

Research article

A novel framework for future drought characterization under ranked-based subset selection and weighted aggregative multi-modal ensemble of global climate models



Muhammad Shakeel ^a, Hussnain Abbas ^a, Zulfiqar Ali ^a, Aqil Tariq ^b, Mansour Almazroui ^{c,d}, Shuraik Kader ^{e,*}

^a College of Statistical Sciences, University of the Punjab, Quaid-e-Azam Campus, Lahore, 54590, Punjab, Pakistan

^b Department of Wildlife, Fisheries and Aquaculture, College of Forest Resources, Mississippi State University, Mississippi State, MS, 39762-9690, USA

^c Center of Excellence for Climate Change Research/Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia

^d Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

^e School of Engineering and Built Environment, Griffith University, 170 Kessels Road, Nathan, Queensland, 4111, Australia

ARTICLE INFO

Keywords:

Global climate models (GCMs)
Drought projection
Gaussian-norm weighted drought index (GNWDI)
Mutual information (MI)
Shared socioeconomic pathways (SSPs)

ABSTRACT

Future drought characterization often relies on Multi-Modal Ensembles (MMEs) of Global Climate Models (GCMs), particularly from the Coupled Model Intercomparison Project Phase 6 (CMIP6). However, the reliability of projections is often hindered by insufficient ranking methodologies for GCMs and inadequate handling of outliers in regional aggregation. This study presents a novel framework to enhance the reliability of drought projections and standardization by introducing innovative ranking, aggregation, and projection methods. The framework is not limited to a specific region but is adaptable to diverse climatic and geographic contexts. The proposed methodology employs Mutual Information (MI) to evaluate the performance of GCM in simulating historical precipitation, followed by comprehensive rating metrics (CRM) to rank models effectively. A novel regional aggregation technique addresses outlier influence, ensuring robust multi-model ensembles. The approach incorporates top-performing GCMs into MMEs using advanced geometric and regression methods, validated using the Kling-Gupta efficiency with knowable moments (KGE_{km}). A Gaussian-Norm Weighted Drought Index (GNWDI) was also introduced, offering enhanced drought standardization within the Standardized Precipitation Index (SPI) framework. Applying this framework in Punjab, Pakistan, using 22 GCMs, enabled the identification of high-performing models such as MIROC-ES2L, CMCC-CM2-SR5, and IPSL-CM6A-LR. Future drought trends for 2015–2100 were projected under three Shared Socioeconomic Pathways (SSPs). Results revealed a rise in extreme droughts and wet conditions under high emission scenarios (SSP5-8.5), highlighting the intensification of drought severity over extended periods. Specifically, under SSP5-8.5, the average probability of extreme droughts (ED) across all time scales is approximately 0.0221, which remains comparable to lower emission scenarios but shows slightly elevated values at longer time scales, such as 48 months (0.025). Additionally, severe wet (SW) conditions notably increase under SSP5-8.5, with the probability rising from 0.044 at 1 month to 0.051 at 12 and 24 months, and peaking at 0.051 again at 48 months, suggesting more frequent extreme hydrological swings under intensified climate forcing. This study significantly advances drought projection techniques by addressing critical gaps in model ranking, aggregation, and standardization. The framework offers a reliable, regionally adaptable tool for policymakers and researchers, enabling proactive drought management and improved climate resilience under varying emission scenarios.

* Corresponding author.

E-mail addresses: 3520235796373@pu.edu.pk (M. Shakeel), hussnain608@gmail.com (H. Abbas), zulfiqar.stat@pu.edu.pk (Z. Ali), at2139@mstate.edu (A. Tariq), mansour@kau.edu.sa (M. Almazroui), s.mohamedabduulkader@griffith.edu.au (S. Kader).

1. Introduction

Climate change has emerged as one of the most important concerns of the twenty-first century (Vicente-Serrano et al., 2020). Anthropogenic activities have intensified hydroclimatic hazards worldwide, including droughts, which now exhibit altered frequency, intensity, and spatial distribution (Bammou et al., 2024; Gao et al., 2023; Zhang et al., 2022). Drought impacts are diverse, affecting agriculture, ecosystems, and human societies. Therefore, precise and frequent assessment of the future drought characteristics is crucial for adequate early warning and mitigation systems.

Numerous drought indices, such as the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974), Palmer Drought Severity Index (PDSI) (Palmer, 1965), Standardized Precipitation Index (SPI) (McKee et al., 1993), Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010) and Reconnaissance Drought Index (RDI) (Tsakiris and Vangelis, 2005), are widely employed for drought assessment and monitoring. These indices leverage varying inputs and address regional drought patterns (Mukherjee et al., 2018; Zargar et al., 2011). However, data availability and source uncertainties remain critical challenges (Ford and Quiring, 2019). Precipitation is central in characterizing drought severity and occurrence (Vicente-Serrano et al., 2022). Global Climate Models (GCMs), particularly those from the Coupled Model Intercomparison Project (CMIP), have become indispensable tools for simulating precipitation and projecting long-term drought simulations (Jones, 2021; Sharma et al., 2021). Despite the advancements in CMIP simulations, GCM-based projections face limitations due to coarse resolution, model biases, and regional variabilities (Ferreira et al., 2023; Wu et al., 2022).

The Multi-Model Ensemble (MME) approach addresses these limitations by integrating multiple GCMs, enhancing reliability through bias mitigation and variability capture (Gholami et al., 2023; Shao et al., 2024). However, including poorly performing GCMs introduces uncertainties, highlighting the need for region-specific GCM ranking and selection. Recent studies have adopted performance-based ranking frameworks using statistical metrics, including correlation coefficient (CC), Nash-Sutcliffe efficiency (NSE), and skill scores, to identify top-performing models (Sa'adi et al., 2024; Sreelatha and Anand Raj, 2021). However, these methods often assume linearity and independence among models, which are restrictive in complex climate systems. This may lead to biased model selection and underestimation of uncertainties in projections, ultimately reducing the reliability of climate change assessments and associated decision-making.

The study aims to improve the reliability of future drought projections under three emission scenarios by ranking GCMs based on historical precipitation performance and integrating top-performing models into MMEs. The objective is to demonstrate that the drought assessment framework can be effectively applied using precipitation data collected from multiple regional stations. This study achieved its overarching aim and objective by introducing a novel framework that uses Mutual Information (MI)-based methods to assess GCM performance by capturing complex, non-linear dependencies between simulated and observed precipitation data, reducing redundancy among models, and demonstrating its effectiveness through application to data from 28 stations in the Punjab region of Pakistan. The proposed framework is designed to be scalable and adaptable for use in other regions and scenarios. Unlike traditional methods such as Pearson correlation, MI can detect linear and non-linear relationships, making it particularly effective for identifying subtle and non-obvious patterns in climate data. The regional aggregation approach further minimizes the influence of outliers, enhancing the robustness of MMEs. The Gaussian-Norm Weighted Drought Index (GNWDI), coupled with K-Component Gaussian Mixture Models (K-CGMM), is proposed for standardized drought categorization. This work contributes significantly to global climate science by providing a robust methodology for drought characterization, offering insights for policymakers and researchers in

proactive drought management under varying emission pathways.

2. Methodological approach

2.1. The proposed framework for weighted regional aggregation and GNWDI

Consider a study space ω consisting of n th grid points, denoted $P = \{p_1, p_2, \dots, p_n\}$. Each k th GCM, represented as $G = \{G_1, G_2, \dots, G_k\}$, provides monthly historical precipitation simulations at each point in the grid. These simulations are expressed as $X_n = \{x_{G1}, x_{G2}, \dots, x_{Gk}\}$, along with the response series Y_n . For future emission scenarios, GCM simulations are denoted as $Sin = \{s_{G1}, s_{G2}, \dots, s_{Gk}\}$, where i indicates shared socioeconomic pathways (SSP).

Multiple evaluation metrics are considered for each n th grid point, denoted as $M = \{M_1, M_2, \dots, M_e\}$. These metrics provide ranking matrices of the order $(n \times k)$, represented as $R1, R2, \dots, Re$. A unified ranking matrix of size $(n \times k)$ is extracted in a suitable comprehensive framework and is denoted as R . Based on this unified ranking matrix, a subset of the top-performing GCMs is selected from the available pool $G = \{G_1, G_2, \dots, G_k\}$. This subset is denoted as $G = \{G_1, G_2, \dots, Gr\}$ and is obtained using the unified classification matrix R using an appropriate group decision-making approach (GDM). The evaluation metrics are discussed in detail under [Appendix A](#). The equations from (A.1) to (A.8) of [Appendix A](#) explain the mathematical working of the evaluation metrics employed. The complete framework of the proposed scheme is built on these notations.

2.1.1. Unified regional ranking and subset selection

The unified ranking matrix R of order $(n \times k)$ is constructed using a comprehensive metric derived from the metrics $R1, R2, R3, R4$ and $R5$. This construction uniquely ranks the k th GCM at each n th grid point. For the subset selection of highly effective GCMs, the rankings are determined by considering each GCM's weaknesses and strengths at every grid point. This bidirectional approach to GCM ranking has been used in multiple studies (Morais and de Almeida, 2012; Rao and Lakshmi, 2021). Consequently, the best-performing GCMs with the highest net strength scores are selected, denoted as $G = \{G_1, G_2, \dots, Gr\}$, and these are further incorporated into regional aggregation, MMEs, and future drought projections under SSP. Further methodological details are provided in [Appendix B](#) (Equations B.1, C.1–C.3).

2.1.2. Weighted regional aggregation

The best-performing subset of GCMs, $G = \{G_1, G_2, \dots, Gr\}$, is selected based on the net strength score of the pool of all GCMs $G = \{G_1, G_2, \dots, G_k\}$. For the historical period, the r th GCM has a simulation $X_n = \{x_{G1}, x_{G2}, \dots, x_{Gr}\}$ with a response Y_n at each n th point in the grid. The framework developed by [Shakeel and Ali \(2024\)](#) is used for regional aggregation. This framework reduces the influence of outliers and enables weighted simulation aggregation across the study region.

The process is repeated for emission scenarios, where the r th GCMs have simulations $Sin = \{s_{G1}, s_{G2}, \dots, s_{Gr}\}$ at each n th grid point. The row-wise median across the n th grid point for each r th GCM is taken as the response Y_n , as it is not available for emission scenarios.

2.1.3. Development and evaluation of MMEs

The top-selected GCMs, $G = \{G_1, G_2, \dots, Gr\}$, are further aggregated in the MMEs. Six MME approaches from three frameworks were used, including baseline comparison, regression-based models, and geometric methods. These frameworks are further evaluated based on the Kling-Gupta Efficiency with Knowable Moments (KGE_{km}) metric ([Pizarro and Jorquerá, 2024](#)) for the historical period in capturing the precipitation trend. The KGE_{km} metric evaluates MMEs considering multidimensional aspects. It provides a more robust assessment by addressing the limitations of traditionally used metrics, such as the Nash-Sutcliffe Efficiency (NSE) ([Nash and Sutcliffe, 1970](#)) and the Kling-Gupta

Efficiency (KGE) (Gupta et al., 2009; Herbei et al., 2024). Appendices D and E provide details and formulations of the MMEs and evaluation metrics (Equations D.1–D.5 and E.1).

2.1.4. Future drought assessment under SPI

Instead of relying on the Gamma-based framework of the SPI or other traditional univariate probability models, the K-CGMM captures the future precipitation trend. The appropriateness of the K-CGMM is evaluated using density plots, Q-Q plots, and Bayesian Information Criterion (BIC) values across all SSPs and seven different time scales,

along with its competitors, 32 univariate probability models. Details of the K-CGMM are provided in Appendix G, and equations (F.1) and (F.2) describe the density functions.

2.1.5. Trend analysis

The Markov chain-based transition probabilities and traditional nonparametric trend analysis methods have also been employed, such as the Mann-Kendall test and Sen's slope (Gebru et al., 2022). Recent studies (Abbas and Ali, 2024; Naz and Ali, 2024) have used steady-state probabilities of the Markov chain to examine trends across seven

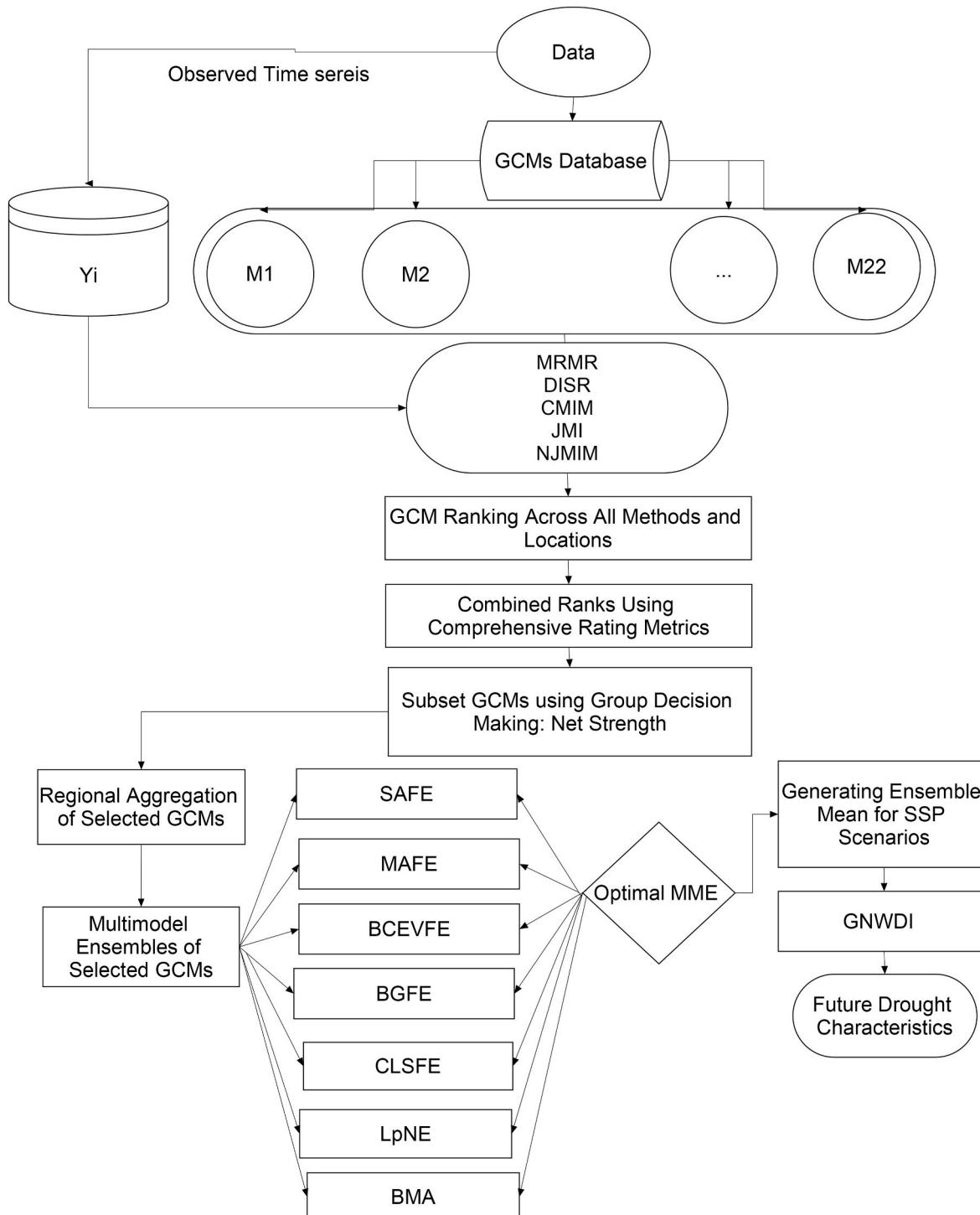


Fig. 1. Design and workflow of the study.

drought states under SSPs. Detailed information on trend steady-state probabilities of the Markov chain is provided in Appendix G. The process starts with equation (G.1), the transition matrix (P) is defined using equation (G.2), and the constant probabilities are obtained by solving equation (G.3). Fig. 1 illustrates the overall methodological approach of the study.

3. Study area and the application

The study is focused on Punjab province, Pakistan, located at a longitude of 72.14° E latitude of 30.82° N, covering an area of 205,344 km 2 (79,284 sq mi). The increasing risk of extreme weather events, including floods and droughts, has become significant. The region suffered from a prolonged drought from 1998 to 2002 (Hina et al., 2021). Consequently, robust frameworks for future drought evaluation are imperative to sustain Punjab's agricultural production amid climate change. This work utilized monthly precipitation data from 28 grid points across Punjab, Pakistan, using 22 GCMs from the CMIP6 ensemble. The historical monthly data span 1950–2014, while monthly projections under three SSPs (SSP1-2.6, SSP2-4.5 and SSP5-8.5) cover the period from 2015 to 2100. All the datasets were retrieved from the Copernicus Climate Data Store (<https://climate.copernicus.eu>).

4. Results

4.1. Performance evaluation of GCMs using mutual information-based methods and CRM scores

Five mutual information-based evaluation methods, MRRMR, JMI, CMIM, DISR, and NJMIM, were integrated in this study. Table 1 shows their ability to reproduce historical precipitation at specific grid points, demonstrating the models' performance across different aspects. Additionally, the CRM score delivers the unified ranking by aggregating them under these methods. The results show that CMCC-CM2-SR5, CNRM-CM6-1, MIROC-ES2L, IPSL-CM6A-LR, and INM-CM5-0 consistently perform well in MRRMR, CMIM, and NJMIM; however, they have low ranks under JMI and DISR.

Despite their lower ranks under JMI and DISR, these models dominate the CRM scores with values of 0.60, 0.57, 0.55, 0.55, and 0.53, corresponding to ranks 1 to 5. In contrast, there are also models with consistently low performance, such as TaiESM1, with their rank scores as follows: MRRMR (Rank 22), JMI (Rank 7), CMIM (Rank 22), DISR

(Rank 6), and NJMIM (Rank 22). Other low-performing models include CanESM5-CanOE, FGOALS-f3-L, and NorESM2-MM, which perform exceptionally well under JMI and DISR but are overshadowed by poor performance in most other ranking methods. Uncertainty in GCM rankings highlighted in this table underscores the significance of multi-criteria evaluation methods. This is the case for one grid point. The heatmap in Fig. S1 provides the ranking of the GCMs in the study region under the unified CRM score.

The GCMs CMCC-CM2-SR5, CNRM-CM6-1, MIROC-ES2L, and IPSL-CM6A-LR emerge as the strongest performers, exhibiting stable and outstanding performance across most locations, with few exceptions. GCMs such as TaiESM1, FGOALS-f3-L, CanESM5-CanOE, and BCC-CSM2-MR are among the worst performers, consistently ranked in the lower level with ranks 19 to 22.

Table 2
Net strength performance score-based comparative ranking of GCMs.

Models Name	Model Strength	Model Weakness	Net Performance Score
ACCESS-CM2	1.5	177.5	-176
AWI-CM-1-1-MR	160	12	148
BCC-CSM2-MR	57	107.5	-50.5
NorESM2-MM	3	147.5	-144.5
CanESM5-CanOE	0	278	-278
CESM2	14	134.5	-120.5
CMCC-CM2-SR5	207.5	0	207.5
MIROC-ES2L	279	0	279
CNRM-CM6-1	196.5	1.5	195
CNRM-CM6-1-HR	75	31.5	43.5
CNRM-ESM2-1	82	14.5	67.5
MIROC6	57.5	36	21.5
UKESM1-0-LL	51.5	121	-69.5
FGOALS-f3-L	0	244	-244
TaiESM1	0	301.5	-301.5
KACE-1-0-G	26	86	-60
GFDL-ESM4	13	134.5	-121.5
MPI-ESM1-2-LR	78	95.5	-17.5
MRI-ESM2-0	85.5	128	-42.5
INM-CM4-8	77.5	76	1.5
INM-CM5-0	152.5	22	130.5
IPSL-CM6A-LR	228	8	220

Table 1
Multi-method evaluation and comprehensive ranking of GCMs.

Model Names	MRRMR Score	JMI Score	CMIM Score	DISR Score	NJMIM Score	CRM Score
ACCESS-CM2	0.57 (13)	17.9 (10)	0.6 (13)	2.83 (10)	0.21 (13)	0.46 (13.5)
AWI-CM-1-1-MR	0.74 (7)	11.88 (15)	0.77 (7)	1.63 (16)	0.25 (8)	0.52 (6.5)
BCC-CSM2-MR	0.57 (14)	18.64 (9)	0.6 (14)	2.98 (9)	0.2 (14)	0.45 (15)
NorESM2-MM	0.44 (18)	21.1 (5)	0.46 (18)	3.38 (5)	0.17 (18)	0.42 (19.5)
CanESM5-CanOE	0.36 (21)	22.06 (1)	0.37 (21)	3.54 (1)	0.13 (21)	0.41 (21)
CESM2	0.43 (19)	21.93 (3)	0.45 (19)	3.51 (3)	0.15 (19)	0.43 (17.5)
CMCC-CM2-SR5	0.98 (1)	0.98 (22)	0.98 (1)	0.98 (19)	0.98 (1)	0.6 (1)
MIROC-ES2L	0.89 (2)	2.01 (21)	0.96 (2)	0.31 (22)	0.31 (2)	0.55 (3)
CNRM-CM6-1	0.8 (4)	9.04 (17)	0.86 (4)	1.15 (18)	0.29 (4)	0.57 (2)
CNRM-CM6-1-HR	0.72 (8)	13.42 (14)	0.75 (8)	2.13 (14)	0.25 (10)	0.51 (8)
CNRM-ESM2-1	0.61 (11)	16.71 (11)	0.64 (12)	2.67 (11)	0.23 (11)	0.49 (10)
MIROC6	0.68 (10)	10.45 (16)	0.72 (10)	1.88 (15)	0.26 (7)	0.47 (12)
UKESM1-0-LL	0.53 (16)	20.88 (6)	0.56 (17)	3.32 (7)	0.18 (17)	0.43 (17.5)
FGOALS-f3-L	0.38 (20)	21.93 (2)	0.41 (20)	3.52 (2)	0.14 (20)	0.42 (19.5)
TaiESM1	0.28 (22)	20.86 (7)	0.29 (22)	3.38 (6)	0.11 (22)	0.28 (22)
KACE-1-0-G	0.61 (12)	15.59 (12)	0.64 (11)	2.46 (12)	0.22 (12)	0.46 (13.5)
GFDL-ESM4	0.53 (17)	21.76 (4)	0.56 (16)	3.46 (4)	0.19 (16)	0.48 (11)
MPI-ESM1-2-LR	0.54 (15)	19.77 (8)	0.57 (15)	3.14 (8)	0.19 (15)	0.45 (16)
MRI-ESM2-0	0.76 (5)	5.68 (19)	0.81 (5)	0.88 (20)	0.27 (6)	0.5 (9)
INM-CM4-8	0.71 (9)	14.88 (13)	0.75 (9)	2.37 (13)	0.25 (9)	0.52 (6.5)
INM-CM5-0	0.76 (6)	7.3 (18)	0.81 (6)	1.42 (17)	0.28 (5)	0.53 (5)
IPSL-CM6A-LR	0.86 (3)	3.92 (20)	0.94 (3)	0.61 (21)	0.31 (3)	0.55 (4)

4.2. Selection of high-performing GCM subsets based on net strength scores

Table 2 shows the aggregated score for identifying the subsets of GCMs that perform best for the study region, considering the GCM rankings in 28 grid points. The net strength of each GCM is calculated by subtracting the weakness from the strength of each GCM in the study region.

The GCMs MIROC-ES2L and CMCC-CM2-SR5 are the most appropriate models, with strengths of 279 and 207.5, without notable weaknesses. The poorly performing GCMs with negative Net Strength, such as MPI-ESM1-2-LR (-17.5), BCC-CSM2-MR (-50.5), and KACE-1-0-G (-60), make them the least reliable for predictions in the study region. Therefore, ten GCMs with net strength—MIROC-ES2L (279), CMCC-CM2-SR5 (207.5), IPSL-CM6A-LR (220), CNRM-CM6-1 (195), INM-CM5-0 (130.5), AWI-CM-1-1-MR (148), CNRM-ESM2-1 (67.5), CNRM-CM61-HR (43.5), MIROC6 (21.5), and INM-CM4-8 (1.5)—exhibit reasonable performance and are considered to be part of the high-performing group. The selected subset of GCMs will be considered for regional aggregation and the development of MMEs.

4.3. Performance of MMEs in capturing historical annual precipitation trends

The preferred choice of GCMs is incorporated into seven MMEs: SAFE, MAFE, BCEVFE, BGFE, CLSFE, LpNE and BMA. **Fig. S2** illustrates the performance of these MMEs in capturing the inherent uncertainty of annual precipitation from 1950 to 2014 for Punjab. The observed series fluctuates consistently with multiple up and down peaks, showing notable inter-annual variability over the entire period. SAFE, MAFE, BMA and BGFE consistently underestimate the observed annual precipitation. These four MMEs fail to capture the peaks, exhibiting a smoother trend than the observed variability, indicating limited effectiveness for the study region.

However, there is a slight improvement under CLSFE, although it still significantly underestimates the observed precipitation. For the year 1993, CLSFE fails to capture annual variability accurately. BCEVFE and LpNE demonstrate greater precision in capturing annual average variability and show closer agreement with the observed series. These two ensembles effectively capture the peaks and troughs. However, the LpNE, due to its higher correlation with observed precipitation, captures most peaks and troughs more accurately than BCEVFE.

The superior performance of LpNE in both magnitude and variability of observed annual precipitation trends is shown quantitatively in **Table 3**. LpNE, with the highest correlation (0.688), zero bias, perfect variability (1.002), and the highest KGE value of 0.688, is the most reliable model to forecast annual precipitation trends in Punjab.

4.4. Evaluation of probability models and drought classification under the SPI framework

Choosing an appropriate probability model under the SPI framework ensures robust drought classification, minimizing errors and biases across multiple timescales. This critical step is illustrated in **Fig. 2**, **Figure S3 (a-d)**, and **Figure S4 (a-d)**, where the probability distributions

Table 3

KGEkm-based assessment of MMEs in the historical period.

Ensemble Methods	r	Beta.2021	Alpha	KGE
SAFE	0.638	-1.319	0.821	-0.38
MAFE	0.602	-1.468	0.798	-0.534
LpNE	0.688	0	1.002	0.688
BMA	0.029	-1.334	0.816	-0.661
BCEVFE	0.618	0	1.002	0.618
BGFE	0.607	-0.964	0.873	-0.048
CLSFE	0.537	-0.425	0.946	0.369

are fitted to the precipitation series in three SSPs (1–2.6, 2–4.5, and 5–8.5). In this study, mixture probability models have been tested in addition to the gamma-based SPI to improve the accuracy of drought prediction.

The density and QQ plots for K-CGMM for all three SSPs demonstrate their superiority over other univariate probability models. It closely aligns with the simulated precipitation series by addressing anomalies such as skewness and heavy tails. Especially for SSP5-8.5, where the variability increases, the univariate probability models fail to capture the tail ends and the overall distribution shape. The accuracy of drought projections is primarily based on the tail behavior of the probability distribution, and K-CGMM provides a robust solution to this end. **Table 4** presents the quantitative selection of appropriate probability distribution models based on BIC values across the three SSPs, evaluated over short-to long-term time scales, including 1, 3, 6, 12, and 48 months.

Across all SSPs and time scales, particularly at longer time scales (24 and 48 months), K-CGMM provides significantly lower BIC values, indicating its efficiency in accurately modeling and identifying prolonged drought conditions. Therefore, K-CGMM is recommended for reliable drought monitoring and forecasting in the study region to minimize the risk of underestimating extreme values.

A standardization process is applied using K-CGMM for the three SSPs on seven different time scales to identify drought conditions. **Fig. 3** **Figure S5**, and **Figure S6** show the dry and wet conditions for each SSP: SSP1-2.6, SSP2-4.5, and SSP5-8.5. For SSP1-2.6, frequent fluctuations between wet and dry conditions are apparent on shorter time scales, 1, 3, and 6 months. However, dry conditions become more dominant and spread over the extended timescales, indicating extreme drought events. Wet events are also noticeable in SSP1-2.6, indicating less severe long-term droughts later in the century.

For SSP2-4.5 and SSP5-8.5, there is higher variability compared to SSP1-2.6, with more frequent red bars for shorter and longer periods. Persistent drought patterns become more noticeable and pronounced with an increasing time scale. The duration of dry periods becomes more severe and intense, with frequent drought events, especially for SSP5-8.5.

The long-term behavior of drought patterns among its different classes is analyzed using the steady-state probabilities. **Table 5** presents the probabilities of remaining within each drought category over time under three SSP scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5) across various short-to long-term timescales, ranging from 1 to 48 months.

For SSP1-2.6, the highest likelihood value indicates a relatively stable condition, as Normal Drought (ND) holds probability values between 0.673 and 0.685. Furthermore, under this low-emission scenario, the increase in time scales shows no significant variation in probabilities between drought categories, making it more favorable for drought management. Similarly, the trend observed for SSP2-4.5 and SSP5-8.5 shows slightly higher variability in drought and wet extremes than SSP1-2.6. However, a noticeable decline in the likelihood of ND can be observed for more extended time scales than the 1-month time scale. However, a slight increase in the likelihood of longer time scales is observed for Extreme Drought (ED) and Severe Wet (SW) in the high-emission scenario. This suggests more frequent extreme droughts and wet conditions under SSP5-8.5 over longer time scales.

5. Discussions

This study advances the state of drought characterization methodologies by integrating innovative approaches for evaluating and ranking GCMs and constructing robust MMEs. The adoption of MI-based methods has proven instrumental in unravelling the variability among GCMs' abilities to replicate historical precipitation patterns, a finding consistent with [Sharma et al. \(2021\)](#), emphasizing the need for rigorous model evaluation. Notably, high-performance models such as CMCC-CM2-SR5, CNRM-CM6-1, MIROC-ES2L, IPSL-CM6A-LR, and INM-CM5-0 emerged as reliable for regional applications. The

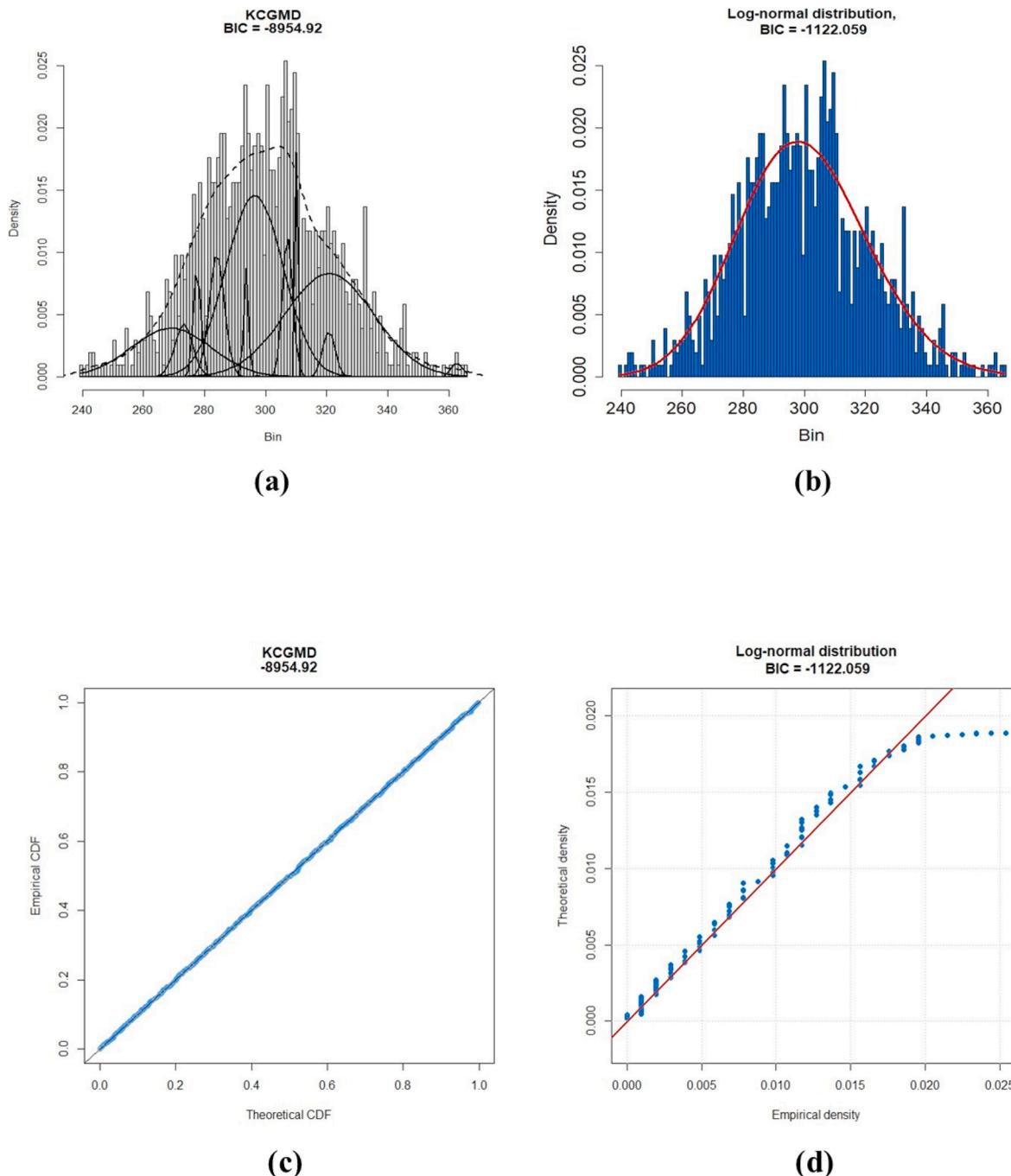


Fig. 2. Visual comparison of univariate and mixture distribution models under SSP1-2.6 at a 9-month timescale: (a) Density curve for the mixture model, (b) Density curve for the univariate model, (c) Quantile-Quantile (QQ) plot for the mixture model, (d) QQ plot for the univariate model.

discrepancies in performance across MI methods, such as JMI and DISR, underscore the necessity of aggregation frameworks like CRM for optimal model selection.

The spatial consistency observed in CRM heatmaps (Fig. S1) and Table 2 highlights the performance disparities among GCMs. The MIROC-ES2L model achieved the highest net strength score of 279, followed by IPSL-CM6A-LR with 220, reinforcing their suitability for regional hydrological modeling and adaptation planning. In contrast, TaiESM1 (-301.5) and FGOALS-f3-L (-244) demonstrated significant underperformance, justifying their exclusion from the MME. These findings align with Fotse et al. (2024), emphasizing the importance of accurate model selection criteria to prevent biased projections.

The novel application of K-CGMM within the SPI framework

addresses critical challenges such as multimodal and heavy-tailed distributions. At time scale 9 for SSP1-2.6, the K-CGMM model achieved the lowest BIC value of -8954.92 , significantly outperforming the univariate probability distribution model with a BIC value of -1122.05 . This substantial difference demonstrates the superior goodness-of-fit of K-CGMM, emphasizing its effectiveness in accurately capturing the underlying data structure. This methodological enhancement improves the accuracy of drought classification, enabling more precise projections under varying SSPs. The study's ability to capture prolonged drought conditions, especially under high-emission scenarios like SSP5-8.5, underscores the escalating risk of extreme droughts and agrees with the intensified trends reported by previous studies (Sharma et al., 2021; Sa'adi et al., 2024).

Table 4

Quantitative evaluation of probability models using BIC across different time scales and SSPs.

SSPs	Time scales	K-CGMM	Univariate models
SSP1-2.6	1	-6747.24	-578.39 (Inverse Gaussian)
	3	-7857.65	-651.71 (Gamma)
	6	-8550.42	-884 (Inverse Gamma)
	9	-8954.92	-1122.06 (Log-normal)
	12	-9254.63	-750.68 (Logistic)
	24	-9852.47	-1003 (Generalized Extreme Value)
	48	-10451.5	-1369.16 (Normal)
	1	-6560.11	-578.21 (Gamma)
	3	-7644.41	-987.39 (Logistic)
SSP2-4.5	6	-8388.43	-866.21 (Gamma)
	9	-8804.7	-1029.02 (Gamma)
	12	-9088.97	-739.64 (Normal)
	24	-9736.86	-856.11 (Triangular)
	48	-10187.9	-1284.66 (Triangular)
	1	-6937.34	-743.13 (Chi-Square)
SSP5-8.5	3	-8102.45	-680.97 (Normal)
	6	-8839.59	-1135.78 (Gamma)
	9	-9229.04	-748.59 (Inverse Gamma)
	12	-9512.52	-902.83 (Log-normal)
	24	-10171	-1126.3 (Logistic)
	48	-10574.5	-667.32 (Laplace)

The comparative analysis of MMEs revealed stark differences in their capabilities to replicate historical precipitation variability. SAFE and BGFE showed significant underestimations with KGEkm values of -0.38 and -0.048, respectively, while LpNE achieved the highest KGEkm (0.688) and correlation ($R = 0.688$). LpNE also maintained a balanced

Beta (0.000) and the highest Alpha (1.002), ensuring better variability capture and reduced bias. This finding resonates with [Battool et al. \(2023\)](#), who emphasized the importance of robust indices for capturing joint climatic extremes. Integrating the multicomponent KGEkm measure, derived from K-moments, for validating MME aligns with the advancements discussed in [Pizarro and Jorquera \(2024\)](#). This approach enhances the robustness of model evaluation, further improving the reliability of drought characterization models.

The projected increase in drought intensity and frequency under high-emission scenarios highlights the critical need for adaptive management and policy interventions. The relatively stable conditions under SSP1-2.6, supporting previous research outcomes ([Battool et al., 2023; Sharma et al., 2021; Wu et al., 2021; Xu et al., 2021](#)), strongly advocate adopting low-emission pathways. Overall, integrating the advanced ranking methods using MI, robust aggregation techniques, and innovative indices like the GNWDI establishes this research as a significant contributor to drought characterization and management.

6. Conclusions

This study presents a robust and innovative framework for future drought characterization, addressing key challenges in GCM evaluation, ranking, and ensemble modeling. The study effectively captures nonlinear relationships and reduces redundancy by employing MO-based methods, enabling a more effective assessment of GCM performance. The CRM further enhances region-specific model selection, minimizing uncertainties in MMEs and improving the reliability of predictions. Key findings reveal the efficacy of the framework in

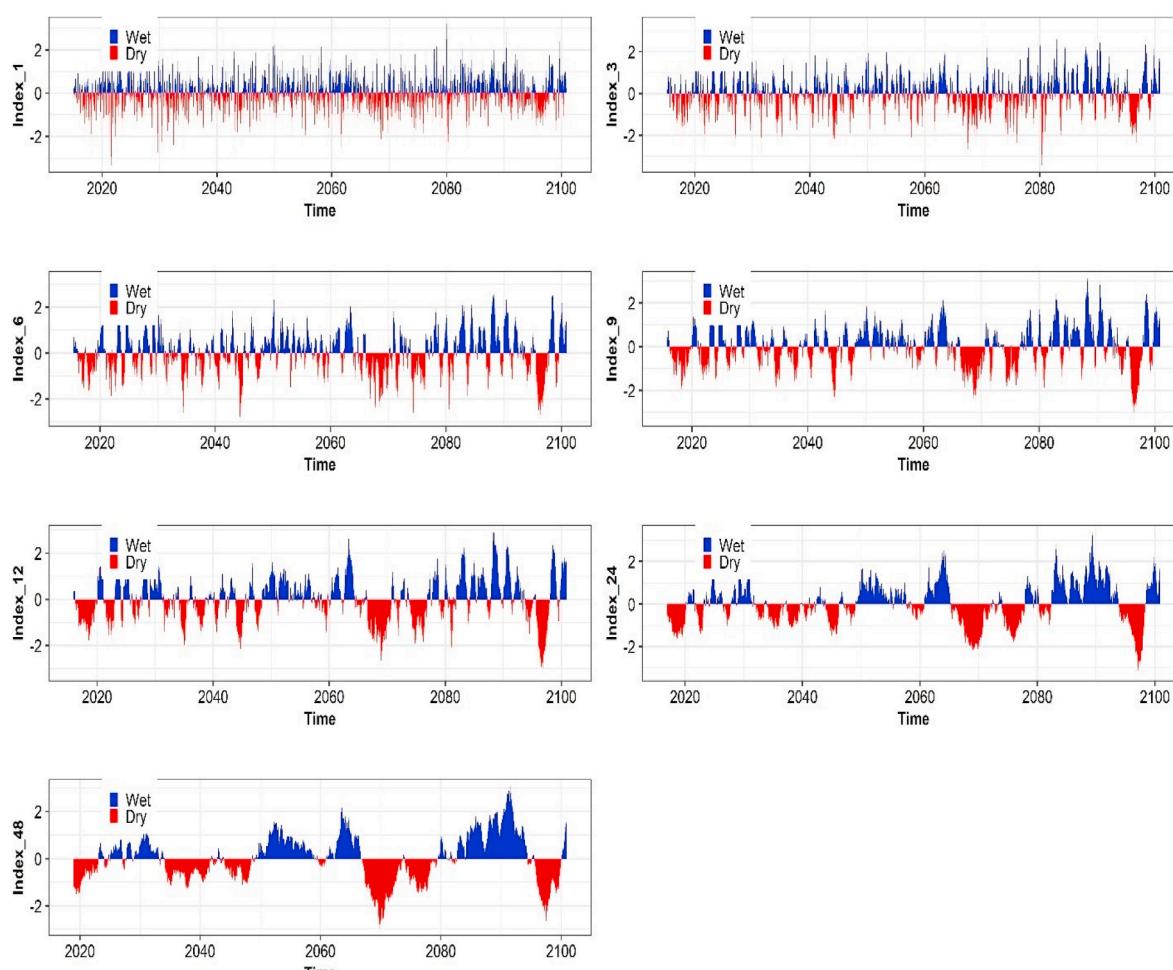


Fig. 3. Multi-timescale analysis of drought classification for SSP1-2.6.

Table 5

Steady-state probability-based evaluation of long-term drought trends across various time intervals and emission scenarios.

Future SSPs	Time Scales	Drought Classification					
		ED	EW	MD	MW	ND	SD
SSP1-2.6	1	0.021	0.023	0.092	0.087	0.685	0.045
	3	0.024	0.023	0.092	0.091	0.681	0.044
	6	0.023	0.023	0.092	0.092	0.679	0.044
	9	0.021	0.030	0.091	0.096	0.676	0.047
	12	0.022	0.022	0.096	0.094	0.679	0.042
	24	0.024	0.024	0.094	0.095	0.673	0.042
	48	0.023	0.028	0.087	0.094	0.679	0.039
							0.050
SSP2-4.5	1	0.022	0.022	0.096	0.094	0.679	0.042
	3	0.023	0.022	0.091	0.094	0.681	0.044
	6	0.022	0.023	0.095	0.092	0.677	0.043
	9	0.021	0.023	0.088	0.094	0.682	0.048
	12	0.022	0.022	0.087	0.089	0.686	0.048
	24	0.021	0.023	0.087	0.092	0.683	0.043
	48	0.022	0.025	0.083	0.084	0.687	0.040
							0.059
SSP3-8.5	1	0.024	0.023	0.092	0.090	0.684	0.043
	3	0.018	0.024	0.090	0.092	0.686	0.047
	6	0.022	0.024	0.094	0.093	0.679	0.043
	9	0.022	0.024	0.093	0.092	0.680	0.043
	12	0.021	0.021	0.091	0.095	0.675	0.046
	24	0.023	0.026	0.093	0.093	0.672	0.043
	48	0.025	0.034	0.094	0.098	0.663	0.041
							0.045

identifying top-performing GCMs, such as CMCC-CM2-SR5, CNRM-CM6-1, MIROC-ES2L, IPSL-CM6A-LR, and INM-CM5-0, which consistently excelled across metrics at 28 grid points. Among MMEs, LpNE emerged as the most robust ensemble, with its high correlation (0.688), minimal bias, and near-perfect variability (1.002). The study also highlights the effectiveness of the K-CGMM within the SPI framework for improving drought classification accuracy, particularly for extended timescales and under high-emission scenarios.

Projections under SSP5-8.5 indicate an intensification of drought severity and frequency, with prolonged dry spells and more frequent extreme events becoming more common. These findings underscore the heightened risks of high-emission pathways and advocate for adopting low-emission pathways (SSP1-2.6) to mitigate future drought impacts. This study's novel contribution includes multiphase advancements in model ranking, aggregation techniques, and drought indices, such as the GNWDI, providing valuable tools for proactive drought management and climate resilience planning.

Despite its advancements, the study has certain limitations. First, the approach primarily emphasizes precipitation-based drought indicators, possibly neglecting the consequences of temperature, evapotranspiration, and soil moisture on drought dynamics. Second, the regional application to Punjab, Pakistan, while demonstrating adaptability, may limit the generalizability to other regions with different hydroclimatic contexts. The framework's computational complexity arises from the grid-based evaluation of multiple GCMs using several performance metrics, the construction of unified ranking matrices, and repeated model selection and aggregation steps for historical and projected scenarios. These steps involve intensive matrix operations and iterative decision-making processes across large datasets, which can be computationally expensive, especially when applied to large regions or long time series.

Future research should expand the framework's application to incorporate multivariable indices like SPEI and PDSI for a more thorough evaluation of drought dynamics. These indices integrate additional environmental variables such as temperature, evapotranspiration, and soil moisture, offering a more detailed characterization of drought conditions. Incorporating such indices using the developed methodology would address the limitations of relying solely on precipitation and enhance the framework's applicability for capturing complex drought dynamics across different climatic regions. Regional studies across diverse climatic zones are needed to validate the adaptability and scalability of the framework. Integrating higher-resolution climate

projections and ML techniques could enhance model performance and computational efficiency. Machine learning models could be trained to approximate specific steps in the ranking and selection pipeline, reducing execution time without sacrificing accuracy. Finally, exploring socio-economic impacts and adaptive capacity within the framework would provide policymakers with actionable insights for holistic climate resilience strategies. Overall, this study lays a strong foundation for advancing global efforts in drought characterization and mitigation under changing climate conditions.

CRediT authorship contribution statement

Muhammad Shakeel: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hussnain Abbas:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis. **Zulfiqar Ali:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Aqil Tariq:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Mansour Almazroui:** Writing – review & editing, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation. **Shuraik Kader:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation.

Consent to participate

Not applicable

Consent to publish

Not applicable

Ethical approval

Not applicable.

Funding

No fundings received for this study.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jenvman.2025.126692>.

Data availability

Data will be made available on request.

References

- Abbas, H., Ali, Z., 2024. A novel statistical framework of drought projection by improving ensemble future climate model simulations under various climate change scenarios. Environ. Monit. Assess. 196 (10), 1–27. <https://doi.org/10.1007/s10661-024-13108-w>.
- Bammou, Y., Benzougagh, B., Igoullan, B., Kader, S., Ouallali, A., Spalevic, V., Kuriqi, A., 2024. Spatial mapping for multi-hazard land management in sparsely vegetated watersheds using machine learning algorithms. Environ. Earth Sci. 83 (15), 447. <https://doi.org/10.1007/s12665-024-11741-9>.
- Batool, A., Ali, Z., Mohsin, M., Shakeel, M., 2023. A generalized procedure for joint monitoring and probabilistic quantification of extreme climate events at regional level. Environ. Monit. Assess. 195 (10), 1223. <https://doi.org/10.1007/s10661-023-11717-5>.
- Ferreira, G.W.d.S., Reboita, M.S., Ribeiro, J.G.M., De Souza, C.A., 2023. Assessment of precipitation and hydrological droughts in South America through statistically downscaled CMIP6 projections. Climate 11 (8), 166. <https://doi.org/10.3390/cl11080166>.
- Ford, T.W., Quiring, S.M., 2019. Comparison of contemporary in situ, model, and satellite remote sensing soil moisture with a focus on drought monitoring. Water Resour. Res. 55 (2), 1565–1582. <https://doi.org/10.1029/2018WR024039>.
- Fotse, A.R.G., Guenang, G.M., Mbenda, A.J.K., Vondou, D.A., 2024. Appropriate statistical rainfall distribution models for the computation of standardized precipitation index (SPI) in Cameroon. Earth Science Informatics 17 (1), 725–744. <https://doi.org/10.1007/s12145-023-01188-0>.
- Gao, Y., Fu, S., Cui, H., Cao, Q., Wang, Z., Zhang, Z., Qiao, J., 2023. Identifying the spatio-temporal pattern of drought characteristics and its constraint factors in the yellow river Basin. Ecol. Indic. 154, 110753. <https://doi.org/10.1016/j.ecolind.2023.110753>.
- Gebru, B.M., Adane, G.B., Park, E., Khamzina, A., Lee, W.-K., 2022. Landscape pattern and climate dynamics effects on ecohydrology and implications for runoff management: case of a dry afromontane forest in northern Ethiopia. Geocarto Int. 37 (26), 12466–12487. <https://doi.org/10.1080/10106049.2022.2068673>.
- Gholami, H., Lotfifrad, M., Ashrafi, S.M., Biazar, S.M., Singh, V.P., 2023. Multi-GCM ensemble model for reduction of uncertainty in runoff projections. Stoch. Environ. Res. Risk Assess. 37 (3), 953–964. <https://doi.org/10.1007/s00477-022-02311-1>.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. Journal of hydrology 377 (1), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>.
- Herbei, M.V., Bădăluță-Minda, C., Popescu, C.A., Horablaia, A., Dragomir, L.O., Popescu, G., Sestrăs, P., 2024. Rainfall-runoff modeling based on HEC-HMS model: a case study in an area with increased groundwater discharge potential. Frontiers in Water 6, 1474990. <https://doi.org/10.3389/fwra.2024.1474990>.
- Hina, S., Saleem, F., Arshad, A., Hina, A., Ullah, I., 2021. Droughts over Pakistan: possible cycles, precursors and associated mechanisms. Geomat. Nat. Hazards Risk 12 (1), 1638–1668. <https://doi.org/10.1080/19475705.2021.1938703>.
- Jones, C.D., 2021. Chapter 4 - numerical modeling of the global climate and carbon cycle system. In: Letcher, T.M. (Ed.), Climate Change, third ed. Elsevier, pp. 67–91.
- McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. Paper Presented at the Proceedings of the 8th Conference on Applied Climatology.
- Morais, D.C., de Almeida, A.T., 2012. Group decision making on water resources based on analysis of individual rankings. Omega 40 (1), 42–52. <https://doi.org/10.1016/j.omega.2011.03.005>.
- Mukherjee, S., Mishra, A., Trenberth, K.E., 2018. Climate change and drought: a perspective on drought indices. Curr. Clim. Change Rep. 4, 145–163. <https://doi.org/10.1007/s40641-018-0098-x>.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. Journal of hydrology 10 (3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Naz, R., Ali, Z., 2024. A novel self-adjusting weight approximation procedure to minimize non-identical seasonal effects in multimodel ensemble for accurate twenty-first century drought assessment. Stoch. Environ. Res. Risk Assess. 38 (6), 2451–2472. <https://doi.org/10.1007/s00477-024-02689-0>.
- Palmer, W.C., 1965. Meteorological Drought, vol. 30. US Department of Commerce, Weather Bureau.
- Pizarro, A., Jorquera, J., 2024. Advancing objective functions in hydrological modelling: integrating knowable moments for improved simulation accuracy. Journal of hydrology 634, 131071. <https://doi.org/10.1016/j.jhydrol.2024.131071>.
- Rao, R., Lakshmi, J., 2021. R-method: a simple ranking method for multi-attribute decision-making in the industrial environment. Journal of Project Management 6 (4), 223–230. <https://doi.org/10.5267/j.pjm.2021.5.001>.
- Rouse, J.W., Haas, R. h., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. Monitoring the Vernal Advancement and Retrogradation (Greenwave Effect) of Natural Vegetation. NASA, Greenbelt, MD, p. 371p. NASA/GSFC Final Report.
- Sa'adi, Z., Alias, N.E., Yusop, Z., Iqbal, Z., Houmsi, M.R., Houmsi, L.N., Muhammad, M.K. I., 2024. Application of relative importance metrics for CMIP6 models selection in projecting basin-scale rainfall over Johor River basin, Malaysia. Sci. Total Environ. 912, 169187. <https://doi.org/10.1016/j.scitotenv.2023.169187>.
- Shakeel, M., Ali, Z., 2024. Integration of Exponential weighted moving average chart in ensemble of precipitation of multiple global climate models (GCMs). Water Resour. Manag. 38 (3), 935–949. <https://doi.org/10.1007/s11269-023-03702-x>.
- Shao, Y., Bishop, C., Abramowitz, G., Hoboichi, S., 2024. Improving multi-model ensembles of climate projections through time variability correction and ensemble dependence transformation. EGU General Assembly. <https://doi.org/10.5194/egusphere-egu24-6796>. Vienna, Austria, 14–19 Apr 2024, EGU24-6796.
- Sharma, S., Hamal, K., Khadka, N., Ali, M., Subedi, M., Hussain, G., Dawadi, B., 2021. Projected drought conditions over southern slope of the central himalaya using CMIP6 models. Earth Systems and Environment 5 (4), 849–859. <https://doi.org/10.1007/s41748-021-00254-1>.
- Seelathala, K., Anand Raj, P., 2021. Ranking of CMIP5-based global climate models using standard performance metrics for Telangana region in the southern part of India. ISH Journal of Hydraulic Engineering 27 (Suppl. 1), 556–565. <https://doi.org/10.1080/09715010.2019.1634648>.
- Tsakiris, G., Vangelis, H., 2005. Establishing a drought index incorporating evapotranspiration. European water 9 (10), 3–11.
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J. Clim. 23 (7), 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>.
- Vicente-Serrano, S.M., Peña-Angulo, D., Beguería, S., Domínguez-Castro, F., Tomás-Burguera, M., Noguera, I., El Kenawy, A., 2022. Global drought trends and future projections. Philosophical Transactions of the Royal Society A 380 (2238), 20210285. <https://doi.org/10.1098/rsta.2021.0285>.
- Vicente-Serrano, S.M., Quiring, S.M., Peña-Gallardo, M., Yuan, S., Domínguez-Castro, F., 2020. A review of environmental droughts: increased risk under global warming? Earth Sci. Rev. 201, 102953. <https://doi.org/10.1016/j.earscirev.2019.102953>.
- Wu, G., Chen, J., Shi, X., Kim, J.S., Xia, J., Zhang, L., 2022. Impacts of global climate warming on meteorological and hydrological droughts and their propagations. Earths Future 10 (3), e2021EF002542. <https://doi.org/10.1029/2021EF002542>.
- Wu, C., Yeh, P.J.F., Ju, J., Chen, Y.Y., Xu, K., Dai, H., et al., 2021. Assessing the spatiotemporal uncertainties in future meteorological droughts from CMIP5 models, emission scenarios, and bias corrections. J. Clim. 34 (5), 1903–1922. <https://doi.org/10.1175/JCLI-D-20-0411.1>.
- Xu, K., Wu, C., Zhang, C., Hu, B.X., 2021. Uncertainty assessment of drought characteristics projections in humid subtropical basins in China based on multiple CMIP5 models and different index definitions. J. Hydrol. 600, 126502. <https://doi.org/10.1016/j.jhydrol.2021.126502>.
- Zargar, A., Sadiq, R., Naser, B., Khan, F.I., 2011. A review of drought indices. Environ. Rev. 19 (NA), 333–349. <https://doi.org/10.1139/a11-013>.
- Zhang, X., Hao, Z., Singh, V.P., Zhang, Y., Feng, S., Xu, Y., Hao, F., 2022. Drought propagation under global warming: characteristics, approaches, processes, and controlling factors. Sci. Total Environ. 838, 156021. <https://doi.org/10.1016/j.scitotenv.2022.156021>.