Forecasting Drought Using the Agricultural Reference Index for Drought (ARID): A Case Study

PREM WOLI
Mississippi State University, Mississippi State, Mississippi

JAMES JONES AND KEITH INGRAM
University of Florida, Gainesville, Florida

JOEL PAZ
Mississippi State University, Mississippi State, Mississippi

(Manuscript received 25 April 2012, in final form 19 October 2012)

ABSTRACT

Drought forecasting can aid in developing mitigation strategies and minimizing economic losses. Drought may be forecast using a drought index, which is an indicator of drought. The agricultural reference index for drought (ARID) was used as a tool to investigate the possibility of using climate indices (CIs) as predictors to improve the current level of forecasting, which is El Niño–Southern Oscillation (ENSO) based. The performances of models that are based on linear regression (LR), artificial neural networks (ANN), adaptive neuron-fuzzy inference systems (ANFIS), and autoregressive moving averages (ARMA) models were compared with that of the ENSO approach. Monthly values of ARID spanning 56 yr were computed for five locations in the southeastern United States, and monthly values of the CIs having significant connections with weather in this region were obtained. For the ENSO approach, the ARID values were separated into three ENSO phases and averaged by phase. For the ARMA models, monthly time series of ARID were used. For the ANFIS, ANN, and LR models, ARID was predicted 1, 2, and 3 months ahead using the past values of the first principal component of the CIs. Model performances were assessed with the Nash–Sutcliffe index. Results indicated that drought forecasting could be improved for the southern part of the region using ANN models and CIs. The ANN outperformed the other models for most locations in the region. The CI-based models and the ENSO approach performed better during the winter, whereas the efficiency of ARMA models depended on precipitation periodicities. All models performed better for southern locations. The CIs showed good potential for use in forecasting drought, especially for southern locations in the winter.

1. Introduction

Drought is a creeping natural hazard, causing tremendous economic and social losses worldwide. In the United States, about $7 billion is lost every year through drought (FEMA 1995; NSTC 2005). Although drought cannot be prevented, its losses can be minimized by mitigation strategies if it is forecast well in advance. Because it evolves slowly, its mitigation plans can be carried out more effectively than those of other natural hazards. The benefits of reliable forecasting are tremendous (Campbell 1973).

Drought may be forecast using drought variables (Hastenrath and Greischar 1993; Chiew et al. 1998; Cordery and McCall 2000) or drought indicators (Nicholls 1985; Mishra and Desai 2005; Cancelliere et al. 2007; Vasiliades and Loukas 2008). A drought variable is a key element that defines drought, such as precipitation, soil moisture, or evapotranspiration. A drought indicator is a measure used to quantify and characterize drought conditions. For forecasting a meteorological drought, forecasting precipitation may suffice. For forecasting an agricultural drought (the shortage of water in the root zone to meet the consumptive use of plants), however, forecasting all associated variables is required. This is difficult, because
these variables are controlled by several plant and soil processes. Therefore, for agricultural drought, the use of a drought indicator, such as a drought index, is more appropriate than using a drought variable.

Various methods exist for forecasting a drought index (Mishra and Singh 2011), one of which is to use its own past values (Panu and Sharma 2002; Mishra and Desai 2005; Mendicino et al. 2008). This approach involves quantifying a relationship between the present and past values of the index by fitting the data to a time series model, such as an autoregressive moving average (ARMA; Box and Jenkins 1970). Another approach is to predict the index as a function of the past values of its predictor variables. Teleconnection indices may be used as drought predictors (Mishra and Singh 2010, 2011) because atmospheric circulation patterns have been found to influence droughts (Stahl and Demuth 1999). The linkage between changes in the atmosphere in one place and effects on the weather in another, distant place is called teleconnection. The indices of large-scale oceanic and atmospheric teleconnection patterns, or oscillations, are called teleconnection or climate indices. Although climate indices are generally calculated on the basis of measurements from a localized area, they can be statistically related with climate variables of other areas around the globe (Bond et al. 2007).

Because teleconnection patterns affect weather variables such as temperature and precipitation, the key inputs of a drought index, a plausible relationship is likely to exist between a drought index and climate indices. Some of the major teleconnection patterns that are reported to have significant impacts on the weather of the southeastern United States are the Atlantic multidecadal oscillation (AMO; Gershunov and Barnett 1998b; Enfield et al. 2001; Stevens 2007), the North Atlantic Oscillation (NAO; Rogers 1984; Yin 1994; Hagemeyer 2006; Stevens 2007), the Pacific decadal oscillation (PDO; Gershunov and Barnett 1998b; Enfield et al. 2001; Stevens 2007), the Pacific–North American (PNA) pattern (Leathers et al. 1991; Yin 1994; Konrad 1998; Hagemeyer 2006; Martinez et al. 2009), sea surface temperature (SST) in the Niño-3.4 region (i.e., the east-central tropical Pacific: 5°N–5°S, 170°–120°W) (Trenberth 1997; Gershunov and Barnett 1998b; Schmidt et al. 2001; Hanley et al. 2003; Hagemeyer 2006), and the SST over the region of 4°S–4°N, 150°–90°W (Hansen et al. 1998).

The AMO index is basically an index of ongoing series of long-duration changes in the SST of the North Atlantic Ocean over 0°–70°N. The NAO index is the difference in sea level pressure between two stations situated close to the Icelandic low and Azores high pressure centers. The PNA index is a linear combination of the normalized geopotential height anomalies at the 700-hPa level at four locations: Hawaii (20°N, 160°W); the North Pacific Ocean (45°N, 165°W); Alberta, Canada (55°N, 115°W); and the Gulf Coast of the United States (30°N, 85°W). The PDO index is the spatial average of the monthly SST of the Pacific Ocean north of 20°N. The Niño-3.4 index is the departure in monthly SST from their long-term means over the Niño-3.4 region. The Japan Meteorological Agency (JMA) index is a 5-month running mean of spatially averaged SST anomalies over the tropical Pacific region (4°S–4°N, 150°–90°W).

The most commonly used approach for forecasting precipitation or drought in the southeastern United States is based on the El Niño–Southern Oscillation (ENSO). The strong teleconnection between ENSO and the weather conditions in this region has enabled skillful forecasting of seasonal temperature and precipitation up to 1 yr in advance (Steinemann 2006; Brolley et al. 2007). The drought-forecasting ability of the ENSO approach, however, varies across months and locations because the impact of ENSO varies across regions and seasons (Ropelewski and Halpert 1986; Gershunov and Barnett 1998a; Smith et al. 1998; Brolley et al. 2007). That is, the ENSO-based approach might produce plausible forecasts for some months of the year and inaccurate forecasts for other months. This situation necessitates the search for other approaches. The study presented here explored whether climate indices could produce more accurate forecasts than those produced using ENSO for the months or locations for which the predictability of ENSO is not good.

For quantifying predictor–predictand relationships, various statistical models have been used, such as linear regression (LR; Kumar and Panu 1997; Panu and Sharma 2002; Wood and Lettenmaier 2006; Vasiliades and Loukas 2008), artificial neural networks (ANN; Wilpen et al. 1994; Zhang et al. 1998; Panu and Sharma 2002; Vasiliades and Loukas 2008), and the adaptive neuro-fuzzy inference system (ANFIS; Nayak et al. 2004; Chang and Chang 2006; Bacanli et al. 2008; Firat and Gungor 2008). An ARMA model can model the periodicity and nonstationarity of time series data effectively (Modarres 2006; Mishra et al. 2007) and can provide systematic procedures for modeling the behavior of uncertain systems (Mishra and Desai 2005). An LR model can explicitly address causal relationships and can show optimal results for a linear mechanism (Belsley et al. 1980). An ANN model is able to map nonlinear, complex, and arbitrary input–output relationships (Hornik et al. 1989; Smith 1996; Zhang et al. 1998; Kim and Valdes 2003; Samarasinghe 2007). An ANFIS (Zadeh 1965; Jang 1993) is a promising tool for modeling complex, nonlinear systems (Marce et al. 2004; Nayak et al. 2004; Bacanli et al. 2008; Altun et al. 2009) and has
the ability to draw conclusions from vague, ambiguous, incomplete, and imprecise information (Lee et al. 2007). Several researchers have used these models, individually or in combination, to forecast various hydrological variables. Some researchers, such as Cutore et al. (2009), have used ANN to forecast a hydrological drought, which is an expression of precipitation shortfall on surface or subsurface water supply. For forecasting an agricultural drought, however, no study has assessed all of these methods using a number of climate indices mentioned above. The current study examined the performance of ANFIS, ANN, ARMA, and LR models relative to that of the ENSO approach in forecasting the agricultural reference index for drought (ARID; Woli et al. 2012), which is a computationally simple, physically and physiologically sound, and generally applicable drought index that can characterize an agricultural drought better than many other drought indices that are applied to agricultural systems (Woli et al. 2012). Although ARID is a simple and generic drought index, it is applicable to a wide range of conditions (Woli et al. 2013).

The ARID is based on the reference grass (Allen et al. 2005) and is expressed as the ratio of plant water deficit to plant water need as

\[
\text{ARID}_i = 1 - \left( \frac{T_i}{ET_{o,i}} \right),
\]

where the subscript \(i\) stands for the \(i\)th day, \(T_i\) is transpiration (mm day\(^{-1}\)), and \(ET_{o,i}\) is the reference grass evapotranspiration (mm day\(^{-1}\)). Values of ARID fall between 0, indicating no water stress, and 1, indicating full water stress. Whereas \(ET_{o,i}\) is computed using the United Nations Food and Agriculture Organization Irrigation and Drainage Paper 56 (FAO-56) Penman–Monteith model (Allen et al. 1998), \(T_i\) is estimated as

\[
T_i = \min (\alpha \xi \Theta_{ad,i-1}, ET_{o,i}),
\]

where \(T_i\) is the transpiration in day \(i\) (mm day\(^{-1}\)), \(\alpha\) is the water uptake coefficient, \(\xi\) is the root-zone depth (mm), and \(\Theta_{ad,i-1}\) is plant available water content after deep drainage at the end of the previous day (mm \(^{-1}\)). The available water content in day \(i\) \(\Theta_{ad,i}\) is estimated as the ratio of the amount of available water in the root zone in day \(i\)—\(W_i\) (mm)—to the root-zone depth: \(\Theta_{ad,i} = W_i/\xi\). The \(W_i\) is estimated from a simple soil water balance:

\[
W_i = W_{i-1} + P_i + I_i - T_i - D_i - R_i,
\]

where subscripts \(i\) and \(i-1\) denote the current and previous day, respectively, and \(P_i\), \(I_i\), \(T_i\), \(D_i\), and \(R_i\) are precipitation, irrigation, transpiration, deep drainage, and surface runoff that occurred in day \(i\) (mm day\(^{-1}\)), respectively. The deep drainage for day \(i\) is estimated as

\[
D_i = \beta \xi (\Theta_{ad,i} - \theta_m),
\]

where \(\beta\) is drainage coefficient, \(\Theta_{ad,i}\) is the available soil water content in day \(i\) before drainage (mm mm\(^{-1}\)), and \(\theta_m\) is plant available water capacity (mm mm\(^{-1}\)). Using the USDA-SCS (1972) method, ARID estimates daily surface runoff (mm day\(^{-1}\)) as

\[
R_i = \frac{(P_i - I_i)^2}{(P_i - I_i + S)},
\]

where \(I_i\) is initial abstraction (mm day\(^{-1}\)) and \(S\) is potential maximum retention (mm day\(^{-1}\)), which is computed as \(S = 25 \, 400/\eta\) (where \(\eta\) is the runoff curve number).

2. Materials and methods

2a. Sites, data, and general computations

Of the five forecasting methods assessed—ANFIS, ANN, LR, ARMA, and ENSO—only the first three methods used climate indices as the predictors of ARID. Because ARMA is a time series model, it only uses the past values of the predictand to predict future values. Although ENSO is also teleconnection based (JMA), the ENSO approach does not directly use the values of the JMA index and the other climate indices used by ANFIS, ANN, and LR. The predictors of ARID for the climate index–based methods were chosen, and further computations and analyses were made in the following steps (which are also summarized in Table 1):

1) Through a literature review, six large-scale oceanic–atmospheric oscillations having significant impacts on the weather of the southeastern United States were identified: AMO, NAO, JMA, Niño-3.4, PDO, and PNA.

2) Five locations in the southeastern United States—Miami (25.5°N, 80.5°W), Bartow (27.9°N, 81.8°W), and Live Oak (30.3°N, 82.9°W) in Florida and Plains (32.7°N, 84.2°W) and Blairsville (34.8°N, 83.9°W) in Georgia—were selected by considering latitude and the availability of historical weather data.

3) Daily weather data spanning 56 yr (1951–2006) for the respective locations were obtained on the Internet from the Florida Automated Weather Network and the Georgia Automated Environmental Monitoring Network.

4) The monthly values of the indices of the oscillations mentioned in step 1, called climate indices, for the years mentioned in step 3 were obtained from the Internet.

5) From the weather data mentioned in step 3, daily values of ARID were computed for each of the five
locations, which later were converted to monthly values by averaging the daily values over a month. For instance, a monthly value for January was computed by averaging all 31 daily values of this month. Computations were carried out monthly because most of the climate indices used in the study recorded only monthly values. Moreover, a monthly time scale is appropriate for monitoring drought from the viewpoints of farm management and crop-growth-stage sensitivity to water deficit (Panu and Sharma 2002).

6) For each climate index, the dataset obtained in step 4 was separated by month, thus creating 12 subsets, one for each month. The monthly separation of the dataset was carried out by assuming that the degree of linkage between large-scale teleconnection patterns and the weather in this region might vary across months.

7) To assess the predictability of ARID for different lead times, ARID was forecast with 1–3-month lead time ($t_1$, $t_2$, and $t_3$) using the current (time $t$) values of its predictors. The predictors' lag time remained the same ($t$) while predicting ARID at different lead times. To forecast ARID at three lead times, three series of ARID were created and each series individually was used. As in step 6, 12 subsets of ARID at each lead-time forecast (ARID$_{t+1}$, ARID$_{t+2}$, or ARID$_{t+3}$) were created for each location.

8) The correlation between a climate index and ARID at each lead time was examined for each climate index, month, location, and lead time; if significant at $\alpha = 0.1$, the index was selected (Tables 2 and 3).

9) The climate indices selected for any location for a month were considered to be the predictors of ARID for the entire region for that month.

10) Cross correlations among the CIs were examined (Table 4).

11) Because most of the climate indices were significantly cross correlated at the significance level of 0.05 (Table 4), principal component analysis was performed on the climate indices chosen for each month and lead-time forecast (Table 3), and the first principal component, hereinafter referred to as X, was used as the predictor of ARID. The X was chosen because (i) the total variance it explained was generally reasonably good (Table 3) and (ii) its use made the number of predictors across months and lead-time

### Table 1. Steps involved in selecting the predictors of ARID and making further computations.

<table>
<thead>
<tr>
<th>Step</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Large-scale oscillations were identified for the southeastern United States</td>
</tr>
<tr>
<td>2</td>
<td>Five locations in the region were selected</td>
</tr>
<tr>
<td>3</td>
<td>Historical weather data spanning 1951–2006 were obtained</td>
</tr>
<tr>
<td>4</td>
<td>Monthly values of the CIs were obtained</td>
</tr>
<tr>
<td>5</td>
<td>Monthly values of ARID were computed for each location</td>
</tr>
<tr>
<td>6</td>
<td>For each CI, the dataset was separated by month</td>
</tr>
<tr>
<td>7</td>
<td>For each location and lead time, the ARID dataset was separated by month</td>
</tr>
<tr>
<td>8</td>
<td>CI–ARID correlation was examined for each index, month, location, and lead time; if significant at $\alpha = 0.1$, the index was selected (Tables 2 and 3)</td>
</tr>
<tr>
<td>9</td>
<td>The CIs selected for any location for a month were considered to be the predictors of ARID for the entire region for that month</td>
</tr>
<tr>
<td>10</td>
<td>Cross correlations among the CIs were examined (Table 4)</td>
</tr>
<tr>
<td>11</td>
<td>PCA was performed on the chosen CIs (Table 3), and the first principal component was used as the predictor of ARID</td>
</tr>
<tr>
<td>12</td>
<td>Models were evaluated using the Nash–Sutcliffe index, and the predictability of ARID was assessed using the RMSE</td>
</tr>
</tbody>
</table>
forecasts consistent (i.e., 1) because the number of climate indices was different for different months and lead times, ranging from 1 to 4 (Table 3).

12) Last, the predictive accuracy of each model was evaluated using the modeling efficiency [ME, also called the Nash–Sutcliffe index (Nash and Sutcliffe 1970)], which is a commonly used measure for evaluating hydrometeorological models (Sevat and Dezetter 1991; ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, Irrigation and Drainage Division 1993; Legates and McCabe 1999; Moriasi et al. 2007). The ME was computed as

\[
ME = 1 - \left[ \frac{1}{n} \sum_{i=1}^{n} (\text{ARID} - \bar{\text{ARID}})^2 \right] / \left[ \frac{1}{n} \sum_{i=1}^{n} (\text{ARID} - \bar{\text{ARID}})^2 \right],
\]

where ARID, \(\bar{\text{ARID}}\), and \(\bar{\text{ARID}}\) are the observed (i.e., originally computed), predicted, and mean observed values of ARID, respectively, and \(n\) is the number of observations. Values of ME range from \(-\infty\) to 1. Whereas a negative value indicates that the observed mean is a better predictor than

### Table 2. CIs selected as predictors of ARID at lead-time \(t + 1\) (1-month-ahead forecasting) for various months and locations in the southeastern United States. The CIs were significantly correlated with ARID at the 90% confidence level (\(\alpha = 0.1\)).

<table>
<thead>
<tr>
<th>Month</th>
<th>Miami</th>
<th>Bartow</th>
<th>Live Oak</th>
<th>Plains</th>
<th>Blairsville</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>JMA, PDO, PNA</td>
<td>—</td>
<td>Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Feb</td>
<td>Niño-3.4, PDO, PNA</td>
<td>Niño-3.4, PDO</td>
<td>Niño-3.4, PDO, PNA</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Mar</td>
<td>AMO, Niño-3.4, PDO</td>
<td>Niño-3.4</td>
<td>—</td>
<td>NAO, JMA</td>
<td>NAO, Niño-3.4, PDO</td>
<td>NAO, Niño-3.4, PNA</td>
</tr>
<tr>
<td>Apr</td>
<td>JMA, PDO</td>
<td>JMA</td>
<td>PDO</td>
<td>—</td>
<td>—</td>
<td>JMA, PDO</td>
</tr>
<tr>
<td>May</td>
<td>—</td>
<td>—</td>
<td>PDO</td>
<td>—</td>
<td>—</td>
<td>JMA</td>
</tr>
<tr>
<td>Jun</td>
<td>AMO, PNA</td>
<td>—</td>
<td>PDO</td>
<td>—</td>
<td>PDO</td>
<td>Niño-3.4, PDO, AMO, Niño-3.4, PNA</td>
</tr>
<tr>
<td>Jul</td>
<td>PNA</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>JMA</td>
<td>JMA, PNA</td>
</tr>
<tr>
<td>Aug</td>
<td>AMO</td>
<td>NAO, Niño-3.4</td>
<td>PDO</td>
<td>NAO</td>
<td>—</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Sep</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>—</td>
<td>NAO</td>
<td>Niño-3.4</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Oct</td>
<td>Niño-3.4</td>
<td>JMA</td>
<td>Niño-3.4, PDO</td>
<td>NAO, Niño-3.4, PDO</td>
<td>NAO</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Nov</td>
<td>Niño-3.4, PDO</td>
<td>Niño-3.4, PDO, PNA</td>
<td>AMO, Niño-3.4, PDO</td>
<td>NAO, Niño-3.4, PDO</td>
<td>—</td>
<td>NAO, Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Dec</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Niño-3.4</td>
</tr>
</tbody>
</table>

### Table 3. CIs used as the predictor of ARID for various months and lead-time forecast \((t + 1, t + 2,\) and \(t + 3\)) for the southeastern United States. Here, “\(X\) var” is the percent of total variance explained by the first principal component \(X\).

<table>
<thead>
<tr>
<th>Month ((t))</th>
<th>(t + 1) forecasting</th>
<th>(t + 2) forecasting</th>
<th>(t + 3) forecasting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb</td>
<td>Niño-3.4, PDO, PNA</td>
<td>Niño-3.4, PDO, NAO</td>
<td>Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Mar</td>
<td>NAO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Apr</td>
<td>JMA, PDO</td>
<td>JMA, PDO</td>
<td>JMA, PDO</td>
</tr>
<tr>
<td>May</td>
<td>JMA</td>
<td>NAO, Niño-3.4</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Jun</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Jul</td>
<td>JMA, PDO</td>
<td>NAO, Niño-3.4, PDO</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Aug</td>
<td>NAO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Sep</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4, PDO</td>
</tr>
<tr>
<td>Oct</td>
<td>NAO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Nov</td>
<td>NAO, Niño-3.4, PDO, PNA</td>
<td>NAO, Niño-3.4</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Dec</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
</tr>
</tbody>
</table>

### Table 4. CIs selected as predictors of ARID at lead-time \(t + 1\) (1-month-ahead forecasting) for various months and locations in the southeastern United States. The CIs were significantly correlated with ARID at the 90% confidence level (\(\alpha = 0.1\)).

<table>
<thead>
<tr>
<th>Month</th>
<th>Miami</th>
<th>Bartow</th>
<th>Live Oak</th>
<th>Plains</th>
<th>Blairsville</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>JMA, PDO, PNA</td>
<td>—</td>
<td>Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Feb</td>
<td>Niño-3.4, PDO, PNA</td>
<td>Niño-3.4, PDO</td>
<td>Niño-3.4, PDO, PNA</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>Niño-3.4, PDO, PNA</td>
</tr>
<tr>
<td>Mar</td>
<td>AMO, Niño-3.4, PDO</td>
<td>Niño-3.4</td>
<td>—</td>
<td>NAO, JMA</td>
<td>NAO, Niño-3.4, PDO</td>
<td>NAO, Niño-3.4, PNA</td>
</tr>
<tr>
<td>Apr</td>
<td>JMA, PDO</td>
<td>JMA</td>
<td>PDO</td>
<td>—</td>
<td>—</td>
<td>JMA, PDO</td>
</tr>
<tr>
<td>May</td>
<td>—</td>
<td>—</td>
<td>PDO</td>
<td>—</td>
<td>—</td>
<td>JMA</td>
</tr>
<tr>
<td>Jun</td>
<td>AMO, PNA</td>
<td>—</td>
<td>PDO</td>
<td>—</td>
<td>PDO</td>
<td>Niño-3.4, PDO, AMO, Niño-3.4, PNA</td>
</tr>
<tr>
<td>Jul</td>
<td>PNA</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>JMA</td>
<td>JMA, PNA</td>
</tr>
<tr>
<td>Aug</td>
<td>AMO</td>
<td>NAO, Niño-3.4</td>
<td>PDO</td>
<td>NAO</td>
<td>—</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Sep</td>
<td>Niño-3.4</td>
<td>Niño-3.4</td>
<td>—</td>
<td>NAO</td>
<td>Niño-3.4</td>
<td>NAO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Oct</td>
<td>NAO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>AMO, Niño-3.4, PDO</td>
<td>—</td>
<td>AMO, Niño-3.4, PDO</td>
</tr>
<tr>
<td>Nov</td>
<td>NAO, Niño-3.4, PDO, PNA</td>
<td>NAO, Niño-3.4</td>
<td>—</td>
<td>NAO, Niño-3.4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Dec</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Niño-3.4</td>
</tr>
</tbody>
</table>
the model, a positive value signifies that using the model estimate is better than using the observed mean. An ME of 0 indicates that the model estimate is as accurate as the mean of the observed data, whereas an ME of 1 corresponds to a perfect match of the modeled values to the observed data. In fact, the closer the ME is to 1, the more accurate is the model.

The predictability of ARID for different lead times was assessed by comparing the 1-, 2-, or 3-month lead-time forecasts of ARID with observed values using the root-mean-square error (RMSE). The comparisons were carried out for each model, month, and location.

b. Computations for ANN models

For ANN, a three-layer multilayer perceptron (MLP) was used (Fig. 1) because it is the most widely used network and provides a general framework for nonlinear input–output mapping (Hornik et al. 1989; Zhang et al. 1998; Samarakkodi 2007; Abbasi 2009).

A single-hidden-layer MLP was chosen because one hidden layer is sufficient to approximate any complex nonlinear function (Cybenko 1989; Hornik et al. 1989; Zhang et al. 1998; Maier and Dandy 2000). To examine how the number of nodes in a hidden layer of the network would affect the forecasting ability of the network and to decide on the number of nodes, we carried out a sensitivity analysis on the number of nodes. The ability of the network was measured in terms of prediction errors (RMSE): the lower the RMSE value is, the better is the forecasting ability. For this, the network was run with different numbers of nodes in a hidden layer ranging from 1 to 10. The results showed that increasing the number of nodes in a hidden layer from 1 did not significantly improve the performance of the network having one input node (Table 5). For simplicity, therefore, we used just one node in the hidden layer. Several other researchers, such as De Groot and Wurtz (1991), Chakraborty et al. (1992), and Tang and Fishwick (1993), are also of the opinion that one node in the hidden layer be used for a network that has just one input node.

![FIG. 1. The architecture of an MLP neural network with one input layer, one hidden layer, and one output layer, each with one neuron.](image-url)
TABLE 5. The RMSE values associated with 1-month-lead-time forecasting of ARID for various months for Miami, using different numbers of neurons in a hidden layer of the neural network having one input node.

<table>
<thead>
<tr>
<th>Month</th>
<th>No. of neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>0.15 0.21 0.20 0.21 0.20 0.22 0.21 0.22 0.20</td>
</tr>
<tr>
<td>Feb</td>
<td>0.16 0.17 0.19 0.17 0.18 0.15 0.19 0.15 0.14 0.17</td>
</tr>
<tr>
<td>Mar</td>
<td>0.11 0.14 0.14 0.12 0.11 0.13 0.13 0.13 0.12 0.14</td>
</tr>
<tr>
<td>Apr</td>
<td>0.10 0.11 0.13 0.10 0.09 0.09 0.09 0.11 0.13 0.10</td>
</tr>
<tr>
<td>May</td>
<td>0.11 0.12 0.13 0.10 0.12 0.14 0.17 0.13 0.11 0.14</td>
</tr>
<tr>
<td>Jun</td>
<td>0.07 0.10 0.07 0.07 0.07 0.11 0.11 0.06 0.12 0.14 0.07</td>
</tr>
<tr>
<td>Jul</td>
<td>0.05 0.07 0.06 0.07 0.08 0.08 0.08 0.09 0.08 0.08</td>
</tr>
<tr>
<td>Aug</td>
<td>0.06 0.05 0.08 0.10 0.07 0.09 0.11 0.08 0.08 0.10</td>
</tr>
<tr>
<td>Sep</td>
<td>0.06 0.07 0.06 0.07 0.06 0.08 0.08 0.05 0.07 0.08 0.07</td>
</tr>
<tr>
<td>Oct</td>
<td>0.12 0.13 0.11 0.10 0.11 0.12 0.13 0.13 0.14 0.11</td>
</tr>
<tr>
<td>Nov</td>
<td>0.17 0.17 0.18 0.19 0.16 0.17 0.15 0.17 0.14 0.18</td>
</tr>
<tr>
<td>Dec</td>
<td>0.16 0.19 0.14 0.18 0.17 0.20 0.16 0.19 0.16 0.19</td>
</tr>
</tbody>
</table>

The MLP network estimated ARID with the following steps (Fig. 1):

1) The network-input \( X \) was fed into the hidden neuron, and its weighted input \( p \) was computed, using its layer weight \( a \) and bias weight \( b \) as

\[
p = aX + b. \tag{7}
\]

2) In the hidden neuron, its output \( h \) was computed using \( p \) in a hyperbolic tangent input transfer function (Klimasauskas 1991; Maier and Dandy 1998; Samarasinghe 2007) as

\[
h = \frac{\exp(p) - \exp(-p)}{\exp(p) + \exp(-p)}. \tag{8}
\]

3) The \( h \) was then fed into the output neuron, and its weighted input \( q \) was computed, using its layer weight \( a \) and bias weight \( b \), as

\[
q = ah + b. \tag{9}
\]

4) In the output neuron, the network output (\( \text{ÁRID} \)) was computed using \( q \) in a linear output transfer function (Rumelhart et al. 1995; Zhang et al. 1998; Sahoo et al. 2005):

\[
\text{ÁRID} = q. \tag{10}
\]

Step 4 concluded forward processing in the network, and, with this, ARID was compared with the target value (ARID) to compute error \( e \), which was then backpropagated through the network to update the parameters \( a, b, \alpha, \) and \( \beta \). For training, the gradient descent method was used because it is efficient and popular (Zhang et al. 1998; ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000; Mishra and Desai 2006; Samarasinghe 2007). Learning rate and momentum were set to 0.01 and 0.9, respectively, assuming that the input–output relationships were complex (Tang et al. 1991; Hagan et al. 1996; Principe et al. 1999; Mishra and Desai 2006; MathWorks 2012). For training, batch learning was used because it is the most preferred method, forces the search in the direction of the true gradient (Maier and Dandy 2000), and avoids instability (Samarasinghe 2007).

Before feeding data into the network, each monthly dataset of 56 input–output combinations was divided into three subsets: 44 for training, 11 for testing, and 1 for validation (Granger 1993). The testing set was used to avoid overfitting the data, and the validation set was used to evaluate the model. The performance of an ANN model was evaluated applying the leave-one-out technique of cross validation. Last, values of RMSE and ME were computed for each month, location, and lead-time forecast using the model-estimated and original values of ARID.

c. Computations for ANFIS models

This study used the Sugeno-type ANFIS (Takagi and Sugeno 1985) because it is computationally efficient and is well suited for modeling nonlinear systems (Jang 1993; Lee et al. 2007; MathWorks 2012). The first-order Sugeno type was used because higher orders often add an unwanted level of complexity (Lee et al. 2007). To partition the input space, grid partitioning was used because it gives more precise modeling (Altun et al. 2009). For training, the hybrid optimization method—the combination of backpropagation and least squares error—was used because of its high learning speed (Jang 1993; Lee et al. 2007). The ANFIS network estimated ARID with the following steps (Fig. 2):

1) In the node of layer 1, the input \( X \) was fuzzified using a generalized bell-shaped (Marja and Esa 1997; Lee et al. 2007; MathWorks 2012) input membership function \( \text{f} \), a curve that maps each element in the input space to a membership value between 0 and 1 as

\[
f_i = \frac{1}{1 + \{|\text{abs}(X-c_i)/\epsilon_i|\}^{2d}} \tag{11}
\]

where the subscript \( i \) stands for the \( i \)th rule and \( c, d, \) and \( \epsilon \) are antecedent parameters that are updated through backpropagation.

2) In the node of layer 2, firing strengths, also called weights \( w \), were computed for rule \( i \). For the single-input variable \( X \), the same values of \( f \) were used for \( w \) as follows,
In the node of layer 3, the normalized weight $w_i^*$ of rule $i$ was computed as the ratio of the $i$th rule’s weight to the sum of the weights of all $n$ rules as:

$$w_i^* = w_i / \sum_{i=1}^{n} w_i.$$  \hfill (13)

An output membership function $g$ was computed from the network-input $X$ and the consequent parameters that are updated through least squares error, $\gamma$ and $\delta$, as:

$$g_i = \gamma_i X + \delta_i.$$  \hfill (14)

In the node of layer 4, the fuzzy values were defuzzified back to crisp output values as:

$$O_i = w_i^* g_i,$$  \hfill (15)

where $O_i$ is the output contributed by the $i$th rule toward the overall ANFIS output.

Last, $\hat{\text{ARID}}$, the overall output of the model, was computed in the fifth layer by summing up the outputs contributed by all rules:

$$\hat{\text{ARID}} = \sum_{i=1}^{n} O_i.$$  \hfill (16)

Before feeding the data into the system, each monthly dataset of 56 input–output combinations was divided into three subsets: a training set (44), a testing set (11), and a validation set (1). An ANFIS model was evaluated by applying the leave-one-out cross-validation technique and computing RMSE and ME for each month, location, and lead-time forecast using the estimated and original values of ARID.

d. Computations for LR models

The $m$-month-ahead forecasts of ARID were generated as follows after estimating the coefficients $u$ and $v$ by regressing the lead-time values of ARID ($\text{ARID}_t$) and the current values of $X$ ($X_t$):

$$\text{ARID}_{t+m} = uX_t + v,$$  \hfill (17)

where $t$ is the month and $m$ is the lead time. Values of $m$ ranged from 1 to 3 months.

For regression and forecasting, respectively, training and validation datasets were used. An LR model was evaluated for each month, location, and lead-time forecast using the leave-one-out technique of cross validation and the RMSE and ME values.

e. Computations for ARMA models

For each location, 12 time series of ARID were created, one for each month. For instance, one time series comprised the ARID values from January 1951 through December 2005, another series comprised the values from February 1951 through January 2006, and the 12th series comprised the values from December 1951 through November 2006. The time series ending in December, for instance, was used to make forecasts for January 2006 (1 month ahead), February 2006 (2 months ahead), and March 2006 (3 months ahead). Similarly, the series ending in January was used to predict for February 2006 (1 month ahead), March 2006 (2 months ahead), and
April 2006 (3 months ahead), and so on. The monthly time series were created with the assumption that weather conditions vary across months. From each monthly series, 16 subseries were created by truncating the monthly series by 1–15 yr. For instance, the first subseries of the series ending in December (mentioned above) ended in December 2005, the second in December 2004, and the 16th in December 1990. Accordingly, 16 ARMA models were developed for each month. The subseries were created for model evaluation purposes (to create \( n \) observations), and 16 was chosen for \( n \) by considering 40 yr as the minimum length of the time series required to reflect any long-term periodicity in the data (56 – 40 = 16).

The ARMA models were developed for each month with the assumption that the weather conditions in this region vary across months. The monthly ARMA models would be consistent and thus be easy to compare with the other methods for which the variation across months in the degree of linkage between large-scale teleconnection patterns and the weather in this region is assumed. A seasonally integrated ARMA model, called the seasonal autoregressive integrated moving average (SARIMA) model can satisfactorily describe a time series that exhibits nonstationarity not only across seasons but also within a season (Mishra and Singh 2011).

To observe the relationships between the present and past values of ARID, autocorrelation functions (ACF) and partial autocorrelation functions (PACF) were calculated for each of the 16 time series of each month and location. In each series, the values of ACF and PACF exhibited periodicities corresponding to the positive correlation between values separated by 12 months, indicating a need for seasonal differencing (Fig. 3a). Thus, each time series was seasonally differenced as ARID\(_{t} - ARID_{t-12}\). To study the behavior of the seasonally differentiated series, its ACF and PACF values were calculated and analyzed to find any correlation among the ARID values. Because the ACF and PACF values of the differenced series did not exhibit any periodicities (Fig. 3b), no further differencing was needed. Thus, the \( D \) in an integrated ARMA model, ARIMA\((j, d, k) \times (J, D, K)\), was set to 1. The parameters \( j \), \( d \), \( k \), \( J \), \( D \), and \( K \) in the model denote the nonseasonal autoregressive (AR) model of order \( j \), the nonseasonal difference of order \( d \), the nonseasonal moving average (MA) model of order \( k \), the seasonal AR model of order \( J \), the seasonal difference of order \( D \), the seasonal MA model of order \( K \), and the seasonal lag \( s \), respectively. The characteristics of the ACF and PACF of the seasonally differenced series showed a strong peak at a lag \( s \) of 12 months in the ACF. Smaller peaks appeared at \( s = 24 \) and \( 36 \) in the ACF, combined with peaks at \( s = 12, 24, 36, \) and 48 in the PACF (Fig. 3a). The cutting off of ACF after lag \( s = 12 \) months (a seasonal lag of 1 yr) and the tailing off of the PACF of the seasonally differenced series indicated a seasonal MA of order \( K = 1 \). Therefore, the orders of \( J, D, \) and \( K \) for the seasonal part were 0, 1, and 1, respectively. The same orders were found for each series, month, and location. For the nonseasonal component, the ACF and PACF values of the seasonally differenced series within seasonal lags \( l = 1, 2, \ldots, 11 \) did not show any periodicities. Therefore, no nonseasonal differencing was needed (\( d = 0 \)). Because the
nonseasonal behavior of ACF and PACF was the same in each series for each month and location, the order of \( d \) was also the same for all cases. For the order of \( j \) and \( k \), the tailing off of ACF and the cutting off of PACF after lag \( l = 1 \) indicated a nonseasonal AR of order \( j = 1 \) and a nonseasonal MA of order \( k = 0 \) (Fig. 3b). Thus, the most appropriate ARMA model for each month and location was ARIMA \((1, 0, 0) \times (0, 1, 1)_{12}\). Because the ARMA model included a seasonal component and differencing, it is hereinafter referred to as the SARIMA model.

Using 16 SARIMA models for each month, ARID was forecast with 1-, 2-, and 3-month lead times. The SARIMA models were then evaluated for each location, month, and lead-time forecast using the RMSE and ME as goodness-of-fit measures.

f. Computations for the ENSO approach

To characterize the ENSO episodes, generally the JMA index is used because it is more sensitive to El Niño and La Niña conditions than are the other indices (Trenberth 1997; Hanley et al. 2003). When the index is \( \geq 0.5^\circ\text{C} (\leq -0.5^\circ\text{C}) \) for six consecutive months, including October–December, an El Niño (La Niña) episode is defined (Hanley et al. 2003). The episode then lasts from October through the following September. The episode for all other values of the index is termed neutral. This study, however, used the modified version of the index because it gives more significant results than does the original index (Keeling 2008; Gérard-Marchant and Stooksbury 2010; Royce et al. 2011). In the modified JMA index, the El Niño (La Niña) episode stops as soon as the temperature conditions are no longer met (Gérard-Marchant and Stooksbury 2010).

The monthly time series of ARID were separated into three ENSO categories. In each category, the ARID values were further separated into 12 months and individually averaged. The monthly average value of ARID in a particular month and ENSO phase was considered to be the forecast of ARID for that month during that ENSO phase.

g. Probabilistic forecasting

A probabilistic forecast is more “honest” than a deterministic forecast (Krzysztofowicz 2001) and is never wrong (Mason 2002). Although deterministic forecasts are more useful, they are more difficult to produce. Therefore, most forecasts are probabilistic.

On the basis of ME, the three best-performing models were selected for making the probabilistic forecasts of ARID. From the 56 values of prediction error \( \epsilon \) (16 in the case of SARIMA), computed as estimated ARID minus original ARID, a cumulative distribution function was created, from which the probability of \( \epsilon \) falling inside the negative and positive values of the prediction error threshold \( E \), denoted as \( P(-E < \epsilon < E) \), was computed as \( P(\epsilon < E) - P(\epsilon < -E) \). The \( P(-E < \epsilon < E) \) value was then used as the probability of \( (F - E) < \text{ARID} < (F + E) \), where \( F \) is the deterministic forecast made by a model. Such computations were made for each location, month, and model. For \( E \), 0.1 was used to indicate a small error.

3. Results and discussion

a. Comparing the forecasting models

The modeling efficiency of each of the four models (ANFIS, ANN, LR, and SARIMA) varied depending on the month and location (Fig. 4). The variation in the efficiencies of the models gave an indication that, for a specific month at a specific location, one specific model may be more appropriate than the others. In general, the efficiencies of the CI-based models (ANFIS, ANN, and LR) and the ENSO approach were large during the winter and small during the summer in most of the locations. Furthermore, these models indicated an acceptable level of performance during the winter, as their ME values were generally positive. During the summer, they indicated an unacceptable level of performance, as their ME values were mostly negative. This seasonal difference was possibly due to the stronger signals of ENSO and large-scale teleconnections during the winter than in the summer. The time series–based model, SARIMA, however, did not show this pattern. Its efficiency largely depended on the periodicity of the time series data. For instance, SARIMA was highly efficient for a month that received a large amount of precipitation in each year of the time series.

In general, the efficiency of each model, including the ENSO approach, was better in southern locations, such as Miami, Bartow, and Live Oak, than in the northern ones, such as Plains and Blairsville (Fig. 4). The modeling efficiencies decreased in the direction of south to north. Miami, in southern Florida, had the largest efficiencies, whereas Blairsville, in northern Georgia, showed the lowest efficiencies. This differential efficiency probably arose because of stronger ENSO and teleconnection signals and larger precipitation periodicities in the south than in the north.

In terms of modeling efficiency, ANN generally demonstrated the best performance for southern locations such as Miami, Bartow, and Live Oak. Of the five approaches compared, the performances of LR and ANFIS were generally poor for all locations. The poor efficiency of LR was likely the result of its being a linear
model, such that it could not represent the nonlinearity that possibly exists between ARID and climate indices. For ANFIS, even though it was supposed to account for nonlinearity in a system, it performed poorly. The poor performance was most probably due to limited data (Costa Branco and Dente 2001; Marce et al. 2004; Sahoo et al. 2005). The total number of parameters associated with an ANFIS model was 15, whereas the number of samples in the training dataset was only 44. Because the training data points were fewer than 4 times the number of model parameters, a threshold value (Marce et al. 2004), the small training dataset probably caused overfitting so that the generality was lost. Because of limited data, ANFIS probably could not generate proper fuzzy rules and derive all the complex information of the system properly. The use of ANN, on the other hand, does not necessarily require a large sample size (Zhang et al. 1998). The ANN models perform well even with a sample size of less than 50 (Kang 1991). Being a universal approximator because of its ability to approximate any nonlinear relationship between inputs and outputs to any degree of accuracy (Hornik et al. 1989; Smith 1996; Zhang et al. 1998; Samarasinghe 2007), ANN performed better than the other CI-based models—namely, ANFIS and LR. The better performance of ANN for southern locations than for northern locations indicated that the relationships between climate indices and ARID are stronger in the southern locations. Probably because of weak teleconnections between the precipitation pattern in northern locations and the climate indices, ANN performed poorly for Plains and Blairsville. Generally, ENSO- and CI-based methods—namely, ANFIS, ANN, and LR—performed poorly for the northern part of the region. The performance of CI-based methods was even poorer than that of the ENSO approach, indicating that the signals of the large-scale teleconnection patterns represented by the climate indices used in this study are even weaker than those of the ENSO (i.e., JMA) in northern locations. Also in southern locations, the ENSO approach performed better than the other models for some months: February and April for Miami; January, April, and October for Bartow; and July–October for Live Oak. These results indicated that the ENSO signal is stronger in these months and locations than are the teleconnections represented by the climate indices. For the other months, signals represented by the climate indices seemed stronger than that of the ENSO. The efficiency of SARIMA models also varied depending upon month and location. Generally, the performance of SARIMA was poorer than those of ANN and ENSO for most of the months in southern locations. The poor performance probably resulted because a SARIMA model assumes that time series are generated from linear processes, whereas the drought time series are often nonlinear and seasonal (Granger and Terasvirta 1993; Zhang et al. 1998). A SARIMA model generally performs better when periodicity exists in the data, as the periodicity is taken into account by the model by differencing the data.
These months. Table 6 also shows the best method in terms of modeling efficiency for each month–location combination. In general, ANN had the highest efficiency for southern locations such as Miami, Bartow, and Live Oak. The SARIMA model performed best for a mid-northern location such as Plains, whereas for Blairsville, the northernmost location in the region, the performance of the ENSO approach was the highest of all methods. The probable causes of these phenomena are explained in the preceding paragraph.

b. Lead-time forecasts

The RMSE associated with 1-, 2-, and 3-month-ahead forecasting for each month and location varied depending on the models used. For the CI-based models—ANFIS, ANN, and LR—the occurrence of RMSE$_{t+1} <$ RMSE$_{t+2} <$ RMSE$_{t+3}$ was less than 25% (Fig. 5). The terms RMSE$_{t+1}$, RMSE$_{t+2}$, and RMSE$_{t+3}$ denote the RMSE associated with 1-, 2-, and 3-month-ahead forecasting, respectively. The percentage of occurrence shows the number of times RMSE$_{t+1} <$ RMSE$_{t+2} <$ RMSE$_{t+3}$ occurred out of the total 60 cases (5 locations × 12 months). For SARIMA, the occurrence of RMSE$_{t+1} <$ RMSE$_{t+2} <$ RMSE$_{t+3}$ was 50% (i.e., 30 out of 60).

Although RMSE$_{t+1} <$ RMSE$_{t+2} <$ RMSE$_{t+3}$ did not occur for 100% of the cases with each model, the results gave an indication that short-term forecasting (e.g., for 1 month in advance) can generally make more accurate estimations than can long-term forecasting (e.g., for 2 or 3 months in advance). The SARIMA models produced less accurate forecasts than did the CI-based models for the months that were more distant because of weaker autocorrelations in the time series data with increasing lag times. The occurrence of RMSE$_{t+1} <$ RMSE$_{t+2} <$ RMSE$_{t+3}$ for fewer cases with the CI-based models than
with the SARIMA models gave an impression that, although ARID is slightly more correlated with climate indices with shorter lags, the correlations between ARID and the climate indices with longer lags are also possible.

c. Probabilistic forecasting

The distribution of $\varepsilon$ in terms of the box-and-whiskers plots for each of the 12 months, five locations, and the three methods (ANN, ENSO, and SARIMA) whose performances were better than those of the others as shown by the ME values presented in section 3a are presented in Fig. 6. In line with the modeling efficiency results, the performance of each method in terms of producing small errors varied depending on months and locations. For instance, the interquartile range of errors for Miami in January was the smallest with the ENSO approach, whereas that for Miami in March was the smallest with ANN. Similarly, ANN produced the smallest errors for Miami in June and July, whereas SARIMA produced the smallest errors in August and September.

The probability of $\varepsilon$ falling inside $-0.1$ and $0.1$, denoted as $P(-0.1 < \varepsilon < 0.1)$ and computed for each location, month, and method, is presented in Table 7. From this table, the probabilistic forecasts of ARID can be made. For instance, $P(-0.1 < \varepsilon < 0.1)$ for Miami in January as estimated by the ENSO approach is 0.39. If the forecast $F$ made by this approach for Miami in January were 0.26, $P(F - 0.1 < \text{ARID} < F + 0.1)$; that is, $P(0.16 < \text{ARID} < 0.36)$ for this month and location would be 0.39. That is, there would be 39% chance that the forecast of ARID would fall inside $0.26 \pm 0.1$.

![Fig. 5. The percentage of occurrence of RMSE_{t+1} < RMSE_{t+2} < RMSE_{t+3} for different forecasting models, where RMSE_{t+1}, RMSE_{t+2}, and RMSE_{t+3} denote the RMSE values associated with 1-, 2-, and 3-month-ahead forecasting, respectively. There were 60 cases in total (5 locations x 12 months), and therefore 50% means 30 cases.](image)

![Fig. 6. Box-and-whiskers plots of errors, computed as the difference between the original and model-estimated values of ARID, for various months, locations, and models. Note that E = ENSO, N = ANN, and S = SARIMA.](image)
Similarly, if the forecast made by an ANN model for Bartow in July were 0.27, \( P(F - 0.1 < \text{ARID} < F + 0.1) \) that is, \( P(0.17 < \text{ARID} < 0.37) \) for this month and location would be 0.82 (Table 7).

The large probability values in Table 7 may be due to extremely wet conditions, extremely dry conditions, or good teleconnections between the large-scale oceanic-atmospheric signals (climate indices) and ARID in this region. Under extremely wet conditions, almost all values of ARID are 0, such as those in January, February, and December for Plains and Blairsville (Figs. 7d,e). Under extremely dry conditions, values of ARID concentrate at 1. Thus, the width of the prediction error is almost zero, and \( P(-0.1 < \varepsilon < 0.1) \) is high in both severely wet and dry conditions. When the distribution of ARID is not close to 0 or 1—such as in July, August, and September for Miami (Fig. 7a), July and August for Bartow (Fig. 7b); and February for Live Oak (Fig. 7c)—the large value of \( P(-0.1 < \varepsilon < 0.1) \) indicates a strong teleconnection between climate indices and ARID and good performance of a forecasting model such as ANN. Figure 7 also illustrates 75% probability that the mean value of ARID as forecast by an ANN model falls inside a certain range. For January in Miami, for instance, there is 75% chance that the forecast of ARID falls inside 0.30 and 0.63 (Fig. 7). The 75% probability was computed as the difference between the 87.5th percentile and the 12.5th percentile.

4. Conclusions

This study attempted to verify whether the use of climate indices—namely, AMO, NAO, JMA, Niño-3.4, PDO, and PNA—could improve the current level of drought forecasting, which is essentially ENSO based, for the southeastern United States and also examined the performance of ANFIS, ANN, ARMA, and LR models relative to that of the ENSO approach in forecasting ARID, an agricultural drought index.

The results indicated that improved forecasting may be possible for several months in most locations of the region with the use of climate indices and forecasting models such as ANN and SARIMA. For instance, ARID forecasting could be improved for January in Miami using the SARIMA models, whereas the index could be better forecast for March–December using the ANN models.

The results showed that the climate indices used in this study have good potential for use in drought forecasting for both winter and summer months, especially for the southern part of the region. Thus, the climate indices can potentially contribute to a better forecast of droughts.

The performance of the ANFIS, ANN, ENSO, LR, and SARIMA models varied depending upon month and location. In general, climate index–based models and the ENSO approach performed better during the winter than during the summer because of stronger signals of the ENSO and large-scale teleconnections in the winter than in the summer. The efficiency of SARIMA models depended largely on the periodicity of precipitation in the data. The ANN models generally outperformed the other models for the southern part of the region, whereas the ENSO approach performed better than the other models for the northern part. The SARIMA models were best for the central part of the region. In general, LR and ANFIS performed poorly, as the former could not represent the nonlinearity between ARID and the climate indices and the latter lost generality mainly because of limited data. The relatively better performance of ANN was due to its ability to approximate the nonlinear relationship between input
and output. The climate index–based models performed worse than the ENSO approach for northern locations, indicating that the signals of the large-scale teleconnection patterns represented by the climate indices are even weaker than those of ENSO (i.e., JMA) in the northern part of the region. In general, the performance of each model decreased in the direction of south to north as a result of weaker ENSO and teleconnection signals and less-distinct precipitation periodicities in the north than in the south.

Acknowledgments. The authors thank the USDA National Institute for Food and Agriculture and the NOAA Regional Integrated Sciences and Assessments program for supporting this work through grants.

REFERENCES


