Earth and Space Science

RESEARCH ARTICLE
10.1029/2022EA002315

A Place-Based Approach to Drought Forecasting in South-Central Oklahoma

Renee A. McPherson[1,2], Irenea L. Corporal-Lodangco[1], and Michael B. Richman[3]

[1]South Central Climate Adaptation Science Center, University of Oklahoma, Norman, OK, USA, [2]Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, OK, USA, [3]School of Meteorology, University of Oklahoma, Norman, OK, USA

Key Points:

- The 2011–2014 drought in the Southern Great Plains stressed water management and tourism
- The authors developed an experimental predictive drought model for south-central Oklahoma
- The model predicted monthly Palmer Drought Severity Index well out to 18 months

Correspondence to:

R. A. McPherson,
renee@ou.edu

Citation:

McPherson, R. A., Corporal-Lodangco, I. L., & Richman, M. B. (2022). A place-based approach to drought forecasting in south-central Oklahoma. *Earth and Space Science*, *9*, e2022EA002315. https://doi.org/10.1029/2022EA002315

Received 27 MAY 2022
Accepted 27 SEP 2022

Author Contributions:

Conceptualization: Renee A. McPherson
Data curation: Irenea L. Corporal-Lodangco
Formal analysis: Irenea L. Corporal-Lodangco
Funding acquisition: Renee A. McPherson
Investigation: Renee A. McPherson
Methodology: Irenea L. Corporal-Lodangco, Michael B. Richman
Project Administration: Renee A. McPherson
Resources: Renee A. McPherson
Software: Irenea L. Corporal-Lodangco
Supervision: Renee A. McPherson
Validation: Irenea L. Corporal-Lodangco, Michael B. Richman

Abstract To assist water managers in south-central Oklahoma prepare for future drought, reliable place-based drought forecasts are produced. Past-, present-, and future-forecasted climate indices (Multivariate ENSO Index, Pacific Decadal Oscillation index, and Atlantic Multidecadal Oscillation index) and past and present Palmer Drought Severity Index (PDSI) are employed as predictor variables to forecast PDSI using a multivariate regression technique. PDSI is forecasted 18 months in advance with sufficient skill to provide water managers early warning of drought. Using a training data set obtained from the period January 1901 to November 2021, a second-order model equation that contains, without any restriction, all the predictors and their interaction terms is built to predict drought intensity. Significant predictors are selected through stepwise regression, with cross-validation producing the simplest restricted model that describes the data well. PDSI values are predicted using 1000 fitted restricted models produced from bootstrapping, then averaged monthly. The technique found the best-fitting model and estimated the model coefficients that minimized the sum of squared deviations between the fitted model and the predictor variables. The adjusted R-squared value of the restricted model is large enough to explain an adequately accurate model, and relatively low values of error measures point to good predictive ability of the model. Although the model slightly overestimates the PDSI forecast maxima and minima, it necessarily captures the timing of the periods of severe to exceptional drought.

Plain Language Summary Natural resource managers in south-central Oklahoma experienced severe-to-exceptional drought conditions during much of 2011–2014. Water managers from the Chickasaw Nation expressed the desire to obtain earlier advisories of impending drought. As a result, the author team developed an experimental predictive drought model using several statistical techniques. Results indicate that the model can adequately predict the timing and intensity of a common drought indicator (i.e., the Palmer Drought Severity Index) for south-central Oklahoma out to 18 months in advance. This lead time is useful for water managers to plan for the future risk of water quantity and quality problems.

1. Introduction

Drought is a hazard across south-central Oklahoma and has affected the Chickasaw Nation, the Chickasaw National Recreation Area, and homes and businesses in this region and beyond. Drought is interwoven with other stressors in a changing climate, such as increasing temperatures and enhanced evapotranspiration. Economic, ecological, and social repercussions of multiyear droughts can be extreme (e.g., Hoerling et al., 2014) and are most severely felt at the local level. Yet to prepare for future drought, natural resource managers and other decision makers must try to negotiate complex information about large-scale, coupled ocean-atmosphere interactions, such as El Niño-Southern Oscillation (ENSO) or the Atlantic Multidecadal Oscillation (AMO). These important climate features are so amorphous to decision makers that they generally cannot tie specific actions to ENSO indices or similar index forecasts from the Climate Prediction Center (of the National Oceanic and Atmospheric Organization, NOAA) or other prediction institutions. Hence, there is a need to develop reliable, place-based drought forecasts that are straightforward for nonclimatologists to interpret and apply in their decision processes.

Prior studies demonstrate that drought conditions in the Southern Great Plains are responsive to sea-surface temperatures (e.g., Seager & Hoerling, 2014) and teleconnection patterns (e.g., McCabe et al., 2004), namely the El Niño-Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and Atlantic Multidecadal Oscillation (AMO). ENSO is a seesaw of atmospheric pressure between the eastern equatorial Pacific and Indo-Australian areas (Glantz et al., 1991). It can be represented by the Southern Oscillation Index (Hanley et al., 2003), Niño 3.4

Visualization: Irenea L. Corporal-Lodangco
Writing – original draft: Renee A. McPherson, Irenea L. Corporal-Lodangco
Writing – review & editing: Renee A. McPherson, Irenea L. Corporal-Lodangco, Michael B. Richman

index (Trenberth et al., 2020), Multivariate ENSO Index (Wolter & Timlin, 2011), or other indices, with high positive (negative) values associated with La Niña (El Niño). ENSO has strong teleconnection with various climate anomalies, such as severe droughts, and is linked with other shifts in global weather patterns (Glantz, 1994). As ENSO's positive phase, La Niña tends to be associated with warm, dry winters in the south-central United States that can result in drought conditions that extend for 1–2 years (Cole et al., 2002), especially when associated with the negative phase of the Pacific Decadal Oscillation (PDO; e.g., Kurtzman & Scanlon, 2007; Newman et al., 2016; Nguyen et al., 2021) and the positive phase of the Atlantic Multidecadal Oscillation (AMO; e.g., Cole et al., 2002; McCabe et al., 2004; Veres & Hu, 2013).

Applying optimized canonical correlation analysis, White et al. (2008) found that the variability of Midwest (including Oklahoma) rainfall was correlated to the maximum sea-surface temperature (SST) variability during April–August in the western tropical Pacific Ocean; during September–December, correlations were stronger in response to maximum SST variability in the eastern/central tropical Pacific Ocean. Using an atmosphere-only global climate model forced by SST anomalies, Seager et al. (2005) determined that tropical Pacific SST variations were the primary driver of persistent droughts across the US Great Plains and Southwest. The interannual relationship between ENSO and the global climate is not stationary and can be modulated by the Pacific Decadal Oscillation (PDO; Wang et al., 2014).

PDO is often described as a long-lived, El Niño-like pattern of Pacific climate variability (Mantua & Hare, 2002). While ENSO cycles usually last 6–18 months, PDO continues in the same phase for 20–30 years. As defined by ocean temperature anomalies in the northeast and tropical Pacific Ocean, PDO has positive and negative phases that alter upper-level atmospheric winds. The PDO is positive when SSTs are anomalously cool across the interior North Pacific and warm along the Pacific Coast of the United States and Canada, with below average sea-level pressures over the North Pacific. The PDO is negative when these climate anomaly patterns are reversed (i.e., warm SST anomalies across the interior North Pacific; cool SST anomalies along the Pacific Coast of North America; and above average sea-level pressures over the North Pacific). Extremes in the PDO pattern can have significant implications for global climate, including global land temperature patterns and droughts (Mantua & Hare, 2002). PDO impacts differ based on their alignment with the ENSO cycle. When the PDO and ENSO are in the same phase, El Niño/La Niña impacts may be magnified. Conversely, when the PDO and ENSO are out of phase, they may offset one another, preventing significant ENSO impacts from occurring (Nguyen et al., 2021; Wang et al., 2014).

McCabe et al. (2004) computed that the Pacific Decadal Oscillation and the Atlantic Multidecadal Oscillation together account for more than half (52%) of the spatial and temporal variance in multidecadal drought frequency over the conterminous United States. AMO is based on the average SST anomalies in the North Atlantic basin, typically over 0–80°N and is identified as a natural variability with warm (positive) and cool (negative) phases that may last for 20–40 years at a time (Alexander et al., 2013). Past occurrences of major droughts in the US Midwest and the Southwest have been linked to AMO. For example, the Dust Bowl of the 1930s and the severe drought of the 1950s occurred in the US Southern Plains during the warm phase of AMO (NOAA-AOML). The warm (cold) phase of the AMO is associated with increased (decreased) drought occurrence over the Southwest and north-central US (McCabe et al., 2004).

Although ENSO is the major driver of severe droughts and persistent wet spells over the United States, studies noted that AMO plays an important role in sustaining drought (Schubert et al., 2004; Seager, 2007; Seager et al., 2005). Mo et al. (2009) concluded that the AMO can modulate the influence of ENSO on drought. The influence is large when the SST anomalies in the tropical Pacific and in the North Atlantic are opposite in phase. Drought is favored over the Southwest, Colorado Basin, Great Plains, East Coast, and Southeast during the co-occurrence of a positive AMO and cold ENSO; the opposite is likely to occur during the negative AMO and warm ENSO. In addition, Enfield et al. (2001) found that the impact of ENSO on winter rainfall over the Mississippi River basin depends on the phase of the AMO, with the warm (cold) phase of the AMO associated with less (more) rainfall. Different phases of the AMO also link to different summer precipitation modes of the North American monsoon (Hu & Feng, 2008).

With these and other relationships between teleconnection patterns and drought, it is understandable that many researchers have worked on ways to predict drought conditions. In most cases, these predictions examine drought severity, timing, and length from monthly to seasonal timescales using drought indices such as the Standardized

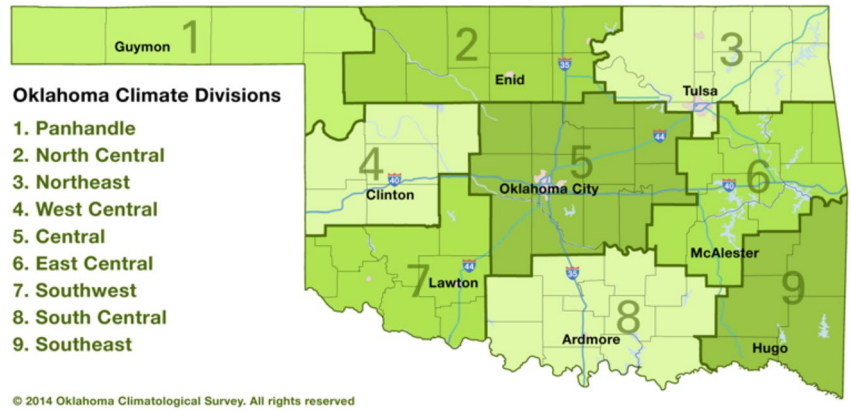

Figure 1. Oklahoma counties separated into nine climate divisions (Oklahoma Climatological Survey, 2014).

Precipitation Index (e.g., Cancelliere et al., 2007; Lavaysse et al., 2020; Singleton, 2012; Yoon et al., 2012). Different techniques have been used for these predictions. For example, Hao et al. (2018) skillfully synthesized the status of using statistical, dynamical, and hybrid methods to predict drought on different timescales and locations worldwide. Statistical approaches require reasonably long time series of climate data to determine relationships between one or more predictors to a selected predictand. Dynamical methods typically rely on the physical processes and computational force of general circulation models, while both dynamical and statistical approaches are blended in the hybrid prediction methods. For this study, we hypothesize that the Palmer Drought Severity Index (PDSI) can predict drought intensity multiple months in advance with sufficient skill to provide water managers an early warning of drought in south-central Oklahoma so that they can take actions to reduce impacts. Because of an existing relationship with the Chickasaw Nation and their continued work on drought and water resources planning, we chose to design an experiment for Oklahoma Climate Division 8 (CD08; Figure 1), which overlaps with nine of the 13 counties within the boundaries of the Chickasaw Nation.

2. Study Area, Data, and Procedure

2.1. Study Area

Oklahoma has nine climate divisions, where boundaries are based on temperature and precipitation averages of reporting stations (Guttman & Quayle, 1996). Corporal-Lodangco et al. (2015) described the characteristics of drought in Oklahoma Climate Division 8 (CD08): (a) On average, droughts last 6 months, with 1 month of severe drought (PDSI ranges from −3.0 to −3.9), and precipitation during a drought averaging 57% of normal precipitation and (b) the most intense drought conditions lasted 40 consecutive months (i.e., from April 1909 to July 1912) and had 18 months of extreme drought (PDSI ranges from −4.0 to −4.9). More recently, severe drought conditions from October 2010 to February 2012 negatively impacted surface flows and shallow aquifer supplies across the Chickasaw Nation, leading to a substantial increase in their water planning efforts for the future. The classification of drought intensity is adopted from the US Drought Monitor (Svoboda et al., 2002).

2.2. Data

Because long-term drought planning is a priority for the Chickasaw Nation, we selected the PDSI (Palmer, 1965) as our forecast variable. Water managers in the Nation are familiar with PDSI, monthly PDSI values for CD08 have been calculated since 1895, and past studies demonstrate that PDSI is good for capturing long-term drought conditions (Burke et al., 2006; Dai, 2011; Dai et al., 1998, 2004). PDSI also has been used widely in state-level plans for drought mitigation or response (Quiring, 2009).

The PDSI is a standardized index that estimates the climate division's relative dryness using temperature and rainfall observations. It can range as low as −10 (extremely dry) to as high as +10 (extremely wet), with values below −2 representing drought and those below −3 representing severe to extreme drought. PDSI is derived from a physical water balance model, allowing it to capture the effect of global warming on drought through potential

Table 1
Data Sets Used in the Study

Monthly data	Length of time	Uniform resource locator (URL)
AMO Index	January 1856–November 2021	https://www.psl.noaa.gov/data/timeseries/AMO/
MEI	January 1871–November 2021	MEI 1871–2005: https://www.esrl.noaa.gov/psd/enso/mei.ext/table.ext.html
		MEI 1950–2018: https://www.esrl.noaa.gov/psd/enso/mei.old/table.html
		MEI 1979–present: https://www.esrl.noaa.gov/psd/enso/mei/
PDO Index	January 1854–November 2021	https://www.ncdc.noaa.gov/teleconnections/pdo/
PDSI	January 1895–November 2021	https://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers/

evapotranspiration (Dai et al., 2019). PDSI has several limitations: It applies a simplistic water-balance equation (Alley, 1984; van der Schrier et al., 2013), values are not spatially comparable (Vicente-Serrano et al., 2011), it is not flexible to examine short-term droughts such as flash drought (Deng et al., 2022), and it assumes that snow and ice melt immediately (van der Schrier et al., 2013). Nevertheless, we believe that PDSI is an acceptable drought indicator for this project because this study focuses on 18-month forecasts for slow-evolving drought concerns, using monthly indices for a single climate division where frozen precipitation melts rapidly.

Global climate indices are quantitative measures of the strength and phase of global teleconnection patterns (McCabe et al., 2004; Wang et al., 2014). Based on past literature regarding droughts, we used the following global climate indices in our model and analysis: (a) Atlantic Multidecadal Oscillation index from NOAA's Earth Systems Research Laboratory (ESRL; Enfield et al., 2001); (b) Pacific Decadal Oscillation index from NOAA's National Centers for Environmental Information (NCEI; Mantua, 1999); and (c) the Multivariate ENSO Index (MEI; Wolter & Timlin, 2011) from NOAA ESRL. Table 1 identifies the period of record for each index and the source from which we obtained data. We built the statistical model using monthly data because of data availability at that interval and to reduce noise in the drought signal (from daily data). Calculations of MEI changed in 1950 (Enfield et al., 2001), resulting in the older data set ranging from 1871 to 2005 and the newer data set ranging from 1950 to present. From 1950 to 2005, these data were highly correlated ($r = 0.991$); hence, we developed a simple linear regression model using data from the overlap period and applied it to all pre-1950 MEIs. From 1950 on, we used the more recent MEI calculation, generating our full 1871-to-present time series.

2.3. Procedure

Figure 2 outlines the order of steps used to develop and validate the forecast models, as well as to produce the drought forecast. Step 1 visualizes and selects the final predictor variables using diagnostic plots, smooths the data sets to reduce noise and develops a time series of lagged PDSI, MEI, PDO index, and AMO index. Lagged values are based on the variable's autocorrelation; hence, for PDSI, we define lag 1 as 18 months, lag 2 as 36 months, etc.; MEI, AMO, and PDO have lag 1 values of 12, 24, and 24 months, respectively, and lag $(n+1) = \text{lag}\ (n) + \text{lag}\ 1$. Step 2 splits the data into distinct training (60% of the data) and testing (40%) data sets. Step 3 builds multivariate regression model to predict PDSI from past PSDI, current and past MEI, PDO, and AMO indices and their combinations, applies a stepwise algorithm to generate the model with the best fit, and conducts this process 1000 times producing a family of models. We validate the selected models (Step 4), forecast the required predictor variables (Step 5), and forecast PDSI using the 1000 models and average the resulting values for the final forecast (Step 6). Finally, we perform hindcasting to assess model performance (Step 7). Details of each step are described in the succeeding sections; an overview is provided in Figure 2.

3. Data Preparation

3.1. Variable Selection

A good way to visualize and select the multiple predictor variables is through diagnostic plots (Figure 2, Step 1). Scatter plots of each pair of predictor variables and density plots (not shown) indicate all four variables exhibit no gross deviation from normality. To test the correlation between the predictor pairs, we sampled observations every 18 months to achieve quasi independence. Table 2 displays the correlation matrix, generated by averaging

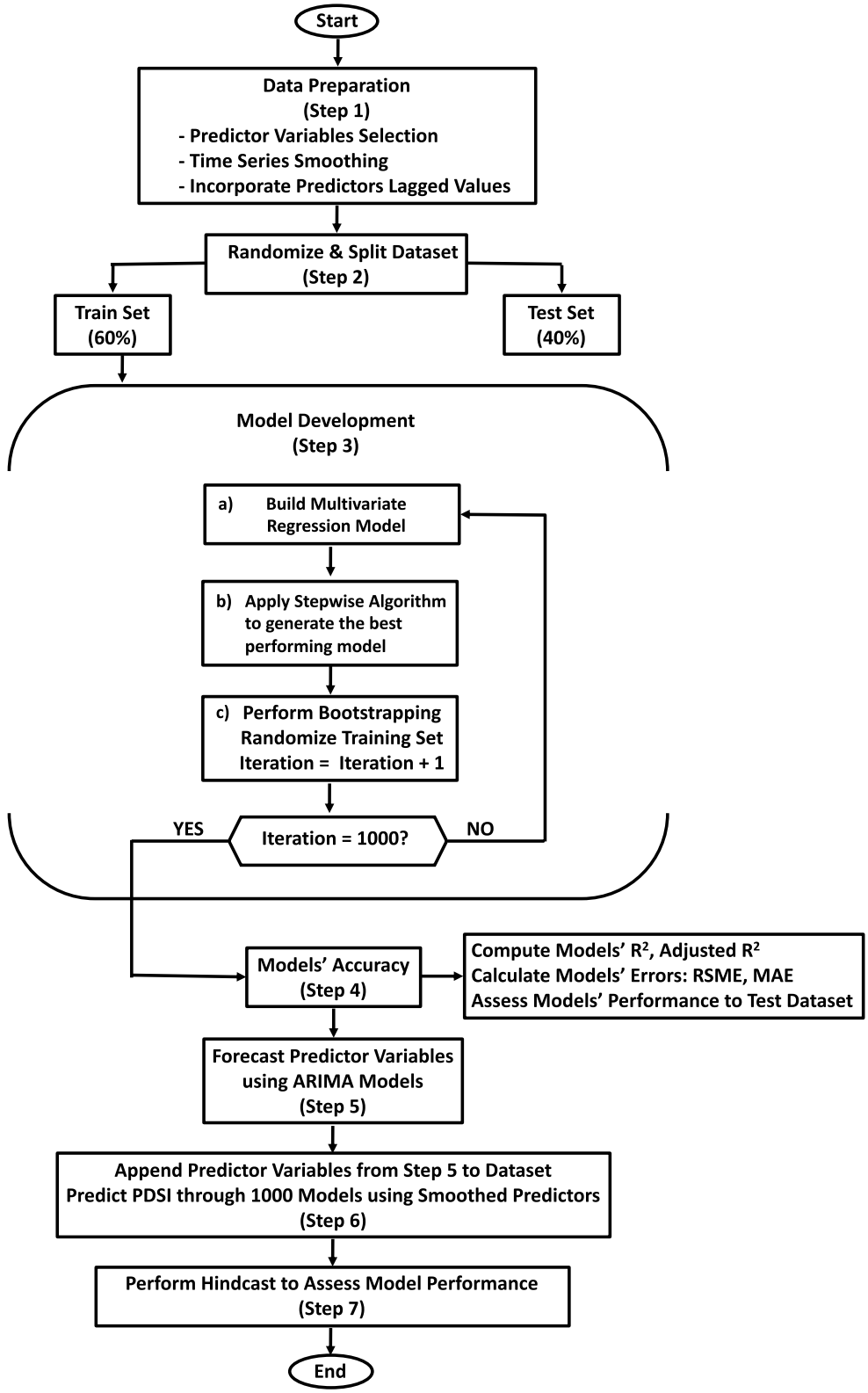

Figure 2. The sequence of processes involved in forecasting drought in Oklahoma Climate Division 8.

Table 2
Pearson Correlations for Each Predictor Pair

	AMO	PDO	MEI
AMO	1.000	−0.054	0.003
PDO	−0.054	1.000	0.461
MEI	0.003	0.461	1.000

the variable values from the 1000 resamplings, by month, then computing the monthly Pearson correlation and, finally, averaging the monthly correlations to yield a single value for each predictor pair. Note that the highest correlation (for PDO and MEI) is only in the moderate range (0.461); thus, model fitting and interpretation should not be problematic.

Figure 3 shows autocorrelation function (ACF) plots, demonstrating how each predictor variable correlates to its past values. As the lag value (number of months between compared values) increases, the correlation decreases. As expected, MEI values are positively significant then uncorrelated after 10 months (Figure 3a) becoming negative during the year afterward, showing the tendency for a La Niña to follow an El Niño in the subsequent year (Chen et al., 2016). In contrast, with a 20–40-year cycle, the AMO (Figure 3b) is highly autocorrelated even after 30 months. Both PDO (Figure 3c) and PDSI (Figure 3d) have positive autocorrelation values decreasing over time, reaching ACF less than 0.1 after about 2 and 1.5 years, respectively. In addition, the partial autocorrelation plots (not shown) show a significant spike only at lag 1, suggesting that all the higher order autocorrelations are effectively explained by the lag 1 autocorrelation.

Plotting the residuals versus the fitted values tests the assumptions of whether the relationship between these variables is linear, whether there is equal and constant variance along the regression line, and whether there are outliers. If the model does not capture an existing nonlinear relationship between the predictor variables and the response variable, this plot will highlight the issue. Figure 4 shows plots of residual versus fitted values of linear regression models (using *lm* in *R*; RStudio Team, 2020): *lm*(PDSI ~ AMO), *lm*(PDSI ~ MEI), *lm*(PDSI ~ PDO), and *lm*(PDSI ~ AMO*MEI*PDO). The plots from the simple linear regression models with PDSI as the response and the AMO and PDO as the predictors (Figures 4a and 4c) clearly indicate a linear relationship between predictors and response and also suggest that there is an equal error variance along the regression line (i.e., there are equally spread residuals around a horizontal line without distinct patterns). In Figures 4b and 4d, the models

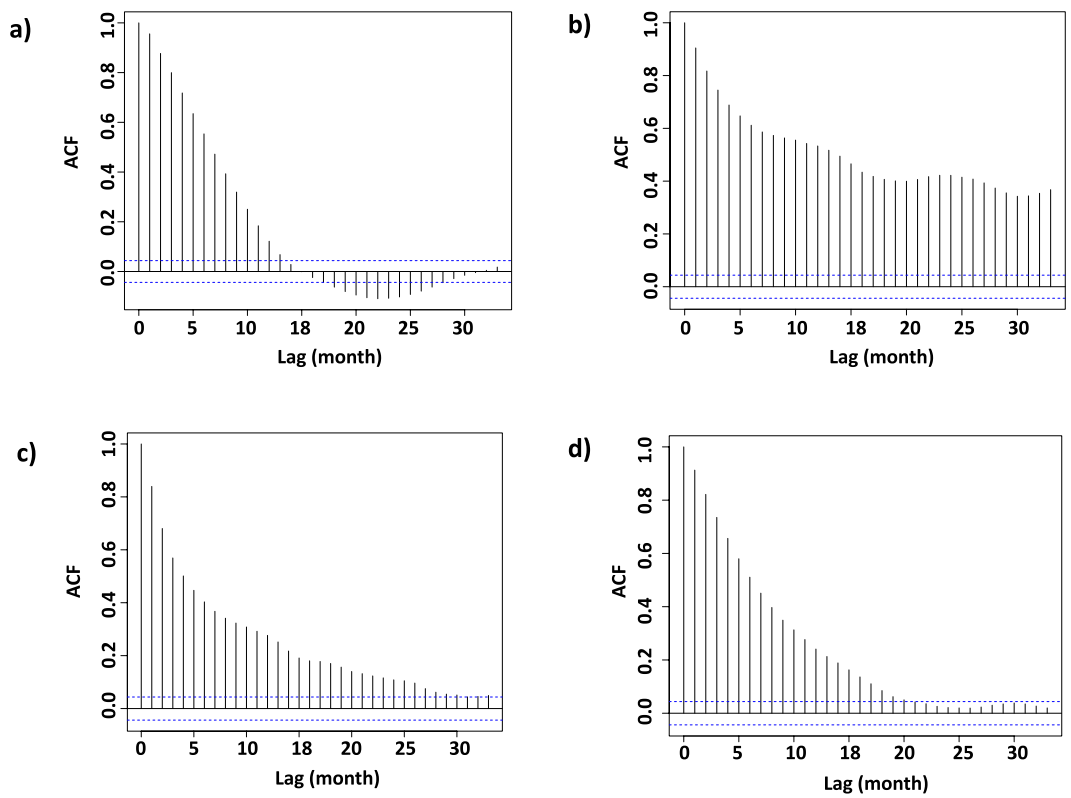

Figure 3. Autocorrelation function plots of predictor variables: (a) Multivariate ENSO Index, (b) Atlantic Multidecadal Oscillation index, (c) Pacific Decadal Oscillation index, and (d) Palmer Drought Severity Index.

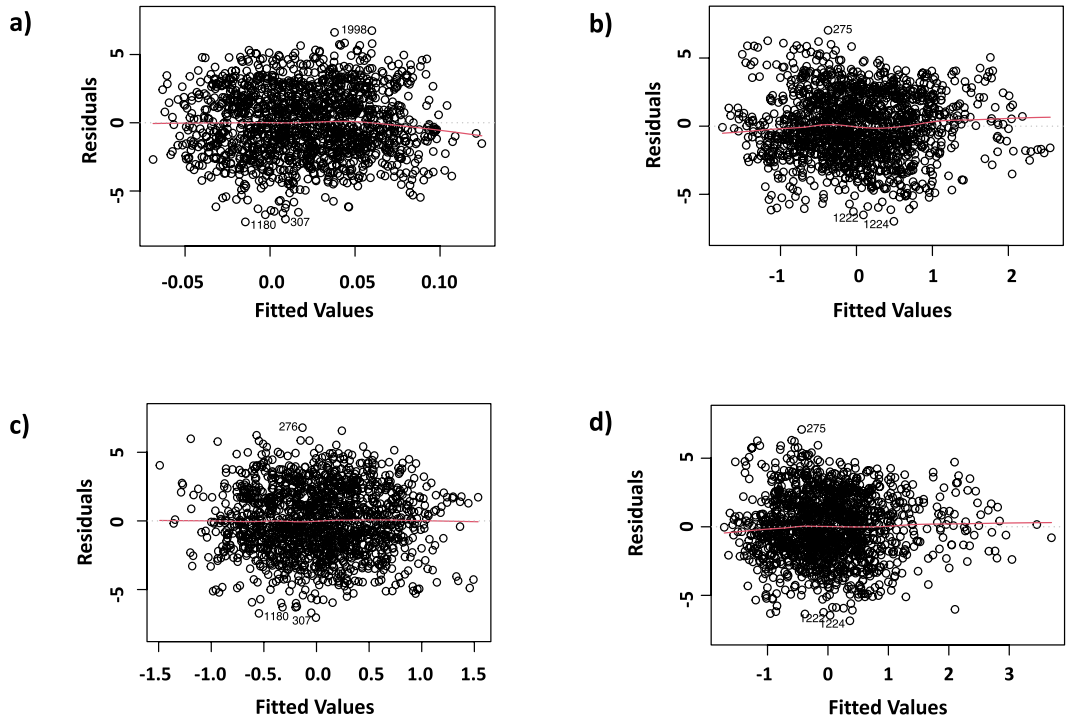

Figure 4. Plots of residual vs. fitted values from fitted regression models: (a) *lm*(PDSI ~ AMO), (b) *lm*(PDSI ~ MEI), (c) *lm*(PDSI ~ PDO), and (d) *lm*(PDSI ~ AMO*MEI*PDO), where PDSI is the response variable and AMO index, MEI, and PDO index are the predictor variables. The red line is the estimated regression line using the *lm* command in *R* that corresponds to the average value of the residuals at each fitted value.

display well-behaved residuals versus the fitted values. All plots show the red line (the average value of the residuals at each fitted value) approximately horizontal at zero, and the residual plots show no fitted patterns. Hence, we will assume a linear relationship between the predictors and PDSI.

We also assessed if residuals are normally distributed using Q-Q plots (i.e., a normal probability plot). All Q-Q plots (not shown) for the same regressions used in Figure 4 showed some nonnormality where standardized residuals exceeded about 1.8 but otherwise followed the diagonal. This suggests that the value of the theoretical quantile is smaller under the empirical cumulative distribution function for the standardized residuals than it is under a normal distribution. Similarly, we examined the residual versus leverage plots (not shown) of the four fitted models, used to detect outliers and identify influential data points in a model. Leverage quantifies the influence that the observed response has on its predicted value. The data point has high leverage if the leverage statistic is above $2(p + 1)/n$ (Bruce & Bruce, 2017), where p is the number of predictors and n is the number of observations. The presence of outliers may affect the interpretation of the model because it increases the relative standard error. In all four models, the standardized residuals were below 3 in absolute value (i.e., the threshold to be a possible outlier; James et al., 2014) and the leverage statistics were below the necessary threshold. The Cook's distance lines are not seen in any of the plots, indicating that there are no outliers influential on the fit of the model.

3.2. Smoothing the Time Series

As a preprocessing step, Fast Fourier Transform (FFT) is applied to time series of AMO index, MEI, PDO index, and PDSI to filter the observed data and reduce their short-term variability (i.e., noise in the time series for our purposes), revealing the important trends (Figure 2, Step 1). FFT is applied with a low-pass filter on variables known to have patterns, such as the MEI, PDO index, and AMO index that may not be statistically different from a red noise process (e.g., Mann et al., 2020). Even if these patterns cannot be distinguished from a red noise process, when they are in a magnitude peak at a given time, t_0, their autocorrelation allows for the extraction of useful signal for several months (e.g., t_1, t_2, …, t_n), where Figure 3 documents n. To implement FFT, the time series is decomposed from its time domain into its constituent frequency domain, undesired frequencies are

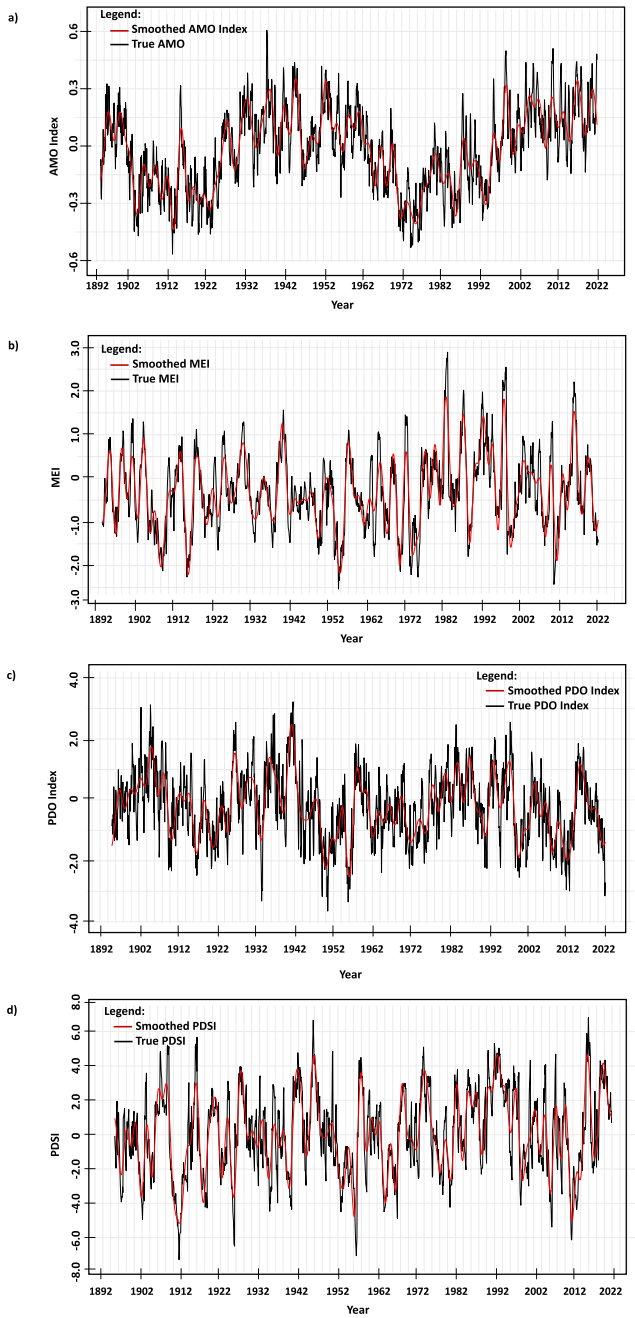

Figure 5. Plots of observed (black) and smoothed (red) values of (a) Atlantic Multidecadal Oscillation index, (b) Multivariate ENSO Index, (c) Pacific Decadal Oscillation index, and (d) Palmer Drought Severity Index. Fast Fourier Transform was applied as a preprocessing step to filter the observed data and reduce its short-term variability, revealing the important trends desired as input for model development.

removed, and inverse FFT is applied to generate the smoothed time series back in time domain. Figure 5 displays the application of FFT for PDSI, MEI, PDO index, and AMO index, resulting in time series that capture major features of evolving teleconnections and drought. The smoothed time series serve as input to the model.

3.3. Incorporating the Lagged Values of Independent and Dependent Variables

The current PDSI depends on the cumulative drought patterns of previous months, so the lagged values of PDSI also is included as one of the predictors. Similarly, the lagged values of AMO index, MEI, and PDO index are also incorporated in the data set (Figure 2, Step 1). The added columns of lagged values in the data set differ and are arranged based on the plots (not shown) of average mutual information of a given time series for a specified number of lags. The variables in the data sets are trimmed so that they have equal lengths.

4. Model Development and Validation

4.1. Model Fitting and Selection

Several statistical models have been used to predict drought, including those using conditional probabilities, Markov Chains, time series, regressions, and artificial intelligence (Hao et al., 2018). For this study, we employed multivariate regression to predict PDSI. Future PDSI values were predicted using the three climate variability indices (MEI, PDO index, and AMO index) and past values of PDSI as the predictor variables. Past patterns in the predictor variables were assumed to continue into the future (i.e., stationarity), as the timescale of the prediction was only 18 months.

The data set consists of the current and lagged values of AMO index, MEI, PDO index, and lagged values of PDSI as the predictor variables. The model development entails randomizing and splitting the data into training and testing sets (Figure 2, Step 2). We applied 60% for training the model, leaving 40% for testing. The model is a second-order model, producing cross-product terms that represent an interaction between the variables. We experimented with a single-order model and several model equations but the forecast results are not as good as the second-order model discussed in McPherson et al. (in preparation). Without any restrictions imposed, all the predictor variables and interaction terms are present in the model equation, and the unrestricted model is fitted to the training data sample (Figure 2, Step 3a). The formula for the model is as follows:

$$Y = (\alpha + \beta_1 X_{1(t)} + \beta_2 X_{1(t-q)} + \beta_3 X_{1(t-2q)} + \beta_4 X_{1(t-3q)} + \beta_5 X_{2(t)} + \beta_6 X_{2(t-q)} + \beta_7 X_{2(t-2q)} + \beta_8 X_{2(t-3q)} + \beta_9 X_{3(t)} + \beta_{10} X_{3(t-q)} + \beta_{11} X_{3(t-2q)} + \beta_{12} X_{3(t-3q)} + \beta_{13} X_{4(t)} + \beta_{14} X_{4(t-q)} + \beta_{15} X_{4(t-2q)} + \beta_{16} X_{4(t-3q)})^2 + \in$$

Y is the predicted value of the dependent variable, PDSI. The parameter α is the intercept of the model. Parameters β_1, β_2, $\beta_3 \ldots$ β_n are the regression coefficients or the "weights" given to the predictor variables. Parameter β_1 represents the change in the mean response corresponding to a unit change in X_1, when X_2, X_3, and X_4 are held constant and so on. X_1, X_2, X_3, and X_4 represent the AMO index, MEI, PDO index, and PDSI, respectively; t is time, q is the lag length, and ϵ is the corresponding random error.

Table 3
The Leading Ten Significant Terms in 1000 Restricted Models With the 100th, 75th, and 50th Percentiles as Well as the Average of Their Corresponding Regression Coefficients

	Absolute value of predictor regression coefficients			
Predictors	100th percentile	75th percentile	50th percentile	Average
AMOI.L1*AMOI.L3	74.36	59.60	56.10	56.01
AMOI*AMO.L3	38.53	28.21	26.39	26.58
AMOI.L2*AMOI.L3	29.29	24.28	22.57	22.44
PDOI.L2*AMOI.L2	22.54	10.38	9.13	9.74
AMOI.L3*MEI.L1	10.80	7.74	7.23	7.27
AMOI.L3*MEI	9.16	7.12	6.63	6.67
AMOI.L2	7.89	6.12	5.73	5.76
PDOI.L2*AMOI.L1	7.02	5.42	5.06	5.09
AMOI*MEI	5.58	4.41	4.11	4.13
PDSI.L2*AMOI	4.13	3.35	3.10	3.10

Note. AMOI is the AMO index and PDOI is the PDO index; L1, L2, and L3 are the lagged values of the associated variable.

Then restrictions are imposed; that is, all regressors whose coefficients are insignificant are excluded from the model equation, keeping only those that have a significant impact on the response variable—predicted PDSI (Figure 2, Step 3b). The restricted model is built using stepwise regression with cross-validation. Stepwise regression uses forward-backward selection to eliminate unnecessary elements of the model until the process results in the simplest model that can no longer be improved. The stepwise regression found the best-fitting model and estimated the model coefficients that minimized the sum of squared deviations between the fitted model and the predictor variables. Model fit was determined by the Bayesian Information Criterion (BIC; Schwarz, 1978), with the final, preferred model having the lowest BIC. Bootstrapping is applied by repeating model development procedures 1000 times creating a family of regression models (Figure 2, Step 3c) and generating random deviates in each iteration. R-squared is 0.96 for the restricted model, demonstrating how well the regression model actually fits the training data set; the corresponding sum of square values are computed in each iteration. Table 3 shows the 100th, 75th, and 50th percentiles and average of the regression coefficients of the leading 10 consistently significant terms of the 1000 model equations. The term AMOI.L1*AMOI.L3, AMO with lag 1 interacting with AMO with lag 3, is the most significant term with a regression coefficient value of 74.36. The interactive terms with AMO index are continually persistent.

4.2. Model Validation

With 1000 model results from 1000 iterations, we computed the average of all the R-squared, adjusted R-squared, root-mean-square error (RMSE), and mean absolute error (MAE) to assess the model performance and quality (Figure 2, Step 4). In this case, both R-squared (not shown) and adjusted R-squared represent the proportion of variation in PDSI that can be explained by the unrestricted (Figure 6a) and restricted (Figure 6b) models, demonstrating model fit. R-squared values close to 1.0 indicate a better fit. For multiple linear regression models, it is common to calculate the adjusted R-squared to address multiple predictor variables. We found the values of the R-squared (not shown) and adjusted R-squared for the unrestricted and restricted models (Figures 6a and 6b) are similar, ranging from 0.92 to 0.96. To assess and validate the performance of the restricted model, the test data set (remaining 40%) serves as input to the restricted model. Both the R-squared and adjusted R-squared values (Figure 6c) average of 0.88 and have a larger spread, ranging from 0.81 to 0.91.

The RMSE is another important measure of model fit, as it accounts for the unexplained variation in the model. Lower RMSE values indicate a smaller amount of deviation and better model fit. As expected, RMSE values for the unrestricted (mean of 0.49; Figure 7a) and restricted (mean of 0.51; Figure 7b) models are smaller than those from the validation of the model equation on the test data set (mean of 0.78; Figure 7c). Although the spread of the RMSE in Figure 8c is larger than the two models, the values are still minimal, suggesting a good model.

The MAE effectively describes the magnitude of the model's residuals. Figure 8 displays histograms of the MAE values for the two model configurations and the MAE from the validation of the model equation on the test data set. Values range from 0.31 to 0.41 (mean of 0.37; Figure 8a) for the unrestricted model, from 0.31 to 0.44 (mean of 0.38; Figure 8b) for the restricted model, and from 0.50 to 0.69 (mean of 0.57; Figure 8c) for the validation of the model equation on the test data set.

For each model, the difference between RMSE and MAE difference is minimal, suggesting that the magnitude of all errors is similar. Importantly, these errors are small compared to the range of PDSI values, generally from −6 to +6 (Figure 5d). Based on the values of RMSE and MAE, the restricted model displays the best model performance.

We analyzed the performance of the drought models from lags 0 to 18 when excluding past PDSI values as one of the predictors (Table 4). Results indicate a general trend toward increased fit (adjusted R-squared) and decreased errors (RMSE and MAE) from lags 0 to 16, then model performance decreases thereafter.

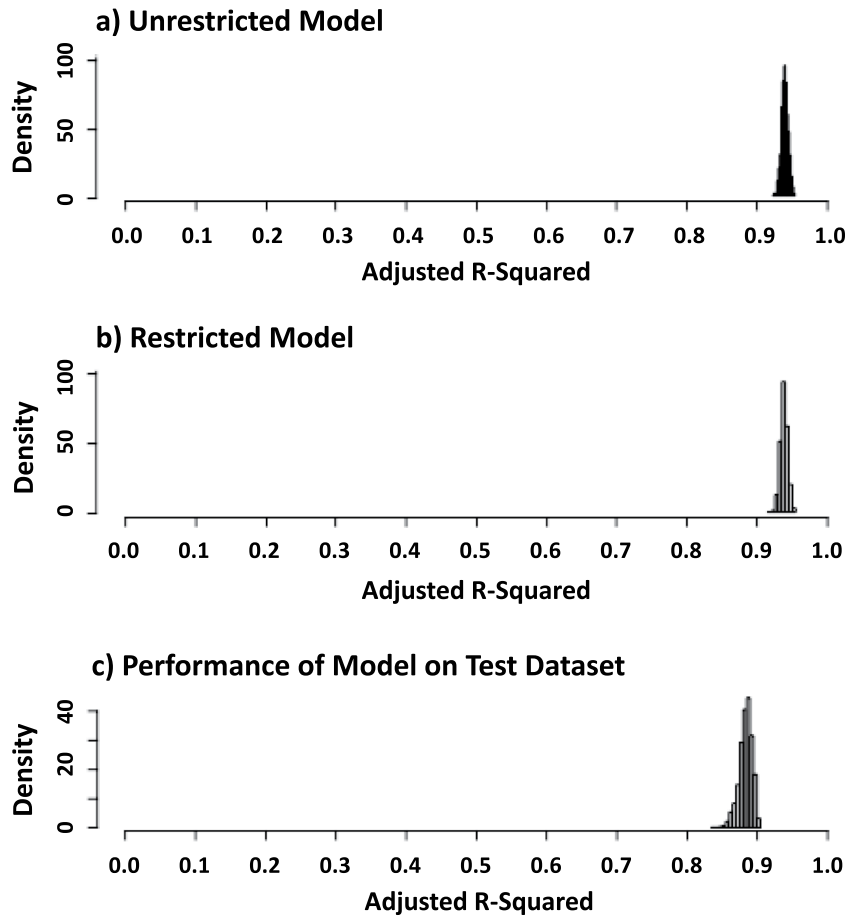

Figure 6. Histograms of adjusted R-squared values for 1000 resamplings of the (a) fit of the unrestricted model using the training data set, (b) fit of the restricted model using the training data set, and (c) performance of the model (using the training data set) as compared to the test data set.

5. Predicting the Global Climate Indices

Based on the autocorrelation values for PDSI (Figure 3a), we chose to develop future PDSI forecasts out to 18 months. Since we are using a regression model, the predictor variables of global climate indices (for AMO, MEI, and PDO) must be predicted out to 18 months for input into the restricted model (Figure 2, Step 5). Although sea-surface temperatures forecasts from seasonal forecast models can be used, our internal tests found better results by using Auto-Regressive Integrated Moving Average (ARIMA). ARIMA acts as a filter that separates the signal in the univariate time series from the noise; the signal is extrapolated into the future to obtain forecasts (Box & Jenkins, 1970). The technique is robust, easy to implement, and has been used in many similar problems (e.g., Lai & Dzombak, 2020; Murat et al., 2018). Nooteboom et al. (2018) employed the ARIMA technique with a modern machine learning approach to predict the El Niño, resulting in predictions that were slightly better than those made by the National Centers for Environmental Prediction. Cuesta and Hunt (2019) used ARIMA and Long Short-Term Memory in predicting the weekly indices of North Atlantic Oscillation and found ARIMA to be the best-performing model. Achite et al. (2022) predicted the standardized precipitation index and the standardized runoff index using both ARIMA and multiplicative seasonal autoregressive integrated moving average. Their results showed the ARIMA models to be the best models, with reasonable accuracy to forecast droughts with up to 12 months of lead time. Other studies have applied ARIMA to predict ENSO (e.g., Rosmiati et al., 2021), AMO (e.g., Frajka-Williams et al., 2017), and PDO (e.g., Diodato et al., 2019).

We applied the Augmented Dickey Fuller (ADF; Dickey & Fuller, 1979) Test to examine stationarity of each predictor variable independently, finding negative ADF statistics and p-values less than or equal to 0.01 (i.e.,

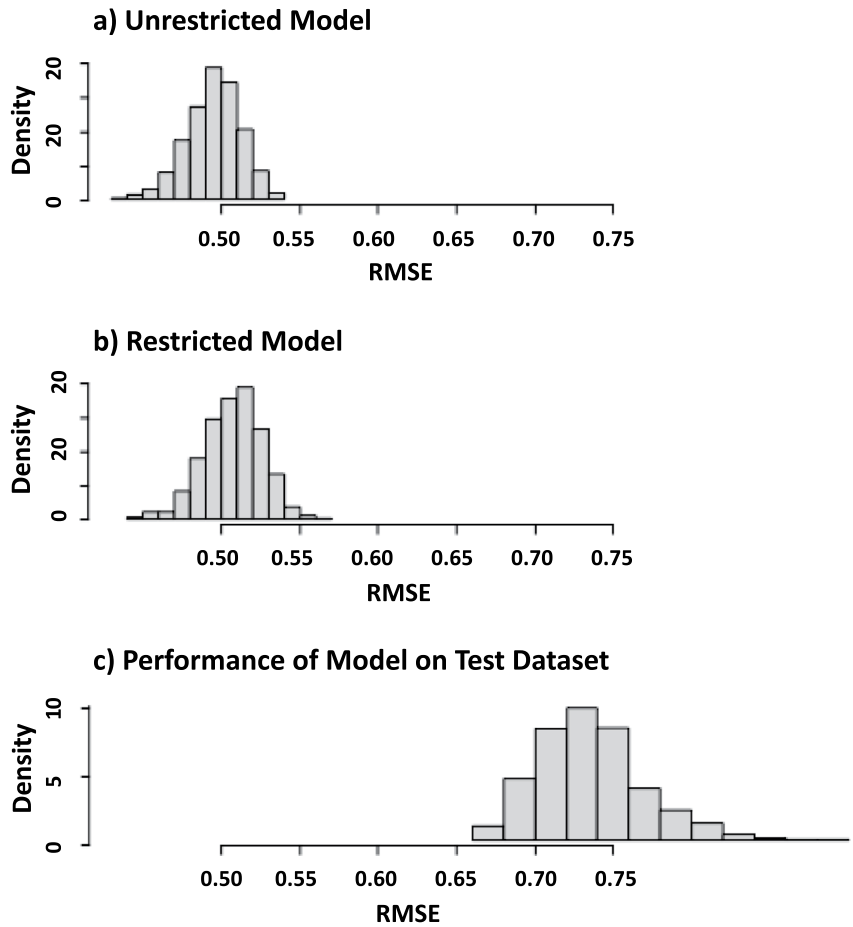

Figure 7. As in Figure 6 except for histograms of root-mean-square error.

stationarity condition is met for all variables). For each teleconnection predictor variable, we decompose the time series, using a 3-month, centered moving average to detrend and obtain a trend component, seasonal component, and the remainder (predominantly random noise). It is important to note the different timescales (not shown) from several years to multiple decades in the individual cycle of each predictor variable.

Once nonstationarity was detected, we examined each series' autocorrelation and partial autocorrelation plots (not shown) to obtain the order parameters (p, q) for the ARIMA model. Spikes at particular lags of the series can help inform the choice of p and q (using *ARIMA* in R). Several model orders are used to fit the ARIMA model; of those, *ARIMA*(1, 1, 38), *ARIMA*(3, 1, 35), and *ARIMA*(1, 1, 25) are the best-fitted model orders for seasonally adjusted time series of AMO index, MEI, and PDO index, respectively, after examining the ACF plots of model residuals and histogram (Figures 9–11) as well as the model summary: p-values, AICs, and model error values (Table 5). The RMSE and MAE values for the PDO index ARIMA model in Table 5 are relatively large, but referring back to Table 3, most of the interactive terms with PDO index are dropped during the stepwise regression process making the PDO index to have less impact on the response variable. All ACF plots (Figures 9b, 10b, and 11b) of model residuals show that all autocorrelations are within the significance limits, indicating that the residuals are behaving like white noise. All predictor variables return large p-values, suggesting that the residuals are white noise. Additionally, the histograms of residuals (Figures 9c, 10c, and 11c) for all predictor variables are approximately normally distributed. The skewness of the residual histograms for AMO index, MEI, and PDO index are −0.034, −0.046, and 0.099, respectively, all near zero indicating symmetric histogram. The kurtosis for the residual histograms are 3.442, 4.446, and 3.702 for AMO index, MEI, and PDO index, respectively, all values are >3 suggesting slightly leptokurtic distributions of the residuals. Finally, we apply the ARIMA models to forecast the AMO index (Figure 12; forecast in red and historical values in black), MEI (Figure 13), and PDO index (Figure 14) out to 18 months.

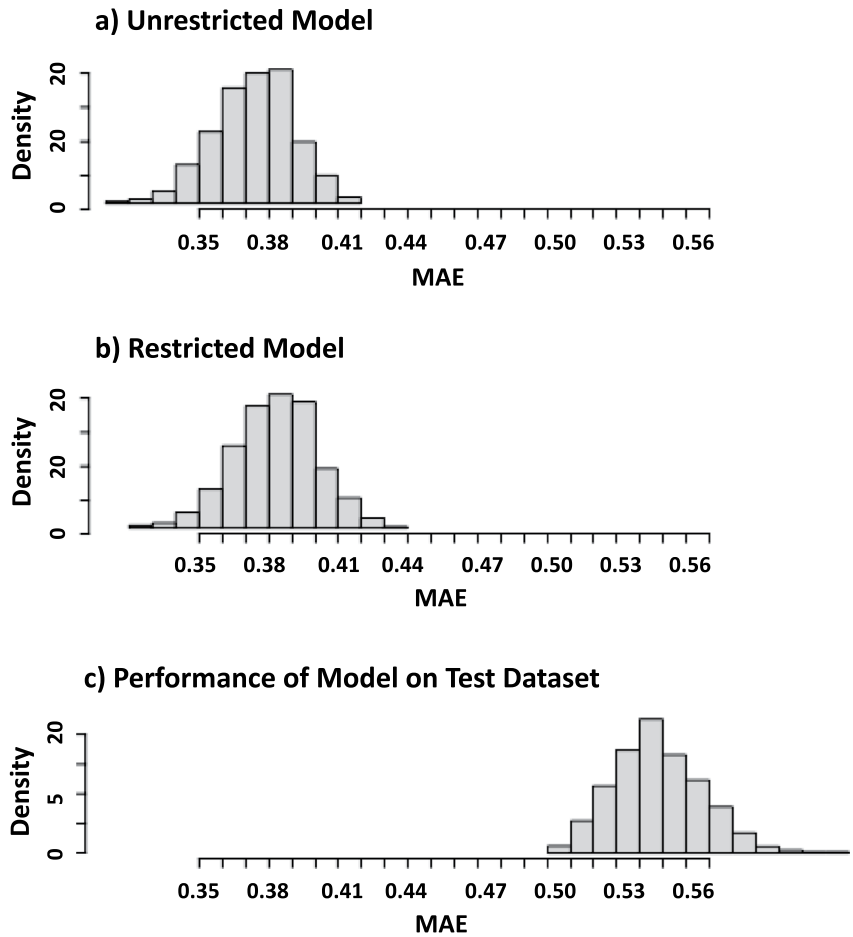

Figure 8. As in Figure 6 except for histograms of mean absolute error.

6. Forecasting and Hindcasting the PDSI

The trimmed historical data set of predictor variables is from January 1901 up to November 2021. To generate an 18-month forecast (December 2021–May 2023) of PDSI (Figure 2, Step 6), ARIMA-predicted values of the predictor variables of those 18 months (from Section 5) are appended to the data set extending it to May 2023. Employing the complete data set as input to the 1000 fitted restricted models (Section 4.1), we predicted PDSI values from January 1901 to May 2023 (Figure 15), averaging the results for each month. Figure 15 shows the PDSI observed (black) and forecast (red) values and the 95% confidence intervals of the bootstrapped forecasts

Table 4
An Assessment of Models' Fit and Errors, Using the Training Data Set, When Only Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), and Multivariate ENSO Index (MEI) Are the Predictors and Lag From 0 Through 18

Restricted models	Lag 0	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9
Adj R-squared	0.132	0.535	0.559	0.566	0.565	0.576	0.598	0.616	0.603	0.631
RMSE	1.931	1.387	1.362	1.330	1.356	1.343	1.295	1.263	1.275	1.239
MAE	1.560	1.081	1.046	1.033	1.055	1.046	1.002	0.978	0.992	0.961
Restricted models	Lag 10	Lag 11	Lag 12	Lag 13	Lag 14	Lag 15	Lag 16	Lag 17	Lag 18	
Adj R-squared	0.652	0.654	0.671	0.685	0.698	0.746	0.759	0.711	0.704	
RMSE	1.203	1.209	1.160	1.165	1.123	1.048	1.015	1.078	1.115	
MAE	0.936	0.940	0.893	0.900	0.871	0.794	0.780	0.829	0.875	

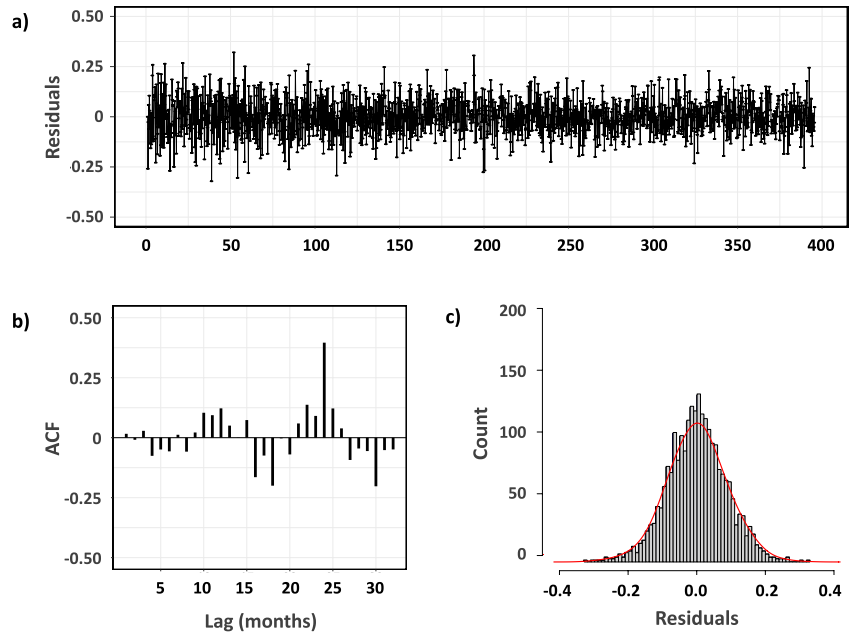

Figure 9. Atlantic Multidecadal Oscillation index ARIMA (1, 1, 38) model plots for (a) residuals, (b) autocorrelation function, and (c) residual histogram with skewness = −0.034 and kurtosis = 3.442.

(lower bound in blue and upper bound in green). Both the upper and lower bounds of the confidence intervals are close to the average forecast, suggesting good performance of the models. The forecast time series shows similar fluctuations to those of the observations on a monthly timescale. The model slightly exaggerates the maxima and minima; however, more importantly for decision makers, the timing of the periods of severe to exceptional drought (PDSI < −3) are captured. The absolute averages of observed and forecast PDSI values are 1.74 and 1.69, respectively—a difference of 0.05, suggesting high accuracy of the forecast.

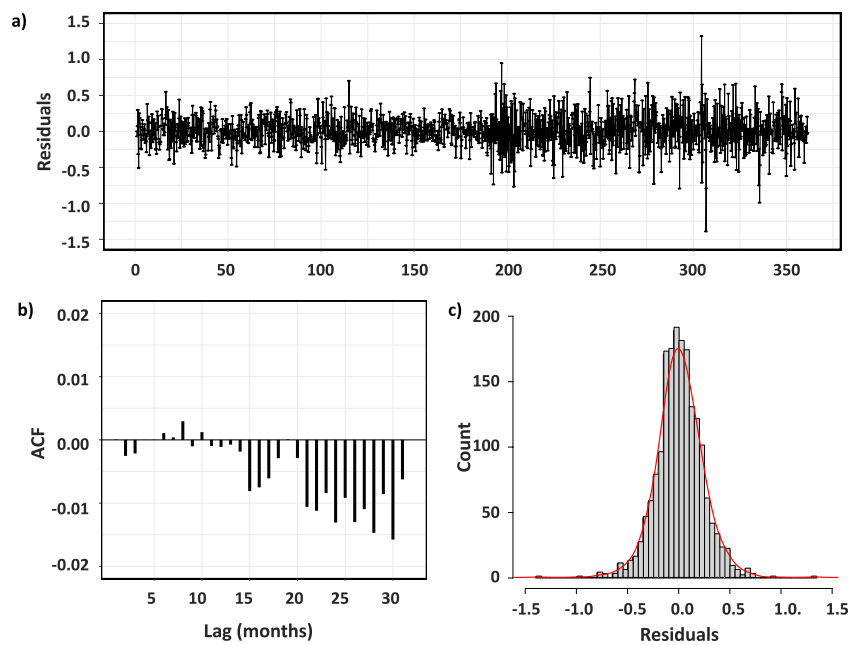

Figure 10. Same as Figure 9 but for the Multivariate ENSO Index ARIMA(3, 1, 35) model. The skewness of the residual histogram is −0.046 and kurtosis is 4.446.

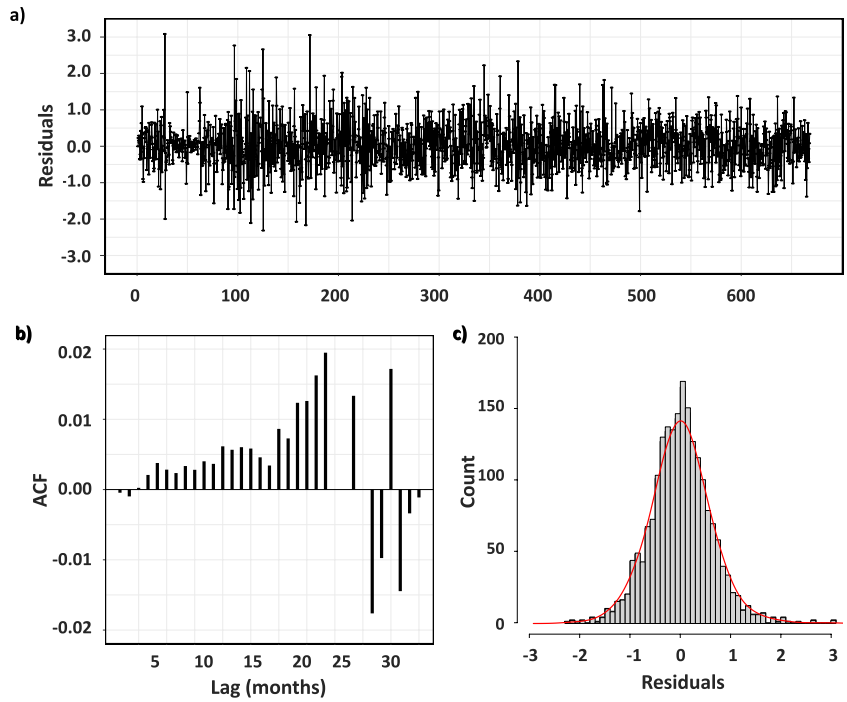

Figure 11. Same as Figure 9 but for Pacific Decadal Oscillation index ARIMA(1, 1, 25) model. The skewness of the residual histogram is 0.099 and kurtosis is 3.702.

To assess the models' performance (Figure 2, Step 7), we performed hindcasting by applying the 1000 drought models to 10 historical drought periods (Table 6). To generate the monthly PDSI forecast, the results of the 1000 models were averaged. R-squared values were calculated to assess the models' fit and RMSE and MAE values documented the models' errors. The drought period from October 1955 to March 1957 generated the lowest adjusted R-squared and the highest values of RMSE and MAE. Across all drought periods, the ensemble mean of R-squared is 0.849—on the left tail of the histogram in Figure 6c, but still is considered a good fit. The ensemble means of RMSE (0.850) and MAE (0.476) are on the right and left tails, respectively, as compared to their counterpart histograms in Figures 7c and 8c. These evaluation statistics are on the tails because we subjectively selected different drought periods rather than a set of 10 random 18-month periods.

Table 5
ARIMA Model Summaries of Atlantic Multidecadal Oscillation (AMO) Index, Multivariate ENSO Index (MEI), and Pacific Decadal Oscillation (PDO) Index

Model summary	AMO index	MEI	PDO index
Model order	(1, 1, 38)	(3, 1, 35)	(1, 1, 25)
p-value	0.025	0.043	0.174
AIC	−3991.54	−216.57	3759.860
RMSE	0.086	0.223	0.604
MAE	0.067	0.168	0.466

7. Conclusions

The study developed a multiple linear regression model to forecast the PDSI out to 18 months in Oklahoma climate division 8. This climate division encompasses a region of the state that is susceptible to drought, with karst topography that supports rapid depletion of surface waters in streams and springs. Much of the region is within the boundaries of the Chickasaw Nation—one of many sovereign tribal nations developing their own drought plans and resources for protecting their waters. Some of these waters are encompassed by the Chickasaw National Recreation Area, part of the National Park Service. During the 2011–2014 drought, tourism was affected by the low water supply in the recreation area. Similarly, small towns in climate division 8 struggled to maintain water supplies during that drought. Hence, the need for a better predictive tool was borne, so that water managers could plan for countermeasures in advance.

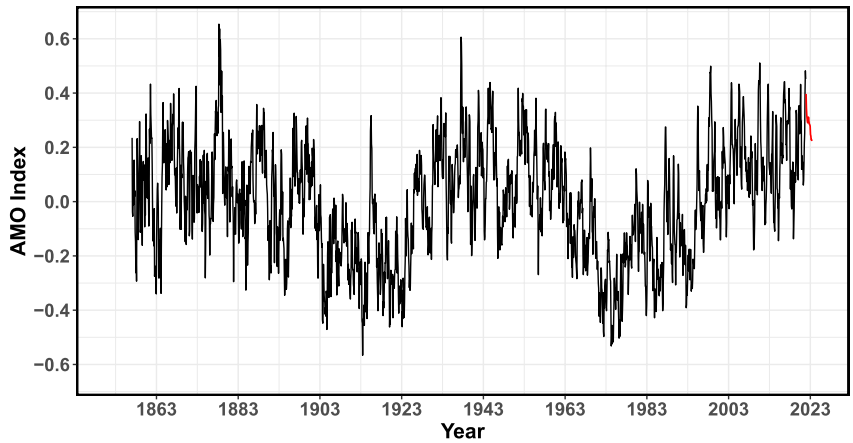

Figure 12. Atlantic Multidecadal Oscillation index actual values (black) and 18-month forecast values (red).

Linkages of teleconnection patterns to the climate system have been established (e.g., Rohli et al., 2022) and, when their indices are in extreme states, their autocorrelation decay is on the order of years, permitting the signal to be leveraged for short-term prediction. However, based on the available observational and modeling evidence, the most plausible explanation for multidecadal peaks seen in the AMO and PDO is that they reflect the response to a combination of natural and anthropogenic forcings during the historical era (Mann et al., 2020). Despite the uncertainty of the source of the patterns, we find their signals useful for short-term drought prediction. In addition, in the Southern Great Plains, prior studies showed that low-frequency climate oscillations, such as AMO, ENSO, and PDO, were related to drought occurrences. We used this information to develop a set of multiple linear regression models to predict the PDSI using past-, present-, and future-forecasted climate indices and past and present PDSI as predictor variables. From the training data set, we built a second-order model equation that contained, without any restriction, all the predictors and interaction terms. Restriction was imposed through stepwise regression, excluding insignificant terms, and a set of 1000 regression models was produced using bootstrapping.

Results showed that the AMO index persistently appeared in each iteration as the leading predictor of PDSI. The MEI and PDO indices only come after the AMO index. The technique identified the best-fitting model (i.e., the "restricted model") that had the minimum sum of squared deviations between the fitted model and the predictor variables. The adjusted R-squared value and relatively lower values of RMSE and MAE of the restricted model

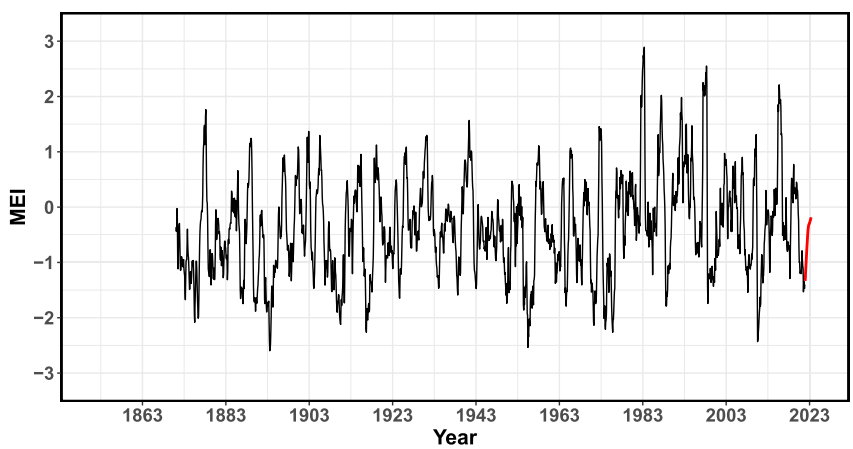

Figure 13. Same as Figure 12 but for Multivariate ENSO Index.

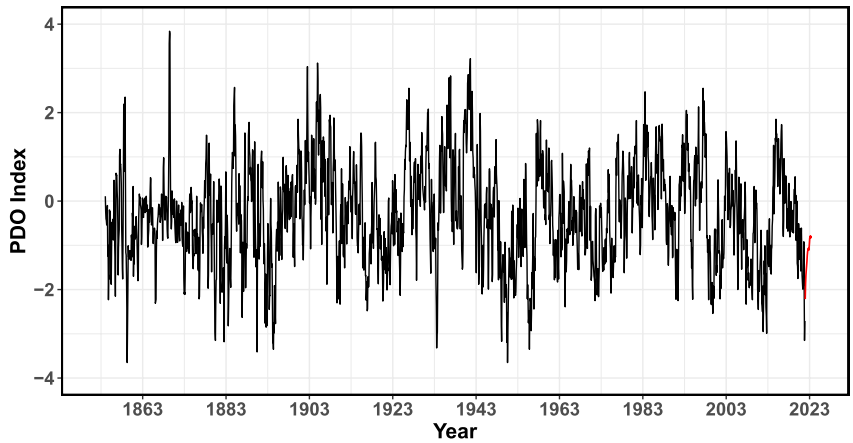

Figure 14. Same as Figure 12 but for Pacific Decadal Oscillation index.

explained the strength of the model and its good predictive ability. To generate a future PDSI forecast, we applied the ARIMA model to predict the climate indices (AMO, MEI, and PDO) out to 18 months. We added these future predictors to the historical data set from January 1901, resulting in the input to the 1000 restricted models. We averaged the results to produce the monthly PDSI forecast. Although the model slightly overestimated the PDSI forecast maxima and minima, it necessarily captures the timing of the periods of severe to exceptional drought. The high squared correlation coefficient of the observed and forecast values suggested that most of the variability of PDSI was explained by the predictors.

One limitation of the technique is the relatively large RMSE and MAE values of PDO index when applied to the ARIMA model to produce the predictors' predicted values for the 18-month forecast period. Considering the relationship of PDO with ENSO, PDO is not an independent predictor and can be dependent upon ENSO (Newman et al., 2016). Prior literature has shown that PDO forecast skill is influenced heavily by ENSO teleconnection (Alexander et al., 2008; Wen et al., 2012). In our study, we found the dependence to be only moderate correlation ($r = 0.461$) and therefore should have little impact on the model. Further, we expect this issue to cause

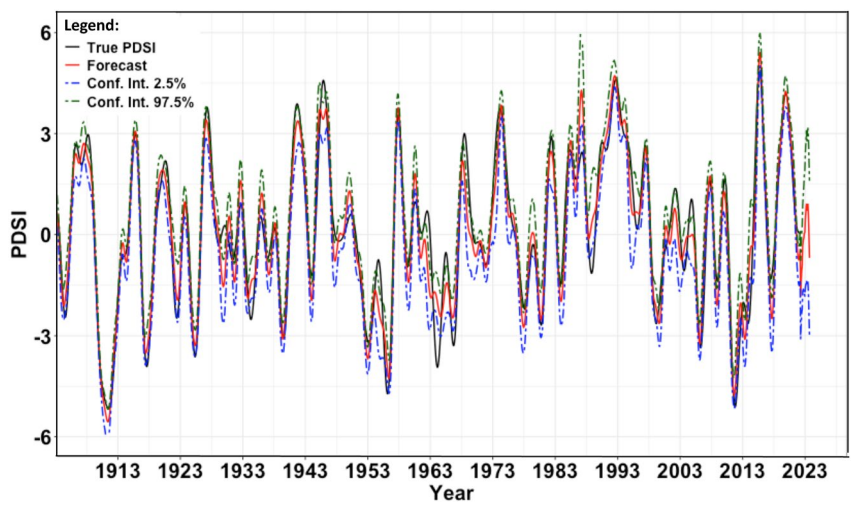

Figure 15. Palmer Drought Severity Index (PDSI) plots for Climate Division 8 (South Central) in Oklahoma from January 1901 to November 2021 for observed PDSI (black) and up to May 2023 for forecast PDSI (red), 2.5% percentile for the forecast (blue), and 97.5% percentile (green).

Table 6
Assessment of Models Performance Through Hindcasting

Forecast period	Adjusted R-squared	RMSE	MAE
October 2020–March 2022	0.802	0.959	0.526
October 1979–March 1981	0.904	0.687	0.519
February 1963–July 1964	0.884	0.796	0.340
October 1955–March 1957	0.681	1.341	0.476
November 1951–April 1953	0.877	0.826	0.347
March 2012–August 2013	0.832	0.877	0.531
April 2005–September 2006	0.922	0.599	0.458
May 1998–October 1999	0.843	0.866	0.525
July 1977–December 1978	0.891	0.741	0.529
November 2016–April 2018	0.857	0.809	0.513
Ensemble mean	0.849	0.850	0.476

Note. R-squared was calculated to assess the models' fit and RMSE and MAE to evaluate the models' errors.

only minimal impact to the overall forecast, as the PDO index is rarely a leading predictor in the restricted models. To address climate drift, we plan to update the model periodically to include the latest observation and forecasts of predictor variables. Future research plans include applying the forecast model to other climate divisions across the south-central United States and examining other climate indicators. We also seek an operational climate unit to host, operate, and maintain the data sets and modeling platform for free and open use by decision makers across the region.

Finally, it is important to note that this model is one of several products developed from years of interactions with staff and citizens of the Chickasaw Nation and the Choctaw Nation of Oklahoma. We intend for our method to have value broadly across the scientific community, but we were initially driven by listening to the needs of those coping with drought in south-central Oklahoma. As climate extremes become more extreme or erratic, it is more important every year that a larger portion of our community's science become actionable and to partner with decision makers (Bamzai-Dodson et al., 2021).

Data Availability Statement

The PDSI data were obtained from the National Centers for Environmental Information of NOAA, via their website: https://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers/. The AMO data were acquired from the Physical Science Laboratory of NOAA/ESRL at this website: https://www.psl.noaa.gov/data/timeseries/AMO/. The PDO data are available from the National Centers for Environmental Information website: https://www.ncdc.noaa.gov/teleconnections/pdo/. MEI data were obtained from the Physical Science Laboratory of NOAA/ESRL through their website: https://psl.noaa.gov/enso/mei/. All computer programs were written using the RStudio software archived at this site: http://www.rstudio.com/.

Acknowledgments
We gratefully acknowledge the Chickasaw Nation for its continued collaboration with the USGS South Central Climate Adaptation Science Center on drought and water resources planning. Funding for this project was provided by the National Aeronautics and Space Administration under Grant NNX15AK02H for the NASA Oklahoma Space Grant Consortium; NOAA-OAR-CPO-2011-2002561, awarded to the National Center for Atmospheric Research; and the Office of the Vice President for Research at the University of Oklahoma through the South Central Climate Adaptation Science Center. Financial support for publication was provided by the University of Oklahoma Libraries' Open Access Fund. The authors thank Virginia G. Silvis, Michael D. Hunter, and Heather Lazrus for their early work on the project and the reviewers of this manuscript for their helpful suggestions. The land on which the University of Oklahoma now resides was the traditional home of the "Hasinais" Caddo Nation and "Kirikir?i:s" Wichita and Affiliated Tribes. We acknowledge, honor, and respect the diverse Indigenous peoples who form all of Oklahoma's 39 sovereign tribal nations. We acknowledge the impacts of drought that these nations have endured while we strove to complete this work following academic practices and pressures, and we apologize for our delays.

References

Achite, M., Bazrafshan, O., Azhdari, Z., Walega, A., Krakauer, N., & Caloiero, T. (2022). Forecasting of SPI and SRI using multiplicative ARIMA under climate variability in a Mediterranean region: Wadi Ouahrane Basin, Algeria. *Climate*, *10*, 36. https://doi.org/10.3390/cli10030036

Alexander, M. A., Kilbourne, K. H., & Nye, J. A. (2013). Climate variability during warm and cold phases of the Atlantic Multidecadal Oscillation (AMO) 1871–2008. *Journal of Marine Systems*, *133*, 14–26. https://doi.org/10.1016/j.jmarsys.2013.07.017

Alexander, M. A., Matrosova, L., Penland, C., Scott, J. D., & Chang, P. (2008). Forecasting Pacific SSTs: Linear inverse model predictions of the PDO. *Journal of Climate*, *21*(2), 385–402. https://doi.org/10.1175/2007jcli1849.1

Alley, W. M. (1984). The Palmer Drought Severity Index: Limitations and assumptions. *Journal of Climate and Applied Meteorology*, *23*(7), 1100–1109. https://doi.org/10.1175/1520-0450(1984)023<1100:tpdsil>2.0.co;2

Bamzai-Dodson, A., Cravens, A. E., Wade, A. A., & McPherson, R. A. (2021). Engaging with stakeholders to produce actionable science: A framework and guidance. *Weather, Climate, and Society*, *13*(4), 1027–1041. https://doi.org/10.1175/wcas-d-21-0046.1

Box, G. E. P., & Jenkins, G. (1970). *Time series analysis: Forecasting and control* (p. 709). Holden-Day.

Bruce, P., & Bruce, A. (2017). *Practical statistics for data scientists: 50 essential concepts* (p. 298). O'Reilly Media.

Burke, E. J., Brown, S. J., & Christidis, N. (2006). Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model. *Journal of Hydrometeorology*, *7*(5), 1113–1125. https://doi.org/10.1175/jhm544.1

Cancelliere, A., Mauro, G. D., Bonaccorso, B., & Rossi, G. (2007). Drought forecasting using the standardized precipitation index. *Water Resource Management*, *21*(5), 801–819. https://doi.org/10.1007/s11269-006-9062-y

Chen, Z., Zhiping, W., Wu, R., Lin, X., & Wang, J. (2016). Relative importance of tropical SST anomalies in maintaining the western North Pacific anomalous anticyclone during El Niño to La Niña transition years. *Climate Dynamics*, *46*(3–4), 1027–1041. https://doi.org/10.1007/s00382-015-2630-1

Cole, J. E., Overpeck, J. T., & Cook, E. R. (2002). Multiyear La Niña events and persistent drought in the contiguous United States. *Geophysical Research Letters*, *29*(13), 1647. https://doi.org/10.1029/2001GL013561

Corporal-Lodangco, I. L., Riley, R., & Fujan, G. (2015). *Historical drought analysis of Oklahoma* (p. 9). Southern Climate Impacts Planning Program. Retrieved from http://www.southernclimate.org/pages/past-research/historical-drought-analysis-by-oklahoma-climate-division

Cuesta, J. E. C., & Hunt, J. K. (2019). Weekly climate indices: Generation and prediction. *Stanford Research Series*. Retrieved from https://medium.com/comet-ml/stanford-research-series-weekly-climate-indices-generation-and-prediction-8e548babb3a7

Dai, A. (2011). Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. *Journal of Geophysical Research*, *116*(D12), D12115. https://doi.org/10.1029/2010jd015541

Dai, A. (Ed.), & the National Center for Atmospheric Research Staff (Eds.). (2019). *The climate data guide: Palmer Drought Severity Index (PDSI)*. Retrieved from https://climatedataguide.ucar.edu/climate-data/palmer-drought-severity-index-pdsi

Dai, A., Trenberth, K. E., & Karl, T. R. (1998). Global variations in droughts and wet spells: 1900–1995. *Geophysical Research Letters*, *25*(17), 3367–3370. https://doi.org/10.1029/98gl52511

Dai, A., Trenberth, K. E., & Qian, T. T. (2004). A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming. *Journal of Hydrometeorology*, *5*(6), 1117–1130. https://doi.org/10.1175/jhm-386.1

Deng, S., Tan, X., Liu, B., Yang, F., & Yan, T. (2022). A reversal in global occurrences of flash drought around 2000 identified by rapid changes in the standardized evaporative stress ratio. *Science of the Total Environment*, *848*, 157427. https://doi.org/10.1016/j.scitotenv.2022.157427

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427–431. https://doi.org/10.2307/2286348

Diodato, N., De Guenni, L. B., Garcia, M., & Bellocchi, G. (2019). Decadal oscillation in the predictability of Palmer Drought Severity Index in California. *Climate*, *7*(1), 6. https://doi.org/10.3390/cli7010006

Enfield, D. B., Mestas-Nunez, A. M., & Trimble, P. J. (2001). The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. *Geophysical Research Letters*, *28*(10), 2077–2080. https://doi.org/10.1029/2000gl012745

Frajka-Williams, E., Beaulieu, C., & Duchez, A. (2017). Emerging negative Atlantic Multidecadal Oscillation index in spite of warm subtropics. *Scientific Reports*, *7*(1), 11224. https://doi.org/10.1038/s41598-017-11046-x

Glantz, M. (Ed.). (1994). *Usable science: Food security, early warning, and El Niño. Proceedings of the workshop on ENSO/FEWS, Budapest, Hungary, October 1993* (p. 250). UNEP, NCAR.

Glantz, M. H., Katz, R. W., & Nicholls, N. (Eds.). (1991). *Teleconnections linking worldwide climate anomalies. Scientific basis and societal impact*. Cambridge University Press.

Guttman, N. V., & Quayle, R. G. (1996). A historical perspective of U.S. climate divisions. *Bulletin of the American Meteorological Society*, *77*(2), 293–303. https://doi.org/10.1175/1520-0477(1996)077<0293:ahpouc>2.0.co;2

Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R., & Spade, E. R. (2003). A quantitative evaluation of ENSO indices. *Journal of Climate*, *16*(8), 1249–1258. https://doi.org/10.1175/1520-0442(2003)16<1249:aqeoei>2.0.co;2

Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, *56*(1), 108–141. https://doi.org/10.1002/2016RG000549

Hoerling, M., Eischeid, J., Kumar, A., Leung, R., Mariotti, A., Mo, K., et al. (2014). Causes and predictability of the 2012 Great Plains drought. *Bulletin of the American Meteorological Society*, *95*(2), 269–282. https://doi.org/10.1175/bams-d-13-00055.1

Hu, Q., & Feng, S. (2008). Variation of the North American summer monsoon regimes and the Atlantic multidecadal oscillations. *Journal of Climate*, *21*(11), 2371–2383. https://doi.org/10.1175/2007jcli2005.1

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). *An introduction to statistical learning: With applications in R*. Springer Publishing Company.

Kurtzman, D., & Scanlon, B. R. (2007). El Niño–Southern Oscillation and Pacific Decadal Oscillation impacts on precipitation in the southern and central United States: Evaluation of spatial distribution and predictions. *Water Resources Research*, *43*(10), W10427. https://doi.org/10.1029/2007WR005863

Lai, Y., & Dzombak, D. A. (2020). Use of the autoregressive integrated moving average (ARIMA) model to forecast near-term regional temperature and precipitation. *Weather and Forecasting*, *35*(3), 959–976. https://doi.org/10.1175/waf-d-19-0158.1

Lavaysse, C., Stockdale, T., McCormick, N., & Vogt, J. (2020). Evaluation of a new precipitation-based index for global seasonal forecasting of unusually wet and dry periods. *Weather and Forecasting*, *35*(4), 1189–1202. https://doi.org/10.1175/waf-d-19-0196.1

Mann, M. E., Steinman, B. A., & Miller, S. K. (2020). Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nature Communications*, *11*(1), 49. https://doi.org/10.1038/s41467-019-13823-w

Mantua, N. J. (1999). The Pacific Decadal Oscillation. A brief overview for non-specialists [Dataset]. Chang. Retrieved from https://www.ncdc.noaa.gov/teleconnections/pdo/

Mantua, N. J., & Hare, S. R. (2002). The Pacific Decadal Oscillation. *Journal of Oceanography*, *58*(1), 35–44. https://doi.org/10.1023/A:1015820616384

McCabe, G. J., Palecki, M. A., & Betancourt, J. L. (2004). Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences*, *101*(12), 4136–4141. https://doi.org/10.1073/pnas.0306738101

Mo, K. C., Schemm, J. E., & Yoo, S. (2009). Influence of ENSO and the Atlantic multidecadal oscillation on drought over the United States. *Journal of Climate*, *22*, 5962–5982. https://doi.org/10.1175/2009jcli2966.1

Murat, M., Malinowska, I., Gos, M., & Krzyszczak, J. (2018). Forecasting daily meteorological time series using ARIMA and regression models. *International Agrophysics*, *32*(2), 253–264. https://doi.org/10.1515/intag-2017-0007

National Oceanic and Atmospheric Administration-Atlantic Oceanographic and Meteorological Laboratory (NOAA-AOML). Retrieved from https://www.aoml.noaa.gov/phod/amo_faq.php

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., et al. (2016). The Pacific Decadal Oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427. https://doi.org/10.1175/jcli-d-15-0508.1

Nguyen, P.-L., Min, S.-K., & Kim, Y.-H. (2021). Combined impacts of the El Niño-Southern Oscillation and Pacific Decadal Oscillation on global droughts assessed using the standardized precipitation evapotranspiration index. *International Journal of Climatology*, *41*(S1), E1645–E1662. https://doi.org/10.1002/joc.6796

Nooteboom, P. D., Feng, Q. Y., López, C., Hernández-García, E., & Dijkstra, H. A. (2018). Using network theory and machine learning to predict El Niño. *Earth System Dynamics*, *9*(3), 969–983. https://doi.org/10.5194/esd-9-969-2018

Oklahoma Climatological Survey. (2014). Map of Oklahoma climate divisions. Retrieved from http://climate.ok.gov/index.php/climate/map/map_of_oklahoma_climate_divisions/oklahoma_climate

Palmer, W. C. (1965). *Meteorological drought. Weather Bureau Research Paper* (Vol. 45, pp. 1–58). U.S. Weather Bureau, NOAA Library and Information Services Division.

Quiring, S. M. (2009). Developing objective operational definitions for monitoring drought. *Journal of Applied Meteorology and Climatology*, *48*(6), 1217–1229. https://doi.org/10.1175/2009jamc2088.1

Rohli, R. V., Snedden, G. A., Martin, E. R., & DeLong, K. L. (2022). Impacts of ocean-atmosphere teleconnection patterns on the south-central United States. *Frontiers of Earth Science*, *10*, 934654. https://doi.org/10.3389/feart.2022.934654

Rosmiati, R., Liliasari, S., Tjasyono, B., & Ramalis, T. R. (2021). Development of ARIMA technique in determining the ocean climate prediction skills for pre-service teacher. *Journal of Physics*, *1731*(1), 012072. https://doi.org/10.1088/1742-6596/1731/1/012072

RStudio Team. (2020). RStudio: Integrated development for R. RStudio [Software]. PBC, Retrieved from http://www.rstudio.com/

Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., & Bacmeister, J. T. (2004). Causes of long-term drought in the U.S. Great Plains. *Journal of Climate*, *17*(3), 485–503. https://doi.org/10.1175/1520-0442(2004)017<0485:coldit>2.0.co;2

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Seager, R. (2007). The turn of the century North American drought: Global context, dynamics, and past analogs. *Journal of Climate*, *20*(22), 5527–5552. https://doi.org/10.1175/2007jcli1529.1

Seager, R., & Hoerling, M. (2014). Atmosphere and ocean origins of North American droughts. *Journal of Climate*, *27*(12), 4581–4606. https://doi.org/10.1175/jcli-d-13-00329.1

Seager, R., Kushnir, Y., Herweijer, C., Naik, N., & Velez, J. (2005). Modeling of tropical forcing of persistent droughts and pluvials over western North America: 1856–2000. *Journal of Climate*, *18*(19), 4065–4088. https://doi.org/10.1175/jcli3522.1

Singleton, A. (2012). *Forecasting drought in Europe with the standardized precipitation index: An assessment of the performance of the European Centre for Medium range weather forecasts variable resolution ensemble prediction system* (Vol. 68). Joint Research Center Scientific and Technical Reports.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., et al. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190. https://doi.org/10.1175/1520-0477-83.8.1181

Trenberth, K. (Ed.), & National Center for Atmospheric Research Staff (Eds.). (2020). *The climate data guide: Nino SST indices (Nino 1+2, 3, 3.4, 4; ONI and TNI)*. Retrieved from https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-Nino-12-3-34-4-oni-and-tni

van der Schrier, G., Barichivich, J., Briffa, K. R., & Jones, P. D. (2013). A scPDSI-based global data set of dry and wet spells for 1901–2009. *Journal of Geophysical Research: Atmospheres*, *118*(10), 4025–4048. https://doi.org/10.1002/jgrd.50355

Veres, M. C., & Hu, Q. (2013). AMO-forced regional processes affecting summertime precipitation variations in the central United States. *Journal of Climate*, *26*(1), 276–290. https://doi.org/10.1175/jcli-d-11-00670.1

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2011). Comment on "Characteristics and trends in various forms of the Palmer Drought Severity Index (PDSI) during 1900–2008" by Aiguo Dai. *Journal of Geophysical Research*, *116*(D19), D19112. https://doi.org/10.1029/2011JD016410

Wang, S., Huang, J., He, Y., & Guan, Y. (2014). Combined effects of the Pacific Decadal Oscillation and El Niño-Southern Oscillation on global land dry–wet changes. *Scientific Reports*, *4*(1), 6651. https://doi.org/10.1038/srep06651

Wen, C., Xue, Y., & Kumar, A. (2012). Seasonal prediction of North Pacific SSTs and PDO in the NCEP CFS hindcasts. *Journal of Climate*, *25*(17), 5689–5710. https://doi.org/10.1175/JCLI-D-11-00556.1

White, W. B., Gershunov, A., & Annis, J. (2008). Climatic influences on Midwest drought during the twentieth century. *Journal of Climate*, *21*(3), 517–531. https://doi.org/10.1175/2007jcli1465.1

Wolter, K., & Timlin, M. S. (2011). El Niño/Southern Oscillation behavior since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *International Journal of Climatology*, *31*(7), 1074–1087. https://doi.org/10.1002/joc.2336

Yoon, J., Mo, K., & Wood, E. (2012). Dynamic-model-based seasonal prediction of meteorological drought over the contiguous United States. *Journal of Hydrometeorology*, *13*(2), 463–482. https://doi.org/10.1175/jhm-d-11-038.1