Journal of Hydroinformatics

© 2024 The Authors

Journal of Hydroinformatics Vol 00 No 0, 1 doi: 10.2166/hydro.2024.227

Check for updates

Intercomparison of machine learning methods for statistical downscaling of daily temperature under CMIP6 scenarios: a case study from Iran

Mohammad Reza Goodarzi (¹^{a,b,*}, Zaynab Hashemipour^{b,c}, Amir Saremi^{b,c} and Majid Niazkar^{d,e}

^a Department of Civil Engineering, Faulty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

^b Department of Civil Engineering, Yazd University, Yazd, Iran

^c Water Resources Management Engineering, Yazd University, Yazd, Iran

^d Euro-Mediterranean Center on Climate Change (CMCC), Porta dell'Innovazione Building, 2nd Floor Via della Libertà 12, Marghera, Venice, Italy

^e Ca' Foscari University of Venice, Department of Environmental Sciences, Informatics, and Statistics, Venice, Italy

*Corresponding author. E-mail: goodarzimr@yazd.ac.ir

(D) MRG, 0000-0002-6296-2315

ABSTRACT

General Circulation Models (GCMs) represent a contemporary and advanced tool designed to simulate the response of climate systems to alterations in greenhouse gas levels. Increasing spatial resolutions of the outputs of GCMs on a regional scale requires a downscaling process. This study applied six Machine Learning (ML) models, named decision tree regression (DTR), support vector regression (SVR), artificial neural networks (ANN), K-nearest neighbors (KNN), Light Gradient-Boosting Machine (LightGBM), and Stochastic Gradient Descent Regressor (SGDRegressor), to downscale daily temperature data from CMIP6 models in Kohgiluyeh and Boyer-Ahmad, Iran. Observations from Nazmakan station were used for training (1995 -2009) and testing (2009 -2015). In addition, future temperature projections during 2015 -2045 were made under SSP2-4.5 and SSP5-8.5 scenarios. Results showed that LightGBM and KNN developed the most reliable results. Mann-Kendall's analysis confirmed a significant upward trend, predicting cooler summers and warmer winters. The predicted data was also validated against observations from the period 2015 -2022. This study highlights the strengths and limitations of nonlinear ML techniques and emphasizes the need for further research to enhance predictive accuracy and spatial resolution in statistical downscaling.

Key words: CMIP6, downscaling, general circulation model, machine learning, temperature

HIGHLIGHTS

- This study employed six machine learning models for downscaling daily temperature of climate models under Coupled Model Intercomparison Project 6 scenarios in southwest Iran.
- Daily temperature for the future period of 2015–2022 was predicted under two Shared Socioeconomic Pathways (SSP2-4.5 and SSP5-8.5) climate scenarios.
- Daily temperature range of future period will decrease compared to the historical period, leading to cooler summers and warmer winters.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).



1. INTRODUCTION

Climate change denotes a sustained alteration in long-term climate patterns. This transformation is a direct outcome of the release of greenhouse gases into the atmosphere, a phenomenon rooted in the pre-industrial revolution era (Zahabiyoun *et al.* 2013). Climate change has evolved into a worldwide challenge with the ability to influence the local scale. Changing climate patterns increase the likelihood of more frequent occurrences of extreme hydrological events, like flash floods or droughts (IPCC 2014). Diverse sectors, such as agriculture, ecosystems, human and food security, as well as natural resource management, face potential threats from these effects, leading to substantial losses. Climate change jeopardizes the sustainable development of nations. Hence, it is crucial to initiate prompt and targeted climate actions aligned with the objectives of sustainable development.

Taking action involves using adaptation as a key strategy. This approach centers on enhancing life adaptability and preparation to mitigate the severity of impacts from future changes (Prathom & Champrasert 2023). To reduce local impacts, an assessment of risks and impacts is conducted by comparing potential future climate scenarios with the present situation. This informs the development of adaptation plans for different sectors (Buras & Menzel 2019; Liu *et al.* 2019; Miao *et al.* 2020; Toot *et al.* 2020). Predicting future climate conditions involves using general circulation models (GCMs) under various scenarios (Goodarzi *et al.* 2022a, b). These predictions account for different scenarios, such as variations in greenhouse gas emissions, land use, and socioeconomic conditions, providing valuable insights for planning and decision-making. This study specifically concentrates on a future scenario known as 'Shared Socioeconomic Pathways' (SSPs), currently employed as an input scenario for the GCM. SSPs are scenarios that simulate how various socioeconomic developments can influence or pose challenges to strategies addressing climate change (Riahi *et al.* 2017). SSPs include five socioeconomic scenarios labeled SSP1–5. The two SSPs selected in this study are regarded to be a combination of moderate social vulnerability with a moderate emission range (SSP2-4.5) and higher emissions that impose high mitigation but low adaptation challenges (SSP5-8.5) (Zhu & Yang 2021; Ma *et al.* 2022). However, due to the coarse spatial resolution, GCM outputs might inadequately depict local data. To enhance spatial resolution, the downscaling process is crucial. GCMs are limited by their coarse spatial resolution (250–600 km), leading to discrepancies between their predictions and actual local and global climate variations. The output might not accurately represent information for a small local area. This poses a challenge for adaptation activities (Gebrechorkos *et al.* 2019). Therefore, to utilize the prediction output, a necessary step involves a process called 'downscaling' to enhance spatial resolution and calibrate the prediction output. Downscaling is a method employed to enhance the resolution of data by capturing the historical relationship between observed data and prediction output, with the goal of calibrating the prediction.

Downscaling methods include statistical and dynamical approaches. The former, which is enhanced through machine learning (ML) techniques, is particularly effective for simulating hydroclimatic variables that exhibit complex nonlinear relationships (Xu et al. 2020). In this context, some ML methods have been applied for statistical downscaling, such as genetic programming (GP) (Coulibaly 2004), artificial neural network (ANN) (Ahmed et al. 2015; Tripathi et al. 2006), multiple linear regression (MLR) (Sachindra et al. 2014), relevant vector machine (RVM) (Okkan & Inan 2015), K-nearest neighbors (KNN) (Liu et al. 2017), support vector machine (SVM) (Goly et al. 2014), gene expression programming (Hashmi et al. 2011), and generalized linear models (GLMs) (Beecham et al. 2014). Furthermore, SVM, ANN, and MLR have been used to simulate minimum and maximum monthly temperatures (Duhan & Pandey 2015). Another study highlighted the applicability of RVM, ANN, GP, and SVM in developing downscaling techniques for drought analysis (Sachindra et al. 2018). The better performance of linear regression compared to ANN with limited available data was recognized (Hatanaka 2022). Also, multi-gene GP was found to provide more valid results than ANN for downscaling daily temperature (Niazkar et al. 2023). In addition, a combination of interpolation approaches, such as inverse distance weighted (IDW) with ML models, e.g., IDW-ANN, was used for downscaling precipitation and temperature (Prathom & Champrasert 2023), while a combination of four ML algorithms (Random Forest, KNN, Extra Trees, and Gradient-Boosting Decision Tree) was employed to predict future precipitation (Wang et al. 2023). Nonetheless, several research gaps still remain. To be more specific, challenges in the current methods include improving spatial resolution, capturing extreme events, and better quantifying uncertainties. In addition, downscaling efforts remained underrepresented for hydroclimatic variables, e.g., extreme precipitation or small-scale phenomena, in certain regions. Moreover, many models lack thorough validation against real-world data, raising concerns about their reliability. Finally, issues related to scalability and generalization across different regions and future scenarios require the development of more robust and adaptable models.

The aim of this study is to employ some ML models to downscale outputs of Coupled Model Intercomparison Project 6 (CMIP6) climate models, i.e., Access-CM2, MPI-ESM1-2-HR, MPI-ESM1-2-LR, and MRI-ESM2-0. In this context, this study assesses the applicability of ANN, support vector regression (SVR), decision tree regression (DTR), KNN, Light Gradient-Boosting Machine (LightGBM), and Stochastic Gradient Descent Regressor (SGDRegressor) for downscaling hydroclimatic variables. Based on the literature review, the last three ML models have been rarely applied for statistical downscaling. In addition, the daily temperature for the future period of 2015–2045 is predicted under SSP2-4.5 and SSP5-8.5 climate scenarios.

2. MATERIALS AND METHODS

2.1. Study area

Kowsar dam basin is located in Kohgiluyeh and Boyer-Ahmad Province, Iran, and has an area of 2,420 km² in the range of latitudes 30° 26′ to 30° 55′ north and longitudes 50° 26′ to 51° 13′ east (Figure 1). This region is primarily characterized by extensive agricultural lands and pastures, making it essential to address water management issues in this province. The average temperatures of the region in winter and summer are 11.5 and 32.4°C, respectively. The average annual rainfall is 421 mm from 1995 to 2015.

In this study, observed data from a synoptic station and simulated data from GCM outputs adopted from CMIP6 were used. The observed data of the daily temperature collected from Nazmakan synoptic station (with latitude 30° 63' north and longitude 50° 74' east) from 1995 to 2015 was obtained by the Meteorological Organization (https://www.irimo.ir), which provides long-term and reliable data about the state of precipitation and temperature in Iran. This database was divided into two parts: (a) training data (from 1995 to 2009) and (b) test data (from 2009 to 2015). The dataset, which is illustrated in Figure 2, was used to evaluate the applicability of four CMIP6 climate models.

Journal of Hydroinformatics Vol 00 No 0, 4



Figure 1 | Location of the study area.



Figure 2 | Temperature observations for the training and testing datasets.

2.2. Models and scenarios

The daily gridded temperature data obtained from four CMIP6 climate models, shown in Table 1, are for 1995–2015 for the historical period and for 2015–2045 for the future period. The data were collected from the Earth System Grid Association (https://esgf-node.llnl.gov/search/cmip6). Furthermore, SSP is a combination of representative concentration pathways and

No	Model	Institution	Country	Resolution
1	Access-CM2	CSIRO-BOM	Australia	$1.25 imes 1.88^{\circ}$
2	MRI-ESM2-0	MRI	Japan	$1.13\times1.13^{\circ}$
3	MPI-ESM1-2-LR	MPI-M	Germany	$0.94\times0.94^{\circ}$
4	MPI-ESM1-2-HR	MPI-M	Germany	$0.94\times0.94^{\circ}$

Table 1 | Characteristics of the CMIP6 models considered in this study

alternative pathways of socioeconomic development (O'Neill *et al.* 2016). The two SSPs selected in this study are SSP2-4.5 and SSP5-8.5.

2.3. ML downscaling models

SVR, ANN, DTR, KNN, LightGBM, and SGDRegressor are applied in this study to effectively downscale GCM outputs, each leveraging different strengths in capturing patterns and relationships within the data.

2.3.1. Support vector regression

An SVM is an ML algorithm that handles both linear and nonlinear regressions. SVR is widely recognized as one of the most commonly used supervised ML techniques (He *et al.* 2022). Basically, it creates an optimal hyperplane to transform input data into a higher dimensional space, where data points become vectors (Bisong 2019). The optimum hyperplane, which divides data into two classes, aims to widen the area between the hyperplane and nearest data points, i.e., support vectors. Finally, the SVM prediction at each point is the hyperplane value at that point plus a bias to avoid overfitting. The performance of an SVM model is significantly influenced by kernel hyperparameters and a regularization term (Leong *et al.* 2021). After comparing linear, polynomial, sigmoid, and radial basis kernel functions, the radial basis function kernel was ultimately chosen for feature research.

2.3.2. Artificial neural network

The ANN, a method proficient in constructing nonlinear models between input and output samples (Piraei *et al.* 2023), derives inspiration from the functioning of the human brain's biological nervous system (Agatonovic-Kustrin & Beresford 2000). It is a mathematically modeled system mimicking biology, consisting of processing elements called neurons (or perceptrons) connected with parameters (weights) assigned to the connections, forming the neuronal structure. In comparison to traditional statistical regression techniques like MLR and GLM, ANN demonstrates superior performance, finding wide-spread application in hydrology and climatology (Kişi 2008; Okkan & Fistikoglu 2014).

In general, a common network in ANN comprises three layers: (i) an input layer including a set of neurons to carry input data, (ii) a hidden layer that enables data flow in the network, and (iii) an output layer holding neuron associated with the output data. Since the input data should be independent, neurons within a layer are not connected to one another, while they are exclusively in connection with the neurons in adjacent layers (Piraei *et al.* 2023). The connection between two neurons relates their values with a mathematical model, whose parameters can be optimized when data flow back and forth in the network. As a result, the layer-based structure provides a relationship between input and output variables regardless of any prior knowledge on the theoretical background of the problem in question.

2.3.3. Decision tree regression

The decision tree algorithm, a commonly used ML method for classification and regression tasks, exhibits a tendency to overfit, leading to less-than-optimal performance on testing datasets (Bisong 2019). With a structure resembling a flowchart composed of nodes and branches, the decision tree is widely embraced in data mining due to its simplicity and ease of understanding. In creating a treelike structure, DTR divides data based on feature values at branch nodes. Each data partition corresponds to the result of a splitting test on the training data. Branch nodes link to leaf nodes, where each leaf signifies a model outcome derived by averaging data points from the training data assigned to that node during the splitting process (Piraei *et al.* 2023).

2.3.4. K-nearest neighbors

The KNN technique, used for both classification and regression purposes, operates as a nonparametric method (Altman 1992). Its core function involves identifying the K closest data points to a designated testing data point and calculating the weighted average of their target values. Assessing similarity between the training and testing data points relies on a distance function. Common distance functions for continuous variables include Euclidean, Manhattan, and Minkowski (Nugrahaeni & Mutijarsa 2016).

Using a desirable distance function, distance metrics can be determined and then sorted to obtain the minimum distance that is nearest, which implies the maximum similarity among neighbor points. In addition, the KNN output can be achieved by the K-nearest data points based on sorting the calculated distances. For this purpose, selecting the K value is crucial on KNN performances (Piraei *et al.* 2023). In essence, a high value for K may incorporate outclass data, whereas a low value may end up in poor training of KNN. Therefore, the optimum number of neighbors (K) should be computed using a cross-validation approach.

2.3.5. Light Gradient-Boosting Machine

LightGBM is a scalable gradient-boosting ML model developed by Microsoft. In comparison to traditional ML models, it exploits a histogram-based technique for enhancing the training process. As a result, it reduces computational efforts and the amount of memory, which enables it to handle large datasets efficiently (Xu *et al.* 2023). Furthermore, LightGBM can tackle not only classification but also regression tasks.

LightGBM employs a specific strategy for creating decision trees, called leaf-wise. Commonly, the decision tree model adopts the level-wise strategy that works on the leaves of the same layer. Conversely, the leaf-wise strategy first searches all leaves to determine the ones with the highest branching gain and then proceed with the branching cycle. Thus, it can not only obtain higher accuracy by conducting more important splits but also decrease the overfitting issue (Fan *et al.* 2019). Moreover, LightGBM can employ various optimization algorithms, like Gradient-based One-Side Sampling and Exclusive Feature Bundling, to improve its precision and run time. To be more specific, these optimization techniques make it a powerful tool for developing robust ML models across various domains.

2.3.6. Stochastic Gradient Descent Regressor

SGDRegressor utilizes stochastic gradient descent (SGD) as its search engine to develop a linear ML model for regression tasks. To be more precise, SGD iteratively conducts parameter estimation for each training example, which helps it to adequately handle large-scale datasets efficiently (Kumar *et al.* 2023). SGDRegressor treats one sample at a time, making it efficient. Thus, the speed of SGDRegressor is one of its key features. Moreover, SGDRegressor incorporates various loss functions to providing flexibility in handling different types of regression problems. In addition, it has several regularization techniques to prevent overfitting (Erdal & Karakurt 2013). Finally, these features make it a versatile ML for a wide range of applications.

2.4. Mann-Kendall test

The Mann-Kendall test, which is a nonparametric test, was first presented by Mann and then the statistical distribution of the test was extracted by Kendall. This test is widely used to find trends in meteorological time series data (Modarres & da Silva 2007; Goodarzi *et al.* 2022a, b). Due to the capabilities of this test in revealing the changes in the time series of climatic variables, it has been widely used in climate change studies. This test examines the null hypothesis of no trend in time series. In this regard, it also identifies sudden change points.

In essence, it consists of two hypotheses: (i) null hypothesis and (ii) alternative hypothesis. The former denotes the lack of pattern or trend in the data series, while the latter suggests the presence of a trend in the data series. For more information on this trend analysis method, interested readers are referred to previous studies (Hırca *et al.* 2022; Niazkar *et al.* 2023).

The implementation of analysis in this test is attributed to the diagrams *UI* and *U'I*. Both graphs originate from a specific point. Moreover, when |U| is greater than 1.96, it indicates that the *UI* graph intersects the two lines Y = 1.96 and Y = -1.96, signifying a significant trend. If *U* is greater than 0, or the overall trend of the UI graph is upward, it discloses a meaningful upward trend. In the case where *U* is less than 0, the trend is both significant and decreasing. Furthermore, if the *UI* and *U'I* graphs intersect within the range of -1.96 to 1.96, it indicates a sudden change.

2.5. Sensitivity analysis

Sensitivity analysis (SA) examines how an independent variable affects the temperature estimation outcomes in statistical downscaling methods based on ML. By employing Equation (1), the SA percentage for the simulated temperature, as determined by downscaling models, was calculated, considering each individual climate model (Zakwan & Niazkar 2021)

$$SA_{i} = \frac{T_{\max}(x_{i}) - T_{\min}(x_{i})}{\sum_{i=1}^{N} [T_{\max}(x_{i}) - T_{\min}(x_{i})]} \times 100$$
(1)

where $T_{\max}(x_i)$ and $T_{\min}(x_i)$ are the maximum and minimum temperature, respectively.

2.6. Performance evaluation

Various metrics were employed to assess the performance of ML-based downscaling models: root mean square error (RMSE), Nash–Sutcliffe efficiency (NSE), the coefficient of determination (R^2), and bias percentage (PBIAS). The metrics are displayed in Table 2, where T_{obs} and T_{est} represent the daily observed and estimated temperatures, respectively, and N denotes the number of data points.

Referring to the optimal values listed in Table 2 for each metric, a greater R^2 value indicates a stronger correlation between the simulation and observed data. Unlike R^2 , a higher value of RMSE signifies a larger deviation between simulated and observed data. A higher NSE indicates more accurate model performance. Finally, the ideal PBIAS value is 0. Negative and positive PBIAS values signify underestimation and overestimation, respectively.

3. RESULTS AND DISCUSSION

This research involves the application of six ML-based downscaling models to forecast the daily temperature for a future period at Nazmakan station under the SSP2-4.5 and SSP5-8.5 climate scenarios. To assess the sparsity between observed and estimated daily temperature values obtained from downscaling models, a comparison was conducted for the training data (1995–2009), as illustrated in Figure 3, and the testing data (2009–2015), as depicted in Figure 4.

The results indicate that LightGBM and KNN have little dispersion for training data and they are close to the observed data line, while DTR displays greater dispersion among all. According to the results of testing data, the lowest dispersion is related

Measure	Equation	Range	Optimal value
RMSE	$\sqrt{rac{\sum\limits_{i=1}^{N}{({T_{{ m obs}_i}} - {T_{{ m est}_i}})^2}}{N}}$	[0, ∞]	0
NSE	$1 - rac{{\sum\limits_{i = 1}^{N} {{{({T_{{{\rm{obs}}}_i}} - {T_{{\rm{est}}_i}}})^2 } }}{{N^{- 2}}}$	[-∞, 1]	1
R ²	$\sum_{i=1}^{N} (T_{\text{obs}_i} - \sum_{i=1}^{N} T_{\text{obs}_i}/N)$ $\int \sum_{i=1}^{N} \left[\left(T_{\text{obs}_i} - \frac{1}{N} \sum_{i=1}^{N} T_{\text{obs}_i} \right) \left(T_{\text{est}_i} - \frac{1}{N} \sum_{i=1}^{N} T_{\text{est}_i} \right) \right] $	[0, 1]	1
PRIAS	$\left(\sqrt{\sum_{i=1}^{N} \left[\left(T_{obs_i} - 1/N \sum_{i=1}^{N} T_{obs_i} \right)^2 \left(T_{est_i} - 1/N \sum_{i=1}^{N} T_{est_i} \right)^2 \right]} \right)$	[0
	$rac{\sum\limits_{i=1}^{n} (T_{\mathrm{obs}_i} - T_{\mathrm{est}_i})}{\sum\limits_{i=1}^{N} (T_{\mathrm{obs}_i})} imes 100$	[,]	

Table 2	Model	evaluation	metrics
	innouci	CVGIGGGOT	111001100

Journal of Hydroinformatics Vol 00 No 0, 8



Figure 3 | Temperature estimations applying ML-based statistical downscaling for the training data: (a) ANN, (b) KNN, (c) DTR, (d) SVR, (e) LightGBM, and (f) SGDRegressor.



Figure 4 | Temperature estimations applying ML-based statistical downscaling for the testing data: (a) ANN, (b) KNN, (c) DTR, (d) SVR, (e) LightGBM, and (f) SGDRegressor.

to LightGBM and KNN and the highest dispersion is related to DTR. Clearly, scattering of data reduces the accuracy of the results.

The temperature estimated by all ML models under both SSP climate scenarios is shown in Figure 5. Compared to the observed temperature shown in Figure 2, all ML models demonstrated a significant increase in the minimum daily

temperature during the coldest days and a decrease in the maximum daily temperature during the hottest days in the study area. LightGBM, KNN, and DTR display changes of about 5–10°C, while SGDRegressor, ANN, and SVR show smaller changes compared to the observed data. The SSP5-8.5 scenario is associated with an increase in the mean temperature compared to the SSP2-4.5 scenario. In general, the daily temperature in the future period versus historical period will lead to warmer winters and cooler summers in the coming years.

3.1. Performance evaluation of downscaling models

The performance of downscaling models was evaluated for training and testing data. At first, R^2 , RMSE, and standard deviation (SD) of six ML-based downscaling methods were calculated and Taylor diagrams were plotted, as shown in Figure 6. The ML performances for predicting temperature differed slightly. According to Figure 6(a), all ML models resulted in very



Figure 5 | Temperature estimations applying ML-based statistical downscaling for 2015–2045: (a) ANN for SSP2-4.5, (b) ANN for SSP5-8.5, (c) DTR for SSP2-4.5, (d) DTR for SSP5-8.5, (e) KNN for SSP2-4.5, (f) KNN for SSP5-8.5, (g) SVR for SSP2-4.5, (h) SVR for SSP5-8.5, (i) LightGBM for SSP2-4.5, (j) LightGBM for SSP5-8.5, (k) SGDRegressor for SSP2-4.5, and (l) SGDRegressor for SSP5-8.5. (*continued*.).

Journal of Hydroinformatics Vol 00 No 0, 11



Figure 5 | Continued.

close R^2 for the training data, and SD of the simulated temperature values were close to the observed ones. To be more specific, the best performance was achieved by LightGBM and DTR with the highest R^2 of about 0.93 and a minor RMSE value. According to Figure 6(b), LightGBM and KNN obtained best performance with an R^2 of about 0.92 for the testing data, and then SGDRegressor, ANN, SVR, and DTR models performed better, respectively. Finally, these results indicate that the efficiency of ML models is different in the training and testing data based on the evaluation criteria considered in this study.

Figure 7 describes the values of different metrics for better performance evaluation of downscaling models. As shown in Figure 7(a), LightGBM and DTR resulted in NSE = 0.93 and showed a better performance compared to the other six models for the training data. Therefore, KNN presented the best performance with NSE = 0.92 for the testing data. Likewise, Figure 7(b) demonstrates LightGBM and DTR as the best models with PBIAS around 0 for the training data and ANN with the lowest PBIAS value for the testing data, respectively. In general, the obtained results showed that LightGBM and KNN achieved the highest accuracy for Nazmakan station.

Journal of Hydroinformatics Vol 00 No 0, 12



Figure 6 | Taylor diagrams of estimated temperature for (a) training and (b) testing data.



Figure 7 | Evaluation of ML-based statistical downscaling performance in temperature estimation for both training and testing datasets using the following metrics: (a) NSE and (b) PBIAS.

3.2. Mann-Kendall test results

To identify the mutation and the initial year of the trend or sudden change, the MK test was executed. Initially, UI and U'I components were graphed on an annual basis for the temperature data throughout the historical and future periods under both climate scenarios. The respective graphs are illustrated in Figure 8.

According to Figure 8(a), the annual changes for 1996–2001 and 2007–2011 showed a negative and decreasing trend, while a positive and increasing trend was observed for 2002–2007 and 2011–2014, with sudden changes occurring in 1996, 2006, and 2008. Based on Figure 8(b), the predicted annual average temperature changes under the SSP2-4.5 scenario indicate a positive and increasing trend, with sudden changes in 2016, 2019, 2021, 2022, and 2025. Similarly, Figure 8(c), associated with the SSP5-8.5 scenario, shows a positive and upward trend from 2015 to 2020 and from 2025 to 2045. Based on Kendall's test, the intersection of the UI and U'I graphs between -1.96 and 1.96 indicates sudden shifts during the period of 2015–2019. These statistical changes suggest that the trends during this period are distinct from others, requiring further investigation into the factors driving these fluctuations.



Figure 8 | Changes of UI and U'I components for annual mean temperature of (a) observed data, (b) SSP2-4.5, and (c) SSP5-8.5 scenarios for 2015–2045.

3.3. SA results

The influence of GCM output variations on daily temperature data generated by ML-based downscaling models was investigated. As shown in Figure 9, the MPI-ESM1-2-HR model had the greatest impact on the SVR, KNN, and DTR results, contributing 34.7, 41.5, and 89.4%, respectively. Conversely, the ACCESS-CM2 outputs exhibited the lowest impacts on

Figure 9 | Results of the SA.

Table 3 | The evaluation between the observed data and the estimated data under both climate scenarios for 2015–2022

ML technique	SSP2-4.5		SSP5-8.5	
	R ²	NSE	R ²	NSE
ANN	0.9150	0.9136	0.9109	0.9102
DTR	0.9028	0.8985	0.8976	0.8945
KNN	0.9193	0.9157	0.9146	0.9117
SVR	0.9145	0.9118	0.9108	0.9089
LightGBM	0.9152	0.8985	0.9135	0.9006
SGDRegressor	0.9152	0.8872	0.9111	0.8867

the KNN and SVR models, with sensitivity analyses of 12.1 and 17.2%, respectively. Also, the SA results for all four GCM models concerning ANN were identical, which limits their reliability. These findings highlight that spatial resolution differences significantly affect model performances, as higher-resolution models capture local climatic features more effectively than their lower-resolution counterparts. Overall, this underscores the critical role that GCMs play in ensuring the accuracy of simulation results.

3.4. Comparison of predicted data with observed data

In this research, a comparison was made between the observed data and the estimated data under both climate scenarios during 2015–2022 for Nazmakan station. The evaluation was calculated based on R^2 and NSE, which is shown in Table 3. KNN obtained the most valid performance with R^2 and NSE of about 0.92 in both scenarios and in other models, between 0.88 and 0.91. In general, it can be concluded that ML-based methods were acceptably precise for down-scaling daily temperature.

4. CONCLUSIONS

In this study, six ML methods, namely, KNN, SVR, DTR, ANNs, LightGBM, and SGDRegressor, were employed to downscale the daily temperature of the Nazmakan station for the period from 1995 to 2015. Training and testing of ML-based

Journal of Hydroinformatics Vol 00 No 0, 15

downscaling models were conducted based on the observed daily temperatures, allowing the outputs of four CMIP6 climate models to be effectively downscaled. The results, evaluated using Taylor diagrams, NSE, and PBIAS, demonstrated that all models performed similarly well for both training and testing data. Based on the study's metrics, all six ML models performed acceptably, with LightGBM achieving the best results in the training dataset, while both KNN and LightGBM performed the best in the testing dataset. The SA indicated that the models were more sensitive to the MPI-ESM1-2-HR climate model. Furthermore, the six ML models were utilized to predict daily temperatures for the future period of 2015–2045 under the SPP2-4.5 and SPP5-8.5 climate scenarios. The results of the Mann–Kendall test illustrated a decreasing trend in the early observed temperatures, followed by a positive and increasing trend. The estimated temperatures also showed an upward trend in most forthcoming years. According to the future projections, there will be a reduction in the daily temperature range compared to the historical period, leading to cooler summers and warmer winters. The evaluation of the performance of ML-based models was supported by the overlapping of available observed and estimated data from 2015 to 2022, which yielded reliable results. Overall, this study provides significant contributions to the field by demonstrating the effectiveness of ML methods in down-scaling climate data. Specifically, the KNN and LightGBM methods exhibited the best performance in simulating daily temperatures and showed strong capabilities in performing nonlinear regression.

To build upon the findings of this study, future research should explore advanced ML techniques, such as ensemble learning and deep learning, to improve downscaling accuracy. Comprehensive validation against observational data is essential for ensuring reliability. Focusing on underrepresented regions and climate variables will provide an enhanced understanding of climate trends. In addition, models must efficiently handle larger datasets. Ensuring that these models can generalize across diverse climatic conditions will strengthen climate predictions and risk assessments.

AUTHOR CONTRIBUTIONS

MG, ZH, AS, and MN conceptualized the study and performed the validation. ZH, AS, and MN contributed to the methodology and formal analysis. AS and MN contributed to software analysis. ZH performed the investigation. ZH and AS wrote the original draft of the article. MG and MN contributed to reviewing and editing the writing of the article, supervision, and project administration.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Agatonovic-Kustrin, S. & Beresford, R. (2000) Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, *Journal of Pharmaceutical and Biomedical Analysis*, **22** (5), 717–727. https://doi.org/10.1016/S0731-7085(99) 00272-1.
- Ahmed, K., Shahid, S., Haroon, S. B. & Xiao-Jun, W. (2015) Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan, Journal of Earth System Science, 124, 1325–1341. https://doi.org/10.1007/s12040-015-0602-9.
- Altman, N. S. (1992) An introduction to kernel and nearest-neighbor nonparametric regression, *The American Statistician*, 46 (3), 175–185. https://doi.org/10.2307/2685209.
- Beecham, S., Rashid, M. & Chowdhury, R. K. (2014) Statistical downscaling of multi-site daily rainfall in a South Australian catchment using a generalized linear model, *International Journal of Climatology*, 34 (14), 3654–3670. https://doi.org/10.1002/joc.3933.
- Bisong, E. (2019) Building Machine Learning and Deep Learning Models on Google Cloud Platform. Berkeley, CA: Apress.
- Buras, A. & Menzel, A. (2019) Projecting tree species composition changes of European forests for 2061–2090 under RCP 4.5 and RCP 8.5 scenarios, *Frontiers in Plant Science*, **9**, 1986. https://doi.org/10.3389/fpls.2018.01986.
- Coulibaly, P. (2004) Downscaling daily extreme temperatures with genetic programming, *Geophysical Research Letters*, **31** (16), L16203. https://doi.org/10.1029/2004GL020075.
- Duhan, D. & Pandey, A. (2015) Statistical downscaling of temperature using three techniques in the Tons River basin in Central India, *Theoretical and Applied Climatology*, **121**, 605–622. https://doi.org/10.1007/s00704-014-1253-5.
- Erdal, H. I. & Karakurt, O. (2013) Advancing monthly streamflow prediction accuracy of CART models using ensemble learning paradigms, *Journal of Hydrology*, **477**, 119–128. https://doi.org/10.1016/j.jhydrol.2012.11.015.

- Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X. & Zeng, W. (2019) Light gradient boosting machine: An efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data, *Agricultural Water Management*, 225, 105758. https://doi.org/10.1016/j.agwat.2019.105758.
- Gebrechorkos, S. H., Hülsmann, S. & Bernhofer, C. (2019) Statistically downscaled climate dataset for East Africa, *Scientific Data*, 6 (1), 31. https://doi.org/10.1038/s41597-019-0038-1.
- Goly, A., Teegavarapu, R. S. & Mondal, A. (2014) Development and evaluation of statistical downscaling models for monthly precipitation, *Earth Interactions*, **18** (18), 1–28. https://doi.org/10.1175/EI-D-14-0024.1.
- Goodarzi, M. R., Abedi, M. J. & Pour, M. H. (2022a) Climate change and trend analysis of precipitation and temperature: A case study of Gilan, Iran. In: Zakwan, M., Wahid, A., Niazkar, M. & Chatterjee, U. (Eds.) *Current Directions in Water Scarcity Research*, Vol. 7. Amsterdam: Elsevier, pp. 561–587.
- Goodarzi, M. R., Mohtar, R. H., Piryaei, R., Fatehifar, A. & Niazkar, M. (2022b) Urban WEF nexus: An approach for the use of internal resources under climate change, *Hydrology*, **9** (10), 176. https://doi.org/10.3390/hydrology9100176.
- Hashmi, M. Z., Shamseldin, A. Y. & Melville, B. W. (2011) Statistical downscaling of watershed precipitation using gene expression programming (GEP), Environmental Modelling & Software, 26 (12), 1639–1646. https://doi.org/10.1016/j.envsoft.2011.07.007.
- Hatanaka, Y. M. (2022) Machine Learning Based Statistical Downscaling for Rainfall on Hawaiian Islands, Honolulu, HI: University of Hawai'i at Manoa.
- He, B., Jia, B., Zhao, Y., Wang, X., Wei, M. & Dietzel, R. (2022) Estimate soil moisture of maize by combining support vector machine and chaotic whale optimization algorithm, *Agricultural Water Management*, **267**, 107618. https://doi.org/10.1016/j.agwat.2022.107618.
- Hırca, T., Eryılmaz Türkkan, G. & Niazkar, M. (2022) Applications of innovative polygonal trend analyses to precipitation series of Eastern Black Sea Basin, Turkey, *Theoretical and Applied Climatology*, 147 (1–2), 651–667. https://doi.org/10.1007/s00704-021-03837-0.
- Intergovernmental Panel on Climate Change. Climate Change. (2014) Synthesis Report. Geneva, Switzerland: IPCC.
- Kişi, Ö. (2008) Stream flow forecasting using neuro-wavelet technique, *Hydrological Processes: An International Journal*, **22** (20), 4142–4152. https://doi.org/10.1002/hyp.7014.
- Kumar, V., Kedam, N., Sharma, K. V., Mehta, D. J. & Caloiero, T. (2023) Advanced machine learning techniques to improve hydrological prediction: A comparative analysis of streamflow prediction models, *Water*, **15** (14), 2572. https://doi.org/10.3390/w15142572.
- Leong, W. C., Bahadori, A., Zhang, J. & Ahmad, Z. (2021) Prediction of water quality index (WQI) using support vector machine (SVM) and least square-support vector machine (LS-SVM), *International Journal of River Basin Management*, **19** (2), 149–156. https://doi.org/10. 1080/15715124.2019.1628030.
- Liu, Y., Yang, Y., Jing, W. & Yue, X. (2017) Comparison of different machine learning approaches for monthly satellite-based soil moisture downscaling over northeast China, *Remote Sensing*, **10** (1), 31. https://doi.org/10.3390/rs10010031.
- Liu, B., Gao, X., Ma, J., Jiao, Z., Xiao, J., Hayat, M. A. & Wang, H. (2019) Modeling the present and future distribution of arbovirus vectors *Aedes aegypti* and *Aedes albopictus* under climate change scenarios in Mainland China, *Science of the Total Environment*, 664, 203–214. https://doi.org/10.1016/j.scitotenv.2019.01.301.
- Ma, Z., Sun, P., Zhang, Q., Zou, Y., Lv, Y., Li, H. & Chen, D. (2022) The characteristics and evaluation of future droughts across China through the CMIP6 multi-model ensemble, *Remote Sensing*, 14 (5), 1097. https://doi.org/10.3390/rs14051097.
- Miao, L., Li, S., Zhang, F., Chen, T., Shan, Y. & Zhang, Y. (2020) Future drought in the dry lands of Asia under the 1.5 and 2.0 C warming scenarios, *Earth's Future*, 8 (6), e2019EF001337. https://doi.org/10.1029/2019EF001337.
- Modarres, R. & da Silva, V. d. P. R. (2007) Rainfall trends in arid and semi-arid regions of Iran, *Journal of Arid Environments*, **70** (2), 344–355. https://doi.org/10.1016/j.jaridenv.2006.12.024.
- Niazkar, M., Goodarzi, M. R., Fatehifar, A. & Abedi, M. J. (2023) Machine learning-based downscaling: Application of multi-gene genetic programming for downscaling daily temperature at Dogonbadan, Iran, under CMIP6 scenarios, *Theoretical and Applied Climatology*, 151 (1–2), 153–168. https://doi.org/10.1007/s00704-022-04274-3.
- Nugrahaeni, R. A. & Mutijarsa, K. (2016) 'Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification', 2016 International Seminar on Application for Technology of Information and Communication (ISemantic).
- Okkan, U. & Fistikoglu, O. (2014) Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M, *Theoretical and Applied Climatology*, **117**, 343–361. https://doi.org/10.1007/s00704-013-1005-y.
- Okkan, U. & Inan, G. (2015) Bayesian learning and relevance vector machines approach for downscaling of monthly precipitation, *Journal of Hydrologic Engineering*, **20** (4), 04014051. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001024.
- O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F. & Lowe, J. (2016) The scenario model intercomparison project (ScenarioMIP) for CMIP6, *Geoscientific Model Development*, 9 (9), 3461–3482. https:// doi.org/10.5194/gmd-9-3461-2016.
- Piraei, R., Niazkar, M., Afzali, S. H. & Menapace, A. (2023) Application of machine learning models to bridge afflux estimation, *Water*, 15 (12), 2187. https://doi.org/10.3390/w15122187.
- Prathom, C. & Champrasert, P. (2023) General circulation model downscaling using interpolation Machine learning model combination Case study: Thailand, Sustainability, 15 (12), 9668. https://doi.org/10.3390/su15129668.
- Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R. & Fricko, O. (2017) The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview, *Global Environmental Change*, 42, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009.

Journal of Hydroinformatics Vol 00 No 0, 17

- Sachindra, D., Huang, F., Barton, A. & Perera, B. (2014) Multi-model ensemble approach for statistically downscaling general circulation model outputs to precipitation, *Quarterly Journal of the Royal Meteorological Society*, **140** (681), 1161–1178. https://doi.org/10.1002/qj. 2205.
- Sachindra, D., Ahmed, K., Shahid, S. & Perera, B. (2018) Cautionary note on the use of genetic programming in statistical downscaling, *International Journal of Climatology*, **38** (8), 3449–3465. https://doi.org/10.1002/joc.5508.
- Toot, R., Frelich, L. E., Butler, E. E. & Reich, P. B. (2020) Climate-biome envelope shifts create enormous challenges and novel opportunities for conservation, *Forests*, **11** (9), 1015. https://doi.org/10.3390/f11091015.
- Tripathi, S., Srinivas, V. & Nanjundiah, R. S. (2006) Downscaling of precipitation for climate change scenarios: A support vector machine approach, *Journal of Hydrology*, **330** (3–4), 621–640. https://doi.org/10.1016/j.jhydrol.2006.04.030.
- Wang, D., Liu, J., Luan, Q., Shao, W., Fu, X., Wang, H. & Gu, Y. (2023) Projection of future precipitation change using CMIP6 multimodel ensemble based on fusion of multiple machine learning algorithms: A case in Hanjiang River Basin, China, *Meteorological Applications*, **30** (5), e2144. https://doi.org/10.1002/met.2144.
- Xu, R., Chen, N., Chen, Y. & Chen, Z. (2020) Downscaling and projection of multi-CMIP5 precipitation using machine learning methods in the Upper Han River basin, Advances in Meteorology, 2020, 1–17. https://doi.org/10.1155/2020/8680436.
- Xu, K., Han, Z., Xu, H. & Bin, L. (2023) Rapid prediction model for urban floods based on a light gradient boosting machine approach and hydrological-hydraulic model, *International Journal of Disaster Risk Science*, 14, 79–97. https://doi.org/10.1007/s13753-023-00465-2.
- Zahabiyoun, B., Goodarzi, M. R., Bavani, A. M. & Azamathulla, H. M. (2013) Assessment of climate change impact on the Gharesou River Basin using SWAT hydrological model, *CLEAN - Soil, Air, Water*, **41** (6), 601-609. https://doi.org/10.1002/clen.201100652.
- Zakwan, M. & Niazkar, M. (2021) A comparative analysis of data-driven empirical and artificial intelligence models for estimating infiltration rates, *Complexity*, **2021**, 1–13. https://doi.org/10.1155/2021/9945218.
- Zhu, Y. & Yang, S. (2021) Interdecadal and interannual evolution characteristics of the global surface precipitation anomaly shown by CMIP5 and CMIP6 models, *International Journal of Climatology*, **41**, E1100–E1118. https://doi.org/10.1002/joc.6756.

First received 20 July 2024; accepted in revised form 8 November 2024. Available online 18 November 2024