

An Integrated Composite Drought Index for Spatial Drought Vulnerability Assessment in the Mahanadi River Basin, India

Lolali Nanda

University of Petroleum and Energy Studies

Pankaj Kumar Srivastava

pankajkgp@gmail.com

University of Petroleum and Energy Studies

Durga Madhab Mahapatra

University of Petroleum and Energy Studies

Research Article

Keywords: Groundwater Drought, Composite Drought Index, Multivariate Drought Vulnerability, Machine Learning, Bagging

Posted Date: October 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11204977/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

An Integrated Composite Drought Index for Spatial Drought Vulnerability Assessment in the Mahanadi River Basin, India

Lolali Nanda^a, Pankaj Kumar Srivastava^{a*}, Durga Madhab Mahapatra^{b*}

^aDepartment of Petroleum Engineering and Earth Sciences, Energy Cluster, School of Advanced Engineering (SoAE), UPES, Bidholi Campus, Dehradun – 248007, Uttarakhand, India

^bDepartment of Chemical Engineering, Energy Cluster, School of Advanced Engineering (SoAE), UPES, Bidholi Campus, Dehradun – 248007, Uttarakhand, India

Lolali Nanda: LOLALI.113618@stu.upes.ac.in

*Pankaj Kumar Srivastava: pankajkgp@gmail.com

*Durga Madhab Mahapatra: dmahapatra@ddn.upes.ac.in

Abstract

Climate variability and increasing water demand are intensifying drought vulnerability in the monsoon-dependent Mahanadi River Basin (MRB), India. Conventional drought assessments often examine individual drought types separately and may overlook the role of groundwater storage in shaping overall drought vulnerability. This study develops an integrated multivariate framework for assessing drought vulnerability by combining meteorological, agricultural, hydrological, and groundwater drought characteristics across the basin. A Composite Drought Index (CDI) was developed using Principal Component Analysis, Copula functions, Random Forest, and a Spatial Neighborhood Aggregation Encoder. Among these approaches, the Copula-based CDI demonstrated the strongest agreement with independent atmospheric moisture stress, with correlations of -0.86, -0.89, -0.92, and -0.81 with Vapor Pressure Deficit at 1-, 3-, 6-, and 12-month timescales, respectively. The Copula-based CDI was subsequently integrated as the exposure component with 17 sensitivity and 11 adaptive-capacity indicators to develop a spatial drought vulnerability framework. A

Bagging ensemble classifier was trained using historical district-level drought observations and evaluated using an independent testing dataset. The model achieved 89.6% accuracy, with precision, recall, and F1-score of 87.5%, 98.1%, and 92.5%, respectively, and an AUC-ROC of 0.945. The resulting vulnerability assessment indicates that approximately 25% of the basin falls within the very-high vulnerability class and 36% within the high vulnerability class, with prominent hotspots across the central, northern, southern, and southeastern sectors. By linking atmospheric and surface drought conditions with subsurface water-storage depletion and socio-environmental coping capacity, the proposed framework provides a comprehensive basis for identifying spatially differentiated drought vulnerability and supporting targeted drought management and water-resource planning in the MRB.

Keywords: Groundwater Drought, Composite Drought Index, Multivariate Drought Vulnerability, Machine Learning, Bagging

Highlight

- Development of composite Drought Index incorporating meteorological, agricultural, hydrological, and groundwater drought
- Bagging-based drought vulnerability assessment integrating CDI in Exposure index
- Western and northern parts of the MRB identified as highly vulnerable to drought

Abbreviations

Mahanadi River Basin: MRB
Bay of Bengal: BOB
Composite Drought Index: CDI
Principal Component Analysis: PCA
Random Forest: RF
Spatial Neighbourhood Aggregation Encoder: SNAE
Meteorological Drought: MD

59 Agricultural Drought: AD
60 Hydrological Drought: HD
61 Groundwater Drought: GD
62 Groundwater: GW
63 Groundwater Storage: GWS
64 Groundwater Level: GWL
65 Standardised Precipitation-Evapotranspiration Index: SPEI
66 Equivalent Standardized Soil Moisture Index: ESSMI
67 Standardized Streamflow Index: SSI
68 Groundwater Drought Index: GDI
69 Drought Vulnerability: DV
70 DV Index: DVI
71 Root Zone Soil Moisture: RZSM
72 Digital Elevation Model: DEM
73 Topography Wetness Index: TWI
74 Normalized Difference Vegetation Index: NDVI
75 Normalized Difference Water Index: NDWI
76 Land surface temperature: LST
77 Exposure Index: EI
78 Sensitivity Index: SI
79 Adaptive Capacity Index: ACI
80 Vegetation Health Index: VHI
81 Vapour Pressure Deficit: VPD
82 Topography Wetness Index: TWI
83 Crop intensity: CI
84 Net Sown Area: NSA
85

86

87

88

89

90

91

92

93

94

1 Introduction

Owing to anthropogenic climate change, the recurrence and intensity of hydroclimatic extremes such as droughts, floods, landslides, and heatwaves have increased worldwide, threatening human beings and natural systems [1]. Among these extremes, drought is the most complex and economically damaging natural hazard due to its higher frequency, longer duration, and severe economic consequences [2], [3]. Recent global assessments indicate that global mean surface temperatures have reached 1.75 °C above the preindustrial climate [4], causing more severe and widespread drought conditions. Consequently, nearly 30% of the global land surface faced drought in 2025 [5]. India is among the countries most vulnerable to drought, with nearly two-thirds of its geographical area susceptible to recurrent drought events [6], [7]. Its hydroclimate is governed by the Indian Summer Monsoon (ISM) [8], [9], which is highly sensitive to land-ocean thermal contrasts, sea surface temperature (SST) patterns, and large-scale ocean-atmosphere interactions [10], [11], [12]. However, anthropogenic warming has raised surface temperatures and altered the SST pattern, modifying atmospheric circulation, contributing to greater variability in the temporal and spatial distribution of monsoon rainfall. Consequently, the frequency and severity of droughts have increased across the country. As an agriculture-dominant country and home to 1.4 billion people, India is highly dependent on ISM, and droughts have profoundly impacted agricultural productivity, water resource management, and socio-economic well-being [13].

Traditionally, drought occurred when there is a precipitation deficit. However, contemporary drought science recognizes drought as a complex and interconnected process that propagates through different components of the hydrological cycle [14]. Initially, drought started as meteorological drought (MD), reducing soil moisture and

leading to agricultural drought (AD), followed by lower levels of streamflow and reservoir levels, which result in hydrological drought (HD) and eventually lead to groundwater storage (GWS) depletion, causing Groundwater Drought (GD) [15], [16], [17], [18]. Several Drought Indices were developed for understanding different types of droughts. For example, the Standardised Precipitation Index (SPI) [19] or the Standardised Precipitation-Evapotranspiration Index (SPEI) [20] incorporated only precipitation or both precipitation and evapotranspiration for MD monitoring. Soil moisture-based Indicators, such as Soil Moisture Deficit Index (SMDI) [21], Soil Moisture Index (SMI) [22], and vegetation-based indicators, including Vegetation Health Index (VHI) [23], are used for AD analysis. Similarly, the Standardized Streamflow Index (SSFI) is used for the HD analysis [24]. In recent decades, GD has become a centre of research as it serves as a hidden water storage during extreme dry periods. For GD studies, the Groundwater Drought Index (GDI) was developed from GRACE GWS anomaly data (GWSA) [25]. However, a single drought index does not capture the full range of drought-related information [26]. Therefore, CDIs have emerged in recent years.

In India, many researchers have developed CDI using multiple drought components for drought assessment. Kanthavel et al. [27] proposed a D-vine copula-based CDI using many hydrometeorological variables to address the nonlinear relationship among MD, AD, and HD in the Upper Tapti River Basin. Prajapati et al. [28] also used PCA-based CDI to monitor MD, AD, and HD in the Marathwada region of Maharashtra and highlighted the advantages of CDI in capturing drought severity and magnitude over individual indicators. Again, the PCA-based CDI method was adopted by Saini et al. [29] integrating MD, AD, and HD in the Banas River basin of Rajasthan to improve drought characterization and reduce the uncertainties

associated with single drought indicators. Tarate et al. [30] utilized the Analytical Hierarchy Process (AHP) to establish a Multi-Criteria Decision-Making (MCDM) CDI using Vegetation Condition Index (VCI), Precipitation Condition Index (PCI), and Terrestrial Water Storage Anomaly Condition Index (TWSACI), and the CDI showed a significant correlation with crop yield data. A notable advancement in composite drought assessment was made by Shah & Mishra [31] who developed a CDI combining MD, AD, GD, and HD for the Sabarmati and Brahmani River Basin. This study highlighted that the integration of GW data significantly enhanced the depiction of drought evolution, especially in cases where GW is the main source of freshwater during prolonged dry events. Recent advances in numerical modelling and Machine Learning (ML) have also improved the complex interaction of the GW system with other components of the hydrological cycle [32].

Although these studies have provided significant value to integrated drought monitoring, the incorporation of GD in CDI has received less attention, and most existing CDIs primarily focus on drought characterization and hazard assessment. However, drought hazard alone does not determine the magnitude of drought impacts, as regions experiencing similar drought severity may exhibit markedly different levels of damage depending on their biophysical, socio-economic, and institutional characteristics [33,34]. Therefore, assessing drought vulnerability (DV) has become equally important for understanding the susceptibility of natural and human systems to drought and for developing effective adaptation and mitigation strategies. The Exposure-Sensitivity-Adaptive Capacity (ESA) approach, as suggested by the Intergovernmental Panel on Climate Change (IPCC), is commonly used to assess DV [35]. Within the ESA framework, the exposure component is commonly represented using drought indicators. However, previous studies have predominantly relied on

single drought indices [36], [37], [38], [39]. Although these indices effectively represent specific dimensions of drought, they do not fully capture its cascading evolution. Recently, many researchers have proposed an integrated DV assessment framework by incorporating a multivariate drought hazard within the ESA framework, demonstrating the advantages of considering multiple drought characteristics over conventional single-index approaches [40], [41], [42], [43]. Nevertheless, even in this integrated framework, GD was not explicitly incorporated into the DV component. To address these knowledge gaps, the present study develops an integrated CDI by combining MD, AD, HD, and GD indicators and subsequently incorporates it as the exposure component within the IPCC-based ESA framework to assess spatial DV across the MRB.

2 Study area

MRB, the 2nd largest basin in peninsular India, covers 4.3% of India's territory [44]. This basin is geographically bounded by 19°8' and 23°32' North latitudes and 80°28' and 86°43' East longitudes, covering an area of 141,589 sq. km (Fig. 1). The Mahanadi River originates in Chhattisgarh, flows through Maharashtra, Jharkhand, and Odisha, and discharges into the Bay of Bengal (BOB). The annual average discharge from the basin is 1800 m³/sec. During summer, it increases up to 6352 m³/sec, while in the dry season, the value reduces to 759 m³/sec [45]. The basin lies in a tropical environment, receiving an annual average rainfall of 1463 mm, with around 85% of the rain occurring during the southwest monsoon (June to September). The basin also experiences high relative humidity, high mean temperatures, and short, mild winters. This raises surface temperature by 1.1°C/century, potentially intensifying climatic extremes and drought phenomena [46]. The basin has experienced many climatic disasters, including floods, droughts, and cyclones, every year due to increased temperatures, urbanization, and

reduced water yield [47], [48]. Even though droughts are not new in this region, their frequency and intensity have increased since 1965. The drought that occurred from September 2008 to November 2010 lasted for 27 months, causing crop failures, affecting farmers, and creating socio-economic problems [49]. Therefore, a DV assessment of this region is essential for strategic planning and management.

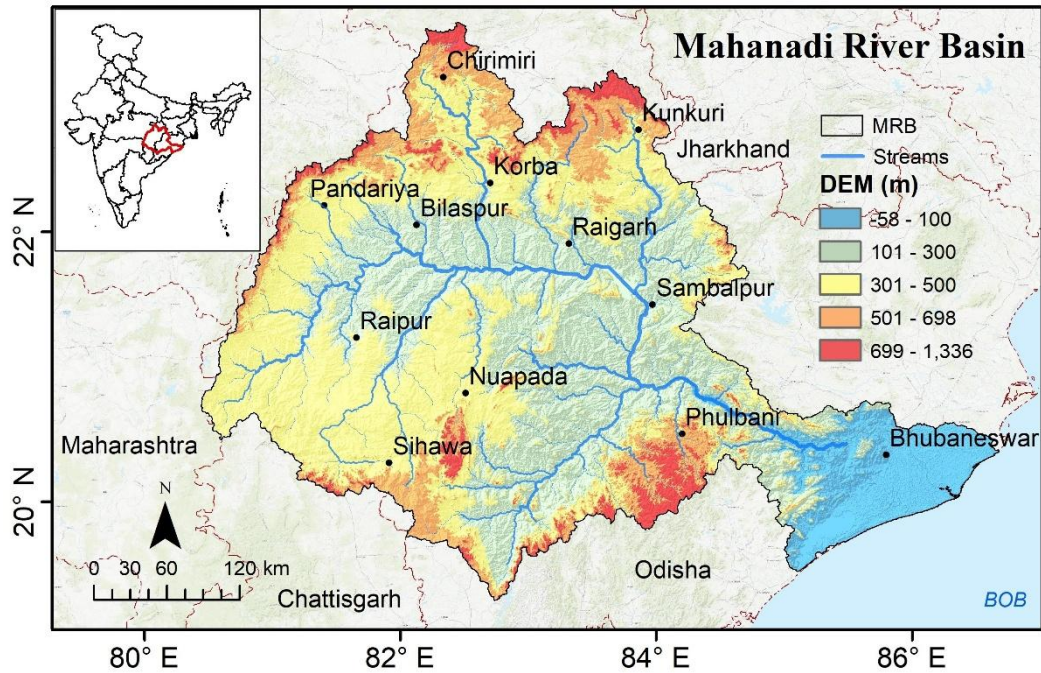


Fig. 1 Location map of the Mahanadi River Basin (MRB). Red dashed lines indicate the State boundaries surrounding the basin. The inset India map represents the geographical position of MRB within India

3 Datasets used

In this study, various hydro-meteorological, environmental, socio-economic, and topographical datasets were considered to develop the CDI and DV map (Table 1). The CDI integrated four key drought components (meteorological, agricultural, hydrological, and GW). For MD, SPEI was considered as it integrates both precipitation (P) and atmospheric evaporative (AET) demand. Thus, it captures climatic water deficits more comprehensively than precipitation alone [50]. AD was represented by Root Zone Soil Moisture (RZSM), as it directly reflects the water

available to vegetation and crops and provides a more direct measure of plant-available water stress [51]. HD was characterized by runoff, as reductions in runoff reflect the response of surface-water availability to sustained precipitation and soil-moisture deficits [52]. GD was represented using GWSA, derived from GRACE observations, because GWSA directly reflects temporal changes in subsurface water storage and enables the detection of GW depletion during prolonged dry conditions. The GWSA was derived from the Terrestrial Water Storage Anomaly (TWSA) from the CSR GRACE-FO RL06 Mascon product data between April 2002 and December 2023 at 0.25° spatial resolution [53]. SPEI was calculated from climatic water deficit series using P and PET. The P data were acquired from the ERA5 reanalysis dataset [54], while the PET, RZSM, air temperature, relative humidity (RH), and dew point temperature data were collected from the ERA5-Land Monthly Aggregated dataset from 2000 to 2023. Runoff data were also derived from the ERA5-Land reanalysis product [55]. All these datasets were collected on the Google Earth Engine (GEE) platform (<https://code.earthengine.google.com/>). Collectively, these indicators represent distinct but interconnected components of drought evolution, from atmospheric water deficit and root-zone moisture stress to surface-water and GWS deficits.

Apart from CDI, many environmental and climatic factors were considered for DV mapping (Table 2). Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) is considered in this study for computing slope, drainage density, distance to streams, and Topography Wetness Index (TWI) (<https://www.earthdata.nasa.gov/data/instruments/srtm>). Soil texture data were taken from the FAO Soil Portal [56] and the lithology data were obtained from the Geo Data India portal (<https://geodataindia.gov.in>). Vegetation and moisture indicators, including

Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI), were derived from the Dynamic World dataset available in GEE for 2023 [57]. The Dynamic World dataset provides near-real-time land cover information generated from Sentinel-2 imagery at 10m spatial resolution and has been widely applied in land surface and environmental studies. Information on groundwater level (GWL) was accessed from India-WRIS (indiaWRIS.gov.in), as a lowering of GWL indicates a rising threat of drought. All the socio-economic, demographic, health, and infrastructure-related indicators used for adaptive capacity assessment were obtained from the Census of India (<https://censusindia.gov.in/census.website/data/census-tables>) and the National Family Health Survey (NFHS-5) (<https://nfhs5.indiagis.org/nfhs5/>).

Table 1 Source and resolution of considered parameters

Data	Sources	Resolution	Time Period
Rainfall	ERA5 Reanalysis	0.25°	
Potential Evapotranspiration			
Dew point temperature	ERA5-Land Monthly Aggregates	0.1°	
Air temperature			
Root Zone Soil Moisture			2000-2023
Total Water Storage	GRACE RL06 Mascons	0.25°	
Land surface temperature	MOD11A1 v6.1	1 km	
Wind Speed	ERA5-Land Monthly Aggregates	0.1°	
NDVI			
NDWI	Sentinel-2 / Dynamic World	10 m	2023
Altitude	SRTM DEM / USGS EarthExplorer	30 m	
Soil texture	FAO Soil Portal	1:5 000 000	
Lithology	GeoData India		
Groundwater Level	India-WRIS	Station	2000-2023
Crop intensity	DES Agriculture Statistics	District	2000-2023

Net Sown Area		
Fallow, Barren,		
Forest land		
Population Density		
Rural population	Census of India	2011
Literacy rate		
Poverty		
Safe Water for drinking		
Sanitation		
Stunting		
Underweight	NFHS-5 (2019–21)	2019
Anaemia		
Clean fuel for cooking		
Electricity		

247

248 **Table 2** Factors selected for DV mapping with causes and their impact on drought

Vulnerability Indicators	Data	Impact Type	Drought Relevance
Exposure Index	Frequency of CDI	Indirect	The increase in drought frequency reduces crop productivity and has an impact on the socio-economic status
	Intensity of CDI	Indirect	High intensity of drought largely influences crop productivity and the economic status of the farmers
	Land surface temperature	Indirect	High LST indicates surface dryness and moisture deficiency
	Wind Speed	Indirect	Strong winds enhance evaporation and soil moisture loss
Sensitivity Index	Altitude	Indirect	Elevation influences rainfall distribution and runoff processes
	Soil texture	Indirect	Coarse soils retain less moisture and increase drought sensitivity
	Lithology	Indirect	Controls infiltration and GWS capacity
	GWL	Direct	Declining GWL increases DV
	Crop intensity	Indirect	High cropping pressure increases agricultural water demand
	Net Sown Area	Indirect	A larger cultivated area increases exposure to AD

Adaptive Capacity Indicators	Fallow, Barren, Forest land	Indirect	Barren/fallow land increases drought susceptibility, while forests reduce it
	NDVI	Indirect	Lower vegetation health reflects drought stress and reduced biomass
	NDWI	Indirect	Lower water index indicates moisture deficiency and surface dryness
	Population Density	Indirect	High population increases pressure on water resources
	Rural population	Indirect	Rural communities are more dependent on climate-sensitive livelihoods
	Literacy rate	Indirect	Higher literacy improves preparedness and adaptive capacity
	Poverty	Indirect	Poor communities have lower resilience to drought impacts
	Safe Water for drinking	Indirect	Limited access to water increases vulnerability during drought
	Sanitation	Indirect	Poor sanitation aggravates drought-related health risks
	Stunting	Indirect	Indicates chronic nutritional vulnerability under drought stress
	Underweight	Indirect	Reflects food insecurity and reduced adaptive resilience
	Anaemia	Indirect	Poor health conditions reduce coping capacity during drought
	Clean fuel for cooking	Indirect	Represents socio-economic resilience and living conditions
	Electricity	Indirect	Better infrastructure improves adaptive capacity during drought

4 Methodology

In the present study, the DV assessment was conducted using a systematic, multi-component geospatial modelling approach that incorporates physical drought indicators alongside socio-economic measures. The overall process involves five major steps (Fig. 2) as follows:

Step 1: Developing a CDI using four key drought components (meteorological, agricultural, hydrological, and GD) to capture the comprehensive drought situation, which further served as the primary input for exposure indices

Step 2: Collection of DV indicators based on literature reviews and data availability. Finally, these were classified into three key components, called Exposure, Sensitivity, and AC, based on the theoretical framework of IPCC. All thematic layers were prepared and processed in the ArcGIS environment.

Step 3: Historical drought-affected areas were compiled from government sources and randomly divided into training and testing datasets (70:30) for model development.

Step 4: Using statistical and ML approaches, Exposure Index (EI), Sensitivity Index (SI), and Adaptive Capacity Index (ACI) were generated. Finally, these three indices were combined to produce the DV Index (DVI).

Step 5: Efficiency of the statistical and ML models was analysed using Area under the ROC curve (AUC-ROC), overall accuracy, precision, recall, F1 score, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

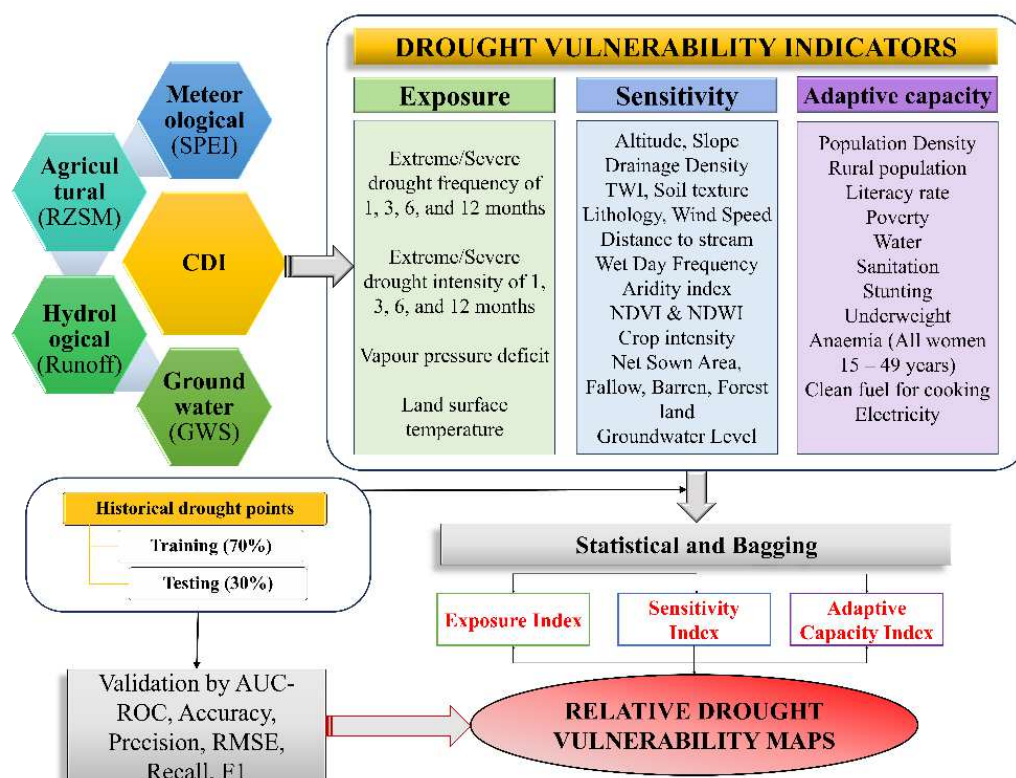


Fig. 2 Methodology flowchart

4.1 Development of Composite Drought Index

To develop a CDI for each grid cell, SPEI, RZSM, GWS, and runoff were considered. SPEI was selected because it incorporates both precipitation and PET, making it a robust indicator of MD [58], whereas RZSM reflects moisture availability in the root zone, making it a direct proxy for AD [51]. Runoff was included to represent SW availability and overall hydrological response [52]. Further, GWS accounts for the changes in subsurface water reserves and helps capture long-term drought conditions [59], [60]. Together, these four variables provide a multifaceted view of drought conditions, covering meteorological, agricultural, surface hydrological, and GW components. This approach considered four main steps: (1) computing the cumulative probability of each variable to transform the original time series into a uniform distribution ranging from 0 to 1; (2) fitting and selecting the most appropriate Copula function, then using it to build the joint cumulative probability series; (3) applying the empirical Kendall distribution function to the joint series to derive the probabilistic drought measure; and (4) standardizing that sequence to produce the final CDI. Climatic water deficit sequences (D_i) were calculated using monthly P and AET as in Eq. (1) [61]:

$$D_i = P_i - AET_i \quad (1)$$

Since the hydrometeorological variables exhibited strong non-normality and variable-specific distributional characteristics, a non-parametric probabilistic standardization approach was adopted instead of fitting predefined parametric probability distributions. To avoid uncertainties associated with fitting predefined parametric probability distributions, a non-parametric Gaussian kernel density estimator (KDE) was adopted [62]. The KDE approach estimates the probability density functions (PDFs) without assuming any predefined distribution for the data. In

this method, all observations contribute to estimating the PDF, with each observation's contribution determined by a kernel density function. Gaussian kernel functions were employed because of their widespread applicability and stable performance in hydrological and climatological studies [63], [64], [65], [66]. The PDF for a given observation x was estimated as:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x - x_i}{h} \right)^2 \right] \quad (2)$$

Where n denotes the number of observations, h represents the bandwidth parameter, and x_i corresponds to the observed values [67]. Observations closer to the target value contribute more strongly to the estimated density, whereas distant observations contribute more strongly to the estimated density, whereas distant observations contribute less. The bandwidth parameter was also estimated using Silverman's rule of thumb as:

$$h = 0.9 \min \left(\sigma, \frac{IQR}{1.34} \right) n^{-1/5} \quad (3)$$

where σ represents the standard deviation, and IQE denotes the interquartile range [67].

The cumulative distribution function (CDF) was subsequently obtained through numerical integration of the estimated PDF as given in eq (4) [62] :

$$F(x) = \int_{-\infty}^x \hat{f}(t) dt \quad (4)$$

The estimated cumulative probabilities were then transformed into standard normal variates using the inverse standard normal function as expressed in eq (5) [19]:

$$Z_i = \Phi^{-1}(F(x_i)) \quad (5)$$

where Z_i represents the standardized drought indices, including SPEI, equivalent Standardized Soil Moisture Index (ESSMI), Standardized Streamflow Index (SSI), and GDI, and Φ^{-1} denotes the inverse standard normal distribution function.

To examine how different statistical, probabilistic, machine-learning, and deep-learning approaches integrate the four drought components into a unified multidimensional drought signal, four complementary CDI construction approaches, such as PCA, RF, Copula, and SNAE (Fig. 3), were evaluated. These approaches were selected to represent different strategies for dimensionality reduction, dependence modelling, nonlinear ML integration, and deep-learning-based feature representation, respectively.

4.1.1 Formulating CDI Using PCA

The standardized drought indicators obtained through KDE were integrated using PCA. It is a multivariate statistical technique that transforms correlated variables into a smaller number of orthogonal principal components while preserving the maximum amount of variance contained in the original dataset. It has been widely adopted in developing CDI to monitor drought assessment [68], [69], [70], [71]. A correlation coefficient matrix of dimension $(p \times p)$ was the 1st step in developing PCA computation, where p is the number of input parameters. In this study, p is taken as four (SPEI, ESSMI, SSI, and GDI); thus, a 4×4 matrix was developed. A correlation matrix was then computed to quantify the relationships among the four drought variables [72]. Kaiser-Meyer-Olkin (KMO) test was adopted to check the suitability of the PCA datasets. It determines whether the observed correlation structure is sufficiently enough to extract reliable PCs. Higher KMO values indicate stronger common

variance among the variables and greater suitability for dimensionality reduction [73].

Further, eigenvalues and their corresponding eigenvectors were determined by eigen decomposition of the correlation matrix. The eigenvalues are the amount of variance explained by each PC, while the eigenvectors (loading coefficients) are the contribution of the original drought indicators to the extracted components. PCs are mutually orthogonal; thus, each component represents an independent source of variability within the dataset. The 1st PC (PC1) that contributed the greatest percentage of the total variance was chosen to estimate the weights of the four drought indicators. The squaring of the loading coefficient of the variable with respect to PC1 gave the relative contribution of the variable to the dominant mode of variability (Fig. 3a). After normalizing the square loadings, the resulting values were then used as the weights of the CDI as –

$$CDI_{y,m} = w_{SPEI,m} \times SPEI_{y,m} + w_{ESSMI,m} \times ESSMI_{y,m} + w_{SSI,m} \times SSI_{y,m} + w_{GDI,m} \times GDI_{y,m} \quad (6)$$

Where y and m represent the year and month, respectively. $CDI_{y,m}$ is the Composite Drought Index for a particular year and month. $w_{SPEI,m}$, $w_{ESSMI,m}$, $w_{SSI,m}$ and $w_{GDI,m}$ are the PCA-derived weight coefficients from the SPEI, ESSMI, SSI, and GDI, respectively [69].

4.1.2 Formulating CDI Using Copula

To account for nonlinear dependence among the drought indicators, a Copula-based CDI was also developed [74], [75], [76]. According to Sklar [77], any multivariate joint distribution can be represented using its marginal

distributions and an appropriate copula function. For an n-dimensional joint probability distribution function $F(x_1, x_2 \cdots x_n)$, the relationship between the joint distribution and marginal distributions can be expressed as :

$$F(x_1, x_2 \cdots x_n) = C[F_{x_1}(x_1), F_{x_2}(x_2), \dots, F_{x_n}(x_n)] \quad (7)$$

Where C denotes the Copula function and $F_{x_1}(x_1), F_{x_2}(x_2), \dots, F_{x_n}(x_n)$ represent the marginal cumulative distribution functions of the variables.

Four marginal probability distributions were utilized in this study, namely Gaussian, Logistic, generalized extreme value (GEV), and t location-scale distributions. The Kolmogorov-Smirnov (K-S) statistic proposed by Massey [78] was used to evaluate the fitting performance of these marginal distributions. Finally, three widely used Copula functions, namely the Gaussian, t, and Frank Copulas, were employed for modelling the joint distribution among drought indices. The cumulative probability sequences of SPEI, ESSMI, SSI, and GDI were integrated using the optimal Copula function to derive the joint cumulative probability sequence (p_i) as:

$$p_i = C(SPEI_i, ESSMI_i, SSI_i, GDI_i) \quad (8)$$

Finally, the copula-fitted cumulative probability sequence was transformed into standardized normal variables using the inverse normal transformation using equation (5) to derive the CDI (Fig. 3b).

4.1.3 Formulating CDI Using RF

RF regression model was employed to develop an ML-based CDI (RF-CDI). It is an ensemble learning technique that uses the output of many decision trees (DT) to make complex nonlinear predictions while minimizing the

possibility of overfitting through bootstrap aggregation and random selection of features. Due to its flexibility and robustness to multicollinearity and nonlinear interaction between variables, RF regression has become one of the most popular techniques used in drought predictions[79], [80], [81]. The predictor matrix (X) for each grid cell was formed as –

$$X_i = [SPEI_i, ESSMI_i, SSI_i, GDI_i] \quad (9)$$

For each grid, the time series data were randomly split into training and testing datasets (80:20) for performance evaluation of the models. Predictor variables were 1st standardized to increase the numerical stability of the training algorithm. The model was initialized using bootstrap sampling with out-of-bag (OOB) validation. The model was trained using 1000 ensemble trees to provide stable estimates. At each iteration, the OOB error, testing coefficient of determination (R^2), and root mean square error (RMSE) were calculated using Eqs. 10-13 to measure the convergence of the model and the optimal ensemble size (Fig. 3c).

$$OOB\ Error = 1 - OOB\ Score \quad (10)$$

$$OOB\ Score = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i^{OOB})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

where y_i and $\hat{y}_i$ denote the observed and predicted CDI values, respectively.

After training, the optimized RF model was run on the entire predictor dataset for each grid to calculate RF-CDI. The predicted CDI values were subsequently standardized to obtain a dimensionless drought index as follows

–

$$\text{RF-CDI} = \frac{\overline{\text{CDI}_{\text{RF}}} - \mu}{\sigma} \quad (14)$$

The relative significance of each drought indicator was assessed using the feature importance values from the trained RF algorithm. These values were calculated based on the decrease in prediction error attributable to each predictor and provide a quantitative estimate of the influence of MD, AD, HD, and GD on prediction.

4.1.4 Formulating CDI Using Spatial Neighbourhood Autoencoder (SNAE)

To construct a DL-based nonlinear CDI, the SNAE approach was considered by integrating KDE-derived drought indicators. While traditional autoencoders analyze each grid separately [82], this method incorporates local spatial information through a neighborhood aggregation layer, enabling the network to learn both the nonlinear relationships among drought indicators and the spatial dependence between neighboring grid cells [83]. The input datasets form a three-dimensional tensor as-

$$X \in R^{T \times N \times 4} \quad (15)$$

Where T is the number of time steps, and N is the number of grid cells.

To understand the effect of network complexity, five different architectures of encoders with depths ranging from one to five hidden layers were considered separately. The architecture with the minimum validation loss value was chosen for the computation of the CDI. The spatial neighborhood of the grid cells was determined by their geographical coordinates. A BallTree algorithm was applied with the Haversine distance measure to determine the neighboring grid cells within a 20 km radius [84]. The set of neighboring grids was used to build the spatial aggregation layer. This layer allows each grid cell to include information about its neighbourhood during feature extraction (Fig. 3d). Following the graph-based neighborhood aggregation principle [85], the aggregated feature representation for the i^{th} grid cell is expressed as

$$Y_i = W \left(\frac{\sum_{j \in N(i)} M_{ij} X_j}{\sum_{j \in N(i)} M_{ij}} \right) \quad (16)$$

where Y_i is the output feature vector of the i^{th} grid cell after neighborhood aggregation, X_j is the input feature vector of the j^{th} neighboring grid cell. $N(i)$ is the set of neighboring grid cells within the 20 km search radius. M_{ij} is the binary neighborhood mask, where $M_{ij} = 1$ if grid j belongs to the neighborhood of grid i , and 0 otherwise. W is the trainable weight matrix learned during model optimization. Following the neighborhood aggregation, channel attention was applied to adjust the importance of the selected spatial features. The attention coefficients were calculated using the fully connected layer along with the Softmax activation function to ensure positivity and normalization of the weights across all channels [86]. The weighted feature representations then served as the input to the next encoder layer. The deepest encoder layer gives the latent bottleneck representation for extracting the essential nonlinear spatial features

from all four drought indicators. The decoder reconstructs the original drought indicators from the latent bottleneck representation by means of fully connected layers along with the skip connections provided by the encoder. The network parameters were optimized by minimizing the reconstruction loss, which is defined as –

$$L_{recon} = \frac{1}{B N C} \sum_{b=1}^B \sum_{i=1}^N \sum_{K=1}^C (X_{b,i,k} - \overline{X_{b,i,k}})^2 \quad (17)$$

B is the batch size, N is the total number of grid cells, and C is the number of input drought indicators, $X_{b,i,k}$ is the original value of the k^{th} drought indicator for the i^{th} grid cell in the b^{th} training sample, and $\overline{X_{b,i,k}}$ are the SNAE model reconstructed values [82].

Adam optimizer with a learning rate of 0.001, a batch size of 28, and a maximum of 1000 training epochs was used to train the SNAE model. The data samples were randomly divided into training and testing sets (80:20). Upon completion of the training process, the encoder was applied to generate the latent space for each grid cell and at each time step. For architectures producing multiple latent feature channels, the bottleneck features were averaged across channels to obtain a single composite drought representation. Finally, the latent space features were normalized to generate the SNAE-CDI.

4.1.5 Validation of CDI models

Multivariate statistical analysis (PCA), probabilistic dependence modelling (Copula), ML-RF, and DL-SNAE approaches represent fundamentally different modelling paradigms. Owing to their distinct mathematical formulations, assumptions, and learning mechanisms, direct comparison based on model-

specific statistical measures is insufficient. Accordingly, the four approaches were evaluated within a common validation framework to determine which modelling paradigm most effectively represents the multidimensional and spatially heterogeneous drought signal of the MRB. Vegetation Health Index (VHI), Land Surface Temperature (LST), and Vapour Pressure Deficit (VPD) were selected to evaluate the CDI model's representation of drought conditions. Because none of these indicators represents an absolute ground-truth measure of drought, their relationships with the CDI were interpreted as measures of external consistency and drought-response agreement rather than direct validation against observations. VHI reflects vegetation stress resulting from moisture deficiency. It combines the Vegetation Condition Index (VCI) and the Temperature Condition Index (TCI), thereby accounting for both moisture availability and temperature-induced stress [87], [88]. The VCI was derived from the NDVI, which was obtained from the MODIS MOD13A2 Version 6.1 product [89]. LST was used to derive TCI, which was obtained from the MODIS Terra LST data (MOD11A1 V6.1) product [90]. All these data were derived in the GEE platform. The VHI was calculated as [91]:

$$VHI = \alpha \times VCI + (1 - \alpha) \times TCI \quad (18)$$

where α ($= 0.5$) is the weighting coefficient. The VCI and TCI were calculated as:

$$VCI = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \times 100 \quad (19)$$

$$TCI = \frac{LST_{max} - LST}{LST_{max} - LST_{min}} \times 100 \quad (20)$$

LST, which served as a proxy for the surface thermal state, and moisture deficiency were also considered. During drought conditions, reduced soil moisture limits evaporative cooling, resulting in elevated surface temperatures. For the atmospheric drought indicator, VPD was incorporated as higher VPD values intensify evapotranspiration and plant water loss, thereby increasing drought conditions. It was calculated from air temperature (T in $^{\circ}\text{C}$) and RH (%) data. VPD was calculated as:

$$VPD = e_s - e_a \quad (21)$$

where e_s and e_a represents saturation vapour pressure and actual vapour pressure, respectively, and is calculated as –

$$e_s = 0.6108 \exp\left(\frac{17.27T}{T + 237.3}\right) \quad (22)$$

$$e_a = e_s \times \frac{RH}{100} \quad (23)$$

The correlation (r) between the CDI models and VPD, LST, and VHI was calculated as –

$$r = \left(\frac{\sum_{i=1}^n (X_i - \bar{X}) (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \right) \quad (24)$$

where X_i is the CDI from different models at the i^{th} observation and Y_i is the VHI, VPD, or LST. $\bar{X}$ and $\bar{Y}$ are the respective mean values. n is the total number of observed values.

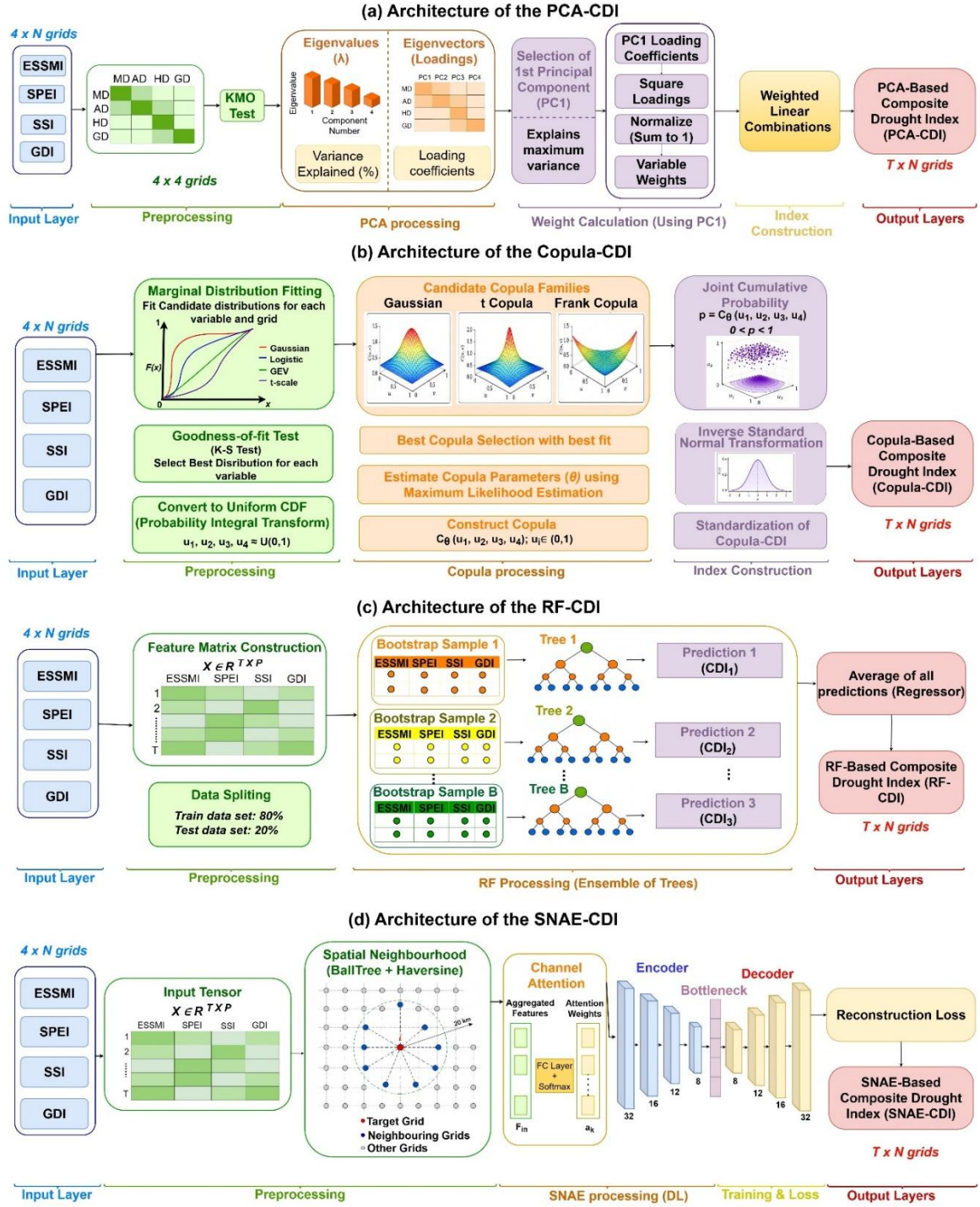


Fig. 3 Schematic workflow illustrating the construction of CDIs using four different approaches: (a) the statistical approach using PCA, (b) the probabilistic approach based on copula theory, (c) the ML approach using an RF regression model, and (d) the DL approach using the SNAE model. All these methods utilize KDE-standardized Precipitation-evapotranspiration data, soil moisture, runoff, and GW storage data as input to produce monthly grid-wise CDI

4.2 Exposure Indicators

El indicates the degree to which a region is subjected to drought [92]. Extreme & severe drought frequency and intensity of 1,3,6,12 months, LST, and VPD are some of the parameters considered as El in this study. The CDI from the best-performing model was used to calculate the mean intensity and frequency of extreme and severe drought. For this, run theory was applied to the time series CDI, and intermittent wet months should not exceed 3 months. The drought characteristics were subsequently analysed at 1,3,6,12 months, to capture short-, medium-, and long-term drought variability. Drought conditions were classified as extreme and severe when the CDI values fell below -2 and ranged between -1.5 and -1.99, based on the drought classification values [19]. Drought frequency represents the occurrence of drought events during the study period, while intensity is the average severity of all identified drought events and is computed using the formulas [93]:

$$DF_i = \left(\frac{N_i}{n} \right) \times 100 \quad (25)$$

$$DI_i = \left(\frac{1}{N_i} \right) \sum_{j=1}^{N_i} CDI_j \quad (26)$$

where DF_i and DI_i represent the drought frequency and intensity, respectively, for time scale i . N_i is the number of drought months at different time scales, and n is the total number of observations.

4.3 Sensitivity Indicators

The SI is an important component of the Vulnerability Index mapping as it measures how susceptible areas are to drought conditions. It serves as a crucial indicator of how a system responds to widespread stressors [94]. Altitude, Slope, Drainage Density, Topography Wetness Index (TWI), Soil texture, Lithology, Wind

Speed, Distance to stream, Wet Day Frequency (WDF), Aridity index (AI), NDVI, NDWI, Crop intensity (CI), Net Sown Area (NSA), Fallow, Barren, forest land, and GWL were considered for SI calculation. Soil texture is an important parameter, as coarse soil won't retain moisture and will allow it to move deeper into the soil. Thus, the upper part of the soil will be dry, becoming prone to drought. Meanwhile, if the soil is fine, it will retain the moisture in the upper layer, preventing drought [95]. Lower TWI indicated reduced soil moisture accumulation and weaker water retention capacity, making an area more vulnerable to drought. The composition of the bedrock also affects DV, as it determines the water storage capacity through weathering processes [96]. Drainage density affects the hydrological properties of a drainage basin. Areas with low drainage density are more vulnerable to droughts, as low drainage density implies less runoff and water availability for consumption [97]. Areas away from rivers are more susceptible to drought conditions, as the supply of water decreases with increasing distance [98]. Areas with lower WDF are more prone to drought conditions. A decline in GWL affects GW availability, causing an increase in drought conditions [99]. NDWI is also an important indicator for drought detection as it represents the water condition of an area [100]. AI can be calculated from potential evapotranspiration (PET) and AET as follows:

$$AI = \frac{PET - AET}{PET} \times 100 \quad (27)$$

As the AI increases, the dryness will also increase, leading to drought-like conditions [101]. CI is the ratio of net cropped area to gross cropped area, and as drought intensity increases, crop intensity also increases [102]. If the NSA increases, then drought will affect agriculture more. Barren land is more likely to be affected by drought, and if it is covered by forest or vegetation, then it can be protected from dryness [103]. Bare terrain is more susceptible to drought. NDVI is the normalised

value of vegetation and ranges from -1 to 1. If it is closer to 1, then it indicates large plant cover, and when it is closer to -1, it represents barren land [104]. They were calculated as :

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (28)$$

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (29)$$

where NIR is near-infrared reflectance.

4.4 Adaptive capacity (AC)

Vulnerability is often regarded as the inability of people to effectively counter a widespread stressor [105]. As a result, AC is the ability of a population or system to cope with, adapt to, and overcome negative consequences of drought through socio-economic, demographic, nutritional, and infrastructural resources [106]. The climate-population interactive effects have been the dominant factors driving recent changes in drought exposures [107]. An increase in population density has posed immense pressure on water security [108]. Therefore, in this study, demographic factors like population density, rural population, literacy rate, and poverty were considered to characterize the human dimensions of DV and exposure. Health indicators like child stunting, underweight, and women's anaemia were also considered as they represent important dimensions of social and health vulnerability under drought conditions. Child stunting and underweight are widely recognized indicators of chronic malnutrition and long-term food security, while anaemia reflects poor nutritional status and reduced physiological resilience [109], [110]. All these indicators were derived from district-wise census data and the National Family Health Survey (NFHS) and then converted into raster layers in ArcGIS. These layers were further standardized. Indicators contributing

positively to adaptive capacity were standardised using the Fuzzy Large function, while those negatively affecting were standardised using the Fuzzy Small function. The standardized layers were subsequently integrated using the Bagging ensemble technique to generate the final Adaptive Capacity Index (ACI).

4.5 Statistical technique for Drought Vulnerability Mapping

A statistical DVI model involved the integration of three main aspects of vulnerability (EI, SI, ACI) as proposed by the Intergovernmental Panel on Climate Change [35]. Before integration of the indicators, all indicators were normalized to ensure uniformity of scale for elimination of dimensional inconsistency. The composite DVI was obtained as [111], [112] -

$$DVI = (EI + SI) - ACI \quad (30)$$

With an increase in exposure and sensitivity, the vulnerability will increase, but it decreases with higher AC. The resulting vulnerability values were classified into five categories that include very low, low, moderate, high, and very high.

4.6 Bagging for Drought Vulnerability Mapping

Bagging (Bootstrap Aggregating) is an ensemble learning method proposed by Breiman [113], which enhances prediction accuracy by decreasing the variance of the individual decision trees through bootstrap sampling and aggregation. This algorithm creates different bootstrap samples from the initial training dataset, then fits each sample with the independent base learner, and finally aggregates their results to achieve a better model. Due to its resiliency to overfitting and capacity to learn complicated non-linear functions, it has been utilized in environmental hazard and susceptibility mapping applications [114], [115], [116].

The drought inventory has been prepared based on the district level information available about the drought of 2016 from the drought affected districts information available from the Government of Odisha

(<https://srcodisha.nic.in/annualReport/oF24XdPzAnnualReport2015-16.pdf>) and JNA report of Chhattisgarh (<https://sphereindiablog.wordpress.com/wp-content/uploads/2016/09/jna-report-chhattisgarh-drought.pdf>). The drought affected district boundary information was spatially overlapped with the 1,896 analysis grids used for the study. All the grids were assigned a drought label based on their spatial presence (drought = 1, non-drought = 0). The binary drought inventory was integrated with the selected EI, SI, and ACI indicators to construct an ML dataset. Consequently, each of the 1,896 grid cells consisted of one binary response variable and the corresponding predictor variables.

The complete dataset was divided randomly into 70% training data and 30% testing data. The training data were used to construct the Bagging classifier, whereas the testing data were used for the evaluation of the constructed model. In the training process, several bootstrap samples were created using the training data via a random sampling with replacement technique. Each bootstrap sample was then used to train a single decision tree, and the final prediction of the DV class was obtained by aggregating the predictions of all trees through majority voting. The final prediction of the Bagging classifier is given by [113]

$$\hat{Y}(x) = mode\{f_1(x), f_2(x), \dots, f_B(x)\} \quad (31)$$

where $\hat{Y}(x)$ is the predicted DV class, $f_b(x)$ is the prediction of the b^{th} decision tree, and B is the total number of trees in the ensemble. The trained model was subsequently applied to all 1,896 grid cells to produce the spatial DVM. Finally, the predicted vulnerability values were classified into five classes (very low, low, moderate, high, and very high vulnerability) using the Jenks Natural Breaks classification method in ArcGIS.

4.7 Model performance evaluation

The predictive accuracies of the statistical and Bagging models were tested on an independent dataset that consisted of 30% of the total data. The model's performance was calculated based on the confusion matrix and several widely adopted evaluation metrics, such as Accuracy, Precision, Recall, F1-score, RMSE, and Area Under the Curve-Receiver Operating Characteristic (AUC-ROC) (Eqs. 32-38). The confusion matrix summarizes the classification results in terms of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), from which the evaluation metrics were computed [117]. The overall classification accuracy, precision, and recall were calculated as -

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (32)$$

$$Precision = \frac{TP}{TP + FP} \quad (33)$$

$$Recall = \frac{TP}{TP + FN} \quad (34)$$

The harmonic mean of precision and recall was determined using the F1-score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (35)$$

The prediction error was quantified using the Root Mean Square Error (RMSE) -

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (36)$$

where y_i and $\hat{y}_i$ denote the observed and predicted classes, respectively, and n is the total number of observations [118]. The ROC curve was used to evaluate the discrimination capability of the developed models by plotting the True Positive Rate

(TPR) against the False Positive Rate (FPR) at different classification thresholds [119].

The TPR and FPR were computed as -

$$TPR = \frac{TP}{TP + FN} \quad (37)$$

$$FPR = \frac{FP}{FP + TN} \quad (38)$$

The AUC provides a threshold-independent measure of classification performance, with values ranging from 0.5 (random classification) to 1.0 (perfect classification). Higher AUC values indicate better discrimination between drought and non-drought classes.

5 Results

5.1 Validation of KDE-based probabilistic Standardization

The Q-Q plots of the original D, RZSM, runoff, and GWS, together with their corresponding standardized indices, including SPEI, ESSMI, SSI, and GDI, showed the effectiveness of the non-parametric KDE-based probabilistic standardization framework in transforming the original hydroclimatic variables into approximately standardized normal variables. The original variables showed pronounced curvature and substantial deviations from the reference line. Among the variables, D and runoff displayed the strongest non-normal characteristics, particularly in the lower and upper quantiles, indicating high skewness and heterogeneous distributional behaviour (Fig. 4 a,e). RZSM and GWS also deviated considerably from the theoretical normal line, although the deviations were comparatively less pronounced (Fig. 4 c,g). These results confirm that these variables exhibit complex statistical characteristics that can't be adequately represented by a single predefined parametric probability distribution. After applying the KDE-based probabilistic transformation, the corresponding standardized indices showed substantial improvement, exhibiting

near-linear behaviour over the central quantile ranges. Compared with the raw variables, these standardized indices demonstrated improved symmetry and reduced distributional irregularities, thereby enhancing their suitability for multivariate drought analysis. The standard deviations of SPEI, ESSMI, SSI, and GDI exhibited values close to 1, indicating effective normalization and improved comparability among drought-related variables (Fig. 4 b,d,f,h).

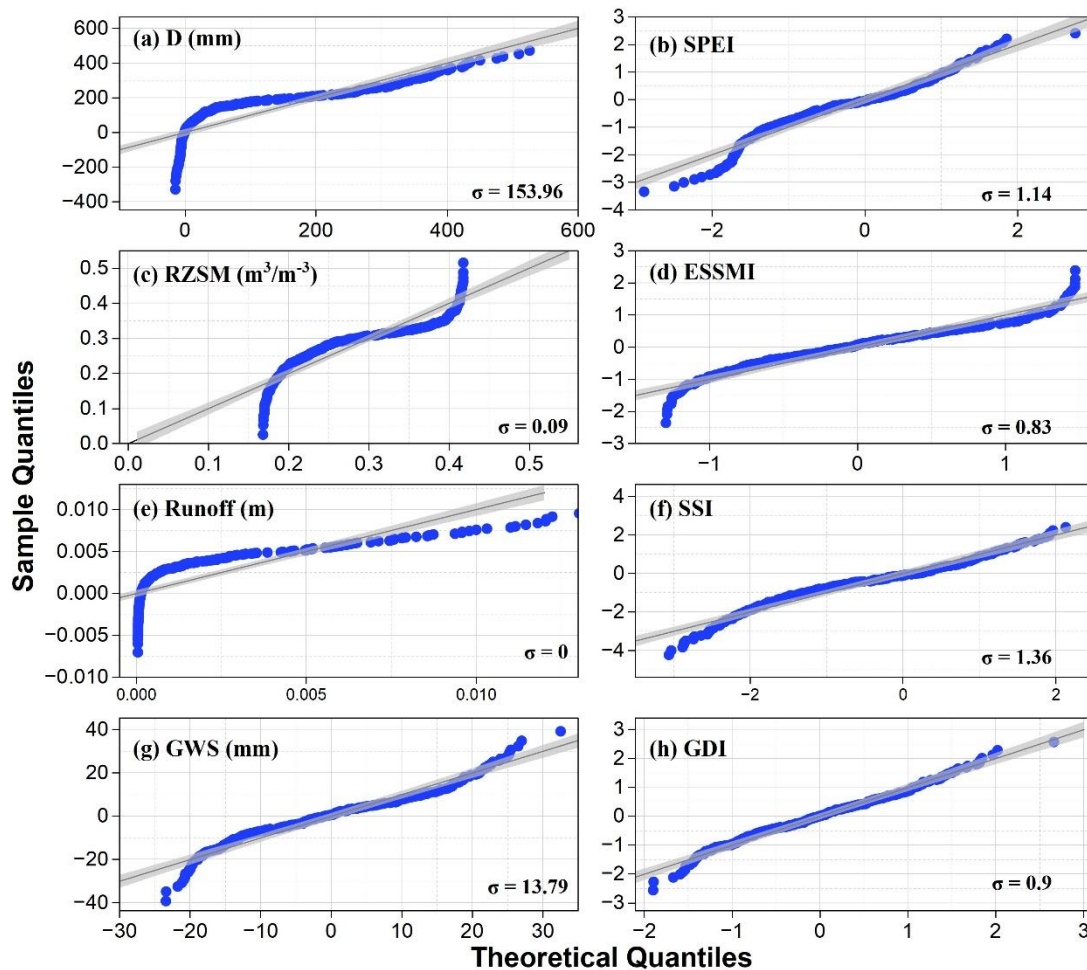


Fig. 4 Normal Q-Q plots showing the probabilistic standardization of the original hydroclimatic variables and their corresponding KDE-based standardized drought indices: (a) climatic water deficit (D), (b) SPEI, (c) root zone soil moisture (RZSM), (d) ESSMI, (e) runoff, (f) SSI, (g) groundwater storage (GWS), and (h) GDI. The original variables exhibit substantial departures from the theoretical normal distribution, whereas the standardized indices show improved agreement with the theoretical normal quantiles, indicating effective probabilistic standardization using the non-parametric KDE framework

5.2 Performance evaluation of CDI models

5.2.1 Evaluation of PCA-based CDI

To construct the CDI using the four KDE-standardized drought indices, the PCA method has been used. Before PCA, the correlation among the drought indicators was analysed, showing a positive correlation (0.31 to 0.89) (Fig. 5a). This indicated the existence of common variability, which is needed for performing PCA. The strongest relationship is observed between AD and HD ($r = 0.89$), while MD and AD have $r = 0.82$ and HD and GD have $r = 0.73$, indicating a strong connection between AD, HD, and GD. In contrast, the lower correlation between MD and GD ($r = 0.31$) corresponds to a delayed response of the GW system to the MD, which can be explained by temporal propagation of drought effects from the atmosphere to GW. The suitability of the datasets was further analyzed with the KMO test for all grids, ranging from 0.57 to 0.73, with nearly 97% of the grids exhibiting >0.60 , suggesting acceptable sampling for PC extraction (Fig. 5b). Further, the eigenvalue distribution (Fig. 5c) showed a sharp decline after 1st principal component (PC1), with mean eigenvalue of 3.7 (greater than the Kaiser criterion), whereas the remaining PCs have eigenvalues <1 . This explains the dominance of PC1 in capturing the dominant variability among the four drought indicators. The cumulative explained variance (Fig. 5d) also illustrates that PC1 alone explained approximately 80% of the total variance, while the cumulative variance is $>95\%$ after the inclusion of PC2. Furthermore, the relation between variables and PCs revealed strong positive contributions from GD and HD, moderate positive contribution from AD, and a contrasting negative contribution from MD, consistent with the typical drought propagation sequence from meteorological deficits to hydrological and GW responses (Fig. 5e). Furthermore, the boxplot of PC1-derived weights (Fig. 5f) highlighted the relative contributions of each

indicator, with AD and HD showing the highest mean weights (approximately 0.30), followed by GD (around 0.19), and MD exhibiting the lowest weight (approximately 0.22), reflecting their differential influence on the dominant drought variability mode. These results collectively validate the use of PC1 scores as a robust CDI that captures the integrated drought signal with minimal information loss, providing an effective tool for comprehensive drought monitoring and assessment across the study region.

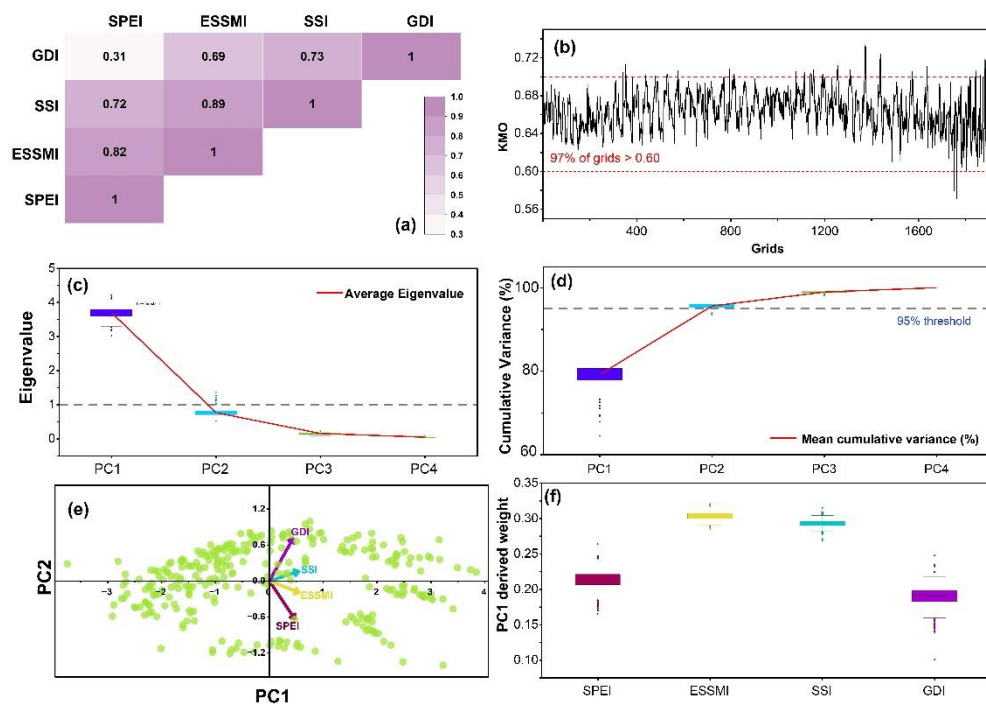


Fig. 5 Principal Component Analysis (PCA) for the development of the CDI. (a) Correlation matrix of the four drought indicators, (b) Kaiser-Meyer-Olkin (KMO) test results, (c) Eigenvalues of the principal components, (d) Cumulative variance explained by successive principal components, (e) PCA loading biplot (PC1-PC2), (f) PC1-derived weights of the drought indicators used for CDI construction

5.2.2 Evaluation of copula-based CDI

The goodness-of-fit of the marginal probability distributions for individual indices was assessed using the K-S statistic for 1896 grids, and based on the results, the optimal probability distribution for each grid cell was selected. Lower K-S values indicate better agreement between the empirical and theoretical cumulative

distribution functions. For SPEI, the Gaussian showed lower median K-S statistics, whereas the Logistic distribution exhibited larger K-S values, exceeding 0.11 in several grids (Fig. 6a). Again, for ESSMI, the Gaussian distribution exhibited lower median values compared to other distributions. The GEV distribution showed several outliers exceeding 0.14, indicating unstable fitting performance across multiple grid cells (Fig. 6b). In the case of SSI, the GEV distribution produced the lowest median K-S statistic (Fig. 6c), while for GDI, the Gaussian, GEV, and t distributions demonstrated relatively similar median K-S statistics (Fig. 6d).

Further, Fig. 6e illustrates the spatial distribution of the optimal copula functions selected for CDI development for each grid. The Frank model dominated in most grid cells (923), particularly across the central and western regions of the area. In contrast, t-copula was selected for 868 grids and showed widespread occurrence across the eastern and southern regions. The Gaussian copula was selected for comparatively fewer grid cells (105 grids). The limited occurrence of the Gaussian copula indicates that the linear dependence structures were insufficient to adequately characterize the complex multivariate drought relationships across the majority of the study area. Therefore, instead of applying a single copula model uniformly across the entire study area, the present study adopted a grid-wise adaptive Copula selection framework.

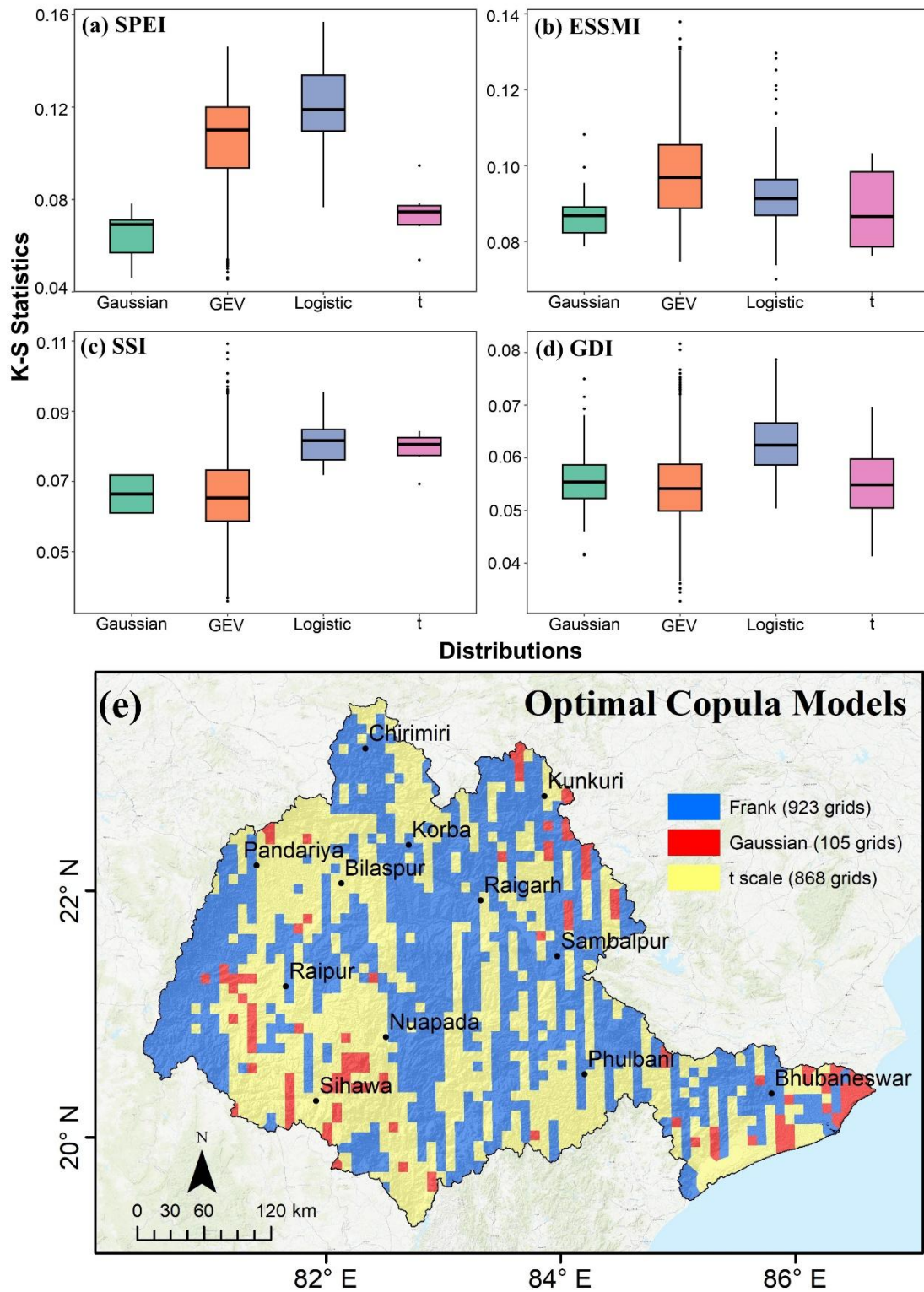


Fig. 6 Marginal distribution fitting and spatial distribution of the optimal copula families used for constructing the Copula-based CDI. Boxplots of the Kolmogorov-Smirnov (K-S) statistics obtained for four candidate marginal distributions (Gaussian, Generalized Extreme Value (GEV), Logistic, and Student's t) fitted to (a) SPEI, (b) ESSMI, (c) SSI, and (d) GDI. Lower K-S values represent superior distributional fit. (e) Spatial distribution of the optimal copula families selected for

1,896 grid cells, illustrating the dominance of the Frank copula, with localized occurrences of the Gaussian and Student's t copulas across the MRB

5.2.3 Evaluation of RF-based CDI

The RF model was developed independently for each of the 1,896 grid cells using the four KDE-standardized drought indicators as predictor variables. Since RF models were developed for each grid cell, the performance of the entire model was evaluated by summarizing the statistical distributions of the training and testing metrics rather than presenting individual model diagnostics. The distribution of the training, testing, and OOB coefficients of determination displayed consistently high predictability within the basin (Fig. 7a). The median value of the training R^2 was higher than 0.99, which means that the RF model was successful in capturing the nonlinear relationships between the four drought indicators used for modeling. The median R^2 for testing also showed high values (around 0.988), suggesting excellent generalization capability and minimal degradation in predictive accuracy for unseen data. Also, the OOB score, which is the internal cross-validation metric derived using bootstrap resampling, had a median value of about 0.983, showing the robustness and consistency of the ensemble learning technique. The relatively small difference among the training, testing, and OOB statistics indicates that the RF model did not exhibit noticeable overfitting despite being trained independently for each grid cell. The prediction errors thus provided additional evidence about the consistency of the RF-based CDI (Fig. 7b). The median value of the training RMSE was roughly 0.075, whereas that of the testing RMSE increased to 0.084, which means that the predictive power was quite consistent during the independent validation process. Although a limited number of grids exhibited relatively larger testing errors, the overall

interquartile range remained narrow, demonstrating that the RF model produced stable predictions throughout most of the study area.

The relative importance of the predictor variables showed considerable disparity in their contributions to the development of CDI (Fig. 7c). ESSMI was found to be the most significant predictor variable, with the highest median value of feature importance (0.50), indicating that SM conditions provided the dominant information for estimating the integrated drought response. The second highest importance level was shown by SSI with a median value of 0.34, which underscores the significant role of streamflow anomalies in defining HD. In comparison, SPEI demonstrated considerably lower importance (0.10), with GDI showing the lowest (0.05). The reduced importance of GD reflects its delayed response relative to MD and HD processes, resulting in greater redundancy with the antecedent drought indicators already represented by ESSMI and SSI.

5.2.4 Evaluation of SNAE-based CDI

Network depth is one of the major determinants governing the representational capability of DL models, as it can capture hierarchical nonlinear features from input data. To examine the effect of network complexity on the representation of the four KDE-based drought indicators, five different SNAE structures with one to five encoder layers were examined. The loss functions for training and validation are provided in Fig. 7d-e. The loss functions showed smooth and stable convergence during the optimization process. This revealed that the utilized neighbourhood aggregation technique along with the attention mechanism was capable of learning the nonlinear interactions among various drought indicators. The training losses (Fig. 7d) experienced a steep decline during the initial training process, followed by a steady decline till convergence. The steep decline showed that the SNAE model was able to

learn the dominant nonlinear interactions between the four drought indicators in a few optimization steps. However, distinct differences were observed in the convergence characteristics among the evaluated architectures. The enlarged inset showed that the three-layer encoder experienced the steepest decline in reconstruction loss, converging much faster than the other architectures. In contrast, the one- and two-layer encoders converged more gradually, suggesting that shallow architectures possess insufficient hierarchical representation to efficiently capture the complex nonlinear interactions and spatial dependence among neighbouring drought grids. The same trend was also noticed for the validation dataset (Fig. 7e). The validation loss decreased consistently throughout the optimization process and closely followed the training loss for all five architectures, indicating excellent generalization capability and the absence of noticeable overfitting despite training for 1000 epochs. The narrow separation between the training and validation curves demonstrates that the spatial neighbourhood aggregation and channel attention mechanisms effectively regularized the network while preserving the essential drought information. At the early stages of optimization, the three-layer encoder again showed the most rapid decline in validation loss. With around 150 training epochs, all architectures converged to the same validation loss level. Considering both convergence behaviour and computational efficiency, the three-layer encoder architecture was selected for the development of the final SNAE-CDI.

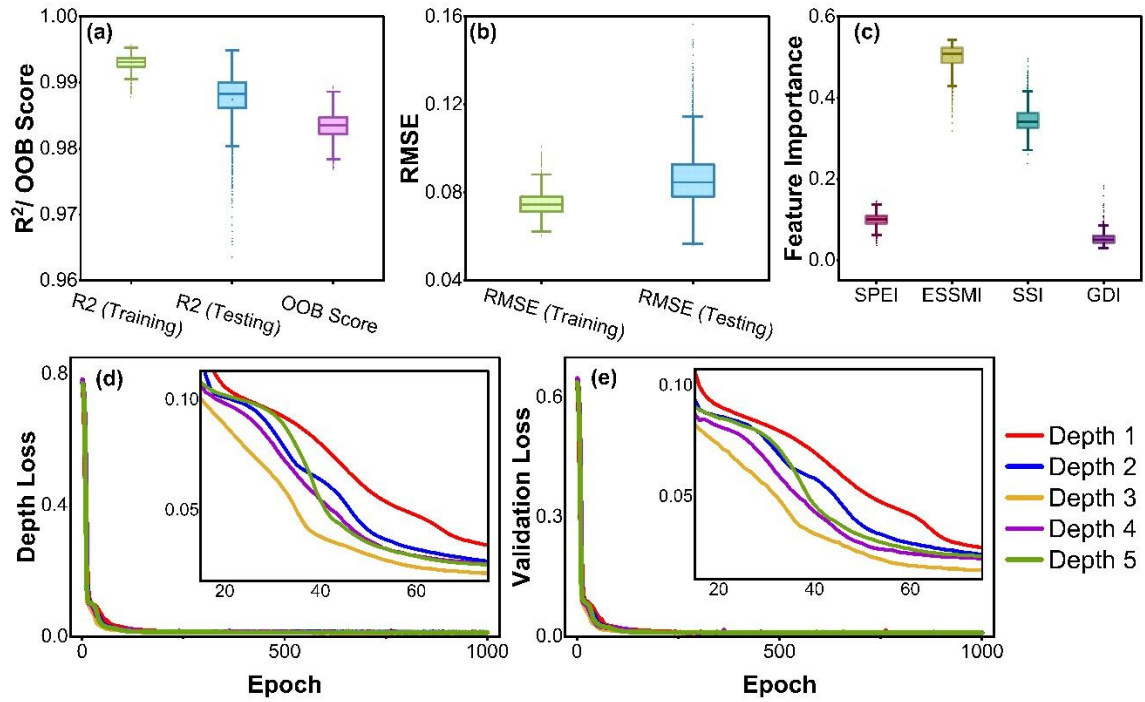


Fig. 7 Optimization and performance evaluation of the RF- and SNAE-based composite drought index models. (a) Training and testing R^2 values together with the out-of-bag (OOB) score of the RF model. (b) Training and testing RMSE of the RF model. (c) Relative feature importance of the input drought indicators (SPEI, ESSMI, SSI, and GDI) used in the RF model. (d) Training loss and (e) validation loss of SNAE architectures with encoder depths ranging from one to five hidden layers

5.2.5 Comparison of CDI models

The performance of the four CDI models was evaluated at both seasonal (MAM, JJAS, ON, and DJF) and accumulation-period (1-, 3-, 6-, and 12-month) scales using VHI, LST, and VPD (Fig. 8). The seasonal analysis was conducted to check whether each CDI model can capture drought conditions in contrasting hydroclimatic conditions throughout the annual cycle or not. Further, the accumulation period analysis was done to evaluate the response of the models to drought persistence over different temporal scales. Higher values of CDI indicate wet conditions and healthy vegetation; a positive correlation of CDI with VHI is anticipated. On the other hand, LST and VPD tend to rise during droughts due to reduced evaporative cooling

and increased evaporative demand. Hence, the negative correlations between CDI and LST and VPD suggest better model performance.

Among the evaluated approaches, the Copula model exhibited the most consistent agreement with VHI throughout the study period (Fig. 8a). Seasonal correlation coefficients were 0.83 (MAM), 0.30 (JJAS), 0.77 (ON), and 0.91 (DJF), while the accumulation-period correlations remained consistently high at 0.84, 0.86, 0.86, and 0.82 for the 1-, 3-, 6-, and 12-month CDIs, respectively. Unlike the other models, the Copula approach maintained relatively uniform performance across both seasonal and temporal scales, suggesting that it effectively captured the dependence structure among all drought components without substantial loss of information. The SNAE model also showed excellent agreement with VHI, although its performance varied more noticeably across the different seasons. Seasonal correlations were 0.79 (MAM), 0.26 (JJAS), 0.58 (ON), and 0.90 (DJF). However, the model exhibited a substantial improvement with increasing accumulation period. The strongest relationship among all evaluated models was observed for the 6-month CDI ($r = 0.96$), closely followed by the 3-month CDI ($r = 0.95$). The RF model also demonstrated strong agreement with vegetation dynamics. Seasonal correlations of 0.84 (MAM), 0.23 (JJAS), 0.80 (ON), and 0.92 (DJF) were obtained, while the corresponding accumulation-period correlations were 0.79, 0.79, 0.78, and 0.78 for the 1-, 3-, 6-, and 12-month CDIs, respectively. Although the RF model performed comparably to the Copula model during the seasonal analysis, particularly during DJF, the correlations gradually declined with increasing accumulation period, indicating relatively lower temporal consistency than the Copula and SNAE models.

The relationship between CDI and VPD was consistently negative for all modelling approaches (Fig. 8b). Similar to the VHI validation, the Copula, SNAE, and

RF models exhibited substantially stronger agreement with VPD than the PCA model. The Copula model showed the strongest and most stable performance, with seasonal correlations of -0.90 (MAM), -0.62 (JJAS), -0.94 (ON), and -0.92 (DJF). The accumulation-period correlations remained consistently high at -0.86, -0.89, -0.92, and -0.81 for the 1-, 3-, 6-, and 12-month CDIs, respectively. The SNAE model also demonstrated strong correspondence with VPD, with seasonal correlations of -0.79, -0.57, -0.72, and -0.86 during MAM, JJAS, ON, and DJF, respectively. Similar to the VHI results, the strongest agreement was observed at the 6-month accumulation period ($r = -0.94$), followed by -0.85 (3-month) and -0.75 (1-month). The RF model also produced consistently strong negative correlations across both seasonal and accumulation-period analyses, exhibiting performance comparable to the Copula model during most seasons but with relatively greater variability across the accumulation periods.

The comparison between CDI and LST also revealed predominantly strong negative correlations across all modelling approaches (Fig. 8c). The Copula model again demonstrated the most consistent performance, producing seasonal correlations of -0.80 (MAM), -0.06 (JJAS), -0.77 (ON), and -0.90 (DJF). The corresponding accumulation-period correlations were -0.62, -0.62, -0.62, and -0.82 for the 1-, 3-, 6-, and 12-month CDIs, respectively. The relatively stable correlations across both seasonal and temporal scales further demonstrate the robustness of the Copula-derived CDI in representing thermal responses to drought. The SNAE model exhibited seasonal correlations of -0.74 (MAM), 0.08 (JJAS), -0.59 (ON), and -0.85 (DJF), while the accumulation-period correlations increased to -0.78, -0.86, -0.90, and -0.76 for the 1-, 3-, 6-, and 12-month CDIs, respectively. The strongest relationship was again observed for the 6-month CDI ($r = -0.90$), confirming the superior

performance of SNAE in representing prolonged drought conditions. The RF model also demonstrated good agreement with LST, particularly during the seasonal analysis. However, the correlation coefficients gradually weakened with increasing accumulation period, indicating comparatively lower consistency than the Copula and SNAE models.

PCA had the lowest correlation values for all validation metrics regardless of season or accumulation period. The lower performance may be due to inherent limitations of PCA, which is based on linear orthogonal transformations and assumes linear relationships between the variables involved [120]. However, drought evolves from complex nonlinear interactions between meteorological, agricultural, hydrological, and GW processes. Thus, PCA does not capture these nonlinear relationships well enough, resulting in lower correlations with VHI, LST, and VPD.

The JJAS season showed the minimum correlation coefficients for VHI, VPD, and LST variables. The southwest monsoon period is associated with heavy precipitation, cloud cover, and soil moisture, reducing drought stress conditions in the basin. During this period, the vegetation quickly recovers, reducing the sensitivity of the VHI to drought changes [121]. At the same time, increased ET reduces LST and VPD [122], [123], decreasing the relationship with drought conditions. Moreover, the dominance of short-term meteorological variability during the monsoon can mask cumulative drought signals, resulting in lower correlations between CDI and the independent validation indicators.

Further grid-wise correlation analysis was also performed (Fig. 8d-f) to evaluate the spatial robustness of the CDI models across 1,896 grids within the basin. The correlation coefficients were classified into four performance categories, namely poor ($r < 0.5$), Moderate ($0.5 \leq r < 0.6$), Good ($0.6 \leq r < 0.7$), and Excellent ($r \geq 0.7$). For

916 VPD and LST, the classes were defined as Poor ($r > -0.5$), Moderate ($-0.6 < r \leq -0.5$),
 917 Good ($-0.7 < r \leq -0.6$), and Excellent ($r \leq -0.7$). The grid-wise comparison with VHI (Fig.
 918 8d) revealed substantial differences among the four CDI models. The PCA model
 919 consistently exhibited the weakest spatial performance (100% of the grids under the
 920 poor category) for the 1-, 3-, and 6-month CDIs, and 98% remained poor even for the
 921 12-month CDI. The Copula model demonstrated a marked improvement with
 922 increasing accumulation period. For the 1-month CDI, nearly 67% of the grids were
 923 classified as poor, whereas 31% exhibited moderate correlations and only 2% were
 924 categorized as good. However, the proportion of good and excellent correlations
 925 increased substantially to 35% for the 3-month CDI, 67% for the 6-month CDI, and
 926 76% for the 12-month CDI. Notably, the 12-month Copula CDI contained 42% of the
 927 grids within the excellent category, including 5% with correlation coefficients exceeding
 928 0.8, representing the highest spatial agreement among all models. The RF model
 929 exhibited a similar trend, with the proportion of poor correlations decreasing
 930 progressively from 55% at the 1-month timescale to 11% for the 12-month CDI.
 931 Concurrently, the percentage of grids classified as good or excellent increased from
 932 only 7% at the 1-month scale to 35%, 56%, and 68% for the 3-, 6-, and 12-month
 933 CDIs, respectively. The RF model therefore showed considerable improvement in
 934 representing vegetation dynamics with increasing drought accumulation. The SNAE
 935 model also demonstrated progressive enhancement in spatial performance with
 936 increasing accumulation period. While the 1-month CDI was dominated by poor
 937 correlations (100%), the proportion of poor grids decreased to 94%, 65%, and 58% for
 938 the 3-, 6-, and 12-month CDIs, respectively. Simultaneously, the percentage of grids
 939 exhibiting good and excellent agreement increased from 5% for the 6-month CDI to
 940 approximately 12% for the 12-month CDI, indicating improved representation of

vegetation conditions at longer accumulation periods, although its spatial performance remained lower than that of the Copula and RF models.

The validation against VPD (Fig. 8e) produced the strongest overall spatial agreement among the three independent drought indicators. The PCA model again exhibited extremely poor performance, with virtually 100% of the grids remaining within the poor category across all accumulation periods. The Copula model consistently demonstrated excellent spatial performance. Approximately 46% of the grids were classified within the excellent category for the 1-month CDI, increasing to 66% and 73% for the 3- and 6-month CDIs, respectively, before remaining high (57%) for the 12-month CDI. More than 95% of the grids were classified within the good and excellent categories for the 6-month CDI, indicating exceptional spatial representation of atmospheric drought conditions. The RF model also exhibited excellent spatial performance, with 41%, 56%, 76%, and 49% of the grids classified within the excellent category for the 1-, 3-, 6-, and 12-month CDIs, respectively. Similar to the Copula model, the 6-month CDI provided the strongest spatial agreement, with more than 94% of the basin classified within the good or excellent categories. The SNAE model exhibited comparatively weaker performance than the Copula and RF models but still showed noticeable improvement with increasing accumulation period. The proportion of poor correlations decreased from 100% for the 1-month CDI to approximately 61% for the 6- and 12-month CDIs, while the proportion of good and excellent correlations increased to 15% and 19%, respectively, at the longer accumulation periods.

The grid-wise validation against LST (Fig. 8f) exhibited trends comparable to those observed for VHI. The PCA model again showed the weakest performance, with nearly all grids remaining within the poor correlation category throughout all accumulation periods. The copula model exhibited the greatest improvement with

increasing accumulation period. The proportion of poor correlations declined substantially from 92.39% for the 1-month CDI to 29.42% for the 12-month CDI, while the percentage of grids classified within the good and excellent categories increased from less than 1% to approximately 42%. The RF model showed a similar spatial pattern, although with slightly lower overall performance than the Copula model. Poor correlations decreased from 92.45% to 37.98%, whereas the proportion of good and excellent correlations increased to approximately 31% at the 12-month accumulation period. The SNAE model also showed progressive improvement with increasing accumulation period, although the enhancement was comparatively smaller than that of the Copula and RF models. Poor correlations declined from 100% at the 1-month timescale to 88% at the 12-month CDI, while only a small proportion of grids attained good agreement.

This validation analysis consistently demonstrated the superior performance of the Copula-derived CDI over the PCA, RF, and SNAE models. The Copula model maintained strong and stable correlations with VHI, VPD, and LST across seasonal and accumulation-period analyses, while also exhibiting the highest proportion of grids classified within the good and excellent correlation categories across the 1,896 grid cells. Although the SNAE model achieved the highest correlations at the 3- and 6-month accumulation periods, the Copula model showed greater overall spatial consistency. Therefore, the Copula-derived CDI was selected for all subsequent analyses.

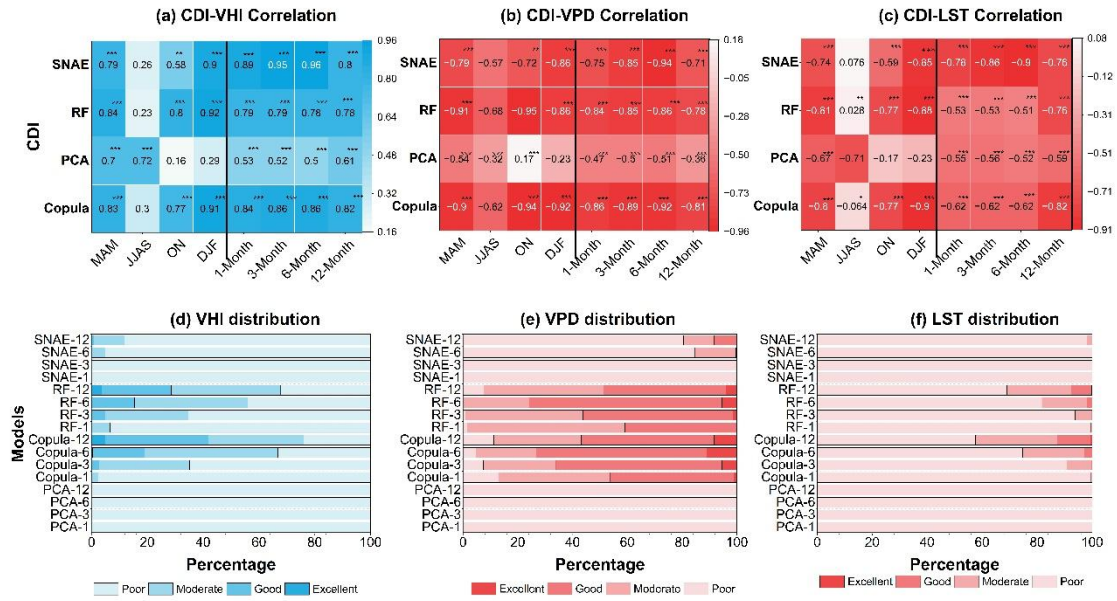


Fig. 8 Performance evaluation of the CDI models using independent drought indicators. Pearson correlation coefficients between the CDI models and (a) Vegetation Health Index (VHI), (b) Vapour Pressure Deficit (VPD), and (c) Land Surface Temperature (LST) at seasonal (MAM, JJAS, ON, and DJF) and accumulation (1-, 3-, 6-, and 12-month) scales. The *** and ** denote p-values < 0.001 and < 0.01, respectively. Grid-wise distribution of correlation performance classes for (d) VHI, (e) VPD, and (f) LST across the study area. For VHI, the correlation classes were defined as Poor ($r < 0.5$), Moderate ($0.5 \leq r < 0.6$), Good ($0.6 \leq r < 0.7$), and Excellent ($r \geq 0.7$). For VPD and LST, the classes were defined as Poor ($r > -0.5$), Moderate ($-0.6 < r \leq -0.5$), Good ($-0.7 < r \leq -0.6$), and Excellent ($r \leq -0.7$)

5.3 Exposure Index (EI)

EI was represented by the frequency and intensity of extreme and severe drought derived from the CDI at 1-, 3-, 6-, and 12-month accumulation periods, together with LST and VPD (Fig. 9). The spatial distribution of drought frequency revealed considerable heterogeneity across the MRB. Extreme drought events were more frequent over the western, south-western, and central parts of the basin, whereas the eastern coastal region generally exhibited lower drought occurrence (Fig. 9a-d). A comparable spatial pattern was observed for severe drought frequency (Fig. 9e-h), although the affected area was more extensive owing to the less stringent drought threshold. As the accumulation period increased from 1 to

12 months, drought frequency became spatially more coherent, indicating the persistence of prolonged hydrological deficits and the cumulative influence of antecedent moisture conditions. The spatial pattern of drought intensity broadly corresponded with that of drought frequency, with the western and south-western regions experiencing the most severe drought conditions (Fig. 9i-p). Greater drought intensity at longer accumulation periods indicates the persistence of moisture deficits and reflects the increasing dominance of long-term hydrological processes over short-term climatic variability. In contrast, the eastern coastal belt generally exhibited lower drought intensity, suggesting the moderating influence of higher rainfall and proximity to the BOB. The climatic exposure variables further supported these observations. LST was highest across the western and central basin, while comparatively lower temperatures prevailed over the eastern coastal region (Fig. 9q). Elevated surface temperatures enhance evapotranspiration and accelerate soil moisture depletion, thereby increasing drought exposure. Similarly, VPD exhibited a pronounced west-to-east gradient, with higher atmospheric moisture demand dominating the western and central regions and lower values occurring in the humid eastern basin (Fig. 9r). The spatial correspondence between elevated LST, high VPD, and recurrent drought hotspots indicates that these regions are subjected to greater atmospheric evaporative demand and consequently experience higher drought exposure. Collectively, the exposure indicators identify the western and central parts of the basin as the most drought-prone regions, while the eastern coastal zone remains comparatively less exposed because of its more humid climatic conditions.

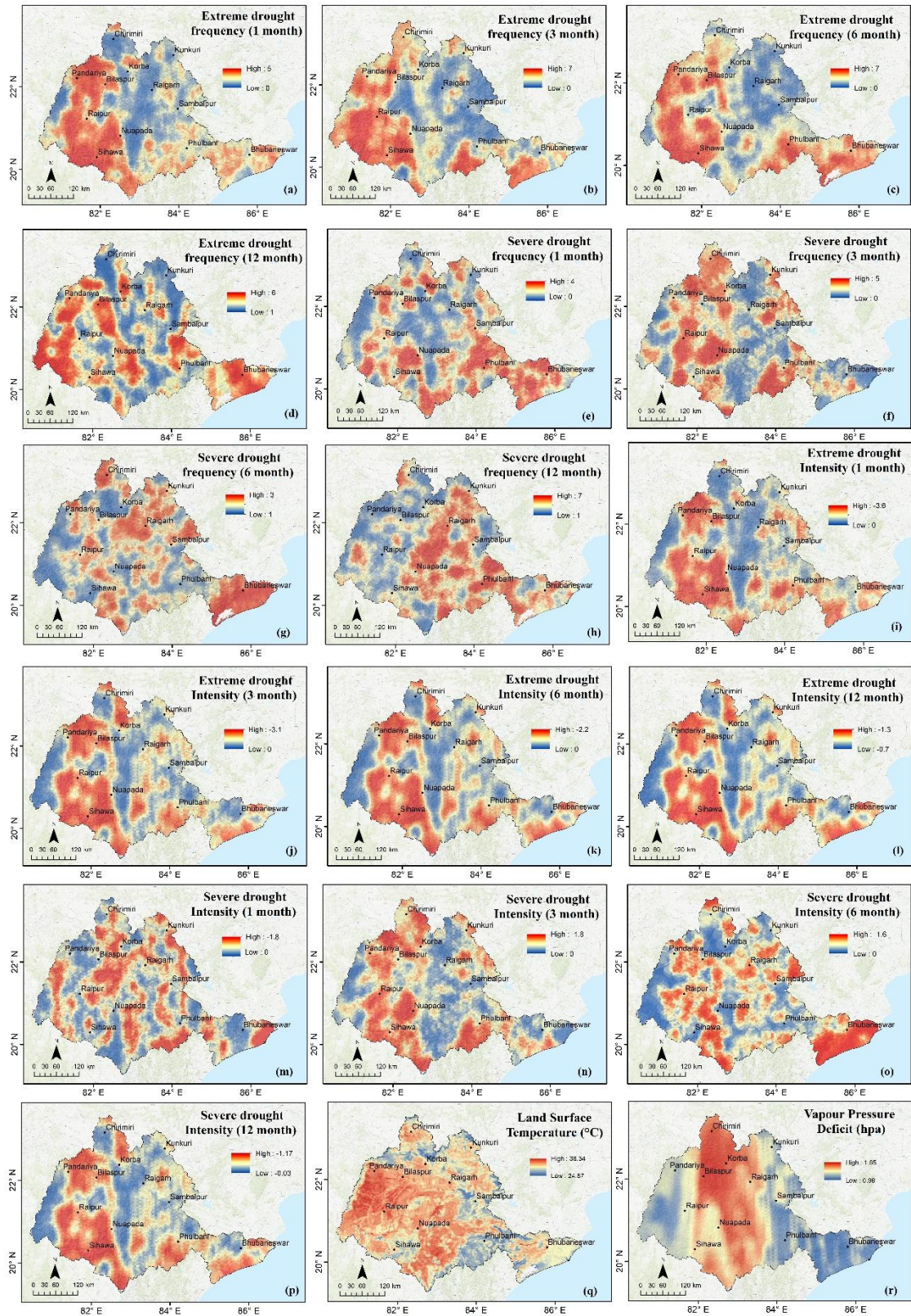


Fig. 9 Spatial distribution of the exposure indicators used for drought exposure assessment in the MRB: (a-d) extreme drought frequency at 1-, 3-, 6-, and 12-month timescales, respectively; (e-h) severe drought frequency at 1-, 3-, 6-, and 12-month timescales, respectively; (i-l) extreme drought intensity at 1-, 3-, 6-, and

12-month timescales, respectively; (m-p) severe drought intensity at 1-, 3-, 6-, and 12-month timescales, respectively; (q) LST; and (r) VP

Finally, the bagging approach was used to create the exposure map using these parameters. The areal distribution (Table 3) indicates that 28% of the basin falls under the high exposure category, followed by 25%, 21%, 14%, and 12% in the moderate, low, very low, and very high categories, respectively. Nearly 40% of the basin is classified within the high to very high exposure categories, signifying a considerable vulnerability to drought-inducing climatic conditions. The spatial distribution of the EI (Fig. 12a) reveals that the western region, particularly Chhattisgarh, exhibits the highest exposure, due to the combined occurrence of frequent and intense drought events together with elevated LST and VPD (Fig. 9a-r). These conditions reflect persistent atmospheric moisture stress and enhanced evaporative demand, thereby increasing the region's climatic exposure (Kumar et al., 2024; Mukherjee & Poonia, 2026; Soubhagya et al., 2026). Additionally, localized high-exposure areas are found in the southern interior and along the coastal regions of Odisha, indicating that both inland drought-prone areas and specific coastal zones are experiencing heightened climatic stress.

5.4 Sensitivity Index (SI)

Sensitivity indicators exhibit marked spatial heterogeneity across the basin. (Fig. 10a-r). In the northern and north-western parts, higher elevation and steeper slopes are found, representing various mountains at the basin border. Meanwhile, the eastern coastal part has low elevations (Fig. 10a-b). Similarly, drainage density is higher in the northern, western, and central regions, and in the eastern coastal plain the value reduces, reflecting the older stage of river (Fig. 10c). In contrast, TWI exhibits limited spatial variation, with slightly higher values confined to valley bottoms and low-lying

areas (Fig. 10d). Most of the basin lies in proximity to stream networks, although comparatively larger distances are observed locally in the eastern coastal region (Fig. 10e). GW levels showed a distinct west-to-east gradient, with comparatively shallow GW occurring in the eastern basin and deeper GW levels towards the western and central regions (Fig. 10f). Wind speed is highest over the south-western and western parts of the basin and gradually decreases towards the eastern coastal belt (Fig. 10g). Similarly, WDF is comparatively higher in the eastern coastal region and lower over the western interior (Fig. 10h), whereas the AI exhibits an opposite pattern, indicating relatively drier climatic conditions across the central and western basin (Fig. 10i). Land-use and vegetation indicators also demonstrate considerable spatial variability. Higher CI is concentrated over the eastern and south-eastern agricultural regions, while relatively lower values occur towards the north (Fig. 10k). Net sown area, barren land, fallow land, and forest cover display fragmented spatial distributions reflecting the heterogeneous land-use practices across the basin (Fig. 10l-o). NDVI indicates comparatively denser vegetation over the eastern and south-eastern parts of the basin, whereas lower vegetation greenness is evident in portions of the western and central regions (Fig. 10p). NDWI reveals that surface water bodies occupy only a limited proportion of the basin and are mainly associated with river channels and reservoirs (Fig. 10q). Soil texture (Fig. 10j) and lithology (Fig. 10r) exhibit substantial spatial variability, reflecting diverse geological formations and soil characteristics that influence infiltration, GW recharge, and soil water retention across the basin.

The individual SIs were subsequently integrated using the Bagging model to generate the composite sensitivity map (Fig. 12b). The SI exhibits an even greater concentration of highly vulnerable areas, with both high and very high sensitivity classes each accounting for 25% of the basin area, while moderate, low, and very low

1071 sensitivity cover 23%, 22%, and 6%, respectively. About half of the basin area is
1072 highly sensitive to drought, indicating that a large proportion of the watershed is highly
1073 responsive to drought conditions (Table 3). The spatial distribution of SI shows a
1074 distinct east-west pattern, whereby the eastern and southeastern parts, largely
1075 covering Odisha state, are more sensitive to droughts. Such a high level of sensitivity
1076 is due to the presence of AI, poor drainage density, fragmented forest cover, more
1077 agricultural lands, and unfavourable soil/lithology. Drought hot spots also exist around
1078 Nuapada and Sihawa in the southern part of the basin because of these physico-
1079 environmental features as well. In contrast, the western and northwestern parts, such
1080 as Raipur, Pandariya, Korba, and Chirimiri, showed very low sensitivity due to
1081 favorable hydro-environmental conditions, good vegetation cover, and higher water
1082 availability.

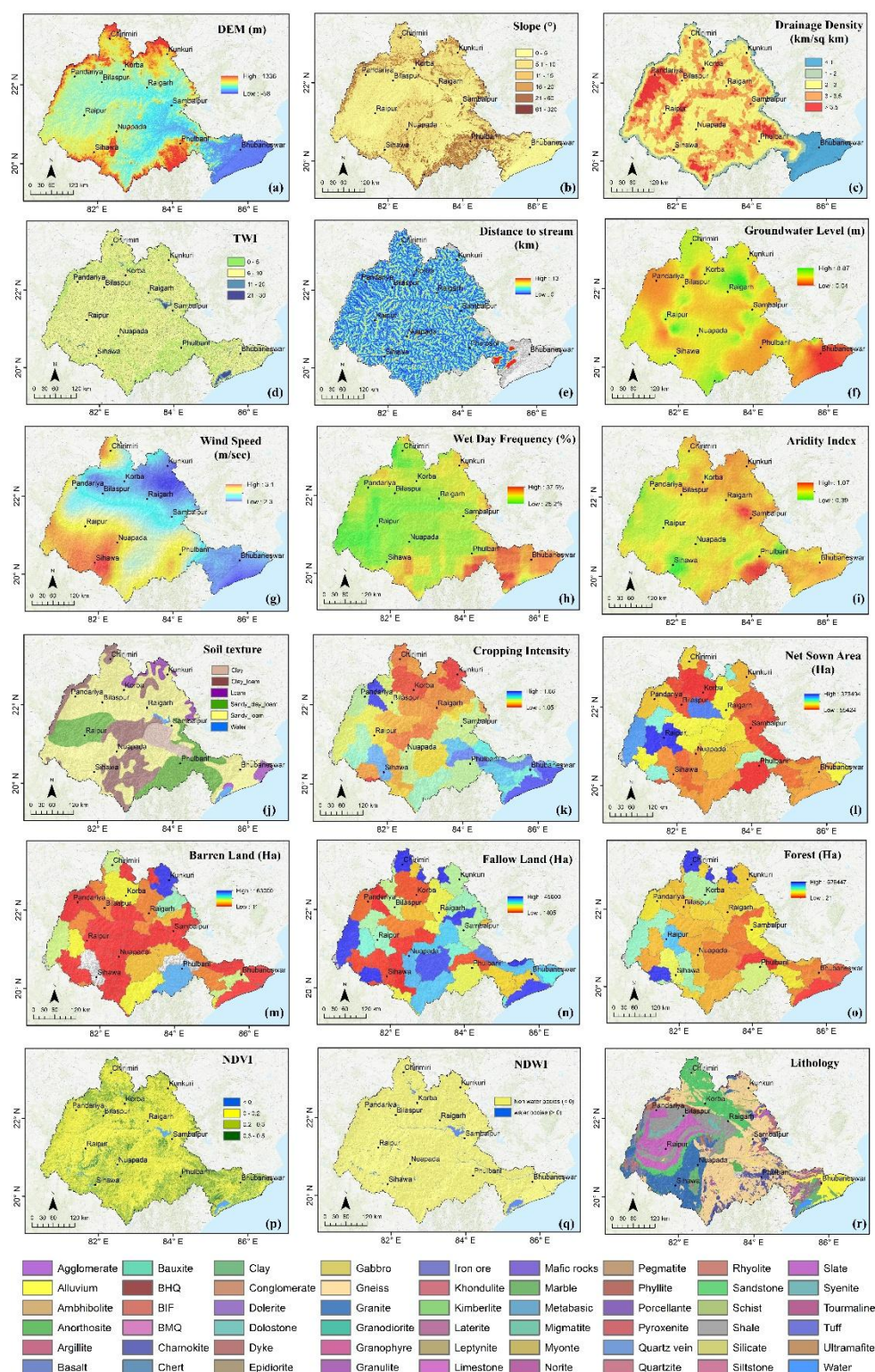


Fig. 10 Spatial distribution of the sensitivity indicators used for drought sensitivity assessment in the MRB: (a) Digital Elevation Model (DEM), (b) slope, (c) drainage density, (d) Topographic Wetness Index (TWI), (e) distance to stream (D2S), (f) groundwater level, (g) wind speed, (h) wet-day frequency, (i) aridity index, (j) soil texture, (k) cropping intensity, (l) net sown area, (m) barren land, (n) fallow land, (o) forest cover, (p) Normalized Difference Vegetation Index (NDVI), (q) Normalized Difference Water Index (NDWI), and (r) lithology

5.5 Adaptive Capacity Index (ACI)

ACI comprised population density, rural population, literacy rate, poverty, sanitation facilities, access to improved drinking water, prevalence of stunting among children under five years, anaemia among women aged 15-49 years, prevalence of underweight children under five years, access to clean cooking fuel, and electricity availability (Fig. 11a-k). These socio-economic indicators exhibit considerable spatial heterogeneity across the MRB, reflecting substantial differences in demographic characteristics, living conditions, public infrastructure, and human well-being. Population density is highest around the major urban centers of Raipur, Bilaspur, and Bhubaneswar, whereas comparatively lower densities are observed in the northern and interior districts (Fig. 11a). Similarly, the rural population is concentrated in the western and central agricultural districts, while relatively lower rural populations occur in the northern uplands (Fig. 11b). Literacy rates are generally higher in the eastern coastal districts surrounding Bhubaneswar and lower across portions of the southern interior (Fig. 11c). Poverty exhibits a contrasting distribution, with comparatively higher levels in the northern and southern boundary districts and lower levels across the western and eastern coastal regions (Fig. 11d). Infrastructure-related indicators also display marked spatial variation. Access to sanitation, improved drinking water, clean cooking fuel, and electricity is comparatively higher in the western urbanized districts and along the eastern coastal belt, whereas lower coverage is evident across the southern interior and parts of the central basin (Fig. 11e, f, j, and k). Health indicators, including childhood stunting, anaemia among women, and underweight children, show elevated prevalence in the eastern interior districts, indicating relatively poorer health conditions and lower adaptive capacity in these regions (Fig. 11g-i). Collectively, these indicators reveal pronounced spatial disparities in socio-economic development,

infrastructure availability, and public health, which strongly influence the basin's capacity to cope with drought.

The individual AC indicators were subsequently integrated using the Bagging model to generate the composite ACI (Fig. 12c). AC is dominated by the moderate class (34%), followed by high (26%), low (16%), very high (13%), and very low (11%) adaptive capacity (Table 3). Overall, 39% of the basin exhibits high to very high adaptive capacity, whereas 27% falls within the low and very low categories, highlighting substantial spatial disparities in drought coping capacity. The spatial distribution of ACI shows almost an inverse pattern to sensitivity. Areas with high to very high AC include the western and south-western basin, especially in and around Raipur and adjoining districts, as well as the eastern coastal belt in and around Bhubaneswar. This is due to relatively higher literacy rates, better access to sanitation, drinking water, power, clean fuel, relatively lower poverty levels, and better health indices, making them more resilient towards the impacts of drought. On the other hand, the southern and southeastern interior part, covering Phulbani, Nuapada, and adjoining areas, showed low to very low AC, because of relatively poor infrastructure, lack of basic amenities, high poverty, and poor health indices. The central basin, covering Bilaspur, Korba, Raigarh, and Sambalpur, is mainly comprised of moderate AC areas, showing an intermediate level of resilience between the highly adaptive western and coastal areas and the less adaptive southern interior.

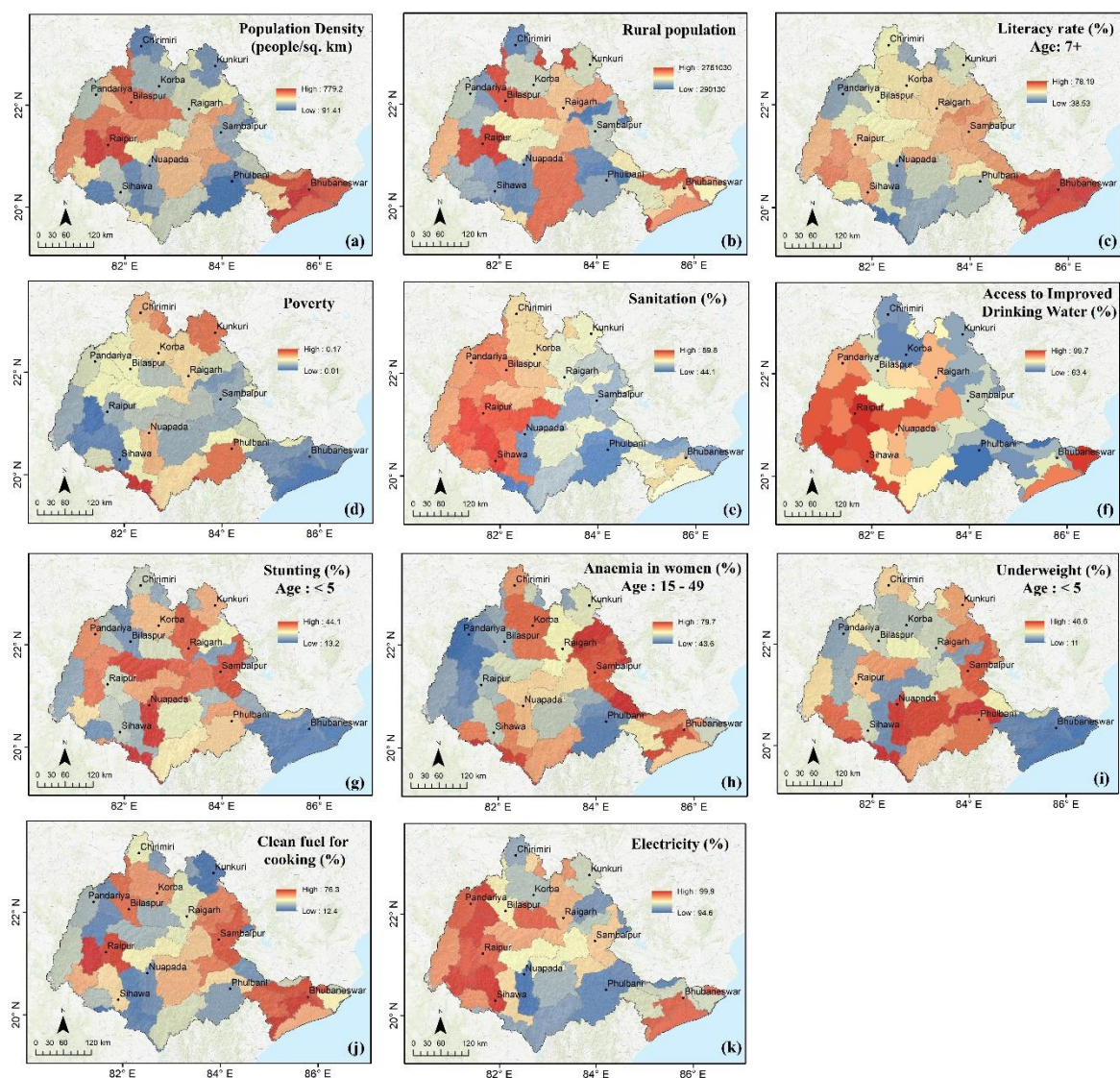


Fig. 11 Spatial distribution of the adaptive capacity indicators used for drought adaptive capacity assessment in the MRB. (a) population density, (b) rural population, (c) literacy rate (population aged ≥ 7 years), (d) poverty, (e) sanitation, (f) access to improved drinking water, (g) prevalence of stunting among children aged < 5 years, (h) prevalence of anaemia among women aged 15-49 years, (i) prevalence of underweight children aged < 5 years, (j) access to clean cooking fuel, and (k) electricity access

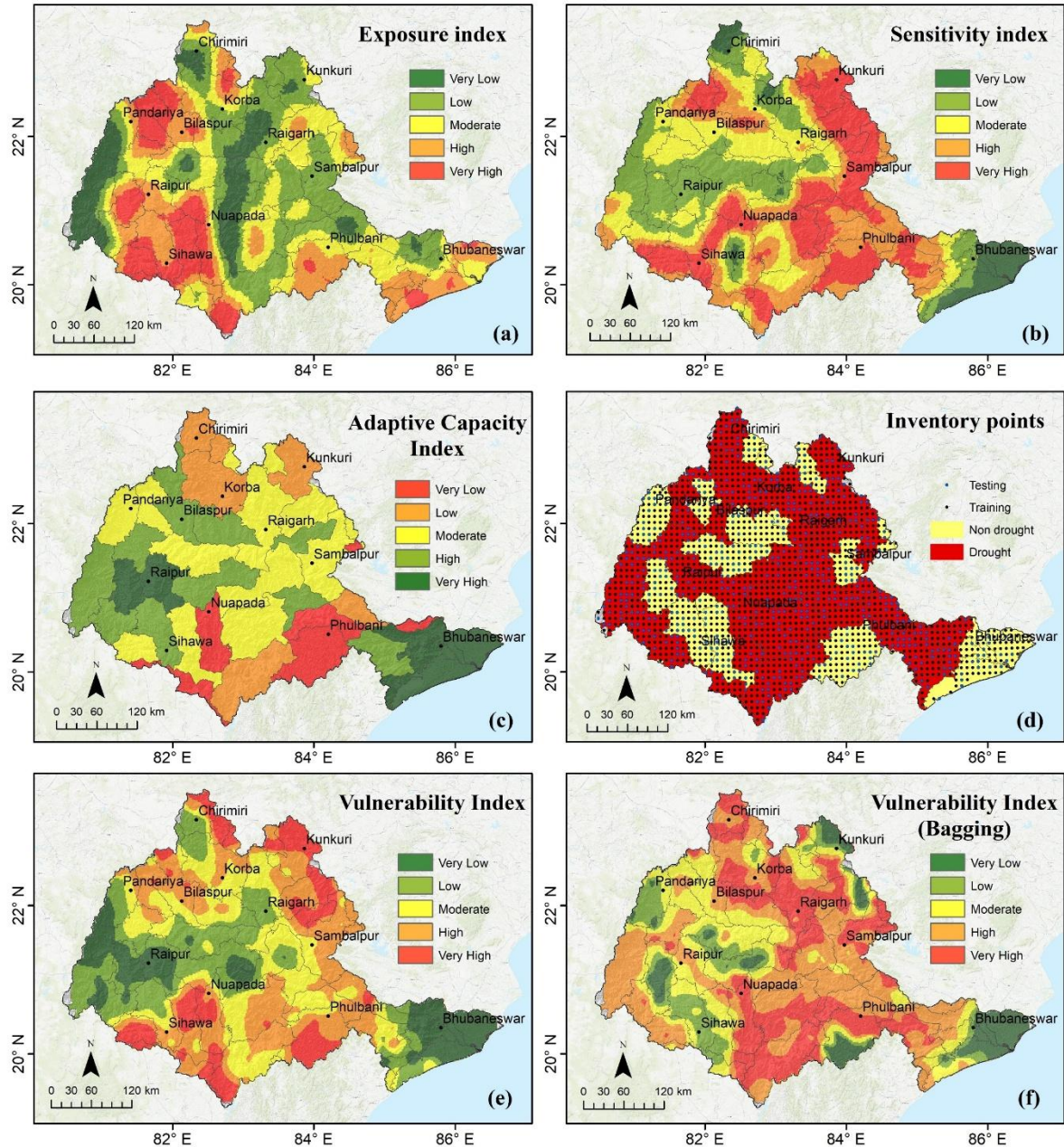


Fig. 12 Spatial distribution of the drought vulnerability (DV) components and vulnerability maps for the MRB. (a) EI, (b) SI, (c) ACI, (d) inventory points showing drought and non-drought samples used for model development and validation, (e) drought Vulnerability Index (DVI) derived using the conventional statistical approach, and (f) Bagging-based drought vulnerability map (DVM)

Table 3 Areal distribution (%) of the EI, SI, ACI, statistical VI, and Bagging-based DVM across the five vulnerability classes (very low, low, moderate, high, and very high) in the MRB.

Indices	Very low	Low	Moderate	High	Very high
Exposure	14%	21%	25%	28%	12%
Sensitivity	6%	22%	23%	25%	25%
AC	11%	16%	34%	26%	13%
Bagging	8%	10%	21%	36%	25%
Statistical	14%	21%	24%	28%	13%

5.6 Drought Vulnerability Assessment using Statistical and Bagging Approaches

The EI, SI, and ACI were integrated to generate DVM using both the conventional statistical approach and the Bagging ensemble model (Fig. 12e-f). The resulting vulnerability maps were classified into five categories, namely very low, low, moderate, high, and very high, using the Jenks natural breaks classification method in ArcGIS. In the statistical DVM, 28% of the basin is highly vulnerable, followed by moderate (24%), low (21%), very low (14%), and a very small portion (13%) belongs to the high vulnerability zone. Overall, 41% of the basin is categorized as high to very high vulnerability. The Bagging-based DVM exhibits a noticeable shift in the spatial distribution of vulnerability. 25% of the basin is very highly vulnerable, while 36% of the basin shows high vulnerability. Consequently, 21% of the area is moderate, 10% is low, and 8% belongs to very low vulnerability (Table 3).

Spatially, the statistical DVM (Fig. 12e) depicts isolated zones of high and very high vulnerability in the northern, southern, and eastern portions of the basin, while most of the central and western areas are in low- to moderate-vulnerability zones. Most of the coastal areas have low vulnerability due to relatively higher AC. Meanwhile, the bagging-based DVM (Fig. 12f) shows a much more consistent and contiguous pattern of vulnerability. Large zones of high to very high vulnerability are

present in the central, northern, southern, and southeastern parts of the basin, specifically around Korba, Raigarh, Sambalpur, Phulbani, Nuapada, and Sihawa. The western boundary near Raipur and the coastal zone showed relatively lower vulnerability due to the higher AC. Compared with the statistical approach, the Bagging model reduces isolated patches of vulnerability and produces smoother transition zones, thereby capturing the nonlinear interactions among exposure, sensitivity, and AC more effectively.

5.7 Performance Evaluation of the Statistical and Bagging-Based Vulnerability Model

The predictive performance of the conventional statistical VI and the Bagging-based DVI was evaluated using Accuracy, Precision, Recall, F1-score, RMSE, and the AUC-ROC for both the training and testing datasets. The ROC curves clearly demonstrate the superior discrimination capability of the Bagging model over the conventional statistical approach. During the training period, the statistical VI had an AUC value of 0.77, while the Bagging model gave an excellent AUC value of 0.97, representing an improvement of approximately 26.1% (Fig. 13a). For the testing data, the AUC value was 0.75 for the statistical method and 0.95 for the Bagging model (Fig. 13b). This represents an increase of 25.6%. In addition, the low gap between the training and testing data of the Bagging model (0.97 vs 0.95) indicates good generalization capability with minimal evidence of overfitting.

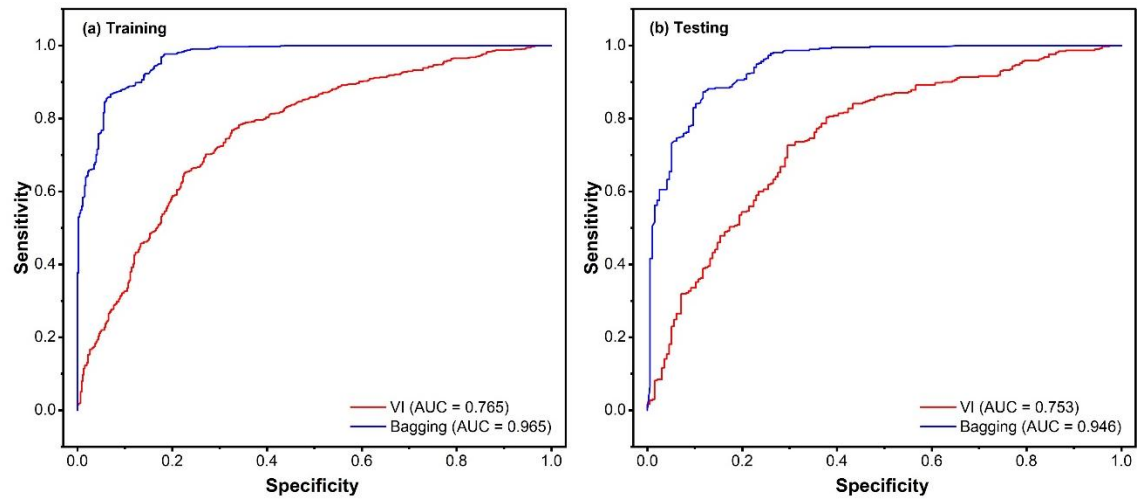


Fig. 13 ROC curves comparing the predictive performance of the conventional statistical DVI (red line) and the proposed Bagging-based DVM (blue line) for the (a) training and (b) testing datasets. The Bagging model consistently achieved higher AUC-ROC values than the statistical model

Statistical significance of the ROC analysis also showed the efficiency of the Bagging model (Table 4). Both the statistical and Bagging models showed very significant ($p < 0.001$) AUC values for the training and testing samples. For the training sample, the 95% confidence interval was extremely narrow (0.96-0.97), while for the testing sample it was 0.93-0.96. The high tightness of the confidence interval suggests the high accuracy and stability of the predictive model. However, the confidence interval of the statistical model was significantly lower, with 0.74-0.79 for training and 0.71-0.80 for testing. The substantially higher AUC values together with the narrower confidence intervals demonstrate that the Bagging model provides a more reliable classification of drought-vulnerable and non-vulnerable regions than the conventional statistical approach.

The superiority of the Bagging model was consistently reflected in the remaining evaluation metrics (Table 5). While training, the Bagging model recorded an accuracy of 92.1%, while the statistical model only recorded 65.3%. Similarly, there

was an improvement in the testing accuracy, which increased from 65.4% in the statistical model to 89.6% in the Bagging model. There was significant improvement in precision. The statistical model had a precision rate of 0.65 for both training and testing, while this value increased to 0.91 and 0.88 for training and testing, respectively, for the bagging model. The recall value of the Bagging model was 0.98 for both the training and testing datasets against 1 in the case of the statistical model. Also, the F1-score improved greatly from 0.79 in the statistical model for both training and testing, while the bagging model had 0.94 and 0.93 for training and testing, respectively. Consistent with these improvements, the RMSE decreased from 0.6 to 0.27 for the training dataset and from 0.60 to 0.3 for the testing dataset, demonstrating that the Bagging model produced considerably lower prediction errors than the conventional statistical vulnerability model.

Table 4 Statistical comparison of the AUC-ROC for the statistical VI and the Bagging-based DVM for the training and testing datasets.

Model	Area (AUC)	Std. Error	Asymptotic Sig. (p-value)	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Training	VI	0.77	0.0138	0.000	0.74 0.79
	Bagging	0.97	0.00	0.000	0.96 0.97
Testing	VI	0.75	0.02	0.000	0.71 0.80
	Bagging	0.95	0.01	0.000	0.93 0.96

Table 5 Performance evaluation of the statistical VI and the Bagging-based DVM

	Model	AUC-ROC	Accuracy	Precision	RMSE	Recall	F1
Training	VI	0.77	0.65	0.65	0.60	1.00	0.79
	Bagging	0.97	0.92	0.91	0.27	0.98	0.94
Testing	VI	0.75	0.65	0.65	0.60	1.00	0.79
	Bagging	0.95	0.90	0.88	0.29	0.98	0.93

6 Discussion

In the last few decades, climate change has caused a variety of natural calamities worldwide, and India is no exception [127]. Climate change has caused many natural hazards, such as droughts, floods, and landslides [128], [129], [130]. Still, drought is the most destructive natural calamity, which can cause significant economic loss and can have an adverse impact on human beings [131]. Unlike other hazards, drought is a cascading event that propagates through different components of the hydrological system, beginning with precipitation deficits and subsequently affecting soil moisture, surface water, GW, ecosystems, and socio-economic systems [132], [133], [134]. Consequently, DV depends not only on climatic factors but also on the dynamics between environmental features, availability of water resources, and the adaptability of human systems to these factors [135], [136], [137]. Developing reliable DV assessments has therefore become a fundamental requirement for sustainable water resource management, agricultural development, and climate change adaptation, especially in monsoon-affected river basins where climatic variability and anthropogenic stresses have increased water insecurity [131], [138], [139].

DV has been studied worldwide through hydro-climatological, environmental, and socio-economic indices by applying multi-criteria decision-making methods [140], fuzzy logic [141], entropy weighting [142], and ML approaches [143]. Even though these analyses have made considerable contributions to drought-risk assessment, they often represent drought conditions through precipitation deficits, vegetation stress, or SW availability, whereas GD remains comparatively underrepresented despite its largest accessible freshwater reserve supporting agriculture, domestic supply, and ecosystems during prolonged dry periods [144]. GD is fundamentally different from MD and AD due to its lagged response to precipitation shortages and

prolonged recovery phase [145], [146]. Aquifer systems often continue to experience water-level depletion even after rainfall conditions recover due to the recharge deficit, GW abstraction, and the properties of aquifer storage. Therefore, the duration of GD is typically greater than the duration of other drought types and represents one of the most critical components of long-term water security under increasing climatic variability. Although recent studies have attempted to integrate multiple drought types into a single composite indicator for drought monitoring [31], [147], [148], [149], their application in regional DV modelling remains limited. This study thus advances previous studies by introducing GD along with MD, HD, and AD through the Copula-CDI, which is subsequently incorporated into an ML framework for regional DVM.

Within the MRB, drought studies have mostly concentrated on MD and AD, rainfall variability, drought frequency, and GIS-based drought risk assessment. For example, Samantaray et al. [150] identified AD hotspots in the middle MRB using a Copula-based severity–duration–frequency framework integrated with ML. Likewise, recent basin-scale drought risk assessments have integrated meteorological, hydrological, agricultural, and socio-economic indicators using the Analytical Hierarchy Process (AHP) and geospatial techniques to delineate drought-prone regions and support integrated drought management strategies [151]. Furthermore, recent investigations have evaluated the future evolution of MD under changing climatic conditions using precipitation-based drought indices and climate projections, highlighting an increasing frequency and severity of drought events under future warming scenarios [152]. Despite these important contributions, such studies have ignored GD as one of the components of DV analysis. GW is an important source of irrigation and drinking water supply in the entire basin. The upper part of the basin, located in Chhattisgarh, is recognized as the Rice Bowl of India [153]. Here, GW is

used as an important supplementary irrigation source during rainfall-deficient years. Consequently, declining GW levels reduce irrigation reliability, increase pumping costs, and threaten agricultural productivity. In contrast, the downstream of MRB (coastal area of Odisha) faces significant vulnerability from the intrusion of saline water [154], [155], where GW depletion leads to a lowering of the fresh water hydraulic heads, resulting in salinization of GW and hence leading to reduced availability and quality of water for agricultural and domestic purposes. Therefore, incorporating GD into the CDI provides a more realistic representation of DV by accounting for both GW quantity and quality stresses, which are inadequately captured by precipitation- or SW-based assessments alone.

Among the evaluated composite drought models, the Copula-based CDI provided the most reliable representation of drought conditions because it explicitly models the nonlinear dependence among all drought types while preserving the marginal characteristics of each drought indicator. Drought propagation within the hydrological cycle is inherently nonlinear, particularly for GW systems where delayed recharge and aquifer storage introduce significant temporal lags between MD and GD [156], [157]. Conventional statistical aggregation methods generally assume simplified relationships among drought indicators and therefore cannot adequately capture these complex interactions [69], [71]. In contrast, the Copula framework effectively characterises compound drought behaviour, resulting in a more realistic representation of drought evolution and improved performance in subsequent vulnerability modelling [74], [76], [158].

The spatial vulnerability pattern further demonstrates the importance of integrating multiple drought dimensions. The Western region of the basin was marked by high climatic exposure due to constant rainfall variability, high evaporation rate, and

rain-fed agriculture. However, most parts of Odisha had higher drought sensitivity, which suggests that environmental characteristics, dependence on GW, land conditions, and socio-economic aspects increase the impact of drought even if the level of climatic exposure is not that high. The adaptive-capacity assessment showed that districts possessing better infrastructure, access to drinking water, sanitation, electricity, and improved socio-economic conditions exhibited greater resilience, whereas interior regions characterised by lower developmental indicators remained less capable of coping with prolonged drought events. Integrating these three dimensions identified extensive areas of high and very high drought vulnerability, demonstrating that DV arises from the interaction between climatic stress, environmental susceptibility, GW availability, and adaptive capacity rather than from precipitation deficits alone.

The Bagging-based DVM substantially outperformed the conventional statistical vulnerability approach, achieving higher Accuracy, Precision, Recall, F1-score, and AUC-ROC while simultaneously reducing prediction errors. Ensemble learning effectively captures nonlinear relationships [159] and complex interactions among EI, SI, and ACI that are difficult to represent through conventional weighted aggregation methods. Consequently, the Bagging model produced a more reliable spatial representation of DV and improved identification of highly vulnerable regions across the basin. These findings demonstrate that integrating a Copula-based CDI with ensemble ML techniques provides a robust framework for regional DVM.

Overall, this study provides a comprehensive framework that links GD monitoring, compound drought characterization, and ML-based DV assessment within a single modelling framework. Such an approach is particularly valuable for large monsoon-dominated river basins where GW increasingly functions as the primary

buffer against prolonged rainfall deficits. The resulting vulnerability maps can support GW recharge planning, conjunctive SW-GW management, irrigation scheduling, drought preparedness, and climate adaptation strategies. Because climate change is expected to increase drought frequency, GW stress, and irrigation demand across peninsular India, incorporating GD into regional vulnerability assessments offers a more realistic basis for long-term water-resource management than assessments based solely on MD or AD indicators.

Conclusion

This study presented a framework for assessing DV in the MRB through MD, AD, HD, and GD using the Copula-based CDI method, followed by coupling it with the Bagging machine-learning method. The Copula-based methodology has been used to capture the non-linear dependence among various types of droughts to create a more realistic compound drought. The vulnerability assessment revealed distinct spatial variability across the basin, where the western region has relatively high climatic exposure, whereas the eastern and coastal region of the basin has higher drought sensitivity due to environmental and socio-economic limitations. Furthermore, the adaptive capacity analysis shows that infrastructural improvement, availability of drinking water, sanitation, electricity, and other socio-economic aspects contribute greatly to increasing drought resilience. Consequently, the highest DV was observed in areas where high exposure and sensitivity coincided with limited adaptive capacity. The Bagging model significantly outperformed the conventional statistical approach, confirming the capability of ensemble learning to capture complex nonlinear interactions among DV components.

A major contribution of this study is the consideration of GD in the evaluation of DV, which has been neglected by previous studies. Considering the contrasting hydrogeological conditions of the basin, the presented approach allows for a more realistic estimation of long-term drought vulnerability and would be useful for GW recharge, conjunctive use of surface and GW, irrigation, and climate adaptation. The proposed methodology can be utilized in other GW-dependent river basins and can assist policymakers and water-resource managers in developing effective drought mitigation and sustainable GW management strategies.

Acknowledgements

L.N., P.K.S., and D.M.M. are grateful to the Department of Petroleum Engineering and Earth Sciences, UPES, for supporting this research. L.N. is thankful to R&D, UPES, for providing a PhD scholarship to support this work.

Author contributions

All authors conceptualized the work. L.N. collected, curated, and analysed the data and prepared the original draft of the manuscript. P.N.N. and D.M.M. supervised the work and contributed to reviewing and editing the manuscript. All authors have read and approved the final version of the manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability

The datasets used in this research were collected from various publicly available open-source platforms, which are listed in Table 1.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

- [1] S. Wang, B. Ancell, Z.-L. Yang, Q. Duan, and E. N. Anagnostou, 'Hydroclimatic extremes and impacts in a changing environment: Observations, mechanisms, and projections', *J. Hydrol. (Amst.)*, vol. 608, p. 127615, May 2022, doi: 10.1016/j.jhydrol.2022.127615.
- [2] Y. Li, J. Liu, X. Li, Z. Wang, and C. Liu, 'Spatiotemporal Characteristics and Hazard Assessment of Drought in Inner Mongolia Based on the MCI', *ISPRS Int. J. Geoinf.*, vol. 15, no. 3, p. 102, Mar. 2026, doi: 10.3390/ijgi15030102.
- [3] G. Geng, W. Zhang, L. Shan, P. Wang, Y. Liu, and M. Chen, 'Intensification and expansion of concurrent droughts in Northern China via a copula-based drought index', *Natural Hazards*, vol. 122, no. 15, p. 562, Aug. 2026, doi: 10.1007/s11069-026-08330-3.
- [4] E. Bevacqua, C.-F. Schleussner, and J. Zscheischler, 'A year above 1.5 °C signals that Earth is most probably within the 20-year period that will reach the Paris Agreement limit', *Nat. Clim. Chang.*, vol. 15, no. 3, pp. 262–265, Mar. 2025, doi: 10.1038/s41558-025-02246-9.
- [5] C. Miao, Q. Zhang, X. Li, J. Hu, M. Sadegh, and A. AghaKouchak, 'Global drought extremes in 2025', *Nat. Rev. Earth Environ.*, vol. 7, pp. 277–279, Apr. 2026, doi: <https://doi.org/10.1038/s43017-026-00785-z>.
- [6] D. S. Chuphal, A. P. Kushwaha, S. Aadhar, and V. Mishra, 'Drought Atlas of India, 1901–2020', *Sci. Data*, vol. 11, no. 1, p. 7, Jan. 2024, doi: 10.1038/s41597-023-02856-y.
- [7] M. H. Dar *et al.*, 'Drought Tolerant Rice for Ensuring Food Security in Eastern India', *Sustainability*, vol. 12, no. 6, p. 2214, Mar. 2020, doi: 10.3390/su12062214.
- [8] H. Varikoden, C. George, P. Ancy, and C. A. Babu, 'A comprehensive review on the recent changes in Indian summer monsoon rainfall and its relations with global climate drivers based on the CMIP6 simulations', *Atmos. Res.*, vol. 326, p. 108302, Nov. 2025, doi: 10.1016/j.atmosres.2025.108302.
- [9] A. Kulkarni, P. V. S. Raju, R. Ashrit, and A. Sagalgile, 'The Hydroclimate Paradox of the Indian Summer Monsoon Projections: Dual Amplification of Deficit and Excess Rainfall in <scp>CMIP6</scp> Models', *International Journal of Climatology*, vol. 46, no. 9, Jul. 2026, doi: 10.1002/joc.70401.
- [10] T. Sakthivel *et al.*, 'Sensitivity of South Asian Summer Monsoon rainfall to Bay of Bengal Sea Surface Temperature over the past 31 kiloyears', *Glob. Planet. Change*, vol. 255, p. 105118, Dec. 2025, doi: 10.1016/j.gloplacha.2025.105118.
- [11] D. J. Goswami, K. Ashok, and B. N. Goswami, 'Local ocean–atmosphere interaction in Indian summer monsoon multi-decadal variability', *Clim. Dyn.*, vol. 60, no. 5–6, pp. 1253–1270, Mar. 2023, doi: 10.1007/s00382-022-06377-z.

- 1399 [12] R. Bhatla, A. Maurya, A. Verma, R. K. Mall, and S. Bist, 'The dynamical influence of
1400 energy fluxes in modulating variability of the Indian summer monsoon', *J. Atmos. Sol.*
1401 *Terr. Phys.*, vol. 274, p. 106568, Sep. 2025, doi: 10.1016/j.jastp.2025.106568.
- 1402 [13] S. Aadhar and V. Mishra, 'Increased Drought Risk in South Asia under Warming
1403 Climate: Implications of Uncertainty in Potential Evapotranspiration Estimates', *J.*
1404 *Hydrometeorol.*, vol. 21, no. 12, pp. 2979–2996, Dec. 2020, doi: 10.1175/JHM-D-19-
1405 0224.1.
- 1406 [14] A. Kumar *et al.*, 'Cascading droughts: Exploring global propagation of meteorological
1407 to hydrological droughts (1971–2001)', *Science of The Total Environment*, vol. 979, p.
1408 179486, Jun. 2025, doi: 10.1016/j.scitotenv.2025.179486.
- 1409 [15] C. Sun, T. Wang, and W. Chen, 'Multi-dimensional response of soil moisture
1410 distribution to different droughts in complex geomorphic area: A case study of the
1411 middle reaches of the Yellow River', *J. Hydrol. Reg. Stud.*, vol. 65, Jun. 2026, doi:
1412 10.1016/J.EJRH.2026.103521.
- 1413 [16] L. Liu *et al.*, 'Global characteristics of drought propagation from surface water to
1414 groundwater', *J. Hydrol. (Amst.)*, vol. 665, p. 134734, Feb. 2026, doi:
1415 10.1016/J.JHYDROL.2025.134734.
- 1416 [17] D. J. Short Gianotti, M. J. Gannon, and D. Entekhabi, 'Meteorological to Agricultural
1417 Drought Transitions Compounded by Heat Waves in Historical and Future Climates',
1418 *Water Resour. Res.*, vol. 62, no. 3, Mar. 2026, doi: 10.1029/2024WR039170.
- 1419 [18] A. Saha, S. C. Pal, I. Chowdhuri, P. Roy, R. Chakraborty, and M. Shit, 'Vulnerability
1420 assessment of drought in India: Insights from meteorological, hydrological,
1421 agricultural and socio-economic perspectives', *Gondwana Research*, vol. 123, pp. 68–
1422 88, Nov. 2023, doi: 10.1016/J.GR.2022.11.006.
- 1423 [19] T. B. Mckee, N. J. Doesken, and J. Kleist, 'The relationship of drought frequency and
1424 duration to time scales', 1993.
- 1425 [20] S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno, 'A Multiscalar Drought
1426 Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration
1427 Index', *J. Clim.*, vol. 23, no. 7, pp. 1696–1718, Apr. 2010, doi: 10.1175/2009JCLI2909.1.
- 1428 [21] B. Narasimhan and R. Srinivasan, 'Development and evaluation of Soil Moisture Deficit
1429 Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought
1430 monitoring', *Agric. For. Meteorol.*, vol. 133, no. 1–4, pp. 69–88, Nov. 2005, doi:
1431 10.1016/j.agrformet.2005.07.012.
- 1432 [22] E. D. Hunt, K. G. Hubbard, D. A. Wilhite, T. J. Arkebauer, and A. L. Dutcher, 'The
1433 development and evaluation of a soil moisture index', *International Journal of*
1434 *Climatology*, vol. 29, no. 5, pp. 747–759, Apr. 2009, doi: 10.1002/joc.1749.

- 1435 [23] F. N. Kogan, 'Remote sensing of weather impacts on vegetation in non-homogeneous
1436 areas', *Int. J. Remote Sens.*, vol. 11, no. 8, pp. 1405–1419, Aug. 1990, doi:
1437 10.1080/01431169008955102.
- 1438 [24] R. Modarres, 'Streamflow drought time series forecasting', *Stochastic Environmental
1439 Research and Risk Assessment*, vol. 21, no. 3, pp. 223–233, Jan. 2007, doi:
1440 10.1007/s00477-006-0058-1.
- 1441 [25] B. F. Thomas, J. S. Famiglietti, F. W. Landerer, D. N. Wiese, N. P. Molotch, and D. F.
1442 Argus, 'GRACE Groundwater Drought Index: Evaluation of California Central Valley
1443 groundwater drought', *Remote Sens. Environ.*, vol. 198, pp. 384–392, Sep. 2017, doi:
1444 10.1016/j.rse.2017.06.026.
- 1445 [26] Z. Hao and A. AghaKouchak, 'Multivariate Standardized Drought Index: A parametric
1446 multi-index model', *Adv. Water Resour.*, vol. 57, pp. 12–18, Jul. 2013, doi:
1447 10.1016/j.advwatres.2013.03.009.
- 1448 [27] P. Kanthavel, C. K. Saxena, and R. K. Singh, 'Integrated Drought Index based on Vine
1449 Copula Modelling', *International Journal of Climatology*, vol. 42, no. 16, pp. 9510–
1450 9529, Dec. 2022, doi: 10.1002/joc.7840.
- 1451 [28] V. K. Prajapati, M. Khanna, M. Singh, R. Kaur, R. N. Sahoo, and D. K. Singh, 'PCA-based
1452 composite drought index for drought assessment in Marathwada region of
1453 Maharashtra state, India', *Theor. Appl. Climatol.*, vol. 149, no. 1–2, pp. 207–220, Jul.
1454 2022, doi: 10.1007/s00704-022-04044-1.
- 1455 [29] D. Saini and O. Singh, 'Comparison of meteorological, hydrological and agricultural
1456 droughts for developing a composite drought index over semi-arid Banas River Basin
1457 of India', *Stochastic Environmental Research and Risk Assessment*, vol. 38, no. 7, pp.
1458 2671–2691, Jul. 2024, doi: 10.1007/s00477-024-02704-4.
- 1459 [30] S. B. Tarate, N. R. Patel, and A. Danodia, 'A combined drought index for monitoring
1460 and assessment of drought severity over India by integrating CHIRPS, MODIS and
1461 GRACE data', *Environ. Dev. Sustain.*, vol. 28, no. 6, pp. 14731–14754, Oct. 2024, doi:
1462 10.1007/s10668-024-05555-9.
- 1463 [31] D. Shah and V. Mishra, 'Integrated Drought Index (IDI) for Drought Monitoring and
1464 Assessment in India', *Water Resour. Res.*, vol. 56, no. 2, Feb. 2020, doi:
1465 10.1029/2019WR026284.
- 1466 [32] L. Nanda, D. P. Panday, P. K. Srivastava, D. M. Mahapatra, U. Bhan, and M. Kumar,
1467 'Evolving Approaches in Groundwater-Surface Water Interaction Studies: From
1468 Numerical Models to Machine Learning', *Curr. Pollut. Rep.*, vol. 12, no. 1, p. 4, Jan.
1469 2026, doi: 10.1007/s40726-026-00395-9.
- 1470 [33] G. Rahman, M.-K. Jung, T.-W. Kim, and H.-H. Kwon, 'Drought impact, vulnerability, risk
1471 assessment, management and mitigation under climate change: A comprehensive
1472 review', *KSCE Journal of Civil Engineering*, vol. 29, no. 1, p. 100120, Jan. 2025, doi:
1473 10.1016/j.kscej.2024.100120.

- 1474 [34] G. Ganjir, M. J. Reddy, and S. Karmakar, 'Assessing agricultural drought hazard,
1475 vulnerability and risk over India', *International Journal of Disaster Risk Reduction*, vol.
1476 130, p. 105830, Nov. 2025, doi: 10.1016/j.ijdr.2025.105830.
- 1477 [35] IPCC, 'Climate change 2013: the physical science basis: Working Group I contribution
1478 to the Fifth assessment report of the Intergovernmental Panel on Climate Change',
1479 Cambridge, U.K., 2014.
- 1480 [36] S. Mahato, G. Mandal, B. Kundu, S. Kundu, P. K. Joshi, and P. Kumar, 'Comprehensive
1481 Drought Vulnerability Assessment in Northwestern Odisha: A Fuzzy Logic and
1482 Analytical Hierarchy Process Integration Approach', *Water (Basel)*, vol. 15, no. 18, p.
1483 3210, Sep. 2023, doi: 10.3390/w15183210.
- 1484 [37] K. K. Mohapatra *et al.*, 'Mapping agricultural vulnerability to climate-driven drought
1485 using the TOPSIS model', *Theor. Appl. Climatol.*, vol. 157, no. 3, p. 171, Mar. 2026, doi:
1486 10.1007/s00704-026-06070-9.
- 1487 [38] M. Masroor, M. H. Rahaman, and H. Sajjad, 'Assessing farmers' perception based
1488 composite drought vulnerability in Godavari Middle Sub-basin, India', *International
1489 Journal of Disaster Risk Reduction*, vol. 92, p. 103747, Jun. 2023, doi:
1490 10.1016/j.ijdr.2023.103747.
- 1491 [39] V. L. Sivakumar, R. Radha Krishnappa, and M. Nallanathel, 'Drought vulnerability
1492 assessment and mapping using Multi-Criteria decision making (MCDM) and
1493 application of Analytic Hierarchy process (AHP) for Namakkal District, Tamilnadu,
1494 India', *Mater. Today Proc.*, vol. 43, pp. 1592–1599, 2021, doi:
1495 10.1016/j.matpr.2020.09.657.
- 1496 [40] M. Upadhyay and M. A. Sherly, 'Multivariate framework for integrated drought
1497 vulnerability assessment – An application to India', *International Journal of Disaster
1498 Risk Reduction*, vol. 85, p. 103515, Feb. 2023, doi: 10.1016/j.ijdr.2022.103515.
- 1499 [41] D. Krishnan, P. Partheeban, M. Ramadoss, and D. Mogaka, 'Assessment of hydrological
1500 drought vulnerability using SPI and SWI indices in Tiruttani block Tamilnadu', *Sci. Rep.*,
1501 vol. 16, no. 1, p. 5795, Jan. 2026, doi: 10.1038/s41598-026-36100-5.
- 1502 [42] Shalwee, R. Dhupper, M. Kumari, A. K. Gupta, and D. Kumar, 'Implementing multi-
1503 criteria decision-making approaches to evaluate and map drought vulnerability with
1504 Geospatial Artificial Intelligence (Geo-AI) based investigations', *Discover Applied
1505 Sciences*, vol. 7, no. 10, p. 1185, Oct. 2025, doi: 10.1007/s42452-025-07189-6.
- 1506 [43] A. I. Abdul Rahim, P. Moorthy, B. Balu, and L. Chokkalingam, 'Future multi-dimensional
1507 drought analysis using machine learning and geospatial approaches: a case study of
1508 Tiruchirappalli District, Tamil Nadu, India', *Environmental Science and Pollution
1509 Research*, vol. 33, no. 20, pp. 9913–9942, Jun. 2026, doi: 10.1007/s11356-026-37940-
1510 4.
- 1511 [44] M. Tyagi, H. Singh, D. A. Thakur, and M. P. Mohanty, 'Compound risks of floods and
1512 droughts over multi-hazard catchments: Revealing association through hydrodynamic-

- 1513 cum-statistical modelling and novel bivariate risk classifier', *Science of The Total*
 1514 *Environment*, vol. 957, p. 177689, Dec. 2024, doi: 10.1016/j.scitotenv.2024.177689.
- 1515 [45] D. K. Panda, A. Kumar, S. Ghosh, and R. K. Mohanty, 'Streamflow trends in the
 1516 Mahanadi River basin (India): Linkages to tropical climate variability', *J. Hydrol.*
 1517 *(Amst)*., vol. 495, pp. 135–149, Jul. 2013, doi: 10.1016/j.jhydrol.2013.04.054.
- 1518 [46] P. G. Rao, 'Effect of Climate Change on Streamflows in the Mahanadi River Basin,
 1519 India', *Water Int.*, vol. 20, no. 4, pp. 205–212, Jan. 1995, doi:
 1520 10.1080/02508069508686477.
- 1521 [47] R. T. Sahu, M. K. Verma, and I. Ahmad, 'Impact of Long-Distance Interaction Indicator
 1522 (Monsoon Indices) on Spatio-Temporal Variability of Precipitation Over the Mahanadi
 1523 River Basin', *Water Resour. Res.*, vol. 59, no. 6, Jun. 2023, doi:
 1524 10.1029/2022WR033805.
- 1525 [48] D. Pandey, A. D. Tiwari, and V. Mishra, 'On the occurrence of the observed worst flood
 1526 in Mahanadi River basin under the warming climate', *Weather Clim. Extrem.*, vol. 38,
 1527 p. 100520, Dec. 2022, doi: 10.1016/j.wace.2022.100520.
- 1528 [49] A. Kumar, S. Paramanik, T. Baghel, M. Pramanik, and M. Hasan, 'Drought Risk in the
 1529 Mahanadi River Basin: A Multidimensional Approach for Integrated Urban-Rural
 1530 Drought Management Strategies', Jul. 08, 2024. doi: 10.32388/O5N5A4.
- 1531 [50] A. Danandeh Mehr, A. U. Sorman, E. Kahya, and M. Hesami Afshar, 'Climate change
 1532 impacts on meteorological drought using SPI and SPEI: case study of Ankara, Turkey',
 1533 *Hydrological Sciences Journal*, vol. 65, no. 2, pp. 254–268, Jan. 2020, doi:
 1534 10.1080/02626667.2019.1691218.
- 1535 [51] K. R. P. Babu and K. R. Bhimala, 'Recent trends in root-zone soil moisture over India
 1536 using the GLEAM data for the period 1980-2020', Mar. 27, 2022. doi:
 1537 10.5194/egusphere-egu22-1663.
- 1538 [52] X. Li, X. Zhang, X. Hong, Z. Ni, and C. Jiang, 'A novel approach for the estimation of
 1539 hydrological drought index incorporating precipitation-runoff relationship', *J. Hydrol.*
 1540 *(Amst)*., vol. 653, p. 132734, Jun. 2025, doi: 10.1016/j.jhydrol.2025.132734.
- 1541 [53] H. Save, S. Bettadpur, and B. D. Tapley, 'High-resolution CSR GRACE RL05 mascons',
 1542 *Journal of Geophysical Research: Solid Earth*. Accessed: May 21, 2026. [Online].
 1543 Available: https://www2.csr.utexas.edu/grace/RL06_mascons.html
- 1544 [54] Copernicus Climate Change Service (C3S), 'ERA5: Fifth generation of ECMWF
 1545 atmospheric reanalyses of the global climate', Copernicus Climate Change Service
 1546 Climate Data Store (CDS). Accessed: May 21, 2026. [Online]. Available:
 1547 [https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-](https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview)
 1548 [levels?tab=overview](https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview)

- 1549 [55] J. Muñoz Sabater, 'ERA5-Land monthly averaged data from 1981 to present.',
1550 *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, 2019, doi:
1551 10.24381/cds.68d2bb30.
- 1552 [56] G. Fischer, F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg,
1553 'Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008)', IIASA,
1554 Laxenburg, Austria and FAO, Rome, Italy. Accessed: May 21, 2026. [Online]. Available:
1555 [https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-](https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/)
1556 [world-soil-database-v12/en/](https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/)
- 1557 [57] C. F. Brown *et al.*, 'Dynamic World, Near real-time global 10 m land use land cover
1558 mapping', Scientific Data. Accessed: May 21, 2026. [Online]. Available:
1559 [https://developers.google.com/earth-](https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_DYNAMICWORLD_V1)
1560 [engine/datasets/catalog/GOOGLE_DYNAMICWORLD_V1](https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_DYNAMICWORLD_V1)
- 1561 [58] M. Mwinjuma, R. Wang, M. Mtupili, and M. Twaha, 'Comparisons of SPI and SPEI in
1562 capturing drought dynamics: A Global assessment across arid and humid regions',
1563 *Atmos. Res.*, vol. 329, p. 108475, Jan. 2026, doi: 10.1016/j.atmosres.2025.108475.
- 1564 [59] C. Herbert and P. Döll, 'Assessing Groundwater Drought Hazard in Groundwater
1565 Depletion Regions: Recommendations for Large-Scale Drought Early Warning
1566 Systems', *Water Resour. Res.*, vol. 61, no. 6, Jun. 2025, doi: 10.1029/2024WR038684.
- 1567 [60] S. Srivastava and O. Dikshit, 'Analysis of groundwater storage (GWS) dynamics and its
1568 temporal evolution for The Indo-Gangetic plain using GRACE data', *Remote Sens.*
1569 *Appl.*, vol. 25, p. 100685, Jan. 2022, doi: 10.1016/j.rsase.2021.100685.
- 1570 [61] S. Lee, D. N. Moriasi, A. Danandeh Mehr, and A. Mirchi, 'Sensitivity of Standardized
1571 Precipitation and Evapotranspiration Index (SPEI) to the choice of SPEI probability
1572 distribution and evapotranspiration method', *J. Hydrol. Reg. Stud.*, vol. 53, p. 101761,
1573 Jun. 2024, doi: 10.1016/j.ejrh.2024.101761.
- 1574 [62] M. Ghamghami and P. Irannejad, 'An analysis of droughts in Iran during 1988–2017',
1575 *SN Appl. Sci.*, vol. 1, no. 10, p. 1217, Oct. 2019, doi: 10.1007/s42452-019-1258-x.
- 1576 [63] S. Gai, X. Zhang, J. Xie, and J. Liu, 'Kernel density estimation method to determine
1577 coincident design day for air-conditioning system design in marine climate of China',
1578 *Build. Environ.*, vol. 269, p. 112487, Feb. 2025, doi: 10.1016/j.buildenv.2024.112487.
- 1579 [64] S. Cao, Y. Diao, J. Wang, Y. Liu, A. Raimondi, and J. Wang, 'KDE-Based Rainfall Event
1580 Separation and Characterization', *Water (Basel)*, vol. 15, no. 3, p. 580, Feb. 2023, doi:
1581 10.3390/w15030580.
- 1582 [65] A. Pavlides, V. D. Agou, and D. T. Hristopulos, 'Non-parametric kernel-based estimation
1583 and simulation of precipitation amount', *J. Hydrol. (Amst)*, vol. 612, p. 127988, Sep.
1584 2022, doi: 10.1016/j.jhydrol.2022.127988.

- 1585 [66] M. Karczewski and A. Michalski, 'A data-driven kernel estimator of the density
1586 function', *J. Stat. Comput. Simul.*, vol. 92, no. 17, pp. 3529–3541, Nov. 2022, doi:
1587 10.1080/00949655.2022.2072503.
- 1588 [67] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Routledge, 1998.
1589 doi: 10.1201/9781315140919.
- 1590 [68] Shalwee, R. Dhupper, M. Kumari, A. K. Gupta, and D. Kumar, 'Multi-dimensional
1591 assessment of drought vulnerability with composite drought index and principal
1592 component analysis for socioeconomic development in Semi-Arid regions', *Discover*
1593 *Applied Sciences*, vol. 8, no. 1, Jan. 2026, doi: 10.1007/s42452-025-07994-z.
- 1594 [69] B. Poudel, D. Dahal, S. Shrestha, R. Sewa, and A. Kalra, 'Developing a Composite
1595 Drought Indicator Using PCA Integration of CHIRPS Rainfall, Temperature, and
1596 Vegetation Health Products for Agricultural Drought Monitoring in New Mexico',
1597 *Atmosphere (Basel)*, vol. 16, no. 7, Jul. 2025, doi: 10.3390/atmos16070818.
- 1598 [70] G. Balaganesh *et al.*, 'Development of composite vulnerability index and district level
1599 mapping of climate change induced drought in Tamil Nadu, India', *Ecol. Indic.*, vol.
1600 113, Jun. 2020, doi: 10.1016/j.ecolind.2020.106197.
- 1601 [71] P. A. Pandya, N. K. Gontia, and H. V. Parmar, 'Development of PCA-based composite
1602 drought index for agricultural drought assessment using remote sensing', *Journal of*
1603 *Agrometeorology*, vol. 24, no. 4, pp. 384–392, Dec. 2022, doi:
1604 10.54386/jam.v24i4.1738.
- 1605 [72] D. S. . Wilks, *Statistical methods in the atmospheric sciences* . Elsevier, 2011.
- 1606 [73] I. Shim, H. Kim, B. Hong, J. An, and T. Hwang, 'Drought Vulnerability Assessment and
1607 Cluster Analysis of Island Areas Taking Korean Island Areas at Eup (Town) and Myeon
1608 (Subcounty) Levels as Study Targets', *Water (Basel)*, vol. 13, no. 24, p. 3657, Dec.
1609 2021, doi: 10.3390/w13243657.
- 1610 [74] X. Yu, X. Zeng, L. Brocca, D. Gui, D. Wang, and J. Wu, 'A Copula-Based Composite
1611 Drought Index for Enhanced Drought Monitoring and Analysis', *Journal of Geophysical*
1612 *Research: Atmospheres*, vol. 130, no. 8, Apr. 2025, doi: 10.1029/2024JD041867.
- 1613 [75] S. Moradian, S. Gharbia, A. AghaKouchak, A. Torabi Haghighi, and A. I. Olbert,
1614 'Integrated multi-index drought monitoring and projection under climate change',
1615 *Atmos. Res.*, vol. 316, Apr. 2025, doi: 10.1016/j.atmosres.2025.107946.
- 1616 [76] Q. Zhao, X. Zhang, C. Li, Y. Xu, and J. Fei, 'Compound ecological drought assessment of
1617 China using a Copula-based drought index', *Ecol. Indic.*, vol. 164, Jul. 2024, doi:
1618 10.1016/j.ecolind.2024.112141.
- 1619 [77] M. Sklar, 'Fonctions de répartition à N dimensions et leurs marges. ', *Annales de l'ISUP*,
1620 vol. VIII (3), pp. 229–231, 1959.
- 1621 [78] F. J. Massey, 'The Kolmogorov-Smirnov Test for Goodness of Fit', *J. Am. Stat. Assoc.*,
1622 vol. 46, no. 253, pp. 68–78, Mar. 1951, doi: 10.1080/01621459.1951.10500769.

- [79] S. Vulova *et al.*, 'Unraveling the response of forests to drought with explainable artificial intelligence (XAI)', *Ecol. Indic.*, vol. 172, p. 113308, Mar. 2025, doi: 10.1016/j.ecolind.2025.113308.
- [80] A. Hussain, R. Niaz, A. Y. Al-Rezami, A. Mohamed Omer, F. S. Al-Duais, and M. M. A. Almazah, 'Application of Random Forest for Identification of an Appropriate Model for Predicting Meteorological Drought', *Advances in Meteorology*, vol. 2025, no. 1, Jan. 2025, doi: 10.1155/adme/7674140.
- [81] A. Hussain, R. Niaz, M. M. A. Almazah, A. Y. Al-Rezami, H. Cheng, and A. Tariq, 'Utilizing Logistic Regression and Random Forest to Model Meteorological Drought Persistence Across Seasonal Transitions', *Earth Systems and Environment*, vol. 10, no. 2, pp. 1389–1412, Apr. 2026, doi: 10.1007/s41748-025-00682-3.
- [82] X. Xing *et al.*, 'A deep learning-based composite agricultural drought index for monitoring and impact assessment in Central Asia', *Agric. Water Manag.*, vol. 323, Feb. 2026, doi: 10.1016/j.agwat.2025.110043.
- [83] Q. Gao, J. Shang, S. Yuan, F. Li, and J. Wang, 'STNMAE: Identifying Spatial Domains from Spatial Transcriptomics Data with Neighbor-Aware Multi-view Masked Graph Autoencoder', *Interdiscip. Sci.*, Jun. 2026, doi: 10.1007/s12539-026-00832-9.
- [84] S. M. Omohundro, 'Five Balltree Construction Algorithms', *Data structures (Computer science)*, vol. 63, 1989.
- [85] T. N. Kipf and M. Welling, 'Semi-Supervised Classification with Graph Convolutional Networks', Feb. 2017.
- [86] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, 'Squeeze-and-Excitation Networks', *Computer Vision and Pattern Recognition*, 2019.
- [87] R. Guria, M. Mishra, C. A. G. Santos, R. M. da Silva, and P. Hatui, 'Spatiotemporal drought drivers in rainfed agroecosystems: A hybrid geospatial and machine learning approach for LULC-specific Vegetation Health Index (VHI) trends and yield loss assessment in the Eastern Ghats region (Odisha), India', *Remote Sens. Appl.*, vol. 42, p. 102072, Apr. 2026, doi: 10.1016/j.rsase.2026.102072.
- [88] J. Zeng *et al.*, 'Improving the drought monitoring capability of VHI at the global scale via ensemble indices for various vegetation types from 2001 to 2018', *Weather Clim. Extrem.*, vol. 35, p. 100412, Mar. 2022, doi: 10.1016/j.wace.2022.100412.
- [89] K. Didan, 'MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061', NASA Land Processes Distributed Active Archive Center.
- [90] Z. Wan, S. Hook, and G. Hulley, 'MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061', NASA Land Processes Distributed Active Archive Center.

- 1659 [91] F. N. Kogan, 'Application of vegetation index and brightness temperature for drought
1660 detection', *Advances in Space Research*, vol. 15, no. 11, pp. 91–100, 1995, doi:
1661 10.1016/0273-1177(95)00079-T.
- 1662 [92] S. Saha, P. Gogoi, A. Gayen, and G. C. Paul, 'Constructing the machine learning
1663 techniques based spatial drought vulnerability index in Karnataka state of India', *J.
1664 Clean. Prod.*, vol. 314, p. 128073, Sep. 2021, doi: 10.1016/j.jclepro.2021.128073.
- 1665 [93] S. Saha *et al.*, 'Spatial assessment of drought vulnerability using fuzzy-analytical
1666 hierarchical process: a case study at the Indian state of Odisha', *Geomatics, Natural
1667 Hazards and Risk*, vol. 12, no. 1, pp. 123–153, Jan. 2021, doi:
1668 10.1080/19475705.2020.1861114.
- 1669 [94] S. T. John and M. S. Philip, 'Development of an agricultural drought vulnerability index
1670 for risk assessment, insurance pricing, and disaster management through Google
1671 Earth engine', *Discover Geoscience*, vol. 3, no. 1, p. 107, Aug. 2025, doi:
1672 10.1007/s44288-025-00223-5.
- 1673 [95] M. Yu, J. Zhang, L. Wei, G. Wang, W. Dong, and X. Liu, 'Impact of soil textures on
1674 agricultural drought evolution and field capacity estimation in humid regions', *J.
1675 Hydrol. (Amst.)*, vol. 626, p. 130257, Nov. 2023, doi: 10.1016/J.JHYDROL.2023.130257.
- 1676 [96] R. P. Callahan *et al.*, 'Forest vulnerability to drought controlled by bedrock
1677 composition', *Nat. Geosci.*, vol. 15, no. 9, pp. 714–719, Sep. 2022, doi:
1678 10.1038/s41561-022-01012-2.
- 1679 [97] S. Kim, S.-K. Yoon, and N. Choi, 'Evaluating the Drainage Density Characteristics on
1680 Climate and Drainage Area Using LiDAR Data', *Applied Sciences*, vol. 13, no. 2, p. 700,
1681 Jan. 2023, doi: 10.3390/app13020700.
- 1682 [98] P. Moser, M. F. Simon, M. B. de Medeiros, A. B. Gontijo, and F. R. C. Costa, 'Interaction
1683 between extreme weather events and mega-dams increases tree mortality and alters
1684 functional status of Amazonian forests', *Journal of Applied Ecology*, vol. 56, no. 12, pp.
1685 2641–2651, Dec. 2019, doi: 10.1111/1365-2664.13498.
- 1686 [99] P. Roy, S. C. Pal, R. Chakraborty, I. Chowdhuri, A. Saha, and M. Shit, 'Climate change
1687 and groundwater overdraft impacts on agricultural drought in India: Vulnerability
1688 assessment, food security measures and policy recommendation', *Science of The Total
1689 Environment*, vol. 849, p. 157850, Nov. 2022, doi: 10.1016/J.SCITOTENV.2022.157850.
- 1690 [100] D. Marusig, F. Petruzzellis, M. Tomasella, R. Napolitano, A. Altobelli, and A. Nardini,
1691 'Correlation of Field-Measured and Remotely Sensed Plant Water Status as a Tool to
1692 Monitor the Risk of Drought-Induced Forest Decline', *Forests*, vol. 11, no. 1, p. 77, Jan.
1693 2020, doi: 10.3390/f11010077.
- 1694 [101] G. Wu *et al.*, 'Interannual variability of ecosystem iso/anisohydry is regulated by
1695 environmental dryness', *New Phytologist*, vol. 229, no. 5, pp. 2562–2575, Mar. 2021,
1696 doi: 10.1111/nph.17040.

- 1697 [102] V. Potopová, O. Lhotka, M. Možný, and M. Musiolková, 'Vulnerability of hop-yields due
1698 to compound drought and heat events over European key-hop regions', *International*
1699 *Journal of Climatology*, vol. 41, no. S1, Jan. 2021, doi: 10.1002/joc.6836.
- 1700 [103] R. Helcoski, L. T. Yonkos, A. Sanchez, and A. H. Baldwin, 'Wetland soil microplastics are
1701 negatively related to vegetation cover and stem density', *Environmental Pollution*, vol.
1702 256, p. 113391, Jan. 2020, doi: 10.1016/J.ENVPOL.2019.113391.
- 1703 [104] M. Kaushik, G. Ch. Kothyari, and A. K. Patidar, 'Morpho-tectonic assessment of a
1704 central Cratonic semi-arid climatic region of India: a geospatial technique', *Zeitschrift*
1705 *für Geomorphologie*, vol. 64, no. 3–4, pp. 181–219, Dec. 2025, doi: 10.1127/zfg/0830.
- 1706 [105] L. Rygel, D. O'sullivan, and B. Yarnal, 'A Method for Constructing a Social Vulnerability
1707 Index: An Application to Hurricane Storm Surges in a Developed Country', *Mitig.*
1708 *Adapt. Strateg. Glob. Chang.*, vol. 11, no. 3, pp. 741–764, May 2006, doi:
1709 10.1007/s11027-006-0265-6.
- 1710 [106] D. P. Johnson, A. Stanforth, V. Lulla, and G. Luber, 'Developing an applied extreme heat
1711 vulnerability index utilizing socioeconomic and environmental data', *Applied*
1712 *Geography*, vol. 35, no. 1–2, pp. 23–31, Nov. 2012, doi:
1713 10.1016/j.apgeog.2012.04.006.
- 1714 [107] T. Li *et al.*, 'Global disparities in rural and urban population exposure to compound
1715 drought and heatwave events', *NPJ Clim. Atmos. Sci.*, vol. 8, no. 1, p. 207, May 2025,
1716 doi: 10.1038/s41612-025-01025-9.
- 1717 [108] J. Das, S. Das, and N. V. Umamahesh, 'Population exposure to drought severities under
1718 shared socioeconomic pathways scenarios in India', *Science of The Total Environment*,
1719 vol. 867, p. 161566, Apr. 2023, doi: 10.1016/j.scitotenv.2023.161566.
- 1720 [109] V. Owino *et al.*, 'The impact of climate change on food systems, diet quality, nutrition,
1721 and health outcomes: A narrative review', *Frontiers in Climate*, vol. 4, Aug. 2022, doi:
1722 10.3389/fclim.2022.941842.
- 1723 [110] S. Kumar, R. Molitor, and S. Vollmer, 'Drought and Early Child Health in Rural India',
1724 *Popul. Dev. Rev.*, vol. 42, no. 1, pp. 53–68, Mar. 2016, doi: 10.1111/j.1728-
1725 4457.2016.00107.x.
- 1726 [111] S. Saha, B. Kundu, G. C. Paul, and B. Pradhan, 'Proposing an ensemble machine
1727 learning based drought vulnerability index using MSP, dagging, random sub-space and
1728 rotation forest models', *Stochastic Environmental Research and Risk Assessment*, vol.
1729 37, no. 7, pp. 2513–2540, Jul. 2023, doi: 10.1007/s00477-023-02403-6.
- 1730 [112] S. Saha, P. Gogoi, A. Gayen, and G. C. Paul, 'Constructing the machine learning
1731 techniques based spatial drought vulnerability index in Karnataka state of India', *J.*
1732 *Clean. Prod.*, vol. 314, p. 128073, Sep. 2021, doi: 10.1016/j.jclepro.2021.128073.
- 1733 [113] L. Breiman, 'Bagging predictors', *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996,
1734 doi: 10.1007/BF00058655.

- 1735 [114] K. Halder *et al.*, 'Application of bagging and boosting ensemble machine learning
1736 techniques for groundwater potential mapping in a drought-prone agriculture region
1737 of eastern India', *Environ. Sci. Eur.*, vol. 36, no. 1, p. 155, Sep. 2024, doi:
1738 10.1186/s12302-024-00981-y.
- 1739 [115] P. G. Nair *et al.*, 'GIS-based flood vulnerability mapping in a tropical river basin using
1740 analytical hierarchy process (AHP) and machine learning approach', *Geocarto Int.*, vol.
1741 40, no. 1, Dec. 2025, doi: 10.1080/10106049.2025.2551261.
- 1742 [116] H. Shahabi *et al.*, 'Flood Detection and Susceptibility Mapping Using Sentinel-1
1743 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of
1744 Bagging Ensemble Based on K-Nearest Neighbor Classifier', *Remote Sens. (Basel)*, vol.
1745 12, no. 2, p. 266, Jan. 2020, doi: 10.3390/rs12020266.
- 1746 [117] D. M. W. Powers, 'Evaluation: From precision, recall and F-measure to ROC,
1747 informedness, markedness & correlation', *Journal of Machine Learning Technologies*,
1748 vol. 2, pp. 37–63, Jan. 2011, doi: 10.9735/2229-3981.
- 1749 [118] C. Willmott and K. Matsuura, 'Advantages of the mean absolute error (MAE) over the
1750 root mean square error (RMSE) in assessing average model performance', *Clim. Res.*,
1751 vol. 30, pp. 79–82, 2005, doi: 10.3354/cr030079.
- 1752 [119] T. Fawcett, 'An introduction to ROC analysis', *Pattern Recognit. Lett.*, vol. 27, no. 8, pp.
1753 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.
- 1754 [120] E. Bergh, 'Interpretable Dimensionality Reduction Using Weighted Linear
1755 Transformation', *Advances in Artificial Intelligence and Machine Learning*, vol. 05, no.
1756 01, pp. 3465–3475, 2025, doi: 10.54364/AAIML.2025.51198.
- 1757 [121] R. Guria, M. Mishra, C. A. G. Santos, R. M. da Silva, and P. Hatui, 'Spatiotemporal
1758 drought drivers in rainfed agroecosystems: A hybrid geospatial and machine learning
1759 approach for LULC-specific Vegetation Health Index (VHI) trends and yield loss
1760 assessment in the Eastern Ghats region (Odisha), India', *Remote Sens. Appl.*, vol. 42, p.
1761 102072, Apr. 2026, doi: 10.1016/j.rsase.2026.102072.
- 1762 [122] N. Dubey and S. Ghosh, 'The relative role of soil moisture and vapor pressure deficit in
1763 affecting the Indian vegetation productivity', *Environmental Research Letters*, vol. 18,
1764 no. 6, p. 064012, Jun. 2023, doi: 10.1088/1748-9326/acd2ef.
- 1765 [123] A. Sampelli, A. Khandare, R. Muvva Venkata, S. Karmakar, and S. Ghosh, 'Atmospheric
1766 aridity perturbs critical soil moisture thresholds of plant water stress over Indian
1767 biomes', *Adv. Water Resour.*, vol. 206, p. 105169, Dec. 2025, doi:
1768 10.1016/j.advwatres.2025.105169.
- 1769 [124] A. Mukherjee and V. Poonia, 'Univariate and Multivariate Analysis of Major Drought
1770 Types Over Mahanadi and Godavari River Basins in Central India', in *Climate Extremes
1771 and Emerging Solutions: Data-Driven Insights and Geospatial Techniques*, Cham:
1772 Springer Nature Switzerland, 2026, pp. 59–75. doi: 10.1007/978-3-032-14457-7_4.

- 1773 [125] L. R. Soubhagya, P. S. Ambika, C. P. Jagadish, M. D. Dwarika, K. R. Sanjay, and P. J.
 1774 Prachi, 'Assessment of spatio-temporal trend in hydro-climatic variables and their
 1775 agricultural implications in the Mahanadi River Basin using innovative trend analysis
 1776 technique', *Plant Science Today*, vol. 13, no. sp1, Jan. 2026, doi: 10.14719/pst.10914.
- 1777 [126] A. Kumar, S. Paramanik, T. Baghel, M. Pramanik, and M. Hasan, 'Drought Risk in the
 1778 Mahanadi River Basin: A Multidimensional Approach for Integrated Urban-Rural
 1779 Drought Management Strategies', Jul. 2024, doi: 10.32388/O5N5A4.
- 1780 [127] A. Saha, S. C. Pal, I. Chowdhuri, P. Roy, R. Chakraborty, and M. Shit, 'Vulnerability
 1781 assessment of drought in India: Insights from meteorological, hydrological,
 1782 agricultural and socio-economic perspectives', *Gondwana Research*, vol. 123, pp. 68–
 1783 88, Nov. 2023, doi: 10.1016/j.gr.2022.11.006.
- 1784 [128] A. Joshi, G. C. Kothiyari, A. K. Patidar, and K. Gupta, 'Landslide susceptibility
 1785 assessment using AHP and RF techniques: Special reference to Kumaun lesser
 1786 Himalaya, Uttarakhand, India', *Physics and Chemistry of the Earth, Parts A/B/C*, vol.
 1787 142, p. 104191, Feb. 2026, doi: 10.1016/j.pce.2025.104191.
- 1788 [129] A. Gogoi, G. C. Kothiyari, and A. K. Patidar, 'Ground subsidence monitoring in India
 1789 using InSAR: A review of current status and future prospects', *Geosystems and
 1790 Geoenvironment*, p. 100436, Jul. 2025, doi: 10.1016/j.geogeo.2025.100436.
- 1791 [130] J. Changmai, A. K. Patidar, G. Ch. Kothiyari, and C. M. Bhatt, 'Geohazard dynamics in
 1792 Kumaon and Garhwal regions of Uttarakhand Himalayas: a critical review of trends
 1793 and socio-environmental impacts', *Spatial Information Research*, vol. 34, no. 1, p. 11,
 1794 Feb. 2026, doi: 10.1007/s41324-025-00662-9.
- 1795 [131] A. Tahasin, M. Haydar, Md. S. Hossen, and H. Sadia, 'Drought vulnerability assessment
 1796 and its impact on crop production and livelihood of people: An empirical analysis of
 1797 Barind Tract', *Heliyon*, vol. 10, no. 20, p. e39067, Oct. 2024, doi:
 1798 10.1016/j.heliyon.2024.e39067.
- 1799 [132] A. Kumar *et al.*, 'Cascading droughts: Exploring global propagation of meteorological
 1800 to hydrological droughts (1971–2001)', *Science of The Total Environment*, vol. 979, p.
 1801 179486, Jun. 2025, doi: 10.1016/j.scitotenv.2025.179486.
- 1802 [133] P. S. Bhunia, Abhishek, S. Mukherjee, and K. C. Patra, 'Drought Characteristics and
 1803 Cascading Dynamics in South India: On the Key Influencing Factors', *Earth Systems and
 1804 Environment*, vol. 10, no. 3, pp. 3719–3746, Jun. 2026, doi: 10.1007/s41748-025-
 1805 00845-2.
- 1806 [134] Z. Qu *et al.*, 'Drought propagation as a nonlinear amplifier of ecohydrological damage',
 1807 *Commun. Earth Environ.*, vol. 7, no. 1, p. 319, Feb. 2026, doi: 10.1038/s43247-026-
 1808 03330-4.
- 1809 [135] V. L. Sivakumar *et al.*, 'An integration of geospatial technology and standard
 1810 precipitation index (SPI) for drought vulnerability assessment for a part of Namakkal

- 1811 district, South India', in *Materials Today: Proceedings*, Elsevier Ltd, Jan. 2020, pp.
1812 1206–1211. doi: 10.1016/j.matpr.2020.08.157.
- 1813 [136] V. Sahana, A. Mondal, and P. Sreekumar, 'Drought vulnerability and risk assessment in
1814 India: Sensitivity analysis and comparison of aggregation techniques', *J. Environ.*
1815 *Manage.*, vol. 299, Dec. 2021, doi: 10.1016/j.jenvman.2021.113689.
- 1816 [137] A. Saha, S. C. Pal, I. Chowdhuri, P. Roy, R. Chakraborty, and M. Shit, 'Vulnerability
1817 assessment of drought in India: Insights from meteorological, hydrological,
1818 agricultural and socio-economic perspectives', *Gondwana Research*, vol. 123, pp. 68–
1819 88, Nov. 2023, doi: 10.1016/j.gr.2022.11.006.
- 1820 [138] T. Wang and F. Sun, 'Integrated drought vulnerability and risk assessment for future
1821 scenarios: An indicator based analysis', *Science of the Total Environment*, vol. 900,
1822 Nov. 2023, doi: 10.1016/j.scitotenv.2023.165591.
- 1823 [139] R. S. Alharbi *et al.*, 'Assessment of Drought vulnerability through an integrated
1824 approach using AHP and Geoinformatics in the Kangsabati River Basin', *J. King Saud*
1825 *Univ. Sci.*, vol. 34, no. 8, Nov. 2022, doi: 10.1016/j.jksus.2022.102332.
- 1826 [140] A. Habibi *et al.*, 'Precision breeding for drought tolerance using a Copeland method:
1827 Integrating receiver operating characteristic analysis and multi-criteria decision
1828 making to identify superior Persian walnut genotypes', *Sci. Hortic.*, vol. 360, p.
1829 114741, Mar. 2026, doi: 10.1016/j.scienta.2026.114741.
- 1830 [141] U. Nigam *et al.*, 'Drought vulnerability assessment using morphometric features and
1831 extreme precipitation indicators to prioritize sub-basins: AI-based Fuzzy Logic
1832 approach', *Results in Engineering*, vol. 29, p. 109317, Mar. 2026, doi:
1833 10.1016/j.rineng.2026.109317.
- 1834 [142] E. N. Bañares, M. S. Mehboob, and M. R. Panda, 'Physical and socio-economic drought
1835 vulnerability in a high-risk local government unit of the Bicol River Basin, Philippines',
1836 *Environmental and Sustainability Indicators*, vol. 30, p. 101177, Jun. 2026, doi:
1837 10.1016/j.indic.2026.101177.
- 1838 [143] W. Sun, L.-C. Chang, J. J. Lin, and F.-J. Chang, 'Integrating deep learning and
1839 groundwater dynamics for drought vulnerability assessment under climate scenarios',
1840 *Groundw. Sustain. Dev.*, vol. 33, p. 101591, May 2026, doi:
1841 10.1016/j.gsd.2026.101591.
- 1842 [144] T. M. V. Maya *et al.*, 'Limnological dynamics in groundwater resources of Bhadra River
1843 Basin, Central Western Ghats, India: Geospatial approach', *Cleaner Water*, vol. 6, p.
1844 100237, Jun. 2026, doi: 10.1016/j.clwat.2026.100237.
- 1845 [145] H. Lai *et al.*, 'Spatiotemporal patterns, abrupt change points and vegetation response
1846 of groundwater drought in China', *Ecol. Indic.*, vol. 189, p. 115145, Aug. 2026, doi:
1847 10.1016/j.ecolind.2026.115145.

- 1848 [146] M. M. Haile, F. Li, Z. Ai, G. Chen, and P. Leng, 'Spatiotemporal characteristics and
1849 driving factors of groundwater drought in Nile River Basin (NRB)', *Catena (Amst.)*, vol.
1850 257, p. 109194, Sep. 2025, doi: 10.1016/j.catena.2025.109194.
- 1851 [147] M. Yuan *et al.*, 'A new multivariate composite drought index considering the lag time
1852 and the cumulative effects of drought', *J. Hydrol. (Amst.)*, vol. 653, p. 132757, Jun.
1853 2025, doi: 10.1016/j.jhydrol.2025.132757.
- 1854 [148] S. Sharafi and M. M. Ghaleni, 'Enhancing drought monitoring and prediction in diverse
1855 climates by using composite drought indices', *Stochastic Environmental Research and
1856 Risk Assessment*, vol. 39, no. 10, pp. 4295–4315, Oct. 2025, doi: 10.1007/s00477-023-
1857 02597-9.
- 1858 [149] A. Srivastava, A. Pandey, and R. Maity, 'Comparative multi-index analysis of existing
1859 drought typology and environmental droughts in a climate-stressed region', *Sci. Rep.*,
1860 vol. 15, no. 1, p. 34349, Oct. 2025, doi: 10.1038/s41598-025-16859-9.
- 1861 [150] A. K. Samantaray, G. Singh, M. Ramadas, and R. K. Panda, 'Drought hotspot analysis
1862 and risk assessment using probabilistic drought monitoring and severity–duration–
1863 frequency analysis', *Hydrol. Process.*, vol. 33, no. 3, pp. 432–449, Jan. 2019, doi:
1864 10.1002/hyp.13337.
- 1865 [151] A. Kumar, S. Paramanik, T. Baghel, M. Pramanik, and Md. M. Hasan, 'Drought risk in
1866 the Mahanadi River basin: a multi-dimensional approach for integrated urban-rural
1867 drought management strategies', *International Journal of River Basin Management*,
1868 pp. 1–19, Jul. 2024, doi: 10.1080/15715124.2024.2368176.
- 1869 [152] A. K. Dandapat, P. K. Panda, S. Sankalp, and M. Jothiman, 'Evaluating drought risks and
1870 metrological hazards in the Mahanadi basin: impact of precipitation deficits and
1871 climate change on the downstream regions of Hirakud dam', *Discover Applied
1872 Sciences*, vol. 7, no. 7, p. 717, Jul. 2025, doi: 10.1007/s42452-025-07310-9.
- 1873 [153] Renuka and S. Das, 'Comparative Spatiotemporal Study of Rice and Wheat Yield in
1874 Aarang, Rasni and Lakholi Blocks of Raipur District, Chhattisgarh, India', *Asian
1875 Research Journal of Agriculture*, vol. 19, no. 2, pp. 269–278, Apr. 2026, doi:
1876 10.9734/arja/2026/v19i2849.
- 1877 [154] S. P. Parija *et al.*, 'Groundwater quality and hydrogeochemical analysis for assessing
1878 saltwater intrusion in coastal aquifers of Eastern Odisha, India', *J. Coast. Conserv.*, vol.
1879 29, no. 6, p. 69, Dec. 2025, doi: 10.1007/s11852-025-01145-7.
- 1880 [155] S. Srujanika, R. Masare, P. Waindeshkar, A. V. Shirke, P. P. Das, and B. Umrikar, 'Spatial
1881 Mapping of Saltwater Intrusion in Coastal Odisha, India using Hydrochemical
1882 Signatures and Machine Learning Tools', *ISPRS Annals of the Photogrammetry, Remote
1883 Sensing and Spatial Information Sciences*, vol. X-5/W2-2025, pp. 671–678, Dec. 2025,
1884 doi: 10.5194/isprs-annals-X-5-W2-2025-671-2025.

1885 [156] P. C. Muraharirao and N. Chinthu, 'Regional scale analysis of the internal propagation
1886 of groundwater droughts in Andhra Pradesh, India', *Theor. Appl. Climatol.*, vol. 157,
1887 no. 1, p. 33, Jan. 2026, doi: 10.1007/s00704-025-05957-3.

1888 [157] S. Salwey, S. Hauswirth, D. Ruijsch, B. van Jaarsveld, J. van Mourik, and N. Wanders,
1889 'Global-scale drought propagation and the drivers and patterns of multi-year
1890 groundwater drought', Apr. 29, 2026. doi: 10.5194/egusphere-2026-2335.

1891 [158] H. Palagiri, N. Sudardeva, and M. Pal, 'Application of ESACCI SM product-assimilated to
1892 a statistical model to assess the drought propagation for different Agro-Climatic zones
1893 of India using copula', *International Journal of Applied Earth Observation and*
1894 *Geoinformation*, vol. 127, p. 103701, Mar. 2024, doi: 10.1016/J.JAG.2024.103701.

1895 [159] X. Deng, Y. Zhang, S. Mao, and P. Liu, 'Non-Linearity Incremental Ensemble Learning
1896 Based Depth Explorations of Latent Individual Contributions', *Electron. Lett.*, vol. 61,
1897 no. 1, Jan. 2025, doi: 10.1049/ell2.70178.

1898

1899

1900