

## A MULTISCALAR GLOBAL DROUGHT DATASET: THE SPEIBASE

### A New Gridded Product for the Analysis of Drought Variability and Impacts

BY SANTIAGO BEGUERÍA, SERGIO M. VICENTE-SERRANO, AND MARTA ANGULO-MARTÍNEZ

**D**rought is a period of deficient precipitation with impacts on agriculture, water resources, and the natural ecosystems. It is the natural hazard that affects more people with the most negative consequences in the world, being responsible for extreme economic loss, famine, epidemics, and land degradation. In many developing countries, drought increases structural problems, causing unemployment, impoverishment, decreases in crop yields, and even forced migrations. Thus, improving our knowledge about the spatial and temporal variability of drought is a fundamental prerequisite to quantify the drought hazard and vulnerability of different systems and regions, with the final purpose of improving drought mitigation and preparedness. The availability of data for characterizing drought conditions over given regions and time periods is therefore in high demand by scientists and managers at different levels. Such datasets must be operative both for analyzing past droughts (their onset and end, magnitude, and spatial extent) and for determining drought probabilities and vulnerabilities across a wide range of systems.

A critical issue in the study of drought vulnerability is the multiscale nature of drought, since the response of the hydrological (soil moisture, groundwater, river discharge, reservoir storage, etc.) and biological (crops, natural vegetation, etc) systems to water shortage varies markedly and with different

response times. This explains why severe drought conditions can be recorded in one system (e.g., low river flows), whereas other systems in the same region (e.g., crops) have normal or even humid conditions. Thus, the time scale over which the water deficit accumulates becomes extremely important, and functionally separates hydrological, environmental, agricultural, and other types of drought.

Recently, a new drought indicator, the Standardised Precipitation-Evapotranspiration Index (SPEI), developed by Vicente-Serrano et al. in a 2010 *Journal of Climate* article, has been proposed to quantify the drought condition over a given area. The SPEI considers not only precipitation but also evapotranspiration (PET) data in its calculation, allowing for a more complete approach to explore the effects of climate change on drought conditions. The SPEI can be calculated at several time scales to adapt to the characteristic times of response to drought of target natural and economic systems, determining their resistance to drought.

Following the development of SPEI, a new global dataset, the SPEIbase, has been made available to the scientific community. The dataset covers the period 1901–2006 with a monthly frequency, and offers global coverage at a 0.5-degree resolution. The dataset consists of the monthly values of the SPEI at time scales from 1 to 48 months. This article describes the SPEIbase and shows some of its potential uses.

A complete description of the data and metadata, and links to download the files, are provided at <http://sac.csic.es/spei>.

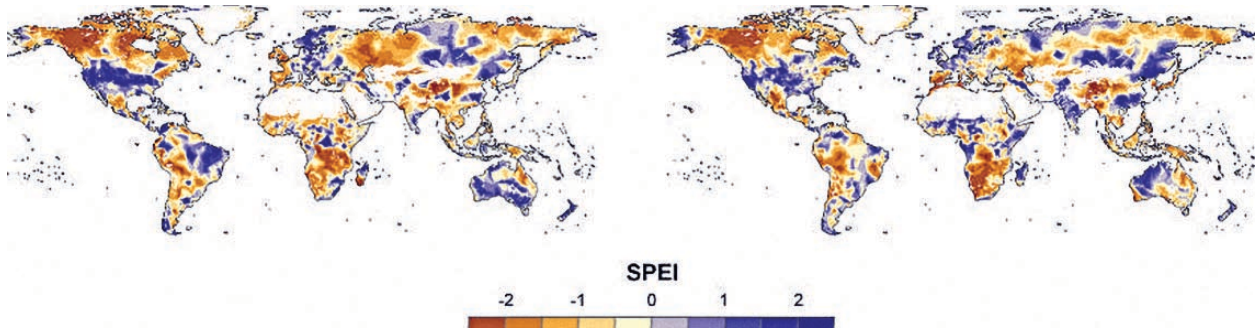
#### THE STANDARDISED PRECIPITATION-EVAPOTRANSPIRATION INDEX (SPEI).

The SPEI is based on a monthly climatic water balance (precipitation minus PET), and it is expressed as a standardized Gaussian variate with a mean of zero and a standard deviation of one. The main advantage of the SPEI over previous multiscale drought indicators

**AFFILIATIONS:** BEGUERÍA AND ANGULO-MARTÍNEZ—Estación Experimental de Aula Dei, CSIC, Zaragoza, Spain; VICENTE-SERRANO—Instituto Pirenaico de Ecología, CSIC, Zaragoza, Spain  
**CORRESPONDING AUTHOR:** Santiago Beguería, Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas, Avda. Montañana 1005, P.O. Box 13034, 50059–Zaragoza, Spain  
E-mail: [sbegueria@eead.csic.es](mailto:sbegueria@eead.csic.es)

DOI:10.1175/2010BAMS2988.1

©2010 American Meteorological Society



**FIG. 1. Global (left) 3-month and (right) 12-month SPEI, Jun 1995.**

like the Standardised Precipitation Index (SPI) is that it includes the effect of the evaporative demand on its calculation, and hence it is suited to explore the effects of global warming on the occurrence of droughts. But, unlike other water balance-based drought indices such as the Palmer Drought Severity Index (PDSI), the SPEI does not rely on the water balance of a specific system (the soil system). The so-called climatic balance used in the SPEI is a theoretical construction in the sense that it does not represent the real moisture deficit experienced by a particular system (a soil column, a river, an ecosystem, etc). To overcome this, the SPEI—like the SPI— can be calculated for different time scales, letting the user find the time scale at which the response to drought of the system under study is highest. Hence the SPEI has a much wider range of applications than the PDSI, since it can be calculated at different time scales.

**DATASET OVERVIEW.** For developing the SPEIbase dataset, the Global 0.5° gridded Climate Research Unit (CRU) TS3 monthly precipitation and mean temperature data were used, since these were the most complete and up-to-date gridded data available with global spatial coverage, as shown by Mitchell and Jones in a 2005 *International Journal of Climatology* article. The Thornthwaite method was used for calculating the PET, and the Kolmogorov-Smirnov test was used to determine the fit between the SPEI data and the log-logistic distribution, showing a very good fit at all time scales across the world (as shown in the 2010 *Journal of Hydrometeorology* article by Vicente-Serrano et al.).

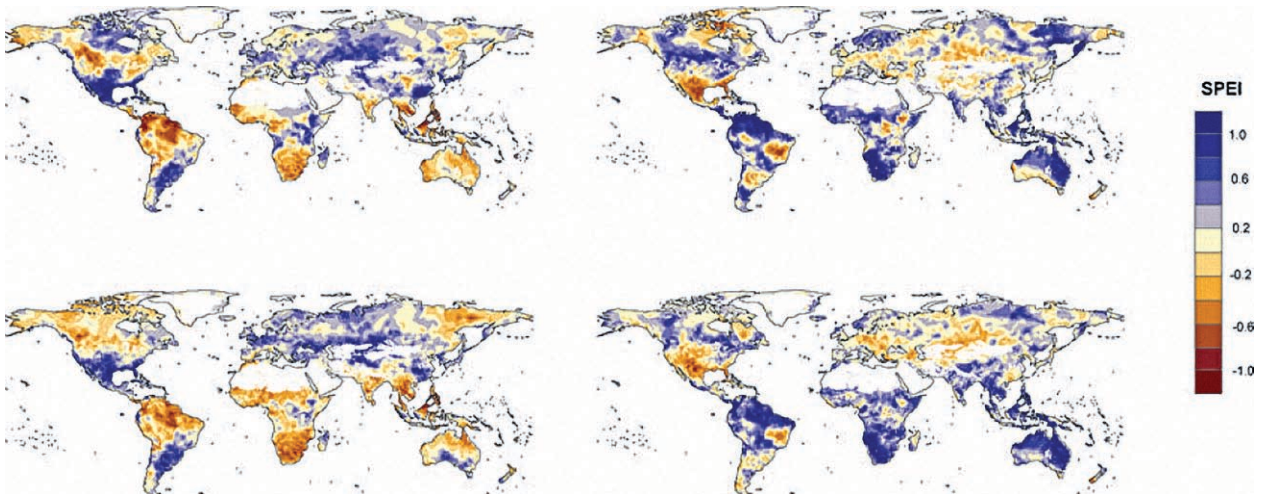
Typical values of the SPEI range between  $-2.5$  and  $2.5$ , corresponding to exceedance probabilities of approximately 0.006 and 0.994, respectively, although the theoretical limits are  $(-\infty, \infty)$ . As an example, the SPEI for June 1995 is shown at two time scales

(Fig. 1). The 3-month SPEI can be considered a short-term drought index. A different spatial pattern was depicted by the 12-month SPEI, indicative of the different scales of the water deficit in different regions of the world. However, very similar SPEI values were recorded in some areas, such as the United States or East China, which showed persistent drought conditions over a long period of time. It is interesting to notice the differences between the long- and short-term SPEI values, as they reveal the trend of the drought conditions over a given region. Noteworthy examples are the expansion of the drought conditions over the United Kingdom or the alleviation of the drought in Southeast Australia.

The SPEIbase has been made available to the public under the Open Database License (ODbL) and the Database Contents License, allowing free right to use, share, and build upon the database under attribution and share-alike conditions. The dataset is served in three alternative formats to facilitate access to different users: netCDF, raw binary, and plain text.

**POTENTIAL APPLICATIONS.** Thanks to its multiscale properties, the potential applications of the SPEIbase are very varied and include such fields as climatology, hydrology, ecology, agronomy, environment, and probably others. Here we show some examples of application in climate studies and ecology:

*Climatology: Impact of El Niño–Southern Oscillation (ENSO) events at the global scale.* The impact of the ENSO on the moisture conditions at the global scale can be assessed using the SPEIbase, for example by computing average SPEI conditions after warm (El Niño) and cool (La Niña) ENSO events (Fig. 2). The multiscale characteristic of the SPEIbase dataset allows for assessing the time scale over which the effects of the ENSO prevail. The example provided in

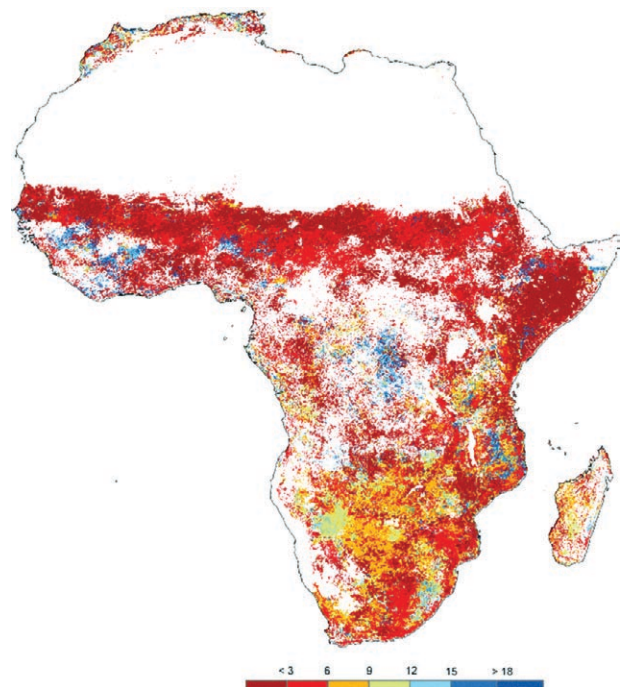


**FIG. 2. Global impact of El Niño/Southern Oscillation (ENSO) on the SPEI: average 3-months Feb SPEI during ENSO's warm years (El Niño, top left) and cool years (La Niña, top right), and average 9-months Aug SPEI during the same years (bottom left and bottom right).**

Fig. 2 shows that the effects of both warm and cool ENSO events on medium-term (9-month) drought conditions are equally well distinguished after the ENSO event.

*Ecology: African biomes' resistance to drought.* The resistance to drought of a given ecosystem can be defined as the time span over which a drought must persist before the ecosystem shows a negative effect. The state of the vegetation over broad regions can be assessed, for example, by the normalized difference vegetation index (NDVI), collected by sensors on board Earth-observation satellites. Thus, it is possible to assess the resistance to drought of an ecosystem by finding the time scale of the SPEI that shows the highest correlation with anomalies of the NDVI. Figure 3 shows the result of performing such an analysis over Africa. The spatial resolution of the figure is  $8 \times 8$  km, corresponding to that of the Global Inventory Modeling and Mapping Studies (GIMMS) NDVI dataset used, which is the most complete and longest remote-sensing dataset with global coverage. For each grid cell in the map, the time series of September–November NDVI and the SPEI at all time scales were obtained, and a correlation analysis was performed. The time scale of the SPEI at which the correlation between the NDVI and the SPEI was highest is shown in the figure. Vegetation of the semiarid areas in the Sahel and the Horn of Africa (Ethiopia, Somalia) showed the lowest resistance to drought, as reflected by the very short time scales (less than six months)

of their response to drought. The tree savannas and subtropical forests, at the other extreme, showed the highest resistance to drought, since the NDVI anomalies had the highest correlation with long-term



**FIG. 3. Time scale (months) at which the SPEI shows the highest correlation with the Sep–Nov normalized difference vegetation index (NDVI) negative anomalies in Africa. Areas without vegetation were omitted from the analysis and are represented here in white.**

SPEI negative anomalies (higher than 15 months in most cases). Other biomes, such as the grass savannas and dry steppes, showed an intermediate response to drought, with response times in the range between 6 and 12 months.

It should be remembered that the SPEI is a standardized variate over the conditions that are normal at a given site, which allows comparing droughts across sites with very different climatology; that is, in absolute terms (mm of rain) a  $-1$  drought in the Sahel region will be very different than a  $-1$  drought in a tropical forest region, but both situations are comparable because they represent the same degree of deviation from the normal conditions at each site, to which the natural vegetation of the area are adapted.

**SUMMARY.** A new multiscalar drought dataset, the SPEIbase, made available to the scientific and management community is presented. The multiscalar character of the dataset allows finding the most appropriate time scale of interest for a specific application, so potential uses of the dataset range over a wide variety of disciplines. The dataset (in three file formats) is distributed under an open license. Further work includes updating the dataset past December 2006, depending on the data availability.

**ACKNOWLEDGMENTS.** We would like to thank the Climate Research Unit of the University of East Anglia (UK) for providing the data used to obtain the SPEIbase dataset. This work has been supported by the research projects CGL2008-01189/BTE and CGL2006-11619/HID, funded by the Spanish Commission of Science and Technology and FEDER; EUROGEOS (FP7-ENV-2008-1-226487) and ACQWA (FP7-ENV-2007-1- 212250), funded by the VII Framework Programme of the European Commission; and “Las sequías climáticas en la cuenca del Ebro y su respuesta hidrológica” and “La nieve en el Pirineo aragonés: Distribución espacial y su respuesta a las condiciones climática,” funded by “Obra Social La Caixa” and the Aragón Government.

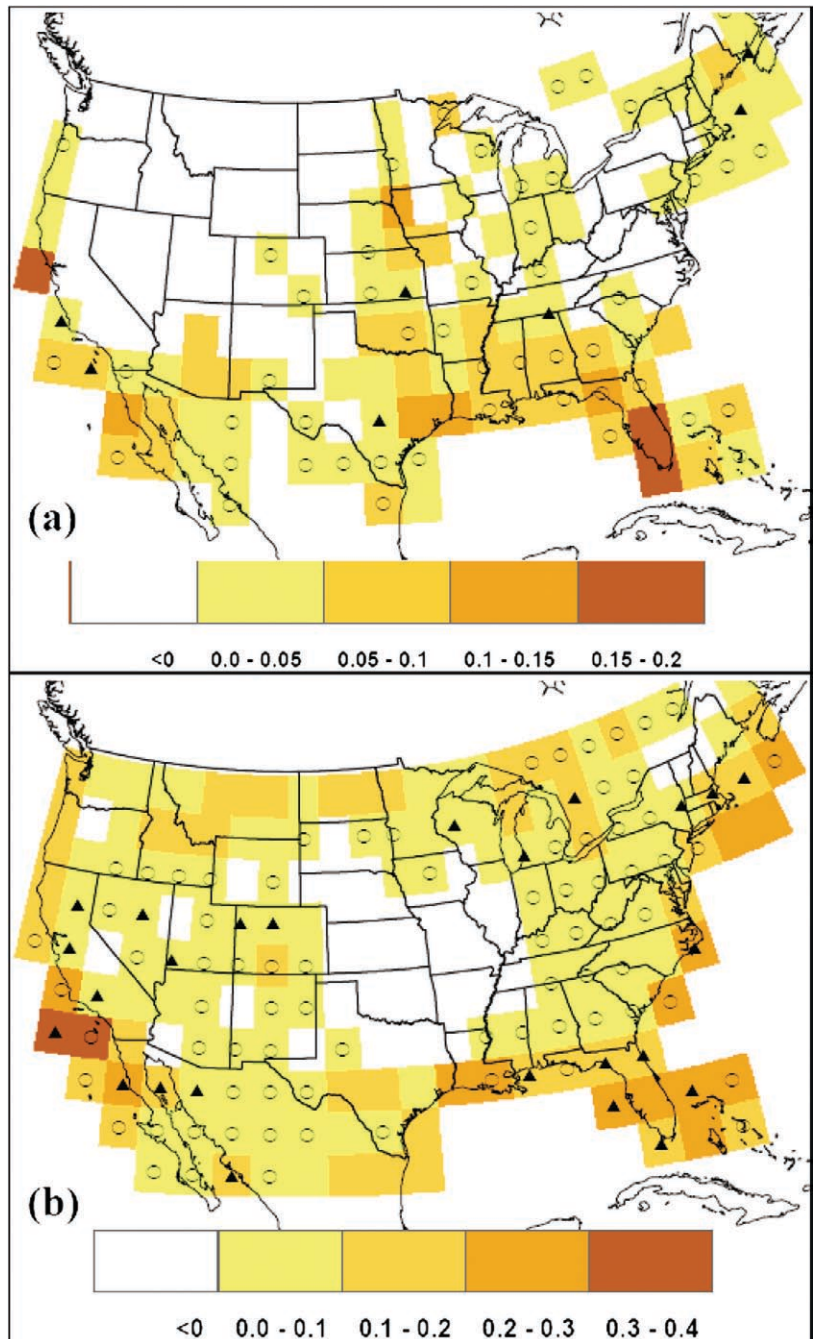
## FOR FURTHER READING

- Dai, A., K. E. Trenberth, and T. Qian, 2004: A global data set of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming. *J. Hydrometeor.*, **5**, 1117–1130.
- Guttman, N. B., 1998: Comparing the Palmer drought index and the Standardized Precipitation Index. *J. Am. Water Resour. Assoc.*, **34**, 113–121.
- Hayes, M., D. A. Wilhite, M. Svoboda, and O. Vanyarkho, 1999: Monitoring the 1996 drought using the Standardized Precipitation Index. *Bull. Amer. Meteor. Soc.*, **80**, 429–438.
- Keyantash, J., and J. Dracup., 2002: The quantification of drought: An evaluation of drought indices. *Bull. Amer. Meteor. Soc.*, **83**, 1167–1180.
- McKee, T. B. N., J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. *Preprints, Eighth Conf. On Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc. 179–184.
- Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.*, **25**, 693–712.
- van der Schrier, G., K. R. Briffa, T. J. Osborn, and E. R. Cook, 2006: Summer moisture availability across North America. *J. Geophys. Res.*, **111**, D11102, doi:10.1029/2005JD006745.
- Vicente-Serrano S. M., S. Beguería, and J. I. López-Moreno, 2010: A multi-scalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index–SPEI. *J. Climate*, **23**, 1696–1718.
- , —, —, M. Angulo, and A. El Kenawy, 2010: A new global 0.5° gridded dataset (1901–2006) of a multiscalar drought index: Comparison with current drought index datasets based on the Palmer Drought Severity Index. *J. Hydrometeor.*, **11**, 1033–1043.
- Wells, N., S. Goddard, and M. J. Hayes, 2004: A self-calibrating Palmer Drought Severity Index. *J. Climate*, **17**, 2335–2351.

### IMPROVING U.S. WINTER FORECASTS USING MULTIMODEL COMBINATIONS AND ENSO

Seasonal climate predictions provide useful information toward short-term planning and management of water supplies and energy consumption in the United States. These forecasts, primarily of precipitation and temperature, are typically issued from coupled General Circulation Models (GCMs) or from Atmospheric General Circulation Models (AGCMs) forced with Sea Surface Temperature (SST) forecasts. However, the skill of climate forecasts could vary substantially depending on the season, the region, and the GCMs themselves. Reducing uncertainties in these forecasts by improving the GCMs with better parameterizations and process representations is time consuming. This issue led recent efforts to focus on an alternate approach of combining multiple GCMs to improve seasonal climatic prediction.

Developing multimodel forecasts is typically performed either by simple pooling of all the ensembles from individual GCMs or by optimal combination under which weights are obtained for each GCM by maximizing the skill metric of the multimodel forecasts. Recent studies on optimal model combination clearly show that multimodel forecasts provide improved reliability (i.e., better correspondence between probability forecasts and their observed relative frequencies) and resolution than individual model predictions. Additionally, our assessment of winter climate forecasts over the continental United States also reveals that the skill of these GCMs varies substantially during El Niño–Southern Oscilla-



**Performance of multimodels and individual models, expressed as average rank probability skill score (RPSS), in predicting DJF winter (a) precipitation and (b) temperature. Grid points having RPSS less than zero are shown in white, which indicates forecast skill that is less than climatology. Grid points with no markers (individual GCMs), open circles (multimodels proposed in this study), and triangles (existing multimodel techniques) indicate the best-performing model having the highest RPSS at that grid point. A positive score for RPSS indicates that the forecast skill exceeds that of the climatological probabilities, with higher scores in progressively darker shading.**

tion (ENSO) conditions. Hence, a new approach to combine multiple GCMs along with climatology is proposed that gives different weights for each GCM contingent on ENSO conditions.

To develop this methodology, we combined historical simulations of winter precipitation and temperature from seven AGCMs by evaluating their skill—represented by Mean Square Error (MSE)—over similar ENSO (represented by Nino3.4) conditions. The MSE estimates are converted into weights for each GCM for developing multimodel tercile probabilities. We considered a total of six multimodel schemes, which included combinations based on pooling of ensembles as well as the long-term skill of the models. To ensure the improved skill exhibited by the multimodel scheme is statistically significant, we also performed rigorous hypothesis tests comparing the skill of multimodels with the individual models' skill.

The proposed multimodel combination approach contingent on Nino3.4 shows improved skill, particularly for regions whose winter precipitation and temperature exhibit significant correlation with Nino3.4. Analyses of weights also show that the proposed multimodel combination methodology assigns higher weights for GCMs and lesser weights for climatology during El Niño and La Niña conditions. On the other hand, due to the limited skill of GCMs during neutral conditions over the tropical Pacific, the methodology assigns higher weights for climatology resulting in improved skill from the multimodel combinations. Thus, analyzing GCMs' skill contingent on the relevant predictor state provides an alternate approach for multimodel combinations such that years with limited skill could be replaced with climatology.

It is important to note that this study has employed historical simulations of precipitation

**ECHOES**

**“ This is the worst I have seen in the almost 17 years I’ve been here.”**

—KEN POWELL, environmental scientist for the Kansas Dept. of Health and Environment, on the death of more than 2,600 cattle in the state over three days in mid-July due to extreme heat and humidity and a lack of wind. Temperatures in the state reached as high as 108°F, in Hill City, and the intense conditions forced ranchers and feedlot workers to labor around the clock to increase the cattle's water consumption, spray them with water, and enhance their straw beds to provide a buffer to the warm ground. Many cattle still succumbed despite these efforts. “There aren't many management strategies to use when you're dealing with those kind of extreme conditions,” noted Todd Domer of the Kansas Livestock Association.

(SOURCE: Reuters)

and temperature from AGCMs to demonstrate the utility of the proposed multimodel combination algorithm. To apply the proposed methodology in a forecasting context, one may have to use the forecast Nino3.4 from multiple coupled GCMs (CGCMs) as the conditioning variable. We are currently working on combining multiple CGCMs contingent on forecast Nino3.4 conditions, so that the proposed methodology could be employed in issuing real-time climate forecasts.—NARESH DEVINENI (NORTH CAROLINA STATE UNIVERSITY) AND A. SANKARASUBRAMANIAN. “*Improving the Prediction of Winter Precipitation and Temperature over the Continental United States: Role of ENSO State in Developing Multimodel Combinations*,” in the June Monthly Weather Review.