (bottom).

### 3.4: Lead times for forecast

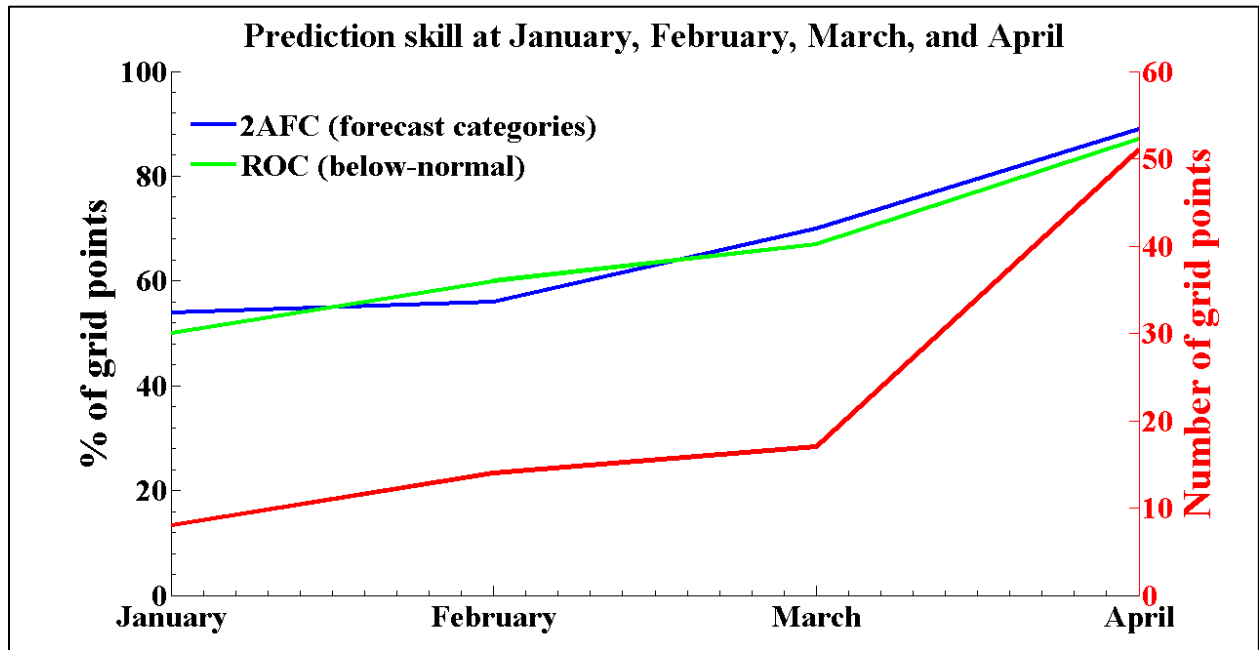
We studied what lead times were feasible for issuing the summer drought forecast using April forecast fields. Using a threshold of 0.6 or 60 percent, the skill level typically acceptable to emergency managers in the State of Texas<sup>2</sup>, we find that the percentage of grid points with the 2AFC (forecast categories) greater than 60 percent increases from 54 percent for forecasts initialized with three-month lead time forecasts of April fields to 70 percent for forecasts initialized with one-month lead forecasts of April fields. This percentage increases to 89 percent for forecasts initialized with realtime April fields. The percentage of grid points with ROC (below-normal) greater than 0.5 increases from 50 percent for forecasts initialized with three-month lead forecasts of April fields, to 67 percent for the forecasts initialized with one-month lead forecasts of April fields, and 87 percent for forecasts initialized with realtime April fields (Figure 7). Thus, it is clear that the categorical forecasts issued using April forecast fields at the three-month lead-time onwards are able to discriminate beyond random guessing over at least 50 percent of the domain. This increases steadily as the forecast lead-time decreases. The region of best skill lies over north-central Texas and Oklahoma. Over much of Texas, the categorical forecasts provide added information on drought susceptibility from the three-month lead time onwards (Figure 8, warm colors). The grid points with the highest ROC and 2AFC scores lie within Texas. The deterministic forecast skill, represented here by the Spearman's correlation coefficient, shows that the number of grid points with correlation values exceeding 0.6 is at 8 for the forecasts initialized with three-month lead April forecast fields, 14 for forecasts initialized with two-month lead April forecast fields, 17 for forecasts initialized with one-month lead April forecast fields, and 51 for forecasts initialized with realtime April fields. Here too, the region of best skill (Figure 8, warm colors) is located over Texas and Oklahoma. The spatial distribution of the region with best skill expands southwards as we move through lead times 3- to realtime. As noted previously, the highest skill is found within the domain of 100°W-90°W and 32°N-38°N.

---

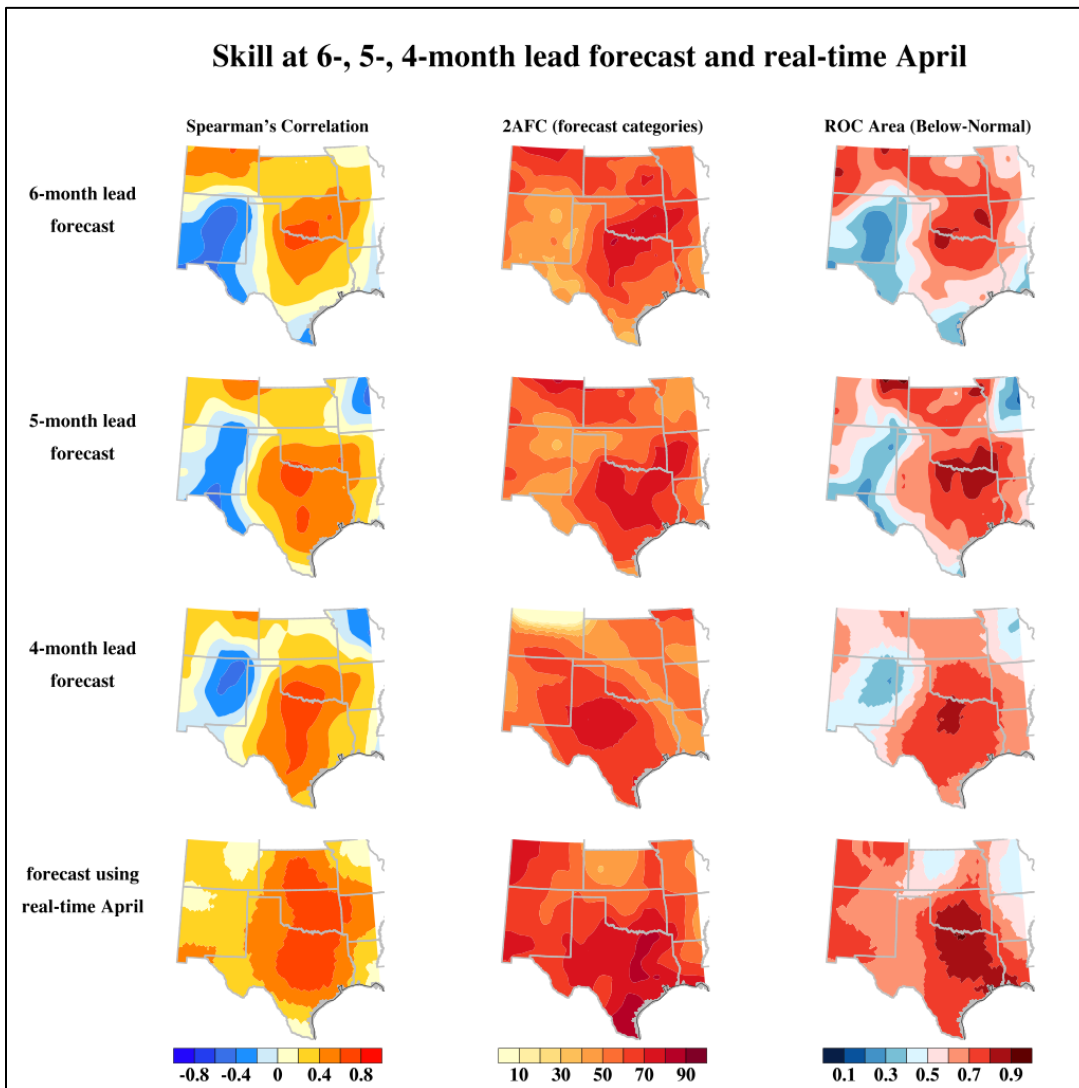
<sup>2</sup> Mike Bewley (personal communication, 2012)



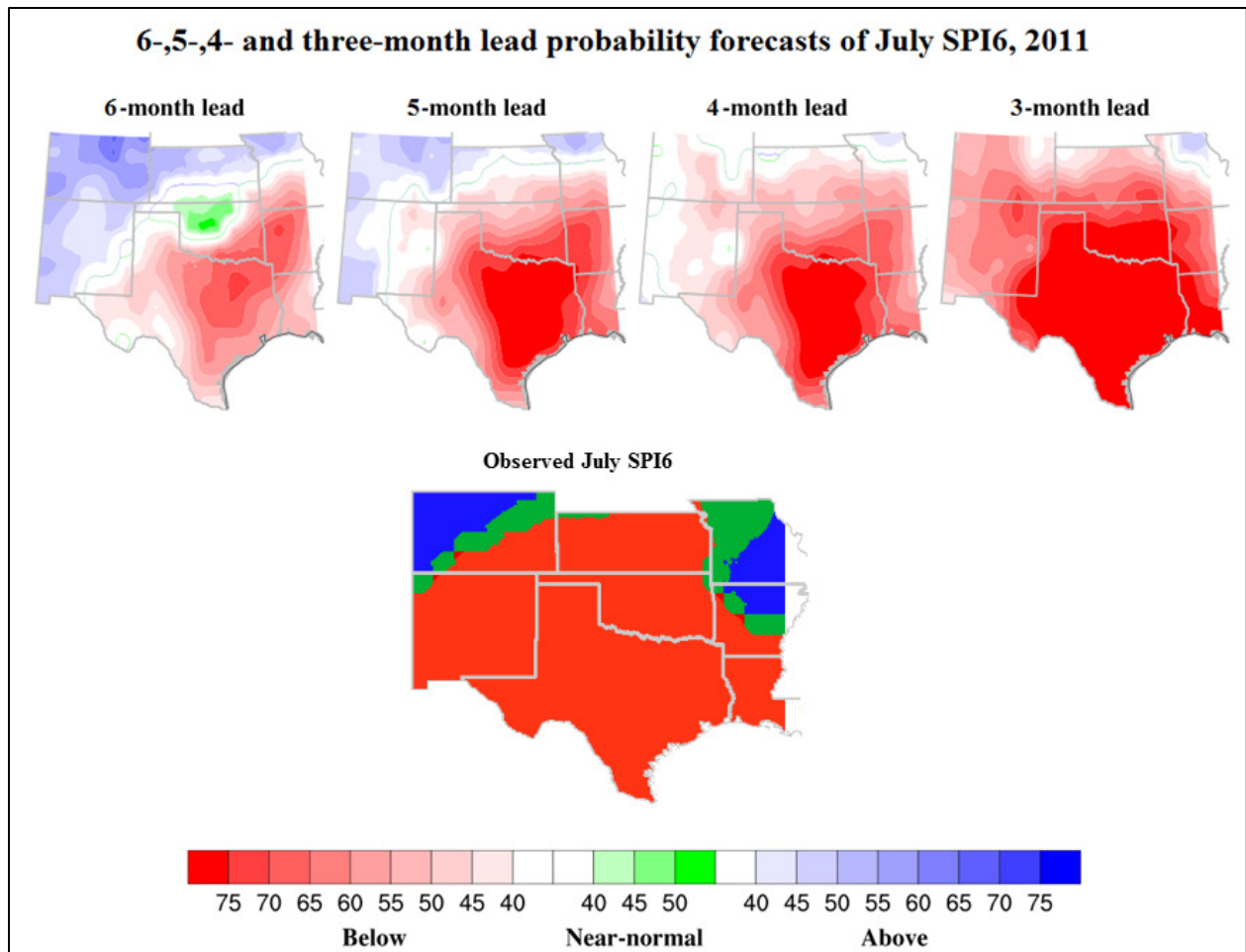
Forecasts of the 2011 drought made using 3-, 2-, one-month lead forecast fields, and realtime observations, of April circulation and soil moisture fields as initial conditions, show how the spatial extent and magnitude of below-normal probabilities increases as the lead time reduces (Figure 9). Of interest is the ability of the indicator to forecast the 2011 drought event even at six-month (January) lead time, using three-month lead forecasts of April predictor fields as initial conditions.



**Figure 7:** Prediction skill for forecasts using April forecast fields initialized with January, February and March observations, and for forecasts made using observed April fields. Prediction skill is expressed as the percentage of grid points with the 2-Alternate Forced Choice (forecast categories) exceeding 50 percent (blue line), percentage of grid points with Relative Operating Characteristics Area (below-normal) exceeding 0.5 (green line), and number of grid points with Spearman's correlation exceeding 0.6 (red line).



**Figure 8:** Forecast skill depicted using Spearman's correlation (left), 2-Alternate Forced Choice (forecast categories) (center), and Relative Operating Characteristics Area (below-normal) (right), using three-month (six-month lead forecast), two-month (five-month lead forecast), one-month (four-month lead forecast) April forecast fields and realtime April observations (three-month lead forecast) as initial conditions.



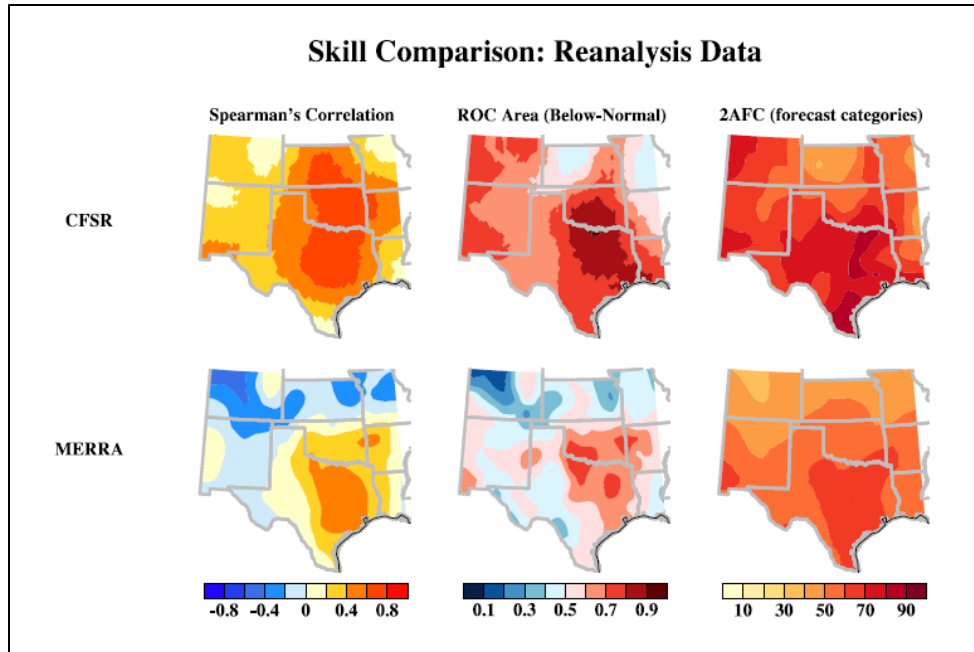
**Figure 9:** Six- to three-month lead categorical forecasts of the 6-monthly Standardized Precipitation Index for July (SPI6) in 2011 (top, from left to right), and observed July SPI6 in 2011 (bottom).

### 3.5: Indicator sensitivity to data input

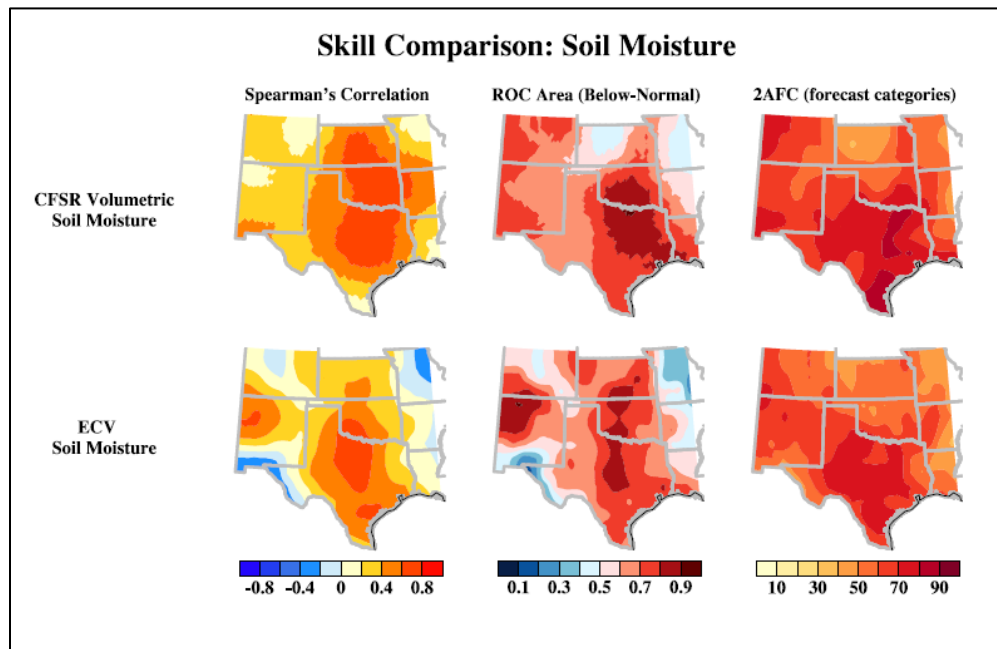
Comparison of indicator skill with April reanalysis predictor fields from MERRA versus CFSR shows that skill levels are higher using CFSR reanalysis data (Figure 10). When we varied the soil moisture field by replacing soil moisture from CFSR with observed soil moisture from ECV-SM soil moisture data set, we find modest improvements in skill over the western and northern regions (Figure 11).

### 3.6: Comparison of the 2014 indicator and official NOAA forecasts from the NMME

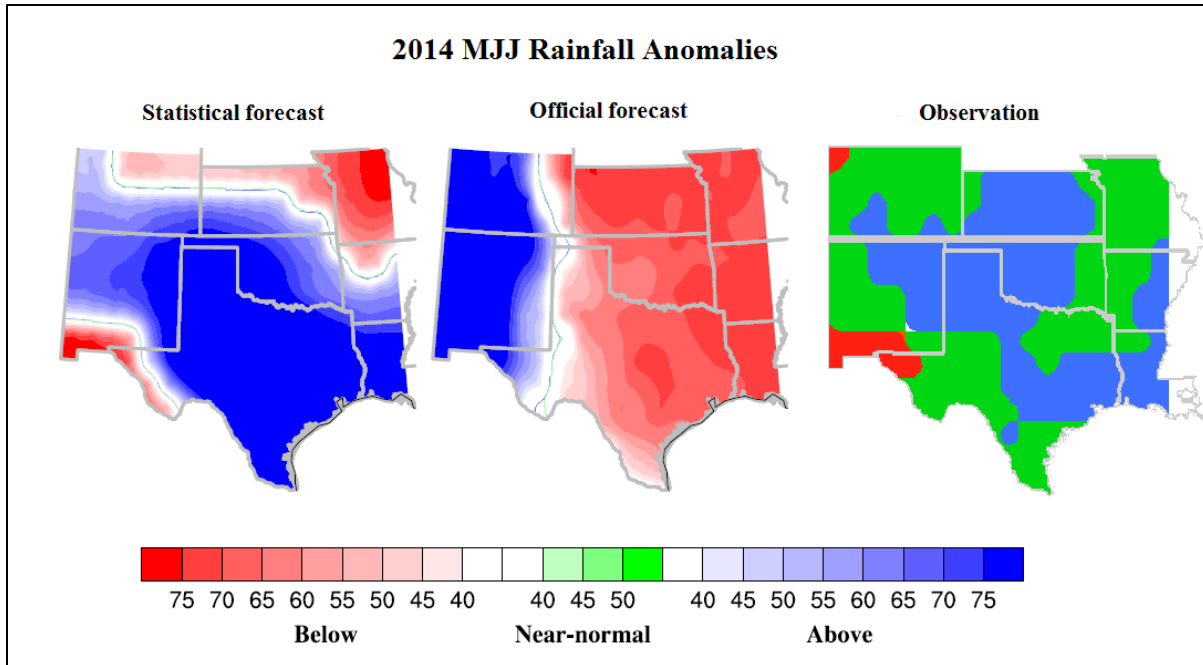
The official precipitation forecast for MJJ issued in April by the National Oceanographic and Atmospheric Administration (NOAA) using the NMME showed higher probabilities of below-normal rainfall over central and eastern Texas, southeastern Oklahoma and Louisiana, and higher probabilities of above-normal precipitation occurring over New Mexico and Colorado (Figure 12, center). Over west Texas and the Panhandle region, the forecast shows higher probabilities of above normal precipitation occurring over much of Texas and below normal precipitation over the western region in New Mexico and Colorado (Figure 12, middle). The observed total precipitation anomaly for MJJ shows northern and southeastern Texas as having received above normal rainfall, with western and southern Texas and southern New Mexico as having received below normal rainfall (Fig. 14, right). Statistics for the southern region from the National Climate Data Center (NCDC) shows that the MJJ season of 2014 had a 1.35 inch positive rainfall anomaly and the season ranked as the “94<sup>th</sup> wettest MJJ” ([http://www.ncdc.noaa.gov/cag/time-series/us/106/00/pcp/3/07/1895-2014?base\\_prd=true&firstbaseyear=1980&lastbaseyear=2010](http://www.ncdc.noaa.gov/cag/time-series/us/106/00/pcp/3/07/1895-2014?base_prd=true&firstbaseyear=1980&lastbaseyear=2010)). In synthesis, the indicator performed better than the official NOAA forecast in getting the nature of the rainfall anomaly correct. This was particularly the case over Texas.



**Figure 10:** Forecast skill depicted using Spearman's correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and 2-Alternate Forced Choice (forecast categories) (bottom), using April predictor fields from the Climate Forecast Reanalysis (left) and from the Modern-Era Retrospective Analysis for Research and Applications (right).



**Figure 11:** Forecast skill depicted Spearman's correlation (top), Relative Operating Characteristics Area (below-normal)(middle), and 2-Alternate Forced Choice (forecast categories) (bottom), where the soil moisture field is varied using Climate Forecast System Reanalysis volumetric soil moisture (left) and Essential Climate Variable-Soil Moisture data (right).



**Figure 12:** Statistical prediction (left), official forecast using the North American Multi-model Ensemble, and observed precipitation anomalies for the 2014 May through July season.

#### 4. Conclusions and recommendations for further study

Summer droughts (depicted using either the six-monthly Standardized Precipitation Index for July or May through July rainfall) over the southern Great Plains region can be predicted with skill levels acceptable to decision makers (~60 percent or higher) using large-scale circulation and land surface moisture fields in April. The best skill is achieved when using April initial conditions of the three predictor variables – i.e. 500 hectopascals geopotential height, difference in temperature between 700 hectopascals and dewpoint at the surface, and soil moisture – to predict July SPI6.

The early warning indicator is able to capture the spatial pattern and magnitude of past drought events and non-drought events well. It exceeds baseline predictability over most of Texas and Oklahoma. Categorical seasonal forecasts (that is, probabilistic estimates of whether a season will be below-, near- or above-normal) from the indicator provide added information on drought susceptibility from the six-month lead time onwards over Texas. In all model realizations, we find that the grid points

with the highest skill scores lie within Texas. Of interest is the ability of the indicator to forecast the 2011 drought event even at six-month (January) lead time.

Higher skill scores were obtained when the statistical forecast was initialized using April fields from the CFSR compared to fields from MERRA. Reasons for such a difference are unclear and warrant further investigation using possibly an ensemble mean of reanalysis fields from CFSR, MERRA and possibly ERA-Interim (Dee and others, 2011)<sup>3</sup>, NCEP 20<sup>th</sup> century reanalysis (Compo and others, 2011), and the North American Regional Reanalysis (Mesinger and others, 2006).

Modest improvement in skill were obtained, over the western and northern regions of the study domain, when we varied the soil moisture field by replacing soil moisture from CFSR with observed soil moisture from ECV soil moisture data set.

The indicator forecast for summer 2014, initialized in April 2014, showed abnormally wet conditions and matches observed conditions better than the forecast from NOAA's Climate Prediction Center.

Given the performance of the drought early warning indicator over Texas, will explore the possibility for providing real-time summer drought forecasts from January onwards to the Drought Preparedness Council, state emergency management initiatives and water planners. Such forecasts would ideally be made available to the public through the TWDB's drought web page and the Water Data for Texas web site.

---

<sup>3</sup> ERA-Interim is the reanalysis of the global atmosphere produced by the European Center for Medium-Range Weather Forecasts (ECMWF) from 1979-present.

## Acknowledgments

Drought early warning research at the University of Texas at Austin is funded by the NASA Indicators for the National Climate Assessment Program (Grant NNX13AN39G), NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections Program (Grant Award NA10OAR4310157) and the Jackson School of Geosciences. Drought research at the Texas Water Development Board has been funded by the U.S. Army Corps of Engineers Texas Water Allocation Assistance Program, and supported by the NASA Grant NNX13AN39G (awarded to The University of Texas at Austin). The research was also supported by the Postdocs Applying Climate Expertise Postdoctoral Fellowship Program, which was partially funded by NOAA's Climate Program Office and administered by the University Corporation for Atmospheric Research Visiting Scientist Programs.



## References

- Barnston, A. G., and Livezey, R.E., 1987, Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns, *Monthly Weather Review*, v. 115, n. 6, p. 1083-1126.
- Barnston, A. G., and van den Dool, H.M., 1993, A degeneracy in cross-validated skill in regression-based forecasts. *Journal of Climate*, v. 6, no. 5, p. 963-977.
- Barnston, A. G., and Ropelewski, C.F., 1992, Prediction of ENSO Episodes Using Canonical Correlation Analysis. *Journal of Climate*, v. 5, n. 11, p. 1316-1345.
- Chen, M., Xie, P., Janowiak, J.E., and Arkin, P.A., 2002, Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations, *Journal of Hydrometeorology*, v. 3, n. 3, p. 249-266.
- Compo, G. P., and others., 2011, The Twentieth Century Reanalysis Project, *Quarterly Journal of the Royal Meteorological Society*, v. 137, n. 654, p. 1-28.
- Dee, D. P., and others, 2011, The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, v. 137, no. 656, p. 553-597.
- Eichler, T., and Higgins, W., 2006, Climatology and ENSO-related variability of North American extratropical cyclone activity, *Journal of Climate*, v. 19, no.10, p. 2076-2093.
- Fernando, D.N., 2014, Assessing the precursors, persistence and predictability of drought over Texas. Texas Water Development Board Contracted Report. 73 p. ([http://www.twdb.texas.gov/publications/reports/contracted\\_reports/doc/1200011490\\_drought.pdf](http://www.twdb.texas.gov/publications/reports/contracted_reports/doc/1200011490_drought.pdf)).
- Fernando, D.N., Mo, K.C. Mo, Fu, R., Bowerman, A., Scanlon, B.R., Solis, R.S., Yin, L., Mioduszewski, J.R., Ren, T., Zhang, K., Mace, R.E. (in-prep for submission to *Climate Dynamics*). What caused the spring intensification and winter demise of the 2011 drought over Texas?

- Hoerling, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., Schubert, S., and Seager, R., 2013, Causes and Predictability of the 2012 Great Plains Drought, *Bulletin of the American Meteorological Society*, v. 95, no. 2, p. 269-282.
- Hong, S.-Y., and Kalnay, E., 2002, The 1998 Oklahoma–Texas Drought: Mechanistic Experiments with NCEP Global and Regional Models, *Journal of Climate*, v. 15, no. 9, p. 945-963.
- Janowiak, J. E., and P. Xie, 1999, CAMS–OPI: A Global Satellite–Rain Gauge Merged Product for Real-Time Precipitation Monitoring Applications, *Journal of Climate*, v. 12, no. 11, p. 3335-3342.
- Kam, J., Sheffield, J., Yuan, X. and Wood, E.F., 2014, Did a skillful prediction of sea surface temperatures help or hinder forecasting of the 2012 Midwestern US drought?, *Environmental Research Letters*, v. 9, no. 3, 034005.
- Kirtman, B. P., and others, 2013, The North American Multi-model Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction, *Bulletin of the American Meteorological Society*, v. 95, no. 4, p. 585-601.
- Koster, R. D., and others, 2004, Regions of strong coupling between soil moisture and precipitation. *Science*, v. 305, no. 5687, p. 1138-1140.
- Kousky, V. E., and Ropelewski, C.F., 1989, Extremes in the Southern Oscillation and their relationship to precipitation anomalies with emphasis on the South American region, *Revista Brasileira de Meteorologia*, v. 4, no. 2, p. 351-363.
- Kumar, A., Chen, M., Hoerling, M., and Eischeid, J., 2013, Do extreme climate events require extreme forcings? *Geophysical Research Letters*, v. 40, no. 13, p. 3440-3445.
- Landman, W. A., and Mason, S.J., 2001, Forecasts of near-global sea surface temperatures using Canonical Correlation Analysis. *Journal of Climate* v. 14, no. 18, p. 3819-3833.

- Liu, Y. Y. and others, 2011, Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrology and Earth System Sciences*, v. 15, no. 2, p. 425-436.
- Liu, Y. Y., and others, 2012, Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sensing of Environment*, v. 123, p. 280-297.
- Lyon, B., and Dole, R.M., 1995, A Diagnostic Comparison of the 1980 and 1988 U.S. Summer Heat Wave-Droughts, *Journal of Climate*, v. 8, no. 6, p. 1658-1675.
- Mason, S. J., and Graham, N.E., 2002, Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation, *Quarterly Journal of the Royal Meteorological Society*, v. 128, no. 584, p. 2145-2166.
- Mason, S. J., and Weigel, A. P., 2009, A Generic Forecast Verification Framework for Administrative Purposes, *Monthly Weather Review*, v. 137, no. 1, p. 331-349.
- Mesinger, F., and others, 2006, North American Regional Reanalysis, *Bulletin of the American Meteorological Society*, v. 87, no. 3, p. 343-360.
- Michaelsen, J., 1987, Cross-validation in statistical climate forecast models, *Journal of Climate and Applied Meteorology*, v.26, no. 11, p. 1589-1600.
- Mo, K. C., 2011, Drought onset and recovery over the United States, *Journal of Geophysical Research: Atmospheres*, v. 116, no. D20.
- Mueller, B., and Seneviratne, S.I. 2012, Hot days induced by precipitation deficits at the global scale, *Proceedings of the National Academy of Sciences*, v. 109, no. 31, p. 12398-12403.
- Myoung, B., and J. W. Nielsen-Gammon, 2010, The Convective Instability Pathway to Warm Season Drought in Texas. Part I: The Role of Convective Inhibition and Its Modulation by Soil Moisture, *Journal of Climate*, v. 23, no. 17, p. 4461-4473.

Namias, J., 1982, Anatomy of Great Plains Protracted Heat Waves (especially the 1980 U.S. summer drought), *Monthly Weather Review*, v. 110, p. 824-838.

Quan, X.-W., Hoerling, M. P., Lyon, B., Kumar, A., Bell, M.A., Tippett, M.K., and Wang, H., 2012, Prospects for Dynamical Prediction of Meteorological Drought, *Journal of Applied Meteorology and Climatology*, v. 51, no. 7, p. 1238-1252.

Richman, M. B., 1986, Rotation of principal components, *Journal of Climatology*, v. 6, no. 3, p. 293-335.

Rienecker, M. M., and others, 2011, MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *Journal of Climate*, v. 24, no. 14, p. 3624-3648.

Ropelewski, C. F., and Halpert, M. S., 1986, North American Precipitation and Temperature Patterns Associated with the El Niño/Southern Oscillation (ENSO), *Monthly Weather Review*, v. 114, no. 12, p. 2352-2362.

Ropelewski, C. F., and Halpert, M.S., 1987, Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation, *Monthly Weather Review*, v. 115, no. 8, p. 1606-1626.

Saha, S. and others, 2010, The NCEP climate forecast system reanalysis, *Bulletin of the American Meteorological Society*, v. 91, no. 8, p. 1015-1057.

Saha, S., and others, 2014, The NCEP Climate Forecast System Version 2, *Journal of Climate*, v. 27, no. 6, p. 2185-2208.

Shabbar, A., and Barnston, A.G., 1996, Skill of seasonal climate forecasts in Canada using Canonical Correlation Analysis. *Monthly Weather Review*, v. 124, no. 10, p. 2370-2385.

Schubert, S., and others, 2009, A U.S. CLIVAR Project to Assess and Compare the Responses of Global Climate Models to Drought-Related SST Forcing Patterns: Overview and Results, *Journal of Climate*, v. 22, no. 19, p. 5251-5272.

Seager, R., Goddard, L., Nakamura, J., Henderson, N., and Lee, D.E., 2013, Dynamical Causes of the 2010/11 Texas–Northern Mexico Drought, *Journal of Hydrometeorology*, v. 15, no. 1, p. 39-68.

von Storch, H., and Zwiers, F.W., 2002. Statistical analysis in climate research, Cambridge University Press. 484 p.

Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., and Ertl, M., 2012, Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v., 1-7, p. 315-321.

Wilks, D. S., 2006, Statistical Methods in the Atmospheric Sciences (2nd ed.), Academic Press, New York, N.Y., U.S.A., 627 p.

## Appendix 1: Table with variance explained by the EOF modes

Input data field	EOF1	EOF2	EOF3
CFS Reanalysis predictor fields and observed July SPI6			
Apr	53.6%	33.9%	12.5%
Mar	83.66%	15.11%	1.23%
MAM	84.21%	15.22%	0.57%
CFSR predictor fields in April and observed rainfall/SPI6			
MJJ			
JJA rainfall			
July SPI6	53.6%	33.9%	12.5%
MERRA predictor fields in April and observed July SPI6			
	47.34%	30.55%	22.11%
CFSR hindcast fields for April and observed rainfall and observed SPI6			
three-month	42.67%	31.77%	25.55%
two-month	39.04%	31.63%	29.34%
one-month	39.84%	34.33%	25.83%
CFSR predictor circulation fields and soil moisture from ECV CCI			
	53.64%	41.45%	4.91%
CFSR predictor circulation fields and soil moisture/soil moisture proxy combined from ECV CCI and SIF			
	41.56%	33%	25.43%

**Table 1:** Percentage variance explained by the three EOF modes for each variation in input data

## Appendix 2: Canonical Correlation Analysis input window in the Climate Predictability Tool

Climate Predictability Tool 14 - CCA\_apr (CCA)

File Edit Actions Tools Options View Help

Explanatory (X) Variables:		Response (Y) Variables:		Forecast Variables:	
File name:	reof1_2apr.txt	File name:	cpt_SPI6.tsv	File name:	reof1_2apr.txt
First data:	FMAMJJ 1979	First data:	FMAMJJ 1979	First data:	FMAMJJ 1979
Last data:	FMAMJJ 2014	Last data:	FMAMJJ 2013	Last data:	FMAMJJ 2014
Start at:	1982	Start at:	1982	Start at:	2006
Number of fields	2	Number of fields	1	Number of fields	2
Number of lags	1	Number of lags	1	Number of lags	1
Number of gridpoints	646	Number of gridpoints	323	Number of gridpoints	646
Number used	646	Number used	323	Number used	646

Length of training period: 24      Length of cross-validation window: 3      Number of forecasts: 9

Actions:

### Appendix 3: Sample Rotated EOF input data for Canonical Correlation Analysis using the Climate Predictability Tool

xmlns:cpt=http://iri.columbia.edu/CPT/v10/																			
cpt:nfields=2																			
cpt:T	2/7/1981	2/7/1982	2/7/1983	2/7/1984	2/7/1985	2/7/1986	2/7/1987	2/7/1988	2/7/1989	2/7/1990	2/7/1991	2/7/1992	2/7/1993	2/7/1994	2/7/1995	2/7/1996	2/7/1997		
	2/7/1998	2/7/1999	2/7/2000	2/7/2001	2/7/2002	2/7/2003	2/7/2004	2/7/2005	2/7/2006	2/7/2007	2/7/2008	2/7/2009	2/7/2010	2/7/2011	2/7/2012	2/7/2013	2/7/2014		
cpt:field=R cpt:T=1982 cpt:nrow=1 cpt:ncol=1! cpt:row=Y, cpt:col=X, cpt:units=U cpt:missing=-999																			
	-110	-109	-108	-107	-106	-105	-104	-103	-102	-101	-100	-99	-98	-97	-96	-95	-94	-93	-92
40	1.69E-02	1.85E-02	1.88E-02	1.02E-02	5.22E-03	1.54E-03	1.31E-04	-5.18E-04	1.95E-04	2.25E-03	7.11E-03	1.14E-02	1.59E-02	2.21E-02	2.27E-02	2.30E-02	2.31E-02	2.31E-02	2.31E-02
39	1.70E-02	1.78E-02	9.35E-03	4.16E-03	2.00E-04	1.76E-03	-5.74E-04	1.60E-03	5.35E-03	6.47E-03	7.75E-03	9.23E-03	1.56E-02	2.15E-02	2.21E-02	2.24E-02	2.25E-02	2.25E-02	2.26E-02
38	1.63E-02	9.23E-03	3.69E-03	2.45E-04	-1.98E-03	1.59E-03	3.40E-05	3.02E-03	7.34E-03	7.41E-03	6.82E-03	8.72E-03	1.27E-02	1.82E-02	2.11E-02	2.15E-02	2.17E-02	2.18E-02	2.18E-02
37	6.48E-03	3.23E-03	5.23E-04	-2.11E-03	-2.60E-03	-1.43E-03	4.66E-04	5.16E-03	6.35E-03	5.69E-03	7.16E-03	8.43E-03	1.40E-02	1.94E-02	2.01E-02	2.05E-02	2.07E-02	2.08E-02	2.09E-02
36	1.74E-03	1.25E-04	-2.53E-03	-3.64E-03	-2.71E-03	-1.59E-03	2.15E-03	3.60E-03	3.63E-03	4.97E-03	5.98E-03	6.72E-03	1.17E-02	1.69E-02	1.77E-02	1.89E-02	1.96E-02	1.97E-02	1.97E-02
35	-1.46E-03	-2.80E-03	-4.57E-03	-4.60E-03	-3.54E-03	-1.72E-03	-1.16E-03	6.69E-04	2.71E-03	3.64E-03	4.57E-03	5.38E-03	6.35E-03	6.91E-03	7.36E-03	1.32E-02	1.76E-02	1.65E-02	1.66E-02
34	-3.78E-03	-4.43E-03	-5.20E-03	-4.10E-03	-3.45E-03	-1.64E-04	-1.75E-03	3.65E-05	1.80E-03	2.50E-03	2.83E-03	3.43E-03	4.34E-03	4.58E-03	4.60E-03	4.86E-03	1.10E-02	6.29E-03	6.33E-03
33	-5.21E-03	-5.34E-03	-4.91E-03	-2.29E-03	-2.66E-03	-2.13E-03	-2.89E-03	-1.08E-03	4.58E-04	1.05E-03	1.63E-03	2.22E-03	2.65E-03	2.45E-03	1.93E-03	1.61E-03	1.57E-03	1.82E-03	1.85E-03
32	-6.31E-03	-6.53E-03	-6.16E-03	-4.49E-03	-5.81E-03	-5.51E-03	-3.27E-03	-1.68E-03	-1.11E-03	-3.68E-04	9.10E-04	1.42E-03	1.16E-03	8.36E-04	6.07E-04	1.82E-04	3.64E-05	4.89E-04	8.73E-04
31	-7.70E-03	-8.43E-03	-8.71E-03	-8.14E-03	-7.73E-03	-7.01E-03	-2.37E-03	-1.48E-03	-2.37E-03	-1.38E-03	2.26E-05	6.18E-04	1.06E-04	-6.59E-04	-8.28E-04	-1.01E-03	-1.17E-03	-8.15E-04	-5.33E-04
30	-9.43E-03	-1.08E-02	-1.11E-02	-1.02E-02	-9.41E-03	-8.62E-03	-5.68E-03	-4.15E-03	-3.72E-03	-2.62E-03	-1.27E-03	-5.70E-04	-1.01E-03	-2.04E-03	-2.77E-03	-3.03E-03	-2.98E-03	-2.56E-03	-2.21E-03
29	-1.10E-02	-1.20E-02	-1.21E-02	-1.14E-02	-1.05E-02	-9.85E-03	-7.00E-03	-5.47E-03	-5.36E-03	-4.35E-03	-3.02E-03	-2.07E-03	-1.93E-03	-2.45E-03	-3.39E-03	-4.20E-03	-4.47E-03	-4.33E-03	-3.94E-03
28	-1.29E-02	-1.30E-02	-1.34E-02	-1.43E-02	-1.39E-02	-1.32E-02	-8.62E-03	-7.21E-03	-6.90E-03	-5.43E-03	-3.03E-03	-1.65E-03	-1.60E-03	-2.74E-03	-3.85E-03	-4.89E-03	-5.56E-03	-5.54E-03	-5.45E-03
27	-1.40E-02	-1.45E-02	-1.60E-02	-1.67E-02	-1.43E-02	-1.43E-02	-1.12E-02	-9.93E-03	-8.51E-03	-6.90E-03	-5.13E-03	-4.24E-03	-4.23E-03	-4.67E-03	-5.17E-03	-5.76E-03	-6.40E-03	-6.86E-03	-6.90E-03
26	-1.55E-02	-1.63E-02	-1.74E-02	-1.86E-02	-1.63E-02	-1.48E-02	-1.26E-02	-1.13E-02	-1.00E-02	-8.71E-03	-7.61E-03	-7.14E-03	-7.17E-03	-7.00E-03	-6.94E-03	-7.22E-03	-7.65E-03	-8.18E-03	-8.21E-03
25	-1.61E-02	-1.67E-02	-1.80E-02	-1.99E-02	-1.89E-02	-1.59E-02	-1.38E-02	-1.22E-02	-1.13E-02	-1.05E-02	-9.55E-03	-9.05E-03	-8.78E-03	-8.59E-03	-8.64E-03	-8.85E-03	-9.10E-03	-9.01E-03	-9.39E-03
24	-1.76E-02	-1.73E-02	-1.82E-02	-2.02E-02	-2.01E-02	-1.69E-02	-1.30E-02	-1.20E-02	-1.26E-02	-1.22E-02	-1.20E-02	-1.14E-02	-1.08E-02	-1.07E-02	-1.10E-02	-1.09E-02	-1.08E-02	-1.10E-02	-1.09E-02



## **Errata**

This technical note was first posted online on 3 February 2015. Subsequently, we realized that an acknowledgement section had not been included. The acknowledgement section is now included on page 32 of the technical note.