Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Projected Precipitation Changes over Iran under SSP Scenarios: A Multi-model Ensemble Assessment Based on Robust Performance Evaluation

Sajad Darzi
University of Shahrood
Emad Mahjoobi
emahjoobi@shahroodut.ac.ir
University of Shahrood

Research Article

Keywords: Adaptive Planning, Climate Change, Climate Models, Entropy, TOPSIS

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11337025/v1

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

Additional Declarations: No competing interests reported.

Projected Precipitation Changes over Iran under SSP Scenarios: A Multi-model Ensemble Assessment Based on Robust Performance Evaluation

Sajad Darzi[1] and Emad Mahjoobi[2]*

[1] *Msc in Water Resources Engineering and Management, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran; Email: sajad.darzi1996@gmail.com, https://orcid.org/0009-0001-3542-2348*

[2] *Assistant Professor, Department of Water and Environmental Engineering, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran; Email: emahjoobi@shahroodut.ac.ir, https://orcid.org/0000-0003-2337-995X*

***Corresponding author:** Emad Mahjoobi, Email: emahjoobi@shahroodut.ac.ir

Abstract

Accurate selection of Global Climate Models (GCMs) is essential for reliable climate projections and water-resource planning. This study evaluates the full CMIP6 ensemble (51 models) for monthly precipitation at 31 synoptic stations across Iran (1970–2014) to identify suitable models, then projects precipitation for 2025–2069 under SSP2-4.5 and SSP5-8.5 at five representative stations. Model performance used five statistical indices: R^2, ANMBD), NRMSE, NNSE, NKGE. Ranking proceeded in two stages: an unweighted Simple Additive Weighting (SAW) baseline, then objective weights from the Entropy method to derive four weighted MCDM methods (WAT, Compromise Programming, Cooperative Game Theory, and TOPSIS). A group decision-making (GDM) approach reconciled method differences to yield a unified ranking. For future projections, a weighted ensemble used the top three GDM-ranked models per station; outputs for 2025–2069 were statistically downscaled via the change factor (CF) method. Evaluation shows NRMSE and R^2 are the most informative objective criteria, while ANMBD contributes least. The final GDM ranking identified IITM-ESM and MIROC-

ES2L as the most reliable GCMs for precipitation in the study area. Among single-method alternatives, WAT correlated most strongly with the integrated GDM results, indicating it can serve as an efficient approximation. Future projections display spatially heterogeneous precipitation trends across stations and scenarios, underscoring the local nature of climate impacts. The results imply no single model is optimal across all regions and support combined MCDM approaches and region-specific GCM selection for robust precipitation projections.

Keywords: Adaptive Planning, Climate Change, Climate Models, Entropy, TOPSIS

Introduction

The twenty-first century is marked by accelerating climate change, necessitating precise scientific inquiry to mitigate its profound environmental and socioeconomic consequences (Mukheef et al. 2024). Driven primarily by the sustained rise in anthropogenic greenhouse gas emissions since the Industrial Revolution (Muhaisen et al. 2024), these changes have fundamentally altered global hydrological cycles. According to the IPCC AR6, the global land-surface temperature for 2011–2020 reached 1.09 °C above the pre-industrial baseline (Masson-Delmotte et al. 2021), while the World Meteorological Organization identified this period as the warmest decade on record (WMO 2021). If current emission trajectories persist, further temperature increases and intensified precipitation anomalies—ranging from severe droughts to flash floods—are projected to undermine the reliability of domestic and agricultural water supplies globally (Cheng et al. 2019; Yilmaz et al. 2024).

As a primary driver of the hydrological cycle, precipitation dictates water quantity and quality, making its reliable projection essential for climate adaptation and water resource management (Roshani and Hamidi 2022). This need is particularly acute in arid and semi-arid regions like Iran, where large climatic gradients and spatially heterogeneous water-use patterns heighten vulnerability to rainfall variability (Ashraf et al. 2021). To address these complexities, General

Circulation Models (GCMs) serve as the principal tools for projecting future hydro-climatic conditions. The latest Coupled Model Intercomparison Project Phase 6 (CMIP6) represents a significant advancement over previous generations, incorporating enhanced algorithmic representations of biogeochemical and oceanic processes to improve simulation fidelity (Zhu and Yang 2020; Carvalho et al. 2022).

Despite the advancements in CMIP6, evaluating model performance remains a critical challenge due to substantial regional uncertainties (Mohsen et al. 2020). Previous evaluations of CMIP6 over Iran have yielded divergent results: for instance, Yazdandoost et al. (2020) observed higher reliability in arid regions compared to humid areas, while others projected varying shifts in precipitation ranging from -12% to +19% depending on the basin and emission scenario (Ansari Mahabadi et al. 2022; Babaeian et al. 2023). Furthermore, model selection remains sensitive to the choice of performance metrics and the reference datasets used (Zareian et al. 2022). These discrepancies highlight a persistent need for a more systematic, multi-dimensional framework that can reconcile conflicting model rankings and reduce predictive uncertainty.

This study addresses these challenges by introducing a comprehensive evaluation and decision-making framework characterized by four distinctive features. First, it utilizes the complete ensemble of 51 CMIP6 GCMs from the Copernicus Data Store, ensuring an exhaustive assessment beyond the scope of preliminary model subsets. Second, the evaluation incorporates an expansive suite of five normalized performance metrics—including coefficient of determination (R2), absolute normalized mean bias deviation (ANMBD), normalized root mean square error (NRMSE), normalized Nash–Sutcliffe efficiency (NNSE), and normalized Kling Gupta efficiency (NKGE)—to provide a holistic capture of bias, error, and efficiency. Third, the research adopts an objective weighting scheme derived from Shannon Entropy, integrated with multiple Multi-Criteria Decision-Making (MCDM) techniques (SAW, WAT, CP, CGT,

and TOPSIS) to rigorously prioritize criteria. Finally, the robustness of the results is underpinned by long-term, continuous records from strategically selected synoptic stations across Iran's diverse geographical regions.

To achieve a consensus ranking, a Group Decision Making (GDM) procedure was employed to aggregate results from the various MCDM methods, thereby identifying the most reliable models for each station. These superior models were then used to form a weighted ensemble (following the Brekke et al. 2012 scheme) to project precipitation changes for the 2025–2069 period under SSP2-4.5 and SSP5-8.5 scenarios. By integrating multi-stage evaluation with an ensemble strategy, this work aims to: (1) identify the most reliable CMIP6 models for precipitation simulation across Iran, (2) quantify future changes at representative stations across diverse cardinal directions, and (3) provide a transparent, reproducible regional model-selection framework to support climate-informed water resource planning.

Materials and Methods

Study area

Iran (≈1,648,000 km^2) in Southwest Asia displays diverse topography and climates—from arid/semi-arid central plateau to Mediterranean and mountainous conditions in the Alborz and Zagros ranges (Heydarizad et al. 2019). The national mean annual precipitation is ≈281 mm, classifying Iran as overall arid (Choobeh et al. 2025). Rainfall is spatially and temporally uneven: highest in the Caspian Sea region and Zagros Mountains, declining generally from west to east and from north to south (Ahmadi and Kamangar 2023). Recent shifts in precipitation regimes have increased vulnerability, challenging water-resource management and agricultural stability. Figure 1 shows Iran's topography and the selected synoptic stations in provincial capitals. Figure 2 summarizes the study workflow.

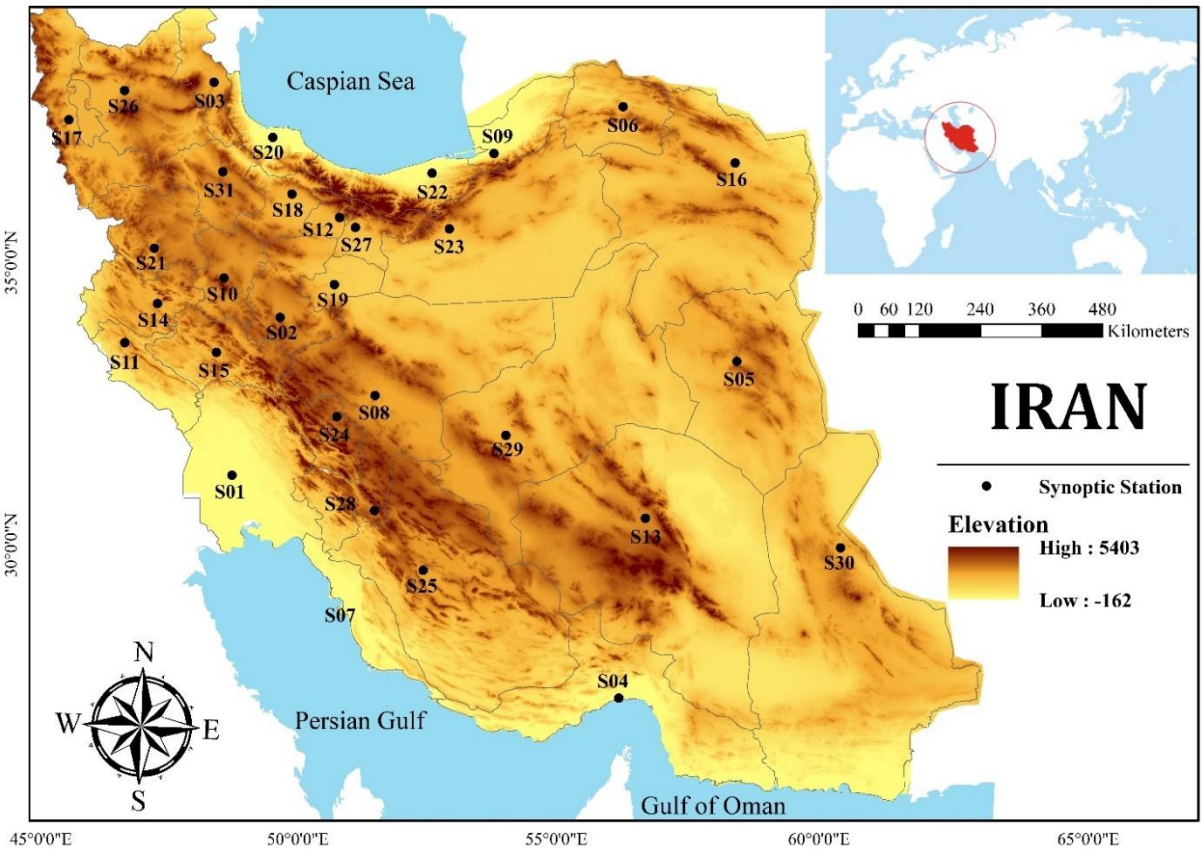


Figure 1. Map of Iran with elevation contours and locations of synoptic stations in provincial capitals.

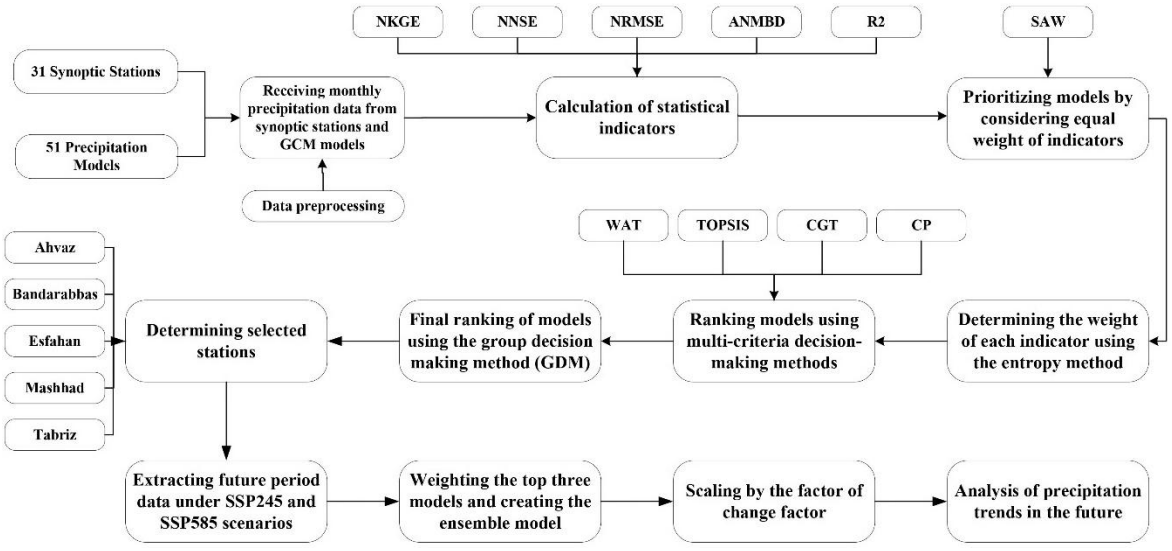


Figure 2. Flowchart of the study's research methodology.

Observational and Simulated data and preprocessing

Monthly precipitation observations from 31 synoptic stations in provincial capitals (primarily Iran Meteorological Organization) were used for April 1970–December 2014 to match available historical GCM simulations. Table 1 lists station details.

Table 1. Characteristics of synoptic stations in provincial capitals.

Station code	Station name	Province	Latitude (N)	Longitude (S)	Height above sea level (m)	Start of the statistical period	Station code	Station name	Province	Latitude (N)	Longitude (S)	Height above sea level (m)	Start of the statistical period
S01	Ahvaz	Khuzestan	31.3	48.7	23	1970	S17	Urmia	West Azerbaijan	37.7	45.1	1328	1970
S02	Arak	Markazi	34.1	49.8	1703	1970	S18	Qazvin	Qazvin	36.3	50	1279	1970
S03	Ardebil	Ardebil	38.2	48.3	1335	1976	S19	Qom	Qom	34.8	50.9	879	1986
S04	Bandarabbas	Hormozgan	27.2	56.4	10	1970	S20	Rasht	Gilan	37.3	49.6	-9	1970
S05	Birjand	South Khorasan	32.9	59.3	1491	1970	S21	Sanandaj	Kurdistan	35.3	47	1373	1970
S06	Bojnourd	North Khorasan	37.5	57.3	1065	1977	S22	Sari	Mazandaran	36.5	53	23	1999
S07	Bushehr	Bushehr	28.9	50.8	8	1970	S23	Semnan	Semnan	35.6	53.4	1127	1970
S08	Esfahan	Esfahan	32.5	51.7	1550	1970	S24	Shahrekord	Chaharmahal Bakhtiari	32.3	50.8	2049	1970
S09	Gorgan	Golestan	36.9	54.4	0	1970	S25	Shiraz	Fars	29.6	52.6	1488	1970
S10	Hamedan	Hamedan	34.9	48.5	1741	1976	S26	Tabriz	East Azerbaijan	38.1	46.2	1361	1970
S11	Ilam	Ilam	33.6	46.4	1337	1986	S27	Tehran	Tehran	35.7	51.4	1419	1970
S12	Karaj	Alborz	35.8	51	1293	1985	S28	Yasuj	Kohgiluyeh and Boyerahmad	30.7	51.6	1821	1987
S13	Kerman	Kerman	30.3	57	1754	1970	S29	Yazd	Yazd	31.9	54.3	1230	1970
S14	Kermanshah	Kermanshah	34.4	47.2	1319	1970	S30	Zahedan	Sistan and Baluchestan	29.5	60.9	1370	1970
S15	Khoramabad	Lorestan	33.4	48.3	1148	1970	S31	Zanjan	Zanjan	36.7	48.5	1659	1970
S16	Mashhad	Razavi Khorasan	36.2	59.6	999	1970							

Simulated monthly precipitation from 51 CMIP6 GCMs were obtained from the Copernicus Climate Data Store (.nc files), preprocessed and converted to station-level time series via ArcGIS spatial processing and time-series scripts. For each station a data matrix of 52 rows (51 GCMs + observations) and up to 537 monthly records (1970–2014) was assembled; Table 2 summarizes model metadata. Datasets were checked for units and temporal alignment and regridded/interpolated to station locations as needed.

Table 2. Characteristics of CMIP6 climate models.

Model Code	Model Name	Country	Resolution (Km)	Model Code	Model Name	Country	Resolution (Km)
G01	ACCESS-CM2	Australia	250	G27	FGOALS-g3	China	100
G02	ACCESS-ESM1-5	Australia	250	G28	FIO-ESM-2-0	China	100

G03	AWI-CM-1-1-MR	Germany	100	G29	GFDL-ESM4	USA	100
G04	AWI-ESM-1-1-LR	Germany	250	G30	GISS-E2-1-H	USA	250
G05	BCC-CSM2-MR	China	100	G31	HadGEM3-GC31-LL	UK	250
G06	BCC-ESM1	China	250	G32	HadGEM3-GC31-MM	UK	100
G07	CAMS-CSM1-0	China	100	G33	IITM-ESM	India	250
G08	CanESM5	Canada	100	G34	INM-CM4-8	Russia	100
G09	CanESM5-CanOE	Canada	500	G35	INM-CM5-0	Russia	100
G10	CESM2	USA	100	G36	IPSL-CM5A2-INCA	France	500
G11	CESM2-FV2	USA	100	G37	IPSL-CM6A-LR	France	250
G12	CESM2-WACCM	USA	100	G38	KACE-1-0-G	South Korea	250
G13	CESM2-WACCM-FV2	USA	100	G39	MCM-UA-1-0	USA	250
G14	CMCC-CM2-HR4	Italy	100	G40	MIROC6	Japan	250
G15	CMCC-CM2-SR5	Italy	100	G41	MIROC-ES2H	Japan	250
G16	CMCC-ESM2	Italy	100	G42	MIROC-ES2L	Japan	500
G17	CNRM-CM6-1	France	250	G43	MPI-ESM1-2-HR	Germany	100
G18	CNRM-CM6-1-HR	France	50	G44	MPI-ESM1-2-LR	Germany	250
G19	CNRM-ESM2-1	France	250	G45	MRI-ESM2-0	Japan	500
G20	E3SM-1-0	USA	100	G46	NESM3	China	100
G21	E3SM-1-1	USA	100	G47	NorCPM1	Norway	250
G22	E3SM-1-1-ECA	USA	100	G48	NorESM2-MM	Norway	100
G23	EC-Earth3-AerChem	Europe	100	G49	SAM0-UNICON	South Korea	100
G24	EC-Earth3-CC	Europe	100	G50	TaiESM1	China	100
G25	EC-Earth3-Veg-LR	Europe	100	G51	UKESM1-0-LL	United Kingdom	250
G26	FGOALS-f3-L	China	100				

Evaluation indices

Model performance was quantified using five statistical indices: coefficient of determination (R^2) for linear association; absolute normalized mean bias deviation (ANMBD) as a normalized bias metric (smaller = less systematic error); normalized root mean square error (NRMSE) for overall error magnitude; normalized Nash–Sutcliffe efficiency (NNSE), obtained by transforming NSE to a bounded normalized from following Nossent and Bauwens (2012) to retain all models for ranking; and normalized Kling–Gupta efficiency (NKGE), defined analogously to bound KGE and facilitate multi-criteria processing.

$$R^2 = \left[\frac{\sum_{i=1}^{n}(o_i - \bar{o})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^{n}(o_i - \bar{o})^2 \sum_{i=1}^{n}(s_i - \bar{s})^2}}\right]^2 \quad (1)$$

$$ANMBD = \left| \frac{\frac{1}{n}(\sum_{i=1}^{n}(s_i - o_i))}{\overline{o}} \right| \quad (2)$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - s_i)^2}}{\overline{o}} \quad (3)$$

$$NNSE = \frac{1}{2 - (1 - \frac{\sum_{i=1}^{n}(o_i - s_i)^2}{\sum_{i=1}^{n}(o_i - \bar{o}_i)^2})} \quad (4)$$

$$NKGE = \frac{1}{2 - \left[1 - \sqrt{(R-1)^2 + (\frac{\mu_S}{\mu_O} - 1)^2 + (\frac{\frac{\sigma_s}{\mu_S}}{\frac{\sigma_O}{\mu_O}} - 1)^2}\right]} \quad (5)$$

s denotes simulated values, $\bar{s}$ the mean of simulated values, o observed values, $\bar{o}$ the mean of observed values, n the number of observations, i the observation index, μ_S and μ_O the simulated and observed means, R the Pearson correlation coefficient, and σ_s and σ_O the standard deviations of simulated and observed series. Explicit formulae for each index and the normalization transformations (Nossent and Bauwens 2012) are provided in Methods (see Index computation and normalization).

Multi-criteria decision making

Criteria were standardized to a common scale using a non-linear normalization method (Ersoy 2020), with separate formulas for benefit (larger is better) and cost (smaller is better) criteria to improve standardization accuracy and flexibility (Zavadskas and Turskis 2008).

$$K_{zij} = \frac{k_{zij}}{k_{zij}^{max^2}} \qquad for\ Benefit\ criteria \quad (6)$$

$$K_{zij} = \frac{k_{zij}^{min}}{k_{zij}{}^{3}} \qquad for\ Cost\ criteria \quad (7)$$

144 K_{zij} denotes the normalized j-th index for the i-th model at station z; k_{zij} is the original value,
145 and k_{zij}^{min} and k_{zij}^{max} are its minimum and maximum. Indices z, i, and j enumerate stations,
146 models, and indices.
147 Objective index weights were derived via the entropy method (Shannon 1948): compute the
148 normalized decision matrix, calculate each criterion's entropy, and determine final weights
149 (Equations 8–10). Greater dispersion yields higher weight; lower dispersion yields lower
150 weight, making the entropy approach a data-driven method for MCDM weighting (Zeleny
151 1998).

$$En_{zj} = -\frac{1}{ln(T)}\sum_{i=1}^{T} K_{zij}\, ln(K_{zij}) \tag{8}$$

$$Dd_{zj} = 1 - En_{zj} \tag{9}$$

$$r_{zj} = \frac{Dd_{zj}}{\sum_{j=1}^{J} Dd_{zj}} \tag{10}$$

152 En_{zj} is the entropy of the j-th index at station z, T the number of models, Dd_{zj} the degree of
153 divergence (importance) of index j at station z, J the total number of indices, and r_{zj} the final
154 weight of index j at station z.
155 Model ranking used SAW as a baseline and four additional MCDM methods: WAT,
156 Compromise Programming (CP), Cooperative Game Theory (CGT), and TOPSIS. In SAW, S_{zi}
157 is the weighted sum of standardized indices; higher S_{zi} denotes better models.

$$S_{zi} = \frac{\sum_{j=1}^{J} K_{zij}}{J} \tag{11}$$

158 In WAT, W_{zi} equals the sume of each criterion's weighted multiplied by its standardized value;
159 higher W_{zi} is preferable (Raju and Kumar 2015).

$$W_{zi} = \left[\sum_{j=1}^{J} r_{zj} K_{zij}\right] \tag{12}$$

CP ranks by distance to an ideal solution, C_{zi} denoting that distance; K_{zj}^{*} is the ideal value and P (commonly 2) the exponent; indices where smaller is better (e.g., NRMSE and ANMBD) were multiplied by −1 so desirable direction matches others, and smaller C_{zi} implies higher rank (Moghaddas et al. 2024).

$$C_{zi} = \left[\sum_{j=1}^{J} r_{zj}^{P} \left|K_{zj}^{*} - K_{zij}\right|^{P}\right]^{\frac{1}{P}} \tag{13}$$

CGT evaluates distance from a non-ideal (worst) solution, G_{zi} denoting that distance; K_{zj}^{**} are non-ideal values (minimum after direction adjustments), and larger G_{zi} is preferred (Vassoney et al. 2021).

$$G_{zi} = \prod_{j=1}^{J} \left|K_{zij} - K_{zj}^{**}\right|^{r_{zj}} \tag{14}$$

TOPSIS ranks by relative distances to ideal and non-ideal solutions: DS_{zi}^{+} (distance to ideal), DS_{zi}^{-} (distance to non-ideal), and CR_{zi} (relative closeness), with larger CR_{zi} indicating higher rank (Sharma and Kale 2022; Raju and Kumar 2015).

$$DS_{zi}^{+} = \sqrt{\sum_{j=1}^{J} {r_{zj}}^{2} (K_{zij} - K_{zj}^{*})^{2}} \tag{15}$$

$$DS_{zi}^{-} = \sqrt{\sum_{j=1}^{J} {r_{zj}}^{2} (K_{zij} - K_{zj}^{**})^{2}} \tag{16}$$

$$CR_{zi} = \frac{DS_{zi}^{-}}{DS_{zi}^{-} + DS_{zi}^{+}} \tag{17}$$

Group decision making

Because different MCDM methods can yield varying rankings, results were consolidated via a Group Decision-Making (GDM) approach to produce a final ranking (Morais and Almeida 2012). Models were partitioned into high-rank and low-rank groups based on the total number

of models T: for even T, the lower bound of the high-rank group X = T/2; for odd T, X = (T + 1)/2; the upper bound of the low-rank group Y = X + 1.

The "power" of each model ST_{zi} at station z quantifies membership in the higher-rank group across M decision methods: for each method h, q_{zi}^{h} = 1 if $1 \le L_i^h \le X_z$ and 0 otherwise; ST_{zi} aggregates these contributions across methods.

$$ST_{zi} = \sum_{h=1}^{M} (X_z - L_i^h + 1) q_{zi}^{h} \qquad for\ 1 \le L_i^h \le X_z \tag{18}$$

The "weakness" WE_{zi} is defined analogously using q_{zi}^{h} = 1 if $Y_z \le L_i^h \le$ T (and 0 otherwise), measuring membership in the low-rank group across methods.

$$WE_{zi} = \sum_{h=1}^{M} (L_i^h - Y_z + 1) q_{zi}^{h} \qquad for\ Y_z \le L_i^h \le T \tag{19}$$

Net power NS_{zi} is computed as $ST_{zi} - WE_{zi}$; models with higher NS_{zi} are deemed more suitable.

Ensemble model development

The Global Decision Model (GDM) was utilized to identify the top-performing models for each station, followed by the extraction of future projections for five representative locations: Ahvaz (southwest), Bandar Abbas (south), Esfahan (central), Mashhad (northeast), and Tabriz (northwest). These specific stations were selected to ensure a comprehensive representation of the country's diverse geographical regions and spatial distribution. Future monthly data for April 2025–December 2069 under SSP2-4.5 and SSP5-8.5 were obtained from the Copernicus Data Store for the top three ranked models per station; the future period length matches the historical period (1970-2014). Outputs from the three top models were preprocessed into time series for ensemble development and analysis.

To reduce uncertainty, ensemble models were constructed at each station by integrating the three top-ranked models using Brekke et al. (2012) rank-based weights. For each superior model, the sum of its ranks across the five indices is divided by ($\hat{T} \times J$) where $\hat{T} = 3$ (number of superior models) and $J = 5$ (number of indices); this quotient is subtracted from one to yield the model weight $R_{\hat{z}\hat{i}}$. Better-ranked models therefore receive higher weights. The same weighting approach was applied to historical data (performance evaluation) and future projections (ensemble forecasts).

Because GCMs lack local-scale resolution, we applied statistical downscaling using the change-factor (CF, delta) method. For precipitation, the multiplicative change factor equals the ratio of the long-term monthly mean for the future period (2025–2069) to that of the base period (1970–2014). For each station and scenario: compute ensemble mean historical ($\overline{BEM}$) and future ($\overline{FEM}$) series; calculate the multiplicative change factor $= \frac{\overline{FEM}}{\overline{BEM}}$; and generate downscaled precipitation $LFEM_i^P$ by applying this factor to observed historical precipitation O_i^p.

Results and Discussion

Model evaluation

We evaluated 51 CMIP6 GCMs' ability to simulate monthly precipitation using five performance indices (R^2, ANMBD, NRMSE, NNSE, NKGE) against observations from synoptic stations in Iranian provincial capitals for April 1970–December 2014. Indices were computed per station for all models; Table 3 reports maximum, mean, and minimum values. Performance is highly heterogeneity across stations due to diverse climates, complex topography, and model differences.

Table 3. Station-wise ranges (Maximum, Mean, Minimum) of climate model evaluation indices.

Station	R^2	ANMBD	NRMSE	NNSE	NKGE
S01	(**0**, 0.063, 0.188)	(0.019, 0.44, 1.262)	(1.585, 1.939, 3.037)	(0.222, 0.419, 0.512)	(0.329, 0.512, 0.614)
S02	(0.002, 0.119, 0.243)	(0.002, 0.376, 1.519)	(1.086, 1.405, 2.653)	(0.158, 0.416, 0.528)	(0.375, 0.552, 0.637)
S03	(0.001, 0.078, 0.176)	(0.042, 1.047, 4.408)	(1.1, 2.033, 5.502)	(**0.029**, 0.231, 0.429)	(0.181, 0.46, 0.595)
S04	(**0**, 0.019, 0.097)	(0.011, 0.419, 0.923)	(2.345, 2.716, 3.392)	(0.324, 0.431, 0.501)	(0.374, 0.489, 0.568)
S05	(0.004, 0.121, 0.264)	(0.002, 0.502, 1.458)	(1.517, 1.915, 2.872)	(0.204, 0.38, 0.479)	(0.388, 0.528, 0.641)
S06	(0.057, 0.151, 0.24)	(0.01, 0.333, 1.173)	(0.906, 1.244, 2.144)	(0.171, 0.393, 0.535)	(0.427, 0.573, 0.651)
S07	(0.001, 0.051, 0.158)	(0.052, 0.399, 0.825)	(1.889, 2.159, 2.994)	(0.29, 0.443, 0.507)	(0.415, 0.517, 0.622)
S08	(0.007, 0.082, 0.174)	(0.017, 1.087, 3.528)	(1.553, 2.754, 5.837)	(0.055, 0.249, 0.451)	(0.217, 0.445, 0.59)
S09	(0.001, 0.048, 0.1)	(0.048, 0.511, 1.463)	(**0.805**, **0.973**, 2.087)	(0.101, 0.348, 0.43)	(0.349, 0.486, 0.575)
S10	(0.001, 0.145, 0.267)	(0.002, 0.419, 1.657)	(1.025, 1.406, 2.684)	(0.133, 0.38, 0.513)	(0.364, 0.554, 0.642)
S11	(0.001, 0.103, 0.236)	(0.064, 0.595, 0.959)	(1.211, 1.404, 1.99)	(0.292, 0.456, 0.527)	(0.365, 0.514, 0.653)
S12	(**0**, 0.102, 0.203)	(0.005, 0.465, 2.436)	(1.21, 1.599, 3.495)	(0.1, 0.366, 0.48)	(0.283, 0.536, 0.642)
S13	(**0**, 0.052, 0.127)	(0.004, 0.687, 4.923)	(1.559, 2.375, 7.776)	(0.039, 0.341, 0.499)	(0.167, 0.48, 0.593)
S14	(0.001, 0.156, 0.296)	(0.003, 0.358, 1.562)	(1.11, 1.366, 2.658)	(0.159, 0.433, 0.52)	(0.377, 0.573, 0.652)
S15	(0.006, 0.144, 0.25)	(0.04, 0.456, 2.33)	(1.109, 1.372, 3.663)	(0.091, 0.434, 0.521)	(0.295, 0.548, 0.654)
S16	(0.028, **0.193**, **0.383**)	(0.015, **0.332**, 0.999)	(1.07, 1.284, 1.976)	(0.269, **0.47**, **0.557**)	(0.466, **0.593**, **0.693**)
S17	(0.007, 0.125, 0.205)	(0.006, 0.547, 2.465)	(1.074, 1.55, 3.809)	(0.074, 0.351, 0.503)	(0.282, 0.531, 0.618)
S18	(0.001, 0.12, 0.225)	(**0.002**, 0.459, 1.651)	(1.055, 1.486, 3.478)	(0.089, 0.377, 0.515)	(0.359, 0.54, 0.636)
S19	(0.001, 0.081, 0.189)	(0.022, 0.995, 2.249)	(1.38, 2.427, 3.871)	(0.104, 0.253, 0.479)	(0.293, 0.453, 0.602)
S20	(**0**, 0.05, 0.207)	(0.151, 0.584, 0.843)	(0.867, 1.05, 1.217)	(0.324, 0.394, 0.485)	(0.41, 0.492, 0.609)
S21	(0.001, 0.172, 0.333)	(0.014, 0.378, 1.704)	(1.004, 1.27, 2.772)	(0.132, 0.44, 0.536)	(0.359, 0.573, 0.665)
S22	(**0**, 0.058, 0.271)	(0.077, 0.596, 1.852)	(0.855, 1.05, 2.738)	(0.07, 0.361, 0.438)	(0.343, 0.485, 0.581)
S23	(0.003, 0.072, 0.128)	(0.002, 0.93, 3.95)	(1.393, 2.317, 5.624)	(0.05, 0.278, 0.461)	(0.199, 0.465, 0.6)
S24	(0.014, 0.118, 0.246)	(0.016, 0.486, 2.189)	(1.227, 1.644, 3.43)	(0.122, 0.4, 0.521)	(0.306, 0.532, 0.659)
S25	(0.001, 0.067, 0.182)	(0.03, 0.433, 1.585)	(1.628, 1.965, 3.12)	(0.215, 0.417, 0.502)	(0.363, 0.517, 0.609)
S26	(0.003, 0.144, 0.23)	(0.019, 0.636, 1.722)	(1.046, 1.514, 2.566)	(0.126, 0.308, 0.464)	(0.349, 0.521, 0.626)
S27	(**0**, 0.099, 0.198)	(0.021, 0.532, 3.102)	(1.183, 1.69, 4.295)	(0.07, 0.354, 0.497)	(0.238, 0.527, 0.618)
S28	(**0**, 0.101, 0.252)	(0.029, 0.595, 0.924)	(1.317, 1.538, 1.969)	(0.348, 0.468, 0.544)	(0.402, 0.511, 0.65)
S29	(0.004, 0.044, 0.159)	(0.147, 2.249, **5.217**)	(2.579, 4.919, **8.789**)	(0.041, 0.149, 0.33)	(**0.159**, 0.332, 0.558)
S30	(**0**, 0.031, 0.119)	(0.018, 0.913, 3.141)	(2.118, 3.365, 6.715)	(0.073, 0.278, 0.442)	(0.236, 0.44, 0.551)
S31	(0.003, 0.134, 0.254)	(0.026, 0.522, 2.746)	(0.991, 1.442, 3.56)	(0.069, 0.344, 0.488)	(0.259, 0.544, 0.642)

R^2 values were often near zero (e.g., Ahvaz, Tehran, Zahedan), with Mashhad showing the highest individual R^2 (0.383) and highest mean R^2 (0.193). ANMBD peaked at Yazd (5.217) while Mashhad had the lowest mean ANMBD (0.332). NRMSE was maximal at Yazd (8.789) and lowest mean at Gorgan (0.973). NNSE ranged from a minimum at Ardabil (0.029) to an individual maximum at Mashhad (0.557), with Mashhad also having the highest mean NNSE (0.470). NKGE spanned from a low at Yazd (0.159) to an individual high at Mashhad (0.693), and Mashhad had the highest mean NKGE (0.593). Overall, Mashhad consistently exhibits the best mean agreement CMIP6 outputs and observations.

Figure 3's spatial maps confirm no uniform spatial pattern in CMIP6 precipitation accuracy across Iran. Mean R^2 is generally low (≈0.02–0.19), highest near Mashhad and Sanandaj and lowest in arid/semi-arid areas (Yazd, Kerman, Zahedan) and parts of the Caspian coast (Gorgan,

Rasht). ANMBD varies markedly (≈0.33–2.25): lowest deviations occure in central-western, southwestern, and some northern regions (including Mashhad, Bojnourd, and the Zagros foothills), while largest biases appear in central-arid areas (Yazd, Esfahan) and parts of the northwest (Ardabil). NRMSE ranges widely (≈0.97–4.92), with lowest errors mainly in the northeast and the western half of the country and highest errors at Zahedan and Yazd.

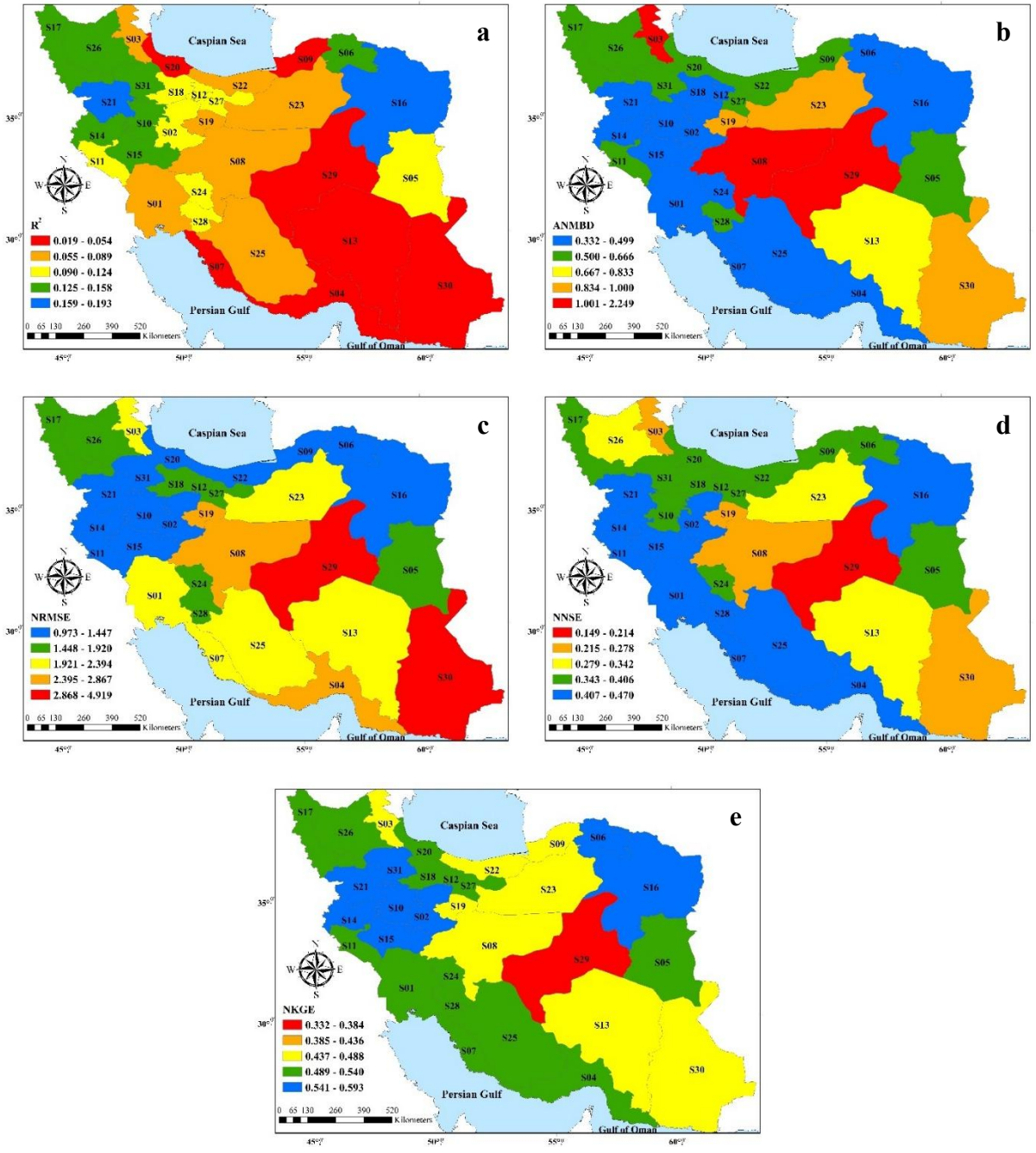


Figure 3. Spatial distribution of mean monthly precipitation evaluation indices from CMIP6 models: (a) R^2, (b) ANMBD, (c) NRMSE, (d) NNSE, (e) NKGE.

Efficiency indices NNSE (Figure 3d) and NKGE (Figure 3e) show generally low performance at most stations but with regional clusters of better performance. High NNSE values occur in the western half (e.g., Kermanshah, Sanandaj) and the Persian Gulf basins. NKGE (≈0.33–0.59) is highest in the northeast (Bojnourd, Mashhad) and parts of the western half, while Yazd exhibits the poorest performance across both efficiency metrics. Overall, CMIP6 models perform relatively better (lower error, higher efficiency) in the northeast (Mashhad/Bojnourd) and central-western regions (Zagros foothills) and worse (lower correlation, higher error) at highly arid central stations (e.g., Yazd) and some coastal areas.

Ranking of models by the SAW method

We ranked the 51 CMIP6 models at each station using the Simple Additive Weighting (SAW) on standardized indices with equal weights for the five performance metrics as a baseline. Table 4 lists the top three models per station: IITM-ESM (G33) is the strongest single performer across the most stations, ranking first at Ardabil, Bandarabbas, Birjand, Shahrekord, and Shiraz and appearing among the top three at seven additional locations (Ahvaz, Ilam, Kermanshah, Khorramabad, Qazvin, Sanandaj, Yasuj). FGOALS-f3-L (G26) ranked first in Bushehr, Ilam, and Yasuj, and MIROC-ES2L (G42) ranked first in Bojnourd, Gorgan, and Mashhad. Figure 4 maps the spatial distribution of the top-ranked precipitation model per station under SAW.

Table 4. Top three precipitation models by the SAW method.

Station	Rank 1	Rank 2	Rank 3	Station	Rank 1	Rank 2	Rank 3
S01	G51	G33	G26	S17	G27	G38	G39
S02	G48	G42	G23	S18	G36	G33	G13
S03	G33	G42	G11	S19	G04	G36	G32

S04	G33	G05	G29	S20	G32	G28	G24
S05	G33	G25	G38	S21	G01	G32	G33
S06	G42	G30	G47	S22	G30	G18	G47
S07	G26	G48	G10	S23	G25	G15	G48
S08	G31	G24	G32	S24	G33	G48	G11
S09	G42	G07	G06	S25	G33	G31	G48
S10	G23	G06	G48	S26	G27	G03	G48
S11	G26	G37	G33	S27	G51	G47	G13
S12	G47	G38	G16	S28	G26	G17	G33
S13	G30	G38	G04	S29	G04	G32	G46
S14	G01	G33	G05	S30	G24	G48	G01
S15	G32	G33	G48	S31	G47	G48	G49
S16	G42	G10	G20				

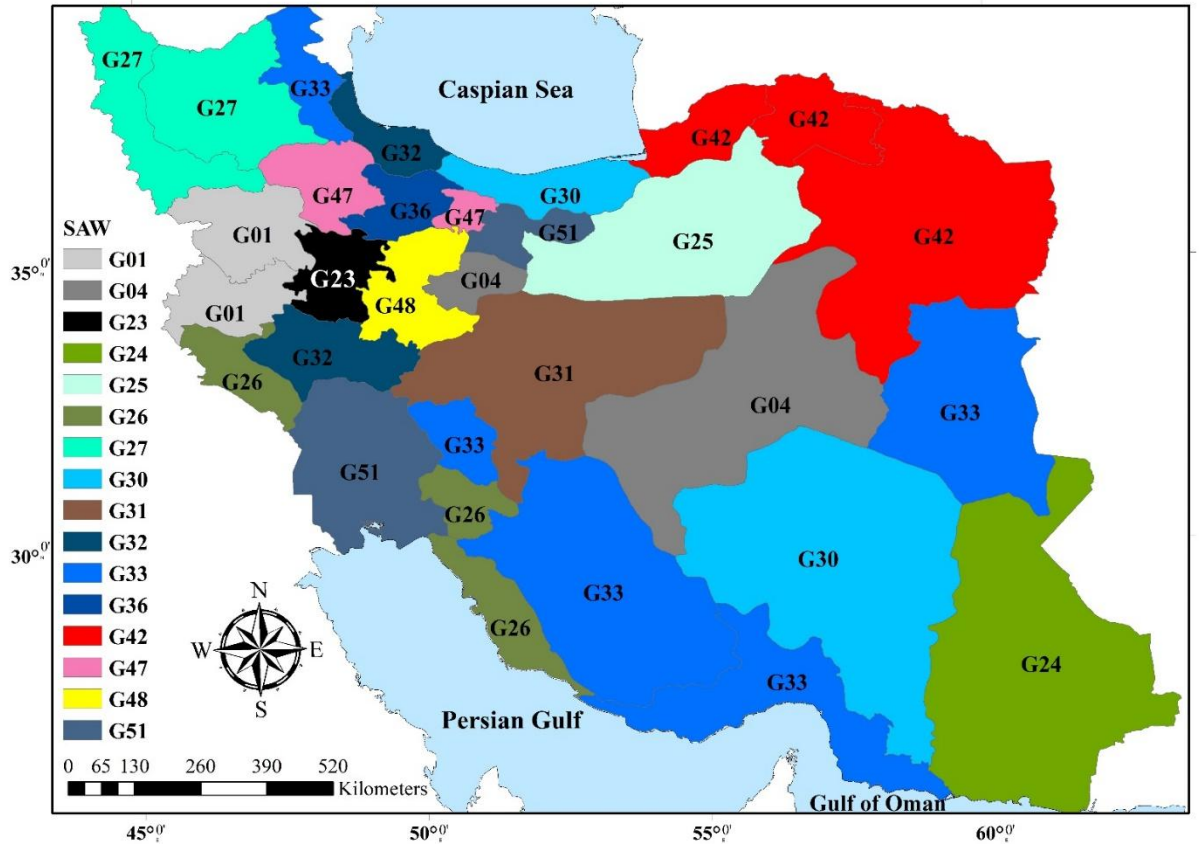


Figure 4. Best precipitation simulation model identified by SAW.

The SAW analysis found 16 different models in the top position across 31 stations; about 31% of the 51 models achieved best performance at least once. This diversity indicates no single model is superior nationwide and that optimal model choice depends on regional characteristics.

Index weighting

Objective weights for the five evaluation indices were computed via the entropy method (Table 5). NRMSE received the highest weight at 17 stations and was the most decisive criterion. R^2 had the highest weight at nine stations. ANMBD obtained the lowest weight at all stations, contributing least to discrimination. NKGE, R^2, NRMSE, and NNSE showed similar weights across most stations, so none of these four overwhelmingly dominated the assessment.

Table 5. Index weight values computed by the Entropy method.

Station	R^2	ANMBD	NRMSE	NNSE	NKGE	Station	R^2	ANMBD	NRMSE	NNSE	NKGE
S01	0.19	**0.08**	**0.27**	0.24	0.23	S17	0.25	**0.056**	**0.26**	0.25	0.18
S02	0.26	**0.06**	**0.26**	0.23	0.2	S18	0.25	**0.057**	**0.25**	0.23	0.21
S03	0.23	**0.06**	**0.25**	0.24	0.22	S19	0.22	**0.055**	0.25	**0.25**	0.23
S04	0.16	**0.07**	**0.28**	0.24	0.25	S20	0.14	**0.116**	**0.26**	0.24	0.24
S05	**0.25**	**0.06**	0.24	0.22	0.24	S21	**0.28**	**0.075**	0.24	0.2	0.2
S06	0.26	**0.06**	**0.26**	0.25	0.18	S22	0.16	**0.09**	**0.26**	0.24	0.26
S07	0.2	**0.1**	0.24	0.21	**0.25**	S23	**0.25**	**0.054**	0.24	0.24	0.22
S08	0.24	**0.06**	**0.24**	0.24	0.22	S24	0.22	**0.072**	**0.25**	0.23	0.23
S09	0.24	**0.07**	**0.25**	0.23	0.21	S25	0.21	**0.081**	**0.25**	0.23	0.23
S10	**0.26**	**0.06**	0.25	0.24	0.2	S26	0.23	**0.059**	**0.26**	0.25	0.2
S11	0.22	**0.08**	0.24	0.21	**0.26**	S27	**0.25**	**0.077**	0.25	0.24	0.19
S12	**0.26**	**0.06**	0.24	0.23	0.21	S28	0.22	**0.063**	0.25	0.21	**0.26**
S13	0.21	**0.06**	**0.26**	0.25	0.23	S29	0.18	**0.072**	0.25	0.24	**0.26**
S14	**0.3**	**0.06**	0.24	0.21	0.2	S30	0.17	**0.072**	**0.26**	0.26	0.24
S15	**0.24**	**0.09**	0.24	0.21	0.22	S31	0.25	**0.073**	**0.25**	0.24	0.19
S16	**0.27**	**0.06**	0.25	0.21	0.21						

Ranking of models using multi-criteria decision-making (MCDM) methods

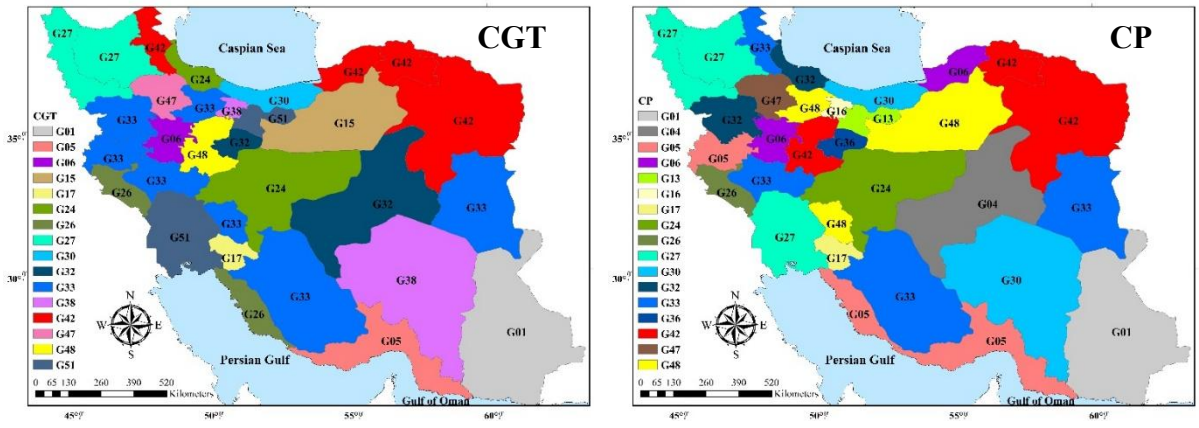
Entropy-derived weights were applied in four MCDM methods (WAT, CP, CGT, TOPSIS) to produce weighted rankings; Table 6 lists the top three models per station. Methods largely agreed though spatial variability in performance persisted.

Table 6. Top three climate models per station based on MCDM.

Station	CP			CGT			TOPSIS			WAT		
	Rank 1	Rank 2	Rank 3	Rank 1	Rank 2	Rank 3	Rank 1	Rank 2	Rank 3	Rank 1	Rank 2	Rank 3
S01	G27	G51	G26	G51	G26	G27	G27	G51	G26	G51	G26	G27

S02	G42	G48	G31	G48	G42	G31	G42	G48	G31	G48	G42	G31
S03	G33	G42	G11	G42	G33	G11	G33	G42	G11	G33	G42	G11
S04	G05	G33	G13	G05	G33	G13	G05	G33	G13	G05	G33	G13
S05	G33	G38	G11	G33	G38	G11	G33	G38	G11	G33	G38	G11
S06	G42	G47	G15	G42	G47	G15	G42	G47	G15	G42	G47	G15
S07	G05	G26	G30	G26	G48	G05	G05	G26	G48	G26	G48	G05
S08	G24	G32	G36	G24	G43	G23	G24	G32	G36	G24	G32	G36
S09	G06	G07	G29	G42	G07	G06	G06	G07	G42	G42	G07	G06
S10	G06	G48	G47	G06	G48	G47	G06	G48	G47	G06	G48	G42
S11	G26	G17	G20	G26	G17	G33	G26	G17	G20	G26	G33	G17
S12	G16	G50	G49	G38	G16	G50	G16	G50	G49	G38	G16	G50
S13	G30	G38	G04	G38	G30	G11	G30	G38	G04	G30	G38	G04
S14	G05	G29	G33	G33	G29	G05	G05	G29	G33	G33	G05	G29
S15	G33	G24	G21	G33	G48	G24	G33	G24	G21	G48	G33	G24
S16	G42	G10	G20	G42	G10	G20	G42	G10	G20	G42	G10	G20
S17	G27	G42	G39	G27	G43	G42	G27	G42	G39	G27	G42	G39
S18	G48	G13	G33	G33	G13	G48	G48	G13	G33	G33	G36	G13
S19	G36	G32	G48	G32	G36	G04	G36	G32	G48	G36	G32	G04
S20	G32	G25	G24	G24	G32	G25	G32	G25	G24	G32	G28	G24
S21	G32	G10	G05	G33	G32	G10	G32	G10	G05	G33	G32	G51
S22	G30	G32	G43	G30	G47	G08	G30	G32	G43	G30	G18	G07
S23	G48	G15	G09	G15	G48	G51	G48	G09	G15	G15	G48	G51
S24	G48	G33	G42	G33	G48	G43	G48	G33	G42	G48	G33	G42
S25	G33	G48	G32	G33	G48	G07	G33	G48	G32	G33	G31	G48
S26	G27	G48	G26	G27	G48	G50	G27	G48	G26	G27	G48	G50
S27	G13	G51	G24	G51	G13	G24	G13	G51	G24	G51	G13	G24
S28	G17	G26	G45	G17	G26	G45	G17	G26	G45	G26	G17	G33
S29	G04	G32	G46	G32	G04	G43	G04	G32	G46	G04	G32	G46
S30	G01	G31	G48	G01	G48	G33	G01	G31	G48	G01	G48	G31
S31	G47	G48	G39	G47	G48	G11	G47	G48	G39	G47	G48	G39

At Mashhad the top three models were identical across all four MCDM methods, and the first-ranked model was consistent across all four methods at nine stations (Birjand, Bojnurd, Ilam, Mashhad, Urmia, Sari, Shiraz, Tabriz, Zanjan). Figure 5 maps these MCDM top-ranked distributions.



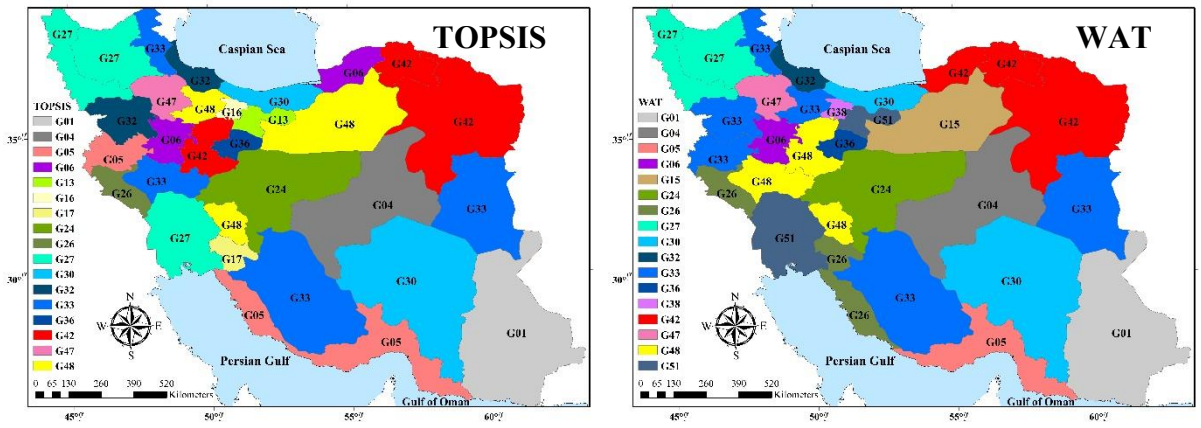


Figure 5. Best precipitation models acroos MCDM methods.

Frequency analysis across the 31 stations identifies a subset of consistently performing CMIP6 models. IITM-ESM (G33) is most frequent: top-ranked six times under WAT and seven times under CGT, and appears four times under CP and TOPSIS. MIROC-ES2L (G42), NORESM2-MM (G48), FGOALS-g3 (G27), and BCC-CSM2-MR (G05), each appearing three times among top ranks. Their repeated selection across MCDM algorithms indicates robust, generalizable precipitation-simulation performance across diverse climatic conditions.

Comparative rankings analysis reveals strong methodological correlations in some cases: CP and TOPSIS (both distance-based) agreed on top-model selections at 28 of 31 stations (exceptions: Bushehr, Gorgan, Semnan) and SAW (unweighted) corresponded with WAT (weighted) at eight stations including Ardabil, Kerman, Mashhad, and Shiraz.

Although CP, TOPSIS, and WAT identified 17 distinct top models (CGT identified 16), consensus was substantial: the same top-ranked model occurred across all four MCDM methods at 13 stations and across all five methods (including SAW) at 10 stations. Notably, the unweighted SAW top-ranked model differed from weighted MCDM results at 21 stations, supporting objective index weighting (entropy) and weighted MCDM for more accurate, spatially robust, and scientifically defensible model selection.

Ranking of models by the Group Decision-Making (GDM) method

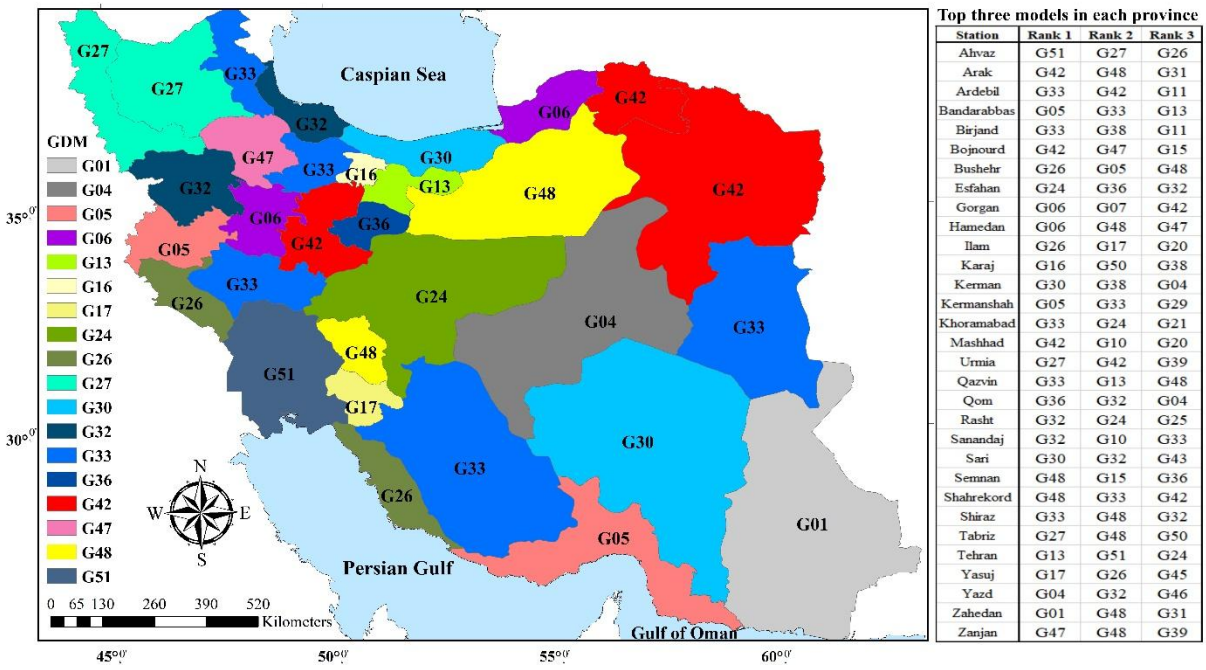
Although individual MCDM methods (WAT, CP, CGT, TOPSIS) produced consistent top-ranked models at some stations, their overall rankings differed due to distinct calculation procedures. To generate a single integrated prioritization, we applied a Group Decision-Making (GDM) technique to aggregate the MCDM results statistically.

GDM combined the individual MCDM rankings for T = 51 models, defining the Power group as ranks 1–26 and the Weakness group as ranks ≥27; final ranks were the Net Power score (Power minus Weakness). Table 7 lists each station's top model and Figure 6 maps the top three GDM-selected models. GDM identified IITM-ESM (G33) and MIROC-ES2L (G42) as overall superior performers: IITM-ESM (G33) placed first at five stations (Birjand, Shiraz, Khorramabad, Qazvin, Ardabil) and MIROC-ES2L (G42) at three stations (Mashhad, Bojnurd, Arak). Overall, GDM selected 18 different top models across 31 stations; IITM-ESM (G33) was second at three additional stations and NorESM2-MM (G48) was second at six stations, indicating a core subset of CMIP6 models that consistently perform well across Iran's climatic regimes.

Table 7. Characteristics of the best model at each station determined by GDM.

Station	Top model	NS_a	Top model rank in each method				Resolution (km^2)	Station	Top model	NS_a	Top model rank in each method				Resolution (km^2)
			WAT	CP	CGT	TOPSIS					WAT	CP	CGT	TOPSIS	
S01	UKESM1-0-LL	104	1	2	1	2	250	S17	FGOALS-g3	106	1	1	1	1	100
S02	MIROC-ES2L	104	1	2	1	2	500	S18	IITM-ESM	102	1	3	1	3	250
S03	IITM-ESM	105	1	1	2	1	250	S19	IPSL-CM5A2-	105	1	1	2	1	500
S04	BCC-CSM2-MR	106	1	1	1	1	100	S20	HadGEM3-GC31-MM	105	1	1	2	1	100
S05	IITM-ESM	106	1	1	1	1	250	S21	HadGEM3-GC31-MM	104	2	1	2	1	100
S06	MIROC-ES2L	106	1	1	1	1	500	S22	GISS-E2-1-H	106	1	1	1	1	250
S07	FGOALS-f3-L	104	1	2	1	2	100	S23	NorESM2-MM	104	2	1	2	1	100

S08	EC-Earth3-CC	106	1	1	1	1	100	S24	NorESM2-MM	104	2	1	2	1	100
S09	BCC-ESM1	102	2	2	2	2	250	S25	IITM-ESM	106	1	1	1	1	250
S10	BCC-ESM1	106	1	1	1	1	250	S26	FGOALS-g3	106	1	1	1	1	100
S11	FGOALS-f3-L	106	1	1	1	1	100	S27	CESM2-WACCM-	104	1	2	1	2	100
S12	CMCC-ESM2	104	2	1	2	1	100	S28	CNRM-CM6-1	105	2	1	1	1	250
S13	GISS-E2-1-H	105	1	1	2	1	250	S29	AWI-ESM-1-1-LR	105	1	1	2	1	250
S14	BCC-CSM2-MR	103	2	1	3	1	100	S30	ACCESS-CM2	106	1	1	1	1	250
S15	IITM-ESM	105	2	1	1	1	250	S31	NorCPM1	106	1	1	1	1	250
S16	MIROC-ES2L	106	1	1	1	1	500								



Top three models in each province

Station	Rank 1	Rank 2	Rank 3
Ahvaz	G51	G27	G26
Arak	G42	G48	G31
Ardebil	G33	G42	G11
Bandarabbas	G05	G33	G13
Birjand	G33	G38	G11
Bojnourd	G42	G47	G15
Bushehr	G26	G05	G48
Esfahan	G24	G36	G32
Gorgan	G06	G07	G42
Hamedan	G06	G48	G47
Ilam	G26	G17	G20
Karaj	G16	G50	G38
Kerman	G30	G38	G04
Kermanshah	G05	G33	G29
Khoramabad	G33	G24	G21
Mashhad	G42	G10	G20
Urmia	G27	G42	G39
Qazvin	G33	G13	G48
Qom	G36	G32	G04
Rasht	G32	G24	G25
Sanandaj	G32	G10	G33
Sari	G30	G32	G43
Semnan	G48	G15	G36
Shahrekord	G48	G33	G42
Shiraz	G33	G48	G32
Tabriz	G27	G48	G50
Tehran	G13	G51	G24
Yasuj	G17	G26	G45
Yazd	G04	G32	G46
Zahedan	G01	G48	G31
Zanjan	G47	G48	G39

Figure 6. Summary map of best precipitation-simulation models from GDM.

A spatial analysis of top-ranked models shows that model spatial resolution alone does not determine performance: the best-performing model had nominal resolution 100 km^2 at 13 stations, 250 km^2 at 14 stations, and 500 km^2 at 4 stations, indicating finer resolution does not consistently guarantee superior precipitation-simulation skill. Other factors—such as physical parameterizations, representation of atmospheric processes, and local climatic and topographic complexity—likely exert stronger influence on predictive accuracy than resolution alone.

Comparing our final GDM rankings with previous studies reveals substantial discrepancies attributable to methodological differences. Examples include ACCESS-CM2 (ranked 30 at Sari here) and MIROC-ES2L (ranked 41 at Birjand in a prior study) previously reported as top performers (Roshani and Hamidi 2022; Chamanehfar et al. 2023; Niroumand et al. 2023), and MRI-ESM2-0 ranking poorly at Kerman in Mianabadi et al. (2023). These divergences stem from variations in ensemble size and composition, model versions, choice and normalization of evaluation indices (e.g., ANMBD, NNSE, NKGE), use of objective weighting (entropy) and GDM aggregation here, and differing historical reference periods, underscoring the need for locally tailored, multi-criteria evaluation frameworks when selecting GCMs for regional impact assessments.

Correlation of Ranking Methods

When resources or time constrain application of multiple MCDM methods, a practical alternative is selecting a single method that approximates the integrated GDM result. To identify such a proxy, we computed correlation coefficients among ranking methods; results are summarized in Table 8 and mapped in Figure 7, which indicates the single best alternative per station based on highest correlation with GDM. Accordingly, the single MCDM method exhibiting the highest correlation with the integrated Group Decision-Making (GDM) ranking can serve as a robust, streamlined proxy for full GDM aggregation at a given station. The correlation analysis found GDM correlated most strongly with WAT ($\rho = 0.982$) and least with CGT ($\rho = 0.924$). CGT generally showed the lowest correlations with other techniques. High correlations included SAW (unweighted) with WAT (weighted) ($\rho = 0.992$) and CP with TOPSIS ($\rho = 0.998$). Station-wise, WAT was the most frequent single-best alternative (16 stations), followed by CP (8) and SAW (5). TOPSIS and CGT were each best at one station.

Table 8. Correlation matrix of ranking results for precipitation climate models.

R	CP	CGT	TOPSIS	WAT	SAW	GDM
CP	1					
CGT	0.836	1				
TOPSIS	**0.998**	0.826	1			
WAT	0.935	0.924	0.932	1		
SAW	0.920	0.917	0.916	**0.992**	1	
GDM	0.974	**0.924**	0.970	**0.982**	0.969	1

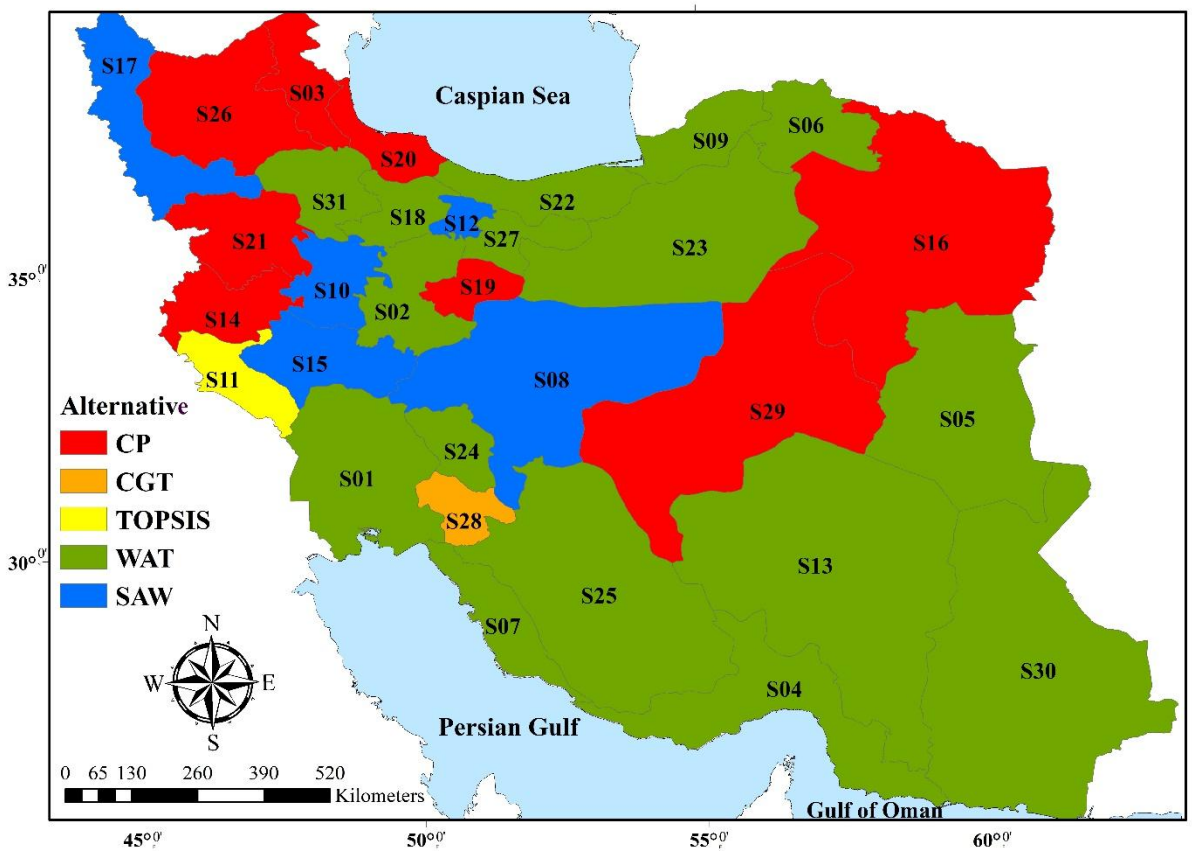


Figure 7. Spatial dispersion of the selected ranking method for precipitation climate models.

Future projection and ensemble trend analysis

Future precipitation analysis focused on five representative synoptic stations (Ahvaz, Bandarabbas, Esfahan, Mashhad, Tabriz) to cover major Iranian climatic regimes while limiting computational load.

Projections cover 2025–2069, matching the historical baseline length (1970–2014), under SSP2-4.5 and SSP5-8.5. For each station, ensembles combined the three top-ranked GDM

models with weights from the rank-based scheme of Brekke et al. (2012), giving greater weight to better historical performers.

GCM outputs were downscaled to station scale via the Change Factor (CF) method, applying projected relative change (future ensemble mean / historical ensemble mean) multiplicatively to station's historical observations. Table 9 summarizes the selected GDM models per station, ensemble weights, and key historical and projected statistics for SSP2-4.5 and SSP5-8.5.

Table 9. Projected precipitation changes for the future period.

Station	Top models	Model weight	Obs		SSP2-4.5			SSP5-8.5		
			Avg	Max	CV	Avg	Max	CV	Avg	Max
Ahvaz	UKESM1-0-LL	0.4	18.87	223.11	1.03	19.42	229.58	1.11	21.01	248.34
	FGOALS-g3	0.27								
	FGOALS-f3-L	0.33								
Bandarabbas	BCC-CSM2-MR	0.47	14.86	263	1	14.83	262.54	0.89	13.16	232.94
	IITM-ESM	0.53								
	HadGEM3-GC31-LL	0								
Esfahan	EC-Earth3-CC	0.6	10.89	94.41	1.1	12.02	104.15	1.19	12.94	112.14
	UKESM1-0-LL	0.2								
	HadGEM3-GC31-LL	0.2								
Mashhad	MIROC-ES2L	0.6	21.21	132	1.06	22.43	139.59	1.06	22.48	139.88
	CESM2	0.27								
	IITM-ESM	0.13								
Tabriz	FGOALS-g3	0.47	22.04	128.43	0.99	21.81	127.09	0.98	21.62	125.96
	NorESM2-MM	0.4								
	TaiESM1	0.13								

According to Table 9, the highest-weighted models in each station ensemble were: UKESM1-0-LL for Ahvaz (weight = 0.40); IITM-ESM (G33) for Bandarabbas (weight = 0.53); EC-Earth3-CC for Esfahan (weight = 0.60); MIROC-ES2L (G42) for Mashhad (weight = 0.60); and FGOALS-g3 (G27) for Tabriz (weight = 0.47). In Bandarabbas, the third-ranked model received zero weight and thus did not contribute to the ensemble. Observed mean monthly precipitation during the baseline (1970-2014) ranged from 10.89 mm at Esfahan to 22.04 mm at Tabriz.

Under SSP2-4.5, Ahvaz, Esfahan, and Mashhad are projected to have increased mean precipitation for 2025–2069; Tabriz is projected to decrease, and Bandarabbas shows no significant change. The largest relative increase under SSP2-4.5 is ~10% at Esfahan. Under SSP5-8.5, Ahvaz, Esfahan, and Mashhad again show increases, while Bandarabbas and Tabriz are projected to decrease; the largest increase is ~19% at Esfahan and the largest decrease is ~11% at Bandarabbas. These directions align with Niroumand et al. (2023).

Conclusion

Accurate assessment and selection of reliable global climate models (GCMs) are essential for credible climate-change projections. This study evaluated the full CMIP6 ensemble (51 models) for monthly precipitation against observations from 31 Iranian synoptic stations (1970–2014) to identify best-performing models and analyze future precipitation trends using multiple statistical indices and advanced multi-criteria decision-making (MCDM) techniques.

The evaluation sequence standardized five performance indices (R^2, ANMBD, NRMSE, NNSE, NKGE), derived objective weights via the Entropy method, applied multiple MCDM ranking methods (SAW, WAT, CP, CGT, TOPSIS), and aggregating results using a Group Decision-Making (GDM) procedure to produce a final, robust ranking.

Entropy weighting identified NRMSE and R^2 as the most influential indices for discriminating model skill, receiving the highest objective weights across stations, while ANMBD consistently received the lowest weight, indicating limited discriminatory power; thus ANMBD is not recommended as a sole evaluation metric. GDM-aggregated rankings show substantial spatial heterogeneity: no single CMIP6 model is uniformly optimal nationwide, although IITM-ESM (G33) and MIROC-ES2L (G42) are the most accurate and robust across the largest number of stations.

A key implication is the need for regionally tailored model selection: local climatic and geographic characteristics (e.g., topography, proximity to moisture sources, regional circulation patterns) strongly modulate model performance; so local evaluation yields more reliable regional impact assessments than reliance on global-average metrics.

For future projections, weighted ensembles were constructed per station from the three top-ranked GDM models. Projections for 2025–2069 (matching the historical baseline 1970–2014) under SSP2-4.5 and SSP5-8.5 were downscaled using the Change Factor (CF) method. Trend analysis shows marked spatial variability-for example, Esfahan projects the largest increase (+19% under SSP5-8.5) and Bandarabbas the largest decrease (−11%)-supporting the station-specific, MCDM-driven selection approach.

Remaining uncertainties stem from intrinsic GCM limitations, observational data quality, downscaling assumptions, and methodological choices. To improve robustness, we recommend: (1) applying objective MCDM approaches (notably GDM) to synthesize multiple rankings; (2) using normalized, complementary performance indices (e.g., NNSE, NKGE, NRMSE, R^2); and (3) constructing weighted ensembles based on historical skill to reduce projection uncertainty.

For operational use, GDM results identify CMIP6 models recommended and not recommended for regional precipitation studies in Iran; this resource aids practitioners and researchers in selecting models likely to yield more reliable regional projections.

Suitable models are: NorESM2-MM (G48); IITM-ESM (G33); MIROC-ES2L (G42); HadGEM3-GC31-MM (G32); FGOALS-f3-L (G26); EC-Earth3-CC (G24); NorCPM1 (G47); BCC-CSM2-MR (G05).

Unsuitable models are: ACCESS-ESM1-5 (G02); CESM2-WACCM (G12); CMCC-CM2-HR4 (G14); CNRM-ESM2-1 (G19); E3SM-1-1-ECA (G22); AWI-CM-1-1-MR (G03); INM-

CM5-0 (G35); MIROC6 (G40); MIROC-ES2H (G41); MPI-ESM1-2-LR (G44); IPSL-CM6A-LR (G37); CanESM5 (G08).

Statements and Declarations

Funding

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

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

Conceptualization: E.M.; Data curation and Visualization: S.D.; Formal analysis and Validation: E.M.; Writing – original draft: S.D.; Writing – review & editing: E.M.

All authors have read and agreed to the published version of the manuscript.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the large size of the file but are available from the corresponding author on reasonable request.

Ethics Approval

Not applicable.

Consent to Participate

Not applicable.

Consent for Publication

All authors consent to the publication of this manuscript in Climatic Change.

References

Ahmadi M, Kamangar M (2023) Projection the Extreme Temperatures of Iran's Cold Season Using a Multi-Model Hamadi System with Socio-Economic Scenarios of CMIP6 Models. Nivar, 47(120-121), 71-88.(In Persian) 10.30467/nivar.2023.389079.1240

Ashraf M, Arshad A, Patel PM, Khan A, Qamar H, Siti-Sundari R, Ghani M, Amin A, Babar JR (2021) Quantifying climate-induced drought risk to livelihood and mitigation actions in Balochistan. Natural Hazards, 109, 2127-2151. 10.1007/s11069-021-04913-4

Babaeian I, Modirian R, Khazanedari L, Karimian M, Kouzegaran S, Kouhi M, Falamarzi Y, Malbusi S (2023). Projection of Iran’s precipitation in 21st Century using downscaling of selected CMIP6 Models by CMHyd. Journal of the Earth and Space Physics, 49(2), 431-449. (In Persian) doi: 10.22059/jesphys.2023.332410.1007436

Brekke L, Kiang J, Olsen J, Pulwarty R, Raff D, Turnipseed D, Webb R, White K (2012) Climate change and water resources management: A federal perspective (U.S. Geological Survey Circular 1331). U.S. Geological Survey. https://doi.org/10.3133/cir1331

Carvalho D, Rafael S, Monteiro A, Rodrigues V, Lopes M, Rocha A (2022) How well have CMIP3, CMIP5 and CMIP6 future climate projections portrayed the recently observed warming. Scientific Reports, 12(1), 11983. 10.1038/s41598-022-16264-6

Chamanehfar S, Mousavi Baygi M, Babaeian I, Modaresi F (2023). Future projection for extreme indices of precipitation and temperature over the period 2026-2100 based on the output

485 of CMIP6 models (Case study: Mashhad), Iranian Journal of Irrigation & Drainage, 16(5), 963-
486 976. (In Persian) DOI: 20.1001.1.20087942.1401.16.5.7.2
487 Cheng TF, Lu M, Dai L (2019) The zonal oscillation and the driving mechanisms of the extreme
488 western North Pacific subtropical high and its impacts on East Asian summer
489 precipitation. Journal of Climate, 32(10), 3025-3050. DOI:10.1175/JCLI-D-18-0076.1
490 Choobeh S, Farrokhzadeh B, Bazrafshan O, Hasanvand H (2025). Impact of Climate Change
491 on Precipitation Extremes Indices over Iran. Watershed Management Research, 37(4), 34-53.
492 (In Persian) doi: 10.22092/wmrj.2024.364521.1565
493 Ersoy N (2021) Selecting the best normalization technique for ROV method: Towards a real
494 life application. Gazi University Journal of Science, 34(2), 592-609. 10.35378/gujs.767525
495 Heydarizad M, Raeisi E, Sorí R, Gimeno L (2019) Developing meteoric water lines for Iran
496 based on air masses and moisture sources. Water, 11(11), 2359. 10.3390/w11112359
497 Mahabadi S, Dehban H, Zareian M, Farokhnia A (2022) Investigation of temperature and
498 precipitation changes in the Iran's basins in the next 20 years based on the output of CMIP6
499 model. IRANIAN WATER RESEARCH JOURNAL, 16(1 (44)), 11-24. SID. (In Persian)
500 https://sid.ir/paper/1057766/en
501 Masson-Delmotte VP, Zhai P, Pirani SL, Connors C, Péan S, Berger N, Scheel Monteiro PM
502 (2021) IPCC (2021) Summary for policymakers. in: Climate change 2021: The physical science
503 basis. contribution of working group i to the sixth assessment report of the intergovernmental
504 panel on climate change https://doi.org/10.1017/9781009157896
505 Mianabadi A, Bateni M, Mohammadi S (2023). Projection of Change in the Distribution of
506 Precipitation and Temperature Using Bias-Corrected Simulations of CMIP6 Climate Models
507 (Case Study: Kerman Synoptic Station), Journal of Climate Change Research, 4(14), 65-84. (In
508 Persian) DOI: 10.30488/CCR.2023.399780.1139

509 Moghaddas P, Mahjoobi E, Gharehchalo S (2024) Prioritization of the CMIP6 general
510 circulation models using multi-criteria decision-making methods in the Nekarood watershed.
511 Iranian Journal of Irrigation and Drainage, 18(1), 15-25. (In Persian)
512 https://idj.iaid.ir/article_183799.html?lang=fa
513 Mohsen KA, Nile BK, Hassan WH (2020) Experimental work on improving the efficiency of
514 storm networks using a new galley design filter bucket. In IOP Conference Series: Materials
515 Science and Engineering (Vol. 671, No. 1, p. 012094). IOP Publishing. 10.1088/1757-
516 899X/671/1/012094
517 Morais DC, de Almeida AT (2012) Group decision making on water resources based on
518 analysis of individual rankings. Omega, 40(1), 42-52.
519 https://doi.org/10.1016/j.omega.2011.03.005
520 Muhaisen NK, Khayyun TS, Al Mukhtar M, Hassan WH (2024) Forecasting changes in
521 precipitation and temperatures of a regional watershed in Northern Iraq using LARS-WG
522 model. Open Engineering, 14(1), 20220567. 10.1515/eng-2022-0567
523 Mukheef RA, Hassan WH, Alquzweeni SJAR (2024) Projections of temperature and
524 precipitation trends using CMhyd under CMIP6 scenarios: A case study of Iraq's Middle and
525 West. Atmospheric Research, 306, 107470. 10.1016/j.atmosres.2024.107470
526 Niroumandfard F, Khashei A, Hashemi R, Ghorbani K (2023). Investigation of Climate Change
527 Projection on Temperature and Precipitation Parameters Using CMIP6 Models (Case Study:
528 Birjand Station), Iranian Journal of Soil and Water Research, 53(9), 2009-2026. (In
529 Persian) DOI:https://doi.org/10.22059/ijswr.2022.343936.669284
530 Nossent J, Bauwens W (2012) Application of a normalized Nash-Sutcliffe efficiency to
531 improve the accuracy of the Sobol'sensitivity analysis of a hydrological model. In EGU general
532 assembly conference abstracts (p. 237).

Raju KS, Nagesh Kumar D (2015) Ranking general circulation models for India using TOPSIS. J Water Clim Change, 6(2):288-299

Roshani A, Hamidi M (2022) Forecasting the effects of climate change scenarios on temperature and precipitation based on CMIP6 models (Case study: Sari station). Water and Irrigation Management, 11(4), 781-795. (In Persian) doi: 10.22059/jwim.2022.330603.920

Shannon CE (1948) A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sharma A, Kale GD (2022) Ranking of general circulation models for Surat City by using a hybrid approach. Water Practice and Technology, 17(10), 2186-2198. 10.2166/wpt.2022.118

Vassoney E, Mammoliti Mochet A, Desiderio E, Negro G, Pilloni MG, Comoglio C (2021) Comparing multi-criteria decision-making methods for the assessment of flow release scenarios from small hydropower plants in the alpine area. Frontiers in Environmental Science, 9, 635100. https://doi.org/10.3389/fenvs.2021.635100

WMO, (2021) State of the Global Climate 2021.

Yazdandoost F, Moradian S, Izadi A, Aghakouchak A (2020). Evaluation of CMIP6 precipitation simulations across different climatic zones: Uncertainty and model intercomparison. Atmospheric Research. 250.10.1016. https://doi.org/10.1016/j.atmosres.2020.105369

Yilmaz B, Aras E, Nacar S (2024) A CMIP6-ensemble-based evaluation of precipitation and temeperature projections. Theoretical and Applied Climatology, 155:7377-7401. https://doi.org/10.1007/s00704-024-05066-7

Zareian M, Dehban E, Gohari S (2022) Evaluation of the Accuracy of CMIP6 Models in Estimating Temperature and Precipitation in Iran Based on Network Analysis. Water and Irrigation Management, 12(4), 783-797.(In Persian) 10.22059/jwim.2022.345975.1006

Zavadskas EK, Turskis Z (2008) A new logarithmic normalization method in games theory. Informatica, 19(2), 303-314. 10.15388/Informatica.2008.215

559 Zeleny M (1998) Multiple criteria decision making: Eight concepts of optimality. Human
560 Systems Management, 17(2), 97-107. https://doi.org/10.1007/978-3-642-60667-0_19
561 Zhu YY, Yang S (2020) Evaluation of CMIP6 for historical temperature and precipitation over
562 the Tibetan Plateau and its comparison with CMIP5. Advances in Climate Change
563 Research, 11(3), 239-251. 10.1016/j.accre.2020.08.001