Short-Term Drought Forecasting in Iran Using Multi-Source Machine Learning: An Assessment of Autoregressive, Teleconnection-Driven, and Hybrid Paradigms

Jun Jian[1, 2], Peyman Mahmoudi[3], Pouria Jafari[4], Alireza Ghaemi[3], Jing Yang[5], Fatemeh Firoozi[6]

[1]Navigation College, Dalian Maritime University, Dalian, China
[2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China
[3]Department of Physical Geography, Faculty of geography and environmental planning, University of Sistan and Baluchestan, Zahedan, Iran
[4]Department of Department of Electronic and Electrical Engineering, Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran
[5]Faculty of Geographical Science, Key Laboratory of Environmental Change and Natural Disaster, Beijing Normal University, Beijing, China
[6]Department of Humanities and Social Science, Farhangyan University, Tehran, Iran

Correspondence to: Peyman Mahmoudi (p_mahmoudi@gep.usb.ac.ir)

Abstract. Meteorological drought forecasting remains a major scientific challenge. This study comparatively evaluates three competing forecasting paradigms—autoregressive (temporal memory), teleconnection-driven (large-scale climate drivers), and hybrid—across Iran. Using 30-year precipitation data from 96 synoptic stations and 19 global climate indices, nine machine and deep learning models were tested to forecast the Standardized Precipitation Index (SPI) at 1-, 2-, and 3-month lead times. The results conclusively reject a universally superior paradigm, showing that the optimal model structure is highly location-dependent. The hybrid approach prevailed in the vast arid and semi-arid regions, while standalone paradigms excelled in specific coastal niches. Among the algorithms, Random Forest (RF) emerged as the most robust. These findings highlight the crucial need to transition from 'one-model-fits-all' approaches toward adaptive, region-centric frameworks for drought early warning systems.

1 Introduction

Skillful drought forecasting, as one of the most destructive and complex natural hazards, is a major scientific challenge and an urgent practical need for policymakers (Deo et al., 2017; Chandrasekara et al., 2021; Oyarzabal et al., 2025). This challenge stems not only from the multi-scale and nonlinear nature of atmosphere-ocean processes (Djerbouai and Souag-Gamane, 2016) but also from a fundamental and unresolved debate in the modeling domain: the strategic choice of the input paradigm (Mishra and Singh, 2011; Fung et al., 2020). In the scientific literature, two main philosophies for predictor selection compete. The first approach, autoregressive, relies on the temporal memory and internal patterns of the drought index time series (Rezaiy and Shabri, 2025; Poudel et al., 2024). The second approach, the teleconnection-driven paradigm,

leverages large-scale climate indices as a source of information for predicting regional conditions (Anshuka et al., 2021; Anochi and Shimizu, 2025). Although both paradigms have shown success in case studies (Ali et al., 2023), a comprehensive assessment evaluating the relative superiority or synergistic potential of these two approaches on a vast geographical scale and using a wide array of modern models is notably absent. This research gap, which relates directly to the heart of forecasting practice, forms the core of the present study.

For quantifying drought in such a comparative assessment, the Standardized Precipitation Index (SPI) has become a standard and global tool due to its statistical robustness, minimal data requirement (only precipitation), and comparability across different spatio-temporal scales (McKee et al., 1993; Oyarzabal et al., 2025). This index allows for a fair evaluation of model performance in regions with diverse climates (Shi et al., 2020; Esquivel-Saenz et al., 2024). Consequently, the ability to accurately predict SPI values is considered the cornerstone of early warning systems and proactive drought risk management (Aldhafeeri et al., 2025).

The evolution of forecasting methods shows a transition from traditional statistical models toward complex machine learning (ML) and deep learning (DL) algorithms (Achite et al., 2023; Gyaneshwar et al., 2023). While standalone models like Artificial Neural Networks (ANN) and Support Vector Machines (SVM) were pioneers in this field (Rezaeianzadeh et al., 2016; Ganguli and Janga Reddy, 2014), the dominant trend in current research is toward hybrid architectures that seek to overcome the limitations of single models by intelligently integrating multiple components (Ozger et al., 2011). These advanced approaches typically combine signal preprocessing techniques (such as wavelet analysis) to extract non-stationary patterns with metaheuristic optimization algorithms for fine-tuning hyperparameters (Rezaiy and Shabri, 2023; Kikon et al., 2023; Alkan, 2023). Numerous studies confirm that these hybrid approaches, by significantly reducing prediction error, decisively outperform standard models (Malik et al., 2021; Komasi et al., 2018) and demonstrate high adaptability to contrasting climates (Moghaddasi et al., 2025).

Despite the effectiveness of autoregressive models, whether standalone (Prodhan et al., 2022) or hybrid (Almikaeel et al., 2022), they inherently overlook the physical dynamics governing the climate system. Drought variability is not a purely random process; rather, it is largely controlled by large-scale ocean-atmosphere interactions known as teleconnection patterns (Wang et al., 2024). These patterns, which are natural fluctuations in the general atmospheric circulation, are the primary source of climate predictability on seasonal to multi-decadal scales (Santos et al., 2024; Craig and Allan, 2022). Patterns such as the El Niño-Southern Oscillation (ENSO), the North Atlantic Oscillation (NAO), and the Indian Ocean Dipole (IOD), by modulating sea surface temperature and atmospheric pressure patterns, exert significant impacts on regional precipitation regimes, including in Iran (Pourasghar and Babaeian, 2023; Cai et al., 2024; Yun et al., 2021). Therefore, understanding and modeling these physical relationships holds the potential to improve the accuracy and extend the horizon of drought forecasts (Yang and Xing, 2022; Liu et al., 2023; Xing et al., 2024).

However, previous research in Iran examining the impact of teleconnections on drought, despite its value, has faced multidimensional limitations in data and methodology that have constrained their practical application. Firstly, the range of predictors has been limited, with most studies confining their analysis to a few well-known patterns (mainly ENSO and

NAO) (Amini et al., 2020; Mahmoudi et al., 2020b). This is while the complex climate of Iran is influenced by a wide network of teleconnections from all three ocean basins (Helali et al., 2022; Ghamghami and Bazrafshan, 2021). Secondly, the geographical scale has often been local (Mokhtar et al., 2021; Tuğrul and Hinis, 2025), which challenges the generalizability of the results. More importantly, many of these studies have relied on linear statistical methods (Seibert et al., 2017) which are incapable of fully explaining the complex, nonlinear, spatially non-stationary, and temporally asymmetric interactions between teleconnections and drought (Dong et al., 2024; Tozer and Kiem, 2017) and may lead to misleading conclusions (Razmjo et al., 2020). The emergence of eXplainable Artificial Intelligence (XAI) tools like SHAP has provided an unprecedented opportunity to open the black box of ML/DL models and gain a deeper understanding of these nonlinear relationships (Dikshit and Pradhan, 2021a, Dikshit and Pradhan, 2021b). In a recent study, Mahmoudi et al. (2025) utilized such an XAI framework to successfully decode the specific teleconnection drivers of Iran's drought, identifying the critical role of nonlinear patterns like the Zonal Wind at 200 hPa (ZWNDz200) and the Atlantic Meridional Mode (AMM). However, while that study focused on the diagnostic interpretation of physical mechanisms, it did not address the prognostic challenge of systematically comparing the autoregressive versus teleconnection-driven paradigms for operational forecasting.

This analysis reveals a prominent and multi-faceted research gap. As noted in the most recent systematic review articles (Oyarzabal et al., 2025; Zellou et al., 2023; Alawsi et al., 2022), a systematic, large-scale comparison to determine whether models based on the teleconnection-driven paradigm can outperform or synergistically enhance conventional autoregressive models has not been conducted. This is a fundamental question in the science of drought forecasting, the answer to which could determine the direction of future research and the design of operational systems.

Aiming to fill this critical gap, this study develops a comprehensive and comparative framework for the short-term forecasting (1-, 2-, and 3-month lead times) of meteorological drought (SPI) across Iran. Leveraging precipitation data from a dense network of 96 synoptic stations over a 30-year period and a diverse suite of nine machine and deep learning models, we critically evaluate three distinct predictor scenarios to answer our fundamental objectives:

- The autoregressive scenario (S1): which uses historical SPI values as a proxy for the temporal memory paradigm.
- The teleconnection-driven scenario (S2): which utilizes a comprehensive set of 19 global climate indices as a proxy for the climate driver-based paradigm.
- The hybrid scenario (S3): which tests the synergistic potential of the two paradigms by integrating both datasets.

By conducting this systematic comparison, we seek to answer the fundamental question: Can our knowledge of the current state of the oceans and atmosphere, either alone or in combination with past information, provide more accurate forecasts of future drought in Iran, and how is this superiority distributed geographically?

The novelty of this research lies in its scale, scope, and systematic approach. By moving beyond local analyses and limited predictor sets, this study provides the first nationwide, multi-model, and multi-paradigm assessment of drought predictability in Iran, directly challenging competing forecasting philosophies in a vast experimental arena. The findings of this research are expected to provide deep physical insights and practical guidance for the development of a new generation of drought

forecasting systems that are not only more accurate but also have a more robust scientific foundation, thereby enhancing national resilience to climate variability and change.

2 Study Area

Iran, with an area of approximately 1.65 million square kilometers in Southwest Asia, is situated between latitudes 25° and 40° N and longitudes 44° and 64° E (Fig. 1a). This geographical location places Iran at the heart of the Northern Hemisphere's arid and semi-arid belt and at the crossroads of large-scale atmospheric circulation patterns. However, Iran's unique characteristic, which makes it an ideal natural laboratory for this research, is its highly complex topography and extreme climatic contrasts. The country's physical landscape is dominated by two massive mountain ranges: the Alborz mountain range in the north, which creates a barrier against the intrusion of moisture from the Caspian Sea, and the Zagros mountain range, which extends from northwest to southeast and acts as a primary obstacle to Mediterranean and westerly rain-bearing systems. These two mountain ranges, alongside the two vast hyper-arid deserts of Dasht-e Kavir and Dasht-e Lut in the central plateau, create exceptionally steep climatic gradients over relatively short geographical distances (Heydarizad et al., 2018; Saemian et al., 2022).

This prominent configuration has transformed Iran into a climatic mosaic where almost all moisture regimes, from hyper-arid to hyper-humid, are found. Analysis of climatic data indicates that approximately 97% of the country's area falls within the realm of hyper-arid, arid, and semi-arid climates. As shown in Figure 1b, based on the UNESCO (1979) classification, vast areas of the central, eastern, and southern parts of the country experience hyper-arid conditions. In stark contrast to these conditions, the southern coastal strip of the Caspian Sea and the northern slopes of the Alborz experience humid and even hyper-humid climatic regimes. This extreme climatic diversity provides a unique opportunity to test the hypothesis of whether different forecasting paradigms (autoregressive, teleconnection-driven, and hybrid) exhibit variable performance across different hydroclimatic regimes.

The spatial distribution of mean annual precipitation (Fig. 1c) quantitatively illustrates these contrasts. While the country's average total precipitation is estimated to be around 225 mm, this value conceals a massive spatial variability: from less than 100 mm in the central deserts to over 1500 and even 2000 mm in parts of the Caspian coasts (Katiraie-Boroujerdy et al., 2020). This precipitation pattern is strongly controlled by the interaction between large-scale atmospheric circulation and local topography. The orographic lift of atmospheric systems over the western slopes of the Zagros and the northern slopes of the Alborz leads to significant precipitation in these regions, while the 'rain shadow' effect causes severe aridity in the central Iranian plateau (Alijani, 1997). This complex interaction makes Iran a region where simple and simultaneous relationships between variables are expected to break down, and nonlinear and time-lagged patterns (which form the core of teleconnection analysis) play a dominant role. Therefore, a comparative assessment of models that rely solely on the system's internal memory (autoregressive) versus models that leverage teleconnection-based drivers, within such a complex setting, can lead to fundamental and generalizable findings for the science of drought forecasting.

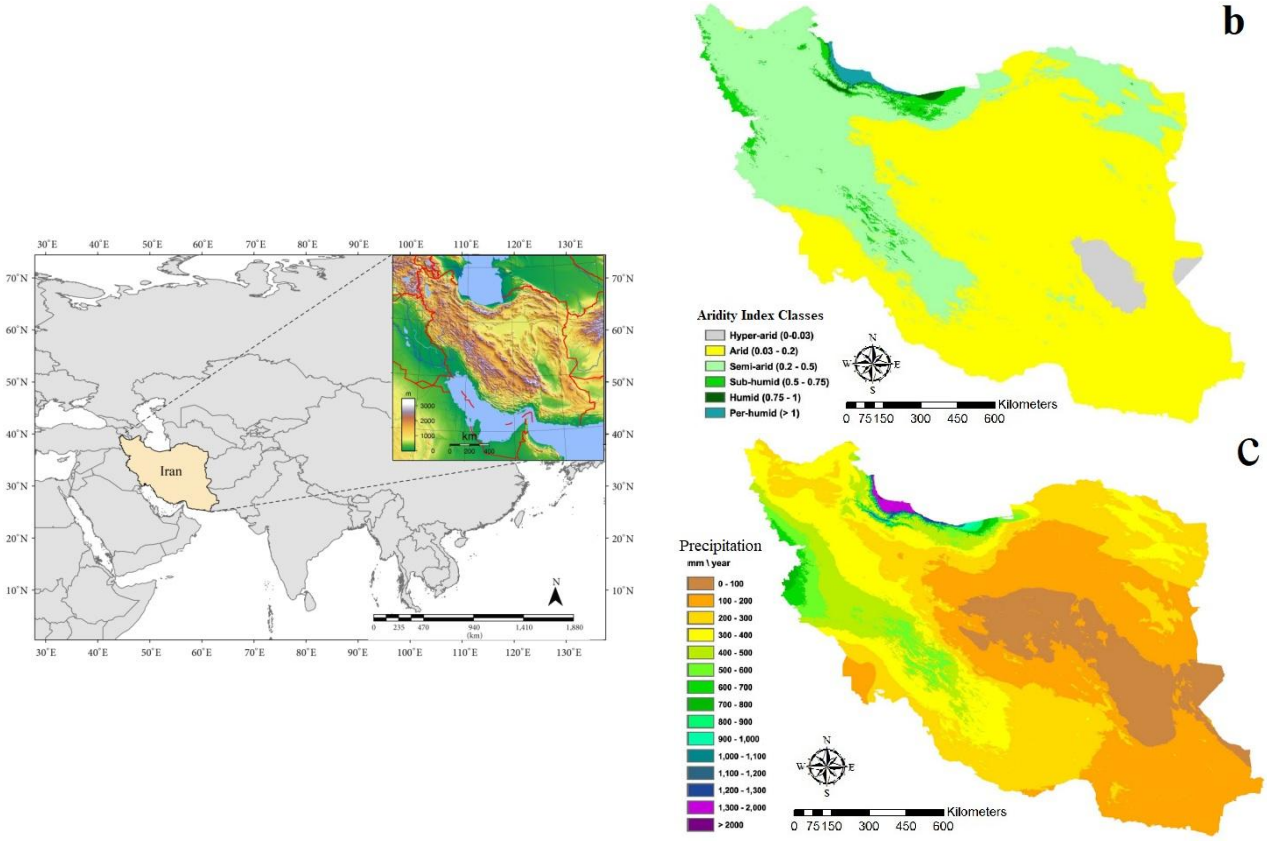


Figure 1: (a) Geographical location of Iran in Southwest Asia, showing the region's topography; (b) Climatic zonation of Iran based on the UNESCO aridity index classes, calculated from the ratio of annual precipitation (P) to potential evapotranspiration (PET) (De Pauw et al., 2018); (c) Spatial distribution of mean total annual precipitation (in mm) across Iran (De Pauw et al., 2018).

3 Data and Methodological Framework

This research is built upon a multi-layered methodological framework designed to comparatively assess competing paradigms in drought forecasting (autoregressive versus teleconnection-driven) within a complex geographical context. This process comprises three main components: (1) the collection and quality control of two fundamental datasets: station-based precipitation data and teleconnection indices; (2) the calculation of the drought index as the target variable; and (3) the development and evaluation of a diverse suite of machine learning models under different input scenarios.

3.1 Research Datasets: From Local Observations to Global Drivers

To address the research questions, two distinct yet complementary datasets were collected and processed for a 30-year period (January 1993 to December 2022):

3.1.1 Station Precipitation Data and Quality Control

The foundation of the drought analysis in this study is monthly total precipitation data spanning a 30-year period from 1993 to 2022, obtained from the observational network of the Iran Meteorological Organization (IRIMO) for 96 selected synoptic stations (Table 1). While the World Meteorological Organization (WMO) officially recommends the 1991–2020 period for computing standard climate normals, the 1993–2022 timeframe was deliberately selected for this research. This specific 30-year window ensures the maximum availability of continuous, gap-free synoptic records across all 96 stations, while crucially incorporating the most recent climatic extremes (up to 2022) necessary for training robust machine learning algorithms. Furthermore, spanning three full decades (360 months), this period is statistically representative of the region's long-term climate variability, capturing essential multi-year drought and wet cycles. These stations (Fig. 2) were chosen to provide optimal spatial coverage of Iran's diverse climatic regions, from the humid northern coasts to the hyper-arid central and southern deserts. Recognizing that the quality of input data determines the validity of modeling results, a rigorous Quality Assurance/Quality Control (QA/QC) protocol was applied to all time series. This protocol involved two essential steps:

a) Data Completeness Check: All 96 time series for the 360-month study period were inspected for missing data, and it was confirmed that there were no gaps in the records.

b) Homogeneity Testing: To ensure that the variability present in the precipitation time series is solely due to climatic processes and not non-natural factors (such as station relocation, instrument changes, or changes in surrounding land use), the homogeneity of each time series was assessed using a suite of four standard statistical tests: the Standard Normal Homogeneity Test (SNHT), the Buishand range test, the Pettitt test, and the Von Neumann ratio test. This set of tests, recommended by Wijngaard et al. (2003), has a high capability to detect inhomogeneities caused by both abrupt shifts and gradual trends. The results confirmed that all utilized time series were homogeneous and suitable for subsequent analyses.

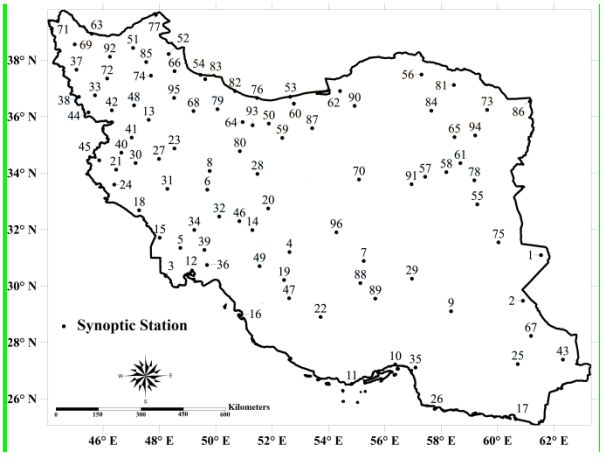


Figure 2. Spatial distribution of the 96 selected synoptic stations across Iran. The names corresponding to the station codes are listed in Table 1.

Table 1. List of the 96 synoptic stations used in this study and their corresponding codes.

Code	Name	Code	Name	Code	Name
1	Zabol	33	Mahabad	65	Kashmar
2	Zahedan	34	Masjed Soleyman	66	Khalkhal
3	Abadan	35	Minab	67	Khash
4	Abadeh	36	Omidiyeh	68	Khorramdareh
5	Ahwaz	37	Oroomieh	69	Khoy
6	Aligudarz	38	Piranshahr	70	Khoor Biabanak
7	Anar	39	Ramhormoz	71	Makoo
8	Arak	40	Ravansar	72	Maragheh
9	Bam	41	Sanandaj	73	Mashhad
10	Bandar Abass	42	Saqez	74	Mianeh
11	Bandar Lengeh	43	Saravan	75	Nehbandan
12	Bandar Mahshahr	44	Sardasht	76	Noushahr
13	Bijar	45	Sar Pol Zohab	77	Parsabad
14	Boroojen	46	Shahre Kord	78	Ghaen
15	Bostan	47	Shiraz	79	Ghazvin
16	Bushehr	48	Takab	80	Ghom
17	Chabahar	49	Yasuj	81	Ghoochan
18	Dehloran	50	Abali	82	Ramsar
19	Doroudzan	51	Ahar	83	Rasht
20	Esfahan	52	Ardabil	84	Sabzevar
21	Eslamabad Gharb	53	Babolsar	85	Sarab
22	Fasa	54	Anzali	86	Sarakhs
23	Hamedan	55	Birjand	87	Semnan
24	Ilam	56	Bojnurd	88	Shahr Babak
25	Iranshahr	57	Boshrooyeh	89	Sirjan
26	Jask	58	Ferdous	90	Shahroud
27	Kangavar	59	Garmsar	91	Tabass
28	Kashan	60	Gharakhil	92	Tabriz
29	Kerman	61	Gonabad	93	Tehran
30	Kermanshah	62	Gorgan	94	Torbate Heydariyeh
31	Khorramabad	63	Jolfa	95	Zanjan
32	Koohrang	64	Karaj	96	Yazd

3.1.2 Large-Scale Climate Indices (Teleconnections)

To evaluate the forecasting paradigm based on physical drivers, a comprehensive set of 19 global teleconnection indices, which quantify natural oscillations in the ocean-atmosphere system, were selected. These indices, extracted as monthly time series from the reputable databases of the National Centers for Environmental Prediction (NCEP) and the Climate Prediction

Center (CPC) at the National Oceanic and Atmospheric Administration (NOAA), represent key dynamics from the three main oceanic basins: the Pacific, Atlantic, and Indian Oceans (Table 2).

The specific rationale for selecting these 19 indices stems from an extensive review of previous climatological literature focused on the Middle East and Iran. Out of numerous global atmospheric and oceanic patterns, these 19 indices were chosen because prior empirical and dynamical studies have repeatedly confirmed their significant influence on the region's precipitation variability and drought occurrences across various temporal scales. Consequently, The selection of this extensive set, going beyond commonly used indices (such as ENSO and NAO), was intended to cover the complex network of potential climatic influences on Iran. These indices are not merely statistical variables; rather, each represents a distinct physical state in the general circulation of the atmosphere and ocean. They modulate global-scale pressure patterns and moisture pathways through atmospheric Rossby waves, thereby creating predictability on seasonal to multi-year timescales (Tozer and Kiem, 2017). The inclusion of indices from the Atlantic Ocean (e.g., AMO, TNA, TSA) and other less-studied patterns (e.g., WP, EA-WR) allows for a more comprehensive identification of the sources of drought variability in Iran.

While incorporating 19 indices introduces the theoretical risk of feeding non-predictive "noise" into the algorithms, we consciously opted not to apply a rigid *a priori* feature selection step (e.g., statistical filtering or dimensionality reduction). This decision was made to avoid prematurely discarding indices that might exhibit complex, highly non-linear, or geographically localized interactions with Iran's diverse climate. Instead, the management of this high-dimensional input space was entrusted to the intrinsic architectures of the selected advanced algorithms. Specifically, tree-based ensembles (such as XGBoost, RF, and CatBoost) inherently perform embedded feature selection during the training phase by splitting only on variables that maximize information gain, effectively ignoring noisy predictors. Similarly, deep learning architectures like LSTM utilize internal gating mechanisms to learn which signals to retain or suppress over time. Coupled with mathematical regularization techniques (such as L2 penalties in BRR and gradient boosting models), these embedded mechanisms naturally penalize uninformative drivers, thereby ensuring model parsimony, mitigating the risk of overfitting, and isolating the true physical signals from the noise.

Table 2: Specifications of the set of 19 global teleconnection indices used in this research as physical predictors. The indices are categorized based on their primary oceanic basin and associated physical phenomenon.

Ocean Basin	Key Phenomenon / Region	Full Name	Abbr.	Primary Data Source	Key Reference
Pacific Ocean	El Niño-Southern Oscillation	Southern Oscillation Index	**SOI**	CPC	Ropelewski & Jones (1987)
	SST Anomaly - Eastern Pacific	Niño 1+2	**Nino1+2**	CPC/NCEP	Reynolds et al. (2002)
	SST Anomaly - East-Central Pacific	Niño 3	**Nino3**	CPC/NCEP	Rayner et al. (2003)
	SST Anomaly - Central Pacific	Niño 3.4	**Nino3.4**	CPC/NCEP	Rayner et al. (2003)
	SST Anomaly - Western	Niño 4	**Nino4**	CPC/NCEP	Rayner et al. (2003)

	Pacific Oceanic Niño Index	Oceanic Niño Index	**ONI**	CPC	Barnston (1997)
	Inter-tropical SST Gradient	Trans-Niño Index	**TNI**	CPC	Trenberth & Stepaniak (2001)
	Atmospheric Pressure Pattern	Pacific/North American Pattern	**PNA**	CPC	Wallace & Gutzler (1981)
	Atmospheric Pressure Pattern	West Pacific Pattern	**WP**	CPC	Wallace & Gutzler (1981)
	Multidecadal Oscillation	Pacific Decadal Oscillation	**PDO**	NCEI	Newman et al. (2016)
	Interdecadal Oscillation	Interdecadal Pacific Oscillation (Tripole Index)	**TPI**	NCEI	Henley et al. (2015)
Atlantic Ocean	Azores-Iceland Pressure Oscillation	North Atlantic Oscillation	**NAO**	CPC	Barnston & Livezey (1987)
	Atmospheric Pressure Pattern	Eastern Atlantic/Western Russia Pattern	**EA-WR**	CPC	Barnston & Livezey (1987)
	SST Anomaly - Tropical North	Tropical Northern Atlantic Index	**TNA**	NCEP	Enfield et al. (1999)
	SST Anomaly - Tropical South	Tropical Southern Atlantic Index	**TSA**	NCEP	Enfield et al. (1999)
Indian Ocean	SST Dipole	Indian Ocean Dipole	**IOD**	NCEP	Saji & Yamagata (2003)
Global / Hemispheric	Western Hemisphere Warm Pool	Western Hemisphere Warm Pool	**WHWP**	NCEP	Wang & Enfield (2001)
	Polar Oscillation	Arctic Oscillation	**AO**	CPC	Higgins et al. (2000)
	Stratospheric Oscillation	Quasi-Biennial Oscillation	**QBO**	NCEP	Baldwin et al. (2001)

Abbreviations: *CPC: Climate Prediction Center; NCEP: National Centers for Environmental Prediction; NCEI: National Centers for Environmental Information; SST: Sea Surface Temperature.*

3.2. Calculation of the Drought Index as the Target Variable

The target variable in this research is the Standardized Precipitation Index (SPI), which is recognized by the World Meteorological Organization (WMO) as a global standard for quantifying meteorological drought (Lorenzo et al., 2024). This index, introduced by McKee et al. (1993), measures the deviation of cumulative precipitation over a specific period (a k-month timescale) from its long-term mean, after normalization to a standard distribution. The main advantage of the SPI lies in its independence from the climatic conditions of a region and the comparability of its results across different temporal and spatial scales (Tsesmelis et al., 2023). The process of calculating the SPI for each station includes the following steps (Ghaemi et al., 2024a):

- Creating cumulative precipitation time series for the desired timescales (e.g., 1-month, 3-month, 6-month, etc.).
- Fitting a probability distribution function, typically the two-parameter gamma function, to each cumulative time series.

- Transforming the cumulative probability obtained from the gamma function into a standard normal variable with a mean of zero and a variance of one. The final value of this variable is the SPI.

This process is performed using standard numerical approximations for the transformation of the gamma distribution to the standard normal distribution (Rezaiy and Shabri, 2025). The resulting SPI values classify the intensity of wet (positive values) and dry (negative values) conditions (Table 3). In this research, the SPI is used as a proxy for drought status and serves as the dependent variable in all forecasting models.

Table 3: Classification of climatic conditions based on the Standardized Precipitation Index (SPI) values (adapted from McKee et al., 1993).

SPI limits	Classes
$SPI \geq 2.00$	Extremely Wet
$1.50 \leq SPI \leq 1.99$	Severely Wet
$1.00 \leq SPI \leq 1.49$	Moderately Wet
$-0.99 \leq SPI \leq 0.99$	Near Normal
$-1.00 \leq SPI \leq -1.49$	Moderate Drought
$-1.50 \leq SPI \leq -1.99$	Severe Drought
$SPI \leq -2.00$	Extreme Drought

3.3. Modeling Paradigms and Experimental Framework

The core of this research is the design and implementation of a comprehensive comparative experiment to evaluate competing paradigms in drought forecasting. This framework is built upon three distinct input scenarios and a diverse set of forecasting algorithms (Figure 3).

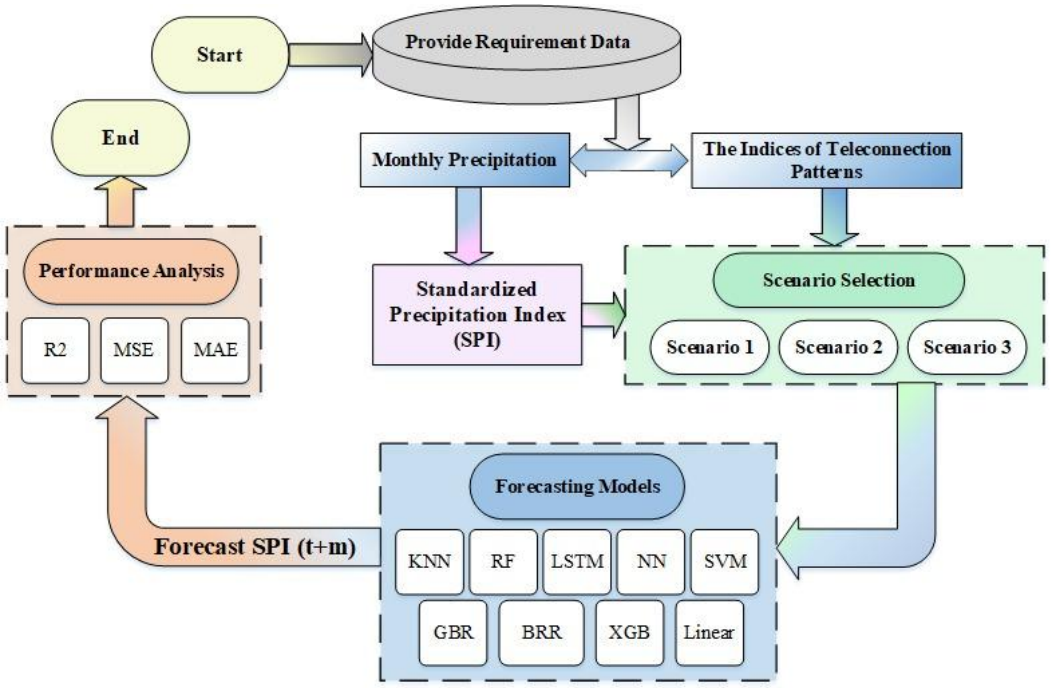


Figure 3: Methodological framework of the research for short-term drought forecasting in Iran. This flowchart illustrates the step-by-step process, from the ingestion of raw data (precipitation and teleconnection indices), calculation of the SPI, design of input scenarios, training of a suite of machine learning models, to the final performance evaluation and selection of the optimal model.

3.3.1 Designing Input Scenarios: From Large-Scale Climate Drivers to Temporal Memory

To address the fundamental research questions and identify the primary sources of predictability, three main families of scenarios were designed for the models' input structure, each representing a different forecasting philosophy.

- **Scenario 1 (S1): Based on Large-Scale Climate Drivers:** This scenario is founded on the hypothesis that the primary source of drought predictability on a monthly scale lies not in local history, but in the current state of the global climate system. In this scenario, the model inputs consist exclusively of the values of the 19 teleconnection indices with time lags of 1 to 3 months ($Tele_{t-1}, Tele_{t-2}, Tele_{t-3}$).
- **Scenario 2 (S2): Autoregressive with Increasing Temporal Memory:** This family of scenarios represents the classic time-series forecasting paradigm, where information is extracted solely from the history of the variable itself (SPI). To test the importance of the system's memory length, this scenario was divided into four subsets, as detailed in Table 4. This design allows us to identify the optimal depth of temporal memory required for forecasting in a data-driven manner.

Table 4: Structure of the autoregressive sub-scenarios (S2 family) for evaluating the impact of temporal memory length.

Scenario Name	Description	Input Variables (previous SPI time lags)

S2-1	Short-term memory	$SPI_{t-1}, SPI_{t-2}, SPI_{t-3}$
S2-2	Medium-term memory	$(SPI_{t-1}, \dots, SPI_{t-6})$
S2-3	Long-term memory	$(SPI_{t-1}, \dots, SPI_{t-9})$
S2-4	Very long-term memory	$(SPI_{t-1}, \dots, SPI_{t-12})$

- **Scenario 3 (S3):** Hybrid: This family of scenarios seeks to test the synergistic potential of the two preceding paradigms. This approach investigates whether adding knowledge of the global climate state (large-scale climate drivers) can improve predictions based on temporal memory (autoregressive). The structure of these scenarios is presented in Table 5.

Table 5: Structure of the hybrid sub-scenarios (S3 family) for evaluating the synergistic effect.

Scenario Name	Description	Input Variables
S3-1	Integration of physical drivers and short-term memory	Combination of inputs from scenarios S1 and S2-1
S3-2	Integration of physical drivers and medium-term memory	Combination of inputs from scenarios S1 and S2-2

The specific time lags assigned to each scenario were determined based on a combination of physical climatological memory, preliminary correlation analyses, and machine learning dimensionality constraints. For teleconnections (S1), a 1-to-3-month lag was chosen as it effectively captures the seasonal atmospheric response to oceanic anomalies (e.g., the atmospheric bridge conveying ENSO signals) before the signal decays into noise, a threshold widely supported by existing climatological literature. In the autoregressive models (S2), an extended range (up to 12 months) was evaluated to fully explore the persistence of the SPI. However, our preliminary exploratory experiments revealed a clear saturation point at 12 months. Extending the temporal memory beyond an annual cycle (e.g., 18 or 24 months) provided no additional predictive skill. Instead, it introduced irrelevant noise and led to model degradation and overfitting, confirming that adding more historical data beyond this point no longer helps. Furthermore, in the hybrid models (S3), preliminary experiments revealed that combining the 57 teleconnection inputs (19 indices × 3 lags) with extended autoregressive lags (e.g., 9 or 12 months) caused a massive expansion of the feature space. This not only increased computational complexity exponentially but also introduced information redundancy, raising the risk of overfitting without yielding any added predictive value. Thus, limiting the hybrid inputs to 3- and 6-month SPI lags provided the optimal balance between capturing temporal memory and avoiding the curse of dimensionality.

For all scenarios, the models' objective is to predict SPI values for forecasting horizons of 1, 2, and 3 months ahead ($\mathrm{SPI}_{t+1}, \mathrm{SPI}_{t+2}, \mathrm{SPI}_{t+3}$).

3.3.2. Suite of Forecasting Algorithms

To ensure that the comparison of scenarios is not dependent on a specific model and that the results are sufficiently robust, a diverse suite of nine machine learning and deep learning algorithms was employed. This collection covers a wide spectrum from simple linear models to highly complex nonlinear structures.

Baseline Models:

- **Multiple Linear Regression (MLR):** Utilized as a simple linear model to establish a baseline for gauging the performance of more complex models (Manriquez-Padilla et al., 2023).
- **K-Nearest Neighbors (KNN):** A non-parametric, instance-based algorithm that is capable of capturing local nonlinear relationships by identifying similar samples in the feature space (Zhou et al., 2024).

Classic Machine Learning Models:

- **Artificial Neural Network (ANN):** A Multilayer Perceptron (MLP) trained with the Levenberg-Marquardt optimization algorithm, recognized as a universal approximator for modeling complex nonlinear relationships (Mohammadi, 2023).
- **Least Squares Support Vector Machine (LS-SVM):** A computationally efficient version of SVM that trains the model by solving a system of linear equations instead of a quadratic programming problem. Using a Radial Basis Function (RBF) kernel, this model can map complex relationships in a high-dimensional feature space (Saravanan et al., 2023; Zhang et al., 2023).

Tree-Based Ensemble Models:

- **Random Forest (RF):** A powerful ensemble model that reduces variance and prevents overfitting by aggregating the predictions of a large number of independent decision trees, each built on bootstrap subsamples of the data (Islam et al., 2023; Ghaemi et al., 2024b).
- **Gradient Boosting Regression (GBR):** A boosting algorithm that sequentially and additively builds models (typically decision trees), such that each new model attempts to correct the errors of the previous one (Zhang et al., 2021; Wu et al., 2024).
- **Extreme Gradient Boosting (XGBoost):** A highly optimized, scalable, and engineered implementation of the gradient boosting framework. XGBoost has gained global renown in data science competitions and practical applications for its high speed and accuracy, achieved through techniques like parallelization, regularization to prevent overfitting, and handling of missing values (Alipour, 2025; Niazkar et al., 2024).

Probabilistic Regression

- **Bayesian Ridge Regression (BRR):** A probabilistic regression model that applies a Bayesian approach to linear regression. Instead of finding a single value for the regression coefficients, it estimates a probability distribution for them. This feature, by imposing a prior on the parameters, naturally regularizes the model and makes it robust against overfitting, particularly when the number of features is large (Reddy et al., 2024; Snaan and Shukur, 2025).

Deep Learning Model:

- **Long Short-Term Memory (LSTM):** An advanced type of Recurrent Neural Network (RNN) specifically designed to learn long-term dependencies in sequential data and time series. Its gated memory architecture (input,

output, and forget gates) allows it to selectively retain or discard relevant information over long sequences, thereby overcoming the vanishing gradient problem of standard RNNs (Li et al., 2023; Tang et al., 2025).

To facilitate a clear understanding of the diverse forecasting algorithms employed and to provide a quick reference for the subsequent performance comparisons, the nine models are categorized and summarized in Table 6 based on their structural families and core characteristics.

Table 6: Summary and categorization of the machine learning and deep learning algorithms utilized for SPI forecasting.

Algorithm Category	Model Name	Abbreviation	Core Characteristic / Primary Advantage
Baseline Models	Multiple Linear Regression	MLR	Establishes a simple linear baseline for performance comparison.
	K-Nearest Neighbors	KNN	Non-parametric, instance-based learning; captures local nonlinearities.
Classic Machine Learning	Artificial Neural Network	ANN	Universal approximator (MLP) for modeling complex nonlinear relationships.
	Least Squares SVM	LS-SVM	Solves a system of linear equations (efficient); effective high-dimensional mapping.
Tree-Based Ensembles	Random Forest	RF	Bagging approach; reduces variance via independent bootstrap decision trees.
	Gradient Boosting Regression	GBR	Sequential boosting; each additive model attempts to correct predecessor errors.
	Extreme Gradient Boosting	XGBoost	Highly optimized boosting; features parallelization and strong regularization.
Probabilistic Regression	Bayesian Ridge Regression	BRR	Estimates probability distributions for coefficients; highly robust against overfitting.
Deep Learning	Long Short-Term Memory	LSTM	Gated recurrent architecture; inherently captures long-term temporal dependencies.

3.3.3. Performance Evaluation and Assessment Metrics

To quantitatively evaluate and compare the performance of the models during the testing phase, three standard statistical metrics were employed. These metrics cover various aspects of the prediction error:

- **Coefficient of Determination (R^2):** A metric that measures the proportion of the variance in the target variable explained by the model. Values approaching 1 indicate a better model fit.
- **Mean Squared Error (MSE):** This metric computes the average of the squared differences between observed and predicted values, thus giving greater weight to larger errors.
- **Mean Absolute Error (MAE):** Represents the average of the absolute prediction errors, disregarding their direction, which makes it easily interpretable.

The formulas for these metrics are as follows, respectively:

$$R^2 = \left[\frac{\sum_{i=1}^{N}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2} \cdot \sqrt{\sum_{i=1}^{N}(P_i - \bar{P})^2}} \right]^2 \quad (1)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2 \quad (2)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i| \quad (3)$$

where O_i are the observed values, P_i are the predicted values, $\overline{O}$ and $\overline{P}$ are the mean of the observed and predicted values, respectively, and N is the total number of observations in the test set (Mesgari et al., 2024).

4 Results

4.1 Predictability Based on Large-Scale Climate Drivers (Scenario S1)

This section evaluates the capability of large-scale climate drivers (teleconnection indices) to forecast drought conditions across increasing lead times. The comprehensive evaluation of nine machine learning algorithms across 96 stations revealed a distinct hierarchy in model performance. As illustrated in Figure 4, ensemble-based learning models, specifically Random Forest (RF), Gradient Boosting Regression (GBR), and Extreme Gradient Boosting (XGBoost), alongside the Long Short-Term Memory (LSTM) deep learning network, consistently outperformed simpler algorithms such as Linear Regression and K-Nearest Neighbors (KNN). Among these, the Random Forest (RF) model demonstrated the highest stability and accuracy across diverse climatic zones—evident in the high R^2 values for the one-month lead time (SPI_{t+1}) shown in the bar charts—and was therefore selected as the representative algorithm for the subsequent spatial analysis.

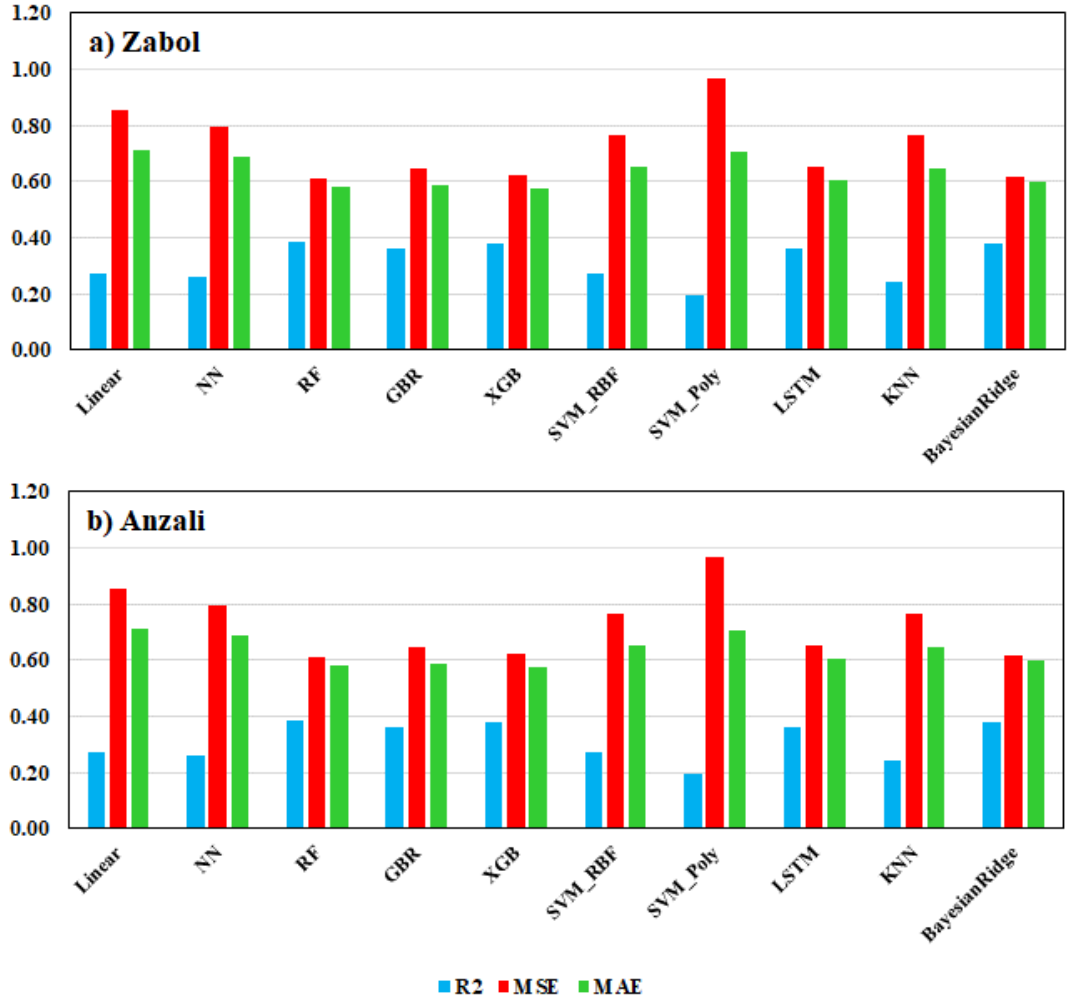


Figure 4: Performance comparison of different machine learning algorithms for one-month-ahead drought forecasting (SPI_{t+1}) under Scenario S1 across selected representative stations. The Random Forest (RF) model demonstrates superior accuracy and lower error rates compared to other models.

The geographical distribution of forecast skill highlights a strong dependency on regional hydroclimatology, particularly in the short-term horizon where teleconnection signals are most potent. Figure 5 illustrates the spatial performance of the RF model for the one-month lead time, representing the best-case scenario for teleconnection-based forecasting. The Coefficient of Determination (R^2) map (Fig. 5a) reveals a high-accuracy band stretching from the northwest across the Caspian Sea coasts and the western slopes of the Zagros Mountains. In these regions, the teleconnection signals are strong enough to explain up to 50% of the variance in the drought index. For instance, at the Anzali station (hyper-humid), where the climate is strongly modulated by large-scale circulation patterns, the accuracy reached an R^2 of 0.57. Conversely, the model's

explanatory power weakens considerably in the central, southern, and southeastern regions, where local convective processes often dominate over large-scale forcing.

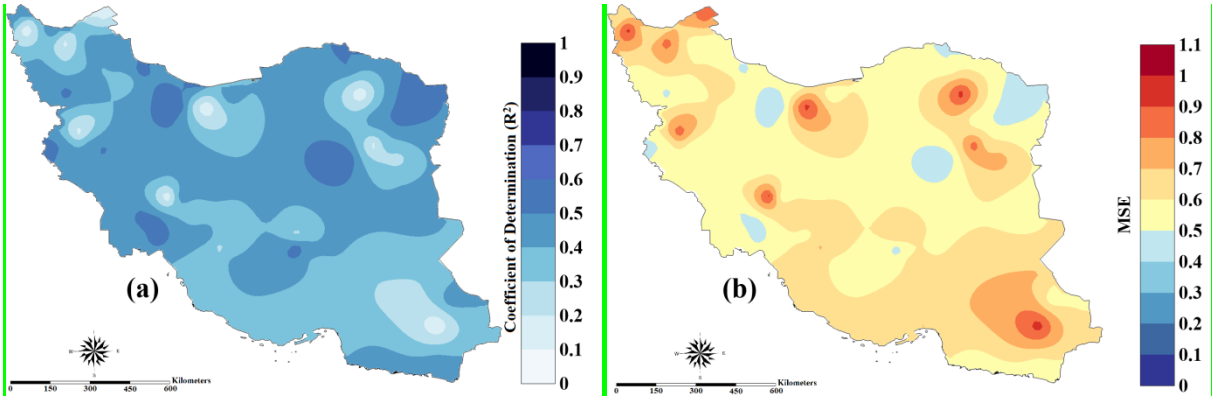


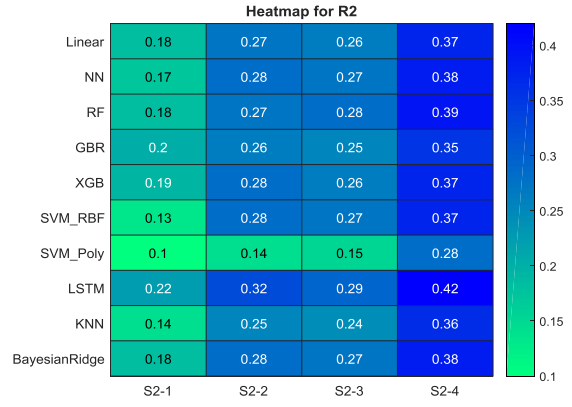
Figure 5: Spatial distribution of the performance of the best-selected model (Random Forest) for one-month-ahead drought forecasting (SPI_{t+1}) under Scenario S1. (a): Coefficient of Determination (R^2). (b): Mean Squared Error (MSE).

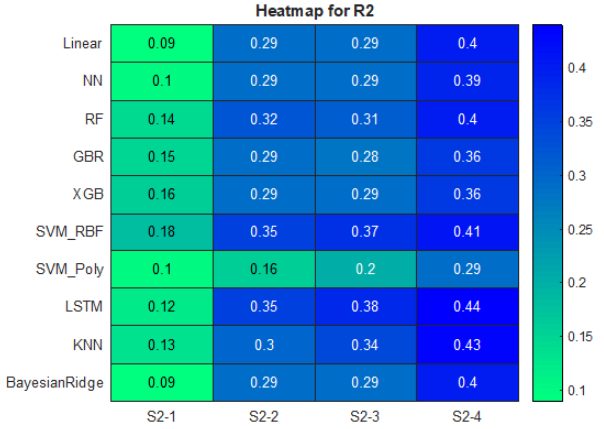
However, the robust initial signal observed at the one-month horizon progressively decays as the forecast lead time extends. While the performance remained relatively stable in humid regions—potentially due to the lagged response of the Caspian Sea climate to large-scale drivers—it deteriorated significantly in arid regions. As detailed in the Supplementary Material, extending the lead time to two and three months exposes the limitations of relying solely on teleconnections. By the three-month horizon (SPI_{t+3}), the spatial coherence observed in Figure 5 begins to fragment, and the area of high predictability shrinks to scattered patches (see Maps in Fig. S2). Furthermore, the competition among models tightens in these longer horizons; while RF maintains its lead, the Bayesian Ridge Regression (BRR) occasionally challenges its superiority (see Bar Charts in Fig. S1), suggesting that probabilistic approaches may better handle the heightened uncertainty of seasonal forecasting when the teleconnection signal fades.

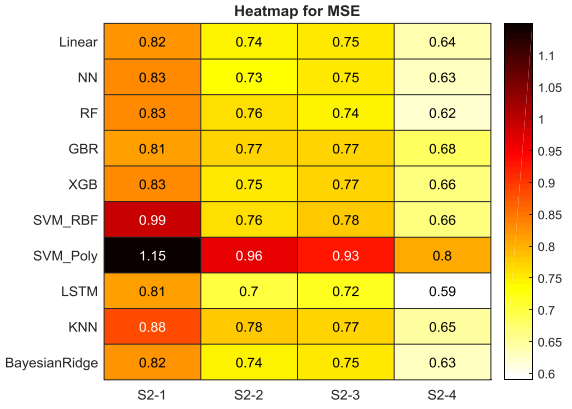
4.2. Impact of Temporal Memory in Autoregressive Modeling (Scenario S2)

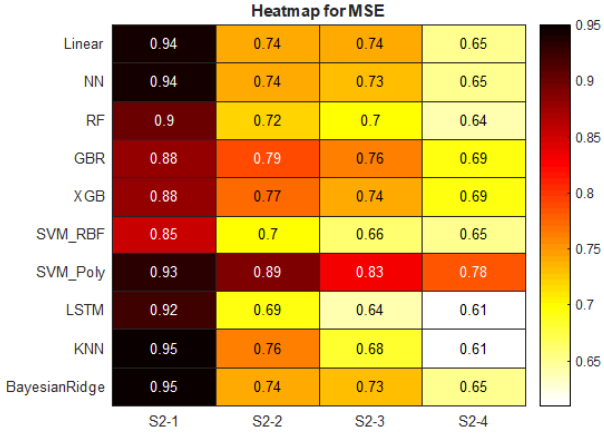
This section assesses the role of internal temporal memory—represented by antecedent SPI values with varying lag lengths (3, 6, 9, and 12 months)—in forecasting drought conditions across increasing lead times. The analysis of autoregressive patterns revealed a robust and consistent trend: increasing the length of the input temporal memory directly leads to improved performance for most machine learning models. Unlike the first scenario where ensemble tree models dominated, here the Long Short-Term Memory (LSTM) network, owing to its specialized architecture for learning long-term dependencies, emerged as the superior algorithm. As visually demonstrated in the heatmaps in Figure 6, the transition from a short-term 3-month memory (S2-1) to a long-term 12-month memory (S2-4) resulted in a substantial leap in accuracy. For

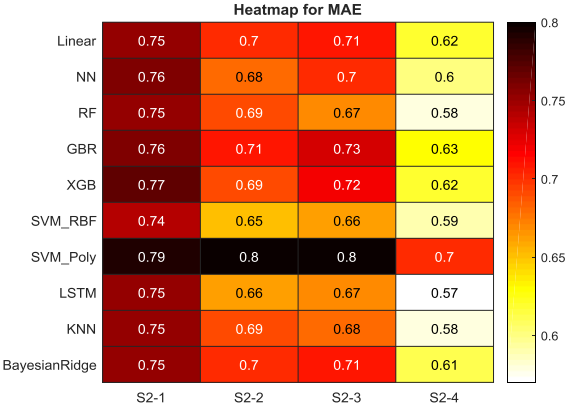
instance, at the Anzali station, the LSTM model's R^2 improved significantly, while errors (MSE and MAE) decreased, confirming that longer historical sequences provide richer contextual information about the persistence and evolution of drought dynamics.











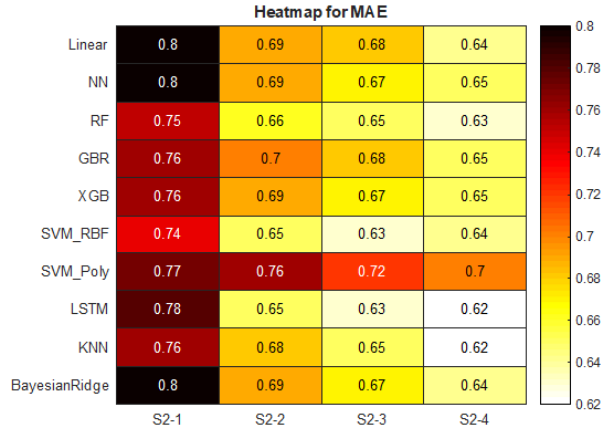


Figure 6: Heatmap of performance metrics (R^2, MSE, MAE) for one-month-ahead SPI prediction (SPI_{t+1}) at the Anzali (right columns) and Zabol (left columns) stations under four temporal memory scenarios (S2-1 to S2-4). The results clearly demonstrate that for most models, particularly LSTM, increasing the memory length from 3 months (S2-1) to 12 months (S2-4) leads to a substantial increase in R^2 and a reduction in error, underscoring the importance of longer-term historical data.

The spatial distribution of this optimal configuration (LSTM with 12-month memory) for the one-month lead time is presented in Figure 7. Similar to Scenario S1, the highest forecast accuracy ($R^2 > 0.50$) is concentrated in the western and northwestern regions, as well as parts of eastern Iran. This indicates that autoregressive signals are highly reliable in these areas for short-term forecasting. However, performance in the hyper-arid southeastern regions remains a challenge, likely due to the high frequency of zero-precipitation months which creates a time series with limited informative signal for the model to learn from.

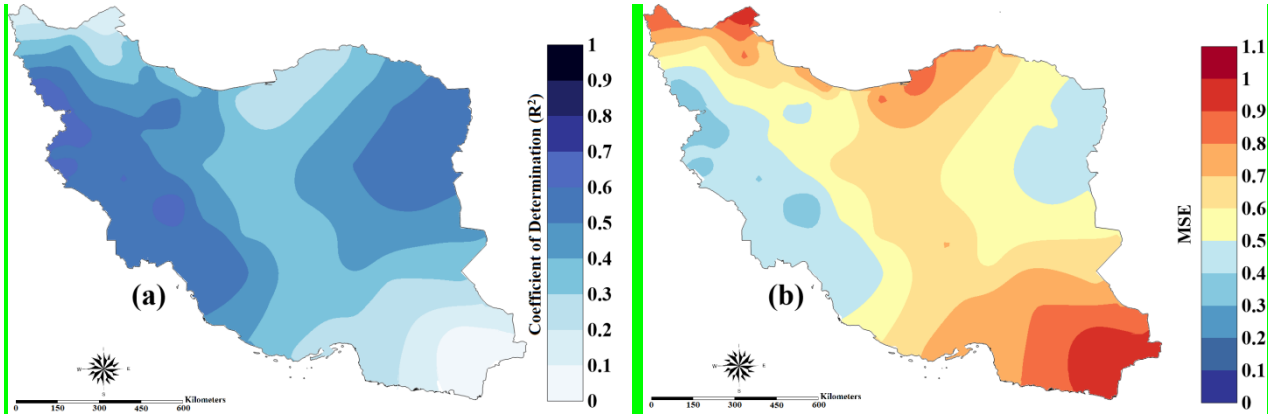


Figure 7: Spatial distribution of the performance of the best-selected model (LSTM) with a 12-month memory input (S2-4) for one-month-ahead drought forecasting (SPI_{t+1}). (a): Coefficient of Determination (R^2). (b): Mean Squared Error (MSE). The model demonstrates robust predictive capability in the west and northwest, while performance decreases in the hyper-arid southeastern regions.

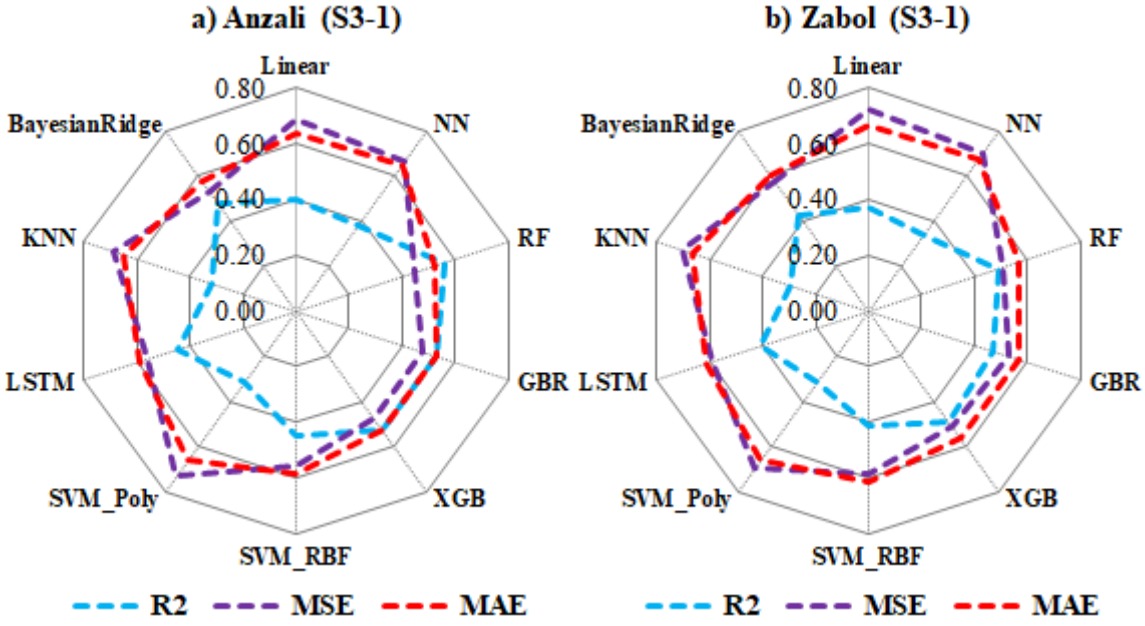
As the forecast horizon extends to two and three months (t+2 and t+3), the reliance on long-term memory becomes even more critical, yet the overall predictability inevitably declines. The comprehensive results provided in the Supplementary Material illustrate this decay. The heatmaps for longer lead times (Fig. S3 in Supplement) show that while the 12-month memory input (S2-4) continues to yield the best results compared to shorter memory inputs, the absolute values of R^2 drop. For example, at the Zabol station for the two-month forecast, the LSTM model in the S2-4 configuration manages to maintain an R^2 of 0.38, a significant recovery from the poor performance observed with shorter memory inputs, but still lower than the one-month horizon. Spatially, this decay manifests as a shrinkage of the high-accuracy zones. As shown in the supplementary maps (Fig. S4 in Supplement), by the third month, the extent of regions with $R^2 > 0.50$ is considerably

reduced compared to Figure 7, and the error values generally increase across the country. This suggests that while extending the input memory helps mitigate the loss of information over time, the persistence signal present in the SPI time series alone is insufficient to support reliable seasonal forecasting in many regions beyond the first month.

4.3. Synergy of Teleconnections and Temporal Memory: Evaluation of Hybrid Modeling (Scenario S3)

The third scenario investigates the synergistic potential of combining large-scale climate drivers with local temporal memory. Two hybrid input structures were evaluated: S3-1 (teleconnections + 3-month memory) and S3-2 (teleconnections + 6-month memory). In contrast to the memory-only scenario where LSTM excelled, the inclusion of climate indices saw the Random Forest (RF) model regain its position as the superior algorithm across all lead times, consistently achieving the highest R^2 and lowest error rates.

The analysis of model performance revealed a critical finding: the optimal balance between external climate signals and internal memory depth is strictly climate-dependent. As illustrated in the radar charts for the first month (Figure 8), there is a trade-off in complexity. At the hyper-humid Anzali station, the simpler structure (S3-1) outperformed the more complex one; increasing memory to 6 months (S3-2) actually degraded performance (lower R^2, higher MSE), suggesting an over-complexity that offers no gain. Conversely, at the hyper-arid Zabol station, the extended memory of S3-2 was beneficial, significantly improving all metrics. This suggests that arid regions, with their erratic rainfall, require a longer history combined with teleconnections to establish a reliable trend.



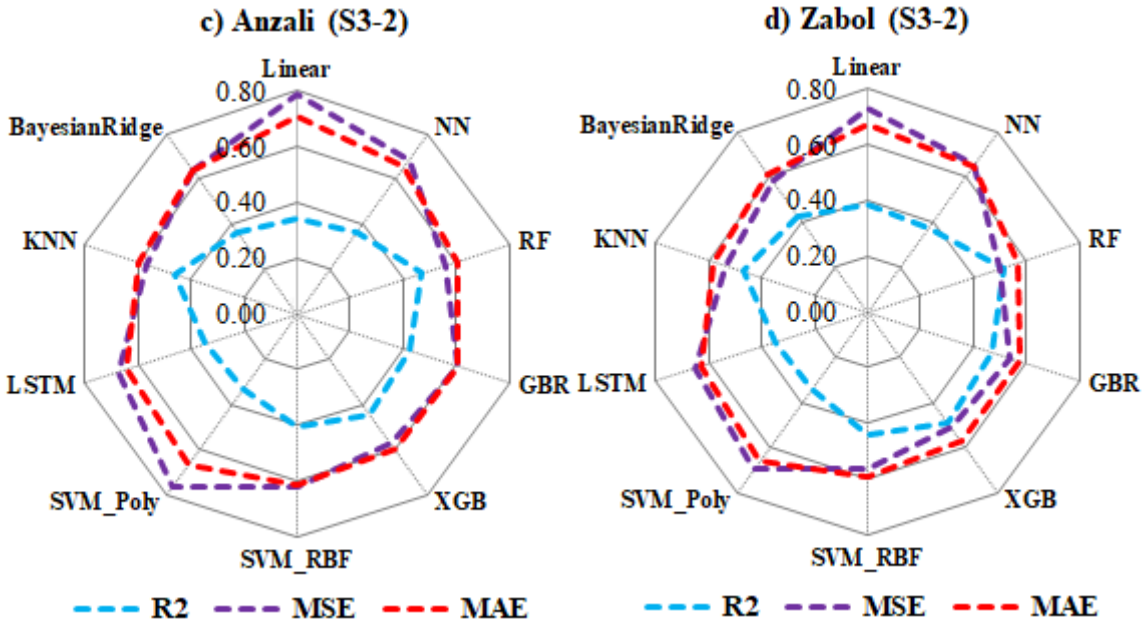
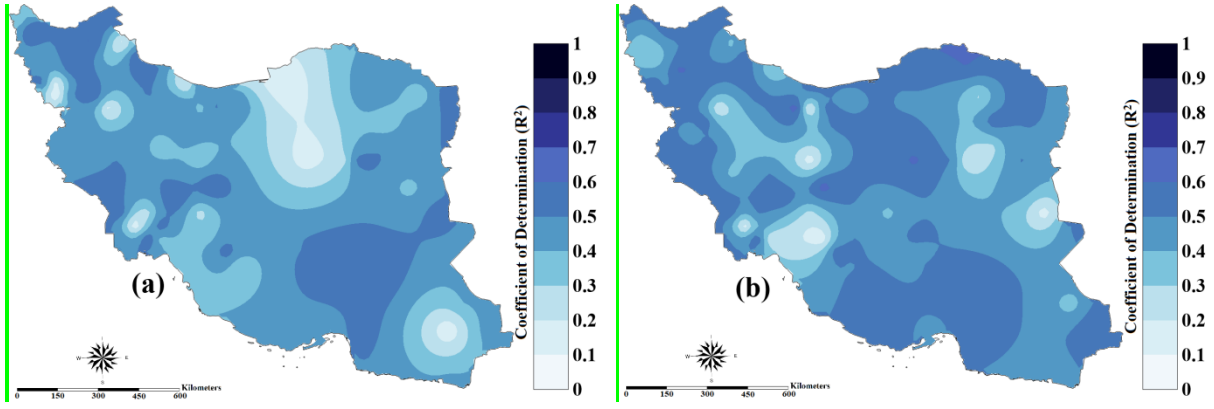


Figure 8: Radar charts comparing the performance (R^2, MSE, and MAE) of nine machine learning models for one-month-ahead forecasting (SPI_{t+1}) at the Anzali (humid) and Zabol (arid) stations. The comparison between S3-1 (3-month memory) and S3-2 (6-month memory) reveals a climatic dependency: while the shorter memory structure (S3-1) is sufficient and more accurate for the humid climate, the arid climate benefits from the extended memory of S3-2. Random Forest (RF) demonstrates superior performance in both cases.

This climatic dichotomy creates a distinct pattern of spatial complementarity across Iran, as mapped in Figure 9. The S3-2 structure (right column) is dominant in the northern and western regions, providing high accuracy ($R^2 = 0.6$). Meanwhile, the S3-1 structure (left column) acts as a necessary complement, offering significantly better predictions in the eastern and central regions where S3-2 struggles. This implies that a single hybrid configuration is not universally optimal; rather, an adaptive selection of memory length based on regional climate is required for maximum accuracy.



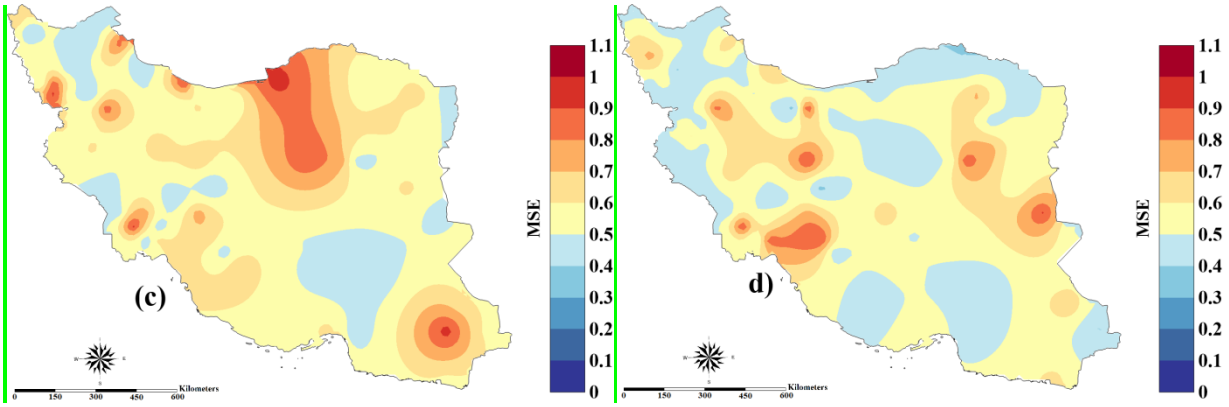


Figure 9: Spatial distribution of the performance (R^2 and MSE) of the superior Random Forest (RF) model for one-month-ahead prediction (SPI_{t+1}) under hybrid scenarios S3-1 and S3-2. The maps reveal a clear spatial complementarity: S3-2 (right) excels in the north and west, while S3-1 (left) provides necessary coverage and higher accuracy in the eastern and central regions.)

Extending the forecast horizon to two and three months (t+2 and t+3) confirms the stability of this spatial pattern, despite the expected general decay in accuracy. The results for longer lead times, provided in the Supplementary Material (Figs. S5 and S6 in Supplement), mirror the findings of the first month. Although the absolute R^2 values decrease—for instance, Anzali's R^2 drops but remains favorable for S3-1—the spatial preference remains consistent: the north/west continues to favor S3-2, while the east favors S3-1. This consistency across time horizons validates the hybrid approach, proving that even as predictive power fades with time, the structural preference of each climatic zone remains robust.

4.4. Final Comparative Analysis: Identifying Regionally Optimal Scenarios

In this final stage, the top-performing models from each of the three scenarios (S1, S2, and S3) were compared side-by-side to identify the definitive optimal input structure for each climatic region. Table 7 summarizes these results for the one-month lead time (SPI_{t+1}). A simultaneous analysis of performance metrics across all lead times reveals a crucial conclusion: no single input structure is universally optimal for Iran; rather, the choice depends heavily on regional hydroclimatology.

Dominance of Hybrid Models (S3) in the Central Plateau: In the arid and semi-arid regions covering the central plateau (e.g., Isfahan, Zabol, Bandar Abbas), the hybrid scenario (S3) demonstrated decisive superiority. For instance, in Isfahan, the S3-2 structure achieved an outstanding ($R^2 = 0.65$) with the lowest error (MSE=0.36), outperforming both the teleconnection-only and memory-only models. This confirms that in water-stressed regions with erratic rainfall, reliable forecasting requires the synergy of physical climate drivers and extended historical memory.

Superiority of Teleconnections (S1) in Coastal and Irregular Regimes: Conversely, in regions with complex atmospheric dynamics or highly irregular precipitation regimes, such as Chabahar and Anzali, the teleconnection-based scenario (S1) frequently emerged as the winner. Notably, in Chabahar, S1 provided a robust ($R^2 = 0.49$), whereas the memory-based

model struggled significantly. This indicates that local temporal memory is unreliable in these areas, and large-scale climate drivers provide a more stable predictive signal.

Role of Persistence (S2) in Specific Locales: In select stations like Ahvaz, the autoregressive scenario (S2) with long-term memory (12 months) outperformed others ($R^2 = 0.55$), highlighting the dominant role of internal drought persistence in this specific region.

Stability across Lead Times: Crucially, this regional preference for specific input structures remains remarkably stable as the forecast horizon extends to two and three months (t+2 and t+3). As detailed in the Supplementary Material (Tables S7 and S8), while the absolute accuracy naturally declines with time (e.g., Chabahar's best R^2 drops but S1 remains the superior choice vs. S2/S3), the winning scenario for each station rarely changes. This temporal consistency validates the proposed regionalization approach, suggesting that operational drought monitoring systems in Iran should adopt a dynamic framework where the input structure is switched based on the target region's climatic characteristics.

Table 7: Performance comparison of the best-selected models from each of the three main scenarios (S1, S2, S3) for one-month-ahead forecasting of the SPI index (SPI_{t+1}) at sample stations. This table presents the statistical metrics R^2, MSE, and MAE to facilitate the selection of the optimal input structure in different climatic regions of Iran and highlights the geographical dependency of the models' performance.

Station	Model	Scenario	R2	MSE	MAE
Zabol	RF	1	0.39	0.61	**0.58**
	LSTM	S2-4	0.42	0.59	**0.57**
	RF	**S3-1**	**0.51**	**0.50**	0.57
Anzali	**RF**	**1**	**0.57**	**0.45**	0.52
	LSTM	S2-4	0.44	0.61	**0.62**
	RF	S3-1	0.56	0.45	**0.52**
Mashhad	RF	1	0.57	0.45	**0.52**
	LSTM	S2-4	0.52	0.49	**0.55**
	RF	**S3-2**	**0.59**	**0.43**	0.50
Esfahan	RF	1	0.47	0.53	**0.58**
	LSTM	S2-4	0.41	0.58	**0.59**
	RF	**S3-2**	**0.65**	**0.36**	0.45
Ahwaz	RF	1	0.49	0.54	**0.55**
	LSTM	**S2-4**	**0.55**	**0.45**	0.47
	RF	S3-2	0.13	0.93	**0.76**
Chabahar	**RF**	**1**	**0.49**	**0.54**	0.55
	LSTM	S2-4	0.06	0.94	**0.78**
	RF	S3-1	0.43	0.57	**0.60**
Bandar Abbas	RF	1	0.31	0.70	**0.66**

	LSTM	S2-4	0.27	0.76	**0.67**
	RF	**S3-2**	**0.47**	**0.56**	0.60

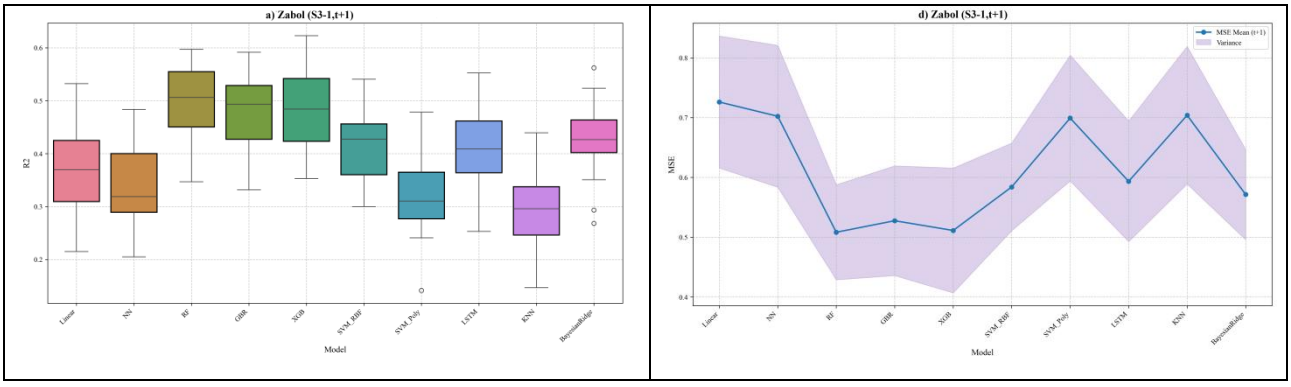
4.5. Uncertainty Analysis and Model Robustness

Given the stochastic nature of machine learning algorithms, evaluating their stability is as crucial as assessing their accuracy. To verify the reproducibility of the results, the models were executed 20 independent times under their optimal scenarios for two contrasting climatic stations: Zabol (arid, Scenario S3-1) and Anzali (hyper-humid, Scenario S1). Figure 10 illustrates the distribution of the Coefficient of Determination (R^2) and the variability of the Mean Squared Error (MSE) for the one-month forecast horizon (t+1). The complete stability analysis for longer horizons (t+2and t+3) is provided in the Supplementary Material (Figures S7 and S8).

Stability in Arid Regions (Zabol): As shown in the top panels of Figure 10, the Random Forest (RF) model exhibits superior stability under the S3-1 scenario. The boxplot (Fig. 10a) reveals that RF not only achieves the highest median R^2 but also possesses the narrowest Interquartile Range (IQR), indicating highly convergent results across different runs. Furthermore, the MSE analysis (Fig. 10b) demonstrates that RF maintains the lowest mean error with a significantly tighter 95% confidence interval (shaded area) compared to deep learning models like LSTM. While LSTM and SVM-Poly show potential for high accuracy in individual runs, their wide variance bands indicate higher uncertainty and sensitivity to initial conditions.

Stability in Humid Regions (Anzali): The robustness of the RF model is further confirmed in the hyper-humid climate of Anzali under the S1 scenario (Fig. 10c, d). Similar to the arid region, RF demonstrates the best balance between accuracy and stability for the first month. However, as detailed in Figure S8 (Supplementary Material), this stability is time-dependent. At the t+3 horizon, the uncertainty for all models increases notably, and the performance gap between RF and LSTM narrows, reflecting the inherent challenge of long-term forecasting in irregular precipitation regimes.

Conclusion on Robustness: Overall, the uncertainty analysis confirms that among the tested algorithms, RF provides the most reliable and computationally stable predictions for operational drought monitoring in Iran, particularly for short-term forecasting (t+1), whereas complex models like SVM-Poly require careful hyperparameter tuning to mitigate instability.



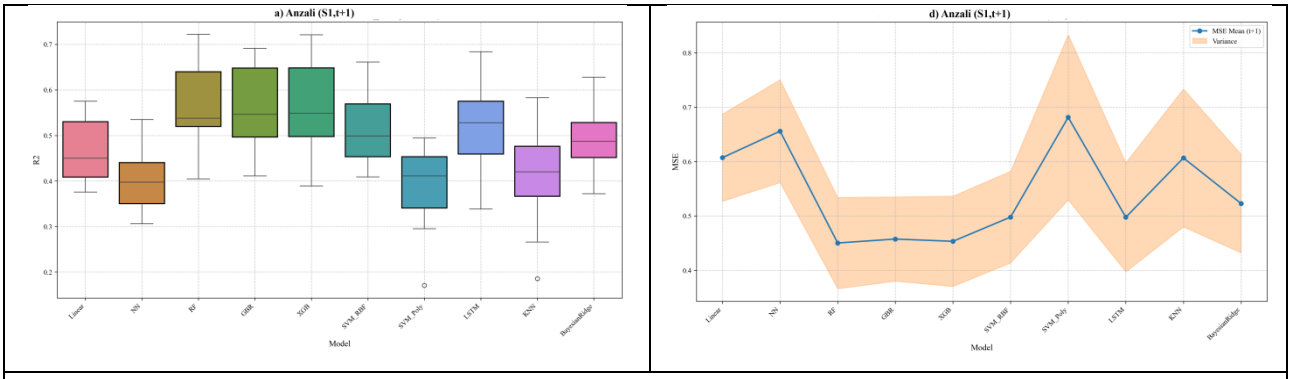


Figure 10. Uncertainty analysis of the models for the one-month forecast horizon (t+1) based on 20 independent runs. Top row (a, b): Zabol station under Scenario S3-1. Bottom row (c, d): Anzali station under Scenario S1. Left panels display the distribution of R^2 (boxplots), and right panels show the mean MSE (solid line) with the 95% confidence interval (shaded area).

5 Discussion

This research was designed and executed to address a fundamental question in the science of drought forecasting: a comparative evaluation of the competing paradigms based on temporal memory (autoregressive), large-scale climate drivers (teleconnections), and a hybrid approach, across the complex climatic landscape of Iran. While the scientific literature is replete with studies that have employed each of these approaches individually (Rezaiy and Shabri, 2025; Anochi and Shimizu, 2025), a comprehensive assessment that simultaneously investigates the relative superiority and synergistic potential of these contrasting philosophies on a national scale, using a diverse array of machine learning models, has been notably absent (Oyarzabal et al., 2025). The findings of the present study, while aligning with the broad global trend towards more complex and hybrid models (Fung et al., 2020; Gyaneshwar et al., 2023), present a far more complex and location-dependent picture of drought predictability dynamics and decisively refute the simplistic notion that one approach is universally superior. Instead, these empirical findings closely resonate with the broader theoretical frameworks of Sub-seasonal to Seasonal (S2S) climate predictability. In this broader context, predictive skill is rarely spatially uniform; rather, it fluctuates based on the delicate balance between slowly varying oceanic boundary conditions (sources of external predictability) and local land-surface memory (sources of internal predictability) (Vitart et al., 2017; Merryfield et al., 2020). One of the stable and consistent findings across all scenarios was the superiority of ensemble learning models, particularly Random Forest (RF), over individual algorithms; a result that fully aligns with the prevailing trend in recent hydroclimatic forecasting literature (Poudel et al., 2024; Sadrtdinova et al., 2024). The inherent ability of the RF model to reduce variance and manage high-dimensional feature spaces makes it an ideal tool for modeling the noisy and non-linear signals of drought. However, with the increase in the forecast horizon and its associated intrinsic uncertainty, the competitive dynamics among the models underwent a shift, and the Bayesian Ridge Regression (BRR) model emerged as a serious contender. This

noteworthy phenomenon reveals a profound insight: in conditions where the predictive signal weakens, the probabilistic approach of a Bayesian model, which quantifies uncertainty within its structure, gains a strategic advantage over models solely based on error optimization. This observation, while emphasizing that the superiority of a specific model is not universal and depends on the problem's conditions (Poudel et al., 2024), highlights the growing importance of probabilistic models for seasonal and longer-term forecasts.

In exploring the autoregressive paradigm (Scenario S2), the results clearly demonstrated that increasing the length of the input temporal memory up to 12 months directly and significantly improves the performance of the models, especially LSTM. This finding, repeated across all three forecast horizons, is more than a mere statistical observation; it is empirical testimony to the long-term memory and considerable inertia within Iran's hydroclimatic system. The persistence of dry or wet conditions, rooted in slow processes such as soil moisture dynamics and groundwater storage, provides a strong signal for forecasting future conditions. The superiority of the LSTM model, specifically designed to learn long-term dependencies, corroborates this hypothesis from an algorithmic perspective as well. This result aligns with a wide range of studies that emphasize the importance of signal preprocessing to unveil complex temporal patterns. Approaches such as wavelet analysis, which have been used in numerous studies as a powerful tool for decomposing the drought signal (Belayneh and Adamowski, 2013; Deo et al., 2017; Alkan, 2023), are, in fact, pursuing a similar objective: extracting information embedded in different time scales. Our research shows that using architectures with inherent long-term memory like LSTM can be an alternative and efficient solution to achieve this same goal.

On the other end of the spectrum, the paradigm based on physical drivers (Scenario S1) proved its capability as a powerful source for prediction, particularly at the short one-month horizon. The significant performance of this scenario in the northern and western regions, which are heavily influenced by large-scale atmospheric systems, confirms the existence of a strong and modelable physical link between general atmospheric circulation and regional precipitation, consistent with studies that have successfully used teleconnection indices as inputs for prediction models (Anshuka et al., 2021). This result empirically validates the physical insights provided by Mahmoudi et al. (2025), who used SHAP analysis to demonstrate that these specific regions exhibit the highest sensitivity to global climate drivers. Yet, the present study extends beyond that diagnostic finding by quantifying how much this physical information can enhance prediction accuracy compared to a memory-based baseline. However, the sharp decline in this scenario's performance at the two- and three-month lead times reveals the inherent limitation of this approach. This performance drop is an empirical manifestation of a phenomenon known in climate science as the predictability wall. Within the broader predictability literature, this rapid decay of skill is widely recognized as the transition from a boundary-forced predictable state (e.g., global teleconnections) to an internal weather-noise dominated state, a dynamic closely tied to the signal-to-noise paradox in climate modeling (Scaife and Smith, 2018). Large-scale climate signals decay over time, and their influence on local conditions is weakened by stochastic and chaotic atmospheric processes. This finding suggests that relying solely on teleconnections for forecasts beyond the intra-seasonal scale may face serious limitations.

Perhaps the most significant and innovative finding of this research lies in the results of the hybrid scenario (S3) and the final comparative analysis. These results explicitly show that the answer to the question Which paradigm is better? does not have a universal answer but is strongly dependent on the geographical location, climatic characteristics, and forecast horizon. From a climatological and physical perspective, this spatial dichotomy is deeply rooted in the physical geography of Iran. In coastal regions (e.g., Anzali on the Caspian coast and Chabahar on the Gulf of Oman), the climate is directly modulated by marine moisture fluxes and unhindered exposure to large-scale atmospheric circulation patterns (such as the Siberian High or the Indian Summer Monsoon). These dominant external forcings make teleconnection signals (S1) highly effective predictors. Conversely, in the arid inland regions and the Central Plateau, external oceanic signals are significantly attenuated by the blocking effect of the Zagros and Alborz mountain ranges. Consequently, the local climate is largely governed by land-atmosphere feedbacks, high evapotranspiration rates, and prolonged dry spells. In these environments, the physical inertia of the system—driven by slow-responding variables like soil moisture deficits and groundwater depletion—makes internal temporal memory (autoregressive models) and hybrid approaches critical for capturing the inherent persistence of drought. The decisive superiority of the hybrid scenario in large parts of the central plateau and the arid and semi-arid regions of Iran transmits a key message: in these areas, neither source of information is sufficient on its own. In arid climates with irregular precipitation, temporal memory alone cannot provide a complete picture, and the response of the local climate to large-scale drivers is either weak or complex. Under such conditions, the synergy resulting from the integration of these two information sources allows the model to achieve a more robust forecast by leveraging the weak but meaningful signal of teleconnections and complementing it with information about drought inertia. This discovery empirically confirms, on a large scale, the importance of hybrid approaches, which have been identified as the forefront of research in recent review articles (Alawsi et al., 2022). Furthermore, the discovery of the spatial complementarity pattern in optimizing memory length within the hybrid scenario—which showed that shorter-term memory is more optimal for humid regions and longer-term memory for arid regions—provides a profound insight into the different dynamics of the climate system in various regimes. In arid climates, where the system is slower and has more inertia, long memory is an asset; whereas in humid climates with faster dynamics, this same long memory can act as noise and lead to overfitting. Finally, the existence of climatic niches such as Chabahar (where the physical paradigm S1 was superior due to the dominance of monsoon dynamics) and Ahvaz (where the autoregressive paradigm S2 was likely dominant due to the prevalence of local processes) brings the complexity of drought forecasting in a country with Iran's climatic diversity to its peak and dooms any attempt to find a single, unified solution to failure.

In sum, by placing the main paradigms of drought forecasting in a comprehensive crucible, this research demonstrates that the future of this field lies not in a dogmatic adherence to one approach, but in the development of adaptive and region-centric modeling frameworks. Such frameworks must be able to dynamically select the optimal combination of predictors, memory length, and model architecture based on the climatic characteristics of each region. From an operational perspective, these findings imply the necessity of moving towards multi-model, multi-paradigm early warning systems that intelligently

aggregate the outputs of models based on different philosophies. The path for future research is also clear: utilizing eXplainable Artificial Intelligence (XAI) techniques to open the "black box" of these relationships and gain a deeper understanding of the governing physical mechanisms (Dikshit and Pradhan, 2021a; Moghaddasi et al., 2025), as well as extending this framework to forecast more impact-oriented indices such as SPEI and agricultural and hydrological drought indices, are the logical next steps toward realizing a comprehensive drought risk management system in Iran.

6 Conclusion

This research was undertaken to resolve a fundamental conundrum in the science of drought forecasting: determining the relative superiority or synergistic potential of paradigms based on temporal memory versus those driven by physical teleconnections, conducted at the national scale of Iran. By executing a comprehensive, multi-model, and multi-paradigm evaluation, this study arrived at a decisive yet nuanced answer to this question: neither philosophy is singularly and universally superior to the other. Instead, their efficacy is distributed in a complex geographical pattern, contingent directly upon the local climatic regime and the forecast horizon.

The foremost contribution of this research is the empirical demonstration that the future of operational drought forecasting systems lies not in advocating for a single approach, but in the intelligent and region-centric integration of information sources. While hybrid models, which amalgamate historical and physical information, emerged as the superior approach across vast portions of Iran's arid and semi-arid climates, in specific "climatic niches," standalone paradigms exhibited more optimal performance (teleconnection-based on the northern and southern coasts, and temporal memory-based in the southwest). This discovery decisively refutes the 'one-size-fits-all' hypothesis and substantiates the necessity of transitioning towards adaptive modeling frameworks.

From an algorithmic perspective, this study confirmed the sustained superiority of ensemble learning models, particularly Random Forest (RF), as a robust tool for modeling the non-linear and noisy signals of drought. However, the close competition from the Bayesian Ridge Regression (BRR) model at longer forecast horizons underscored the growing importance of probabilistic models for managing uncertainty in seasonal forecasts.

Ultimately, by providing a detailed roadmap of "which paradigm performs best where," this research takes a fundamental step toward the operationalization of forecasting knowledge. The findings strongly suggest that the next generation of drought early warning systems must transcend static and uniform architectures, evolving into dynamic, multi-paradigm systems capable of optimally selecting the most suitable modeling strategy based on regional characteristics. This transition will not only enhance forecast accuracy but also provide a more robust scientific foundation for risk management and national resilience against one of the most devastating climatic hazards. Future research should focus on extending this framework to other drought indices and leveraging eXplainable Artificial Intelligence (XAI) to deepen the physical understanding of these complex patterns.

Acknowledgments

This study was successfully completed thanks to the backing of several institutions. Financial sponsorship was provided by the National Natural Science Foundation of China (NSFC; Award No. 42261144671), Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (Award No. 311024001), National Key R&D Program of China (Award No. 2024YFE0103200), and the Iran National Science Foundation (INSF; Award No. 4013097). The analysis herein would not have been possible without the critical precipitation data made available by the Iran Meteorological Organization (IRIMO), whose contribution is sincerely appreciated.

CRediT Authorship Contribution Statement

Jun Jian: Formal analysis, Investigation, Writing – review & editing. Peyman Mahmoudi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. Jing Yang: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – review & editing. Pouria Jafari: Methodology, Investigation, Data curation, Investigation, Software, Validation, Writing – review & editing. Alireza Ghaemi: Data curation, Resources, Software, Writing – original draft, Writing – review & editing. Fatemeh Firoozi: Resources, Visualization. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement:

The datasets utilized and analyzed during this investigation are available upon reasonable request from the corresponding author.

Conflicts of Interest:

The authors declare that they have no competing interests

References

Achite, M., Elshaboury, N., Jehanzaib, M., Vishwakarma, D. K., Pham, Q. B., Anh, D. T., Abdelkader, E. M., and Elbeltagi, A.: Performance of machine learning techniques for meteorological drought forecasting in the Wadi Mina Basin, Algeria, Water, 15, 765, doi:10.3390/w15040765, 2023.

Aldhafeeri, A. A., Ali, M., Khan, M., and Labban, A. H.: SPI-informed drought forecasts integrating advanced signal decomposition and machine learning models, Water, 17, 2747, doi:10.3390/w17182747, 2025.

Alkan, A.: Drought forecasting using Palmer Drought Severity Index with wavelet transform technique and machine learning methods, Int. J. Res. Publ. Rev., 4, 2177–2185, doi:10.55248/gengpi.2023.4158, 2023.

Ali, S., Khorrami, B., Jehanzaib, M., Tariq, A., Ajmal, M., Arshad, A., Shafeeque, M., Dilawar, A., Basit, I., Zhang, L., Sadri, S., Niaz, M. A., Jamil, A., and Khan, S. N.: Spatial downscaling of GRACE data based on XGBoost model for improved understanding of hydrological droughts in the Indus Basin Irrigation System (IBIS), Remote Sens., 15, 873, doi:10.3390/rs15040873, 2023.

Alijani, B.: Climate of Iran, Payam Noor University Press, Tehran, 1997 (in Persian).

Alipour, M. H.: Streamflow prediction in ungauged basins located within data-scarce areas using XGBoost: role of feature engineering and explainability, Int. J. River Basin Manag., 23, 71–92, 2025.

Almikaeel, W., Čubanová, L., and Šoltész, A.: Hydrological drought forecasting using machine learning – Gidra River case study, Water, 14, 387, 2022.

Alawsi, M. A., Zubaidi, S. L., Al-Bdairi, N. S. S., Al-Ansari, N., and Hashim, K.: Drought forecasting: A review and assessment of the hybrid techniques and data pre-processing, Hydrology, 9, 115, doi:10.3390/hydrology9070115, 2022.

Amini, M., Ghadami, M., Fathian, F., and Modarres, R.: Teleconnections between oceanic–atmospheric indices and drought over Iran using quantile regressions, Hydrol. Sci. J., 65, 2286–2295, doi:10.1080/02626667.2020.1802029, 2020.

Anochi, J. A., and Shimizu, M. H.: Precipitation forecasting and drought monitoring in South America using a machine learning approach, Meteorology, 4, 1, doi:10.3390/meteorology4010001, 2025.

Anshuka, A., Buzacott, A. J. V., Vervoort, R. W., and van Ogtrop, F. F.: Developing drought index-based forecasts for tropical climates using wavelet neural network: An application in Fiji, Theor. Appl. Climatol., 143, 557–569, doi:10.1007/s00704-020-03446-3, 2021.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K., and Takahashi, N.: The quasi-biennial oscillation, Rev. Geophys., 39, 179–229, doi:10.1029/1999RG000073, 2001.

Bamston, A. G., Chelliah, M., and Goldenberg, S. B.: Documentation of a highly ENSO-related SST region in the equatorial Pacific: Research note, Atmos.–Ocean, 35, 367–383, doi:10.1080/07055900.1997.9649597, 1997.

Barnston, A. G., and Livezey, R. E.: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns, Mon. Weather Rev., 115, 1083–1126, doi:10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2, 1987.

Belayneh, A., and Adamowski, J.: Drought forecasting using new machine learning methods, J. Water Land Dev., 18, 3–12, doi:10.2478/jwld-2013, 2013.

Cai, F., Liu, C., Gerten, D., Yang, S., Zhang, T., Li, K., and Kurths, J.: Sketching the spatial disparities in heatwave trends by changing atmospheric teleconnections in the Northern Hemisphere, Nat. Commun., 15, 8012, doi:10.1038/s41467-024-52254-0, 2024.

Chandrasekara, S. S., Kwon, H.-H., Vithanage, M., Obeysekera, J., and Kim, T.-W.: Drought in South Asia: A review of drought assessment and prediction in South Asian countries, Atmosphere, 12, 369, 2021.

Craig, M., and Allan, R. P.: The role of teleconnection patterns in the variability and trends of growing season indices across Europe, Int. J. Climatol., 42, 1072–1091, doi:10.1002/joc.7290, 2022.

Deo, R. C., Tiwari, M. K., Adamowski, J. F., and Quilty, J. M.: Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model, Stoch. Environ. Res. Risk Assess., 31, 1211–1240, 2017.

De Pauw, E., Dehkordi, V. R. G., and Ghaffari, A.: Agroecological zones, in: The Soils of Iran, Springer, 163–173, doi:10.1007/978-3-319-69048-3_9, 2018.

Dikshit, A., and Pradhan, B.: Explainable AI in drought forecasting, Mach. Learn. Appl., 6, 100192, doi:10.1016/j.mlwa.2021.100192, 2021a.

Dikshit, A., and Pradhan, B.: Interpretable and explainable AI (XAI) model for spatial drought prediction, Sci. Total Environ., 801, 149797, doi:10.1016/j.scitotenv.2021.149797, 2021b.

Djerbouai, S., and Souag-Gamane, D.: Drought forecasting using neural networks, wavelet neural networks, and stochastic models: Case of the Algerois Basin in North Algeria, Water Resour. Manag., 30, 2445–2464, doi:10.1007/s11269-016-1298-6, 2016.

Dong, H., Huang, S., Wang, H., Shi, H., Singh, V. P., She, D., Huang, Q., Leng, G., Gao, L., and Wei, X.: Effects of multiple large-scale atmospheric circulations on precipitation dynamics in China, Sci. Total Environ., 923, 171528, doi:10.1016/j.scitotenv.2024.171528, 2024.

Enfield, D. B., Mestas-Nuñez, A. M., Mayer, D. A., and Cid-Serrano, L.: How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures?, J. Geophys. Res.–Oceans, 104, 7841–7848, doi:10.1029/1998JC900109, 1999.

Esquivel-Saenz, P. J., Ortiz-Gómez, R., Zavala, M., and Flowers-Cano, R. S.: Artificial neural networks for drought forecasting in the central region of Zacatecas, Mexico, Climate, 12, 131, doi:10.3390/cli12090131, 2024.

Fung, K. F., Huang, Y. F., Koo, C. H., and Soh, Y. W.: Drought forecasting: A review of modelling approaches 2007–2017, J. Water Clim. Change, 11, 771–799, doi:10.2166/wcc.2019.236, 2020.

Ganguli, P., and Janga Reddy, M.: Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach, Hydrol. Process., 28, 4989–5009, doi:10.1002/hyp.9966, 2014.

Ghaemi, A., Hashemi Monfared, S. A., Bahrpeyma, A., Mahmoudi, P., and Zounemat-Kermani, M.: Spatiotemporal variation of projected drought characteristics in Iran under climate change scenarios using CMIP5-CORDEX product, J. Water Clim. Change, 15, 1054–1075, 2024a.

Ghaemi, A., Hashemi Monfared, S. A., Bahrpeyma, A., Mahmoudi, P., and Zounemat-Kermani, M.: Exploitation of ensemble-based machine learning strategies to elevate the precision of CORDEX simulations in precipitation projection, Earth Sci. Inform., 17, 1373–1392, 2024b.

Ghamghami, M., and Bazrafshan, J.: Relationships between large-scale climate signals and winter precipitation patterns over Iran, J. Hydrol. Eng., 26, doi:10.1061/(ASCE)HE.1943-5584.0002066, 2021.

Gyaneshwar, A., Mishra, A., Chadha, U., Vincent, P. M. D. R., Rajinikanth, V., Ganapathy, G. P., and Srinivasan, K.: A contemporary review on deep learning models for drought prediction, Sustainability, 15, 6160, doi:10.3390/su15076160, 2023.

Heydarizad, M., Raeisi, E., Sori, R., and Gimeno, L.: Identification of Iran's moisture sources using a Lagrangian particle dispersion model, Atmosphere, 9, 408, doi:10.3390/atmos9100408, 2018.

Helali, J., Mohammadi Ghaleni, M., Hosseini, S. A., Lotfi Siraei, A., Saeidi, V., Safarpour, F., Mirzaei, M., and Lotfi, M.: Assessment of machine learning model performance for seasonal precipitation simulation based on teleconnection indices in Iran, Arab. J. Geosci., 15, 1343, doi:10.1007/s12517-022-10640-2, 2022.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., and Folland, C. K.: A tripole index for the Interdecadal Pacific Oscillation, Clim. Dyn., 45, 3077–3090, doi:10.1007/s00382-015-2525-1, 2015.

Higgins, R. W., Leetmaa, A., Xue, Y., and Barnston, A.: Dominant factors influencing the seasonal predictability of U.S. precipitation and surface air temperature, J. Clim., 13, 3994–4017, doi:10.1175/1520-0442(2000)013<3994:DFITSP>2.0.CO;2, 2000.

Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T., and Pegion, P.: On the increased frequency of Mediterranean drought, J. Clim., 25, 2146–2161, doi:10.1175/JCLI-D-11-00296.1, 2012.

Huguet, F., Gaetani, M., Pohl, B., and Douville, H.: Ocean–atmosphere drivers and teleconnection patterns of Indian summer monsoon rainfall extremes, Clim. Dyn., 55, 543–561, doi:10.1007/s00382-020-05292-y, 2020.

Huntsman, C., and Kalra, A.: Long-term drought forecasting using deep learning and artificial neural networks, J. Hydroinformatics, 22, 149–164, doi:10.2166/hydro.2019.064, 2020.

Jones, P. D., Jonsson, T., and Wheeler, D.: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland, Int. J. Climatol., 17, 1433–1450, doi:10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P, 1997.

Khedun, C. P., Mishra, A. K., Singh, V. P., and Giardino, J. R.: A copula-based precipitation forecasting model: Investigating the interdependencies between precipitation and climate indices in the Texas–Gulf region, Water Resour. Res., 50, 1381–1403, doi:10.1002/2013WR014133, 2014.

Kiladis, G. N., and Diaz, H. F.: Global climatic anomalies associated with extremes in the Southern Oscillation, J. Clim., 2, 1069–1090, doi:10.1175/1520-0442(1989)002<1069:GCAAWE>2.0.CO;2, 1989.

King, M. P., Hell, M., and Keenlyside, N.: Investigation of the atmospheric mechanisms related to the autumn NAO using a nudging hindcast approach, Clim. Dyn., 44, 1609–1629, doi:10.1007/s00382-014-2182-7, 2015.

Kofi, D. F., Mani, A., Rousta, I., Khozeymeh-Nezhad, H., Bhattacharya, P., and Viterbo, P.: Evaluation of multiple machine learning models in forecasting drought over semi-arid regions using meteorological drought indices, Sensors, 24, 2154, doi:10.3390/s24062154, 2024.

Labban, A. H., Albadarin, A. B., Aldhafeeri, A. A., and Ali, M.: A robust framework for meteorological drought forecasting: SAGD-XGBoost hybrid model applied to Kuwaiti environments, Environ. Monit. Assess., 196, 479, doi:10.1007/s10661-024-12445-8, 2024.

Latif, M., Dommenget, D., Dima, M., and Grötzner, A.: The role of Indian Ocean sea surface temperature in forcing East African rainfall anomalies during the 1997/98 ENSO, J. Clim., 12, 3497–3504, doi:10.1175/1520-0442(1999)012<3497:TROIOT>2.0.CO;2, 1999.

Lau, N.-C., and Nath, M. J.: Impact of ENSO on SST variability in the North Pacific and North Atlantic: Seasonal dependence and role of extratropical atmosphere–ocean coupling, J. Clim., 14, 2846–2866, doi:10.1175/1520-0442(2001)014<2846:IOEOVI>2.0.CO;2, 2001.

Lim, Y.-K.: The East Atlantic/West Russia (EA/WR) teleconnection in the North Atlantic: Climate impact and relation to Rossby wave propagation, Clim. Dyn., 44, 3211–3222, doi:10.1007/s00382-014-2381-4, 2015.

Lim, Y.-K.: The impact of the North Atlantic Oscillation on European climate: A review, Rev. Geophys., 54, 199–244, doi:10.1002/2015RG000520, 2016.

Liu, Y., Ren, G., Song, C., Shen, S., Liang, Y., and Yanchen, S.: Seasonal drought forecasting using an ensemble machine learning model: An application for southeast China, Water Resour. Manag., 36, 4743–4758, doi:10.1007/s11269-022-03268-x, 2022.

Mahmoudi, P., Jafari, P., Ghaemi, A., Jian, J., Firoozi, F., and Yang, J.: Decoding Iran's drought drivers: An explainable AI approach to unraveling global teleconnection impacts, Earth Syst. Environ., doi:10.1007/s41748-025-00925-3, 2025.

Marx, C., and Paeth, H.: Combining teleconnection patterns and machine learning for statistical seasonal climate forecast, Clim. Dyn., 58, 1207–1224, doi:10.1007/s00382-021-05981-y, 2022.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, in: Proceedings of the Eighth Conference on Applied Climatology, Anaheim, CA, 17–22 January 1993, American Meteorological Society, Boston, MA, 179–184, 1993.

Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., Danabasoglu, G., Dirmeyer, P. A., Doblas-Reyes, F. J., Domeisen, D. I. V., Ferranti, L., Ilynia, T., Kumar, A., Müller, W. A., Rixen, M., Robertson, A. W., Smith, D. M., Takaya, Y., Tuma, M., Vitart, F., White, C. J., Álvarez, M. S., Ardilouze, C., Attard, H., Baggett, C., Balmaseda, M. A., Beraki, A. M., Bhattacharjee, P. S., Bilbao, R., Andrade, F. M. de, DeFlorio, M. J., Díaz, L. B., Ehsan, M. A., Fragkoulidis, G., Gonzalez, A. O., Grainger, S., Green, B. W., Hell, M. C., Infanti, J. M., Isensee, K., Kataoka, T., Kirtman, B. P., Lee, J.-Y., Mayer, K., McKay, R., Mecking, J. V., Miller, D. E., Neddermann, N., Ólason, E. Ó., Pepler, A. S., Perlwitz, J., Recalde-Coronel, G. C., Reintges, A., Renkl, C., Solaruja-Murali, B., Spring, A., Stan, C., Sun, Y. Q., Tozer, C. R., Vigaud, N., Woolnough, S., and Yeager, S.: Current and Emerging Developments in Subseasonal to Decadal Prediction, Bull. Amer. Meteorol. Soc., 101, E869–E896, doi:10.1175/BAMS-D-19-0037.1, 2020.

Miller, R. D., and Modarres, R.: Climate controls on daily precipitation in Iran: Application of canonical correlation analysis (CCA), Int. J. Climatol., 59, 139–150, doi:10.1002/joc.8480, 2025.

Mishra, A. K., and Singh, V. P.: A review of drought concepts, J. Hydrol., 391, 202–216, doi:10.1016/j.jhydrol.2010.07.012, 2010.

Montazeri, M., and Gharib, M.: OCI–SPI modeling and drought prediction based on ocean climate indices using multiple linear regression and ANN models (case study: Mazandaran Province, Iran), Theor. Appl. Climatol., 135, 1461–1474, doi:10.1007/s00704-018-2458-0, 2019.

Nam, W.-H., Hong, E.-M., and Kim, T.: Machine learning-driven drought prediction using multi-seasonal climate patterns, Environ. Res., 236, 115708, doi:10.1016/j.envres.2023.115708, 2023.

Nury, A. H., Hasan, S., Alam, M. J., Paul, S. K., and Das, D. K.: Drought index forecasting in Bangladesh using various neural network models: A wavelet-based approach, Water, 13, 673, doi:10.3390/w13050673, 2021.

Panda, S. N., Mishra, A., and Shikha, D.: SPI-based drought modeling and forecasting using wavelet neural network supplemented with decomposition methods in eastern India, Stoch. Environ. Res. Risk Assess., 32, 1041–1056, doi:10.1007/s00477-017-1445-8, 2018.

Rahmandad, A., and Kerachian, R.: A data-driven model for robust drought forecasting under uncertainty: Application to the Lake Urmia Basin, Iran, J. Hydrol., 613, 128383, doi:10.1016/j.jhydrol.2022.128383, 2022.

Rao, B., Nguyen, V.-H., Chua, S. L., Adem, A. A., Le, M.-H., and Ahn, K.-K.: Effects of large-scale climatic patterns on dry conditions in Korea, Water Resour. Manag., 31, 2165–2181, doi:10.1007/s11269-017-1635-y, 2017.

Rêgo, J. C., Fernandes, L. L., and Rocha, A. M.: Evaluation of machine learning methods for drought forecasting in Brazil, Theor. Appl. Climatol., 152, 111–124, doi:10.1007/s00704-022-04134-3, 2023.

Ropelewski, C. F., and Halpert, M. S.: Global and regional precipitation patterns associated with the El Niño/Southern Oscillation, Mon. Weather Rev., 115, 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARAPO>2.0.CO;2, 1987.

Saeed, F., and Mazouz, N.: Improved short-term drought prediction for semi-arid climates using hybrid machine learning–SARIMA models, J. Arid Environ., 215, 104874, doi:10.1016/j.jaridenv.2023.104874, 2024.

Salman, S. A., Yaseen, Z. M., and Al-Ansari, N.: Smarter teleconnection indices for drought prediction: A novel hybrid SWI–PSO model, Environ. Model. Softw., 149, 105338, doi:10.1016/j.envsoft.2022.105338, 2022.

Sattar, M. N., and Mahmoudi, P.: Automatic flood detection using WALI algorithm and graph cut fusion in cloud-prone environments, J. Hydrol., 614, 128520, doi:10.1016/j.jhydrol.2022.128520, 2022.

Scaife, A. A., and Smith, D.: A signal-to-noise paradox in climate science, npj Clim. Atmos. Sci., 1, 28, doi:10.1038/s41612-018-0038-4, 2018.

Sehgal, V., Samui, P., Kim, D. M., and Kumar, M.: Three timescale drought forecasting using a wavelet-based random vector functional-link neural network model: A case study in eastern India, Hydrol. Sci. J., 65, 2073–2086, doi:10.1080/02626667.2020.1788113, 2020.

Sen, Z.: Practical and applied hydrogeology, Elsevier, Amsterdam, 2015.

Shahvari, N., and Modarres, R.: Region-based study of wet and dry season precipitation of Iran, Theor. Appl. Climatol., 111, 255–264, doi:10.1007/s00704-012-0636-z, 2013.

Sharifinia, S., Modarres, R., and Fuladlu, Y.: Teleconnection controls on winter precipitation variability in Iran from 1979 to 2018, Int. J. Climatol., 43, 1708–1726, doi:10.1002/joc.7797, 2023.

Shiferaw, H., and Adamowski, J.: Ensemble drought prediction using SWAT and machine learning models in Ethiopia, J. Hydrol., 564, 123–139, doi:10.1016/j.jhydrol.2018.07.045, 2018.

Shiru, M. S., Shahid, S., Chung, E.-S., and Alias, N.: Projection of meteorological droughts in Nigeria during growing seasons under climate change, Sci. Rep., 10, 10107, doi:10.1038/s41598-020-66972-9, 2020.

Sobhani, P., Modarres, R., and Jahanbakhsh-Asl, S.: Assessing relationships between teleconnection patterns and precipitation in western Iran using generalized linear models, Theor. Appl. Climatol., 151, 537–553, doi:10.1007/s00704-022-04112-9, 2023.

Subhakar, D., Panda, S. N., and Mishra, A.: Short-term drought forecasting using multi-resolution forecasting methods with a novel activation function, Stoch. Environ. Res. Risk Assess., 36, 4517–4534, doi:10.1007/s00477-022-02299-y, 2022.

Tabari, H., Some'e, B. S., and Zadeh, M. R.: Shift changes and monotonic trends in rainfall and temperature series over Iran, Theor. Appl. Climatol., 109, 95–108, doi:10.1007/s00704-011-0551-3, 2012.

Tao, X., Wang, H., Liu, J., Zhang, Y., and Chen, X.: Fusion of machine learning models for drought forecasting under complex monsoon climates, Clim. Dyn., 60, 2871–2888, doi:10.1007/s00382-023-06742-y, 2023.

Trenberth, K. E., and Hurrell, J. W.: Decadal atmosphere–ocean variations in the Pacific, Clim. Dyn., 9, 303–319, doi:10.1007/BF00204745, 1994.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index, J. Clim., 23, 1696–1718, doi:10.1175/2009JCLI2909.1, 2010.

Vicente-Serrano, S. M., Azorin-Molina, C., Peña-Gallardo, M., Tomas-Burguera, M., Martín-Hernández, N., Beguería, S., El Kenawy, A., and Domínguez-Castro, F.: The complex influence of ENSO on droughts in Europe, Remote Sens., 12, 1622, doi:10.3390/rs12101622, 2020.

Vicente-Serrano, S. M., Miralles, D. G., Domínguez-Castro, F., and Yang, L.: A review of drought variability and change across Europe, Weather Clim. Extremes, 42, 100536, doi:10.1016/j.wace.2023.100536, 2024.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.-S., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A. W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R., and Zhang, L.: The Subseasonal to Seasonal (S2S) Prediction Project Database, Bull. Amer. Meteorol. Soc., 98, 163–173, doi:10.1175/BAMS-D-16-0017.1, 2017.

Wang, S.-Y., Yoon, J.-H., Becker, E., and Gillies, R. R.: California winter precipitation linked to atmospheric teleconnections, Nat. Commun., 8, 14306, doi:10.1038/ncomms14306, 2017.

Wang, L., Lei, T., Wan, Y., Li, T., Wang, Y., and Gu, H.: Short-term drought forecasting in China using transformer-based deep learning models, Agric. Water Manag., 296, 108011, doi:10.1016/j.agwat.2024.108011, 2024.

Wang, Y., Zhang, L., Zhang, Q., Sun, P., and Singh, V. P.: Influence of multiscale oceanic–atmospheric teleconnections on precipitation variability in China, J. Hydrol., 612, 128268, doi:10.1016/j.jhydrol.2022.128268, 2022.

Wilhite, D. A., Sivakumar, M. V. K., and Pulwarty, R.: Managing drought risk in a changing climate: The role of national drought policy, Weather Clim. Extremes, 3, 4–13, doi:10.1016/j.wace.2014.03.002, 2014.

Wu, H., Hayes, M. J., Weiss, A., and Hu, Q.: An evaluation of the standardized precipitation index, J. Clim., 18, 2466–2482, doi:10.1175/JCLI-14-00552.1, 2005.

Xavier, D. D., Haynes, P. H., and Martin, Z. S.: Large-scale circulation dynamics driving Eurasian winter droughts, Clim. Dyn., 62, 4551–4570, doi:10.1007/s00382-023-06925-7, 2024.

Yamoah, M. A., Li, X., Jehanzaib, M., Abbasi, S. H., and Pham, Q. B.: A hybrid machine learning model for meteorological drought forecasting in West Africa, Water, 16, 1097, doi:10.3390/w16081097, 2024.

Yuan, F., Khan, M., Ali, M., and Khan, S. M.: Teleconnection-based drought prediction using advanced deep learning models: A case study of South China, Clim. Dyn., 61, 2111–2128, doi:10.1007/s00382-023-06711-5, 2023.

Zhang, Q., and Singh, V. P.: Teleconnection impacts on hydrological variability in China: A review, Earth Sci. Rev., 205, 103199, doi:10.1016/j.earscirev.2020.103199, 2020.

Zhao, T., Chen, X., Zheng, H., and Li, J.: Multilevel drought early warning using multi-source deep learning models over China, Remote Sens., 14, 2919, doi:10.3390/rs14122919, 2022.

Zhou, S., She, D., Huang, Q., Leng, G., Wang, H., and Wang, L.: Exploring multi-timescale drought characteristics in China with a deep learning-based data fusion approach, Agric. For. Meteorol., 350, 109742, doi:10.1016/j.agrformet.2024.109742, 2024.

Zhou, Y., Lu, Y., Wang, H., Wang, J., and Guo, S.: Hierarchical AI-based forecasting of hydrological droughts: A case study in the Yangtze River Basin, J. Hydrol., 630, 130738, doi:10.1016/j.jhydrol.2024.130738, 2024.

Zhou, Y., Xu, Y., Zhang, L., and Singh, V. P.: Machine learning-based SPI drought forecasting using GRACE satellite products in China, Water, 16, 2469, doi:10.3390/w16132469, 2024.

Zou, X., Mo, K., Arsenault, K. R., and Ek, M. B.: Evaluation of soil moisture drought recovery metrics in the United States, J. Hydrometeorol., 22, 2681–2698, doi:10.1175/JHM-D-21-0064.1, 2021.

Zubaidi, S. L., Sumi, T., and Al-Ansari, N.: Hybrid data-driven drought prediction using wavelet–MLP and signal decomposition methods, Water, 12, 620, doi:10.3390/w12030620, 2020.

Zhou, X., Zhang, Y., Long, D., and Ma, S.: Drought detection using integrated GRACE terrestrial water storage and machine learning methods in China, Adv. Water Resour., 151, 103908, doi:10.1016/j.advwatres.2021.103908, 2021 (Retracted).

Zuo, L., Li, Y., Zhang, F., and Zhang, J.: Sub-seasonal forecasting of meteorological drought in arid regions using LSTM and wavelet decomposition, J. Arid Environ., 213, 104833, doi:10.1016/j.jaridenv.2023.104833, 2023.