

Research papers

Quantifying thresholds for advancing impact-based drought assessment using classification and regression tree (CART) models

Anoop Valiya Veettil^{a,b}, Ashok K. Mishra^{b,*}

^a Cooperative Agricultural Research Center, College of Agriculture and Human Sciences, Prairie View A&M University, Prairie View, TX 77446, USA

^b Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA

ARTICLE INFO

Keywords:

Drought indices

Decision tree

Threshold

Impact-based assessments

ABSTRACT

Drought events can significantly impact water resources, agriculture, and socioeconomic sectors. Droughts are often quantified using numerous indices, which are widely used for forecasting, risk assessments, and spatio-temporal analysis. Quantifying the appropriate linkage between drought indices and their impact on hydrological and agricultural indicators will improve hazard communication with stakeholders and further advance impact-based forecasting tools. This study aims to quantify the thresholds associated with drought indices that can capture the intricate connection between droughts and impact-specific indicators. We investigated the performance of several drought indices, such as Palmer-based (e.g., PDSI, PMDI, and PHDI) and multiscale drought indices (e.g., SPI and SPEI) with the impact-specific indicators of hydrological (e.g., streamflow and reservoir level) and agricultural (soil moisture and crop yield) indicators. The 'threshold' values associated with drought indices with impact-specific indicators are quantified using the classification and regression tree (CART) algorithm. Our results suggest that the decision tree approach is suitable for identifying critical thresholds of drought indices associated with a range of hydrological and agricultural indicators relevant to the stakeholder's application to different sectors. The drought characteristics derived from selected indices broadly vary during extreme events. The results further indicate that the threshold associated with each drought index and the impact-specific indicator varies drastically within the same climate division. We argue that quantifying the drought indices thresholds can provide valuable information for impact-based monitoring and forecasting.

1. Introduction

Drought is a complex natural hazard, considered the costliest natural disaster affecting society and the environment (Mukherjee et al., 2018; Mishra and Singh, 2010). Climate change significantly alters the spatiotemporal distribution of precipitation, evapotranspiration, and soil moisture (IPCC, 2021; Sheffield et al., 2012; Veettil and Mishra, 2020; Konapala et al., 2020). Such change in hydrologic fluxes will further enhance the impact of drought on different sectors. In addition, extreme drought cascades into multiple socio-environmental systems, such as wildfire (Long et al., 2021; Mukherjee et al., 2021), water scarcity (Veettil and Mishra, 2020; Veettil et al., 2022b), and loss of crop yield and livestock (Chawla et al., 2020; Rajsekhar et al., 2015), and conflicts (García-Herrera et al., 2010; Hsiang et al., 2013).

Droughts are generally categorized into meteorological, agricultural, hydrological, and socioeconomic droughts (Wilhite, 2000; Mishra and Singh, 2010). These droughts are classified based on the extended

anomalies (deficit) in precipitation as meteorological droughts, soil moisture as agricultural drought, and surface or subsurface water supply (e.g., streamflow and aquifer storage) as hydrological drought. Socio-economic drought occurs when the water supply from the regional water resources system cannot meet the water demands of economic goods (Rajsekhar et al., 2015). Quantifying the evolution of a drought event is often challenging because of the complex interaction between climate and catchment variables (Konapala and Mishra, 2020; Van Loon and Laha, 2015). Traditionally, different drought indices are commonly used as a proxy to study drought events (Pedro-Monzonís et al., 2015; Mishra and Singh, 2010). These drought indices are based on long-term anomalies using a combination of different hydrologic fluxes (Mishra and Singh, 2010).

Over the past decades, over a hundred drought definitions and indices have been implemented in different parts of the world. Among them, the commonly used indicators include Palmer Drought Severity Index (PDSI, Palmer, 1968), Palmer Modified Drought Index (PMDI,

* Corresponding author.

E-mail address: ashokm@clemson.edu (A.K. Mishra).

Bhalme and Mooley, 1979), Palmer Hydrologic Drought Index (PHDI, Guttman, 1991), Z-Index (ZNDX, Karl, 1986), Standardized Precipitation Index (SPI, Guttman, 1999; McKee et al., 1993), Standardized Precipitation and Evapotranspiration Index (SPEI, Vicente-Serrano et al., 2010), multivariate drought index (Hao and AghaKouchak, 2013; Rajsekhar et al., 2015). These indices are commonly used for different types of drought investigation, such as drought forecasting (Cancelliere et al., 2007; Christian et al., 2021), spatio-temporal drought analysis (Veetil et al., 2022a; Mishra et al., 2021), non-stationary frequency analysis (Das et al., 2020), climate change impact assessments (Mukherjee et al., 2018; Kreibichet al., 2022), multivariate drought analysis, agricultural yields, and reservoir management (Labędzki and Bąk, 2014; Wu et al., 2017; Gavahi et al., 2020). Developing drought monitoring tools using existing drought indices can be critical in managing water resources for different sectors (Wilhite et al., 2007; Mishra and Singh, 2011).

Although the drought indices (e.g., SPI, SPEI, PDSI) are valuable and useful proxy for a variety of applications (e.g., monitoring, forecasting, risk assessments), there are several limitations associated with such indicators, such as:

- (a) *Lack of integration of drought indices and impacts*: existing drought indices are derived using a suitable probability distribution, and the drought events are classified (e.g., extreme, severe, moderate) based on standardized normal variates without capturing the actual drought impacts on different sectors (Mishra and Singh, 2010, 2011).
- (b) *Lack of proper communication with stakeholders*: the numeric values associated with drought indices do not quantify the level of drought impacts on different sectors (e.g., agriculture, reservoir management); as a result, there is difficulty in communicating with stakeholders, for example, interpreting SPEI = 2 for a farmer (rancher) can be challenging.
- (c) *Lack of impact-based forecasting tools*: There are inconsistencies and uncertainties while comparing drought indices based on different variables and data sets for impact assessments. Generating drought impact information ahead of time is a prerequisite for drought management is still missing, and the operational drought early warning systems should include drought hazards and impacts (Sutanto et al., 2019). A limited number of studies evaluated the links between reported drought impacts and drought hazards derived from SPI and SPEI (Blauhut et al., 2015; Bachmair et al., 2017; Stage et al., 2015).

To address the above limitations, we argue that the drought indices can be valuable for stakeholders by identifying appropriate 'thresholds' that directly link with the drought impact-based assessment for advancing drought monitoring and forecasting. Quantifying the drought threshold can advance systematic estimation of the linkage between drought severity and drought impact associated with different types of droughts (e.g., hydrological and agricultural droughts) that are often connected through hierarchical and non-linear relationships. The specific objectives of this study are 1) to investigate the relationship between the multiple drought indices and hydrological and agricultural impact-specific drought indicators and 2) to identify the critical threshold that directly links drought indices and a range of hydrological and agricultural indicator variables. Impact-based indicators are quantified based on streamflow, reservoir level (hydrological response), soil moisture, and corn yield (agricultural response). We applied the classification and regression tree (CART) concept for identifying drought thresholds that can be useful for the stakeholders and decision-makers to develop necessary action for minimizing drought impacts.

The manuscript is organized as follows: section 2 discusses the study area and Data; Section 3 discusses the methodology, which includes an explanation of different drought indices (Palmer-based indices, standardized precipitation index, and standardized precipitation and runoff

index) and classification and regression trees; section 4 describes the results of the relationship between drought indices and the threshold limits of each drought index. Section 5 provides the discussion and concluding remarks.

2. Study area and data

The state of South Carolina (SC) is located in the Southeastern United States (US) with an area of 82,930 km² (32,020 mi²) (Fig. 1). The primary land cover of the state consists of forest (56%), agriculture (17%), and developed land (13.5%). South Carolina receives an annual average rainfall of 1220 mm of annual precipitation (Billah and Goodall, 2011; SCDHEC, 2010), and the coastal region receives relatively more rainfall during the summer, while the remaining state receives more rain during the spring. The distribution of monthly precipitation from 1950 to 2017 in each climate division is shown in Fig. 1. The annual average temperature varies from 10 °C in the mountains to 16 °C along the coast. Usually, January is the coldest month, and July is the hottest month in South Carolina.

In recent decades, the state has been going through a periodic water shortage due to drought and growing water consumption (Veetil and Mishra, 2016). For example, the southeast United States experienced significant drought from 1965 to 1971, 1980 to 1982, 1985 to 1988, 1998 to 2002, and 2006 to 2009 (Veetil and Mishra, 2016, 2020). The major rivers, such as Savannah, Pee Dee, Santee, and Edisto, play a critical role in meeting the state's water demand. Most of these river basins share a boundary with neighboring states except the Edisto River Basin (Badr et al., 2004).

There are 14 major reservoirs providing water for different sectors such as irrigation, domestic, and industry. Lake Strom Thurmond (area: 290 km²), located in the Savannah River Basin, is the largest reservoir in the state. According to the SC Department of Agriculture, the major crops in the state are corn, soybean, cotton, peanuts, tobacco, and wheat, grown over 1.3 million acres. Four general soil groups found in the state are (1) loamy, clayey soils of the wet lowlands; (2) wet, sandy soils found in strips near the coast; (3) well-mixed soils of river floodplains, underlain with sandy, loamy sediments; and (4) clayey, sandy soils of coastal-area salt marshes and dunes.

2.1. Data

We obtained two types of data sets, drought indices, and impact-based indicators. We evaluated the performance of five drought indices which include the Palmer Modified Drought Index (PMDI), Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Standardized Precipitation Index (SPI), and Standardized Precipitation and Evapotranspiration Index (SPEI). The Palmer drought indices and the SPI data sets for the seven climate divisions were obtained from National Oceanic and Atmospheric Administration (NOAA). NOAA provides a long-term drought index at a divisional scale across the United States. The SPEI data obtained from Climatic Research Unit (CRU), available online at <http://badc.nerc.ac.uk/data/cru/> (Mitchell and Jones, 2005) for the same period at a spatial resolution of 0.5°. The streamflow records for the 18 locations were obtained from US Geological Survey's (USGS) current water data for the nation. The reservoir level data for three major reservoirs (Hartwell, Russel, and Thurmond) were obtained from the US Army Corps of Engineers. The higher water level is observed in Hartwell Reservoir, with an average of 200.25 m, and the lower level of the reservoir is observed in Thurmond, with an average of 108.8 m. The yearly corn-yield data for each climate division was obtained from USDA National Agricultural Statistics Service (NASS), available at <https://quickstats.nass.usda.gov/>. The observed soil moisture data was obtained from Soil Climate Analysis Network (SCAN) and US Climate Reference Network (CRN). The monthly drought indices are analyzed from 1950 to 2017. The impact-specific indicators used in this study are: (a) streamflow (1950 –

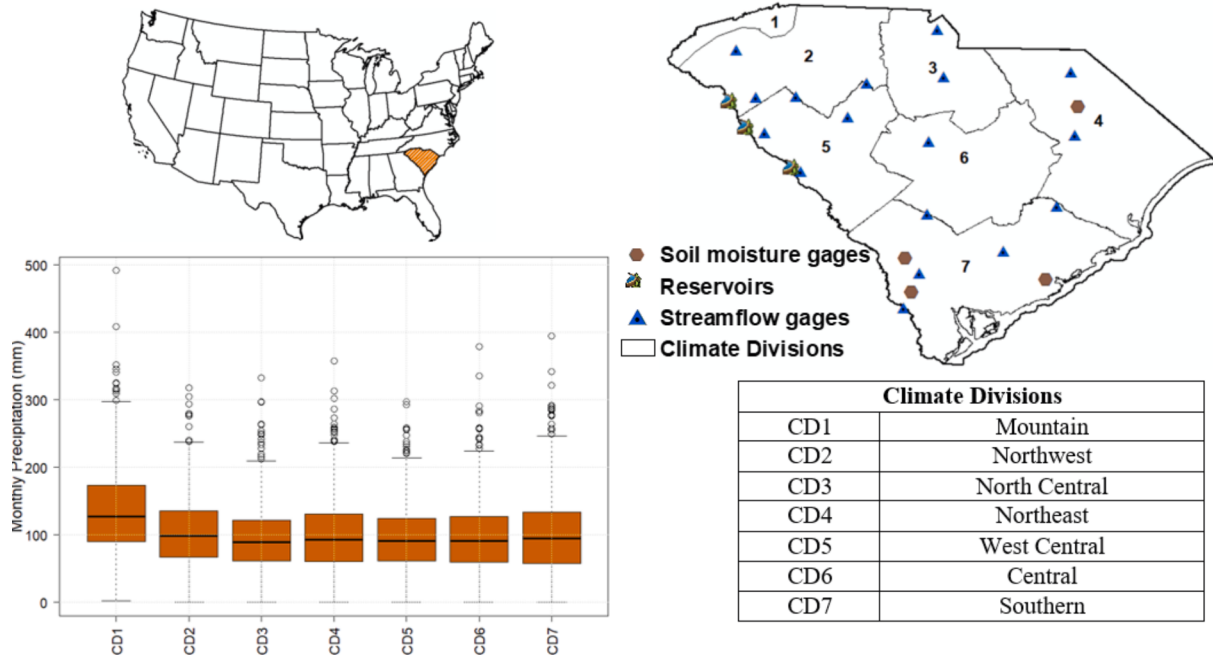


Fig. 1. Location of seven climate divisions within the state of South Carolina (USA). The boxplot represents the monthly precipitation patterns from 1950 to 2017 across the selected climate divisions. The table provides the name of each climate division corresponding to the number.

2017); (b) Three reservoirs level used in this study are Hartwell Reservoir (1962 – 2017), Thurmond Reservoir (1954 – 2017), Russel Reservoir (1985 – 2017); (c) four soil moisture stations used in this study are SM1 (1999 – 2017), SM2 (1999 – 2017), SM3 (2009 – 2017), and SM4 (2009 – 2017), and (d) Corn yield data (1950 – 2017).

3. Methodology

3.1. Palmer drought indices: ZNDX, PDSI, PHDI, and PMDI

Three Palmer's drought indices considered in this study include Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Palmer Modified Drought Index (PMDI), and Z-Index (ZNDX). PDSI can measure both wetness (positive value) and dryness (negative values) based on the supply and demand concepts of the water balance equation (Palmer, 1965), and it is useful for quantifying and comparing drought across different regions. PDSI is calculated based on temperature, precipitation, and the soil's available water capacity (AWC) (Jacobi et al., 2013). Before calculating monthly values of PDSI, several parameters are derived from the input data sets by calculating climatically appropriate quantities for an existing condition. Other information used in the Palmer indices are the soil's available water capacity (AWC) and the Thornthwaite potential evapotranspiration.

$$\widehat{ET} = \alpha(PE) \quad (1)$$

$$\widehat{R} = \beta(PR) \quad (2)$$

$$\widehat{RO} = \gamma P(RO) \quad (3)$$

$$\widehat{L} = \delta(PL) \quad (4)$$

$$P = \widehat{ET} + \widehat{R} + \widehat{RO} - \widehat{L} \quad (5)$$

where ET is the evapotranspiration, PE is the potential evapotranspiration, R is the soil water recharge, PR is the potential recharge, RO is the runoff, L is the water loss from the soil, PL is the potential water loss from the soil, and P is the precipitation. The parameters α , β , γ , and δ for each of the 12 months at each location are determined as follows, where

the overscore quantity within parenthesis denotes the monthly average throughout the record.

$$\alpha = (\overline{ET})(\overline{PE})^{-1} \quad (6)$$

$$\beta = (\overline{R})(\overline{PR})^{-1} \quad (7)$$

$$\gamma = (\overline{RO})(\overline{PRO})^{-1} \quad (8)$$

$$\delta = (\overline{L})(\overline{PL})^{-1} \quad (9)$$

These coefficients are used to compute the differences, d, for each month between the actual precipitation for the month, P, and the $\widehat{P}$ for the existing condition. Here d is regarded as moisture departed from the normal.

$$d = P - \widehat{P} \quad (10)$$

$$= P - (\alpha_j PE + \beta_j PR + \gamma_j PRO - \delta_j PL) \quad (11)$$

The standardized index is calculated to account for regional differences (d), as shown below,

$$Z = d * K \quad (12)$$

Here, Z is the Z-index (ZNDX), which is regarded as the soil moisture anomaly index. The ZNDX is a short-term drought measure without additional memory from previous months. Where K is a weighting factor for a given month used to adjust the soil moisture departures (d), to make it comparable among different areas and months.

The value of the PDSI is computed from the moisture anomaly index. Also, PDSI may abruptly change from dry to near-normal conditions, even though rivers or groundwater storage may be considerably below their normal levels. The PDSI for the i^{th} month is computed as follows,

$$PDSI_i = PDSI_{i-1} + \frac{1}{3}Z_i - 0.103PDSI_{i-1} \quad (13)$$

The Palmer Hydrological Drought Index (PHDI) was derived from the PDSI to quantify the long-term impact of drought on hydrological systems. PHDI values tend to be negative for up to several months after PDSI has returned to normal levels (i.e., it usually returns to near-normal

levels more gradually than the PDSI (Karl et al., 1986)). Therefore, the PHDI is considered as a measure of long-term hydrological drought since streamflow, reservoir storage, and groundwater tend to stay below normal values for some time after a meteorological drought ends.

The PMDI is a modification of the PDSI for real-time operational purposes (Jacobi et al., 2013), which is obtained from the sum of the wet and dry terms weighted by probability values. The PMDI has the same value as the PDSI during dry or wet spells but differs during transition periods. Overall, the PDSI and PMDI are appropriate for monitoring meteorological drought.

3.2. Standardized precipitation index (SPI)

The standardized precipitation index (SPI) introduced by McKee et al. (1993) is derived based on the long-term precipitation record. The following steps are used to derive SPI for a region: a) long-term monthly precipitation is extracted from available sources (e.g., in-situ, remote sensing), b) a suitable probability distribution (e.g., gamma distribution) is identified to capture the long-term precipitation distribution, c) once the probability density function is determined, the cumulative probability of the precipitation time series at different time scale (e.g., 3, 6, 9 months) is computed; and (d) the inverse normal Gaussian function, with mean zero and standard deviation one, is then applied to the cumulative probability distribution function, which results in the SPI. Positive SPI values indicate greater than median precipitation, and negative values indicate less than median precipitation. McKee et al. (1993) used the gamma distribution to transform precipitation series into standardized units.

The advantages of using SPI are 1) simplicity: the SPI is based on rainfall; 2) its standardization, which ensures that the frequency of extreme events at any location and on any time scale is consistent; and 3) variable time scale, which is helpful for the analysis of drought dynamics, especially the determination of onset and cessation (Angelidis et al., 2012). On the other hand, the length of data and different probability distributions are likely to affect SPI values (Mishra and Singh, 2010).

3.3. Standardized precipitation and evapotranspiration index (SPEI)

The SPEI (Vicente-Serrano et al., 2010) is calculated based on the precipitation minus potential evapotranspiration time series derived over multiple time scales. The procedure implemented to derive the SPEI series is almost similar to SPI. Like the PDSI, SPEI addresses the warming-related impacts on different hydrological and agricultural systems by considering the temperature variation (via potential evapotranspiration) in the calculation. In contrast, PDSI lacks multiscale characteristics, which is essential for addressing the relationship between temporal drought and hydrological/ agricultural systems. The SPEI addresses changes in evaporation demand caused by temperature fluctuations and trends. This is an advantage over the SPI, where the calculation is only based on the precipitation records.

3.4. Statistical analysis

We investigated the linear relationship between various drought indices and impact-specific indicators (i.e., streamflow, reservoir level, soil moisture, and crop yields) using Pearson correlation (r) analysis. Relationship values range from -1 to $+1$, and there is no correlation between two variables when this value equals 0. For example, the Pearson correlation between reservoir level (RL) and drought indices (DI) is calculated as,

$$r = \frac{\sum_{k=1}^n (RL - RL_m)(DI - DI_m)}{\sqrt{\sum_{k=1}^n (RL - RL_m)^2 \sum_{k=1}^n (DI - DI_m)^2}} \quad (14)$$

where RL_m and DI_m represent the average reservoir level and drought

indices, respectively, and n is the number of data. Similarly, the Pearson correlation is calculated for streamflow, soil moisture, and crop yield.

3.5. Classification and regression trees

Classification and regression trees (CART) algorithms are used for classification and regression learning tasks. CART is a powerful approach commonly used in machine learning, where the data are partitioned into subsets with homogeneous values of the dependent variable (Krzywinski and Altman, 2017). The CART algorithm uses a hierarchical (tree structure), which includes a root node, branches, internal nodes, and leaf nodes (James et al., 2013). The CART algorithm has several advantages: (a) the hierarchical structure of CART models mimics the heuristics of decision-making and is more intuitive than traditional regression models (Drakopoulos, 1994), (b) the model outcomes are better than standard regression models, specifically during the presence of non-linear relationships and interactions (Varian, 2014). Also, CART produces easy-to-understand models with any combination of continuous/ discrete variables.

CART effectively represents the stepwise decision-making process of a complex system (Solomatine, 2002) by stratifying the predictor space into many simple regions based on the output variable (James et al., 2013, Veetil and Mishra, 2020). This supervised learning algorithm is easy to interpret and can be considered one of the most appropriate approaches for solving complex problems (James et al., 2013). Further, splitting rules that divide the variable input space into various classes are characterized as a tree. Therefore, these approaches are known as decision (or classification) tree methods. In this tree-like structure, the nodes generate a threshold for each drought index, which can capture a range of drought indicator variables. The CART algorithms are useful for threshold analysis, where the training set (root node) is split into two based on the best attribute and threshold value.

This study developed the CART models to identify the associated threshold between drought indices and impact-specific indicators (i.e., hydrological and agricultural indicators). A recursive partitioning algorithm, which classifies the space defined by the input variables (i.e., drought indices) based on the output variables (e.g., streamflow, soil moisture, and crop yield), is used for developing the CART. The decision trees for quantifying the threshold of drought indices associated with hydrological and agricultural indicators are formulated based on the following steps. (1) The response variables (i.e., hydrological and agricultural indicators) and input variables (i.e., drought indices) are first selected for a particular climate division of the South Carolina State; (ii) divide the input variable space $X_1, X_2, \dots, X_p$ into J discrete and non-intersecting regions, such as $R_1, R_2, \dots, R_J$; and (iii) for every variable that falls into the region R_j , the tree makes the same prediction, which is the mean of the drought indicator values in the region R_j . The purpose of dividing predictor space into different regions is to reduce the residual sum of squares (RSS). The RSS is calculated as follows,

$$RSS_{Min} = \sum_{j=1}^J \sum_{i \in R_j} (y_i - \hat{y}_{R_j})^2 \quad (15)$$

Where, $\hat{y}_{R_j}$ is the mean response of the response variable (hydrological and agricultural indicator values) within the j^{th} region. Generally, considering every possible partition of predictor feature space into distinct regions is computationally challenging. Therefore, the decision tree algorithm utilizes recursive binary splitting, which works based on a top-down approach by successfully splitting the predictor space represented by two new branches further down the tree. To perform the recursive binary splitting, we first selected the predictor X_j and the threshold ' s ' (equations (15) and (16)) so that splitting the predictor space can lead to the maximum reduction in RSS, given by equation (15).

$$R_1(j, s) = \{X | X_j < s\} \quad (16)$$

$$R_2(j, s) = \{X|X_j \geq s\} \quad (17)$$

The optimal thresholds are calculated based on the lowest RSS for the resulting decision tree. The tree output divides data into a series of nodes, and each node represents the range of hydrological and agricultural indicator variables in the form of a boxplot. The final tree provides thresholds associated with drought indices and significance (p-value) value. The threshold refers to the value at which a splitting condition is applied to partition the data into different branches. The threshold associated with each drought index may vary based on the RSS used to choose the split at each node. Therefore, this approach can quantify the threshold associated with each drought index and the corresponding range of hydrological and agricultural indicators using boxplot representation.

4. Results

4.1. Comparison between drought indices and the need for threshold approach

The CART algorithm was implemented to quantify threshold values of drought indices and their association with impact-specific indicators. The drought indices considered in the present study are PDSI, PMDI, PHDI, ZNDX, SPI (SPI3, SPI6, and SPI12), and SPEI (SPEI3, SPEI6, and SPEI12) for the period 1950 to 2017. To investigate the limitations associated with the drought indices, first, we quantified drought characteristics, such as average drought duration (DD), average drought severity (DS), and drought frequency (DF) for each index and climate division (Table 1). We did not include climate division 1 (CD1) in our

Table 1

The average drought duration (DD), severity (DS), and frequency (DF) for the selected climate divisions located in South Carolina, USA.

DT = -1		CD2	CD3	CD4	CD5	CD6	CD7
PDSI	DD	9.59	7.36	5.11	7.78	7.17	5.65
	DS	15.54	10.89	6.02	10.74	8.98	6.82
	DF	0.51	0.67	0.66	0.61	0.61	0.78
PHDI	DD	10.77	8.53	5.81	9.26	8.13	6.02
	DS	16.83	12.25	6.53	12.54	10.11	7.30
	DF	0.52	0.64	0.70	0.57	0.60	0.82
ZNDX	DD	1.84	1.83	1.63	1.91	1.75	1.67
	DS	1.95	1.89	1.53	1.73	1.60	1.48
	DF	2.19	2.24	2.16	2.16	2.16	2.30
PMDI	DD	6.38	5.63	3.90	5.07	4.64	4.23
	DS	10.58	8.18	4.53	7.07	5.74	5.47
	DF	0.75	0.85	0.88	0.88	0.91	0.97
SPI3	DD	2.42	2.24	1.90	2.05	1.97	1.95
	DS	1.34	1.16	0.87	1.00	1.00	0.94
	DF	0.79	0.93	0.94	0.94	0.94	0.93
SPI6	DD	3.81	2.89	2.87	2.80	2.58	2.95
	DS	1.95	1.50	1.41	1.29	1.53	1.52
	DF	0.55	0.69	0.57	0.69	0.64	0.61
SPI12	DD	5.55	4.97	3.91	4.09	3.97	3.47
	DS	2.58	2.42	2.45	2.04	2.26	2.09
	DF	0.43	0.46	0.34	0.48	0.46	0.45
SPEI3	DD	2.34	2.20	2.17	2.19	2.32	2.33
	DS	1.00	0.95	0.89	0.90	0.90	0.94
	DF	0.88	0.97	0.90	0.96	0.88	0.85
SPEI6	DD	3.18	2.69	2.92	3.24	2.88	3.38
	DS	1.42	1.18	1.38	1.26	1.26	1.40
	DF	0.67	0.76	0.54	0.69	0.63	0.60
SPEI12	DD	5.57	4.93	5.43	4.66	4.92	4.06
	DS	2.46	2.27	2.69	2.05	2.40	1.85
	DF	0.42	0.45	0.31	0.48	0.37	0.46

*DT: Drought Threshold; DD: Drought Duration; DS: Drought Severity; DF: Drought Frequency.

Table 2

The maximum drought duration calculated based on different drought indices for selected climate divisions from 2010 to 2014.

DT = -1	CD2	CD3	CD4	CD5	CD6	CD7
PDSI	25	34	23	34	25	23
PHDI	34	38	33	36	33	32
ZNDX	8	5	4	6	5	7
PMDI	25	38	21	35	28	22
SPI3	3	3	3	3	3	3
SPI6	5	5	3	10	2	14
SPI12	13	13	11	20	11	13
SPEI3	3	5	4	3	4	4
SPEI6	7	8	6	9	6	12
SPEI12	4	11	20	26	20	20

analysis due to its small geographical area, dominated by forests, and lack of impact-specific indicator data. Here, the drought characteristics are calculated based on the run theory (Yevjevich, 1967) for the events which exceed severity -1 . Drought duration is the duration of consecutive time series when the drought index is below the threshold value (-1). Severity is defined as the summation of the drought index below the threshold. Drought frequency can be defined as the number of occurrences of drought events exceeding a certain threshold (Veetil et al., 2018; Loucks and Van Beek, 2017). Here, we expressed drought frequency as the ratio of the total number of drought events to the total number of years (i.e., 67 years).

It was observed that the drought indices provide different information based on their characteristics (DD, DS, and DF) within the same climate division. For example, in climate division 2, the DD, DS, and DF values for PDSI were 9.59, 15.54, and 0.51, respectively. Whereas in the case of ZNDX, it was 1.84, 1.95, and 2.19, respectively. Similarly, the multivariate indices showed a noticeable difference in their drought characteristics. Throughout 2010–2014, most of the southeastern United States went through a severe to extreme drought event (Rippey, 2015). We demonstrated the difference between drought indices based on the drought durations calculated for climate divisions during the severe to extreme drought period (2010–2014) (Table 2). A significant variation in maximum drought duration within each climate division is observed. For example, the maximum drought duration in CD2 varied from three (SPI3 and SPEI3) to 34 (PHDI) months. These considerable disparities in drought characteristics will show distinct drought risk outcomes on various impact-specific indicators. The analysis based on Table 1 and Table 2 indicates that monitoring drought using different indices leads to various duration, severity, and frequency results. Also, multiple time steps of drought indices make it difficult to decide which time step best shows the drought condition. Therefore, quantifying the thresholds associated with drought indices for specific impact (i.e., streamflow, reservoir level, soil moisture, and crop yields) can advance the preparedness and planning to cope with the adverse effects of a regional drought event. The following sections describe the potential influence of drought quantified based on Palmer and multiscale indices on different hydrological (streamflow and reservoir level) and agricultural drought indicators (soil moisture and crop yield). Also, a trend analysis of drought indices and impact-specific indicators based on the Mann-Kendall (MK) test (Mann, 1945; Kendall, 1975) is performed, and the results are discussed in Appendix A.

4.2. Thresholds associated with drought indices for streamflow

Fig. 2 shows the boxplot illustrating the correlation coefficients between the monthly streamflow and seven drought indices. The correlation is calculated for the 18 flow-gaging stations distributed in the

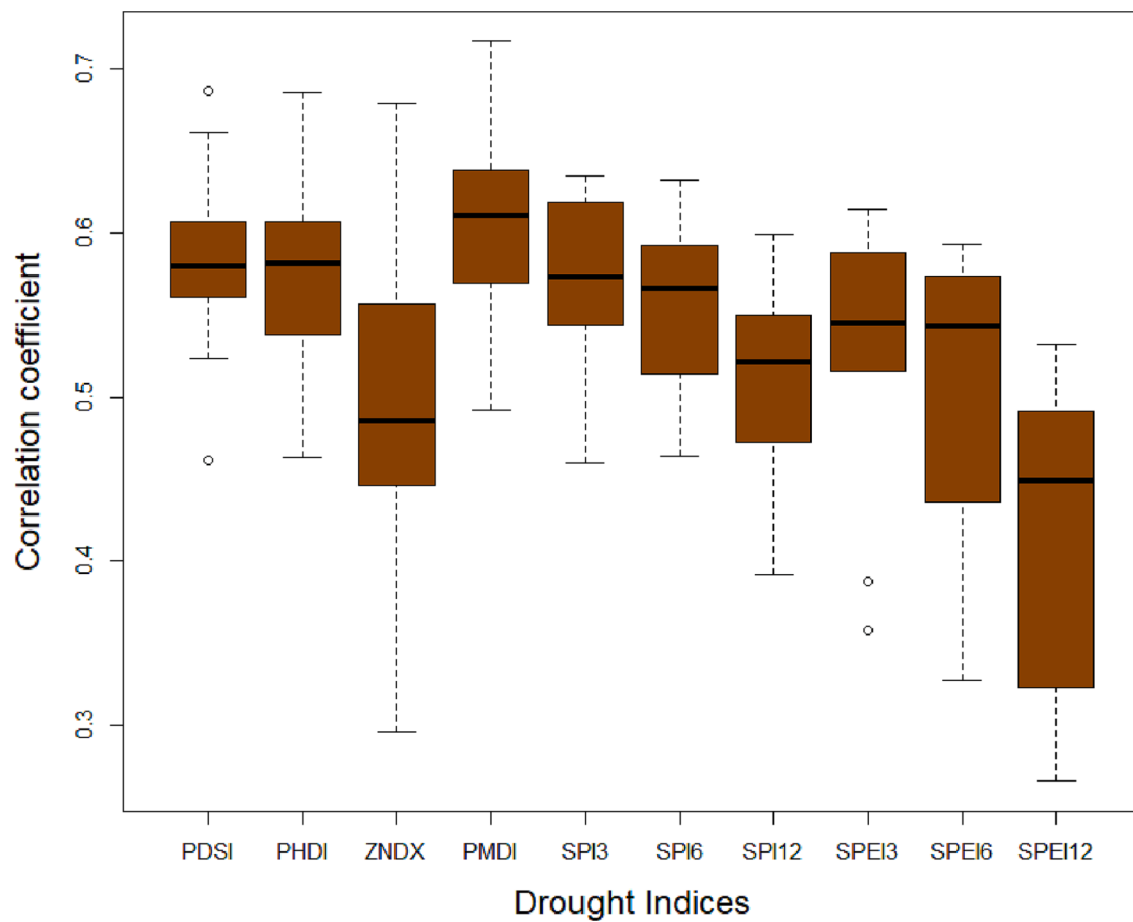


Fig. 2. Boxplot showing the correlation coefficient between streamflow and drought indices. Overall 18 streamflow gaging stations are selected from all the seven climate divisions for the correlation analysis.

different climatic divisions of the state. In general, it was observed that the correlation is higher for the Palmer indices, excluding the ZNDX, and PMDI tended to be higher among them. The median correlation value for PMDI was 0.62, followed by PHDI, with a correlation coefficient of 0.58. The SPI and SPEI showed a relatively low linkage between the streamflow records. For instance, SPI3 showed a median correlation of 0.57, and SPI6 and SPI12 showed a median value less than SPI3. In the case of SPEI, SPEI3 showed maximum correlation with the streamflow, whereas SPEI12 showed the least correlation in the analysis. The ZNDX and SPEI showed a median correlation of less than 0.5 with the streamflow over the climate divisions.

In the next step, we applied the classification and regression tree (CART) concept to identify the threshold associated with different drought indices for the monthly streamflow values. Using this approach, we generated four decision tree models for the monthly streamflow at the gage station USGS 02186000, located in the Northwest climate division (i.e., CD2), which has minimum anthropogenic impact and is not influenced by large water storage structures. In addition, these streamflow records showed the highest correlation, particularly for PHDI (0.66), PMDI (0.68), SPI3 (0.62), and SPEI3 (0.63). The model output of CART based on the drought indices that showed a higher correlation in the analysis is illustrated in Figs. 3 and 4. The figures summarize the process of estimating each drought index's threshold and the streamflow range (i.e., response variable) with respect to the threshold.

The correlation of PMDI with the monthly streamflow is higher than other drought indices considered in the analysis (Fig. 3a). Therefore, we

initially investigated the threshold of PMDI on streamflow. The threshold of the first split in the decision tree is 0.7. For instance, if the PMDI is less than or equal to 0.7, the tree's growth is towards the left. The tree advances towards the right side if it is higher than 0.7. Here $p < 0.001$ represents the significance of the correlation between the split based on PMDI and the monthly streamflow. The second split is based on a threshold of -2.51 , followed by a threshold of -3.94 , which indicates if the PMDI is less than or equal to -3.94 , the monthly average streamflow will vary from 1 to $5 \text{ m}^3/\text{s}$ with a median value of $3 \text{ m}^3/\text{s}$. Whereas, when the PMDI is greater than or equal to -3.94 , the growth of the tree is towards the right, and it showed a variation of flow from 2 to $6 \text{ m}^3/\text{s}$. Node 6 splits the tree with a threshold of -1 . For example, if the PMDI is less than or equal to -1 , the monthly average streamflow will vary from 3 to $8 \text{ m}^3/\text{s}$. Similarly, when the PMDI is greater than 0.7, the tree's growth is towards the right side. Maximum flow is observed when the PMDI is greater than 3.18, where the flow varies from 6 to $18 \text{ m}^3/\text{s}$ with a median flow of $12 \text{ m}^3/\text{s}$.

A similar analysis is performed for the drought indices PHDI (Fig. 3b), SPI3 (Fig. 4a), and SPEI3 (Fig. 4b). The maximum threshold for PHDI, SPI3, and SPEI3 were 3.15, 1.45, and 1.59, respectively. The total number of nodes in the PMDI and PHDI decision tree output is 13, whereas the SPI3 and SPEI3 split the variable's space to 11 (i.e., the total number of nodes = 11). In addition, the initial split of the tree is 1.48 based on PHDI and 0.37 based on SPI3, and the SPEI3 decision tree showed the least value of the initial split in the analysis.

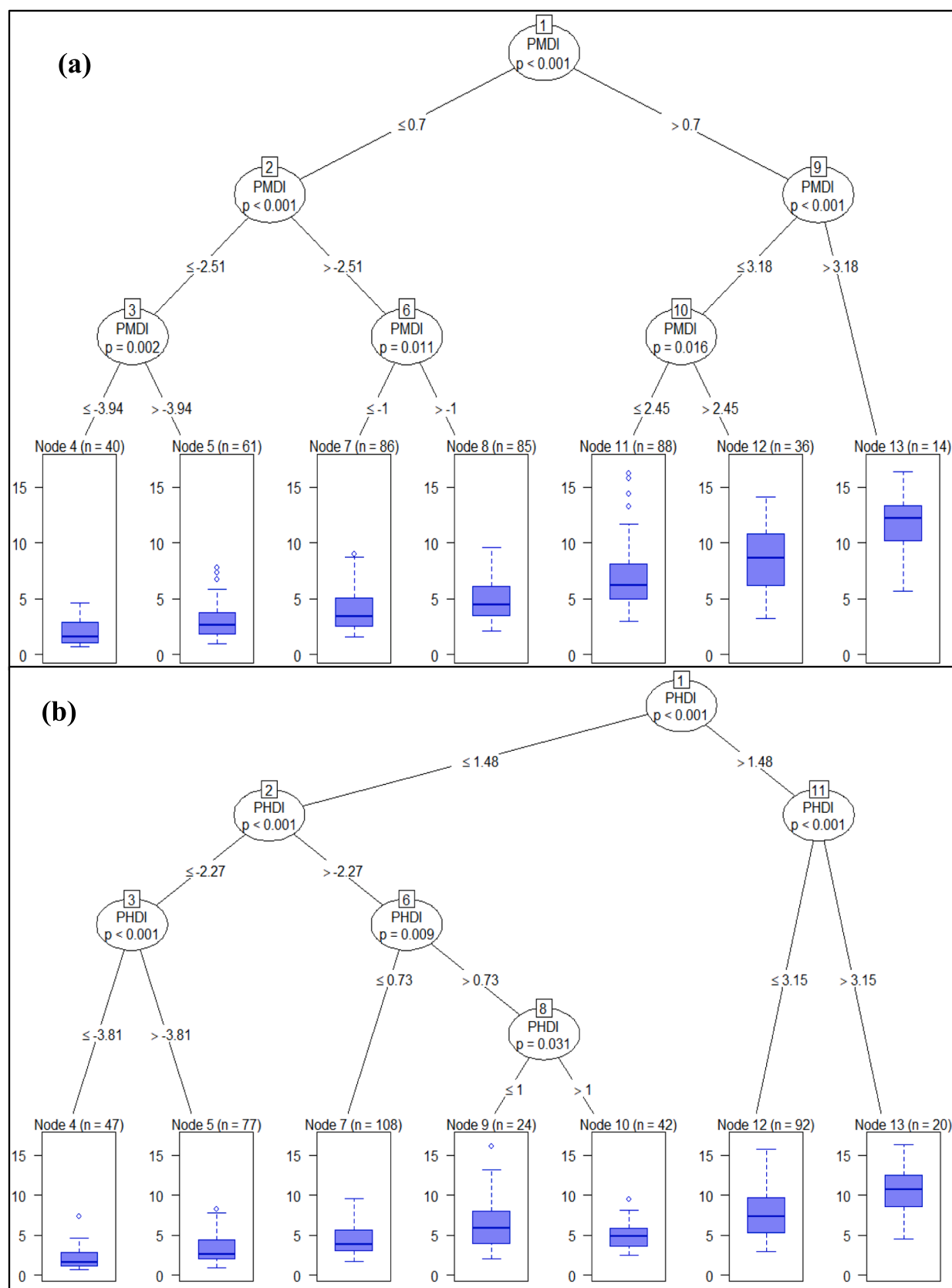


Fig. 3. The decision trees show the thresholds associated with drought indices (a) PMDI and (b) PHDI. The boxplot shows the streamflow (m^3/s) based on different drought thresholds. The threshold analysis is performed for the gage station located in Northwest Climate Division, which is less affected by human influence.

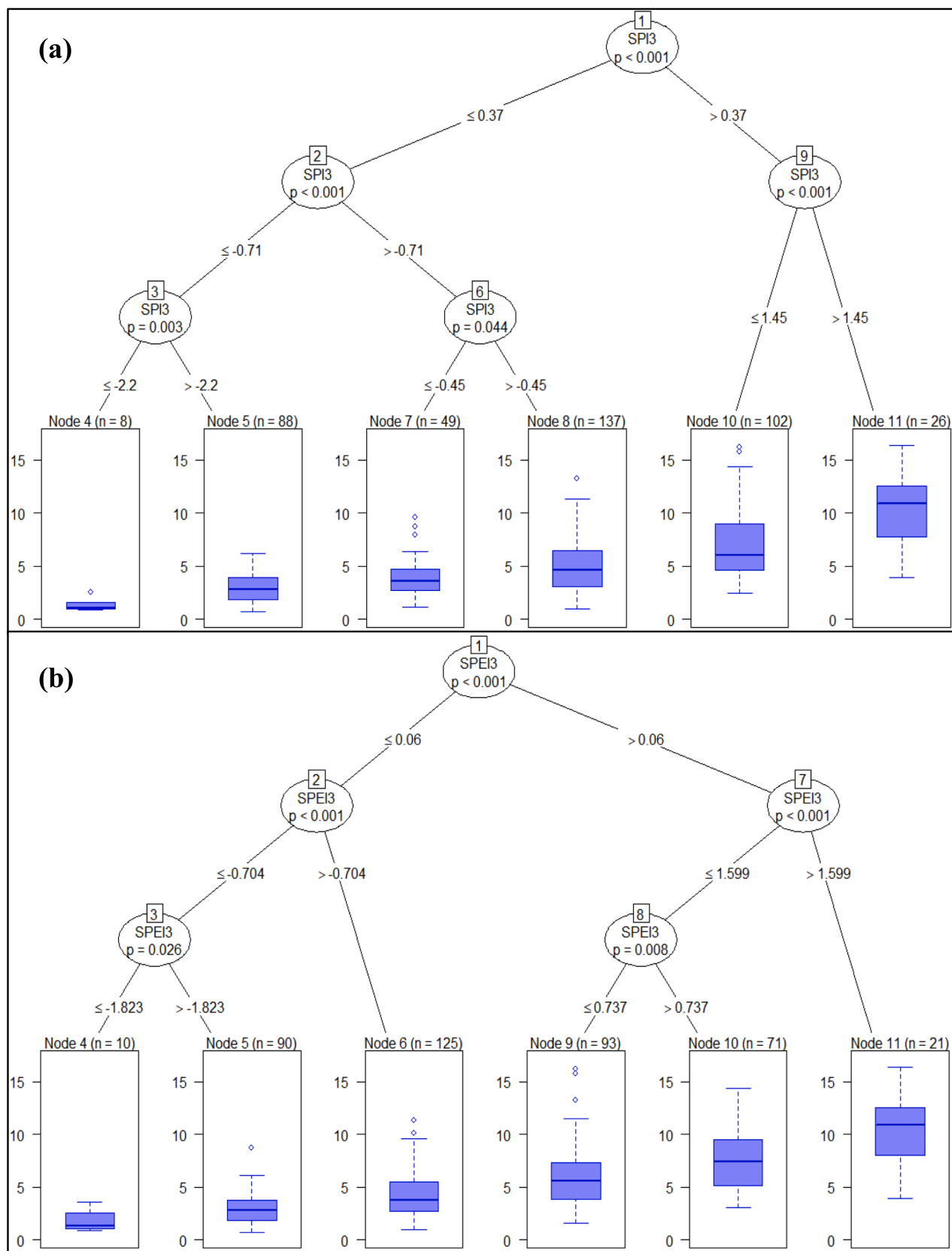


Fig. 4. The decision trees show the thresholds associated with two drought indices (a) SPI3 and (b) SPEI3.

4.3. Thresholds associated with drought indices for reservoir level

The three reservoirs considered in the analysis are Hartwell, Russel, and Thurmond, which are located in the Savannah River basin. The average storage level in the Hartwell, Russel, and Thurmond reservoirs is 200, 144.5, and 99 m, respectively. The Pearson correlation coefficient between the drought indices and monthly reservoir storage is provided in Fig. 5. Unlike streamflow, PHDI showed the highest correlation with reservoir storage level, where Hartwell showed a correlation of 0.7, and Russel and Thurmond showed a correlation of 0.59 and 0.61, respectively. In addition, the Palmer-based drought indices (except the ZNDX) showed a higher correlation with reservoir storage. In the case of PMDI, the correlation was 0.67, 0.54, and 0.57 for the Hartwell, Russel, and Thurmond reservoirs, respectively. In addition, it is interesting to observe that the long-term drought indices (i.e., SPI12 and SPEI12) have a strong relationship with the reservoir level. For instance, SPI12 showed a correlation of 0.68, 0.52, and 0.60 for the Hartwell, Russel, and Thurmond reservoirs, respectively. All the drought indices showed a higher correlation with the Hartwell reservoir storage. This may be because of the least influence of anthropogenic factors over the inflow to the reservoir. Russel reservoir storage, located downstream of the Savannah River Basin, showed less correlation with the drought indices.

We quantified the threshold associated with drought indices for reservoir storage using the CART analysis. Here, we selected Hartwell reservoir data to perform threshold analysis as the anthropogenic activities have the least influence on its drainage area. Using the CART approach, we generated separate decision trees for the drought indices that are highly correlated with reservoir storage levels, which include PHDI, PMDI, SPI12, and SPEI12. The model output of CART analysis using PHDI is schematically represented in Fig. 6a. Here the primary split of PHDI is with a threshold value of -0.89 . For example, if the PHDI is less than or equal to -0.89 , the tree's growth is towards the left, and the second and third split evolved at a threshold of -3.2 and -4.0 , respectively. The result shows that if the PHDI is less than -4.0 , the Hartwell reservoir level likely varies between 196 m and 199.5 m, with a median value of 198 m. Whereas, if the PHDI is higher than -4.0 and less than -3.2 , the reservoir level will vary from 197 m to 201 m. The final split of the decision tree is with a threshold of 1.76, where we observed the highest range of reservoir storage levels. The decision tree generated for PMDI (Fig. 6b) initiated with a threshold of -2.02

followed by -4.0 . For instance, if the PMDI is less than or equal to -4.0 , the Hartwell Reservoir level varies from 196 m to 199.8 m, with a median of 198 m. In addition, the total number of nodes in the case of the PHDI decision tree was 13, whereas PMDI resulted in 11 nodes in the output tree.

Similarly, the decision tree is generated for investigating the threshold of SPI12 and SPEI12 on the Hartwell reservoir level (Fig. 7a and 7b). The number of nodes in the SPI12 and SPEI12 decision trees was 11 and 9, respectively. In the case of SPI12, the first split of the decision tree was with a threshold of -0.36 . The output tree shows that if the SPI12 is less than or equal to -1.75 , the Hartwell reservoir level varies from 196 m to 198.5 m. The maximum threshold of SPI12 observed in the decision tree was 0.38, suggesting that if the SPI12 is higher than 0.38, the reservoir level varies from 199.5 to 202.5 m. In the case of SPEI12, the initial split threshold was 0.564, followed by -1.367 . For instance, SPEI12 is less than or equal to -1.367 , and the Hartwell Reservoir level varies from 195 m to 201.5 m. The maximum SPEI12 threshold in the analysis was 0.858.

4.4. Thresholds associated with drought indices for soil moisture

This section analyzed the relationship between the drought indices and observed soil moisture. The in-situ soil moisture for the Northeast and Southern Climate Divisions is obtained from the US CRN and SCAN databases. The Pearson correlation analysis is performed for the topsoil layer (5 cm) and the soil moisture below the root zone (100 cm). However, soil moisture below the root zone is weakly correlated with all the drought indices. Therefore, we focused on the topsoil moisture in further analysis.

Fig. 8 illustrates the correlation coefficient between the four soil moisture stations (i.e., SM1, SM2, SM3, and SM4) and drought indices. Here SM1, SM3, and SM4 are located in the Southern climate division of the state, and SM2 is located in the Northeast climate division. The correlation between the soil moisture and drought indices was relatively less than the streamflow and reservoir level analysis. However, we noticed a higher correlation for the soil moisture station SM1. For example, in the case of SM1, ZNDX and PMDI showed a value higher than 0.5. It was also observed that SM3 showed comparatively less correlation in the analysis. In the case of SPI and SPEI, short-duration drought showed a relatively better correlation with soil moisture. In

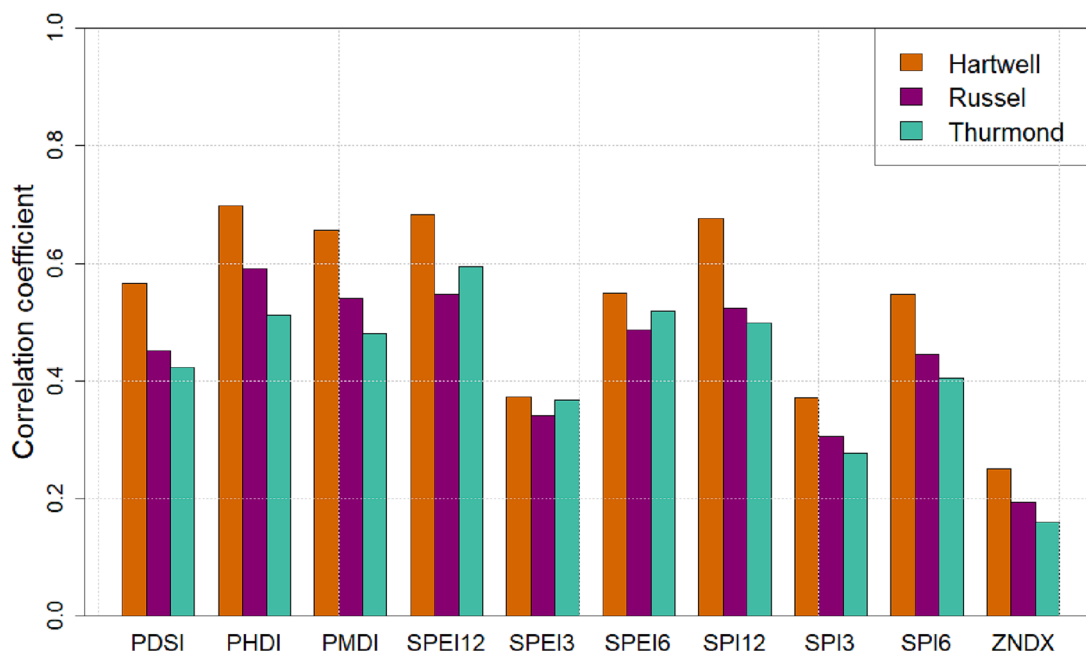


Fig. 5. Correlation between Reservoir level with drought indices. A high correlation is observed for PHDI, SPEI12, and SPI12.

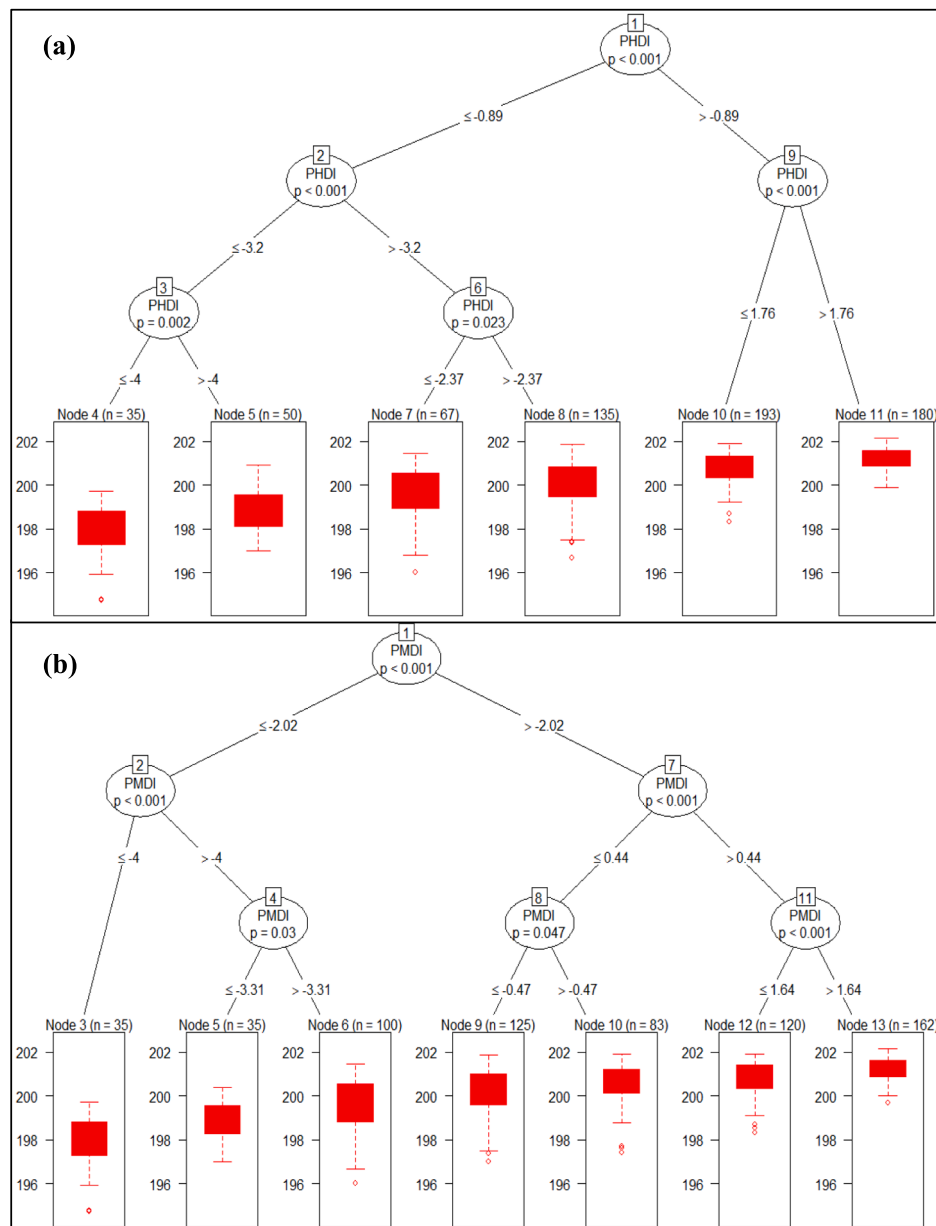


Fig. 6. The decision trees show the drought index threshold of (a) PHDI and (b) PMDI. The boxplot shows the Reservoir level (Unit: Meters) for Hartwell Reservoir based on different drought thresholds.

contrast, the long-duration drought showed less correlation with the in-situ data analyzed.

The highest correlation between drought indices and soil moisture is observed for the soil moisture station SM1, located in the Southern climate division. Therefore, we performed the drought threshold analysis using the decision tree between SM1 and ZNDX, PMDI, SPI3, and SPEI3, and the output of the tree is illustrated in Figs. 9 and 10. For example, in the case of ZNDX, the threshold was 0.49, indicating that if the ZNDX is less than or equal to 0.49, the topsoil moisture in the southern climate division varies from 2.5 to 20 percent, with a median value of 8 percent. Whereas, if the threshold is higher than 0.49, the topsoil moisture varies from 4.5 to 28 percent.

4.5. Thresholds associated with drought indices for crop yields

The crop yield of each climate division is represented based on the yearly corn yield data obtained from the USDA NASS. However, no clear records of corn crop management (e.g., growing season) are available

for South Carolina State's climate divisions. Therefore, we considered the growing season from 1st of March to 31st of October for all climate divisions, and the drought indices also averaged across the crop growing season. The boxplots based on the correlation coefficients between corn yield across the climate divisions and drought indices are shown in Fig. 11. Overall, the highest correlation (median) is observed for the ZNDX (0.28), followed by PDSI (0.27). In contrast, other Palmer-based indices showed relatively less correlation with the corn yield across the state. In the case of SPI and SPEI, the short-term drought showed a comparatively higher correlation value than SPI12 and SPEI12.

We quantified the threshold associated with drought indices for corn yield. The correlation between corn yield and drought indices, particularly ZNDX and PDSI, was relatively higher for the Northwest climate division. Therefore, the threshold analysis is performed for the Northwest climate division. For instance, the PDSI showed a correlation value of 0.52, and ZNDX showed a correlation value of 0.47 for the Northwest climate division. The decision tree analysis is performed for PDSI and ZNDX, and the critical threshold of drought indices and corresponding

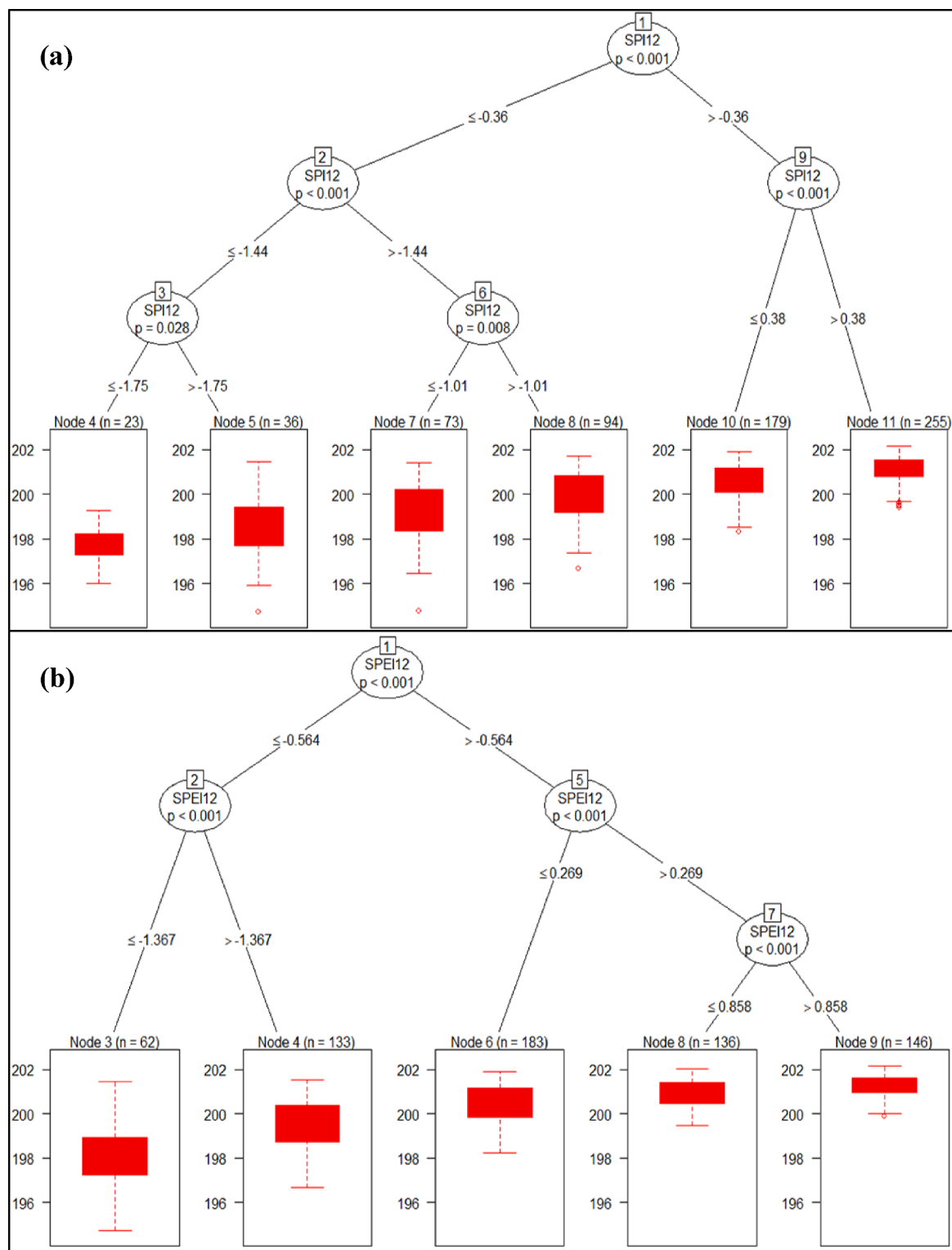


Fig. 7. The decision trees show the drought index threshold of (a) SPI12 and (b) SPEI12. The boxplot shows the Reservoir level (Unit: Meters) for Hartwell Reservoir based on different drought thresholds.

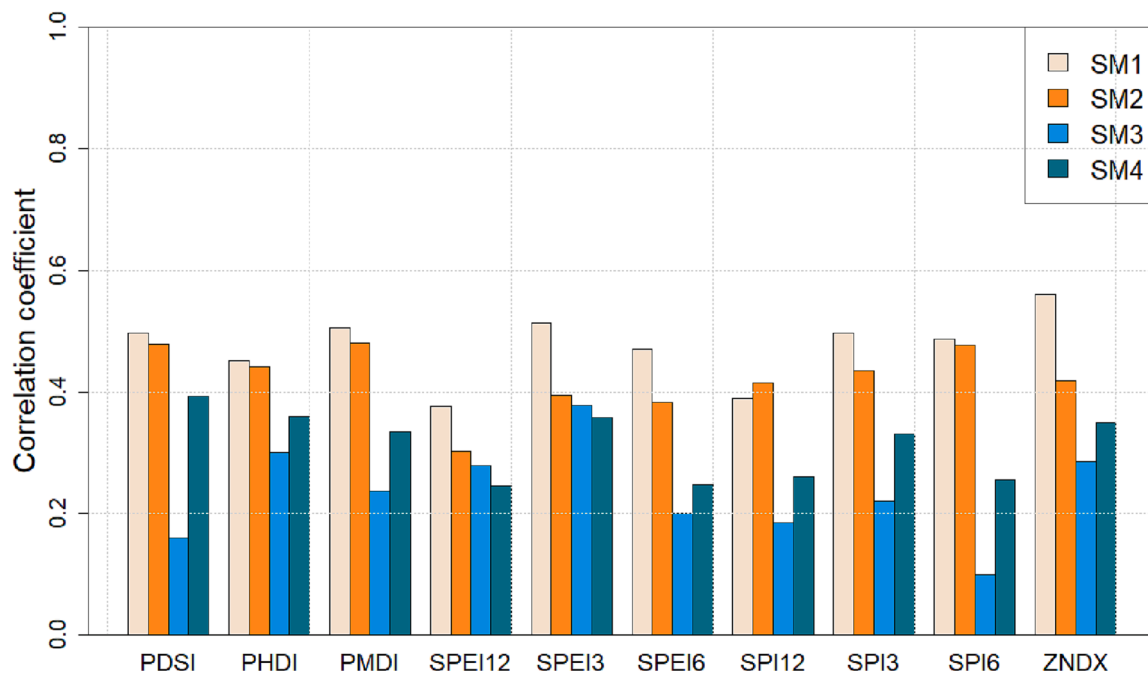


Fig. 8. Correlation between soil moisture with drought indices. The topsoil moisture for the initial 5 cm is analyzed, and a higher correlation is observed for ZNDX, followed by SPEI3, PMDI, and SPI3 for station SM1. The soil moisture gaging stations SM1, SM3, and SM4 are in the Southern climate division, and SM2 is in the Northeast climate division.

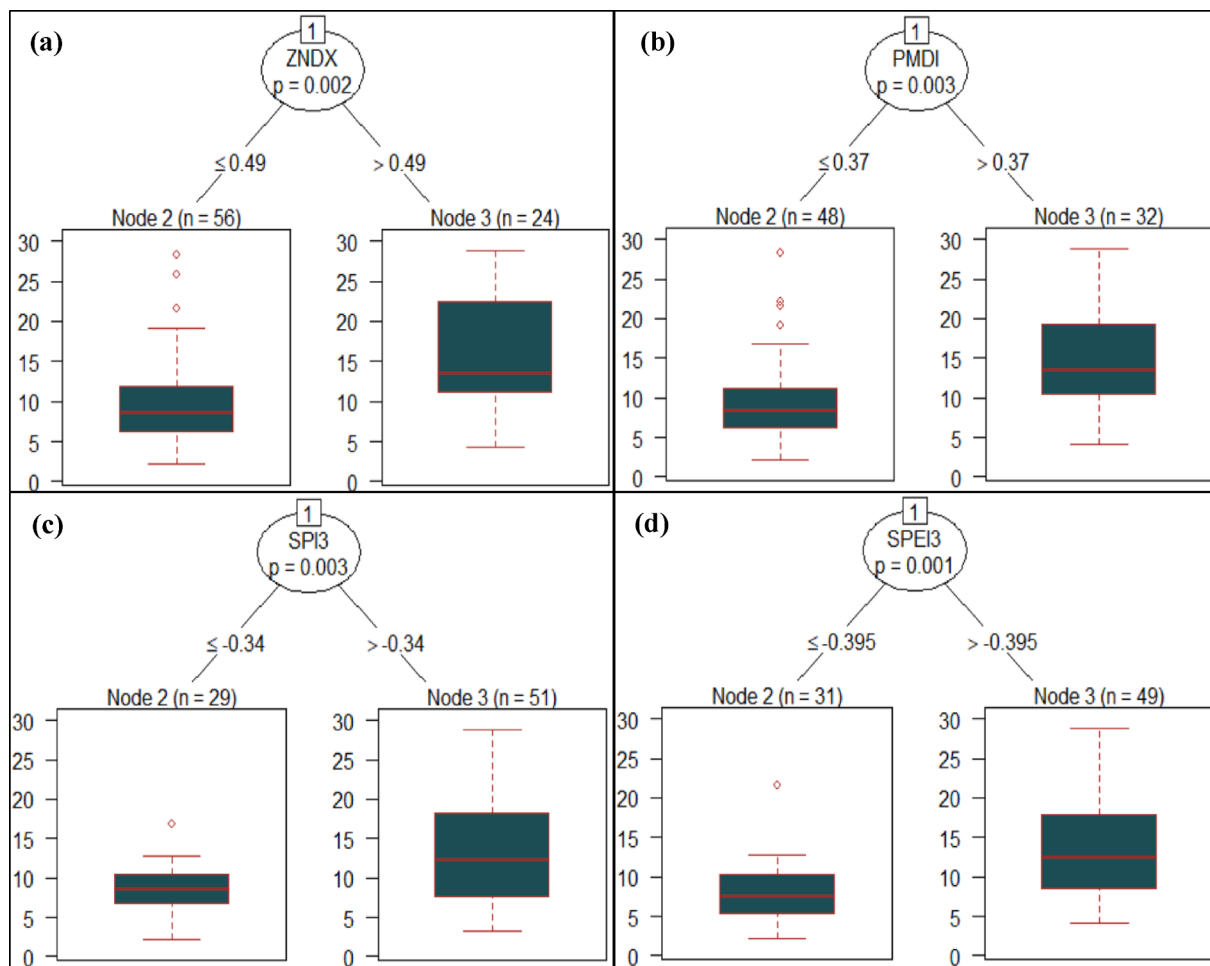


Fig. 9. The decision trees show the threshold of the drought index based on (a) ZNDX, (b) PMDI, (c) SPI3, and (d) SPEI3. The boxplot shows the soil moisture (Unit: Percentage) for soil moisture station SM1 based on different drought threshold.

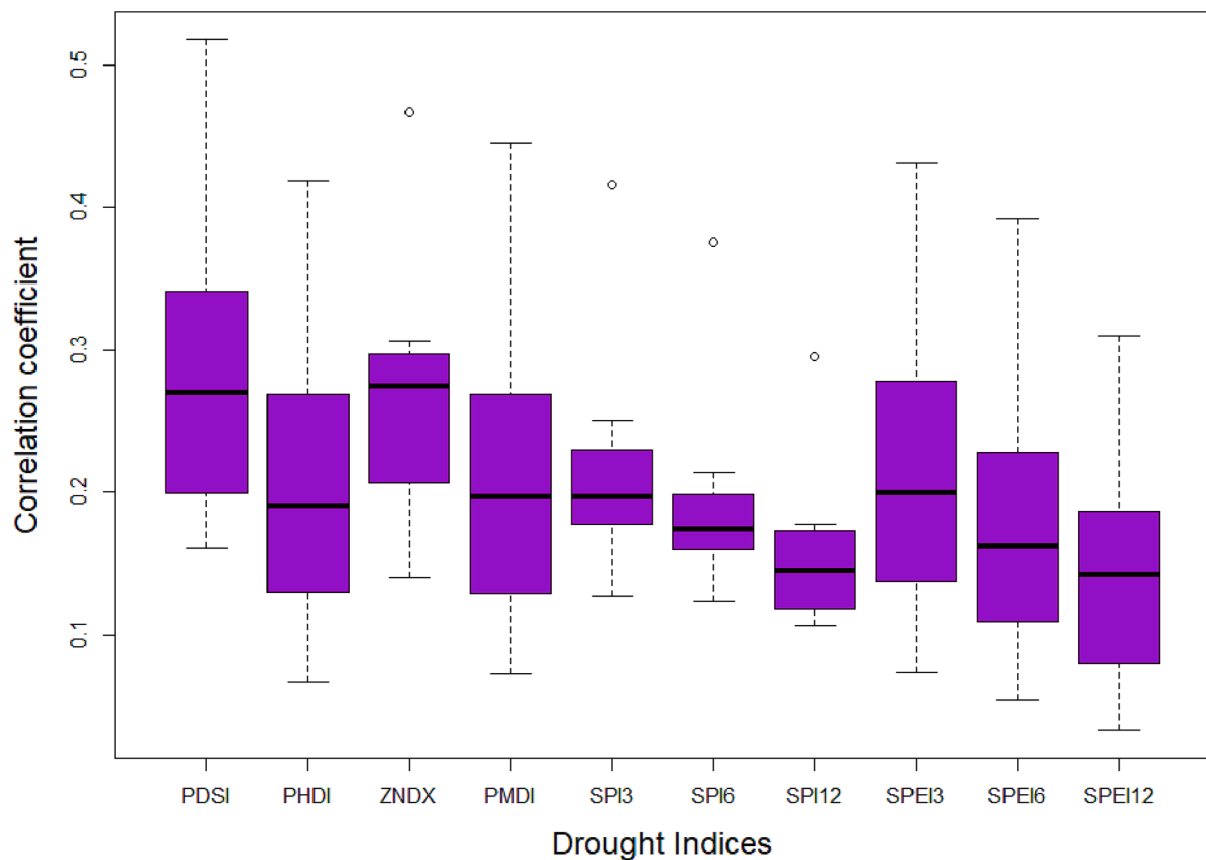


Fig. 10. Boxplot showing the correlation between corn yield across the climate divisions and drought indices. Higher values of correlation are observed for PDSI and ZNDX. We considered the growing season from March to October for all climate divisions, and the drought indices also averaged across the crop growing season.

yield is summarized in Fig. 11. It is observed that the PDSI split the corn yield with a threshold of 0.241. For example, when the PDSI of the climate division is less than or equal to 0.241, the tree's growth is towards the left and results in a median corn yield of 6 Tonnes per Acre. Whereas, if the PDSI is greater than 0.241, the corn yield in the climate division increases to 10 Tonnes per Acre. In the case of ZNDX, the threshold was 0.022. Similarly, the threshold analysis is performed for SPI3 and SPEI3, and the output tree is illustrated in Fig. 11c and 11d.

5. Discussion and concluding remarks

For operational drought management, it is essential to identify the region-specific drought indices and their relevant thresholds for the stakeholders. However, to date, little effort has been made to explore which index (or indices) best represents drought impacts of a specific region using threshold concepts. Also, imprecise definitions, slow onset, and multiple socio-ecological interactions make drought quantification challenging and complicate its impacts on different sectors. This leads to inconsistency and uncertainty in communicating drought indices for early preparedness with stakeholders. This study addresses the knowledge gap by identifying the appropriate thresholds associated with drought indices and impact-specific indicators. We compared the performance of drought indices, such as Palmer-based (e.g., PDSI, PMDI, and PHDI) and multiscale drought indices (e.g., SPI and SPEI) at 3, 6, and 12 months with the impact-specific indicators, such as streamflow, reservoir level, soil moisture, and crop yield. Palmer drought indices, which are widely used and implemented in the drought monitoring system (Zhou et al., 2022; Jacobi et al., 2013); SPI, recognized by the World Meteorological Organization as an effective drought monitoring tool for climate risk management (Hayes et al., 2012; Yaseen et al., 2021); SPEI, which considers the potential evapotranspiration on

drought severity quantifications (Vincente-Serrano et al., 2010; Liu et al., 2021).

The SPI and SPEI are analyzed for multiple time scales (i.e., three months, six months, and 12 months). Here, the short-term multiscale indices (i.e., SPI3 and SPEI3) showed a relatively higher correlation with the streamflow than medium and long-term multiscale indices. This can be interpreted as the short-term fluctuations in the precipitation pattern likely to control streamflow variations for the selected watersheds. Whereas, long-term drought indicators (SPI 12 and SPEI 12) have a higher correlation with reservoir level, which suggests that the contribution of surface and baseflows to the reservoir storage are better represented by these two drought indicators. The baseflow (delayed flow) may not be captured by the short-term drought indicators. In the case of soil moisture and crop yield, short-term drought indicators showed a relatively higher correlation.

PMDI showed the highest correlation with streamflow, as they are derived based on the sum of the wet and dry terms weighted by probability values. In the case of reservoir level, PHDI showed a relatively higher correlation. PHDI is a modified version of PDSI, and it is designed to capture longer-term dryness that is likely to affect water storage, reservoirs, and groundwater. PHDI responds more slowly to regional weather changes. The advantage of this delayed response is that while the weather may return to normal, a deficiency in lake levels may still exist. Palmer Z index responds to short-term conditions enabling it to capture rapidly developing drought conditions, as a result, this indicator likely to have a relatively higher correlation in the case of surface soil moisture and crop yield.

The CART algorithm was used to identify the thresholds associated with drought indices and impact-specific indicators. The threshold refers to a splitting criterion used to divide the response variables (i.e., impact-specific indicators) based on the drought index at each internal node of

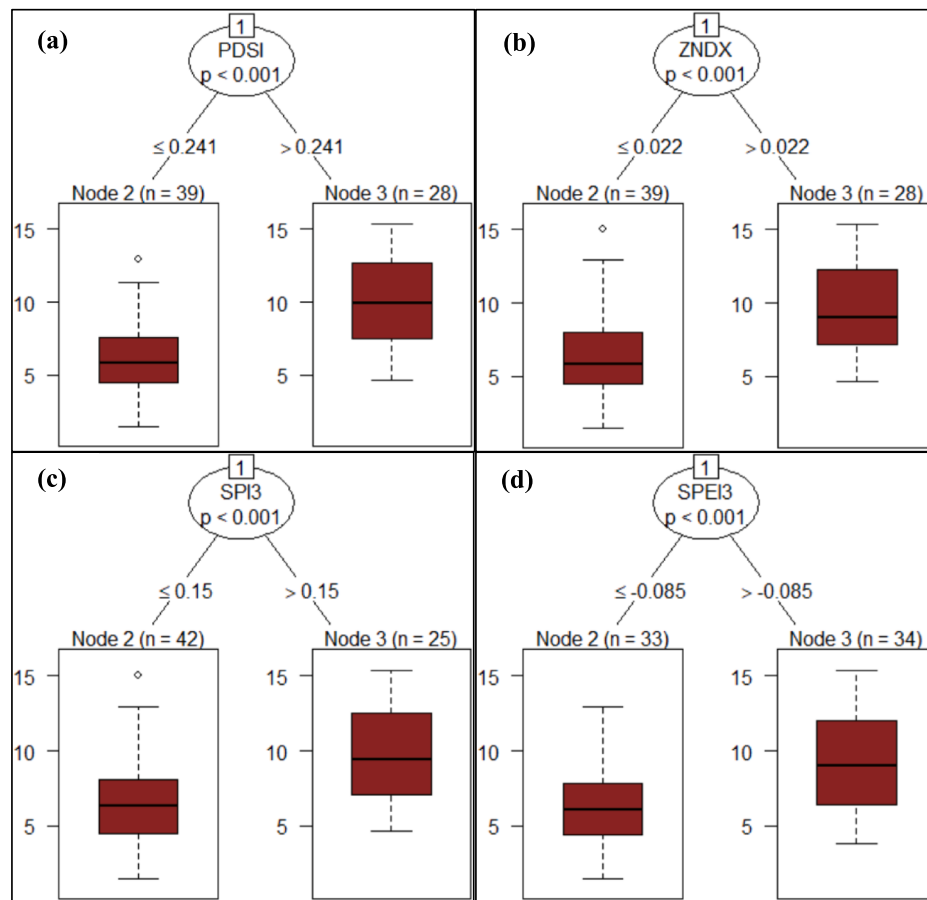


Fig. 11. The decision trees show the threshold of the average drought index based on (a) PDSI, (b) ZNDX, (c) SPI3, and (d) SPEI3. The boxplot shows the crop yield (Unit: Tonnes per Acre). A maximum correlation is observed between the crop yield from Northwest Climate Division and corn yield. Therefore, the decision tree analysis is performed for Northwest Climate Division.

the tree. The drought thresholds create binary splits and create a distribution of impact-specific indicators using a boxplot representation. This threshold can properly address the numerical value associated with each drought index and its influence on impact-specific indicators. The decision trees provide the mean of the responses as an output of the tree to generate the boxplots of impact-specific indicators associated with the drought indices threshold. The boxplots show an overlap in the resulting distribution (range) of impact-specific indicators based on the model performance analysis (Appendix B). For example, the trees developed for streamflow and reservoir level showed comparatively less overlap, resulting in more branches due to model accuracy. However, the median values showed explicit differences in each boxplot, resulting in all the impact-specific indicators.

Many factors, such as supplementary irrigation, influence corn yield. Our study area is located in South Carolina, where rainfed agriculture is predominant. Also, rainfed crops are generally more sensitive to drought, and performing separate analyses for irrigated crops can separate the influence of drought on major crop yields (Lu et al., 2020). The following conclusions are drawn from this study,

- (a) The characteristics of drought vary substantially across different climate divisions. For example, in climate division 2, the average drought duration (DD) from 1950 to 2017 was 9.59 for the PDSI, while it was 1.84 for ZNDX and 6.38 for PMDI. The average drought severity (DS) also demonstrated wide variations within the climate divisions. For instance, in climate division 5, the PDSI showed a DS of 10.74, while it was 1.73 for ZNDX. The DS for SPI3 and SPEI3 was 1.0 and 0.9, respectively.

- (b) Drought indicators behave significantly differently during extreme drought conditions. The calculation of the maximum drought duration using different drought indices during the 2011 drought showed significant differences within the climate divisions. For example, In climate division 2, the maximum drought duration for PDSI and PMDI was 25 months during the 2011 drought, while it was 34 months for PHDI and 8 months for ZNDX. Additionally, SPI3 and SPEI3 had the shortest drought duration during the 2011 drought in climate division 2.
- (c) The correlation between drought indices and impact-specific indicators also demonstrated significant variations in the analysis. The PMDI had the strongest correlation with streamflow, with a median correlation coefficient of 0.62, followed by PHDI, with a coefficient of 0.58. The correlation between SPI and SPEI and streamflow was relatively low. For example, SPI3 had a median correlation of 0.57, while SPI6 and SPI12 had a median value lower than that of SPI3. In the case of SPEI, SPEI3 showed the highest correlation with streamflow, while SPEI12 had the weakest correlation in the analysis. Unlike streamflow, PHDI had the strongest correlation with the reservoir storage level, with Hartwell showing a coefficient of 0.7, and Russell and Thurmond having a coefficient of 0.59 and 0.61, respectively. In terms of soil moisture and crop yield, ZNDX had the highest correlation.
- (d) Palmer indices generally showed improved performance compared to multiscalar index in the selected study area; however, the indicator's performances will vary across reason and corresponding impact-specific sectors. The multiscale indices strongly correlate with various system responses in different studies (Vicente-Serrano et al., 2012). Additionally, Haslinger

et al. (2014) found similar correlations between the Palmer drought index and SPEI and streamflow data in their study in Austria. The relationship between drought indices and system responses may vary depending on the time scale and region's climate-catchment interaction (Veetil and Mishra, 2020). These findings highlight the need to test and compare the regional performance of different drought indices to relevant impact-specific indices.

- (e) Drought shocks can have significant socio-economic impacts. Therefore, understanding the influence of drought severity levels on impact-specific indicators is crucial for regional economic development. The decision tree approach quantifies critical thresholds of drought indices that trigger a range of hydrological and agricultural indicators (system response levels). The results indicate that the threshold associated with each drought index and the impact-specific indicator varies drastically within the climate division. For instance, the PMDI threshold for streamflow was roughly half the PHDI threshold for the same streamflow measurement.

CRedit authorship contribution statement

Anoop Valiya Veetil: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Ashok K. Mishra:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This study was supported by the National Science Foundation (NSF) award # 1653841.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhydrol.2023.129966>.

References

- Angelidis, P., Maris, F., Kotsovinos, N., Hrisanthou, V., 2012. Computation of drought index SPI with alternative distribution functions. *Water Resour. Manag.* 26 (9), 2453–2473.
- Bachmair, S., Svensson, C., Prosdociimi, I., Hannaford, J., Stahl, K., 2017. Developing drought impact functions for drought risk management. *Nat. Hazards Earth Syst. Sci.* 17, 1947–1960.
- Badr, A. W., Wachob, A., & Gellici, J. A. (2004). South Carolina Water Plan, South Carolina Department of Natural Resources. Land, Water and Conservation Division, Columbia, SC.
- Bhalme, H.N., Mooley, D.A., 1979. On the performance of modified Palmer index. *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B* 27 (4), 281–295.
- Billah, M.M., Goodall, J.L., 2011. Annual and interannual variations in terrestrial water storage during and following a period of drought in South Carolina, USA. *J. Hydrol.* 409 (1–2), 472–482.
- Blauhut, V., Gudmundsson, L., Stahl, K., 2015. Towards pan-European drought risk maps: quantifying the link between drought indices and reported drought impacts. *Environ. Res. Lett.* 10.
- Cancelliere, A., Mauro, G.D., Bonaccorso, B., Rossi, G., 2007. Drought forecasting using the standardized precipitation index. *Water Resour. Manag.* 21 (5), 801–819.
- Chawla, I., Karthikeyan, L., Mishra, A.K., 2020. A review of remote sensing applications for water security: Quantity, quality, and extremes. *J. Hydrol.* 585.
- Christian, J.I., Basara, J.B., Hunt, E.D., Otkin, J.A., Furtado, J.C., Mishra, V., Xiao, X., Randall, R.M., 2021. Global distribution, trends, and drivers of flash drought occurrence. *Nat. Commun.* 12 (1), 1–11.
- Das, J., Jha, S., Goyal, M.K., 2020. Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India. *J. Hydrol.* 580.
- Drakopoulos, S.A., 1994. Hierarchical choice in economics. *J. Econ. Surv.* 8 (2), 133–153.
- García-Herrera, R., Díaz, J., Trigo, R.M., Luterbacher, J., Fischer, E.M., 2010. A review of the European summer heat wave of 2003. *Crit. Rev. Environ. Sci. Technol.* 40 (4), 267–306.
- Gavahi, K., Abbaszadeh, P., Moradkhani, H., Zhan, X., Hain, C., 2020. Multivariate assimilation of remotely sensed soil moisture and evapotranspiration for drought monitoring. *J. Hydrometeorol.* 21 (10), 2293–2308.
- Guttman, N.B., 1991. A sensitivity analysis of the Palmer hydrologic drought index 1. *JAWRA J. Am. Water Resour. Assoc.* 27 (5), 797–807.
- Guttman, N.B., 1999. Accepting the standardized precipitation index: a calculation algorithm 1. *JAWRA J. Am. Water Resour. Assoc.* 35 (2), 311–322.
- Hao, Z., AghaKouchak, A., 2013. Multivariate standardized drought index: a parametric multi-index model. *Adv. Water Resour.* 57, 12–18.
- Haslinger, K., Koffler, D., Schöner, W., Laaha, G., 2014. Exploring the link between meteorological drought and streamflow: Effects of climate-catchment interaction. *Water Resour. Res.* 50 (3), 2468–2487.
- Hayes, M. J., Svoboda, M. D., Wardlaw, B. D., Anderson, M. C., & Kogan, F. (2012). Drought monitoring: Historical and current perspectives.
- Hsiang, S.M., Burke, M., Miguel, E., 2013. Quantifying the influence of climate on human conflict. *Science* 341 (6151), 1235367.
- IPCC, C. W. T. (2021). Allan, R. P., Hawkins, E., Bellouin, N., & Collins, B. IPCC, 2021: summary for Policymakers.
- Jacobi, J., Perrone, D., Duncan, L.L., Hornberger, G., 2013. A tool for calculating the Palmer drought indices. *Water Resour. Res.* 49 (9), 6086–6089.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: Springer.
- Karl, T.R., 1986. The sensitivity of the Palmer Drought Severity Index and Palmer's Z-index to their calibration coefficients including potential evapotranspiration. *J. Clim. Appl. Meteorol.* 77–86.
- Konapala, G., Mishra, A., 2020. Quantifying climate and catchment control on hydrological drought in the continental United States. *Water Resour. Res.* 56 (1), e2018WR024620.
- Konapala, G., Mishra, A.K., Wada, Y., Mann, M.E., 2020. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nat. Commun.* 11 (1), 3044.
- Kreibich, H., Van Loon, A.F., Schröter, K., Ward, P.J., Mazzoleni, M., Sairam, N., Abeshu, G.W., Agafonova, S., AghaKouchak, A., Aksoy, H. and Alvarez-Garreton, C., 2022. The challenge of unprecedented floods and droughts in risk management. *Nature*, 608(7921), pp.80–86.
- Krzywinski, M., Altman, N., 2017. Classification and regression trees. *Nat. Methods* 14, 757–758. <https://doi.org/10.1038/nmeth.4370>.
- Łabędzki, L., & Bąk, B. (2014). Meteorological and agricultural drought indices used in drought monitoring in Poland: a review. *Meteorology Hydrology and Water Management. Research and Operational Applications*, 2.
- Liu, C., Yang, C., Yang, Q., Wang, J., 2021. Spatiotemporal drought analysis by the standardized precipitation index (SPI) and standardized precipitation evapotranspiration index (SPEI) in Sichuan Province, China. *Sci. Rep.* 11 (1), 1–14.
- Long, J.W., Lake, F.K., Goode, R.W., 2021. The importance of Indigenous cultural burning in forested regions of the Pacific West, USA. *For. Ecol. Manage.* 500.
- Loucks, D.P., Van Beek, E., 2017. Water resource systems planning and management: An introduction to methods, models, and applications. Springer.
- Lu, J., Carbone, G.J., Huang, X., Lackstrom, K., Gao, P., 2020. Mapping the sensitivity of agriculture to drought and estimating the effect of irrigation in the United States, 1950–2016. *Agric. For. Meteorol.* 292.
- McKee, T. B., Doesken, N. J., & Kleist, J. (1993, January). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (Vol. 17, No. 22, pp. 179–183).
- Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. *J. Hydrol.* 391 (1–2), 202–216.
- Mishra, A.K., Singh, V.P., 2011. Drought modeling—A review. *J. Hydrol.* 403 (1–2), 157–175.
- Mishra, V., Thirumalai, K., Jain, S., Aadhar, S., 2021. Unprecedented drought in South India and recent water scarcity. *Environ. Res. Lett.* 16 (5).
- Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.* 25 (6), 693–712.

- Mukherjee, S., Mishra, A., Trenberth, K.E., 2018. Climate change and drought: a perspective on drought indices. *Curr. Clim. Change Rep.* 4 (2), 145–163.
- Mukherjee, S., Mishra, A.K., Mann, M.E., Raymond, C., 2021. Anthropogenic warming and population growth may double US heat stress by the late 21st century. *Earth's Future* 9 (5) e2020EF001886.
- Palmer, W. C. (1965). Meteorological droughts. US Department of Commerce Weather Bureau Research Paper 45, 58 pp.
- Palmer, W. C. (1968). Keeping track of crop moisture conditions, nationwide: the new crop moisture index.
- Pedro-Monzónis, M., Solera, A., Ferrer, J., Estrela, T., Paredes-Arquiola, J., 2015. A review of water scarcity and drought indexes in water resources planning and management. *J. Hydrol.* 527, 482–493.
- Rajsekhar, D., Singh, V.P., Mishra, A.K., 2015. Multivariate drought index: An information theory based approach for integrated drought assessment. *J. Hydrol.* 526, 164–182.
- SCDHEC, 2010. South Carolina Department of Health and Environmental Control, Watershed Water Quality Assessment, <www.scdhec.gov/water>.
- Sheffield, J., Wood, E.F., Roderick, M.L., 2012. Little change in global drought over the past 60 years. *Nature* 491, 435–438.
- Solomatine, D.P., 2002. July). Data-driven modelling: paradigm, methods, experiences. In: *In Proc. 5th international conference on hydroinformatics*, pp. 1–5.
- Stagge, J.H., Kohn, L., Tallaksen, L.M., Stahl, K., 2015. Modeling drought impact occurrence based on meteorological drought indices in europe. *J. Hydrol.* 530, 37–50.
- Sutanto, S.J., van der Weert, M., Wanders, N., et al., 2019. Moving from drought hazard to impact forecasts. *Nat. Commun.* 10, 4945. <https://doi.org/10.1038/s41467-019-12840-z>.
- Van Loon, A.F., Laaha, G., 2015. Hydrological drought severity explained by climate and catchment characteristics. *J. Hydrol.* 526, 3–14.
- Varian, H.R., 2014. Big data: new tricks for econometrics. *J. Econ. Perspect.* 28 (2), 3–27.
- Veettil, A.V., Mishra, A.K., 2016. Water security assessment using blue and green water footprint concepts. *J. Hydrol.* 542, 589–602.
- Veettil, A.V., Mishra, A.K., 2020. Multiscale hydrological drought analysis: Role of climate, catchment and morphological variables and associated thresholds. *J. Hydrol.* 582.
- Veettil, A.V., Fares, A., Awal, R., 2022a. Winter storm Uri and temporary drought relief in the western climate divisions of Texas. *Sci. Total Environ.* 835 <https://doi.org/10.1016/j.scitotenv.2022.155336>.
- Veettil, A.V., Mishra, A.K., Green, T.R., 2022b. Explaining water security indicators using hydrologic and agricultural systems models. *J. Hydrol.* 607.
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscale drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J. Clim.* 23 (7), 1696–1718.
- Vicente-Serrano, S. M., Beguería, S., Lorenzo-Lacruz, J., Camarero, J. J., López-Moreno, J. I., Azorin-Molina, C., ... & Sanchez-Lorenzo, A. (2012). Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.* 16 (10), 1–27.
- Wilhite, D.A., Svoboda, M.D., Hayes, M.J., 2007. Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. *Water Resour. Manag.* 21 (5), 763–774.
- Wilhite, D. A. (2000). drought as a natural hazard: concepts and definitions.
- Wu, J., Chen, X., Yao, H., Gao, L., Chen, Y., Liu, M., 2017. Non-linear relationship of hydrological drought responding to meteorological drought and impact of a large reservoir. *J. Hydrol.* 551, 495–507.
- Yaseen, Z.M., Ali, M., Sharafati, A., Al-Ansari, N., Shahid, S., 2021. Forecasting standardized precipitation index using data intelligence models: Regional investigation of Bangladesh. *Sci. Rep.* 11 (1), 1–25.
- Zhou, Y., Zhou, P., Jin, J., Wu, C., Cui, Y., Zhang, Y., Tong, F., 2022. Drought identification based on Palmer drought severity index and return period analysis of drought characteristics in Huaihei Plain China. *Environ. Res.* 212 <https://doi.org/10.1016/j.envres.2022.113163>.