Drought Task Force Assessment Protocol

August 8, 2013

Overview and Objectives

NOAA's Drought Task Force (DTF) has developed a capability assessment protocol to guide researchers toward quantifying the benefits of their activities with respect to existing drought monitoring and prediction capabilities. The DTF was established in October 2011 with the goal of achieving significant new advances in the ability to understand, monitor and predict drought over North America. The Task Force is an initiative of NOAA's Climate Program Office Modeling, Analysis, Predictions, and Projections (MAPP) program in partnership with NIDIS. It brings together over thirty-five leading drought scientists, primarily but not exclusively MAPP-funded, from multiple academic and federal institutions. The group is comprised of scientists from research laboratories and/or operational centers from NOAA, NASA, and the U.S. Department of Agriculture; and partners from the National Drought Mitigation Center, the National Center for Atmospheric Research (NCAR), and other groups. Their concerted research effort builds on individual MAPP research projects and related drought-research sector developments. The projects span the wide spectrum of drought research needed to make fundamental advances, from those aimed at the basic understanding of drought mechanisms to those evaluating new drought monitoring and prediction tools for operational and service purposes, and as part of NCEP's Climate Test Bed.

A major thrust of the DTF has been to develop a drought test-bed framework that individual research groups can use to develop and evaluate methods and ideas. Central to this is a focus on four high-profile North American droughts that are key areas for NIDIS early warning system development (1998-2004 western US drought, 2006-2007 SE US drought, the 2010- 2012 Tex-Mex drought over the Southern Plains, and the 2012 summer Midwestern US drought). The intent is to develop a framework that facilitates collaboration among projects, defines metrics to assess the quality of monitoring and prediction products, and helps to develop an experimental drought monitoring and prediction system that incorporates and assesses recent advances.

The DTF test-bed framework has been developed through the initial foundational efforts of three working groups (WG) that helped to address the major aspects of the test-bed: 1) WGI - Metrics: to define and apply metrics to evaluate advances in drought monitoring and prediction 2) WGII - Case Studies: to analyze drought cases by integrating all aspects of drought research and 3) WGIII - Experimental System: to incorporate research advances in an experimental drought monitoring and prediction system and assess improvements. More recently, a Research-to-Capability (RtC) activity has been initiated as part of the DTF with the goal of assessing recent progress in drought monitoring and prediction, with an eye towards advancing operational and service capabilities, building on the metrics and case studies framework developed by the WGs.

The DTF Assessment Protocol is presented as a reference for the groups that will engage in the DTF RtC activity, establishing guidelines for this assessment activity. The protocol may also be useful for drought researchers beyond the DTF effort. 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 should be viewed as a first step toward providing a community approach to such a capability assessment, and one that can expand to be more comprehensive as needed.

Protocol Principles

The protocol should include research performance measures that are:

- Specific to drought and define thresholds or criteria that separate drought conditions from other system states and phenomena
- A description of key geophysical features of drought that are of interest to decision makers in applications sectors and motivated by societal impacts. Examples include the onset, severity, duration, and change in intensity of a drought variable.
- Centered on the drought event case studies selected by the DTF and include the application of statistically robust metrics.

Protocol Elements

1. Assessment Metrics

As part of the protocol, researchers should apply the metrics in the table below to determine the ability to detect (for monitoring) and to forecast (for prediction) drought, respectively. Metrics should be assessed by lead time for prediction, but not monitoring; and other conditional factors should be considered where warranted. The metrics can be reported in presentations, project reports and publications, and expressed in terms that address the basic DTF question posed above.

Key predictand (s) for drought variable (e.g., P, T, soil moisture, streamflow)	**Metric(s) and skill scores comparing**
Onset and recovery of drought condition	Lead time of prediction Error of identification
Duration and severity of drought condition	Error, bias, correlation (time, value)
Indication (detection, prediction) of drought condition: deterministic	Categorical metrics: Critical Success Index (CSI), Equitable Threat Score (ETC) Probability of Detection (POD), False Alarm Rate (FAR), and others.
Probability of drought condition: probabilistic	Brier Skill Score (binary); secondarily, Brier decompositions for reliability and resolution
Value, overall Value given drought occurring in the observed or forecast period	1. Error, bias, correlation (of ensemble mean or median for probabilistic) 2. Ranked Probability Score (CRPS)

2. Verification data

Many verification data in drought categories and hydrologic fields are indices or ad hoc products. There is a need to be cautious on the uncertainties of all those products.

- **Precipitation**: surface rain gauge observations and blended precipitation analyses where appropriate (e.g., satellite, gage, radar blends of sufficient period coverage, extent and quality).

- **Temperature**: station observations and gridded analyses derived from station data and other sources, where appropriate.

- **Drought categories**: US Drought Monitor (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, e.g., 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 North American Soil Moisture Data Base, snow water equivalent from SNOTEL or USHCN, snow cover from IMS, MODIS or Landsat, and streamflow from USGS gauge observations. For predictions, verification fields may also include observation-driven analyses or simulations (e.g., from NLDAS-2), or quality controlled input fields to the USDM. In general, verifying with monitoring simulations on other simulations is discouraged.

3. Verification periods and Case Studies

- The **four case studies** selected for drought capability evaluation are the following:

    1) Winter 2001-Spring 2002 severe western US drought event.
        - Focus roughly on an area consisting of the 6-states CA, NV, UT, AZ, NM, and CO for Dec. 2001 thru May 2002, the primary wet season for most of that region
        - Evaluation of the overall 1998-2004 drought is also encouraged.
    2) Fall 2005-Summer2008 sustained southeast US drought period
        - Focus roughly on an area consisting of the 4-states TN, MS, AL, GA, for which precipitation was mostly below average season-over-season beginning in Fall 2005 thru summer 2008. Rain began recovering in Fall 2008
    3) The 2010- 2011 water-year drought over the Southern Plains,
        - Focus roughly on Texas, for the period beginning abruptly in Oct 2010 and continuing thru Sept 2011
    4) The 2012 summer drought over the Central Great Plains,

  - Focus roughly on a 6-state region of WY, CO, NE, KS, MO, IA for the period beginning abruptly in May 2012 and ending Sept 2012
- Forecast capability evaluation over a **30-year (1981-2010) period** or longer is encouraged if relevant and feasible following the NMME protocol (See details on NMME Protocol in the Appendix). Researcher's analyses should focus on one or more of these events to facilitate comparison with other community research. Hindcasts or retrospective simulations of these events should be utilized, including, for example, the CFSRR; the NCEP/ESRL GEFS Reforecast; NARR and MERRA.

4. Baselines and benchmarking

The **baselines** against which research efforts are to be measured reflect existing operational or research capabilities. Primary baselines include:

- For monitoring or assessment capabilities
  - US Drought Monitor (USDM)
  - NLDAS Drought Monitor
  - SNOTEL-based analyses, e.g., SWSI
  - NCDC PDSI
  - VegDRI
- For prediction capabilities
  - CFSv2 or IRI's SPI forecast for atmospheric drought features (without further pre- or post- processing)
  - CPC Monthly and Seasonal Drought Outlooks
  - Streamflow predictions created via the Ensemble Streamflow Prediction (ESP) approach or by statistical water supply forecasting procedures (e.g., principle components regression), both of which represent current operational capabilities. Operational center datasets are preferred if available.
  - NCDC's PDSI forecasts, if appropriate

The **benchmarking** activities apply the metrics defined in the protocol to the selected verification period or case studies to assess the baseline capabilities and define baseline performance in terms of the drought metrics. Specifically, the benchmarking assessment will be for the following variables, periods, and regions.

- **Variables**: precipitation, temperature, snow water equivalent, soil moisture, evaporative variables, runoff, streamflow
- **Periods**: the four case study periods or the NMME hindcast period (defined in Section 3)
- **Regions**: drought region in each case or CONUS

Assessments of future new capabilities will follow the same approach as the benchmarking procedure 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.

5. Data Resource Links

The following data resources are either mentioned in the sections above or related to the drought assessment topic.

- **NIDIS Portal (a range of products):**
  Forecasting: :http://www.drought.gov/drought/content/products/forecasting
  Impacts: http://www.drought.gov/drought/content/products/impacts
  Monitoring: http://www.drought.gov/drought/content/products/current-drought-and-monitoring.
- **US Drought Monitor (USDM)**: http://droughtmonitor.unl.edu/dmshps_archive.htm.
- **CPC Drought Information Site**: http://www.cpc.ncep.noaa.gov/products/Drought/
- **NLDAS Drought Monitor**: http://www.emc.ncep.noaa.gov/mmb/nldas/drought
- **NCDC PDSI**: http://www.ncdc.noaa.gov/oa/climate/research/prelim/drought/palmer.html
- **NRCS Mountain Snowpack Maps**: http://www.wcc.nrcs.usda.gov/cgibin/ms.pl
- **NRCS Surface Water Supply Forecasts**: http://www.wcc.nrcs.usda.gov/wsf/wsf.html
- **Vegetation Drought Response Index (VegDRI)**: http://vegdri.unl.edu
- **NRCS SNOTEL and SCAN sensor data**: http://www.wcc.nrcs.usda.gov/products.html
- **North American Soil Moisture Data Base**: http://soilmoisture.tamu.edu
- **US Historical Climatology Network**:
  http://cdiac.ornl.gov/epubs/ndp/ushcn/ushcn_map_interface.html
- **CFSv2**: http://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2seasonal.shtml
- **IRI SPI forecast**:
  http://iridl.ldeo.columbia.edu/maproom/.Regional/.N_America/.Drought/
- **CPC Monthly Drought**:
  http://www.cpc.ncep.noaa.gov/products/expert_assessment/monthly_drought.html
- **CPC Seasonal Drought**:
  http://www.cpc.ncep.noaa.gov/products/expert_assessment/seasonal_drought.html
- **NMME web site:** http://www.cpc.ncep.noaa.gov/products/NMME
- **National Weather Service Seasonal-Scale Streamflow Predictions**:
  Only a few River Forecast Centers maintain sufficient forecast archives to provide an operational baseline for skill assessment (one is the CBRFC: http://www.cbrfc.noaa.gov). Contact RFCs directly to inquire about their data resources.

•

Appendix

NMME Phase-I Hindcast and Real-time Experimental Prediction Protocol

The CY2011 NMME experimental predictions have been made in real-time since August 2011. As part of the development of the real-time capability, the NMME partners agreed on a hindcast and real-time prediction protocol. Some of the key elements of this protocol include:

- Real-time ISI prediction system must be identical to the system used to produce hindcasts. This necessarily includes the procedure for initializing the prediction system. The number of ensemble members per forecast, however can be larger for the real-time system.

- Hindcast start times must include all 12 calendar months, but the specific day of the month or the ensemble generation strategy is left open to the forecast provider.

- Lead-times up to 9 months are required, but longer leads are encouraged.

- The target hindcast period is 30 years (typically 1981-2010).

- The ensemble size is left open to the forecast provider, but larger ensembles are considered better.

- Data distributed must include each ensemble member (not the ensemble mean). Total fields are required (i.e., systematic error corrections to be coordinated by MME combination lead, NOAA/CPC). Forecast providers are welcome to also provide bias-corrected forecasts and to develop their own MME combinations.

- Model configurations – resolution, version, physical parameterizations, initialization strategies, and ensemble generation strategies – are left open to forecast providers.

- Required output is monthly means of global grids of SST, T2m, and precipitation rate. More fields will be added based on experience and demand. It is also recognized that higher frequency data are desirable and this will be implemented as feasible.

- Routine real-time forecast data must be available by the 8th of each month.