Water and Soil Management and Modeling

Online ISSN: 2783 - 2546

University of Mohaghegh Ardabili

Performance of the XGBoost algorithm in downscaling temperature and relative humidity in a temperate climate: a case study in Kermanshah

Mohammad Fouladi Nasrabad [1], **Mohsen Pourreza Bilondi** *[2], **Mahdi Amirabadizadeh**[2], **Mahna Javaheri** [3]

[1] Ph.D. Candidate in Water Resources, Department of Water Engineering, University of Birjand, Birjand, Iran
[2] Associate Professor, Department of Water Engineering, University of Birjand, Birjand, Iran
[3] M.Sc student in Water Resources, Department of Water Engineering, University of Birjand, Birjand, Iran

Abstract

Climate change requires a precise analysis of local data, and statistical downscaling using machine learning algorithms such as XGBoost can enhance the spatial accuracy of General Circulation Models (GCMs). This study examined the performance of XGBoost in the daily downscaling of temperature and relative humidity at the synoptic station of Kermanshah during the period from 1990 to 2014. In order to investigate the performance of the XGBoost model in downscaling the climate variables of temperature and relative humidity, local and large-scale data were divided into two training and testing sections, so that the years 1990 to 2007 were considered as the training section and the years 2007 to 2015 were considered as the testing section. The present research showed that the XGBoost algorithm, as one of the advanced machine learning methods, performed very well in downscaling climatic parameters, especially temperature, and to some extent relative humidity. To evaluate the model's performance, the metric values were examined separately in two sections: training and testing. For temperature using ECMWF-ERA5 data, the KGE, NSE, and R^2 values in the training section were 0.98, 0.98, and 0.99, respectively. For temperature using MPI-ESM1-2-HR data, the values of these metrics in the training section were 0.93, 0.94, and 0.97, and in the testing section, they were 0.91, 0.88, and 0.94, respectively. Additionally, for relative humidity using ECMWF-ERA5 data, the metric values in the training section were 0.82, 0.93, and 0.97, and in the testing section, they were 0.7, 0.72, and 0.87. Finally, for relative humidity using MPI-ESM1-2-HR data, the values of these metrics in the training section were 0.72, 0.67, and 0.82, and in the testing section, they were 0.70, 0.65, and 0.81. Graphical analyses confirmed the superiority of the model in simulating intermediate and extreme temperature values, especially with ECMWF-ERA5, but limitations in reproducing extreme values were observed in relative humidity.

Keywords: Correction of statistical errors, Climate change, General circulation models, Statistical indices, Machine learning

Article Type: Research Article

Academic Editor: Zainab Hazbavi

*Corresponding Author, E-mail: mohsen.pourreza@birjand.ac.ir

Citation: Fouladi Nasrabad, M., Purreza Bilondi, M., Amirabadizadeh, M., & Javaheri, M. (2025). Performance of XGBoost algorithm in downscaling temperature and relative humidity in temperate climate: a case study in Kermanshah, 5 (Special Issue: Climate Change and Effects on Water and Soil), 215-232.
doi: 10.22098/mmws.2025.18179.1661
Received: 29 August 2025, Received in revised form: 25 September 2025, Accepted: 26 September 2025, Published online: 01 Novamber 2025
Water and Soil Management and Modeling, Year 2025, Vol. 5, Special Issue, pp. 215-232
Publisher: University of Mohaghegh Ardabili
© Author(s)

1. Introduction

Climate change is considered one of the most significant global challenges of the 21st century, with its effects on water resources, agriculture, and human settlements becoming increasingly evident (IPCC, 2021). General Circulation Models (GCMs) are common and reliable tools for predicting future climate conditions, providing climatic information on a global scale. However, one of the fundamental limitations of these models is their low spatial resolution, which results in insufficient accuracy at regional or local scales, particularly for hydrological analyses or Water resource planning may lack the necessary accuracy (Maraun et al., 2019; Chen et al., 2020; Fouladi Nasrabad et al., 2024). To overcome this limitation, downscaling techniques are used. Downscaling methods are generally divided into two categories: dynamic and statistical. The dynamic method includes the use of regional climate models (RCMs) that convert large-scale GCM data to smaller scales and can model physical processes more accurately. In contrast, the statistical method is based on statistical relationships between large-scale and local variables and includes techniques such as regression, weather typing methods, and weather generators. These methods are flexible and low-cost and can be applied to various climatic variables (Wilby & Wigley, 1997; Maraun et al., 2010). Nevertheless, downscaling methods have limitations; dynamic methods require high computational resources and may transfer the errors of GCMs to local scales, while statistical methods rely on the assumption of stationary statistical relationships in the future, which may be violated under conditions of severe climate change, and may also be inefficient in reproducing extreme events (Maraun & Widmann, 2018; Gutiérrez et al., 2019). In this context, statistical downscaling (SD) has been widely used in climate studies due to its lower computational resource requirements and high flexibility (Anandhi et al., 2018; Maraun et al., 2019). The fundamental assumption in SD is that the statistical relationships between GCM predictors and local variables will remain stable in the future. Although this assumption may be violated under conditions of extreme climate change, SD remains a valuable tool in producing local climate data (Maraun & Widmann, 2018). Statistical techniques range from simple methods like linear regression to more complex methods such as artificial neural networks (ANN), support vector machines (SVM), and random forests (RF) (Okkan & Kirdemir, 2018; Nourani et al., 2019). In the past decade, with the expansion of data science, machine learning (ML) models have gained a special position in downsizing. These models are capable of modeling nonlinear and complex relationships between climatic variables and can outperform traditional methods (Reichstein et al., 2019; Giri et al., 2021). Extreme Gradient Boosting (XGBoost) is a machine learning algorithm based on gradient boosting that improves model performance sequentially using decision trees as base learners. Introduced by Chen and Guestrin (2016), this algorithm offers scalability for large datasets with features such as scatter-aware algorithms for sparse data and a weighted quantile sketch for approximate tree learning. XGBoost supports parallel processing to accelerate computations, regularization techniques to reduce overfitting, and automated handling of missing data, making it suitable for complex tasks such as regression, classification, and ranking (Chen and Guestrin, 2016). In the field of water resources engineering, Niazkar and colleagues (2024) reported that XGBoost outperformed traditional models in 74% of the 72 studies reviewed between 2018 and 2023. However, there are limitations such as dependence on data quality and quantity, inability to provide explicit mathematical equations due to its tree structure, and the need for hyperparameter tuning. Downscaling climatic parameters to improve the spatial resolution of satellite data and numerical prediction models is essential, especially in hydrological applications and water resource management. XGBoost has increasingly been used in this field due to its ability to model complex nonlinear relationships. The systematic review by Niazkar et al. (2024) examines the applications of XGBoost in water resources engineering and highlights its use in the statistical downscaling of hydrometeorological variables, such as water storage, precipitation, and soil moisture. In the literature examining XGBoost's performance for downscaling temperature and relative humidity in temperate climates (illustrated by the Kermanshah case study), several key studies demonstrate its effectiveness across related environmental applications. Ali et al. (2023) used XGBoost to downscale GRACE-derived terrestrial water storage anomaly (TWSA) from 1° to 0.25° in the Indus Basin Irrigation System, outperforming artificial neural networks with metrics such as NSE = 0.99, Pearson R = 0.99, RMSE = 5.22 mm, and MAE = 2.75 mm, and enabling calculation of WSDI and WSD for precise drought identification (2010–2016). Dong et al. (2023) applied XGBoost for multivariate bias correction of GEFSv2 numerical weather prediction forecasts across seven climatic regions in China, improving short-term precipitation forecasts—especially in winter—and surpassing traditional methods like empirical distribution function matching (EDCDFm). Zhang et al. (2023) downscaled satellite-derived soil moisture obtained via GNSS-R, refining SMAP resolution from 36 km to 3 km in a humid subtropical area (≈834.45 mm annual precipitation) to support local agricultural and hydrological management. Georgiades et al. (2025) developed an XGBoost-based model to temporally downscale daily climate data to hourly Temperature–Humidity Index (THI) for global dairy cattle heat-stress assessments, training on ERA5 and projecting under CMIP6 scenarios (SSP2-4.5 and SSP5-8.5) with strong validation (MAE ≈ 3.4, MSE ≈ 19, R^2 ≈ 0.94) while noting limitations in complex mountainous terrain. Zhu et al. (2025) provided a comprehensive review highlighting XGBoost's growing role in spatial downscaling of satellite precipitation and its ability to capture nonlinear relationships among hydrometeorological variables—often enhanced by post-processing such as residual correction for regional applications like flood prediction in temperate zones. Khosravi et al. (2025) proposed a stacking ensemble that incorporates XGBoost to improve CMIP6-based projections of maximum/minimum temperature and precipitation under multiple SSPs, reporting very high R^2 values (up to 0.99 for temperatures) and projecting substantial warming (up to +8°C) and precipitation shifts for Iran and nearby arid regions, collectively underscoring XGBoost's utility for localized downscaling in climate-vulnerable temperate areas. These studies demonstrate the adaptability of XGBoost in downscaling climatic parameters. The ability of this algorithm to manage nonlinear relationships and complex data has made it a powerful tool for generating high-resolution climatic data that is essential for water resource planning, agricultural management, and climate change adaptation strategies. However, the success of these applications depends on the quality of input data and the appropriate selection of independent variables, such as elevation, land cover type, and climatic data. In the field of Iran's climate, numerous studies have focused on downscaling precipitation and temperature using machine learning algorithms, but they have often been limited to older methods such as ANN and SVM (Zare et al., 2021; Ghahreman et al., 2022; Daneshkhah et al., 2020). A review of available resources indicates that the use of XGBoost for downscaling climatic parameters in different climatic regions of Iran (arid and humid) has not yet been widely implemented and is considered a significant research gap (Parsa et al., 2023).

Accordingly, the present research aims to fill this gap by examining the efficiency of the XGBoost model for downscaling climatic parameters of temperature and relative humidity, using outputs from GCMs in a region with a dry climate. In this study, the performance of the XGBoost model in downscaling temperature and relative humidity parameters is assessed through statistical analyses to accurately evaluate its capability in reproducing climate variables.

2. Materials and Methods

2.1. Study Area:

The study area of the synoptic station of Kermanshah is located in Kermanshah province in the west of the country, with an area of 64.24 square kilometers. This province extends between latitudes 33° 54′ to 35° 15′ north and longitudes 45° 24′ to 48° 30′ east. The Kermanshah synoptic station is situated at the latitude and longitude of 34° 37′ and 47° 9′, and its location is shown in Figure 1 (Almasi et al., 2024).

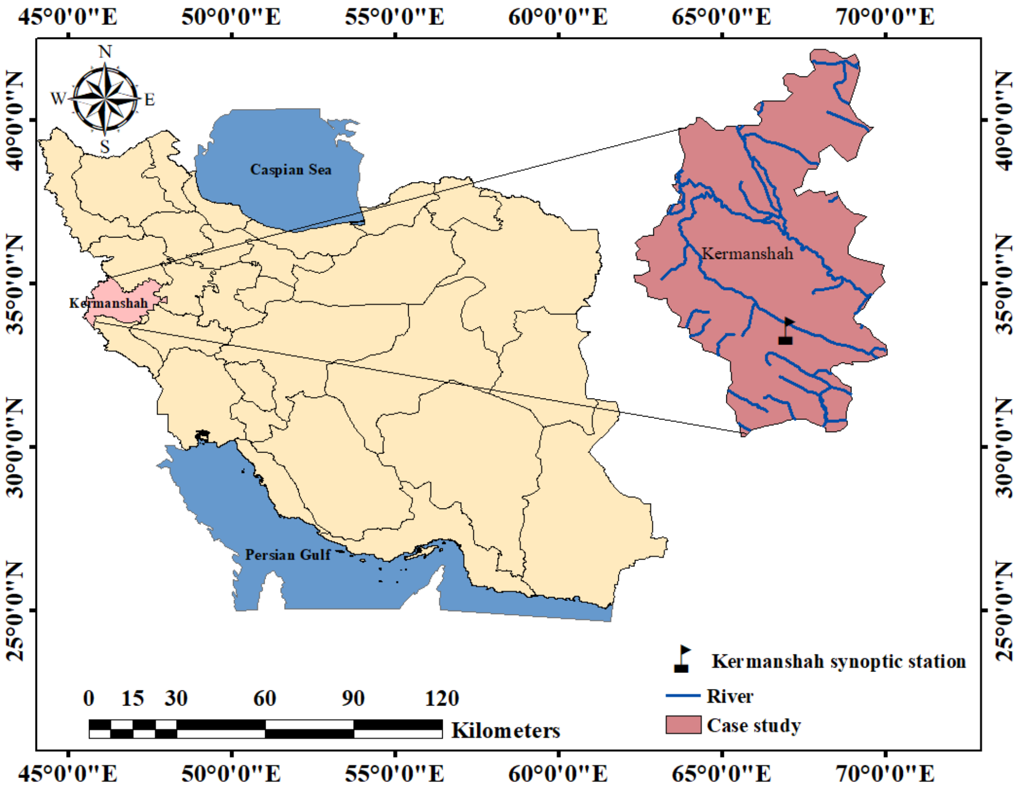


Figure 1. Study area Kermanshah synoptic station

2.2. Research method:

This study aimed to investigate the efficiency of machine learning models in downscaling the climate parameters of temperature and relative humidity on a daily scale at the Kermanshah synoptic station. To extract large-scale features, historical data from the general atmospheric circulation models ECMWF-ERA5 and MPI-ESM1-2-HR were used and downloaded from the site https://climate-scenarios.canada.ca, appropriate for the location of the Kermanshah synoptic station. In this study, daily relative humidity and temperature data from the Kermanshah synoptic station between 1990 and 2014 were used. In order to investigate the performance of the XGBoost model in downscaling the climate variables of temperature and relative humidity, local and large-scale data were divided into two training and testing sections, so that the years 1990 to 2007 were considered as the training section and the years 2007 to 2015 were considered as the testing section. Then, in order to reduce the systematic and inherent errors in the large-scale data, quantile mapping was applied separately for each of the 23 GCM parameters and for the historical reference period to approximate the statistical distribution of the model to the statistical distribution of the observations. This operation was performed separately for each of the two climate variables, temperature and humidity. The model performance evaluation was also carried out in two stages: training and testing, based on criteria such as Kling-Gupta efficiency (KGE), Nash-Sutcliffe efficiency (NSE), normalized root mean square error (NRMSE), and R^2 coefficient of determination Gupta et al., 2009; Nash and Sutcliffe, 1970). The implementation of machine learning models and data processing was done using the XGBoost, scikit-learn, and numpy packages in the Python programming environment.

2.2.1. General circulation models of the atmosphere

General Circulation Models (GCMs) serve as important tools for climate prediction and understanding global changes. The ECMWF-ERA5 model is one of the climate prediction models that provides large-scale data for downscaling precipitation. Data from GCMs were downloaded from the website https://climate-scenarios.canada.ca/, and then the performance of the XGBOOST model was evaluated using each of the ECMWF-ERA5 and MPI-ESM1-2-HR models at the synoptic station of Kermanshah based on the evaluation criteria of NRMSE, NSE, R2, and KGE in both training and testing phases. This model is commonly used in climate studies due to its availability and global coverage (Wang et al., 2021). The ECMWF-ERA5 data includes various parameters such as sea level pressure, temperature, relative humidity, and wind speed, which are used as inputs for the downscaling model. These parameters are generally used as predictor variables in climatic downscaling models. (Sachindra et al., 2018)

The MPI-ESM1-2-HR (Max Planck Institute Earth System Model) is one of the advanced GCMs developed by the Max Planck Institute for Meteorology in Germany. This model, utilizing sub-models of the atmosphere, ocean, lakes, sea ice, and biosphere, is capable of simulating the complex interactions of Earth's climate. MPI-ESM1-2-HR comes in different versions with various spatial resolutions and is used for long-term climate change studies and future climate scenario simulations. The output data of this model includes various parameters such as temperature, humidity, pressure, and atmospheric circulation, which are also applicable in statistical downscaling studies (Muller et al., 2018). The accuracy and stability of the MPI-ESM1-2-HR model have led it to be recognized as one of the reference models in many international projects, such as CMIP5 and CMIP6. In Table 1, the information and specifications related to the 23 parameters of the GCMs used in this research are provided.

Table 1. Specifications of 23 parameters of standardized GCMs

Number	Predictor	Description	Acronym Type	Level (hpa)	Unit
1	Mean sea level pressure	Mslp	Circulation	Surface	hPa
2	Wind speed	p1_f		1000	m/s
3	Zonal wind component	p1_u		1000	m/s
4	Meridional wind component	p1_v		1000	m/s
5	Relative vorticity of true wind	p1_z		1000	s^{-1}
6	Divergence of true wind	p1_zh		1000	degrees
7	Air temperature at 2 m	temp	Temperature	Surface	s^{-1}
8	Specific humidity	Shum	Humidity	1000	K
9	Precipitation	prep		Surface	kg/kg
10	Wind speed	p5_f	Circulation	500	$kg/m^2/s$
11	Zonal wind component	p5_u		500	m/s
12	Meridional wind component	p5_v		500	m/s
13	Relative vorticity of true wind	p5_z		500	m/s
14	Divergence of true wind	p5_zh		500	s^{-1}
15	Specific humidity	Shum	Humidity	500	degrees
16	Geopotential	p500	Circulation	500	s^{-1}
17	Wind speed	p8_f		850	kg/kg
18	Zonal wind component	p8_u		850	m
19	Meridional wind component	p8_v		850	m/s
20	Relative vorticity of true wind	p8_z		850	m/s
21	Divergence of true wind	p8_zh		850	m/s
22	Specific humidity	Shum	Humidity	850	s^{-1}
23	Geopotential	p850	Circulation	850	degrees

2.2.2. XGBoost model

The extreme gradient boosting (XGBoost) algorithm is selected as an advanced machine learning model for exponential downscaling of temperature and relative humidity. XGBoost is suitable for this study due to its ability to model complex nonlinear relationships and its high accuracy in spatiotemporal predictions (Chen and Guestrin, 2016). XGBoost uses a set of weak decision trees and achieves a robust model by gradually optimizing them. The mathematical relationships in XGBoost are based on the loss function and gradient optimization, which are expressed as follows:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (1)$$

In which $l(y_i, \hat{y}_i)$ is the loss function, $\hat{y}_i$ is the model's prediction, and $\Omega(f_k)$is the cost of each tree. In this relationship, $l(y_i, \hat{y}_i)$ is typically represented by the squared error function, which is $(y_i - \hat{y}_i)^2$. $\Omega(f_k)$ also includes two main components: $\gamma T + \frac{1}{2}\lambda||w||^2$, where γ is the cost for each node, T is the number of nodes, λ is the regularization coefficient, and w is the weight of each node (Chen & Guestrin, 2016).

2.2.3 Evaluation criteria

The KGE, NSE, NRMSE, and R² criteria are widely used in hydrological studies to assess the accuracy of models. These criteria help us evaluate the accuracy of the model's predictions in comparison with observational data. Next, detailed explanations regarding each of these criteria will be provided.

KGE: This metric is defined based on three components: scale error, distribution error, and delay error, and is calculated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (2)$$

In which r is the correlation coefficient, α is the ratio of variances, and β is the mean of the relative values (Gupta et al., 2009). The ideal value for KGE is close to 1.

NSE: This criterion is calculated as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (3)$$

where y_i is the observed data, $\hat{y}_i$ is the predicted data, and $\bar{y}$ is the mean of the observed data (Nash & Sutcliffe, 1970). The ideal value for NSE is between 0 and 1, with values close to 1 indicating high accuracy of the model.

RMSE: This criterion is calculated as follows:

$$RMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}}{\bar{O}} \quad (4)$$

where y_i is the observed data and $\hat{y}_i$ is the predicted data, and $\bar{O}$ is the mean of the observed data. (Willmott & Matsuura, 2005). The ideal value for NRMSE is close to 0.

R^2: This criterion is calculated as follows:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2}} \right)^2 \quad (5)$$

where y_i represents the observed data, $\hat{y}_i$ represents the predicted data, $\bar{y}$ is the mean of the observed data, and $\bar{\hat{y}}$ is the mean of the predicted data (Cohen et al., 2013). The ideal value for R² ranges from 0 to 1, and values close to 1 indicate high accuracy of the model.

3. ResultsandDiscussion

The present study, utilizing historical data from climate models, showed that the XGBoost algorithm, as one of the advanced machine learning methods, performs very well in downscaling climatic parameters, especially temperature and, to some extent, relative humidity. The results obtained from evaluating the model using statistical criteria in both training (1990–2007) and testing (2007–2014) phases indicate the desirable accuracy and stability of this algorithm. In particular, the superior performance of XGBoost in reproducing intermediate and extreme temperature values compared to relative humidity and the relative advantage of ECMWF-ERA5 data over the MPI-ESM1-2-HR model are noteworthy.

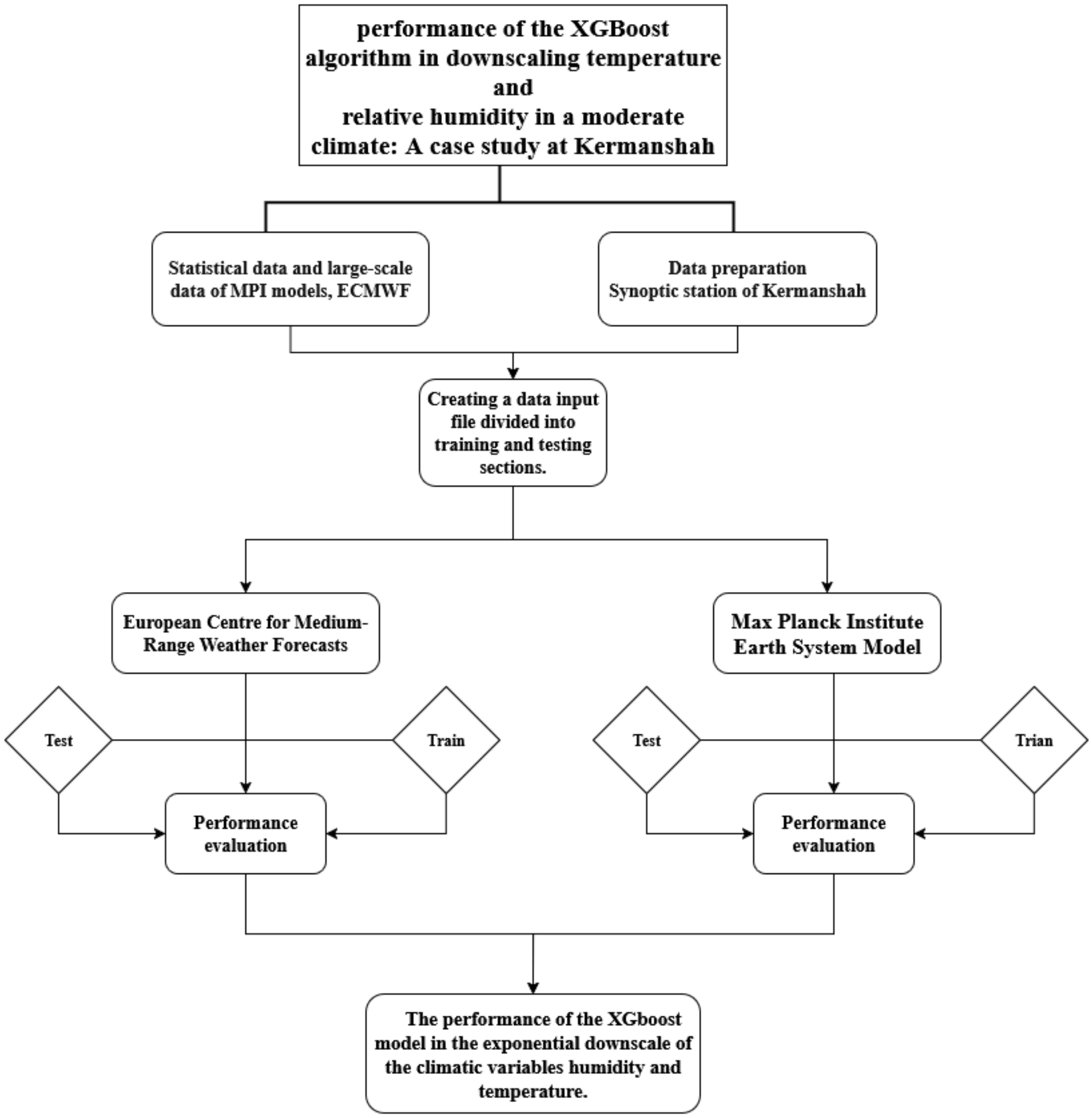


Figure 2. Flowchart of research steps

Although the model's performance in simulating extreme values of relative humidity has encountered some limitations, the overall results confirm the use of XGBoost as an effective tool in improving the spatial and temporal accuracy of climatic data in the dry regions of Iran. On the other hand, the sensitivity of the model's performance to the type of input variables from GCMs highlight the importance of accurately selecting predictor parameters in the downscaling process. Therefore, it is recommended that in future studies, the evaluation of the XGBoost model in more diverse climatic regions, the use of multiple stations, the combination with deep learning methods, and the examination of the model's response to various climate change scenarios be considered (Mauritsen et al., 2019).

Table 2. values of the evaluation criteria in the training and testing section of relative humidity and mean temperature parameters

(Evaluation criteria)	MPI-ESM1-2-HR-ESM1-2-HR				ECMWF-ERA5-ERA5			
	(Temperature)		(Relative humidity)		(Temperature)		(Relative humidity)	
	Train	Test	Train	Test	Train	Test	Train	Test
NRMSE	5.04	7.46	14.07	13.81	2.28	2.81	5.95	13.17
R^2	0.97	0.94	0.82	0.81	0.99	0.99	0.97	0.87
NSE	0.94	0.88	0.67	0.65	0.98	0.98	0.93	0.72
KGE	0.93	0.91	0.72	0.70	0.98	0.98	0.82	0.7

Examining the values of evaluation criteria from two perspectives: the percentage changes of criteria in the training and testing sections and the assessment of criterion values in these two domains, provides precise information and analysis. Investigating the percentage changes in criteria reflects the stability and effective training of the model, while assessing the criterion values allows for a judgment regarding the model's performance adequacy. Therefore, based on the above table, the XGBoost model has demonstrated good performance in the exponential scale of the temperature parameter using the model MPI-ESM1-2-HR parameters in both training and testing sections. Additionally, the evaluation metric values in the training and testing sections showed only slight changes, indicating good training and stability of the XGBoost model. In the simulation of the relative humidity parameter, the model has also shown stable and good performance in both training and testing sections, according to the calculated evaluation criteria. A noteworthy point is the differing performance of the model in the exponential scale of the temperature and humidity parameters using ECMWF-ERA5 model parameters compared to the MPI-ESM1-2-HR model. According to the evaluation criteria, the XGBoost model has performed much better than the MPI-ESM1-2-HR model in the exponential downscale of the temperature parameter, and the evaluation metrics in both training and testing phases show slight changes, indicating good training and high stability of the model. However, in the exponential downscale of the humidity parameter, the model performed better in the training phase compared to the testing phase, and was accompanied by an increase in error, which can be attributed to the type of error metric used in model training as well as the tendency of the XGBoost model to simulate intermediate values. Since in the training and testing process, the data for evaluating the model's performance has been set in a non-sequential manner, graphs such as Swarm, Q-Q, Scatter, and Violin have been used over the entire simulation range for a more detailed analysis of the model's performance.

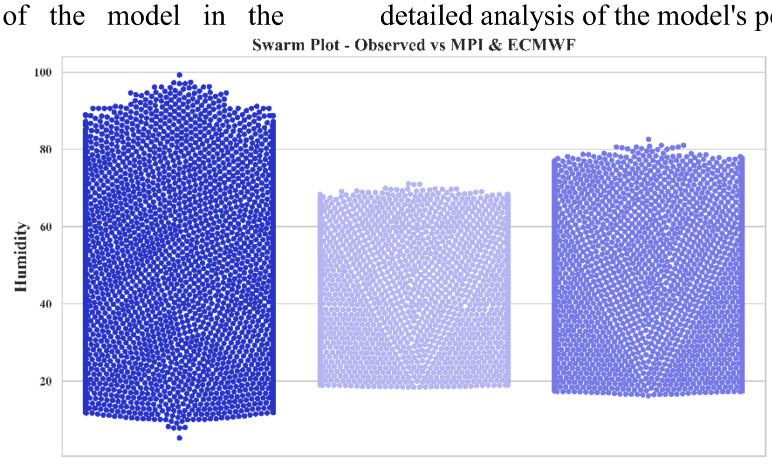


Figure 3. Swarm Plot of relative humidity variable (Observed vs simulated of MPI-ESM1-2-HR & ECMWF-ERA5)

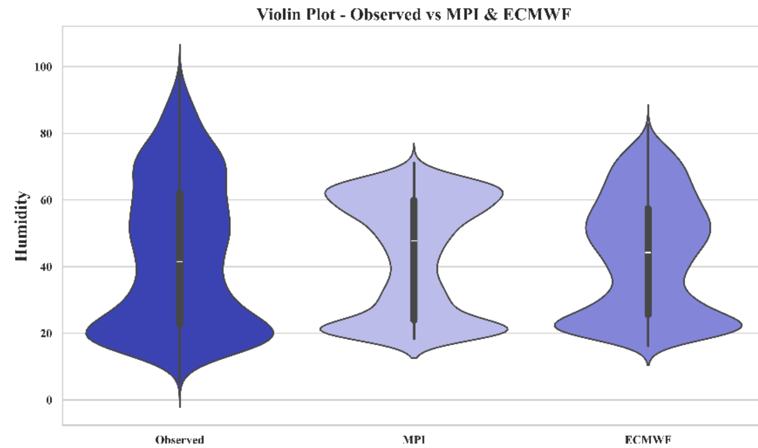
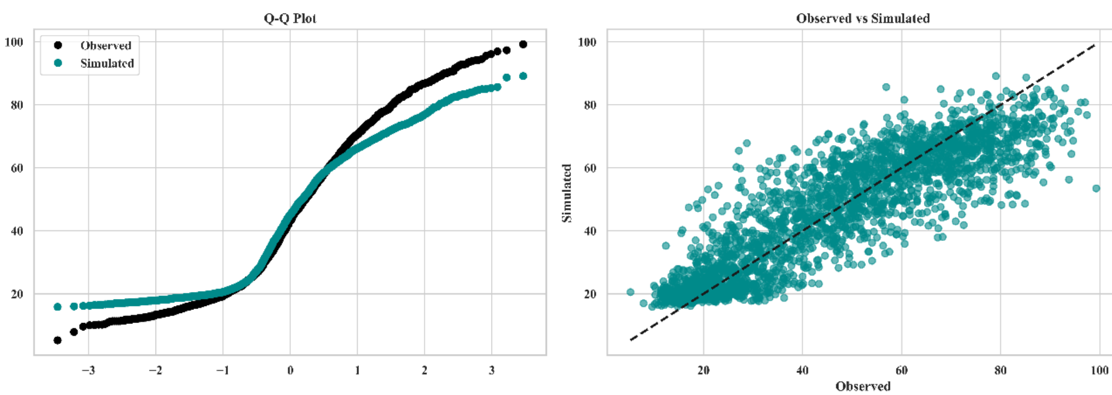
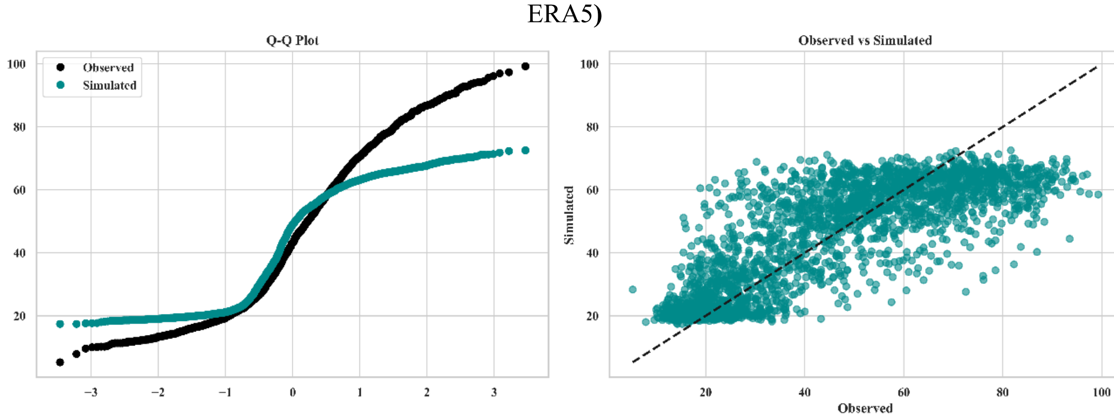


Figure 4. Violin Plot of relative humidity variable (Observed vs simulated of MPI-ESM1-2-HR & ECMWF-ERA5)



a b

Figure 5. QQ-Plot (a) and Scatter plot (b) of the relative humidity variable (Observed vs simulated of ECMWF-ERA5**)**



a b

Figure 6. QQ-Plot (a) and Scatter plot (b) of the relative humidity variable (Observed vs simulated of MPI-ESM1-2-HR)

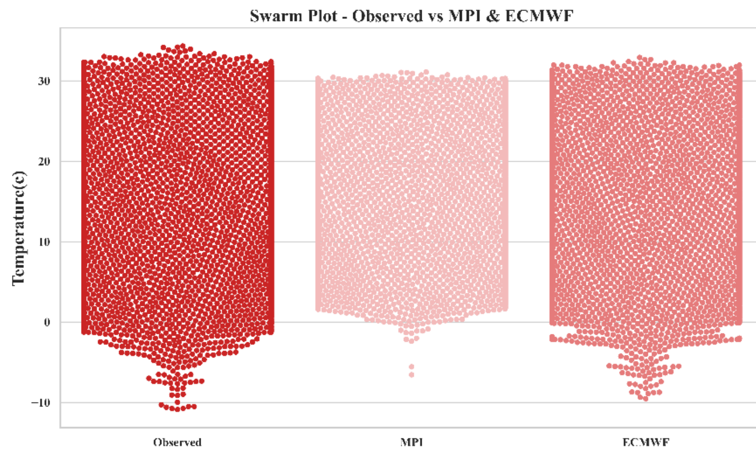


Figure 7. Swarm Plot of temperature variable (Observed vs simulated of MPI-ESM1-2-HR & ECMWF-ERA5)

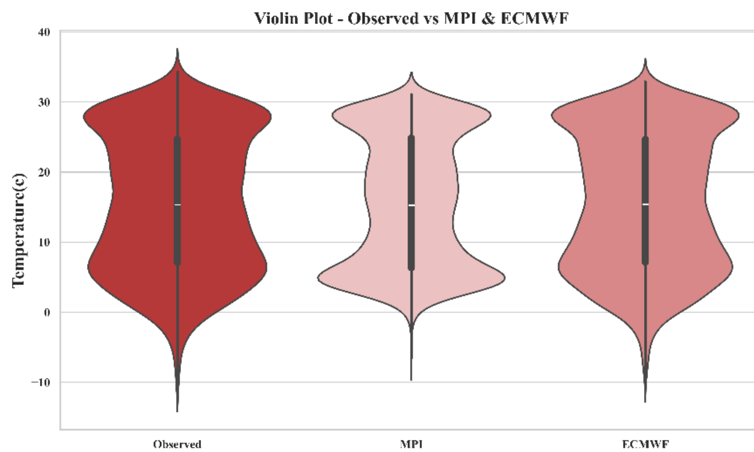
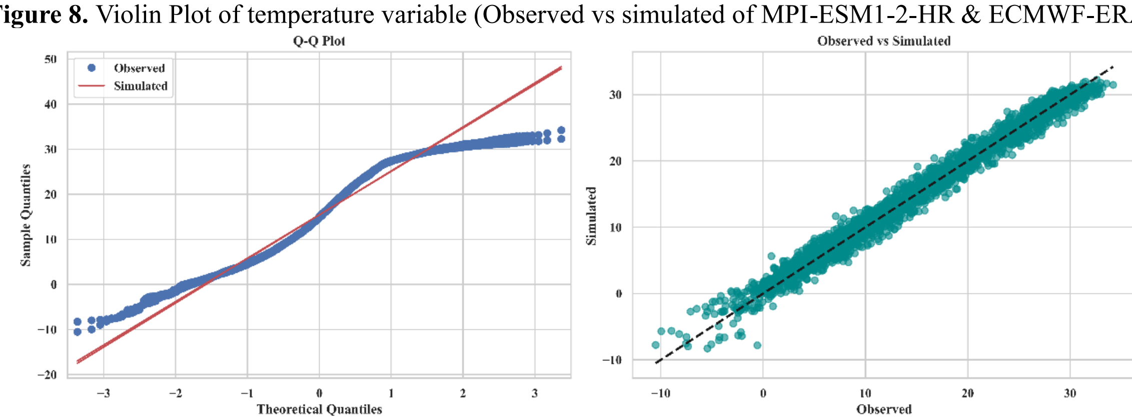
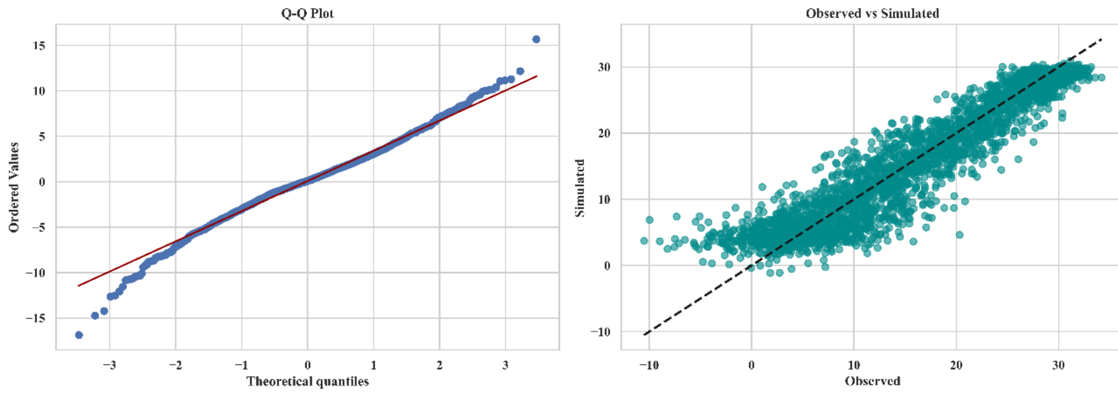


Figure 8. Violin Plot of temperature variable (Observed vs simulated of MPI-ESM1-2-HR & ECMWF-ERA5)



a b

Figure 9. QQ-Plot (a) and Scatter plot (b) temperature variable (Observed vs simulated of ECMWF-ERA5)



a b

Figure 10. QQ-Plot (a) and Scatter plot (b) of the relative humidity variable (Observed vs simulated of MPI-ESM1-2-HR)

3.1 Examination and analysis of the distribution of observed and simulated relative humidity and temperature data

Considering Figures 3 to 10, the Xgboost model, using the ECMWF-ERA5 model parameters, has performed better in the downscaling of humidity parameters compared to the MPI-ESM1-2-HR model. This is evident as the analysis of the SWARM and Scatter plots shows that the model tends to generate median values more effectively, although it struggles with producing relative humidity values above 80% and below 10%. However, an analysis of the violin plot indicates that the XGBoost model has similarly produced values with means comparable to observational data, simulated upper and lower threshold values, as well as showsa good performance in terms of the density of the simulated data using the ECMWF-ERA5 model parameters. The analysis of the XGBoost model's performance in the downscaling of temperature parameters, regarding statistical characteristics and the distribution of observational and simulated values, shows that the model has performed well in the downscaling of this climatic parameter using the ECMWF-ERA5 model, demonstrating excellent accuracy in producing both intermediate and threshold values.

The analysis of the violin chart shows that although the model using the parameters of both the MPI-ESM1-2-HR and ECMWF-ERA5 models performs similarly in simulating temperature values with an average comparable to the observed values, in terms of the density of the simulated values over the entire simulation range, especially in the ranges of 10- to 0 degrees and 20 to 35 degrees (extreme values), it performs significantly better using the parameters of the ECMWF-ERA5 model compared to the MPI-ESM1-2-HR model. The examination of data dispersion and the Q-Q chart also confirms that the XGBoost model has performed more successfully in the fine-scale temperature parameter, particularly with the use of parameters from the ECMWF-ERA5 model.

In this study, to improve the accuracy and reliability of the climate downscaling model based on the XGBOOST algorithm, a sensitivity analysis was performed based on the internal parameters of the machine learning model using the CRROS validation method, separately for each of the GCMs and each of the climate variables of temperature and relative humidity.

Figure 11 shows the results of the sensitivity analysis of the XGBoost model parameters separately for each of the GCMs and each of the climate variables of temperature and relative humidity of the GCMs.

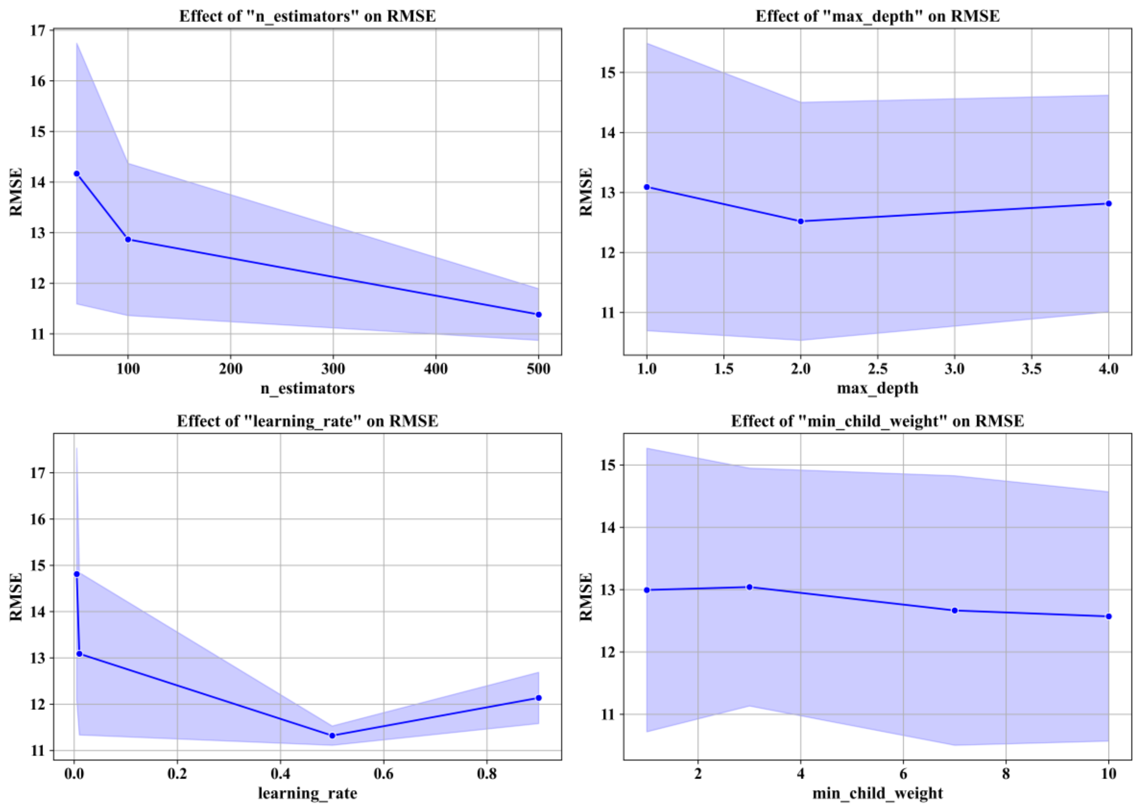


Figure 11. Results of sensitivity analysis of XGBoost model parameters separately for the ECMWF-ERA5 GCMs model and relative humidity climate variabless

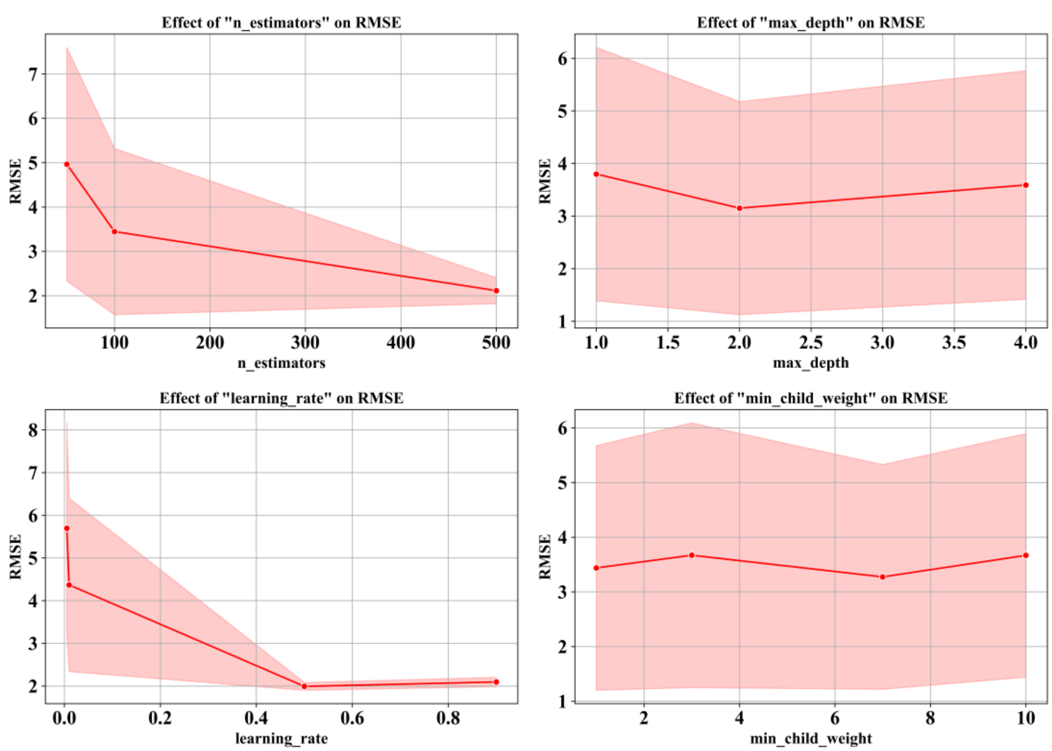


Figure 11. cont. Results of sensitivity analysis of XGBoost model parameters separately for the ECMWF-ERA5 GCM and temperature climate variables

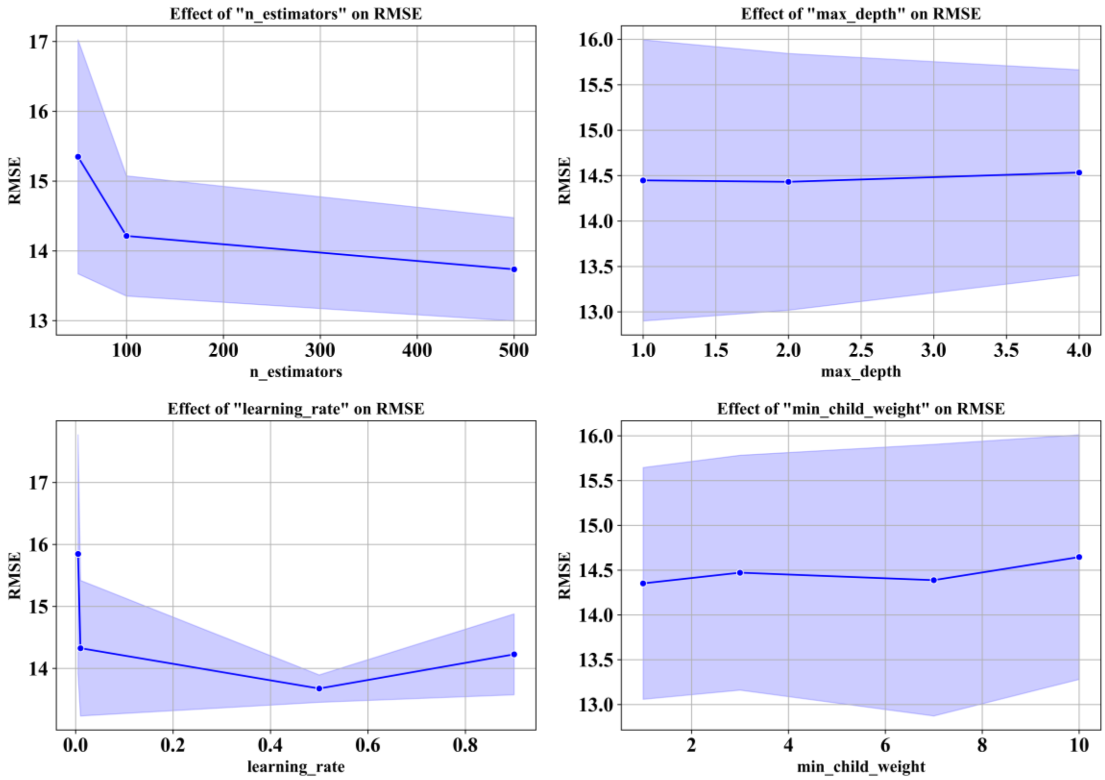


Figure 11. cont. Results of sensitivity analysis of XGBoost model parameters separately for the MPI-ESM1-2-HR GCM and relative humidity climate variables

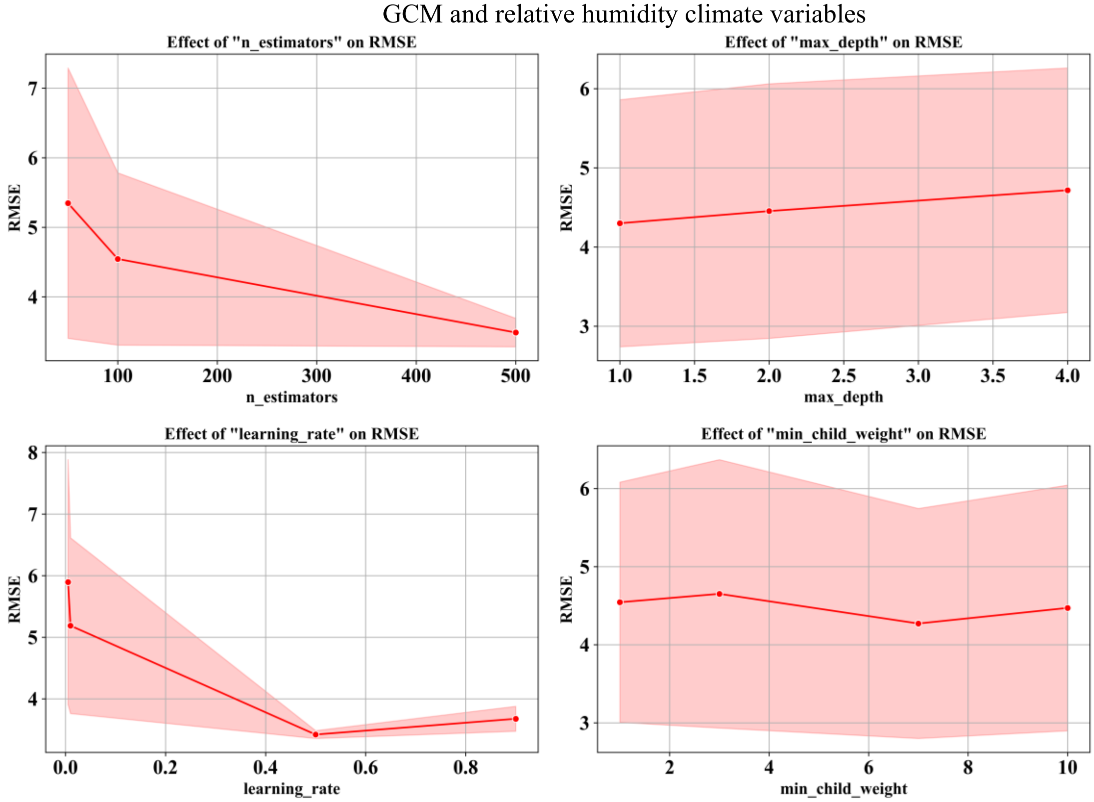


Figure 11. cont. Results of sensitivity analysis of XGBoost model parameters separately for the MPI-ESM1-2-HR GCM and temperature climate variables

Sensitivity analysis using the Cross-Validation method can, in addition to examining the sensitivity of each parameter, also show the optimal range of each parameter. So that the bandwidth indicates the changes in the error function value (RMSE) in each run, assuming that the value of that parameter remains constant. Therefore, a greater bandwidth indicates greater changes in the error function and, as a result, a greater sensitivity of the model to that parameter. On the other hand, the line in the graph shows the changes in the error function with respect to the change in the parameter value. Therefore, the lower the value of the error function, the more optimized the value of that parameter is. As a result, by simultaneously examining these two analyses, the optimal range of each parameter can be identified (the optimal range is located in a location with low bandwidth and minimum error function value). Therefore, examining the sensitivity analysis graphs shows that the XGboost model has shown the greatest sensitivity in the exponential downscaling of both climate variables, temperature and relative humidity, and in both GCMs to the parameters max_dept and min_Child because they have the greatest bandwidth compared to other parameters. Based on the analyses and reviews conducted, the results of this research indicated that the XGBoost model performs well at the sub-sampling level for temperature and humidity parameters. Another point of discussion is the varying performance of the XGBoost model at the sub-sampling level for temperature and humidity parameters. Since the XGBoost model, with the RMSE error function, tends to generate median values, and considering that the intrinsic nature of temperature values differs from humidity (in terms of the intensity of daily variations), the model showed better performance at the sub-sampling level for temperature parameters compared to humidity parameters. Additionally, another finding from this research highlights the significant impact of the type of input parameters to the XGBoost model (differences in GCM parameters) on the model's performance at the sub-sampling level for climatic parameters.The successful performance of machine learning models, including XGBoost, in downscaling climate parameters in various regions of the world has been reported. For example, in a study in North Africa (a region with predominantly dry and semi-arid climate), machine learning models, including XGBoost, were used to downscale hourly temperature of reanalysis products, and the results showed that these models significantly reduce RMSE error compared to classical methods (El Htiti et al., 2023). This finding is consistent with the results of the present study regarding the high performance of XGBoost for temperature. Additionally, in another study in China, the performance of the XGBoost model in estimating soil temperature in different climatic zones was evaluated, and the results demonstrated its superiority over Random Forest (RF) and M5P models (Zhao et al., 2022). This also confirms the high capability of XGBoost in modeling temperature-related variables.

In the field of downscaling modeling using XGBoost for multiple parameters, a study in China evaluated this model for correcting the bias of numerical weather predictions for daily precipitation and noted its satisfactory performance; the same research also mentioned the successful application of XGBoost for temperature (Dong et al., 2022). Although the present study does not directly address precipitation, the success of XGBoost in various climatic parameters indicates its broad potential.

One of the limitations of this study is the use of data from a single synoptic station. Although the Kermanshah station has been selected as a representative of humid regions, examining the model's performance in other stations with different topographical and climatic features in both dry and humid areas could lead to a more comprehensive understanding of the capabilities of the XGBoost model. The observed limitation in this study regarding the simulation of extreme values of temperature and high humidity has also been reported in some other machine learning studies. For example, it has been noted that machine learning models may struggle to generalize to environmental conditions that have not been adequately represented in the training data (Ma et al., 2021b, cited by Zhang et al., 2024). This could be one of the reasons for oversmoothing and the inability to fully reproduce the ranges of observational variability. Additionally, as mentioned in the results, the model's performance for relative humidity has room for improvement; using more predictor variables or combining XGBoost with other deep learning models could be explored in this regard. Future research could also investigate the impact of various climate change scenarios on the outputs downscaled by the XGBoost model. Comparing XGBoost with other novel deep learning algorithms, such as transformer-based models, could provide new insights.

4. Conclusions

The present study showed that the XGBoost algorithm, as one of the advanced machine learning methods, has a very good performance in downscaling of climate parameters, especially temperature, and to some extent relative humidity. To evaluate the performance of the model, the values of the metrics in the training and testing sections were examined separately. For temperature with MPI-ESM1-2-HR data, the values of NRMSE, R^2, NSE, and KGE in the training section were 5.04, 0.97, 0.94, and 0.93, respectively, and in the testing section were 7.46, 0.94, 0.88, and 0.91, respectively. For relative humidity with MPI-ESM1-2-HR data, the values of these metrics in the training section were 14.07, 0.82, 0.67, and 0.72, respectively, and in the testing section were 13.81, 0.81, 0.65, and 0.70, respectively. In addition, for temperature with ECMWF-ERA5 data, the values of the metrics in the training section were 2.28, 0.99, 0.98, and 0.98, respectively, and in the testing section were 2.81, 0.99, 0.98, and 0.98, respectively. Finally, for relative humidity with ECMWF-ERA5 data, the values of these metrics in the training section were 5.95, 0.97, 0.93, and 0.82, respectively, and in the testing section were 13.17, 0.87, 0.72, and 0.7, respectively. The results of the model evaluation using statistical metrics in both the training and testing sections indicate the favorable accuracy and stability of this algorithm. In particular, the superior performance of XGBoost in reproducing the mean and extreme values of temperature compared to relative humidity and the relative superiority of ECMWF-ERA5 data over the MPI-ESM1-2-HR model are noteworthy. Also, the results of the sensitivity analysis showed that the XGBoost model showed the highest sensitivity to the two parameters max-depth and min-child in the exponential downscaling of the two climate parameters, relative humidity and temperature, using data from both ECMWF-ERA5 and MPI-ESM1-2-HR models. Although the model performance in simulating extreme values of relative humidity has faced some limitations, the overall results confirm the use of XGBoost as an effective tool in improving the spatial and temporal accuracy of climate data in the arid regions of Iran. On the other hand, the sensitivity of the model performance to the type of input variables from GCMs shows the importance of careful selection of predictive parameters in the downscaling process. The main limitations of this study are: limited access to long-term and high-quality data from other synoptic stations, which hinders multi-station validation and generalizability of the results to more diverse regions. Despite these limitations, the research results provide practical suggestions for practical applications: As a powerful tool in the process of downscaling climate variables such as temperature and relative humidity, the XGBoost model can effectively model complex nonlinear relationships between large-scale GCM data and local observations, which leads to the production of local data with high resolution and higher accuracy. The main application of this downscaling in water resources issues is to improve hydrological models for river flow forecasting, dam management, and drought risk assessment. For example, with accurate downscaling of temperature and humidity, potential evapotranspiration can be estimated with higher accuracy, which plays a key role in water allocation planning in semi-arid watersheds of Iran, such as the Kermanshah basin, and helps reduce uncertainty in climate change scenarios. For future research, it is suggested to investigate the performance of the XGBoost model in other climates of Iran (such as coastal humid regions or central deserts) to assess its generalizability to diverse climatic conditions. And to increase the accuracy of the model in downscaling the relative humidity parameter, use ensemble methods such as stacking with other models (such as random forest or neural networks) to reduce the error of simulation of extreme values, apply advanced calibration such as Gaussian discriminant analysis (GDA) after downscaling, and integrate with dynamic downscaling techniques to improve nonlinear relationships. These approaches can help develop more powerful tools for climate change adaptation.

Author Contributions:
Mohammad Fouladi Nasrabad: Conceptualization, methodology, formal analysis and investigation, visualization, resources, writing-original draft preparation.
Mohsen Pourreza Bilondi: Conceptualization, supervision, formal analysis and investigation.
Mehdi Amirabadizadeh: Conceptualization, methodology, formal analysis and investigation.
Mahna Javaheri: formal analysis and investigation, manuscript editing**.**

Conflicts of interest
The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data availability statement:
The data of this research is available through correspondence with the corresponding author.

References

Almasi, A., Fatemi, S. E., & Eghbalzadeh, A. (2024). The prediction of monthly rainfall in Kermanshah Synoptic Station under the social-economic scenarios of the sixth climate change report. *Advanced Technologies in Water Efficiency*, *4*(1), 40-64. doi:10.22126/atwe.2024.10245.1097

Ali, S., Khan, S. D., Haq, M. U., Li, J., Virrantaus, K., & Chen, Y. (2023). Spatial downscaling of GRACE data based on XGBoost model for improved understanding of hydrological droughts in the Indus Basin Irrigation System (IBIS). *Remote Sensing*, 15(4), 873. doi:10.3390/rs15040873

Anandhi, A., Frei, A., Pierson, D. C., Schneiderman, E. M., Zion, M. S., Lounsbury, D., & Matonse, A. H. (2018). Examination of change factor methodologies for climate change impact assessment. *Water Resources Research, 54*(2), 1067-1086. doi:10.1029/2010WR009104

Sobhani, B., Eslahi, M., & Babaeian, I. (2017). Comparison of statistical downscaling in climate change models to simulate climate elements in Northwest Iran. *Physical Geography Research*, *49*(2), 301-325. doi:10.22059/jphgr.2017.62847

Sachindra, D. A., Huang, F., Barton, A., & Perera, B. J. C. (2018). Statistical downscaling of precipitation using machine learning techniques. Atmospheric Research, 212, 240–258. doi:10.1016/j.atmosres.2018.05.022

Chen, J., Brissette, F. P., Lucas-Picher, P., & Caya, D. (2020). Impacts of spatial resolution of global climate models on the statistical downscaling of precipitation. *Climate Dynamics, 55*(7-8), 1815-1837. doi:10.1007/s00382-020-05347-8

Chen, S., Wen, Z., Yang, P., Zhang, T., & Chen, J. (2022). Challenges and perspectives for high-resolution precipitation downscaling. *Earth and Space Science, 9*(11), e2022EA002453. doi:10.1029/2022EA002453

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). ACM. doi:10.1145/2939672.2939785

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). doi:10.1145/2939672.2939785

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. doi:10.1145/2939672.2939785

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). Applied multiple regression/correlation analysis for the behavioral sciences. Routledge. doi:10.4324/9780203774441

Daneshkhah, A., Ghorbani, M. A., Naganna, S. R., & Ghazvinian, P. H. (2020). Statistical downscaling of precipitation using machine learning techniques: A case study of Urmia Lake basin, Iran. *Theoretical and Applied Climatology, 140*(3-4), 1215-1231. doi:10.1007/s00704-020-03122-8

Diouf, I., Tramblay, Y., & Vischel, T. (2019). Predictor selection for statistical downscaling of rainfall in senegal using general circulation model outputs. *Theoretical and Applied Climatology, 137*(3-4), 2757-2771. doi:10.1007/s00704-019-02796-9

Dong, J., Zeng, W., Wu, L., Huang, J., Gaiser, T., & Srivastava, A. K. (2023). Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with XGBoost in different regions of China. Engineering Applications of Artificial Intelligence, 117, 105579. doi:10.1016/j.engappai.2022.105579

El Htiti, M., Ouagabi, A., Lazaar, M., Bouziane, M., Hayani, A., & Guenoun, J. (2023). Machine-Learning-Based Downscaling of Hourly ERA5-Land Air Temperature over Mountainous Regions. Atmosphere, 14(4), 610. MDPI. doi:10.3390/atmos14040610

Fouladi Nasrabad, M., Amirabadizadeh, M. and Dastourani, M. (2024). Performance Evaluation of Two General Circulation Models for Downscaling Average Temperature in Birjand County. *Integrated Watershed Management*, *4*(1), 30-45. doi:10.22034/iwm.2024.2013786.1109

Ghahreman, B., Daneshvar, M. R. M., Zare, H., & Ebrahimi, M. (2022). Evaluating the performance of machine learning algorithms for seasonal precipitation downscaling in Iran. *Water Resources Management, 36*(1), 157-177. doi:10.1007/s11269-021-03004-5

Giri, R. K., Swain, S., Pingale, S. M., & Meshram, C. (2021). Statistical downscaling and projection of future temperature and precipitation using SVM, relevance vector machine and gaussian process regression over Narmada River basin, India. *Stochastic Environmental Research and Risk Assessment, 35*(6), 1189-1213. doi:10.1007/s00477-020-01949-x

Dong, J., Zeng, W., Wu, L., Huang, J., Gaiser, T., & Srivastava, A. K. (2023). Enhancing short-term forecasting of daily precipitation using numerical weather prediction bias correcting with XGBoost in different regions of China. *Engineering Applications of Artificial Intelligence*, *117*, 105579. doi:10.1016/j.engappai.2022.105579

Georgiades, M., Boucher, O., Lamarque, J.-F., & Quaas, J. (2025). Global projections of heat stress at high temporal resolution using machine learning. Earth System Science Data, 17(3), 1153–1172. doi:10.5194/essd-17-1153-2025

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance metrics: Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2), 80-91. doi:10.1016/j.jhydrol.2009.08.003

Gutiérrez, J.M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S. and San Martín, D. (2019). An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. *International Journal of Climatology, 39*(9), 3750-3775. doi:10.1002/joc.5462

Hassanzadeh, E., Zhang, H., Murphy, J. M., & Matthews, A. J. (2021). Improving precipitation downscaling using deep learning: A study with focus on the Maritime Continent. *Journal of Geophysical Research: Atmospheres, 126*(23), e2021JD034869. doi:10.1029/2021JD034869

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. doi:10.1016/j.quaint.2017.06.020

Khosravi, A., Safavi, H. R., & Mirnezami, H. (2025). Developing an ensemble machine learning framework for enhanced climate projections using CMIP6 data in the Middle East. npj Climate and Atmospheric Science, 8(1), Article 103. doi:10.1038/s41612-025-01033-9

Li, X., Chen, Y., Duan, Z., & Liu, J. (2022). Improving satellite precipitation estimates using XGBoost algorithm: A comparative study across different climatic regions. Remote Sensing of Environment, 268, 112778. doi:10.3390/w14142150

Maraun, D., & Widmann, M. (2018). *Statistical downscaling and bias correction for climate research*. Cambridge University Press. doi:10.1017/9781107588783

Maraun, D., Huth, R., Gutiérrez, J. M., Martín, D. S., Dubrovsky, M., Fischer, A., Hertig, E., Soares, P. M. M., Bartholy, J., Pongracz, R., Widmann, M., Casado, M. J., Ramos, P., & Bedia, J. (2019). The VALUE perfect predictor experiment: evaluation of temporal variability. *International Journal of Climatology, 39*(9), 3786-3818. doi:10.1002/joc.5222

Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water, 10*(11), 1536. doi:10.3390/w10111536

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M. and Fast, I. (2019). Developments in the MPI-ESM1-2-HR-M Earth System Model version 1.2 (MPI-ESM1-2-HR-ESM1.2) and its response to increasing CO2. Journal of Advances in Modeling Earth Systems, 11(4), 998–1038. doi:10.1029/2018MS001400

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F. et al (2018) A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). Journal of Advances in Modeling Earth Systems, 10 (7). pp. 1383-1413. ISSN 1942-2466. doi:10.1029/2017ms001217

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. Journal of Hydrology, 10(3), 282-290. doi:10.1016/0022-1694(70)90255-6

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Gonzalez, S., & Laudon, H. (2024). Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). *Environmental Modelling & Software*, 174, 105971. doi:10.1016/j.envsoft.2024.105971

Nourani, V., Razzaghzadeh, Z., Baghanam, A. H., & Molajou, A. (2019). ANN-based statistical downscaling of climatic parameters using decision tree predictor screening method. *Theoretical and Applied Climatology, 137*(3-4), 2111-2126. doi:10.1007/s00704-018-2722-5

Okkan, U., & Kirdemir, U. (2018). Statistical downscaling of monthly precipitation using linear regression, conditional metric-based models and spline interpolation. *International Journal of Climatology, 38*(5), 2421-2439.doi:10.1002/joc.5344

Parsa, M., Dehghani, M., Rezaei, M., & Klove, B. (2023). Downscaling precipitation using machine learning algorithms over Lake Urmia Basin, Iran. *Water, 15*(7), 1383. doi:10.3390/w15071383

Prasad, R. K., Ahmad, S., Ali, S., Adhikary, S. K., & Mohanty, U. C. (2018). Statistical downscaling of temperature using machine learning techniques over the Himalayan region. *Theoretical and Applied Climatology, 131*(3-4), 1175-1190. doi:10.1007/s00704-016-2004-z

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*(7743), 195-204. doi:10.1038/s41586-019-0912-1

Sachindra, D. A., Huang, F., Barton, A., & Perera, B. J. C. (2014). Statistical downscaling of general circulation model outputs to precipitation—part 1: calibration and validation. *International Journal of Climatology, 34*(13), 3264-3281. doi:10.1002/joc.3915

Sachindra, D. A., Ng, A. W. M., Muthukumaran, S., & Perera, B. J. C. (2018). Impact of predictor selection on performance of statistical downscaling models. *International Journal of Climatology, 38*(2), 923-945. doi:10.1002/joc.5224

Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., Kargar, K., Mosavi, A., Nabipour, N., & Chau, K. W. (2020). Predicting standardized precipitation index using machine learning models. *Engineering Applications of Computational Fluid Mechanics, 14*(1), 1192-1206. doi:10.1080/19942060.2020.1800921

Wilby, R. L., & Wigley, T. M. L. (1997). Downscaling general circulation model output: A review of methods and limitations. *Progress in Physical Geography, 21*(4), 530-548. doi:10.1177/030913339702100403

Zhu, X., Li, W., Chen, J., Ma, Y., Liu, H., & Zhu, D. (2025). Exploring machine learning approaches for precipitation downscaling. Geo-spatial Information Science. Advance online publication. doi:10.1080/10095020.2025.2477547

Zare, H., Ghahreman, B., Daneshvar, M. R. M., & Ebrahimi, M. (2021). Performance assessment of support vector machine and artificial neural network models in statistical downscaling of daily precipitation (case study: Urmia Lake basin, Iran). *Water Supply, 21*(5), 2139-2155. doi:10.2166/ws.2021.008

Zhang, M., Wang, K., Liu, Y., & Li, Y. (2022). Monthly streamflow forecasting based on XGBoost using decomposition-integration strategy. *Water Resources Management, 36*(10), 3659-3676. doi:10.1007/s11269-022-03203-8

Zhang, X., Tang, G., Wang, X., Song, Z., & Hong, Y. (2024). Downscaling satellite-derived soil moisture in the Three North region using ensemble machine learning and multiple-source knowledge. Hydrology and Earth System Sciences Discussions, hess-2024-129. Copernicus Publications. doi:10.5194/hess-2024-129

Zhang, Y., Wu, Z., Liu, K., Lan, T., Chen, J., & Li, Z. (2023). Enhancing spatial resolution of GNSS-R soil moisture retrieval through XGBoost algorithm-based downscaling approach: A case study in the Southern United States. *Remote Sensing*, 15(18), 4576. doi:10.3390/rs15184576

Zhao, L., Feng, T., Li, X., Chen, S., & Zhang, J. (2022). Modelling Soil Temperature by Tree-Based Machine Learning Methods in Different Climatic Regions of China. Applied Sciences, 12(10), 5088. MDPI. doi: 10.3390/app12105088