SPECIAL
Advances in Drought Monitoring
COLLECTION

Prospects for Advancing Drought Understanding, Monitoring, and Prediction

ERIC F. WOOD,[*] SIEGFRIED D. SCHUBERT,[+] ANDREW W. WOOD,[#] CHRISTA D. PETERS-LIDARD,[@] KINGTSE C. MO,[&] ANNARITA MARIOTTI,[**] AND ROGER S. PULWARTY[++]

[*] *Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey*
[+] *Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[#] *Research Applications Laboratory, NCAR, Boulder, Colorado*
[@] *Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[&] *NOAA/NCEP/Climate Prediction Center, College Park, Maryland*
[**] *NOAA/OAR/Climate Program Office, Silver Spring, Maryland*
[++] *NOAA/OAR/Climate Program Office, Silver Spring, Maryland, and NOAA/ESRL/Physical Sciences Division, Boulder, Colorado*

(Manuscript received 29 August 2014, in final form 20 January 2015)

ABSTRACT

This paper summarizes and synthesizes the research carried out under the NOAA Drought Task Force (DTF) and submitted in this special collection. The DTF is organized and supported by NOAA's Climate Program Office with the National Integrated Drought Information System (NIDIS) and involves scientists from across NOAA, academia, and other agencies. The synthesis includes an assessment of successes and remaining challenges in monitoring and prediction capabilities, as well as a perspective of the current understanding of North American drought and key research gaps. Results from the DTF papers indicate that key successes for drought monitoring include the application of modern land surface hydrological models that can be used for objective drought analysis, including extended retrospective forcing datasets to support hydrologic reanalyses, and the expansion of near-real-time satellite-based monitoring and analyses, particularly those describing vegetation and evapotranspiration. In the area of drought prediction, successes highlighted in the papers include the development of the North American Multimodel Ensemble (NMME) suite of seasonal model forecasts, an established basis for the importance of La Niña in drought events over the southern Great Plains, and an appreciation of the role of internal atmospheric variability related to drought events. Despite such progress, there are still important limitations in our ability to predict various aspects of drought, including onset, duration, severity, and recovery. Critical challenges include (i) the development of objective, science-based integration approaches for merging multiple information sources; (ii) long, consistent hydrometeorological records to better characterize drought; and (iii) extending skillful precipitation forecasts beyond a 1-month lead time.

1. Introduction

The National Integrated Drought Information System's (NIDIS) Implementation Plan states that "[d]rought is among the most damaging and least understood of all natural hazards" (NISDIS 2007, p. ii; www.drought.gov/drought/). It is well understood among climatologists that drought is a naturally occurring phenomenon. Its slow onset; often long-duration, cumulative impacts; and widespread extent results in droughts being the costliest of natural disasters (NCDC 2012; Below et al. 2007) with profound economic, social, and environmental impacts. The National Climatic Data Center (NCDC) has recorded droughts in the United States having severe economic impacts (more than $1 billion in damages) during 16 of the 21 years from 1980 to 2011, with an estimated annual average direct drought loss of $9.5 billion (adjusted to 2011 dollars; Smith and Katz 2013). It should be noted that these estimates do not take into account secondary impacts affecting water resources, recreation economies, energy, and ecosystems.

Recognizing the economic and social impacts from drought, the U.S. Congress in 2006 passed the National

Corresponding author address: E. F. Wood, Civil and Environmental Engineering, Princeton University, Olden Street, EQUAD E415, Princeton, NJ 08544.
E-mail: efwood@princeton.edu

DOI: 10.1175/JHM-D-14-0164.1

Integrated Drought Information System Act of 2006 (Public Law 109-430) with NOAA as the lead agency. The subsequent NIDIS Implementation Plan was developed to ''[f]oster, and support, a research environment that focuses on risk assessment, forecasting, and management,'' among other goals (NIDIS 2007, p. iii). Given its widespread support, NIDIS was reauthorized and signed into law by President Obama in March 2014 (U.S. Government 2014). Major foci of the reauthorized NIDIS Act include the identification of research, monitoring, and forecasting needs to enhance the predictive capability of drought early warnings on ''(i) the length and severity of droughts; (ii) the contribution of weather events to reducing the severity or ending drought conditions; and (iii) regionally specific drought impacts'' (U.S. Government 2014).

NIDIS has partnered with NOAA's Climate Program Office, which set up a Drought Task Force (DTF) with the overarching goal of advancing drought understanding, monitoring, and prediction through coordinated research activities that address a number of NIDIS-relevant scientific objectives. These include (i) the scientific understanding of the weather and climatic mechanisms that lead to the onset, maintenance, and recovery of drought; (ii) improving drought prediction skill by identifying and exploiting sources of drought predictability and related aspects such as the dependence on time scales, regions, seasons, and variables, and improvements in forecast models and procedures; (iii) improving current drought monitoring capabilities, including the exploitation of new data, methodologies, and metrics that would improve society's capability to manage drought; and (iv) improving drought information systems through incorporating the latest advances in monitoring and prediction, objective metrics relevant to various societal sectors, and advanced information delivery platforms. The DTF, as part of the Modeling, Analysis, Prediction, and Projections (MAPP) program, involves scientists from academia, other agencies, and across NOAA (cpo.noaa.gov/MAPP/DTF). The DTF both leverages and contributes to drought research across the federal government as part of the U.S. Global Change Research Program and international research programs under the World Climate Research Programme. Initiatives such as the Drought Task Force will be key in advancing current national drought capabilities toward the development of the Global Drought Information System (Pozzi et al. 2013).

At the outset, the DTF adopted a test bed framework centered around three working groups (WGs) related to Metrics, Case Studies, and Drought Early Warning Systems (DEWS). The focus of the Metrics WG was to identify and apply metrics for evaluating scientific and technological advances in monitoring and prediction. These include metrics that enable a systematic and comprehensive evaluation of the quality of existing drought monitoring and prediction services, their performance in a number of drought case studies, and their potential to support national and global DEWS. The Case Studies WG focused on identifying and analyzing several high-profile case studies that appear to have different drought mechanisms, feedbacks, and potential predictability. These cases consist of the southeastern U.S. drought during 2006/07, the Texas drought of 2011, the central Great Plains drought of 2012, and the western U.S. drought from 1998 to 2002. The DEWS WG focus is on supporting the continued development and evaluation of drought monitoring and prediction tools, such as the North American Land Data Assimilation System (NLDAS) and the North American Multimodel Ensemble (NMME) system. Note that DEWS is used here to indicate a research focus on monitoring and prediction. Comprehensive early warning information systems, such as those intended as part of NIDIS, also include risk assessments, communication, and engaging planning and preparedness components in communities and places at risk (Pulwarty and Verdin 2013).

Building off the foundational work of its WGs, the DTF has performed drought case study analyses and developed explanatory narratives that together have provided insights into the physical mechanisms underpinning drought, our ability to represent and predict them using models and observational datasets, and the extent to which new research has shown noteworthy potential to advance operational drought capabilities. Focusing on the Research to Capability (RtC) pathway, the DTF has developed a Drought Assessment Protocol as the framework to evaluate new experimental drought monitoring and prediction tools using common evaluation standards with the objective of informing strategies for improving operational drought services (http://cpo.noaa.gov/sites/cpo/Reports/MAPP/drought/DTF_Assessment_Protocol.pdf). The protocol is described in detail in the appendix.

To synthesize our current understanding of and capabilities for drought monitoring and prediction, the DTF has convened a special collection of journal articles published in the *Journal of Hydrometeorology* (http://journals.ametsoc.org/page/droughtMonitoring) on the drought-related research undertaken by its working groups. This paper provides a synthesis of those contributions and specifically tries to summarize and assess the research being carried out that addresses the objectives of the monitoring and prediction WGs. We set the stage for the synthesis by first presenting in section 2 some results from the 2012 drought illustrating current

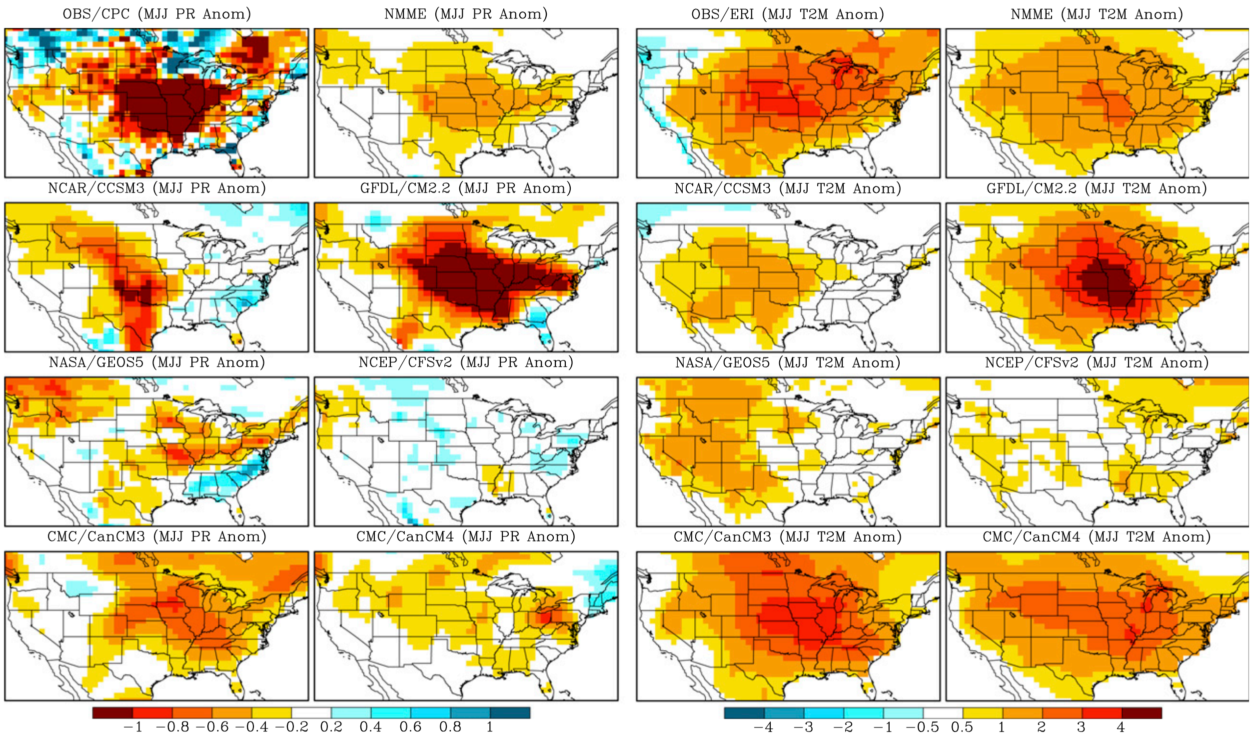

FIG. 1. (left) May–July precipitation PR (mm day^{-1}) and (right) 2-m air temperature T2M (°C) anomaly forecasts from six NMME models (ensemble average). Top row is observed anomaly and NMME average forecast using the individual model averages. Precipitation observations from the Climate Prediction Center are given as OBS/CPC; temperature observations based on ERA-Interim are given as OBS/ERI.

capabilities and the challenges faced by the DTF in improving our ability to monitor and predict such extreme events. Section 3 then summarizes current operational drought monitoring practices and advances from ongoing research. Section 4 describes advances in drought prediction and the understanding of drought predictability. Section 5 summarizes the current status and research advances and discusses remaining challenges, research opportunities, and prospects for improving drought understanding, monitoring, and prediction.

2. An exemplary case: 2012 drought

It has been estimated by Joseph LaVorgna, the chief U.S. economist at Deutsche Bank Securities Inc. (see www.bloomberg.com/news/2012-11-12/u-s-drought-may-cut-gdp-by-one-percentage-point-deutsche-says.html), that the 2012 U.S. drought may have caused damage estimated to range between \$75 billion and \$150 billion, with losses of over \$30 billion in agriculture alone, and reducing the U.S. gross domestic product by ~1%. The U.S. Drought Monitor showed that over 60% of the continental United States (CONUS) was under drought conditions in 2012, with over 40% of the regions under severe to exceptional drought conditions. Through monthly teleconferences, the DTF drew together and scrutinized the latest research community assessments of the drought, some of which are summarized briefly below.

As an overall assessment, it can be noted that the nation's current drought monitoring and prediction enterprise did not accurately predict the hydroclimatic severity and rapid development of the 2012 drought. In particular, the forecast skill of the latest dynamical seasonal climate forecast models was limited beyond one month lead time, and the forecasts across models varied significantly. Figure 1 shows, for example, the 3-month precipitation and surface air temperature forecasts initialized in early May 2012 from the NMME, including NOAA's operational Climate Forecast System, version 2 (CFSv2; Saha et al. 2010). Most NMME models predicted a precipitation and temperature anomaly in the central United States, but with marked intermodel variability in their extent, location, and intensity. For some other regions (e.g., the wet and cool Pacific Northwest), the observed anomalies were absent in virtually all model predictions.

The ensemble-averaged forecast from NOAA's operational seasonal climate forecast model CFSv2 was close to neutral, yet a closer inspection of the ensemble set from the last 10 days of April 2012 shows that the four most skillful and four least skillful ensemble members had

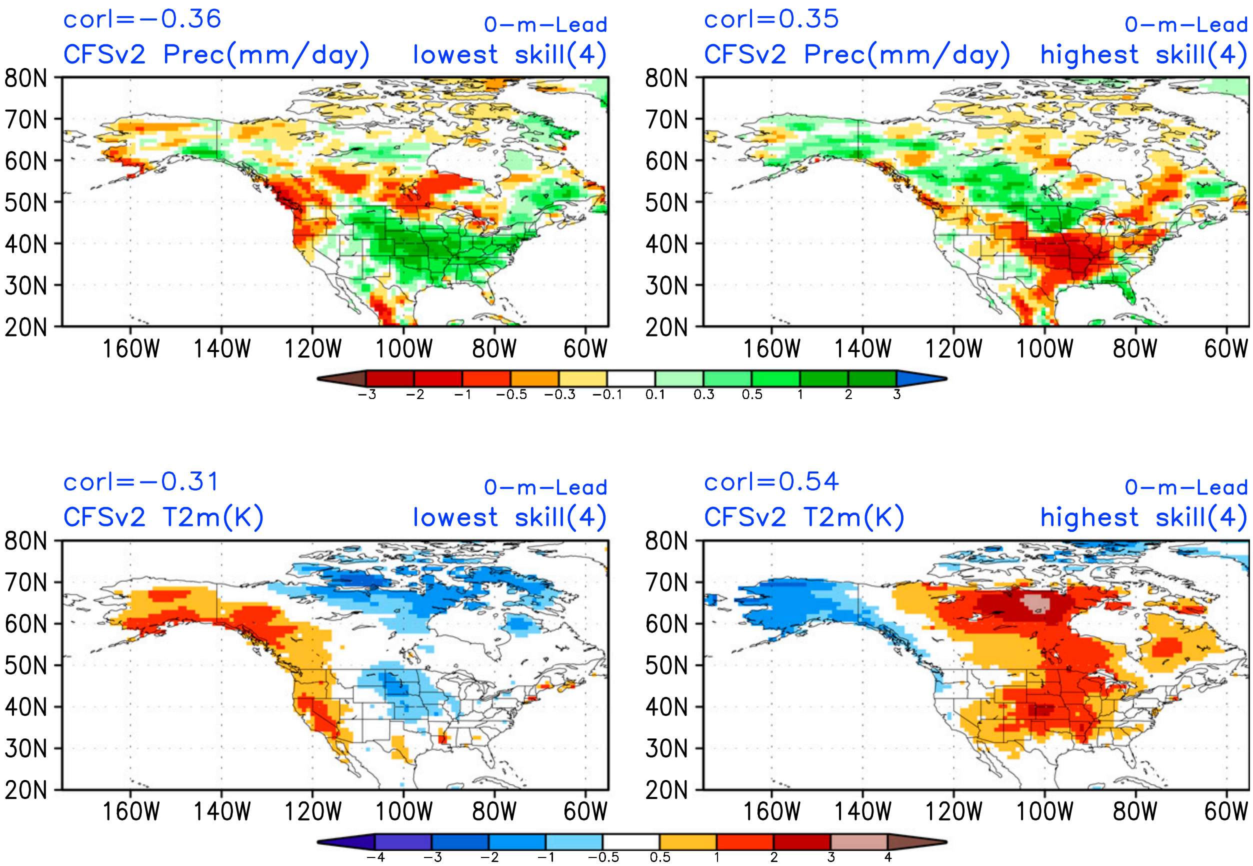

FIG. 2. (top) PR (mm day^{-1}) and (bottom) T2M (°C) anomaly forecasts for May 2012 based on the four (left) lowest and (right) highest skill ensembles based on the 40 CFSv2 ensembles during 20–30 Apr 2012. (Figures courtesy of A. Kumar and M. Chen, NCEP/CPC; note that the color bars are different than in Fig. 1).

dramatically different forecasts—almost mirror images (see Fig. 2). Given that the SST patterns for these most and least skillful members were almost identical (A. Kumar 2012, personal communication), this suggests that the level of atmospheric noise (at least in this model) was significant and contributed to the wide spread among its 40 ensemble members. In fact, all the NMME models had very similar SST patterns, as can be seen in Fig. 3, which points to the limited role of SST in the occurrence of this drought and hence its limited predictability using traditional seasonal forecasting methods.

The 2012 drought presented a sobering snapshot of the challenges faced by the drought management community in leveraging science and technology to better anticipate and respond to drought. In turn, the drought also highlighted the remaining challenges for the science community to improve our understanding of the fundamental predictability of droughts and the tools we use to predict them. Specific research issues include identifying relevant sources of drought predictability and determining if models are exploiting such sources to the greatest extent possible, and whether drought-related model processes (and related prediction skill) are consistent with observations. To further illustrate this, Fig. 4 presents the correlation between the first principal component (PC) of averaged June–August (JJA) SST, labeled S1(SST), and SST itself (Fig. 4, left) and the average correlation between S1(SST) and JJA precipitation anomalies over CONUS from the ensembles (Fig. 4, right) using observational data and NMME hindcasts from 1982 to 2012. The numbers of ensembles used are 10, 9, and 24 for GFDL CM2.1, NASA GMAO, and NCEP CFSv2, respectively. Much of the variability in tropical Pacific SST (Niño-3 region, western Pacific, and tropical Atlantic) is captured by S1(SST), regions believed to have teleconnections to precipitation over CONUS. Also shown are the S1(SST)–SST patterns for three NMME models, demonstrating very similar correlation patterns to the observed SSTs and among themselves. The correlation pattern between S1(SST) and observed CONUS precipitation is shown in Fig. 4 (top right). The average correlation patterns between S1(SST) and model-predicted precipitation are significantly different among the models themselves, with

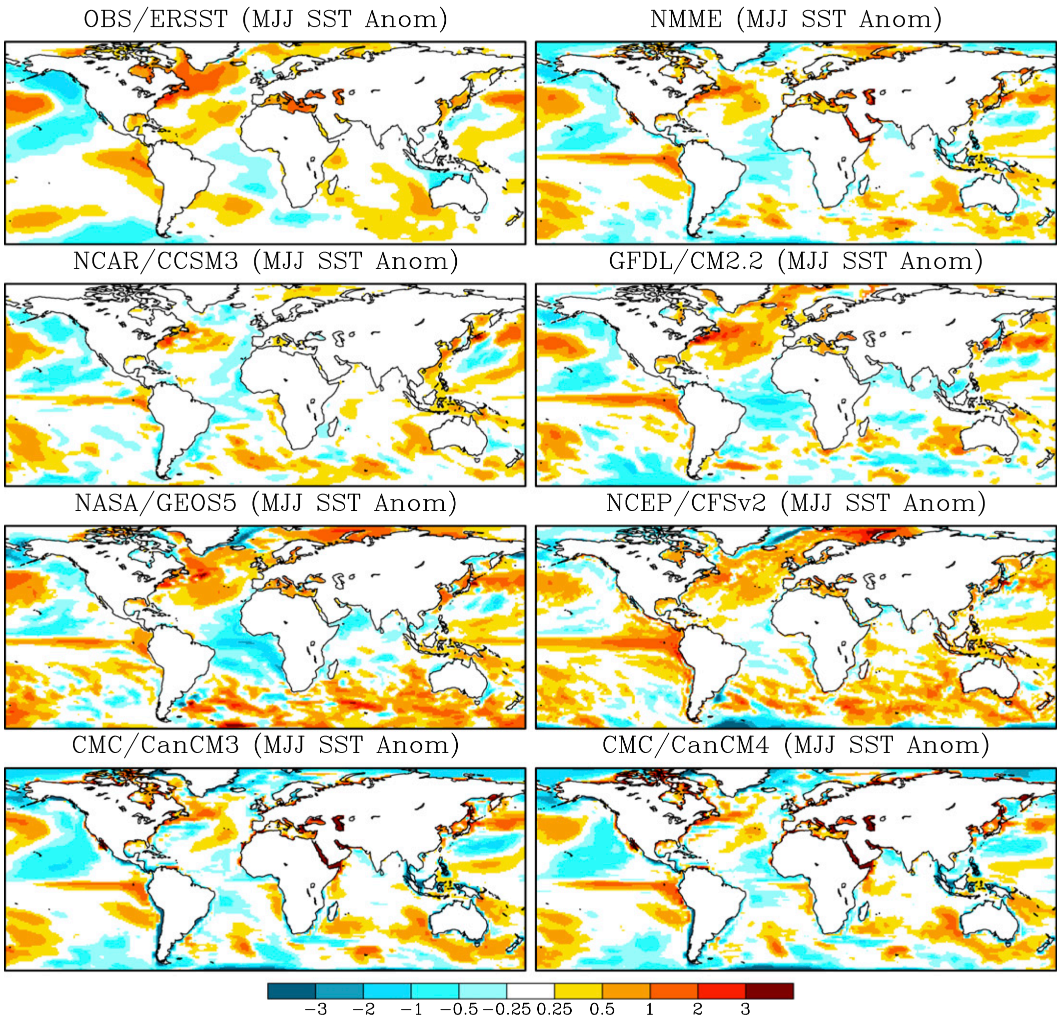

FIG. 3. As in Fig. 1, but for SST (°C).

NASA GMAO having the highest similarity with the observed patterns and NCEP CFSv2 having the lowest. The different number of ensemble members for each model makes it harder to generalize across models, suggesting a need for greater coordination and consistency in designing multimodel experiments. Nonetheless, these results imply that while all the models reasonably capture SST variability, the intermodel-predicted precipitation variability is large, which shows that our understanding of the sources of predictability is incomplete, especially against the background of intrinsic atmospheric variability, and that our understanding of which model parameterizations (and features such as resolution and reductions in overall model structural uncertainty including initial and boundary conditions) are most likely to lead to more skillful models is limited. This highlights the challenges in developing multimodel ensemble systems.

These results underscore those from a number of recent studies (e.g., Seager et al. 2014; Hoerling et al. 2014; Wang et al. 2014; see section 4 for a discussion of these studies), demonstrating the challenges that remain in understanding the mechanisms that lead to drought and improving the predictions of drought onset, duration, severity, and recovery.

3. Drought monitoring

Assessing research to improve drought monitoring is particularly challenging because drought represents a combination of numerous geophysical phenomena, and the monitoring and management of drought involves a synthesis of disparate information sources whose characteristics vary. This variability in information includes time, space, quality, format, availability, and in relative importance depending on the objectives and

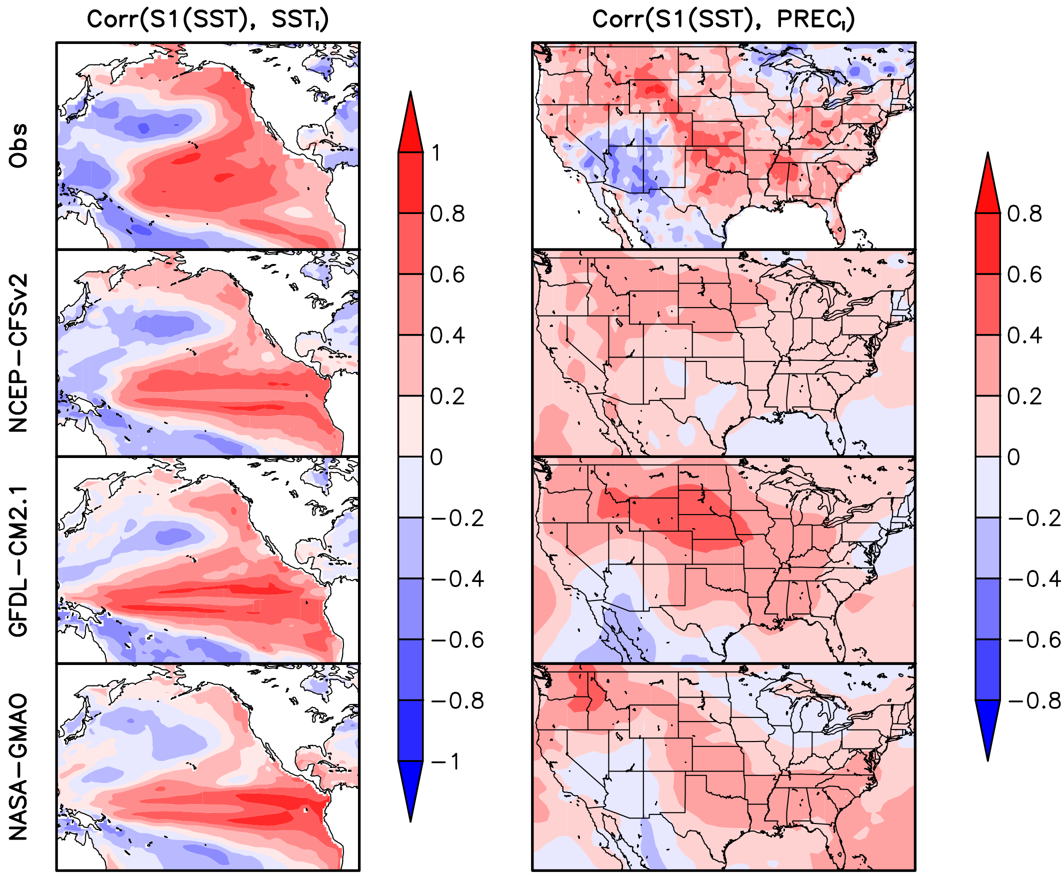

FIG. 4. Correlation between (left) first PC of SST [labeled S1(SST)] and SST and (right) S1(SST) and ensemble mean precipitation for JJA 1982–2012. Top row contains observations. Lower rows contain NMME models. See text for the number of ensembles in each model. (Courtesy of J. Kam, Princeton University.)

perspectives of users in different drought-affected sectors. DTF members address the monitoring challenge through research to improve objective analyses that contribute to a consistent, accurate, and reliable determination and quantification of drought—including, for example, assessments of soil moisture, river discharge, temperature anomalies, and depiction of vegetation health. Additionally, by carrying out objective historical reanalyses, a climatology for drought variables such as precipitation or soil moisture can be developed that aids in depicting current conditions within a risk framework. Finally, objective analyses can assist in improving the process of integrating multiple indicators of drought. This section first discusses current operational monitoring capabilities and then provides an overview of the DTF special collection articles related to monitoring-relevant research supported by the NOAA MAPP program.

a. *Operational U.S. drought monitoring*

Improving operational drought monitoring in the United States offers a major opportunity for DTF drought monitoring research. At the national level, operational drought monitoring is led by four primary groups: the National Drought Mitigation Center (NDMC) at the University of Nebraska–Lincoln, the U.S. Department of Agriculture (USDA), the NCEP Climate Prediction Center (CPC), and the NCDC. Climatologists from these groups alternate to produce the nation's drought monitoring information product, the U.S. Drought Monitor (USDM; http://droughtmonitor.unl.edu and drought.gov) map of current drought conditions. Established in 1999, the weekly map uses a ranking/percentile system to facilitate the integration of numerous input analyses and indices having unique periods of record and units of measurements. The resulting classification scheme includes one abnormally dry category (labeled D0) and four drought categories (D1, moderate; D2, severe; D3, extreme; and D4, exceptional) that reflect dry conditions below the 30th, 20th, 10th, 5th, and 2nd percentiles, respectively. A rotating lead author uses his/her best judgment to reconcile differences in input analyses from a broad range of sources in

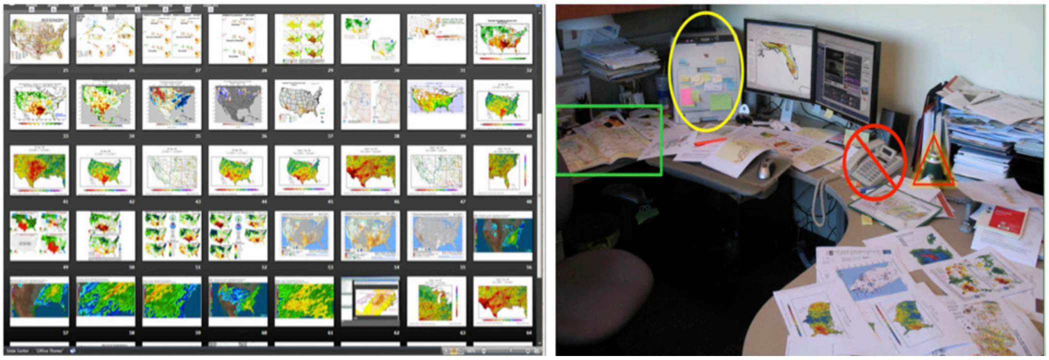

FIG. 5. USDM map construction involves surveying numerous science-based objective Earth system monitoring data products and integrating them into a single indicator map (Svoboda et al. 2002).

constructing a draft USDM map (Fig. 5 illustrates the integration effort). The draft map is reviewed by over 350 local- to national-level drought coordinators, agency leads, and experts, and their feedback is incorporated by the lead author, who targets a ''convergence of evidence'' consensus indicating a single drought severity category. The resulting final USDM map depicts this category, either for only one (specially noted) type of impact or for all facets of drought combined (i.e., meteorological, hydrological, and agricultural are widely accepted drought aspects). The USDM does not attempt to quantify uncertainty in the final map categories.

The use of indices is a central feature of drought and climate monitoring because indices provide a readily communicable description of relative severity and rarity, supporting the intercomparison of drought across a range of physical aspects, geography, and seasonality (Heim 2002). Some indices, such as the Palmer drought severity index (PDSI; Palmer 1965), the standardized precipitation index (SPI; McKee et al. 1993), and the surface water supply index (SWSI; Shafer and Dezman 1982), are traditionally derived from direct, in situ measurements. Others, such as total soil moisture percentiles (SMPs; Mo 2008; Sheffield and Wood 2008), standardized soil moisture index (SSI; Hao and AghaKouchak 2013), and the standardized runoff index (SRI; Shukla and Wood 2008) rely on simulation model outputs. More recently, satellite-based indices such as the evaporative stress index (ESI; Anderson et al. 2013) are developed from satellite imagery.

In parallel to the consensus-based integration process described above, the USDM also merges multivariate drought analysis inputs using two prescribed-weight blends of the input analyses: one depicts short-term drought conditions aimed at meteorological, environmental, and agricultural impacts, and the other is for long-term conditions aimed at hydrological impacts (example shown in Fig. 6). The fixed input weights in these blends were established through expert judgment rather than through statistically objective analytical procedures. The blends are reproducible, but their quantification of drought is nonetheless strongly influenced by the subjective choice of their components and weights.

Policymakers and media use the USDM in discussions of drought and in allocating drought relief. Since 2012, in fact, governmental disaster declarations have become nearly automatic for a county shown in severe drought on the USDM for eight consecutive weeks. A number of states also use the USDM or its input indices to trigger their local drought task force activities and drought declaration processes. For many stakeholders, the USDM and similar agency synthesis efforts are the mission-critical face of drought monitoring, and USDM maps provide a categorical measure of the existence of drought. As an alternative, the experimental research systems present drought through a collection of separate, univariate outputs of objective hydroclimatic analysis systems or of satellite-based indices. Integrated multivariate and location-specific analyses, which are central to USDM-type efforts, have only lately become a research focus, as is highlighted in the following section. Notable intersections of the operational and research spheres are found in the expanding consideration of research system results as input in USDM formation that is being facilitated by the production of those results in USDM category terms and the increased use of USDM results as a ground truth for validating experimental monitoring system performance where reliable observations are lacking.

b. *Research to advance operational drought monitoring*

The overarching objective of drought monitoring research is to develop accurate, reliable, high-resolution characterizations of the geophysical variables involved

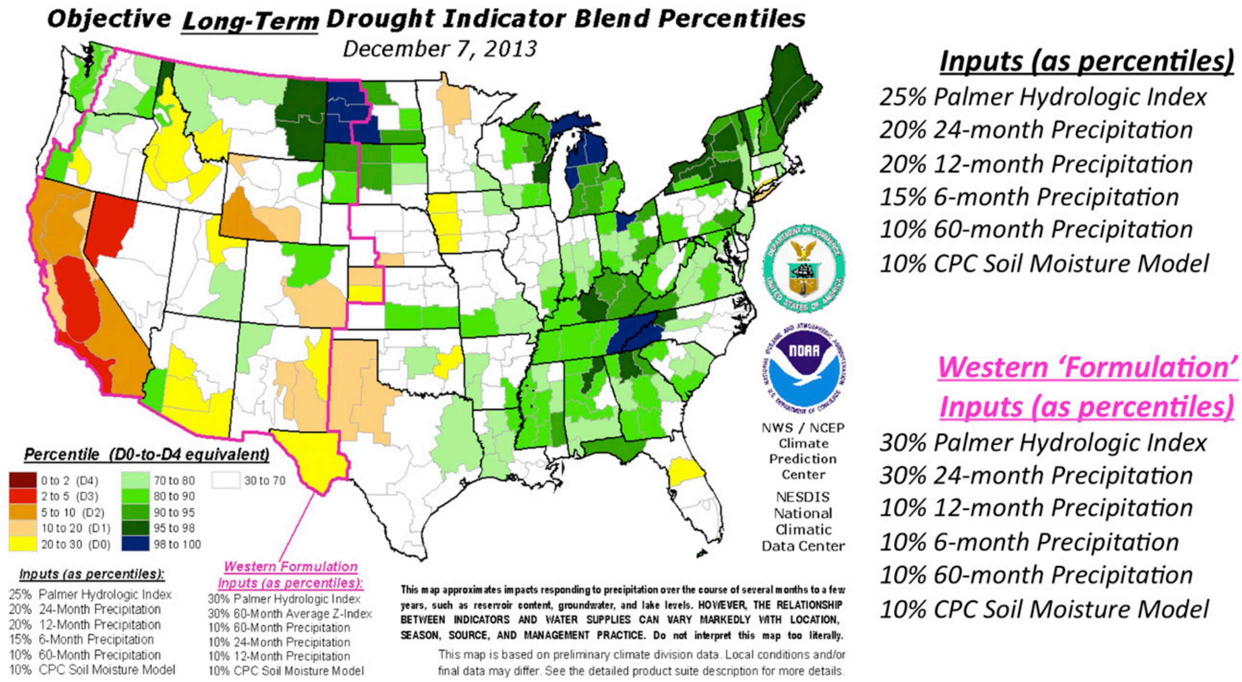

FIG. 6. (left) Objective long-term blend USDM input analysis with (right) two blending formulas, enlarged. (For more detail, see http://droughtmonitor.unl.edu/SupplementalInfo/CurrentConditions.aspx.)

in drought through objective science-based methods and data—the development of Drought Early Warning Systems. NOAA's DTF research pursues this objective through supporting (i) the real-time execution of macroscale land surface models that objectively, quantifiably, and reproducibly depict surface conditions using operational, real-time forcings; (ii) observational surface analyses based on satellite remote sensing retrievals of drought-relevant parameters; (iii) development of long-term retrospective climate system datasets and reanalyses; and (iv) research toward understanding the fundamentals of drought processes. Because the USDM relies on the objective geophysical analyses that benefit from advances in these areas, the DTF work has arguably benefitted the official USDM mapping process, and hence management activities. Yet, quantifying this improvement is difficult: the integration of map inputs is subjective, and there is no objective standard by which to assess the relative accuracy of current USDM ratings versus DEWS ratings, or versus ratings that would have emerged from earlier versions of drought monitor inputs. More measurable are improvements in the geophysical input analyses and the growing availability and sophistication of DEWS. This section summarizes efforts in the first three general areas above, and the fourth area is discussed in more detail in the context of prediction (section 4b).

1) LAND SURFACE MODELING AND INDICES

Phases 1 and 2 of NLDAS (NLDAS and NLDAS2, respectively; Mitchell et al. 2004; Xia et al. 2012) were initiated in 1999 and have since been steadily enhanced through NOAA and NASA research programs. Housed at the NCEP Environmental Modeling Center, NLDAS runs four land surface models at an hourly time step for a region enclosing CONUS at 1/8°. The associated forcing inputs (e.g., precipitation, temperature, humidity, wind speed, and radiation) and land surface model outputs (e.g., soil moisture, snow water equivalent, and evapotranspiration) represent a central thrust of science-based advances in drought monitoring and the core component of an effort to advance a national DEWS in support of NIDIS. For example, the USDM blend products (cf. Fig. 6) currently make use of the CPC soil moisture analysis, which is based on a leaky bucket soil moisture accounting formulation (Huang et al. 1996) that is surpassed in physical realism by the NLDAS modeling efforts. Similarly, the USDM climate division precipitation analysis is now at a coarser resolution than the NLDAS precipitation input; thus, the NLDAS data products can support a finer-resolution and higher-quality version of the USDM objective blends and other inputs.

One of the first such NLDAS-derived drought applications was the University of Washington Experimental

Surface Water Monitor (SWM; Wood 2008; www.hydro.washington.edu/forecast/monitor/), launched a decade ago to generate both real-time (1-day lag) analyses of CONUS-wide land surface moisture field anomalies and 3-month predictions oriented toward drought. Follow ups included the development of experimental continental- to global-scale systems such as the Princeton University Drought Monitor (Luo and Wood 2007; http://hydrology.princeton.edu/forecast/) and the University of California, Irvine, Global Integrated Drought Monitoring and Prediction System (GIDMaPS; Hao et al. 2014; http://drought.eng.uci.edu/). Princeton University has leveraged their work over CONUS to develop an African drought monitoring and forecasting system (Sheffield et al. 2014) that has been extended to Latin America (http://stream.princeton.edu). The NLDAS effort itself now supports a drought monitor (www.emc.ncep.noaa.gov/mmb/nldas/drought/) that provides similar moisture anomaly maps and other near-real-time (3–4 day lag) analyses. SWM and NLDAS have also come to incorporate a small ''poor-man's ensemble'' of land surface models, motivating a new research thrust of understanding and integrating intermodel differences in drought depiction. These systems differ in their approach to monitoring drought from USDM in that they provide real-time univariate depictions related to drought (i.e., of soil moisture or of snow) or bivariate analyses (as in the GIDMaPS combination of precipitation and soil moisture, described further below), rather than the broad integration of factors represented in the USDM. In contrast to the USDM, however, they provide long, objective, and reproducible retrospective analyses that allow for assessment of drought through recent history and clear definitions of frequency statistics. Such systems are the foundation of current efforts toward science-based DEWS.

A major current challenge for DEWS is the integration of multiple geophysical facets of drought to provide a more comprehensive depiction, similar to the paradigm of the USDM. To this end, Xia et al. (2014) developed the Objective Blended NLDAS Drought Index (OBNDI), which is an optimal blend of drought indices (monthly mean evapotranspiration, total runoff, top 1-m soil moisture, and total column soil moisture) from the NCEP NLDAS multimodel ensemble. The variable weights are formed by regressing the drought indices onto USDM drought area percentages for different drought categories. A reconstructed OBNDI achieved closer agreement with the USDM than any individual variable. The resulting blended index represents one promising, explicit pathway from the objective NLDAS drought analyses to the USDM map categories and also provides an increased spatial and temporal resolution product relative to the USDM. Future research directions include tailoring the weights spatially and seasonally and using the optimal blend framework to incorporate a broader range of independent analyses, for example, the ESI or USGS streamflow percentiles.

Several other DTF efforts also address the integration challenge through developing methods to combine different indices or model results into a single indicator. Hao and AghaKouchak (2014) propose the multivariate standardized drought index (MSDI), which is a multi-index approach to combine the SPI and SSI. This index will have higher severity values in the period when both indices show below-average precipitation and soil moisture, giving the index a better chance to capture the timing and intensity of a drought event that is influenced by more than one physical factor. The joint-probability estimation framework is extendable to a broader set of analyses and offers potential for emulating the convergence of evidence USDM philosophy using locally tailored, objective combinations. Mo and Lettenmaier (2014) used an averaging approach for combining multiple monitoring indices (in this case, the SPI, SRI, and soil moisture SM percentiles from the NLDAS) and the aforementioned SWM. The ensemble mean index is the sum of all indices, transformed to a uniform distribution by using the climatology of the ensemble (percentile) averages for each of the component variables. To assess uncertainties in the classifications, Mo and Lettenmaier (2014) introduce a concurrence measure showing the extent to which the different indices agree, which could provide an indication of confidence in the resulting metric—that is, the classification scheme provides information about drought severity as well as the representativeness of the ensemble mean index. The grand mean index is also given in the D0–D4 categories used by the USDM, to increase its relevance to the drought monitoring community.

2) Remotely sensed observational analyses

In addition to land surface model (LSM)-derived drought-related analyses, DTF research includes the development of new strategies for using satellite data to monitor drought (and floods), which can provide an assessment of drought characteristics independent of land surface model analyses. Anderson et al. (2013) compare the ESI with NLDAS model-based estimates of SM, evapotranspiration ET, and runoff anomalies, and with other empirical indices such as the vegetation health index (VHI) and SPI, using the USDM classifications as a reference. The ESI uses the thermal infrared (TIR) satellite-based Atmosphere–Land Exchange Inverse (ALEXI) energy balance model to estimate ET deficits. The authors found that while NLDAS ensemble-averaged SM anomalies correlated best with USDM classes, the ESI had the strongest relationship of the satellite-based indices. Furthermore, the study found that ESI and SM in combination provided a skillful indicator of drought severity changes, often preceding USDM class deterioration by several weeks, suggesting that in addition to a monitoring resource, ESI may enhance drought prediction capabilities for rapidly evolving ''flash drought'' conditions.

Like most current LSMs, the NLDAS models do not include a dynamical vegetation component and therefore do not capture the reduction in evaporation that can arise from vegetation changes caused by drought (e.g., crop damage or delay). In this regard, while this limitation is being remedied by NLDAS land model advances, the ESI may provide complementary information to the NLDAS. Figure 7 shows the comparison among the real-time ESI, the fraction green vegetation cover from MODIS, and the ET anomaly from the Noah model for approximately the same period (June 2013). The ESI captures the negative ET anomalies over California and the Four Corners region. The negative ESI anomalies over Minnesota and Iowa resemble the anomalies in vegetation cover fraction because of the delayed growth of crops there. On the other hand, ESI did not capture the vegetation signal in the Southeast (Fig. 7b), probably because the tree-dominated vegetation does not have as large a surface temperature signal as do crops.

Otkin et al. (2013) focused in detail on rapid-onset droughts and used the meteorology from the North American Regional Reanalysis (NARR) to show that these events are typically driven by the combination of warm temperatures, low rainfall, strong winds, and below-normal cloud cover that together act to enhance evaporation and rapidly dry the soil. The study underscored the findings of Anderson et al. (2013) in showing that the remotely sensed ESI captures these phenomena and can provide an early warning of drought impacts on agricultural systems.

Finally, Dong et al. (2014) focused on quantifying errors in MODIS fractional snow cover (FSC) datasets, which have been a useful input for hydrological analyses related to drought. The quantitative uncertainty assessment and a 34-yr high-resolution model climatology enhance hydrological assimilation and applications. Comparing MODIS FSC from 2000 to 2005 over the CONUS with an extensive observational network, the authors found that the more recent MODIS Collection 6 product generally improves over the prior version (Collection 5) in detecting the presence of snow cover, ranging from a 30% increase in probability of detection

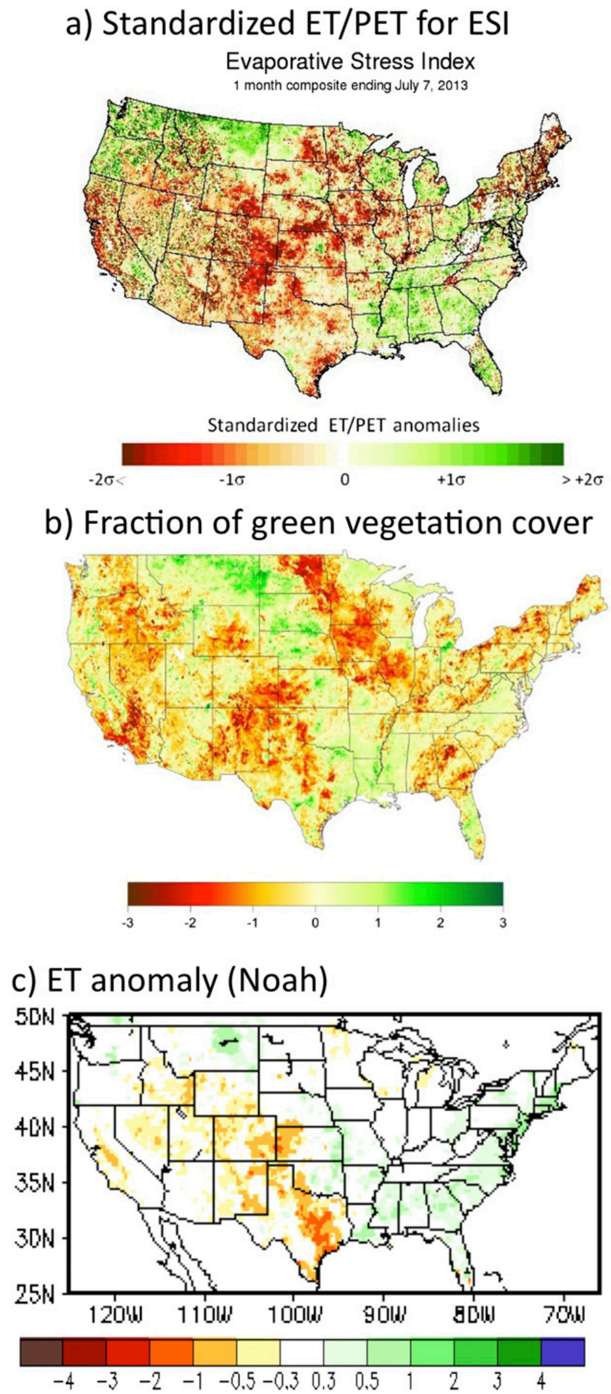

FIG. 7. For the month ending on 7 Jul 2013 (a) mean standardized ET/potential ET anomaly (ESI) and (b) the fraction of green vegetation cover from MODIS. (c) June 2013 ET anomaly from the NCEP NLDAS Noah model simulation. (Source is CPC July drought briefing.)

(POD) in Nevada to a relatively small improvement over Colorado (2% POD increase). The authors also demonstrate a relationship between the MODIS FSC retrieval errors and temperature, which can become a

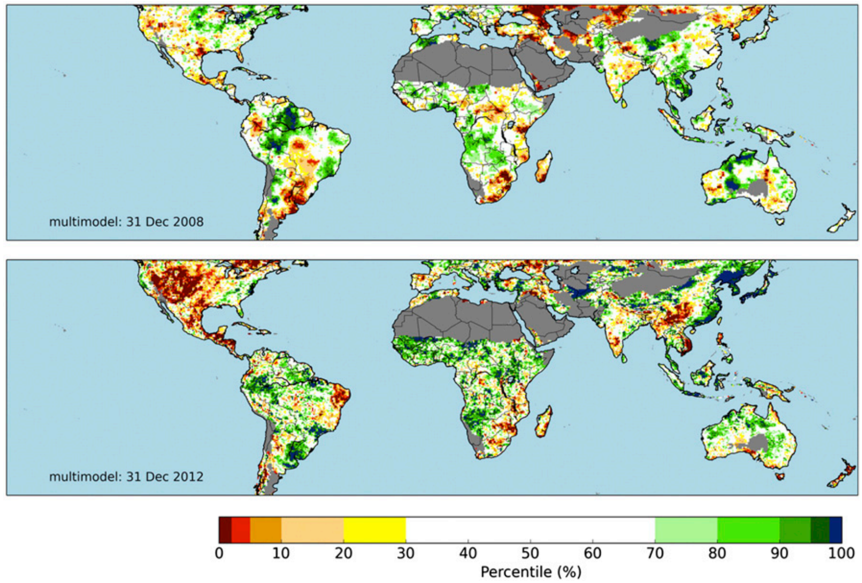

FIG. 8. GDIS total moisture percentiles on 31 Dec (top) 2008 and (bottom) 2012. [Source is Nijssen et al. (2014)].

useful index for filtering out the misclassification of MODIS snow cover pixels.

3) Global drought monitoring systems

The land surface modeling and satellite-based assessments have potential utility beyond CONUS, though their extensibility depends on the quality of any in situ measurement networks required for forcing or calibrating. The DTF included efforts toward the development of global drought monitoring approaches, such as the Global Drought Information System (GDIS) described in Nijssen et al. (2014) and a satellite-based SPI approach in AghaKouchak and Nakhjiri (2012). In Nijssen et al. (2014), the authors extended the SWM (Wood 2008) and NLDAS multimodel monitoring (Mo and Lettenmaier 2014) systems to a near-global system (GLDAS; covering from 50°S to 50°N) and used multiple LSM outputs to form ensemble drought-related indices, such as SMP (Fig. 8). The effort identified significant fundamental challenges in the establishment of such a system. For example, there is no global, long-term, consistent, homogeneous precipitation analysis from the historical period to near–real time that can be used as forcing for the LSM (mainly because the station networks used to form these analyses change over time). Such key model forcings require careful reconstruction of long-term records to ensure consistency with near-real-time records so that anomalies will not be due to the changes in the observing system. Clearly, the lack of data for both input and verification of monitoring is a major challenge in many regions of the world, which adds to the scientific challenges of drought research. Reliance on satellite data will be particularly important for the development of a GDIS, including areas where data are sparse, but compounds the challenge of developing a temporally consistent analysis system. The authors note that developing methodologies to exploit such data in regions with better data coverage (such as the United States) may be particularly useful.

4. Advances in drought prediction

As mentioned in the introduction, the DTF has overarching goals that include improving our understanding of the physical mechanisms of drought and improving drought prediction skill through full utilization of predictability sources by advanced prediction systems. Implicit in these goals is that improvements in understanding can ultimately lead to improvements in prediction skill, by improving prediction systems to adequately capture key linkages underpinning predictability. An important aspect of these goals is the identification of the sources of predictability (the signal) as well as the nature and magnitude of the unpredictable noise. It is now generally accepted that both aspects (the signal and noise) can be functions of time scale, season, region, and quantity of interest (e.g., Schubert et al. 2009). An added complication is that predictability itself can vary with time so that, for example, periods with large El Niño–Southern Oscillation (ENSO) variability can potentially have greater predictability than more quiescent periods in the tropical Pacific (e.g., Pegion and Kumar 2013).

Specifically for drought, important questions remain regarding our ability to predict various aspects of drought, including onset, duration, severity, and recovery. These uncertainties largely reflect our differing abilities to predict precipitation, temperature, soil wetness, snow, and runoff, which in turn reflect the basic mechanisms by which any predictable signal, say, in a slow component of the climate system (e.g., SST), is propagated through the climate system in the presence of weather and short-term climate noise to impact regional climate. As such, it is arguable that understanding the nature of (and possible changes in) the unpredictable noise is just as important as understanding the nature of the potentially predictable signal. A telling example of that was already provided in section 1 in the context of the CFSv2 hindcasts for April 2012, where basically the same SST forecasts led to very different outcomes (an example of large intraensemble spread) in terms of the precipitation and temperature forecasts over North America. This example also brings up the issue of what really matters in terms of the SST signal (spatially and temporally) insofar as what aspect of the SST drives the atmospheric response over North America: how accurately do we need to predict the SST, and are there particular regions (or even ocean basins) where the large-scale atmospheric response is particularly sensitive to SST anomalies? Other issues include the need to better understand the importance/role of land–atmospheric feedbacks and land initial conditions, as well as large-scale atmospheric variability, in the life cycle of drought, and key technical issues related to the need for higher model resolution and the impacts of model bias on prediction skill.

The DTF has attempted to make the very daunting task of advancing drought understanding and prediction manageable by developing a more limited framework that focuses on specific major drought events over North America. The basic idea is that these types of events are highly relevant since they have the greatest impacts on society, while at the same time the large magnitude of such events makes them the best candidates for identifying the important physical mechanisms and for understanding the key elements of successful drought predictions. Specifically, the DTF has developed a drought test bed framework that individual research groups can use to test/evaluate methods and ideas. As mentioned in section 1, central to this is a focus on four high-profile North American droughts that are key areas for NIDIS early warning system development (the 1998–2004 western U.S. drought, the 2006/07 southeastern U.S. drought, the 2011 Texas–northern Mexico drought over the southern plains, and the 2012 drought over the central Great Plains). To provide a more general assessment of prediction skill, the DTF has also embraced the NMME protocol for forecast evaluation covering the 30-yr (1981–2010) period as described in Kirtman et al. (2014).

In the following, we attempt to review, highlight, and synthesize the key outcomes of the DTF special collection in the context of the prediction issues discussed above. The focus is on how the relevant DTF contributions have advanced our understanding and contributed to improvements in drought prediction. A key aspect of the synthesis is to also identify any remaining gaps in our understanding as well as to identify aspects of problems that constitute "low hanging fruit," in the sense that one can take advantage of significant gaps between our current prediction capabilities and the limits to predictability. The section is broadly divided into subsections on (i) an assessment of current prediction capabilities and (ii) a synthesis of our understanding of drought and its predictability.

The section begins with a review of the contributions that address state-of-the-art drought prediction capabilities.

a. *Current drought prediction capabilities*

Successful drought prediction critically depends on skill in forecasts of both temperature and precipitation, knowledge of the current state of drought, and the ability to accurately model related changes in drought-relevant moisture stores such as soil moisture, groundwater, and snowpack. The U.S. Seasonal Drought Outlook (SDO; www.cpc.ncep.noaa.gov/products/expert_assessment/sdo_summary.html)[1] produced by NOAA CPC relies on forecaster judgment to combine sources such as the official CPC temperature and precipitation outlooks, long-lead forecasts including CFSv2, short-term forecasts from GFS and ECMWF, and initial conditions from the USDM. This subjective process currently produces a forecast map for monthly and seasonal means, such as that shown in Fig. 9. This example also illustrates the aforementioned challenges in drought operational prediction, such as the lack of skill forecasting the 2012 drought in the upper Great Plains, as discussed by Hoerling et al. (2014). For the 2011 Texas–northern Mexico drought, the SDO generally performed better but predicted improvements in areas where they were not observed, as shown in Fig. 10.

Since 2011, the NOAA MAPP program has been supporting the evaluation of the experimental NMME

[1] See also the monthly outlooks: www.cpc.ncep.noaa.gov/products/expert_assessment/mdo_summary.html.

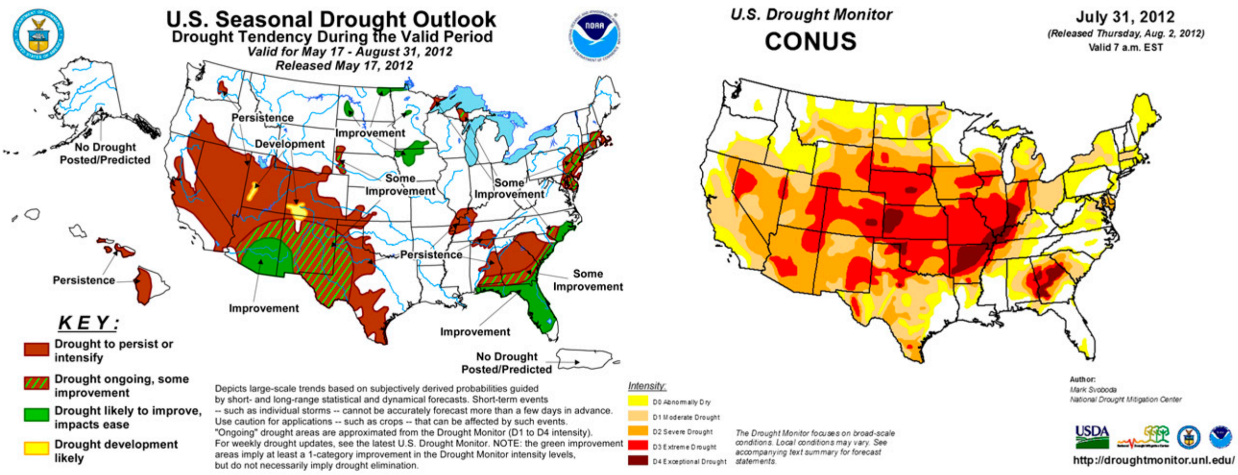

FIG. 9. (left) Seasonal drought outlook and (right) observed drought from USDM for summer 2012.

seasonal forecast system as part of the NOAA Climate Test Bed (CTB) and Climate Prediction Task Force research (http://cpo.noaa.gov/ClimatePrograms/ModelingAnalysisPredictionsandProjections/MAPPTaskForces/ClimatePredictionTaskForce.aspx). The NMME (Kirtman et al. 2014) leverages the considerable research and development activities on coupled model prediction systems carried out at universities and various research laboratories throughout North America. As described by Infanti and Kirtman (2014), the southeastern U.S. precipitation forecast skill of the NMME system typically equals or surpasses that of individual models throughout most seasons and lead times, but the skill tends to be low overall for summer seasons. There is a tendency for the southeastern U.S. region to show more skill in winter seasons versus summer seasons, and NMME was able to predict winter season variability. During the 2006/07 southeastern U.S. drought, the NMME showed moderate skill at short leads during more extreme seasonal phases of this drought, but a lack of skill at long leads, particularly during the driest phase of the drought.

Skill in 2-m air temperature and precipitation prediction is the foundation for NMME drought prediction skill. For example, Fig. 11 shows 1–3-month lead NMME temperature and precipitation forecast skill in 2011 and 2012. As the figure shows, the ensemble skill (shown as the green "mme" bar) is robust despite one or more low-skill members. As expected, temperature skill is generally higher than precipitation skill, and the ensemble skill for 2012 is reduced overall relative to 2011. In addition to the forecasts of temperature and precipitation, NMME forecasts have been developed of drought indices including the 1-, 3-, 6-, and 12-month SPI (www.cpc.ncep.noaa.gov/products/Drought/Monitoring/spi_outlooks_3.shtml).

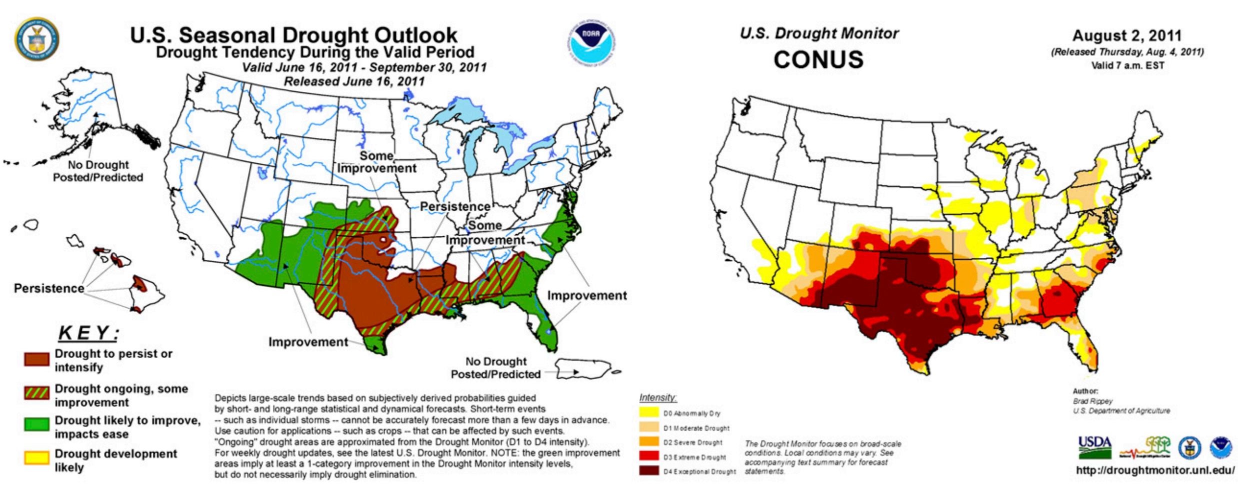

FIG. 10. As in Fig. 9, but for summer 2011.

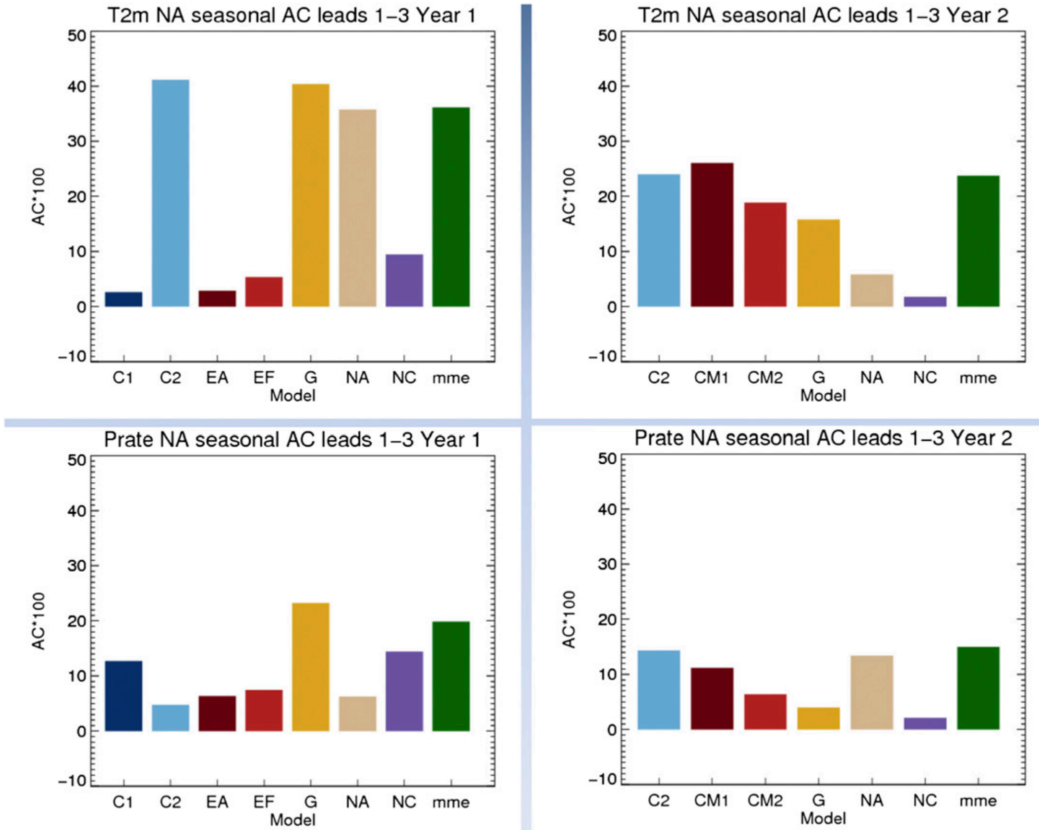

FIG. 11. NMME North American (top) temperature and (bottom) PR rate skill (anomaly correlation AC; %) for 1–3-month leads in year (left) 1 (2011) and (right) 2 (2012) for NMME participating models (C1, CFSv1; C2, CFSv2; EA, ECHAM4a; EF, ECHAM4f; G, GFDL CM2.1; NA, NASA GMAO; NC, NCAR; mme, multimodel average). (Additional information on the participating models can be found at www.cpc.ncep.noaa.gov/products/NMME/Phase1models.png.)

In an effort to improve land surface and drought forecasting, CTB activities also included the transition to a seasonal hydrological forecast system using CFSv2 and the VIC LSM (Liang et al. 1994) to NCEP. Yuan et al. (2013) demonstrated that Bayesian downscaled CFSv2 precipitation forecasts have higher correlations and smaller errors for monthly precipitation than the traditional ensemble streamflow precipitation (ESP). Further conditioning by ENSO yields skill out to 4 months. Streamflow forecasts using the CFSv2 precipitation as input to the VIC LSM yield limited skill beyond 1 month, but soil moisture–based drought frequency skill over the central United States from 2 months out to 6 months is possible when the precipitation forecasts are skillful. Overall, this study indicates that climate models can provide better seasonal hydroclimatic forecasts than ESP through appropriate downscaling procedures, but significant improvements are dependent on the variables, seasons, and regions. Closely related work by Madadgar and Moradkhani (2013) finds that probabilistic forecasts of spring streamflows in the upper Colorado River basin are more reliable at predicting drought flows than ESP.

The use of ensembles in NMME generally improves our temperature and precipitation prediction skill, although significant precipitation skill beyond 1-month lead is still a challenge. The results with the CFSv2-VIC system indicate that ENSO conditioning combined with good initial conditions can provide skill in land surface soil moisture forecasts for up to a 6-month lead time. This, and other related work, suggests that slowly evolving components of the land surface, such as soil moisture and groundwater, may provide additional information for drought forecasting. In the next section, recent DTF insights into drought mechanisms and predictability are discussed.

b. *Drought mechanisms and predictability*

Several DTF papers focused on improving our understanding of various hydrological processes (land, ocean, and atmosphere) and how these contribute to the development of drought. Here too there was considerable focus on the more recent 2010–12 period of intense droughts over the United States, with Texas and northern Mexico experiencing record drought during 2010/11 and the U.S. central and upper Great Plains feeling the grip of intense heat and drought during the summer of 2012. These studies were largely focused on providing insights into the physical processes that could lead to useful prediction skill (e.g., the role of La Niña conditions), as well as the fundamental limitations (predictability limits) imposed on prediction skill at seasonal and longer time scales by the role of internal atmospheric variability that is unforced by SST. The latter was especially true for the summer of 2012, where internal atmospheric variability appears to have played a key role in the rapid development of the heat and drought conditions in the central Great Plains.

Specifically, Seager et al. (2014) analyzed the causes of the 2010/11 drought in Texas and northern Mexico. They concluded that La Niña conditions in the tropical Pacific Ocean initiated the drought but also found that a very negative North Atlantic Oscillation (NAO) contributed to the dryness in the southern plains and southeastern United States. An important finding was that intensification of the drought in summer 2011 was not forced by SST but was most likely due to internal atmospheric variability. They noted that this was reflected in the skill of the models used by the International Research Institute for Climate and Society (IRI) for producing seasonal forecasts, which did predict drought onset in fall 2010 but did not predict drought intensification in summer 2011.

Hoerling et al. (2014) examined the central Great Plains drought of 2012. They noted that seasonal forecasts did not predict the intensity or the rapid development of the drought at lead times longer than about 1 month (see also Fig. 1). An assessment of the drought causes indicated an important role for natural variations in weather, including a reduction in atmospheric moisture transport from the Gulf of Mexico and the absence of processes that would provide airmass lift and condensation. They further concluded that neither ocean surface temperatures nor changes in greenhouse gases induced a substantial reduction in summertime rainfall over the central Great Plains during 2012. In trying to understand whether there existed some large-scale factors that might enhance the probability of such an extreme event, they conducted climate model simulations that revealed a regime shift toward warmer and drier summertime Great Plains conditions during the recent decade. This shift, most likely due to natural decadal variability, is such that the probability for a severe summer Great Plains drought may have increased fivefold in the last decade compared to the 1980s and 1990s.

Wang et al. (2014) compared the roles played by SST forcing in the evolution of the 2011 and 2012 U.S. droughts. They found that the pronounced winter and early spring temperature differences between the two years primarily reflect differences in the contributions from the Atlantic and Indian Oceans, with both acting to cool the east and upper Midwest during 2011, while during 2012 the Indian Ocean reinforced the Pacific-driven continental-wide warming and the Atlantic played a less important role. During late spring and summer 2011, the tropical Pacific SST forced a continued warming over the southern United States, with the Atlantic acting to extend the warming northward. The observed anomalies were, however, considerably stronger than the ensemble mean, though they fell well within the model's ensemble spread [consistent with the Seager et al. (2014) conclusion regarding the important role of internal atmospheric variability in driving the intensification of the drought]. Also consistent with Hoerling et al. (2014), they found that during June and July 2012, the rapid development of the intense heat and drying over the central United States was largely the result of internal atmospheric processes with only weak controls from (primarily Atlantic) SST forcing.

Additional research carried out by the DTF focused on basic questions concerning the usefulness of higher-resolution precipitation information for streamflow prediction and the nature of evaporative sources that supply moisture to the continental United States. In particular, Koster et al. (2014) quantified the degree to which errors in the initial soil moisture degrade a streamflow forecast and examined how the information content of high-resolution precipitation data translates to streamflow forecast skill. In particular, the linearity found between imposed soil moisture initialization error and the degradation of streamflow forecast skill allowed them to estimate the increase in forecast skill attainable with improved soil moisture measurement, for example, as expected from the NASA Soil Moisture Active Passive (SMAP) satellite mission. Using their land modeling system, they demonstrated that high-resolution precipitation forecasts will only be effective in improving (large scale) streamflow forecasts in areas with soil moisture–limited evaporation.

Dirmeyer et al. (2014) applied a quasi-isentropic back-trajectory scheme to output from the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) to estimate surface evaporative sources of moisture supplying precipitation over land for the period 1979–2005. Their methodology allows estimating moisture recycling and the partitioning of local precipitation between terrestrial and oceanic sources. They compared the evaporative sources for extreme situations like droughts or wet intervals to the corresponding climatological distributions using the method of relative entropy. A key finding of that study was that changes in local and remote surface evaporation sources of moisture supplying precipitation over land are more a factor in droughts than in wet periods over much of the globe, though further work is needed to differentiate between dynamical and hydrological factors causing the changes.

The above results highlight the fact that there are still considerable gaps in our understanding of the role of the SST in the different ocean basins (Pacific, Atlantic, and Indian Oceans) and how they reinforce or counteract each other to impact the hydrological conditions over North America. Current models do poorly in predicting SST in the Atlantic and Indian Oceans, suggesting that improvements in the predictions of SST in these ocean basins could help improve precipitation and temperature forecasts over North America on seasonal time scales. The results also emphasize that short-term (on roughly monthly time scales) extremes can have a substantial, unforced (by SST) component and that more work needs to be done to understand the nature and predictability of that variability including such phenomena as the Madden–Julian oscillation (MJO) and summertime extratropical stationary Rossby waves.

The papers also draw attention to the lack of operational skill for precipitation in summer and suggest that this deserves more attention to assess if progress can be made, for example, by better exploiting antecedent soil moisture conditions as well as large-scale atmospheric variability to improve forecast skill at subseasonal time scales. Results from the current studies suggesting a limited role for soil moisture in initiating the 2012 summer drought must be tempered by the fact that current models are still deficient in their representation of land–atmosphere coupling. Atmospheric reanalyses are becoming an increasingly important tool for analyzing the moisture budgets associated with drought. Unfortunately, despite substantial improvements in the latest products, there are still differences between them, and also apparent inconsistencies with observed precipitation anomalies, that limit progress in understanding the links between hydrological anomalies and circulation anomalies.

5. Discussion and conclusions

The investments in drought-related science, technology and information systems over the past decade have clearly enhanced and expanded the quality and range of drought data products, the number of people engaged in drought-related activities (such as the NIDIS drought early warning pilot projects; Pulwarty and Verdin 2013), and our understanding of drought as a phenomenon in the United States. This synthesis paper has highlighted areas in drought monitoring and prediction that can be considered successes, that remain as challenges, and that represent opportunities for the community.

a. Monitoring

In drought monitoring, for instance, a key success of the last decade has been the application and refinement of a modern class of hydrological models toward objective drought analysis, including extended retrospective forcing datasets to support hydrologic reanalyses that are nearly a century long. Objective drought analysis is critical for developing retrospective drought indices and forecasts of drought because they provide objective consistency that is not available from the interpretive approaches behind the USDM. Thus, while one goal of objective LSM-based drought monitors aligns with the USDM mission (i.e., to provide real-time drought analyses), the LSM-based efforts also strive to provide scientific insights into drought trends and variability and to serve real-time monitoring systems globally, reaching areas in which the resources for a USDM-style convergence of evidence approach are lacking. Numerous drought products and innovations have emerged from these simulation/assimilation efforts, among which are newly derived indices and new objective strategies for integrating indices and multiple sources of information. Drought monitoring system websites, drought information clearinghouses (e.g., www.drought.gov), outreach efforts, and web services make such products ever more accessible to the drought management community, including the USDM authors. A second key area of success has been the expansion of near-real-time satellite-based analyses that are relevant to drought, particularly those describing vegetation and evapotranspiration. These products further add to the information resources that can be used for characterizing current droughts and have been shown to often complement land model–based approaches.

These scientific and technological advances suggest that there may be commensurate advances in the accuracy of official USDM category maps that integrate those advances, yet it is difficult to quantify the presumed increases in skill. This difficulty arises in part from the multifaceted, poorly defined nature of drought's geophysical and social impacts, but it is also a feature common to operational activities in monitoring and forecasting that in the United States historically relies on subjective consensus processes to shape information outputs. The human effort arguably adds value, but it also obscures clear linkages to objective products and baselines and reduces the personnel time available for learning about, assessing, and developing strategies for integrating new data resources or methods. Objective approaches for data integration, without losing the advantages of the popular USDM product, are thus a critical need and an outstanding challenge. Numerous geophysical analyses of different facets of drought now exist—including, for example, ensembles of precipitation, soil moisture, and runoff anomalies at subcounty-scale resolutions, accompanied by estimates of uncertainty—but the operational USDM product does not yet reflect many of these potential enhancements. Investments that have spurred the broad expansion of drought information systems and products in the science community need to be matched by investments in operational arenas to enable the uptake of these advances, and secondarily to support greater interaction between operational and research entities [the so-called research-to-operations (R2O) and operations-to-research (O2R) interface]; this should not happen at the expense of foundational drought research. The DTF and the CTB are an ideal context for such enhanced "transition" investments. A critical institutional challenge is to create a pathway for development and testing (or ingesting) improvements into USDM—for example, the OBNDI of Xia et al. (2014) could provide one baseline for strategies to estimate USDM assessments.

Another institutional challenge is maintaining the nation's in situ, gauge-based observing networks for meteorological, climate, and hydrologic variables. A number of key measurements are either sparse or declining (see http://water.usgs.gov/streamgaging/), which impairs the evaluation and implementation of LSM-based simulations. Critical drought variables such as soil moisture and evaporation are not well observed, though several initiatives, such the establishment of the AmeriFlux network in 1996 to provide observations of water, energy, and momentum on an hourly basis; the National Resources Conservation Service (NRCS) SNOTEL network for snow and meteorological variables; and the NRCS Soil Climate Analysis Network (SCAN) and NOAA U.S. Climate Reference Network (USCRN) for soil moisture have helped. A multiagency and multistate national forum on the 2012 drought, convened by NIDIS, highlighted the need to sustain and even expand these efforts, rather than scaling back (NIDIS 2012). A national soil moisture network was also discussed as part of NIDIS (www.drought.gov/drought/news/developing-coordinated-national-soil-moisture-network-meeting).

Temporal or spatial coverage limitations of our observing networks and suboptimal reporting characteristics lead to scientific challenges as well. Many measurement stations that are active in the historical period do not report in real time because of insufficient gauge automation: as an example, data from more than one-third of the active precipitation stations are not accessible for days or months later than their recording time. Other data products, such as from radars or satellite platforms, lack long historical records. Unfortunately, drought is almost universally described in relative terms—for example, by departure from a mean, by return frequency, or by severity percentiles—thus, long, consistent hydrometeorological records are needed to characterize such metrics. Nearly all real-time monitoring and prediction systems struggle to enforce consistency between real-time analyses or prediction and historical analyses, such that real-time estimates of anomalies are accurate. The SWM of Wood (2008) and the GDIS described in Nijssen et al. (2014) grapple directly with this challenge, suggesting that the issue is resolvable, yet more investigation is clearly needed.

To the extent that monitoring challenges are institutional, they may be addressable through support via concerted agency programs and infrastructure development and greater integrated support for R2O that includes building operational capacity. Now that satellite-based remote sensing platforms are available or are about to be launched for soil moisture [Advanced Microwave Scanning Radiometer 2 (AMSR2), Soil Moisture Ocean Salinity (SMOS), and SMAP], precipitation [Global Precipitation Measurement (GPM)], and water levels [Surface Water Ocean Topography (SWOT) mission], one challenge that needs to be met is the near-real-time merging of these satellite data with in situ observations, perhaps using techniques developed in Chirlin and Wood (1982) and utilized in Chaney et al. (2014) to develop a high-resolution meteorological dataset over sub-Saharan Africa.

b. *Prediction*

The development of the NMME suite of seasonal model forecasts is a significant success in demonstrating the potential of having an international collaboration of operational and research groups focusing on both the generation of forecasts and their analysis. Of particular note are the ongoing scientific assessments from the project's participants as well as DTF and Climate Prediction Task Force scientists that have deepened our knowledge about the intermodel variability of forecasts and the strengths and challenges of using a multimodel ensemble system for drought forecasting. Key issues that remain to be resolved are when and why certain members of NMME are more skillful than the ensemble mean, and whether we can form an optimal ensemble conditioned on the phase of teleconnection patterns such as ENSO, the Pacific decadal oscillation (PDO), or

NAO. Current plans for the NMME include archiving high temporal (6 hourly) hindcasts and forecasts. This will provide both technological challenges of processing over 2 PB of model outputs and scientific challenges of analyzing the ensembles forecasts to answer drought predictability research questions (www.cpc.ncep.noaa.gov/products/ctb/nmme/NMME_Phase2_data_description.pdf). These archives being open and accessible to the community represent a great opportunity for drought research that, it is hoped, will result in important progress in the coming years in drought forecasting.

An important goal for NOAA and the drought impact community is the concept of ''seamless'' monitoring and forecasting of drought. A notable success is the implementation at NCEP of the NLDAS multi-LSM drought monitoring system discussed earlier. Also successful was the transition from Princeton University to NCEP via the Climate Test Bed of the CFSv2-VIC LSM seasonal hydrological forecasting system that offers a monitoring–forecasting capability as described earlier. The next logical goals in the seamless monitoring–prediction system would be implementing the four NLDAS drought monitoring LSMs into the CFSv2 seasonal forecasting system, and then the NMME suite of seasonal forecasting models together with the suite of NLDAS LSM to offer a comprehensive multimodel seasonal drought monitoring–forecasting system.

Substantial progress has been made in our understanding and quantification of the role of SST in producing drought over North America. The importance of La Niña in the southern Great Plains is now well established, and there are new results to suggest that the other oceans (Indian and Atlantic) can play an important role in either enhancing or suppressing the role of the Pacific. There is now also a better appreciation of the role of internal atmospheric variability in producing some of the most extreme droughts, limiting the predictability of such events on seasonal and longer time scales. Despite such progress, there are still important limitations in our understanding and ability to predict various aspects of drought, including onset, duration, severity, and recovery. A key challenge in that regard is to extend skillful precipitation forecasts beyond 1-month lead, or more generally, beyond lead times at which initial conditions (atmosphere and land) control the skill. Current results suggest that skillful precipitation forecasts combined with a well-initialized land model provide longer lead skill for predicting soil moisture than for streamflow, presumably reflecting differing predictability characteristics of these two aspects of drought.

As suggested above, improving the prediction of the full life cycle of droughts requires a better understanding of how any predictable signals propagate through the ocean–atmosphere–land system. This in turn should shed light on the necessary model improvements for improving drought predictions as well as the fundamental predictability limitations imposed on our ability to produce skillful forecasts of the various facets of drought involving precipitation, temperature, soil wetness, snow, and runoff. These limitations manifest themselves as signal-to-noise ratios that depend on quantity, region, season, forecast lead time, and potentially on the climate state itself. Quantifying such dependences requires better understanding the nature of (and changes in) the potentially predictable signal (e.g., that associated with SST anomalies) as well as the unpredictable noise (e.g., that associated with internal atmospheric variability).

Key challenges regarding potentially predictable signals beyond 1-month lead time involve better isolating what matters in terms of the SST signal (spatially and temporally) insofar as what aspect of the SST drives the atmospheric response over North America: that is, how accurately do we need to predict the SST, and are there particular regions (or even ocean basins) where the large-scale atmospheric response is particularly sensitive to SST anomalies? At these time scales the unpredictable signal is typically dominated during the cold season by well-known atmospheric teleconnections (e.g., the NAO, Arctic Oscillation, and Pacific–North American patterns), while during the warm season there is now mounting evidence that large-scale planetary (Rossby) waves (also largely driven by processes internal to the atmosphere) play an important role in contributing to some of the most extreme events (e.g., limiting our ability to predict the 2012 central Great Plains drought more than 1 month in advance). Land initialization (soil moisture and snow) is another key source of predictability at 1–2-month lead times. In addition to improving land–atmosphere coupling in climate models, there are uncertainties about the sensitivities to the land models, including how the skill lead times vary with LSM and how the skill in both soil moisture and streamflow depend on the model physics. Recent advances in the development of ultrahigh-resolution global climate models (so-called cloud-permitting models run at 10-km and higher resolution) offer new capabilities for addressing these challenges.

c. *Assessment of drought capabilities and research*

The challenge of assessing advances in operational drought monitoring and prediction through research has motivated the DTF to establish a Drought Capability Assessment Protocol. The fundamental goal of the protocol is to guide researchers toward evaluating their research outcomes through performance metrics that provide insights on their competitiveness with existing operational or state-of-the-art capabilities, which are treated as baselines or benchmarks. The protocol encourages U.S.-focused researchers to assess their efforts in the context of four recent drought case studies, the better to allow for intercomparison of research efforts from different and support synthesis. The protocol is described in more detail in the appendix.

This synthesis paper summarizes recent work that appears in this special collection, and the activities of the NOAA Climate Program Office's Drought Task Force, which through its working groups has supported and coordinated improvements in drought monitoring, drought prediction, and the development of metrics to assess these improvements. As witnessed over the last few years, the United States is highly vulnerable to droughts—2011 over Texas, 2012 over the Great Plains, and 2014 in the western United States and particularly in California—that have severe social and economic impacts. Over the last decade, there have been significant improvements in monitoring droughts, particularly the development of objective monitoring systems leveraged off the North American Land Data Assimilation System (NLDAS). Currently, a multimodel NLDAS runs at NCEP, with its information being integrated into the U.S. Drought Monitor system along with their observer-based information. Understanding drought mechanisms—the hydroclimate drivers that lead to drought—that can lead to improved predictive skill of droughts on seasonal time scales has proven to be a larger challenge. This is a research area that will require sustained support and effort for achieving progress. Nonetheless, as discussed earlier in this paper, progress is occurring, and with advances from the North American Multimodel Ensemble (NMME) project (Kirtman et al. 2014), there is the expectation that understanding of drought mechanisms will advance significantly. Supporting the monitoring and prediction work critically includes developing assessment metrics, verification datasets, and benchmarking so progress can be measured objectively. Starting in 2014, the Drought Task Force will enter a new 3-yr phase based on a new set of drought research projects that will address several high-priority research areas. Despite these new investments, it is expected that the broad challenges laid out in the synthesis paper will remain and will necessitate sustained research and programmatic attention.

APPENDIX

Drought Capability Assessment Protocol

The DTF Drought Capability Assessment Protocol was established to guide researchers toward quantifying the benefits of their research with respect to existing drought monitoring and prediction capabilities. Scientists should be able to apply the common protocol to help provide quantitative answers to the basic question: Is my research effort improving upon current capabilities to monitor or predict drought, and by how much? The protocol centers attention on four high-profile North American droughts and requires the use of drought-specific performance metrics that are applied, where appropriate, to standard evaluation periods and datasets. The elements of the protocol are described below.

a. Assessment metrics

As part of the protocol, researchers should apply the metrics in Table A1 to assess their work's ability to quantitatively detect (for monitoring) or forecast (for prediction) drought. The performance metrics can be extended, but in general should

- define criteria that separate drought conditions from other system states and
- describe key geophysical drought features that are of interest to decision makers in applications sectors and that are motivated by societal impacts. Examples include the onset, severity, duration, and change in intensity of a drought variable.

Metrics should be assessed by lead time for prediction, but not monitoring, and other conditional factors should be considered where warranted.

b. Verification datasets

A central part of drought capability assessment is the use of verification, and the protocol defines the following guidance on the use of verification datasets.

- Precipitation and temperature: Station observations and gridded analyses where appropriate (e.g., satellite, gauge, radar blends of sufficient period coverage, extent, and quality).
- Drought categories: USDM categories may be used as verifying observations for categorical estimates or predictions unless other impact-based quantifications of drought existence or severity are available. In some cases it may be appropriate to verify categorical drought against univariate percentiles, for example, from NLDAS soil moisture.
- Hydrologic fields: In situ observations or derived analyses are a primary verification resource. Examples include soil moisture from NRCS SCAN or the NOAA USCRN soil moisture networks; snow water equivalent from SNOTEL or USHCN; snow cover from IMS, MODIS, or Landsat; and streamflow from

TABLE A1. DTF protocol assessment metrics.

Key predictand(s) for drought variable (e.g., PR, temperature, SM, streamflow)	Metric(s) and skill scores comparing
Onset and/or recovery of drought condition	Lead time of prediction Error of identification
Duration and severity of drought condition	Error, bias, and correlation (time and value)
Indication (detection and prediction) of drought condition: deterministic	Categorical metrics: Critical success index, equitable threat score, POD, false alarm rate, and others
Probability of drought condition: probabilistic	Brier skill score (binary); secondarily, Brier decompositions for reliability and resolution
Value, overall	1. Error, bias, and correlation (of ensemble mean or median for probabilistic)
Value given drought occurring in the observed or forecast period	2. Continuous rank probability score

USGS gauge observations. For predictions, verification fields may also include observation-driven analyses or simulations (e.g., from NLDAS2) or quality-controlled input fields to the USDM. In general, verifying monitoring simulations on other simulations is discouraged.

c. *Verification periods and case studies*

Researchers' analyses should focus on one or more of the following four case studies to facilitate comparison with other community research.

1) From winter 2001 to spring 2002 severe western U.S. drought: Focus roughly on an area consisting of the six states of California, Nevada, Utah, Arizona, New Mexico, and Colorado from December 2001 through May 2002; evaluation on the overall 1998–2004 is also encouraged.
2) From fall 2005 to summer 2008 sustained southeastern U.S. drought: Focus roughly on an area consisting of the four states Tennessee, Mississippi, Alabama, and Georgia from fall 2005 through summer 2008.
3) The 2010/11 water year drought over the southern plains: Focus roughly on Texas, from October 2010 through September 2011.
4) The 2012 summer drought over the central Great Plains: Focus roughly on a region consisting of the six states Wyoming, Colorado, Nebraska, Kansas, Missouri, and Iowa from May to September 2012.

Forecast capability evaluation over a 30-yr (1981–2010) period or longer is encouraged if relevant and feasible. Hindcasts or retrospective simulations of these events should be utilized, including, for example, the Climate Forecast System Reanalysis and Reforecast (CFSRR), the NCEP ESRL GEFS reforecast, and NARR and MERRA.

d. *Baselines and benchmarking*

The use of familiar operational or current capability baselines is critical to making drought research relevant for potential transition to operational usage. Primary baselines include but are not limited to following:

- For monitoring or assessment capabilities: USDM, NLDAS2 drought monitor, SNOTEL-based analyses, NCDC PDSI, and Vegetation Drought Response Index (VegDRI).
- For prediction capabilities: CFSv2 or IRI SPI forecast; CPC monthly and seasonal drought outlooks; streamflow predictions created via the ESP approach or by statistical water supply forecasting procedures (operational center datasets) are preferred if available; and NCDC's PDSI forecasts, if appropriate.

The benchmarking activities apply the assessment metrics over the selected verification period or case studies, focusing on variables including precipitation, temperature, snow water equivalent, soil moisture, evaporative variables, runoff, and streamflow for the periods, case studies, or regions described above. Assessments of future new capabilities should follow the same approach but apply the metrics to new methods or models to the variables, periods, and regions defined in this protocol. The improvements and impacts will be compared to the benchmark performance values.

REFERENCES

AghaKouchak, A., and N. Nakhjiri, 2012: A near real-time satellite-based global drought climate data record. *Environ. Res. Lett.,* **7,** 044037, doi:10.1088/1748-9326/7/4/044037.

Anderson, M., C. Hain, J. Otkin, X. Zhan, K. Mo, M. Svoboda, B. Wardlow, and A. Pimstein, 2013: An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. Drought Monitor classifications. *J. Hydrometeor.,* **14,** 1035–1056, doi:10.1175/JHM-D-12-0140.1.

Below, R., E. Grover-Kopec, and M. Dilley, 2007: Documenting drought-related disasters: A global reassessment. *J. Environ. Dev.,* **16,** 328–344, doi:10.1177/1070496507306222.

Chaney, N. W., J. Sheffield, G. Villarini, and E. F. Wood, 2014: Development of a high-resolution gridded daily meteorological dataset over sub-Saharan Africa: Spatial analysis of trends in climate extremes. *J. Climate,* **27,** 5815–5835, doi:10.1175/JCLI-D-13-00423.1.

Chirlin, G. R., and E. F. Wood, 1982: On the relationship between kriging and state estimation. *Water Resour. Res.,* **18,** 432–438, doi:10.1029/WR018i002p00432.

Dirmeyer, P. A., J. Wei, M. G. Bosilovich, and D. M. Mocko, 2014: Comparing evaporative sources of terrestrial precipitation and their extremes in MERRA using relative entropy. *J. Hydrometeor.,* **15,** 102–116, doi:10.1175/JHM-D-13-053.1.

Dong, J., M. Ek, D. Hall, C. Peters-Lidard, B. Cosgrove, J. Miller, G. Riggs, and Y. Xia, 2014: Using air temperature to quantitatively predict the MODIS fractional snow cover retrieval errors over the continental United States. *J. Hydrometeor.,* **15,** 551–562, doi:10.1175/JHM-D-13-060.1.

Hao, Z., and A. AghaKouchak, 2013: Multivariate Standardized Drought Index: A parametric multi-index model. *Adv. Water Resour.,* **57,** 12–18, doi:10.1016/j.advwatres.2013.03.009.

——, and ——, 2014: A nonparametric multivariate multi-index drought monitoring framework. *J. Hydrometeor.,* **15,** 89–101, doi:10.1175/JHM-D-12-0160.1.

——, ——, N. Nakhjiri, and A. Farahmand, 2014: Global integrated drought monitoring and prediction system. *Sci. Data,* **1,** 140001, doi:10.1038/sdata.2014.1.

Heim, R. R., Jr., 2002: A review of twentieth-century drought indices used in the United States. *Bull. Amer. Meteor. Soc.,* **83,** 1149–1165, doi:10.1175/1520-0477(2002)083<1149:AROTDI>2.3.CO;2.

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.

Huang, J., H. M. van den Dool, and K. G. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *J. Climate,* **9,** 1350–1362, doi:10.1175/1520-0442(1996)009<1350:AOMCSM>2.0.CO;2.

Infanti, J. M., and B. P. Kirtman, 2014: Southeastern U.S. rainfall prediction in the North American Multi-Model Ensemble. *J. Hydrometeor.,* **15,** 529–550, doi:10.1175/JHM-D-13-072.1.

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.

Koster, R. D., G. K. Walker, S. P. P. Mahanama, and R. H. Reichle, 2014: Soil moisture initialization error and subgrid variability of precipitation in seasonal streamflow forecasting. *J. Hydrometeor.,* **15,** 69–88, doi:10.1175/JHM-D-13-050.1.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.,* **99,** 14 415–14 428, doi:10.1029/94JD00483.

Luo, L., and E. F. Wood, 2007: Monitoring and predicting the 2007 U.S. drought. *Geophys. Res. Lett.,* **34,** L22702, doi:10.1029/2007GL031673.

Madadgar, S., and A. Moradkhani, 2013: A Bayesian framework for probabilistic seasonal drought forecasting. *J. Hydrometeor.,* **14,** 1685–1705, doi:10.1175/JHM-D-13-010.1.

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

Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.,* **109,** D07S90, doi:10.1029/2003JD003823.

Mo, K., 2008: Model-based drought indices over the United States. *J. Hydrometeor.,* **9,** 1212–1230, doi:10.1175/2008JHM1002.1.

——, and D. Lettenmaier, 2014: Objective drought classification using multiple land surface models. *J. Hydrometeor.,* **15,** 990–1010, doi:10.1175/JHM-D-13-071.1.

NCDC, 2012: Summer 2012 drought update. [Available online at www.ncdc.noaa.gov/news/summer-2012-drought-update.]

NIDIS, 2007: The National Integrated Drought Information System Implementation Plan. NIDIS Rep., 34 pp. [Available online at www.drought.gov/media/imageserver/NIDIS/content/whatisnidis/NIDIS-IPFinal-June07.pdf.]

——, 2012: National Drought Forum summary report and priority actions: Drought and U.S. preparedness in 2013 and beyond. NIDIS Rep., 53 pp. [Available online at www.drought.gov/media/pgfiles/2012-droughtForumFullReport.pdf.]

Nijssen, B., S. Shukla, C. Lin, H. Gao, T. Zhou, J. Sheffield, E. F. Wood, and D. P. Lettenmaier, 2014: A prototype global drought information system based on multiple land surface models. *J. Hydrometeor.,* **15,** 1661–1676, doi:10.1175/JHM-D-13-090.1.

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

Palmer, W. C., 1965: Meteorological drought. Research Paper 45, U.S. Weather Bureau, 58 pp.

Pegion, K., and A. Kumar, 2013: Does an ENSO-conditional skill mask improve seasonal predictions? *Mon. Wea. Rev.,* **141,** 4515–4533, doi:10.1175/MWR-D-12-00317.1.

Pozzi, W., and Coauthors, 2013: Towards global drought early warning capability: Expanding international cooperation for the development of a framework for global drought monitoring and forecasting. *Bull. Amer. Meteor. Soc.,* **94,** 776–785, doi:10.1175/BAMS-D-11-00176.1.

Pulwarty, R., and J. Verdin, 2013: Crafting integrated early warning systems: The case of drought. *Measuring Vulnerability to Natural Hazards: Towards Disaster Resilient Societies,* 2nd ed. J. Birkmann, Ed., United Nations University Press, 720 pp.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24,** 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

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

Schubert, S., and Coauthors, 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. *J. Climate,* **22,** 5251–5272, doi:10.1175/2009JCLI3060.1.

Seager, R., L. Goddard, J. Nakamura, N. Henderson, and D. E. Lee, 2014: Dynamical causes of the 2010/11 Texas–northern Mexico drought. *J. Hydrometeor.,* **15,** 39–68, doi:10.1175/JHM-D-13-024.1.

Shafer, B. A., and L. E. Dezman, 1982: Development of a Surface Water Supply Index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. *Proc. 50th Annual Western Snow Conf.,* Reno, NV, Western Snow Conference, 164–175. [Available online at www.westernsnowconference.org/node/932.]

Sheffield, J., and E. F. Wood, 2008: Global trends and variability in soil moisture and drought characteristics, 1950–2000, from observation-driven simulations of the terrestrial hydrologic cycle. *J. Climate,* **21,** 432–458, doi:10.1175/2007JCLI1822.1.

——, and Coauthors, 2014: A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bull. Amer. Meteor. Soc.,* **95,** 861–882, doi:10.1175/BAMS-D-12-00124.1.

Shukla, S., and A. W. Wood, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.,* **35,** L02405, doi:10.1029/2007GL032487.

Smith, A., and R. Katz, 2013: U.S. billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Nat. Hazards,* **67,** 387–410, doi:10.1007/s11069-013-0566-5.

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.

U.S. Government, 2014: National Integrated Drought Information System Reauthorization Act of 2014. H. R. 2431, 113th Congress.

Wang, H., S. Schubert, R. Koster, Y.-G. Ham, and M. Suarez, 2014: On the role of SST forcing in the 2011 and 2012 extreme U.S. heat and drought: A study in contrasts. *J. Hydrometeor,* **15,** 1255–1273, doi:10.1175/JHM-D-13-069.1.

Wood, A. W., 2008: The University of Washington Surface Water Monitor: An experimental platform for national hydrologic assessment and prediction. *22nd Conf. on Hydrology,* New Orleans, LA, Amer. Meteor. Soc., 5.2. [Available online at https://ams.confex.com/ams/88Annual/techprogram/paper_134844.htm.]

Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for 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.

——, M. B. Ek, D. Mocko, C. D. Peters-Lidard, J. Sheffield, J. Dong, and E. F. Wood, 2014: Uncertainties, correlations, and optimal blends of drought indices from the NLDAS multiple land surface model ensemble. *J. Hydrometeor.,* **15,** 1636–1650, doi:10.1175/JHM-D-13-058.1.

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