Assessment of C3S Monthly to Seasonal Climate Forecast Models for Mean and Extreme Precipitation Over Iran

Mohammad Saeed Najafi (*Corresponding author***);** Department of Water Resources Study and Research (WRR), Water Research Institute (WRI), Tehran, Iran;
Email: ms.najafi@wri.ac.ir
Contact number: +989122163282
ORCID: (https://orcid.org/0000-0002-2333-1641)

Sajad Akbari Moghaddam Sani; Department of Meteorology and Climatology, Doctoral School of Exact and Natural Sciences, Nicolaus Copernicus University, Toruń, Poland
Email: sajad.akbari@doktorant.umk.pl

Razieh Noroozian; Department of water and hydraulic structures, civil engineering, K. N. Toosi University of Technology, Tehran, Iran
Email: razieh.noroziyan76@gmail.com

Samin Danandeh Saribaglou; Department of water and hydraulic structures, civil engineering, K. N. Toosi University of Technology, Tehran, Iran
Email: Danandehsamin@gmail.com

Abstract

Reliable seasonal climate forecasts are critical for effective water resource management and disaster preparedness, especially in regions like Iran that experience high climatic variability and extreme precipitation events. This study evaluates the performance of seven models from the Copernicus Climate Change Service (C3S) and their multi-model ensemble (MME) in predicting mean and extreme precipitation indices across Iran. Based on data from 148 synoptic stations across diverse climatic zones (1997–2016), forecast skill at one- to three-month lead times was evaluated using Nash–Sutcliffe Efficiency (NSE) and Pearson correlation (Corr). MME forecasts consistently outperformed individual models across all climate zones and lead times. Among individual models, the ECMWF and UKMO systems achieved the highest skill, especially for short lead times, while other models (e.g. Météo-France, CFSv2) generally performed worse. Skill was notably higher in western and northeastern Iran, whereas forecasts were weakest along the humid Caspian coast. Forecast skill declined as lead time increased. These results quantify the strengths and limitations of current C3S seasonal forecasts for Iran, highlighting that multi-model ensembles and top-performing models (ECMWF/UKMO) offer the most reliable guidance for water resource and risk management.

Keywords: Seasonal climate forecasting, Copernicus Climate Change Service (C3S), Multi-model ensemble (MME), Extreme precipitation indices, Model evaluation, Iran.

1. Introduction

Reliable and accurate seasonal climate forecasts are critical tools for disaster risk reduction, agricultural planning, and sustainable water resource management (UNDRR, 2022). Their importance is particularly pronounced in regions vulnerable to climate extremes—such as floods, droughts, and temperature anomalies—where forecast information supports timely decision-making (Alexander & Block, 2022; Yeditha et al., 2023). Typically covering lead times of one to three months, these forecasts help mitigate the societal and economic impacts of hydrological extremes through early warning systems and proactive planning (Kundzewicz & Kaczmarek, 2000). This need is especially urgent in developing and climate-sensitive countries such as Iran, where agriculture employs nearly 18% of the population and consumes over 90% of the nation's freshwater resources (Vaghefi et al., 2019).

Iran with complex topography, limited water availability, and history of climate-related disasters make it an ideal case for evaluating the effectiveness of seasonal prediction systems. Assessing the skill of seasonal forecast models is therefore vital to improving their reliability and identifying regional limitations (Krishnamurthy, 2019; Sikder & Hossain, 2016). Regionally specific assessments of the skill associated with diverse seasonal forecasting techniques have been reported in the scientific literature, with notable examples from e.g., Europe (Mishra et al., 2019), northeastern Brazil (Pinheiro et al., 2023) and Australia (Wang, 2019); but, achieving reliable forecasts over mountainous and arid regions remains a major challenge due to the nonlinear, intermittent, and localized nature of precipitation (Han et al., 2019; Pepler et al., 2015).

To address this challenge, a variety of modeling approaches have been developed. Statistical models rely on empirical relationships between past climate observations but may struggle with non-stationary and data limitations (Li et al., 2023; Xu et al., 2022). In contrast, dynamical models simulate the atmosphere–ocean system using physical laws, offering improved robustness in changing climate regimes (Chan et al., 2021). These global models are often downscaled through regional climate models (RCMs) to improve spatial accuracy (Pan et al., 2015; Wu et al., 2020). However, both approaches are subject to uncertainties stemming from their initialization, structure, and physical parameterizations (Calvo-Olivera et al., 2024; Li et al., 2019; Tippett et al., 2019).

Reliable seasonal to sub-seasonal forecasts are essential for climate services and climate change adaptation. Within this modeling systems, the Copernicus Climate Change Service (C3S) provides operational seasonal forecasts using a multi-model ensemble (MME) system

developed by leading meteorological agencies, including European Centre for Medium-Range Weather Forecasts (ECMWF), UK_Met Office (UKMO), Japan Meteorological Agency (JMA), Deutscher Wetterdienst (DWD) and others. MME have been developed to reduce structural uncertainty and improve skill by leveraging probabilistic consensus across models (Doblas-Reyes et al., 2005; Vitart, 2014). Although these systems are widely applied in Europe and Africa (Almazroui et al., 2022; Gebrechorkos et al., 2022; Nana et al., 2024), their evaluation in regions with diverse climate regimes—such as Iran—remains limited. Moreover the skill of a MME derived from the best C3S project models has not been adequately examined for these specific areas. Evaluating this ensemble could be a key step toward improving forecast accuracy at longer lead times.

While assessing seasonal forecasting models has been previously applied in Iran, most existing studies rely on statistically post-processed outputs, bias adjustment methods to increase accuracy of models or examine a limited number of models (see, e.g., Alizadeh-Choobari, 2019; Yazdandoost et al., 2023). For example, Nobakht et al. (2021) examined only three C3S models (UKMO, Meteo_France, and ECMWF) across eight climatic regions, and Nasseri et al. (2025) focused exclusively on ECMWF's fifth-generation system (SEAS5).

Notably, Iran spans a wide range of climatic zones, from the humid Caspian coast to the hyper-arid interior plateau, making it a challenging environment for seasonal prediction models designed primarily for temperate regions. While previous studies have assessed global or regional models over various regions of Iran (see, e.g., Nobakht et al., 2021; Shirvani et al., 2023; Najafi & Kuchak, 2024), there is a clear gap in evaluating the performance of all C3S forecasts—particularly for extreme precipitation indices—at varying lead times. Therefore, a comprehensive spatial analysis of all C3S models is necessary, given the existence of a more diverse range of models and the significant climatic diversity of Iran. This study aims to fill that gap by providing a detailed, station-by-station assessment of model performance, which can offer a more accurate and operationally relevant understanding of forecasts for the region.

It is necessary to explain that, following the approach of Nikraftar et al. (2024), although statistical post-processing can significantly improve forecast skill, our goal was to examine the raw model outputs in order to better understand the intrinsic strengths and limitations of each forecasting system. By avoiding statistical corrections, we focused on evaluating the core dynamical capabilities of the models, which is particularly important in the context of increasing climate variability and extreme events, where statistical approaches may struggle to extrapolate beyond the training dataset. This approach allowed us to discern model performance

in representing both total precipitation and precipitation extremes. Present work offers a substantial benefits for climate services and forecast users in Iran by identifying the most effective models, thereby helping them select the most reliable seasonal precipitation forecasts.

2. Data and methods

2.1. Data

2.1.1. Observational dataset

For this study, we selected 148 synoptic stations across the country (as shown in Figure 1), each possessing at least 20 years of qualified daily and monthly precipitation data. The data, spanning from 1997 to 2016, were obtained from the Iran's Meteorological Office (IRIMO). Iran experiences an arid and semi-arid climate, receiving less than a third of the world's average annual precipitation (Dehghani et al., 2020; Najafi et al., 2014). While Iran's average annual precipitation is approximately 245.5 mm (Najafi & Kuchak, 2024), this national average conceals profound regional variations. For example, arid desert areas receive as little as 50 mm of rainfall per year, in stark contrast to the southwestern Caspian Sea coast, which witnesses significantly higher amounts exceeding 1700 mm annually (Dehghani et al., 2020). Figure 1 illustrates the topography of Iran, the locations of the stations, and variation of the mean monthly precipitation averaged across the country.

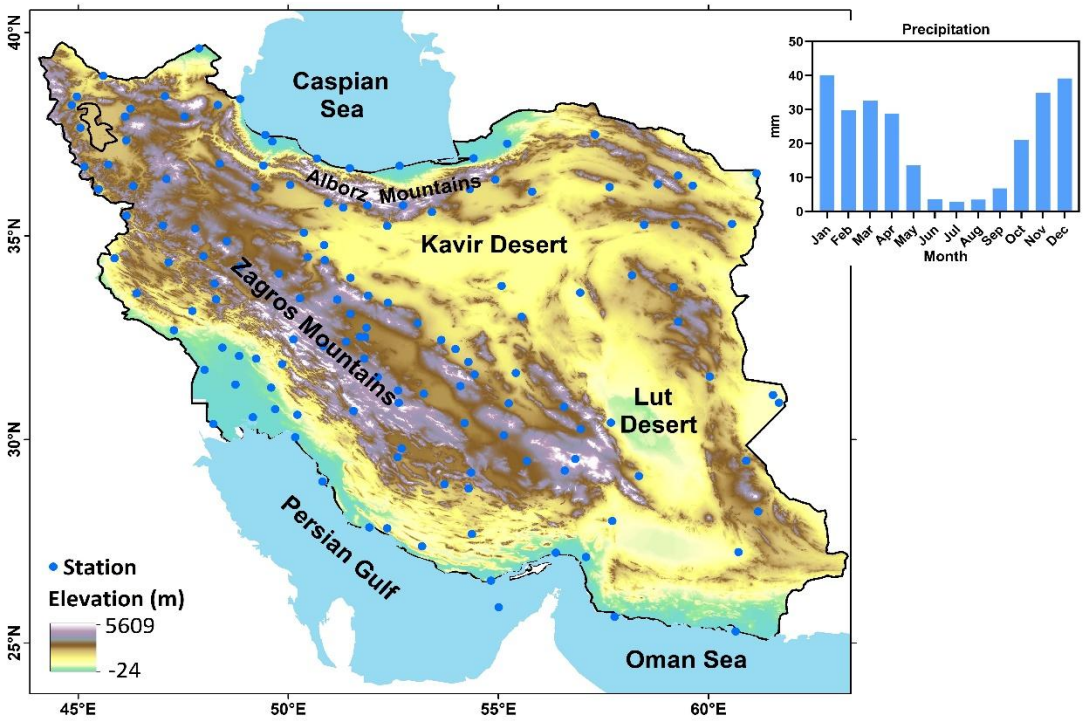


Figure 1: Topographic map of the study area, showing the locations of the 148 synoptic stations used in this study (blue dots). The bar chart in the top right corner illustrates the mean monthly precipitation averaged across all stations for the study period.

2.1.2. Seasonal forecast datasets

This study utilized seven seasonal forecast models (listed in Table 1) from the C3S, accessible via the C3S Climate Data Store (https://cds.climate.copernicus.eu/datasets/seasonal-original-single-levels?tab=download). These models produce daily global precipitation forecasts at a 1.0° spatial resolution, initialized on the first of each month (Gebrechorkos et al., 2022) up to 7 month ahead. This study utilized daily C3S model data from 1997-2016, covering up to three month lead times. From this data, we extracted accumulated monthly precipitation and extreme precipitation indices for all three lead times. To assess the models' predictive skill at 148 specific stations, we adopted a point-pixel evaluation technique, following the methodology of Najafi et al. (2025). This method involved comparing the precipitation recorded at each station with the value of the nearest grid cell from each model dataset. The nearest grid cell was identified using the nearest neighbor approach, calculating the distance between each station and the centroid of every grid cell. Where multiple stations resided within a single grid cell, their precipitation data were averaged before comparison with the model's grid cell value. To correct for the large grid sizes of forecast models in areas with varying altitudes, we derived a regression relationship between observed precipitation and altitude in Iran and applied it to the predicted values.

As part of this study, we examines the relationship between eight SST-driven teleconnection (SSTDT) indices and Iranian precipitation. We'll use correlation calculations to analyze indices such as the Atlantic Multidecadal Oscillation (AMO), Indian Ocean Dipole (IOD), Oceanic Niño Index (ONI), and Pacific Decadal Oscillation (PDO), as well as sea surface temperatures (SST) in the Niño 1+2, 3, 3.4, and 4 regions. After establishing these correlations, we'll assess the importance of these indices for the accuracy of our predictive models. The teleconnection index data was sourced from the NOAA website.

Table 1: Description of selected models of the Copernicus Climate Change Service (C3S) Seasonal Forecast Models

Model	Time range (days)	Forecast initial condition	Hindcast initial condition	Forecast ensemble size members	Hindcast ensemble size	Hindcast period	Spatial resolution of atmospheric/ocean models
ECMWF	215	1st of month	1st of month	51	25	1981-2016	~0.32°–0.72°/0.25°
DWD	180	1st of month	1st of month	50	30	1993-2023	~100 km/0.4°
CMCC	184	1st of month	1st of month	50	40	1993-2016	0.5°/0.25°
Météo-France	215	1st of month	1st of month	25	24	1993-2018	~0.5°/0.25°

JMA	215	every day	2 start dates lagged by 15 days	5	5	1993-2016	~55 km/0.25°
NCEP (CFSv.2)	215	each day of month at every 6 hours	every 5 days at every 6 hours	4/day	4/start date	1993-2016	~1°/0.25°(equator) to 0.5°
UKMO	215	each day of month	1st, 9th, 17th, 25th of month	2/day	7/start time	1993-2016	0.83° - 0.56°/0.25°

2.2. Evaluation

This study assesses the performance of C3S models for monthly to seasonal precipitation forecasts, considering lead times up to three months. This analysis provides an overview of the models' ability to predict both mean and extreme precipitation events across Iran throughout the study period. To rigorously assess the seasonal precipitation predictability of the C3S models across Iran, we employed several common deterministic verification metrics. For evaluating each model's predictive ability, we specifically utilized the Nash-Sutcliffe Efficiency (NSE) metric. The NSE, represented by Equation (1), quantifies the goodness of fit of model predictions to observed data. It is calculated as:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \qquad (1)$$

where O_i represents the i-th observed value, P_i represents the i-th forecasted value, $\bar{O}$ is the mean of the observed values, and n(240) is the total number of data (12×20). The linear relationship between forecasted and observed precipitation was quantified using Pearson's correlation coefficient (Corr) (Equation 2):

$$\mathrm{Corr} = \frac{\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2 \sum_{i=1}^{n}(P_i - \overline{P})^2}} \qquad (2)$$

where n denotes the number of observations (240), P_i is the i-th modeled values, O_i is the i-th observed values in synoptic stations, $\bar{O}$ ($\bar{P}$) is the average of the observed (modeled) precipitation values. For a comprehensive evaluation, NSE and Corr calculations were performed systematically across our dataset (1997-2016), specifically for each individual C3S model and every specific station throughout Iran. It's important to note that in evaluating C3S model performance for precipitation forecasting, the index 'i' denotes the month, with separate analyses for each lead time. Conversely, for extreme indices, 'i' represents the year, as these indices are computed on an annual basis; the analysis in this section is conducted on an annual basis because the output of the extreme indices (Table 2) is calculated as the number of days per year or mm per year. Therefore, the analyses in this section are based on the annual output of the indices predicted by the model and their comparison with annual observed values.

2.3. **Multi-model forecasts**

Seven selected models were combined to generate multi-model forecasts using multivariate regression (MRMM) technique (Eq. 3; Zhi et al., 2012):

$$MRMM = \overline{O} + \frac{1}{N}\sum_{i=1}^{N} a_i (p_i - \overline{p_i}) \qquad (3)$$

where $\bar{o}$ is the mean observed value, p_i is the *i*-th model forecast value, $\bar{p}$ is the mean of the *i*-th model forecast value, a_i is the weight of the i-th model, which can be calculated by following method, and N is the number of models participating in the MRMM.

To determine the weight *(ai)* in MRMM technique for each of the seven C3S models, we first assign a model reliability factor *Ri* to each one, following the method by Chen et al. (2011). This factor quantifies a model's skill in precipitation forecasting for the historical period of 1997-2016. The *Ri* for each model is calculated as a simple sum of its ranks based on the NSE metric. A higher *Ri* value indicates better forecasting skill. The final weight (*ai*) for each model is derived by normalizing its *Ri* value so that the sum of all seven weights equals 1.

$$R_i = \frac{\sum_{i=1}^{N} S_i}{S_i} \qquad (4)$$

$$W_i = \frac{R_i}{\sum_{i=1}^{N} R_i} \qquad (5)$$

We calculated a separate MRMM for each station, determining a unique weight for each model based on its performance at that specific station.

2.2.1. Extreme precipitation events

To assess the C3S models' ability to forecast extreme precipitation events, we selected 6 extreme precipitation indices defined by the CCl/WCRP-Clivar/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI; Alexander et al., 2006). Our process involved first extracting daily precipitation values for each station from the C3S datasets for lead times up to three months. Subsequently, these indices were calculated for all 148 station and compiled for further assessment (Table 2). Next, we used the NSE metric to compare observed extreme climate index values against C3S model predictions from 1997-2016. This allowed us to determine each model's overall skill at every station across Iran.

Table 2: Information for extreme precipitation indices defined by ETCCDI used in this study.

	Index	Name	Definition	Unit
1	CDD	Consecutive dry days	Maximum number of consecutive days with PR < 1 mm	Days
2	CWD	Consecutive wet days	Maximum number of consecutive days with PR >= 1 mm	Days
3	R10	Heavy precipitation days	No. of days with Precipitation >= 10 mm/day	Days
4	Rx1day	Max 1-day precipitation	Monthly maximum 1 day precipitation	mm
5	Rx5day	Max 5-day precipitation	Monthly maximum consecutive 5 days precipitation	mm
6	R95p	Very wet days	Annual precipitation total from events greater than 95th percentile of (1997-2016) daily precipitation.	mm

3. Results

3.1. Evaluation of Model Performance for Monthly Precipitation

This study evaluates the forecast capability of seasonal C3S and MME models for precipitation across 148 stations in Iran. Following, we analyze the relationship between these models and sea surface temperature (SST), a critical input known to drive seasonal forecast accuracy.

3.1.1. Multi-Model Ensemble (MME) Evaluation

This study not only investigates the individual performance of the C3S models but also examines an ensemble prediction approach, derived from these models, and its efficiency in seasonal precipitation forecasting in Iran. Combining multiple models into an ensemble leverages their individual strengths and mitigates the influence of their biases and errors (Mamgain et al., 2024). For seasonal precipitation forecasting in Iran, MME forecasts consistently showed superior performance, as indicated by their higher NSE values (Figure 2) across all lead times. At the first lead time (L1), the MME achieved an NSE of 0.55, significantly outperforming individual models like UKMO (best) and NCEP-CFSv2 (weakest, NSE 0.04). While still the most efficient for lead times 2 and 3, the ensemble's NSE slightly decreased (L2: 0.36, L3: 0.29). ECMWF and JMA were the best individual models for L2 and L3, respectively, but NCEP-CFSv2 remained consistently the weakest. While Pakdaman et al. (2022) reported improved performance from a European multi model ensemble approach using ECMWF, CMCC, Meteo-France, and DWD for monthly and seasonal precipitation forecasting, our study, utilizing seven C3S models and a weighted multi-model ensemble approach, further confirms the robustness and accuracy of the MME approach for seasonal precipitation

forecasting in Iran, even with increasing lead times. In line with previous studies (Xu et al., 2020; Manzanas et al., 2014), the overall results in this section indicate that multi-model ensembles are effective in enhancing the predictive skill of individual models.

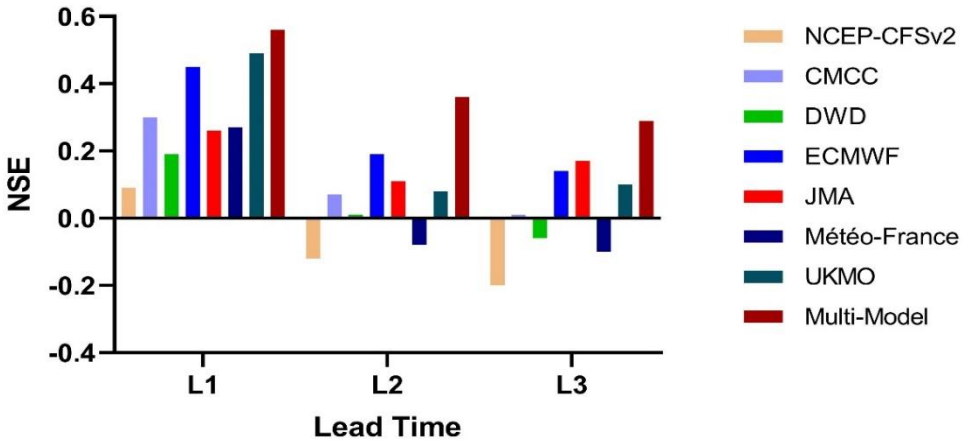


Figure 2: Average NSE scores for monthly precipitation forecasts over Iran (1997–2016). The figure compares the performance of individual C3S models and the Multi-Model Ensemble (MME) against observed monthly precipitation for one-, two-, and three-month lead times.

3.1.2. Skill of C3S Models in Seasonal Precipitation Forecasting

This section initially examines the models' skill in forecasting seasonal precipitation over Iran up to three month lead time. Model performance is assessed by computing the Nash-Sutcliffe Efficiency (NSE) between model outputs and observed data. Figure 3 illustrates the spatial distribution of annual precipitation, with values derived from the average C3S model forecasts across one-, two-, and three-month lead times for the 1997–2016 period. All models generally agree that maximum precipitation in Iran occurs along the southwestern coast of the Caspian Sea. While the models' absolute precipitation outputs diverge considerably from actual observations, they nevertheless consistently estimate this particular region of Iran as having the highest rainfall in terms of spatial distribution. Similar to observed values, the models estimate the lowest precipitation (<100 mm) in central, eastern, and southeastern Iran. It's important to note that specific regions within the middle Zagros Mountains record annual precipitation exceeding 1000 mm. However, the C3S models consistently underestimate these precipitation values when compared to observational data for these areas.

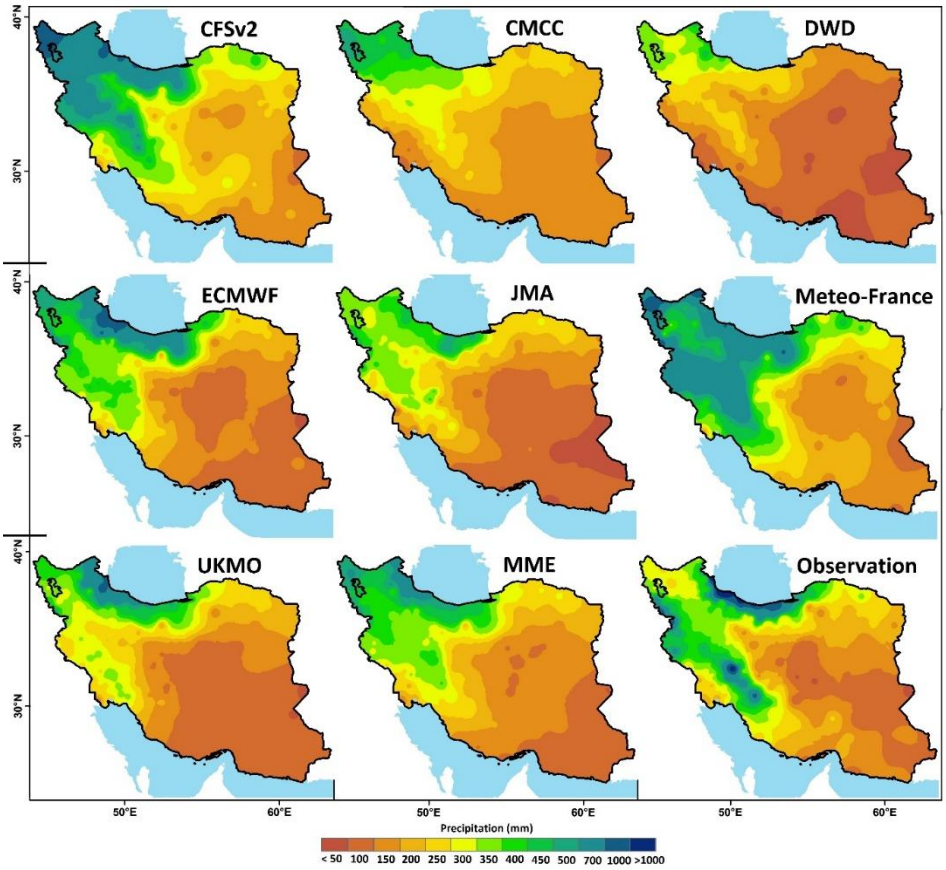


Figure 3: Spatial distribution of annual precipitation over Iran: Observed versus C3S model outputs. Model precipitation values derived from the average forecasted precipitation from C3S models, considering all three lead times during 1997-2016.

To quantify the accuracy of the C3S models, the NSE index was employed. Figure 4 shows the spatial distribution of the NSE values for precipitation forecasts with a one-month lead time. Overall, the ensemble model demonstrates the highest performance for precipitation prediction, consistently outperforming the individual models. However, at a one-month lead time, MME and all individual models showed a common tendency toward lower forecast accuracy in the southern coastal regions along the Sea of Oman and the Persian Gulf, as well as in the northwest of Iran. The CFSv2 model demonstrates the highest efficiency in the southern Zagros in the southwest of Iran, as well as in the northeastern regions of Iran. The CMCC model shows the highest efficiency in the west and northeast of Iran. Although the DWD model had more stations with NSE values in the 0.1 to 0.25 range compared to the two previous models that demonstrated its higher overall efficacy, similarly its highest efficiency was likewise found in the northeastern and southwestern parts of Iran. Notably, all three mentioned models do not demonstrate adequate efficiency in the Caspian Sea coasts, which is one of the wettest regions in Iran. The ECMWF model demonstrated greater efficiency in forecasting precipitation at the

one-month lead time than the three preceding models. A total of 33 stations showed NSE values above 0.4 for this model. These stations are mainly concentrated in parts of western Iran and the central Zagros in the southwest, as well as in parts of eastern and northeastern Iran. Furthermore, it demonstrates reasonably good performance in the southern highlands of central Iran and along the Caspian Sea coast.

The JMA model achieves its highest efficiency in the northern Zagros along the western border of Iran and, furthermore, shows acceptable efficiency along the eastern border. Across the majority of synoptic stations, the Meteo-France model demonstrates low efficiency (NSE < 0.1) for lead time-1, indicating its limited utility for most of the country. However, at 15 stations, primarily located along the western border, in the southern Zagros (southwest), and in the northeast, the model achieves higher NSE values above 0.4. Like the ECMWF model, the UKMO model also shows reasonable efficiency in forecasting one-month lead time precipitation across Iran. This model has 25 stations with NSE values above 0.4, primarily situated in the western, southwestern, eastern, and northeastern regions of the country. This model, along with ECMWF, shows the highest efficiency along both the northern and southern coasts of the country, and their results can be more reliably used for precipitation forecasting at lead time-1 compared to other C3S models. Finally, it should be noted that the CMCC, DWD, and JMA models show higher efficiency in the southern slopes of the Alborz Mountains compared to other models. Therefore, these models can be utilized for monthly precipitation forecasting in these areas.

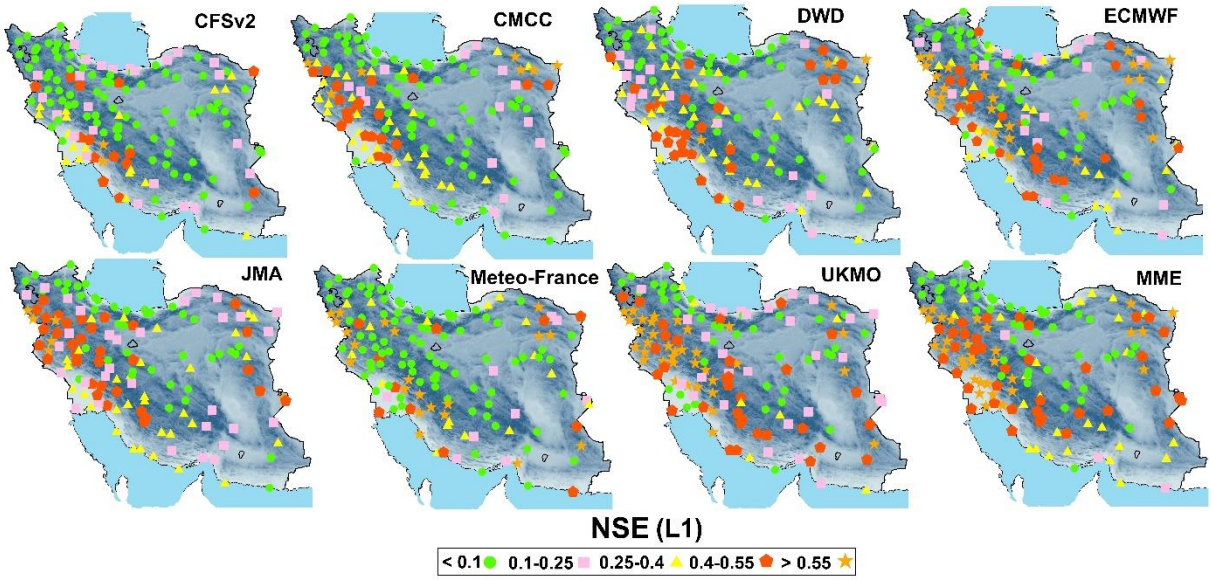


Figure 4: Spatial distribution of the NSE score for the MME and seven individual C3S models at a one-month lead time, based on 148 synoptic stations across Iran (1997–2016).

Figure 5 shows the models efficacy at the second lead time. The considerable drop in precipitation forecasting accuracy at this lead time across all models is evident, with only 13 stations nationally recording NSE values above 0.4. Among these few higher-performing stations, 7 are associated with the ECMWF model. For the MME model, 9 stations achieved NSE scores greater than 0.4, indicating the superior performance of the ensemble model compared to individual models. The CMCC, DWD, and JMA models demonstrate moderate skill in the west to southwest and northeast. Specifically, DWD's performance is weaker than CMCC's in the western region, while JMA performs best at the eastern stations. The JMA model exhibits the highest efficiency in the eastern stations compared to other models. As mentioned earlier, in the second lead time, the ECMWF model demonstrates the best performance in across most regions of the country, with NSE values primarily ranging between 0.25 and 0.4, except in the northern, northwestern, and southeastern parts of Iran. Météo-France shows weak performance with NSE < 0.1 in 128 stations. UKMO also shows reduced skill, with NSE > 0.25 in only 10 stations. Overall, based on the NSE values across the station network, the ECMWF model more effectively forecasts precipitation at the second lead time than the other models.

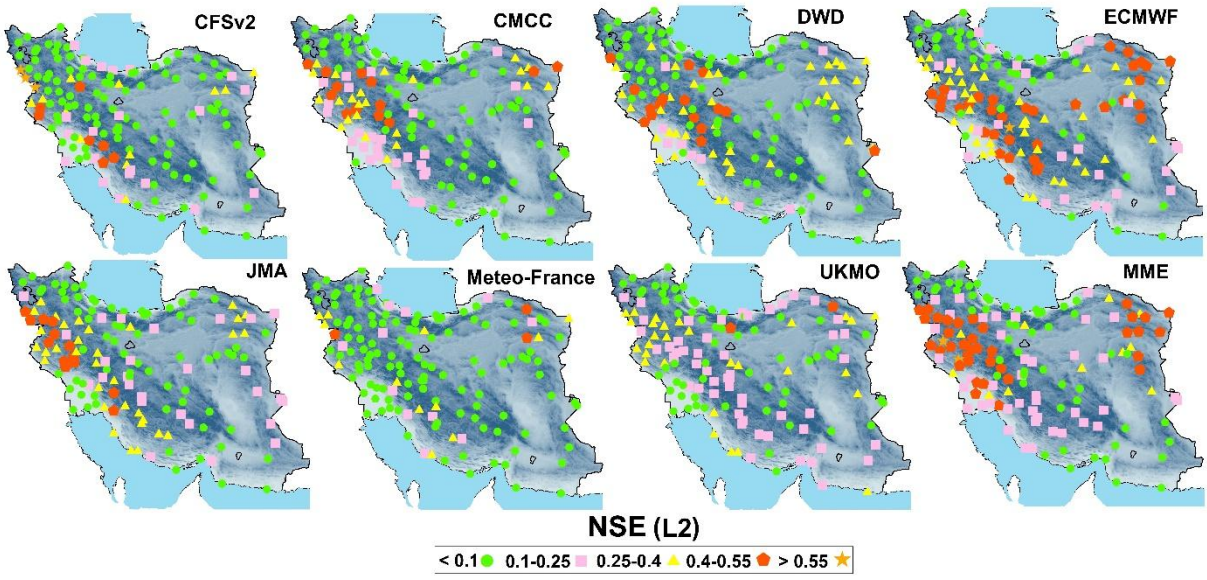


Figure 5: Caption is similar to figure 4, but for two-month lead time.

As anticipated for the third lead time, a further decline in model performance was observed (Figure 6). For the CFSv2 model, ~90% of the stations recorded NSE values below 0.1. The CMCC model also showed a substantial number of stations (100, or 67%) below this 0.1 NSE mark. CMCC performs slightly better in western, southwestern, and northwestern areas than CFSv2. The DWD model exhibits poor performance, with over 85% of stations showing NSE

values below 0.1. Higher scores are found only in the western and southern fringes and parts of the northeast. The Meteo-France model, similar to the DWD, displayed low efficiency throughout much of Iran, but performed better along the western border. The ECMWF, JMA, and UKMO models exhibited higher efficiency across most of Iran compared to the other models. The percentage of stations with NSE > 0.25 is 20% (ECMWF), 19% (JMA), and 12% UKMO. Only the ECMWF and JMA models achieved NSE values exceeding 0.4 at a small percentage of stations, specifically 8% and 2% respectively. Overall, the model's forecasts for the third lead time were generally unreliable.

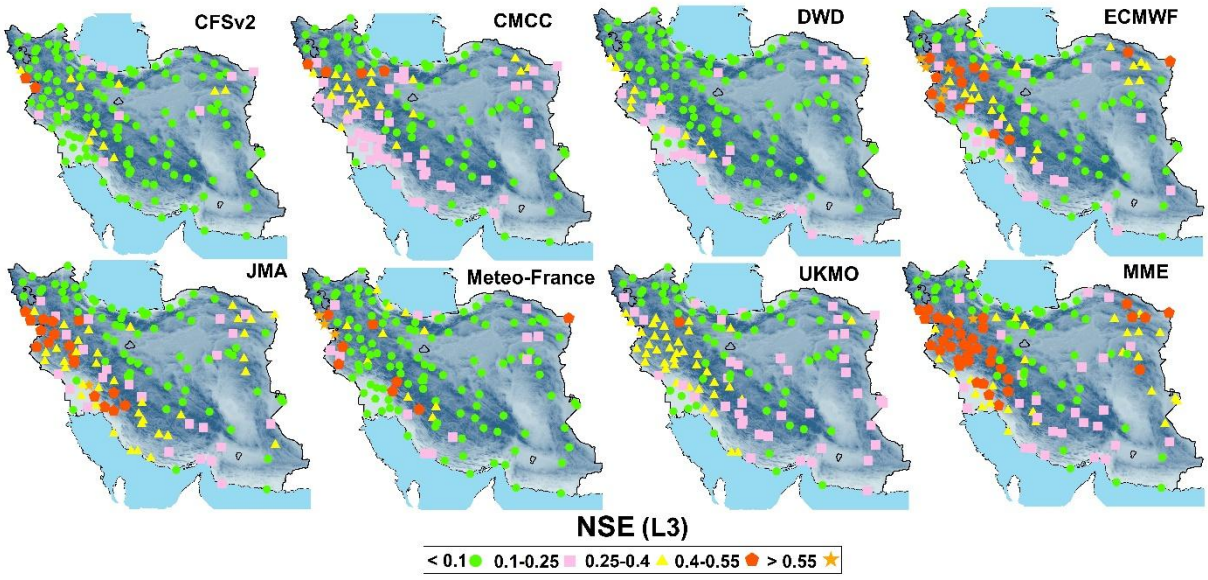


Figure 6: Caption is similar to figure 4, but for three-month lead time.

Generally, the models exhibit enhanced skill in predicting precipitation across Iran's western regions, consistent across all three lead times. Precipitation in these areas primarily occurs during the colder months (October to April) and is largely influenced by Mediterranean and Sudanese low-pressure systems. This higher accuracy is likely due to the models' greater efficiency in forecasting these systems and their associated moisture transport into Iran. Gebrechorkos et al. (2022) found that C3S models, particularly ECMWF and UKMO, demonstrate strong skill in predicting monthly and seasonal precipitation in East Africa. This region is notably where Sudanese low-pressure systems originate, which are significant drivers of precipitation in areas like western Iran. Further, Giuntoli et al. (2022) linked the models' effectiveness in forecasting Mediterranean precipitation regimes to ENSO phases, noting higher accuracy in regions sensitive to these phases. This aligns with findings by Dehghani et al. (2020), who demonstrated a statistical relationship between autumn and winter precipitation in western and southwestern Iran and ENSO signals, suggesting that ENSO cycles can significantly influence precipitation patterns in this region. Although Shirvani et al. (2022)

demonstrated that ENSO is not the only source of skill in seasonal forecast models, their influence remains a key factor.

In contrast, the models show lower efficiency in the northern and northwestern regions of Iran. Precipitation in these areas relies on small-scale and convective systems, indicating a general limitation of the models in simulating precipitation influenced by such localized phenomena. This diminished predictive capability for small-scale precipitation is likewise observed in southeastern Iran, a region under the influence of monsoon systems. This finding aligns with research by Najafi and Kuchak (2024), who demonstrated that even regional seasonal precipitation prediction models are less effective at forecasting small-scale precipitation events.

One of the most crucial predictors in seasonal forecasts is SST anomalies. This is because seasonal prediction is built upon changes in the probability of weather events, and these changes themselves are a result of slowly varying forcing, such as SST anomalies. SST patterns in the global oceans, for instance, have proven to be an influential tool for predicting regional precipitation variations (L'Heureux et al, 2015; Gibson et al, 2021). So, precipitation predictability at seasonal timescales is primarily linked to large-scale variability in sea surface temperature (SST) (Doblas-Reyes et al, 2013).

This study will first analyze eight SSTDT indices and their association with precipitation in Iran. We will then assess the statistical relationship between the precipitation forecasts for Iran from each C3S model and these indices. The main objective of this section is to statistically examine the relationship between one of the most crucial input factors for seasonal prediction models (i.e., SSTDT) and precipitation in Iran. Subsequently, we will investigate the relationship between these indices and the precipitation predicted by the models, aiming to determine which SSTDT has a stronger statistical association with the predicted precipitation for each model and lead time. Naturally, if models can predict the relationship between forecasted precipitation and SSTDT to be similar or close to what's observed in real data, they will achieve higher accuracy in precipitation forecasting. In this section, observed and model-predicted monthly precipitation data were examined alongside monthly index values during the 1997-2016 period.

Figure 7 reveals that three SSTDT indices—Nino1+2, AMO, and PDO—exhibit a statistically significant relationship with monthly precipitation variations across Iran. Specifically, Nino1+2 and AMO show significance at the $p<0.01$ level, while PDO is significant at the $p<0.05$ level. Of these, Nino1+2 and PDO demonstrate a positive correlation with precipitation, whereas AMO shows a negative correlation. Additionally, although Nino3 and ONI also display a

positive correlation with precipitation across Iran, their correlation are not statistically significant.

Models that accurately simulate the correlation between predicted precipitation and these SSTDT indices, consistent with actual observational data from Iran, will be significantly more effective for seasonal precipitation forecasting across the country. Our evaluation of C3S models revealed that UKMO and ECMWF are particularly effective at the first lead time, and UKMO is effective for the second lead time. These models demonstrate a superior ability to simulate the connection between SSTDT indices and precipitation, showing strong alignment with observational data. While other models show behaviors that are roughly consistent with observations, they fall short in a crucial area: the varying strength of their correlation coefficients. This variability prevents them from accurately representing the true impact of these indices on precipitation forecasting within each model. Consequently, this difference in correlation strength fundamentally alters how these models simulate SSTDT and their relationship with precipitation in Iran. For the third lead time, all models demonstrated an inability to effectively capture the relationship between precipitation and SSTDT across Iran. This could be a contributing factor to the decrease in model efficiency as lead time increases.

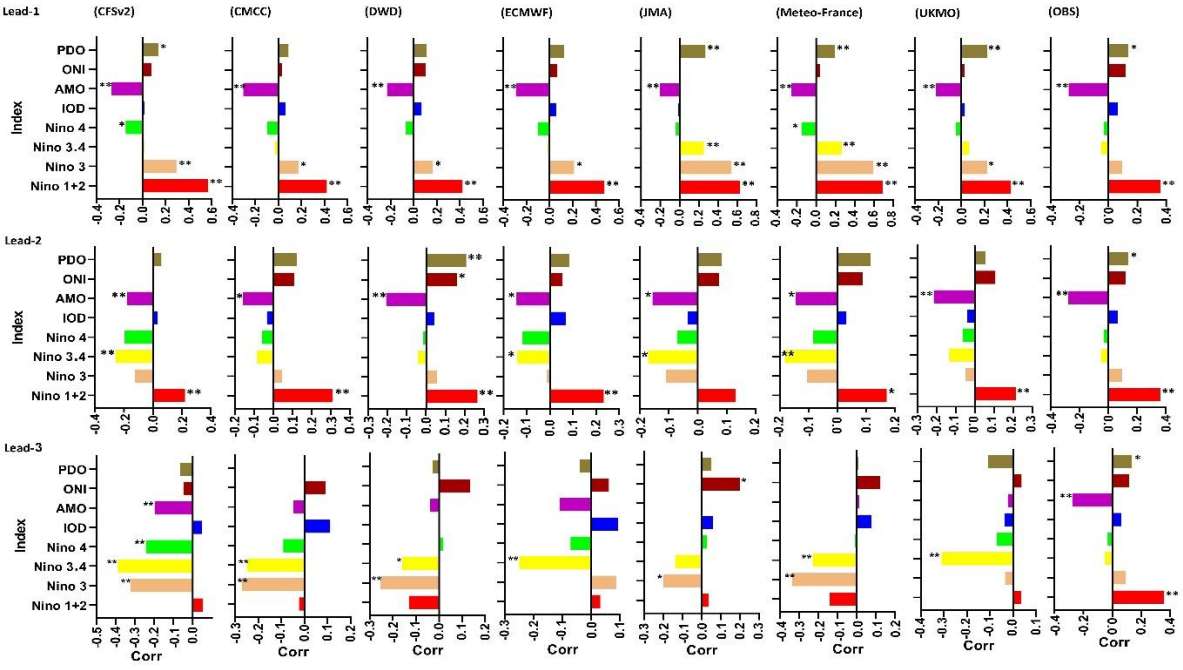


Figure 7: The Corr between SSTDT indices and Iran's monthly precipitation. The figure compares observed data (OBS) with forecasts from C3S models for 1997–2016, with a lead time of 1 to 3 months (Lead-1 to Lead-3). Significant correlations are indicated by a single asterisk for the 0.05% level and a double asterisk for the 0.01% level.

As previous studies (e.g., Kowal et al., 2023; Nobakht et al., 2021) indicated that model performance varied significantly across different climatic zones. To comprehensively assess

the variability in model performance across Iran's diverse climatic zones, we conducted an analysis of model efficiency within six distinct geographical regions spanning the country (Figure 8). This investigation, informed by the climatic zone delineations established by Najafi et al. (2025), focused on evaluating the average model performance across three different lead times.

In the Caspian Sea coastal areas of northern Iran (R1), a region defined by its humid and temperate climate, we didn't find any single model to be significantly better than the others. However, the MME and UKMO models did show slightly higher median correlations in this area. Specifically, the MME achieved the highest performance with an average correlation of 0.44. Among the individual C3S models, UKMO demonstrated the best performance, with an average correlation of 0.39. In contrast, other C3S models—including CFSv2 (Corr=0.32), CMCC (Corr=0.319), Meteo-France (Corr=0.317), JMA (Corr=0.26), DWD (Corr=0.24), and ECMWF (Corr=0.21)—showed considerably lower average correlation coefficients. The notably poor performance of the ECMWF model in this region is particularly striking, suggesting its limited suitability for the humid and temperate coastal climate of Iran. The boxes in this region are moderately wide, suggesting a moderate level of variability in the correlation coefficients for individual models. While the median correlations aren't particularly high, the spread around them is somewhat contained for the middle 50% of cases. The models performed poorly in the northern parts of Iran. This aligns with Nobakht et al.'s (2021) research, which also found that ECMWF, UKMO, and Meteo-France models showed their poorest performance in Northern Iran. Based on research by Shestakova and Toropov (2021), large-scale atmospheric systems, including weather fronts and pressure systems, are responsible for 50% of the precipitation in this region. The other half of the precipitation is driven by regional factors, primarily the combined influence of the Caspian Sea's lake effect and the local topography (orography). The simultaneous interaction of these two elements significantly boosts the total precipitation in the area. This is a key challenge for large-scale climate models, such as C3S models, which often have difficulty accurately simulating these regional factors. Their coarse resolution is not fine enough to capture the complex, smaller-scale details that are crucial for predicting local weather and precipitation with precision.

In the western parts (R2), characterized by a Mediterranean climate (cold, wet winters and warm, dry summers), the MME exhibited the highest efficacy in precipitation forecasting, with an average correlation of 0.64. Among the individual C3S models, both UKMO and ECMWF demonstrated the strongest performance, showing average correlations of 0.57 and 0.53,

respectively. Conversely, the DWD model yielded the weakest performance in this region, with an average correlation coefficient of 0.45.

Overall, the correlation coefficients for models in this region are notably higher than those observed in northern Iran, indicating a generally enhanced forecasting capability of C3S models within western Iran. Furthermore, the interquartile ranges (IQRs) of the box plots for this region were generally narrower compared to the Caspian Sea region, particularly for higher-performing models such as UKMO and MME, with IQRs frequently less than 0.1. This narrower spread suggests greater consistency in the models' performance within this Mediterranean climatic zone. These findings are further supported by previous research. Alizadeh et al. (2019) demonstrated that CFSv2 model predictions exhibit higher efficacy in western Iran compared to other regions. Similarly, Nobakht et al. (2021) concluded that forecast models generally show superior performance in the western regions of Iran.

In the northwestern parts of Iran (R3), characterized by a semi-arid climate with mild summers and cold winters, the models generally exhibit moderate correlation values, predominantly ranging from 0.3 to 0.6. The MME shows the highest performance with an average correlation coefficient of 0.54. Notably, in this region, the JMA, CMCC, and Meteo-France models, which had previously demonstrated weaker performance, showed better efficacy. Conversely, the UKMO model, despite its strong performance in other regions, exhibited comparatively lower efficacy than the aforementioned three models in this region. Narrower IQRs indicate good consistency in forecast skill within this climate zone.

In the eastern, southeastern, and central parts of Iran (R4), characterized by an extremely arid and warm climate, models generally exhibit lower correlation coefficients compared to the western and northwestern regions. Most median correlations in these areas fall between 0.2 and 0.4, suggesting that predicting precipitation in these regions might be more challenging for the models. Within this extremely arid zone, the MME shows the highest efficacy, with an average correlation coefficient of 0.44. Among the individual C3S models, UKMO demonstrates the best performance in this region, while the other models show comparable, often lower, effectiveness. The larger IQRs highlight significant variability in model performance, indicating that this region appears to be more challenging for consistent prediction. While Yazdandoost et al. (2020) identified the CMCC and CFSv2 models (from the NMME series) as most suitable for precipitation forecasting in Eastern Iran's hydrological basins, our results indicate that these two models actually show lower performance compared to other C3S models we examined, with others providing more accurate predictions. This region is typically arid but

can experience summer monsoon systems, and overall, models tend to exhibit lower efficacy here than in other regions.

In the southwestern region and along the Persian Gulf coast (R5), an area characterized by a semi-arid and exceptionally hot climate, the MME demonstrated the highest performance in precipitation forecasting. Following the MME, the Meteo-France, JMA, UKMO, and ECMWF models exhibited comparable skill, with median correlations clustering around 0.43. Other models, such as CMCC and DWD, showed comparatively lower median correlation values. Analysis of the IQRs for this region revealed varying degrees of consistency among the models. While certain models displayed reasonable reliability in their predictions, others exhibited greater variability in their predictive skill within this challenging climatic zone.

In northeastern Iran (R6), a region defined by its arid to semi-arid climate with warm summers and cold winters, the models generally exhibit moderately good correlations, typically ranging from 0.4 to 0.6. Both the MME and UKMO consistently show higher median correlations in this region. The box plots for this region are characteristically moderate to narrow, similar to those observed in the northwest. Their IQRs frequently fall between 0.1 and 0.15, which suggests a relatively consistent performance among the models in the northeast.

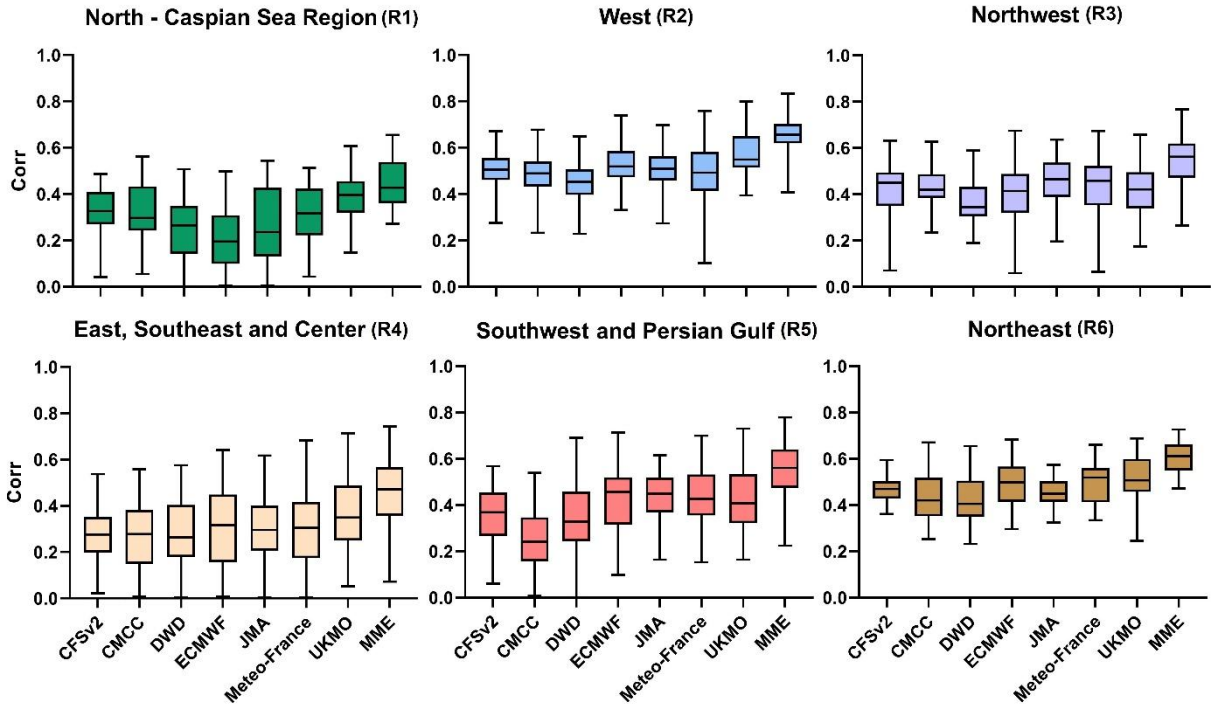


Figure 8: Boxplots of the Corr between observed and forecasted monthly precipitation across six distinct climate zones in Iran. Each subplot corresponds to a different region (R1-R6), and the boxplots compare the performance of eight forecasting models (CFSv2, CMCC, DWD, JMA, ECMWF, Meteo-France, UKMO) and MME. The values represent the average performance across three different lead times.

4.2 Efficiency of C3S models in forecasting monthly precipitation

In this section, the correlations were calculated on a monthly basis for each station in the 1997–2016 time series (Figure 9). This approach, which allows for the analysis of seasonal variability, differs from the previous section where a single correlation was computed for the entire period without separating the months. At the L1, all models exhibit their highest forecast skill. The MME consistently outperforms individual C3S forecasting systems across all regions, with average correlation values ranging from 0.419 in R4 to 0.564 in R6. Peak MME performance varies monthly and regionally, with maxima of 0.661 (October, R1), 0.641 (January, R2), 0.654 (January, R3), 0.542 (April, R4), 0.720 (May, R5), and 0.723 (February, R6). The model's lowest efficiency occurs during the summer, with scores of 0.351 in June and 0.176 in August. This happens because most parts of Iran receive little to no rainfall during the summer. While the actual precipitation is zero, the model often estimates a small, non-zero value. This discrepancy between the model's prediction and the observed data (zero) leads to a significant drop in its efficiency during these dry months.

The UKMO system stands out as the best-performing individual model in regions R1, R2, and R6. Its forecasts are particularly strong during late winter and autumn, achieving high correlations from October to November. However, performance declines sharply in summer (e.g., <0.21 in July, R5). Our results are consistent with a previous study by Manzanas et al. (2014), which demonstrated that global ensemble seasonal forecasting models—such as those that combine outputs from systems like ECMWF, UKMO, and IFM-GEOMAR—typically exhibit weak performance during the summer months (June to August). This finding aligns with a similar study by Shirvani and Landman (2016), who used different models for Iran and found that the highest performance was achieved in the fall season. They also noted that winter skill was similar to that of the spring, but with lower overall scores.

ECMWF and DWD are competitive in some regions: ECMWF achieves an average of 0.452 in R1 and 0.501 in R6, with correlations exceeding 0.60 in spring and autumn. DWD demonstrates high skill in winter and early spring (e.g., >0.55 in April, R1), but generally declines to below 0.30 during August. In contrast, JMA and CMCC display weaker and more erratic performance; JMA's correlations drop below 0.20 during summer months (e.g., 0.13 in R2, August), while CMCC falls below 0.17 across R4–R6 during June–July. Spatially, regions R2, and R6 exhibit the highest predictive skill and largest seasonal amplitude (MME range ~0.30), whereas R1 and R5 demonstrate both lower average skill and reduced seasonal variation (MME range ~0.20).

At the L2, forecast skill declines by approximately 0.10–0.15 correlation points. Nevertheless, the overall seasonal pattern—characterized by winter and autumn peaks and summer minima—

remains evident. MME continues to best in all regions, with average correlation values between 0.316 (R4) and 0.477 (R6). Monthly maxima range from 0.637 (August, R1) to 0.672 (July, R6), while minima fall as low as 0.061 (February, R3) and 0.088 (August, R5). The model's efficiency for a two-month lead time increases during the summer, which suggests that the models are more reliable for longer lead times during the dry season.

ECMWF ranks second in R1 (average 0.321) and R6 (0.402), with peak skill during spring (~0.33 in April, R1) and autumn (>0.54 in October). UKMO performs competitively in R1–R3 (e.g., ~0.338 in R1), particularly in late summer (e.g., ~0.655 in August, R1), but shows marked degradation during winter in R4–R6 (<0.24 in January). DWD maintains moderate winter skill (January ≥0.33), yet its performance declines across most regions during spring (<0.25). JMA and CMCC remain consistently low-performing, seldom exceeding 0.30 during summer months and performing particularly poorly in R5 and R6. Regionally, R6 continues to demonstrate the highest overall predictability (MME >0.45 in most months), whereas R1, R4 and R5 consistently record the lowest skill, with year-round MME values remaining below 0.34.

At the L3, the monthly forecast skill declines further across all systems. MME correlations now average between 0.086 (R4) and 0.461 (R1), and the primary predictability window shifts toward the summer and early autumn months. Maximum MME correlations occur in July (0.691, R1), November (0.535, R3), and March (0.556, R6), while minimum values drop to 0.070 (July, R5) and 0.086 (December, R6).

ECMWF remains the most skillful individual system in R1 (0.369) and R3 (0.324), with peak performance in midsummer (e.g., >0.63 in July, R1). However, its skill drops near 0.20 in late spring. CFSv2 provides the best winter forecasts in R4 and R5 (e.g., ~0.337 and 0.373 in January, respectively). CMCC exhibits the lowest overall performance, with regional means as low as 0.146 (R6) and 0.165 (R2). UKMO shows moderate skill during summer but declines to <0.20 in spring across multiple regions. In terms of spatial variability, R1 and R6 retain relatively higher skill (MME averages of 0.409 and 0.400, respectively), while R4 and R5 remain the most challenging forecast domains (MME ~0.26–0.29). These results suggest that summer and early autumn (JJA–SON) provide the most favorable window for extended-range prediction at this lead time, while winter retains marginal skill above the spring baseline. Spring—particularly May—emerges as the least predictable season across all models and regions.

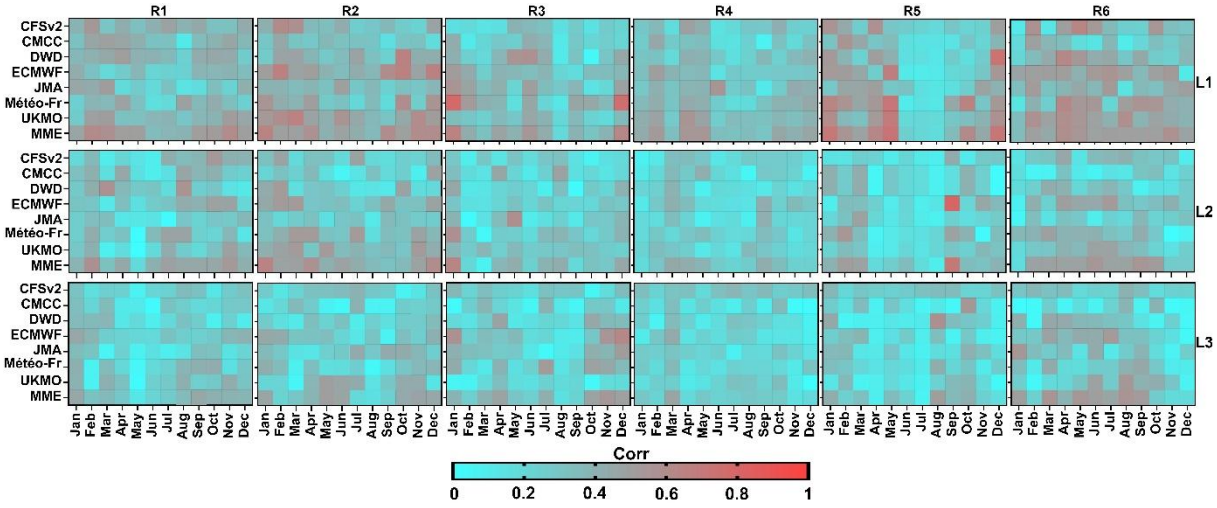


Figure 9: Heatmap of the Corr between observed and forecasted average monthly precipitation across six distinct climate zones in Iran (R1–R6). The rows represent three different lead times (L1, L2, L3), while the columns show the monthly performance of individual C3S models and the MME. The color scale at the bottom indicates the correlation value, with warmer colors (red) representing a higher correlation and cooler colors (turquoise) representing a lower correlation.

4.3 Model Skill in Forecasting Extreme Precipitation Indices

The subsequent analysis focuses on evaluating the efficacy of the seven examined models in forecasting precipitation extreme indices across Iran. In this section, the performance of the seven C3S and MME models under investigation was assessed using the average correlation coefficient, computed across three lead times and six precipitation extreme indices. Figure 10 illustrates the spatial variability of model performance across Iran, showing how the efficiency of forecasting extreme precipitation events differs by location. Overall, the MME model outperformed other models in predicting extreme precipitation. For the CDD index, the ECMWF model demonstrates the highest forecast efficiency, yielding an average Corr of 0.34 across the three lead times. The UKMO model follows, with an average Corr of 0.29. The remaining five models exhibit comparable performance, with average Corr approximating 0.2. Notably, most models display greater efficiency in forecasting this index in the western part of the country. Concerning the CWD index, the ECMWF model demonstrates superior predictive capability, registering an average Corr of 0.24. Spatially, model performance for this index tends to be higher in the western regions of the country. Furthermore, the ECMWF and UKMO models exhibit noteworthy forecast skill at stations located in the northeastern sector.

Regarding the Rx1Day index, the UKMO model exhibits the highest efficiency, yielding an average Corr of 0.2 across the three considered lead times. The performance of the other models is comparable,

with the highest correlation values generally observed at stations in the western part of the country. However, the CMCC model demonstrates notably weaker performance compared to the others, even in the western regions. For this index, the Météo-France model shows acceptable forecast skill in the northeastern part of the country. For the Rx5Day index, the ECMWF and UKMO models exhibit the highest efficiency, with Corr of 0.21 and 0.204, respectively. Notably, the CFSv2, CMCC, and DWD models did not demonstrate acceptable efficiency at any station across the country. Conversely, the Météo-France model shows superior performance compared to the other models in the east of the country.

Another index examined is the frequency of days with precipitation exceeding 10 mm (R10mm). For this index, the UK-Met model exhibits the highest average Corr (0.22). The UKMO model demonstrates more acceptable forecasting performance for this threshold in the western, southwestern, and northeastern regions of the country relative to the other models. However, in the western regions, the DWD, CMCC, and ECMWF models exhibit comparable predictive skill to the UKMO model. For the R95p index, the ECMWF and UKMO models demonstrate the highest efficiency in forecasting this extreme index, both with a Corr of 0.23. The Météo-France model also shows acceptable performance, particularly in the western part, while the efficiency of other models is not acceptable. Overall, the ECMWF and UKMO models show good efficiency, especially for the western regions and some parts of the northeast, and can be used for forecasting the extreme index. In other regions, none of the seven examined models are recommended for predicting extreme precipitation indices.

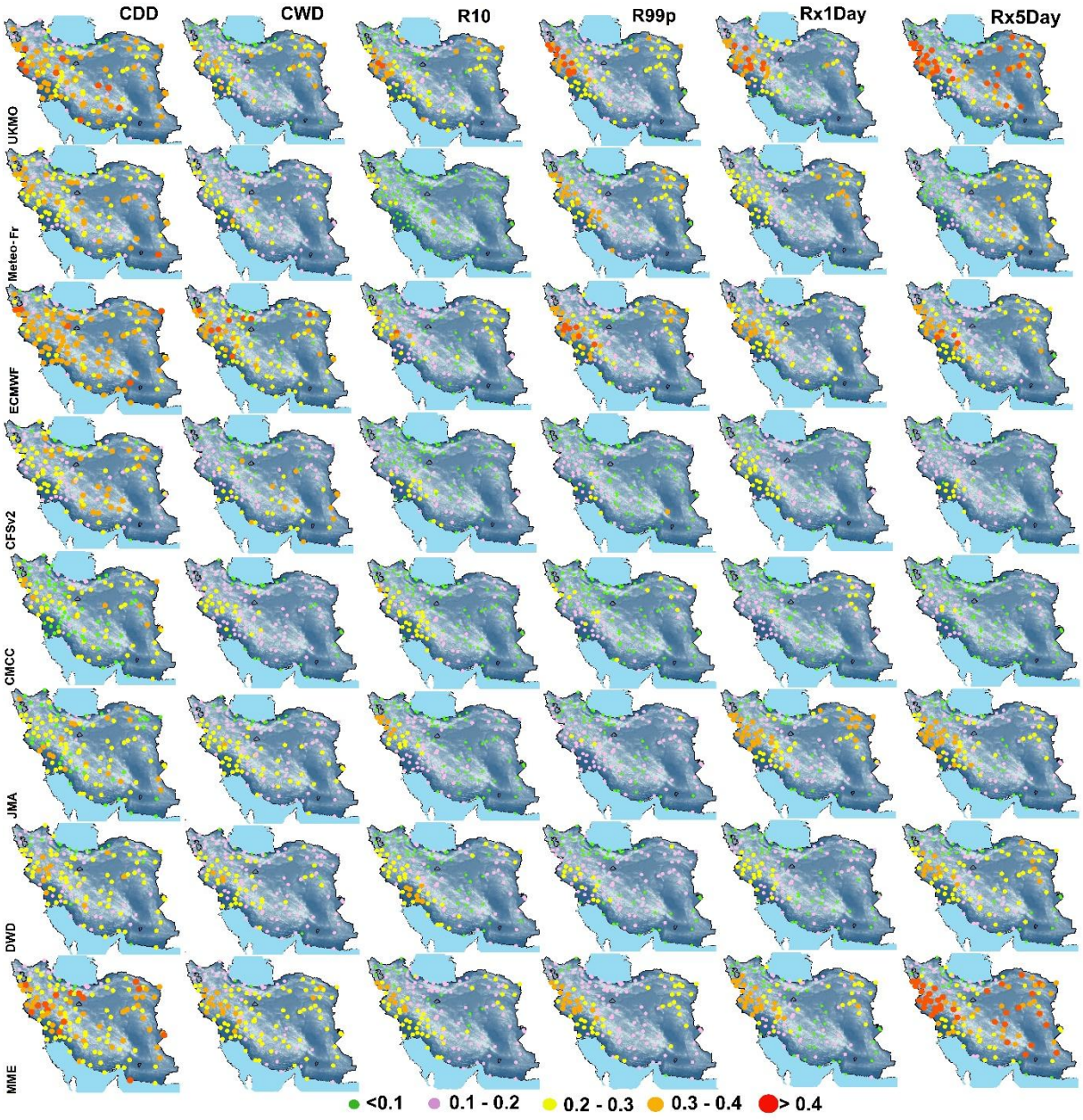


Figure 10: Spatial distribution of the average correlation between observed and forecasted extreme precipitation indices across 148 synoptic stations in Iran. Each row represents a different forecasting model (UKMO, Meteo-Fr, ECMWF, CFSv2, CMCC, JMA, DWD, and the MME), while the columns show the results for six extreme precipitation indices (CDD, CWD, R10, R99p, Rx1Day, Rx5Day). The correlations displayed are averaged over three lead times.

For the R95p index, the ECMWF and UKMO models demonstrate the highest efficiency in predicting this extreme index, both with a correlation coefficient of 0.23. The Météo-France model also shows acceptable performance for this index, particularly in the western part of the country, while the efficiency of other models is not acceptable for the R95p index in any region of the country. Overall, the ECMWF and UKMO models show good efficiency, especially for the western regions and parts of

the northeast, and can be used for predicting the extreme index. In other regions of the country, none of the seven examined models are recommended for predicting extreme precipitation indices.

4. Conclusion

This study evaluated the performance of seven global seasonal forecasting models from the Copernicus Climate Change Service (C3S)—CFSv2, CMCC, DWD, ECMWF, JMA, Météo-France, and UKMO— as well as a multi-model ensemble (MME) in predicting precipitation across Iran during 2000–2019 using observational data from 148 synoptic stations. The primary aim was to identify the most reliable models for seasonal precipitation forecasting to support water resource management.

The main findings indicate that model skill varies substantially across different climatic regions and lead times. Overall, the MME had better skill than the individual models at all three lead times. CFSv2 showed better performance in southern Zagros and northeastern Iran, CMCC demonstrated skill in western and northeastern areas, and JMA performed well in northern Zagros and eastern margins. Most models underperformed in the southern coastal and northwestern regions. ECMWF and UKMO generally exhibited higher skill, particularly at shorter lead times, with forecast accuracy declining as lead time increased. Seasonally, ECMWF excelled in winter, DWD and UKMO in spring, JMA and ECMWF in summer, and UKMO in autumn.

Regarding extreme precipitation indices, the MME generally performed best at forecasting extreme precipitation in Iran. The ECMWF and UKMO models were the most reliable individual models, consistently showing high accuracy, especially in the western and northeastern parts of the country. In contrast, the CFSv2, CMCC, and DWD models were generally less effective. Ultimately, the effectiveness of any C3S model is highly dependent on both the specific extreme precipitation index being forecast and the climate region. For many parts of the country (e.g., northern parts with humid climate or southeastern parts with hyper-arid climate), none of the seven models were found to be reliable enough for forecasting extreme precipitation.

In conclusion, while raw outputs from some C3S models provide valuable guidance for precipitation forecasting, spatial and temporal variability in model skill remains. Future improvements are likely to arise from combining dynamical forecasts with statistical post-processing (e.g., deep-learning methods) and ensemble techniques, particularly for medium-range seasonal predictions in complex terrains.

References

- Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, … Vazquez-Aguirre JL (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res Atmos 111(D5). https://doi.org/10.1029/2005JD006290
- Alexander S, Block P (2022) Integration of seasonal precipitation forecast information into local-level agricultural decision-making using an agent-based model to support community adaptation. Clim. Risk Manag 36:100417. https://doi.org/10.1016/j.crm.2022.100417
- Alizadeh-Choobari O (2019) Dynamical downscaling of CSIRO-Mk3.6 seasonal forecasts over Iran with the regional climate model version 4. Int J Climatol 39:3313–3322. https://doi.org/10.1002/joc.6021
- Alizadeh-Choobari O, Qadimi M, Marjani S (2019) Evaluation of 2-m temperature and precipitation products of the Climate Forecast System version 2 over Iran. Dyn Atmos Oceans 88: 101105. https://doi.org/10.1016/j.dynatmoce.2019.101105
- Almazroui M, Khalid S, Kamil S, Ismail M, Islam MN, Saeed S, Abid MA, Ehsan MA, Hantoush AS (2022) Skill assessment of Saudi-KAU and C3S models in prediction of spring season rainfall over the Arabian Peninsula. Atmos Res 280: 106461. https://doi.org/10.1016/j.atmosres.2022.106461
- Calvo-Olivera C, Guerrero-Higueras ÁM, Lorenzana J, García-Ortega E (2024) Real-Time Evaluation of the Uncertainty in Weather Forecasts Through Machine Learning-Based Models. Water Resour Manage 38(7): 2455–2470. https://doi.org/10.1007/s11269-024-03779-y
- Chan KTF, Dong Z, Zheng M (2021) Statistical seasonal forecasting of tropical cyclones over the western North Pacific. Environ Res Lett 16(7): 074027. https://doi.org/10.1088/1748-9326/ac05f1
- Chen W, Jiang Z, Li L (2011) Probabilistic projections of climate change over China under the SRES A1B scenario using 28 AOGCMs. J Clim 24(17): 4741-4756. https://doi.org/10.1175/2011JCLI4102.1
- Dehghani M, Salehi S, Mosavi A, Nabipour N, Shamshirband S, Ghamisi P (2020) Spatial Analysis of Seasonal Precipitation over Iran: Co-Variation with Climate Indices. ISPRS Int J Geo-Inf 9(2): 73 https://doi.org/10.3390/ijgi9020073
- Doblas-Reyes FJ, García-Serrano J, Lienert F, Biescas AP, Rodrigues LR (2013) Seasonal climate predictability and forecasting: status and prospects. Wiley Interdiscip Rev Clim Change 4(4): 245-268. https://doi.org/10.1002/wcc.217
- Doblas-Reyes FJ, Hagedorn R, Palmer TN (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. Tellus A: Dyn Meteorol Oceanogr 57(3): 234. https://doi.org/10.3402/tellusa.v57i3.14658
- Gebrechorkos SH, Pan M, Beck HE, Sheffield J (2022). Performance of State-of-the-Art C3S European Seasonal Climate Forecast Models for Mean and Extreme Precipitation Over Africa. Water Resour Res 58(3): e2021WR031480. https://doi.org/10.1029/2021WR031480
- Gibson PB, Chapman WE, Altinok A, Delle Monache L, DeFlorio MJ, Waliser DE (2021) Training machine learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. Commun Earth Environ 2(1): 159. https://doi.org/10.1038/s43247-021-00225-4
- Giuntoli I, Fabiano F, Corti S (2022) Seasonal predictability of Mediterranean weather regimes in the Copernicus C3S systems. Clim Dyn 58(7): 2131-2147. https://doi.org/10.1007/s00382-021-05681-4
- Han T, Wang H, Hao X, Li S (2019) Seasonal Prediction of Midsummer Extreme Precipitation Days over Northeast China. J Appl Meteorol Climatol 58(9): 2033–2048. https://doi.org/10.1175/JAMC-D-18-0253.1
- Kowal KM, Slater LJ, Garcia Lopez A, Van Loon AF (2023) A comparison of seasonal rainfall forecasts over Central America using dynamic and hybrid approaches from Copernicus Climate Change Service seasonal forecasting system and the North American Multimodel Ensemble. Int J Climatol 43(5): 2175-2199. https://doi.org/10.1002/joc.7969
- Krishnamurthy V (2019) Predictability of Weather and Climate. Earth Space Sci 6(7): 1043–1056. https://doi.org/10.1029/2019EA000586

- Kundzewicz ZW, Kaczmarek Z (2000) Coping with hydrological extremes. Water Int 25(1), 66–75. https://doi.org/10.1080/02508060008686798
- L'Heureux ML, Tippett MK, Barnston AG (2015) Characterizing ENSO coupled variability and its impact on North American seasonal precipitation and temperature. J Clim. 28, 4231–4245. https://doi.org/10.1175/JCLI-D-14-00508.1
- Li W, Duan Q, Ye A, Miao C (2019) An improved meta-Gaussian distribution model for post-processing of precipitation forecasts by censored maximum likelihood estimation. J Hydrol 574: 801–810. https://doi.org/10.1016/j.jhydrol.2019.04.073
- Li Y, Xü K, Wu Z, Zhu Z, Wang QJ (2023) A statistical–dynamical approach for probabilistic prediction of sub-seasonal precipitation anomalies over 17 hydroclimatic regions in China. Hydrol Earth Syst Sci 27(22): 4187–4203. https://doi.org/10.5194/hess-27-4187-2023
- Mamgain A, Prasad SK, Sarkar A, Shanker G, Dube A, Mitra AK (2024) Evaluating short-range forecasts of a 12 km global ensemble prediction system and a 4 km convection-permitting regional ensemble prediction system. Pure Appl Geophys 181(7): 2217-2241. https://doi.org/10.1007/s00024-024-03524-x
- Manzanas R, Frías MD, Cofiño AS, Gutiérrez JM (2014) Validation of 40 year multimodel seasonal precipitation forecasts: The role of ENSO on the global skill. J Geophys Res 119: 1708−1719, https://doi.org/10.1002/2013JD020680
- Mishra N, Prodhomme C, Guemas V (2019) Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe. Clim Dyn 52(7): 4207-4225. https://doi.org/10.1007/s00382-018-4404-z
- Najafi MS, Kuchak VS (2024) Ensemble-based monthly to seasonal precipitation forecasting for Iran using a regional weather model. Int J Clim 44(12): 4366-4387. https://doi.org/10.1002/joc.8586
- Najafi MS, Alizadeh O, Sauter T (2025) Evaluation of gridded precipitation datasets over Iran. J Hydrol Reg Stud 58: 102234. https://doi.org/10.1016/j.ejrh.2025.102234
- Najafi MS, Khoshakhllagh F, Zamanzadeh SM, Shirazi MH, Samadi M, Hajikhani S (2014) Characteristics of TSP Loads during the Middle East Springtime Dust Storm (MESDS) in Western Iran. Arab J Geosci, 7(12): 5367–5381. https://doi.org/10.1007/s12517-013-1086-z
- Nana HN, Tamoffo AT, Kaissassou S, Djiotang Tchotchou LA, Tanessong RS, Kamsu-Tamo PH, Kenfack K, Vondou DA (2024) Performance-based evaluation of NMME and C3S models in forecasting the June–August Central African rainfall under the influence of the South Atlantic Ocean Dipole. Int J Clim 44(7): 2462–2483. https://doi.org/10.1002/joc.8463
- Nasseri M, Schoups G, Taheri M (2025) Statistical Assessment and Augmentation of European Centre for Medium-Range Weather Forecasts Monthly Precipitation Forecast (SEASonal Prediction of Precipitation). Int J Clim 45(3): e8723. https://doi.org/10.1002/joc.8723
- Nikraftar Z, Mbuvha R, Sadegh M, Landman WA (2024) Impact-Based Skill Evaluation of Seasonal Precipitation Forecasts. Earth's Future 12(11): e2024EF004936. https://doi.org/10.1029/2024EF004936
- Nobakht M, Saghafian B, Aminyavari S (2021) Skill Assessment of Copernicus Climate Change Service Seasonal Ensemble Precipitation Forecasts over Iran. Adv Atmos Sci 38(3): 504–521. https://doi.org/10.1007/s00376-020-0025-7
- Pakdaman M, Babaeian I, Javanshiri Z, Falamarzi Y (2022) European multi model ensemble (EMME): a new approach for monthly forecast of precipitation. Water Resour Manage 36(2): 611-623. https://doi.org/10.1007/s11269-021-03042-8
- Pan X, Li X, Cheng G, Li H, He X (2015) Development and Evaluation of a River-Basin-Scale High Spatio-Temporal Precipitation Data Set Using the WRF Model: A Case Study of the Heihe River Basin. Remote Sens 7(7): 9230–9252. https://doi.org/10.3390/rs70709230
- Pepler AS, Díaz LB, Prodhomme C, Doblas-Reyes FJ, Kumar A (2015) The ability of a multi-model seasonal forecasting ensemble to forecast the frequency of warm, cold and wet extremes. Weather Clim Extrem 9: 68–77. https://doi.org/10.1016/j.wace.2015.06.005
- Pinheiro E, Ouarda TBMJ (2023) Short-lead seasonal precipitation forecast in northeastern Brazil using an ensemble of artificial neural networks. Sci Rep 13: 20429. https://doi.org/10.1038/s41598-023-47841-y

- Shestakova AA, Toropov PA (2021) Orographic and lake effect on extreme precipitation on the Iranian coast of the Caspian Sea: a case study. Meteorol Atmos Phys 133(1): 69-84. https://doi.org/10.1007/s00703-020-00735-4
- Shirvani A, Landman WA, Barlow M, Hoell A (2023) Evaluation of the forecast skill of North American Multi-Model Ensemble for monthly and seasonal precipitation forecasts over Iran. Int J Climatol 43(2):1141-1166. https://doi.org/10.1002/joc.7900
- Shirvani A, Landman WA (2016) Seasonal precipitation forecast skill over Iran. Int. J. Climatol. 36: 1887−900. https://doi.org/10.1002/joc.4467
- Sikder S, Hossain F (2016) Assessment of the weather research and forecasting model generalized parameterization schemes for advancement of precipitation forecasting in monsoon-driven river basins. J Adv Model Earth Sys 8(3): 1210–1228. https://doi.org/10.1002/2016MS000678
- Tippett MK, Ranganathan M, L'Heureux M, Barnston AG, DelSole T (2019) Assessing probabilistic predictions of ENSO phase and intensity from the North American Multimodel Ensemble. Clim Dyn 53(12): 7497–7518. https://doi.org/10.1007/s00382-017-3721-y
- UNDRR (2022) Global Assessment Report on Disaster Risk Reduction 2022. United Nations. https://doi.org/10.18356/978921001505
- Vaghefi SA, Keykhai M, Jahanbakhshi F, Sheikholeslami J, Ahmadi A, Yang H, Abbaspour KC (2019) The future of extreme climate in Iran. Sci Rep 9(1): 1464. https://doi.org/10.1038/s41598-018-38071-8
- Vitart F (2014) Evolution of ECMWF sub-seasonal forecast skill scores. Q J R Meteorol Soc 140(683): 1889–1899. https://doi.org/10.1002/qj.2256
- Wang QJ, Shao Y, Song Y, Schepen A, Robertson DE, Ryu D, Pappenberger F (2019) An evaluation of ECMWF SEAS5 seasonal climate forecasts for Australia using a new forecast calibration algorithm. Environ Model Softw 122: 104550. https://doi.org/10.1016/j.envsoft.2019.104550
- Wu M, Nikulin G, Kjellström E, Belušić D, Jones C, Lindstedt D (2020) The impact of regional climate model formulation and resolution on simulated precipitation in Africa. Earth Sys Dyn 11(2): 377–394. https://doi.org/10.5194/esd-11-377-2020
- Xu L, Chen N, Yang C, Yu H, Chen Z (2022) Quantifying the uncertainty of precipitation forecasting using probabilistic deep learning. Hydrol Earth Syst Sci 26(11): 2923–2938. https://doi.org/10.5194/hess-26-2923-2022
- Xu L, Chen N, Zhang X, Chen Z (2020) A data-driven multi-model ensemble for deterministic and probabilistic precipitation forecasting at seasonal scale. Clim Dyn 54(7): 3355-3374. https://doi.org/10.1007/s00382-020-05173-x
- Yazdandoost F, Zakipour M, Izadi A (2023) Statistical refinement of the North American Multi-Model Ensemble precipitation forecasts over Karoon basin, Iran. J Water Clim Change 14(8): 2517-2530. https://doi.org/10.2166/wcc.2023.277
- Yazdandoost F, Moradian S, Zakipour M, Izadi A, Bavandpour M (2020) Improving the precipitation forecasts of the North-American multi model ensemble (NMME) over Sistan basin. J Hydrol 590: 125263. https://doi.org/10.1016/j.jhydrol.2020.125263
- Yeditha PK, Anusha GS, Nandikanti SSS, Rathinasamy M (2023) Development of Monthly Scale Precipitation-Forecasting Model for Indian Subcontinent using Wavelet-Based Deep Learning Approach. Water 15(18): 3244. https://doi.org/10.3390/w15183244
- Zhi X, Qi H, Bai Y, Lin C (2012) A comparison of three kinds of multimodel ensemble forecast techniques based on the TIGGE data. Acta Meteorol Sin 26(1): 41–51. https://doi.org/10.1007/s13351-012-0104-5