Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2025, 31 (2) : 447 – 469 DOI: 10.15832/ankutbd.1509731

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

Evaporation and precipitation prediction for future time frames via combined machine learning-climate change models: Quri Gol Wetland Case

Mohammad Reza Abdollahpour Azad[a], Mohammad Reza Jalali[ba], Mohammad Taghi Sattari[b,c*], Reza Mastouri[a]

[a]*Department of Civil Engineering, Arak Branch, Islamic Azad University, Arak, IRAN*

[b]*Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, IRAN*

[c]*Department of Agricultural Engineering, Faculty of Agriculture, Ankara University, Ankara, 06110, TURKEY*

ARTICLE INFO

Research Article
Corresponding Author: Mohammad Taghi Sattari, E-mail: mtsattar@gmail.com
Received: 03 July 2024 / Revised: 26 November 2024 / Accepted: 05 December 2024 / Online: 25 March 2025

Cite this article

Azad M R A, Jalali M R, Sattari M T, Mastouri R (2025). Evaporation and precipitation prediction for future time frames via combined machine learning-climate change models: Quri Gol Wetland Case . *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(2):447-469. DOI: 10.15832/ankutbd.1509731

ABSTRACT

Evaporation is a critical component in the management of water resources. Due to the complex interactions between various meteorological variables involved in evaporation calculations, numerous nonlinear models have been developed. The applicability and performance of these models vary depending on the specific climatic conditions of each region. This study evaluates the impacts of climate change on evaporation and precipitation patterns in the Quri Gol Wetland, located in East Azerbaijan, Iran, using machine learning models and climate change projections. Evaporation values for the present period (1991-2020) were estimated using six machine learning models: Random Forest (RF), Gradient Boosted Tree (GBT), Generalized Linear Model (GLM), Support Vector Machine (SVM), Gaussian Process Regression (GPR), and deep learning (DL). Future projections (2021-2050, 2051-2080, 2081-2100) were based on the LARS-WG and SDSM models under three climate scenarios (RCP 2.6, RCP 4.5, and RCP 8.5). The performance of the machine learning models was assessed using statistical metrics including R^2, Scatter Index (SI), Mean Absolute Error (MAE), Willmott's Index (WI), and Kling-Gupta Efficiency (KGE). The RF and DL models provided the most accurate predictions, with RF achieving an R^2 of 0.821 and an MAE of 0.902, while DL reached an R^2 of 0.822 and an MAE of 0.915 in the validation phase. Results from climate change projections indicated a significant increase in evaporation over the next century, with cumulative evaporation rising by up to 50.01% under the RCP 8.5 scenario by 2081-2100. In contrast, the projected increase in precipitation was much smaller, reaching a maximum of 16% in the same period. This imbalance between evaporation and precipitation highlights the potential for increasing water stress in the Quri Gol Wetland. The findings emphasize the need for adaptive water management strategies to mitigate the effects of increased evaporation and maintain ecological stability in the region.

Keywords: Climate change, Meteorological variables, Evaporation, Machine learning, Statistical indices

1. Introduction

Climate change is a critical challenge facing our planet in the 21st century, necessitating effective evaluation and prediction of its impacts, particularly on water resources. The adverse effects of climate change on water availability not only threaten ecosystems but also have significant environmental, economic, and social consequences. Each year, vast quantities of freshwater are stored in reservoirs; however, evaporation contributes to water quality degradation by increasing salinity levels. To address these challenges, understanding and accurately estimating surface water evaporation is crucial for sustainable water management in a changing climate.

Various methods have been developed for estimating surface water evaporation, typically classified into four main categories: (1) Pan evaporation measurement, (2) water balance, (3) energy balance, and (4) mass transfer (Gianniou & Antonopoulos 2007). Among these, the pan evaporation method is often preferred due to its simplicity, cost-effectiveness, and relatively high accuracy compared to more complex energy balance methods (Hamel 2009; Anderson 2004). This preference underscores the importance of reliable evaporation estimation techniques, particularly in light of the escalating impacts of climate change on hydrological cycles.

As noted, evaporation prediction is critical due to the increasing challenges posed by climate change and water resource depletion. Early research, such as Terzi (2010), employed genetic programming to model daily evaporation from Lake Egirdir, Turkey, showing its potential as a replacement for traditional methods. Kim et al. (2012) applied generalized regression neural networks (GRNN), support vector machines (SVM), and multilayer perceptron (MLP) to predict pan evaporation in arid and temperate zones, with artificial neural network (ANN) models outperforming empirical and multiple linear regression (MLR)

methods. Similarly, Guven & Kisi (2013) compared the Stephens-Stewart (SS) model, ANN, adaptive neuro-fuzzy inference system (ANFIS), fuzzy genetic (FG), and linear genetic programming (LGP) for estimating pan evaporation. They found LGP to be the most accurate model using meteorological data from Turkey. Goyal et al. (2014) extended the use of machine learning by comparing ANFIS, fuzzy logic, ANN, and least squares support vector regression (LS-SVR) for daily pan evaporation in India's subtropical climate, with LS-SVR and fuzzy logic providing more accurate predictions. Kisi (2015) employed M5 Model Trees (M5Tree), multivariate adaptive regression splines (MARS), and LSSVM at Turkish stations, with MARS proving superior. Similarly, Keshtegar et al. (2016) modeled pan evaporation using meteorological inputs in Iran, with their proposed models outperforming M5Tree and ANFIS. Wang et al. (2017a) applied six heuristic computing techniques across various climates in China, finding MLP superior to other models in most cases. Lu et al. (2018) also demonstrated the effectiveness of gradient boosted decision trees (GBDT) for daily pan evaporation in China's Poyang Lake Basin, with GBDT outperforming RF and M5Tree models.

Further advancements in machine learning and hybrid techniques are evident in Ghorbani et al. (2018), who integrated MLP with the firefly algorithm (MLP-FFA) for daily pan evaporation, showing it outperformed single models. Zounemat-Kermani et al. (2019) explored optimization algorithms, such as genetic algorithm (GA), particle swarm optimization (PSO), and others based on ANFIS, finding GA and PSO most effective. Shiri (2019) assessed ANN and ANFIS for pan evaporation in the U.S., recommending supplementary information for areas lacking local climatological data. Kisi & Heddam (2019) compared M5Tree, MARS, and MLP, with MARS proving to be the most accurate. Majhi et al. (2020) demonstrated that deep neural networks (DNN) provided superior accuracy in estimating daily pan evaporation in India, while Ghaemi et al. (2019) coupled M5Tree and MARS with the maximum overlap discrete wavelet transform (MODWT) to improve prediction performance in Turkey. Wang et al. (2020) introduced a hybrid technique combining kernel-based nonlinear Arps decline (KNEA) and the salp swarm algorithm (SSA) for evaporation prediction in China's arid regions, showing SSA-KNEA's high potential compared to MARS and M5Tree. Seifi & Soroush (2020) employed whale optimization algorithm (WOA), grey wolf optimization (GWO), and GA in combination with ANNs to predict daily pan evaporation across Iranian stations, finding hybrid models superior to single methods in most locations. Shaker Sureh et al. (2024) used data-driven models and atmospheric circulation indices to improve meteorological drought prediction in the Lake Urmia basin, finding that the M5 Tree model provided the most accurate results.

Climate change significantly influences hydrological processes, as evidenced by various studies utilizing climate models to evaluate its effects. Zarghami et al. (2011) investigated the impact of climate change on runoff in East Azerbaijan province using General Circulation Models (GCM) alongside statistical data from six meteorological stations. By applying the LARS-WG method and simulating flow through artificial neural networks, they demonstrated a substantial projected reduction in flow across three basins in the region. Similarly, Helfer et al. (2012) examined temperature and evaporation trends in Australia, revealing that rising temperatures during spring and summer correspondingly increase annual average air temperature and evaporation rates. Goyal & Ojha (2014) utilized GCMs to assess the effects of climate change on temperature extremes and evaporation in the Rajasthan River Basin, India. Their analysis of historical data from 1948 to 2000 and projections for 2001 to 2100 indicated that the modified M5 decision tree model exhibited the highest accuracy in predicting both minimum and maximum temperature.

In the context of Lake Urmia, Azizi et al. (2017) evaluated the role of climate change in reducing water levels, highlighting the combined effects of increased agricultural water demands, which contributed to a 46% reduction in total inflow, alongside a 16% impact from climatic variables. Ahmadaali et al. (2018) further explored the consequences of climate change on sustainable agricultural development in the Lake Urmia region by employing the WEAP21 model to assess five water management scenarios. Their findings indicated that the scenario integrating improved crop patterns and enhanced irrigation efficiency yielded the most favorable environmental indicators for agricultural sustainability. These studies underscore the critical need for robust climate models to inform water resource management strategies amid the ongoing challenges posed by climate change.

In light of the ongoing global climate crisis, this study aims to address the significant challenge of accurately predicting evaporation rates in sensitive ecological zones, specifically the Quri Gol Wetland in Iran. The primary goal is to assess the impacts of climate change on evaporation by employing GCM and various diffusion scenarios (A2, B1, A1B). Furthermore, the research investigates the efficiency of statistical downscaling methods (SDSM and LARS-WG) for microscale evaporation simulation, a crucial step in understanding localized climate impacts. In addition, this study develops advanced data mining methods such as Gaussian process regression and support vector machine, comparing their performance against other robust machine learning models like random forest, gradient boosted tree, generalized linear model, and deep learning methods. The evaluation of these models through multiple performance metrics represents a comprehensive effort to improve the accuracy and reliability of evaporation forecasts.

The novelty of this study stems from the integration of advanced machine learning techniques with GCM projections to predict future evaporation under varying climate scenarios. Unlike previous studies that predominantly relied on empirical models or singular machine learning approaches, this research leverages a hybrid methodology that integrates GCM output with the best-performing machine learning model, offering a novel approach for regional climate adaptation strategies. This dual-method comparison contributes to filling a gap in the existing literature where the combination of statistical downscaling and cutting-edge data mining techniques has not been thoroughly explored. By focusing on the Quri Gol Wetland, the study not only provides valuable insights for local environmental and water management but also offers a model that can be adapted to other

regions globally. This research advances the science of hydrology by offering a scalable framework that can enhance climate resilience planning, especially in regions vulnerable to evaporation changes driven by climate variability.

2. Methods and Material

2.1 Study area

Quri Gol Wetland is located in the East Azerbaijan province of Iran, approximately 40 kilometers southeast of the provincial capital, Tabriz. Situated at an altitude of 1890 meters above sea level, the wetland covers an area of about 200 hectares. Geographically, it lies between 37°55′N latitude and 46°42′E longitude. This wetland is surrounded by the Sahand Mountain range, which significantly influences its climate and hydrology.

Quri Gol Wetland is located in a semi-arid to arid climate zone, characterized by cold winters and warm, dry summers. The region experiences significant temperature variations due to its altitude and geographic location. Annual precipitation is relatively low, averaging around 300-400 mm, with the bulk of rainfall occurring in the spring and autumn months. Winter precipitation is primarily in the form of snow, contributing to the wetland's water supply as it melts in the warmer months. The wetland is primarily fed by seasonal rainfall, snowmelt, and a few small streams originating from the surrounding mountains. Evaporation rates are high during the summer, significantly influencing the water balance of the wetland, especially under the pressures of climate change.

Quri Gol is a freshwater, endorheic (closed basin) wetland, meaning it does not drain into any sea or ocean. Its hydrology is heavily dependent on the balance between precipitation, runoff, and evaporation. The wetland's water level fluctuates seasonally, with higher levels in spring and early summer due to snowmelt and precipitation, followed by a reduction during the hot, dry summer months due to increased evaporation. One of the key environmental challenges facing Quri Gol Wetland is water loss through evaporation. Due to its relatively shallow depth and open surface area, the wetland is highly susceptible to evaporative water loss, which is exacerbated by rising temperatures and changing precipitation patterns associated with climate change.

Quri Gol Wetland's unique hydrological and climatic conditions, coupled with its ecological significance and vulnerability to climate change, make it a crucial area for studying evaporation and water resource management. Understanding the impacts of climate variability on this wetland will contribute valuable insights for managing similar ecosystems facing the pressures of changing environmental conditions. In this study, the meteorological data of maximum and minimum temperature, solar radiation, precipitation, and evaporation on a daily scale from Tabriz weather station were used over a 30-year period (1991-2020) to provide a comprehensive understanding of the evaporation dynamics in the wetland. Figure 1 illustrates the study area, highlighting the geographical context of this research.

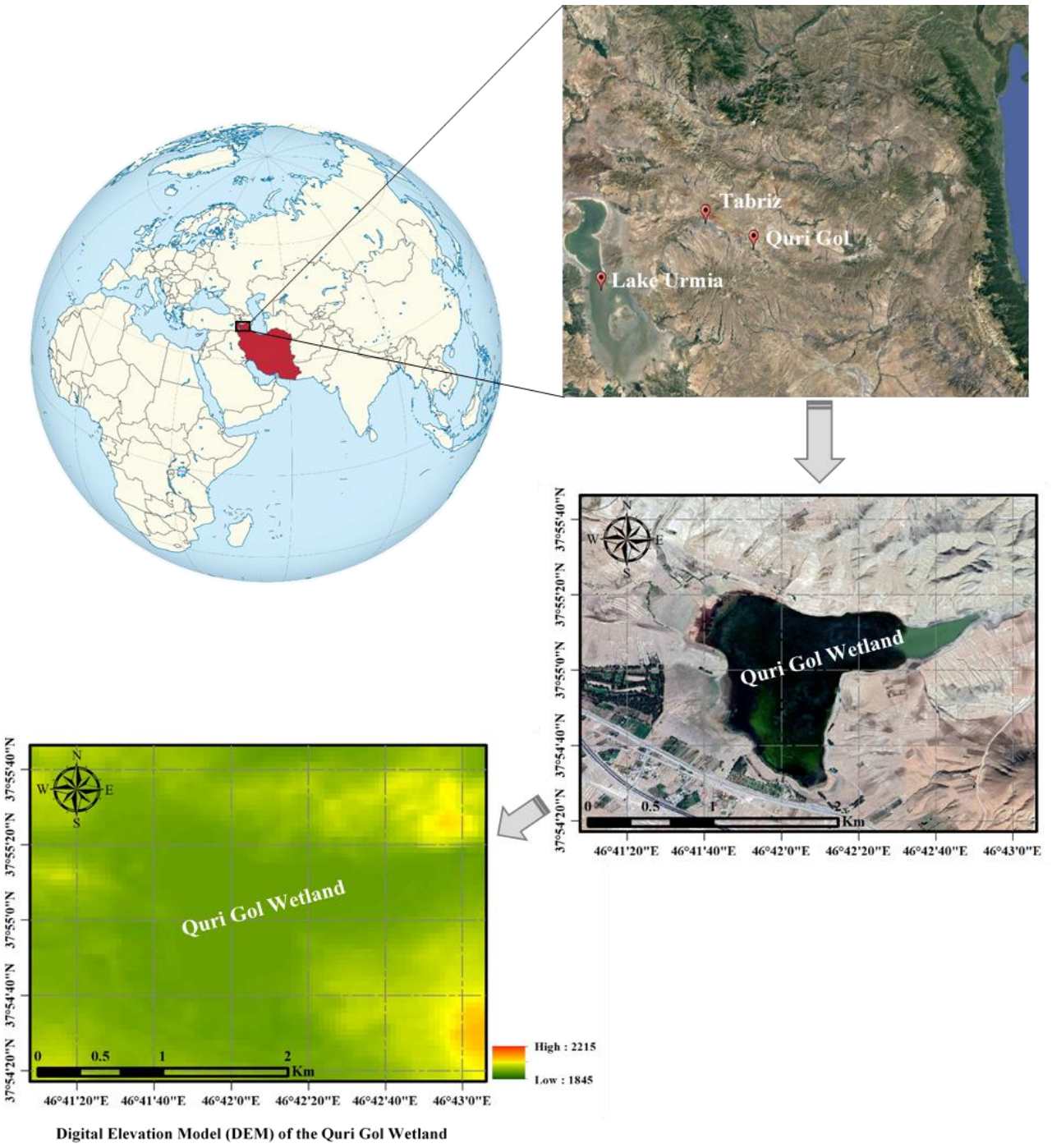

Figure1- Geographical context of the study area, (Sourced from ASTER GDEM)

In this study, the input parameters for evaporation modeling were selected based on their correlation with evaporation rates, ensuring that the most relevant meteorological variables were chosen to optimize model accuracy. The combinations of input parameters are structured to progressively increase complexity, starting with the most basic meteorological data and expanding to include additional variables that influence evaporation. The first combination (I) uses maximum (Tmax) and minimum (Tmin) temperatures as inputs, which are essential for evaporation as temperature directly influences energy availability for the evaporation process. In the second combination (II), solar radiation (Radi.) is added to account for the additional energy input from the sun, a critical factor driving the evaporation rate. The third and most comprehensive combination (III) includes Tmax, Tmin, solar radiation, and precipitation (Prec.), which introduces the role of water availability and atmospheric moisture. The combination of these parameters was determined by evaluating their correlation coefficients with evaporation, ensuring that the selected variables have the highest predictive power for the evaporation process. This approach allows for a systematic exploration of how adding more meteorological data impacts the accuracy of the model, providing a thorough understanding of the most significant contributors to evaporation in the study area. The specific combinations of input variables are detailed in Table 1.

Table1- Input combinations for evaporation modeling using machine learning methods

Combination number	*Input*	*Output*
I	T_{max}, T_{min}	Evap.
II	T_{max}, T_{min}, Radi.	Evap.
III	T_{max}, T_{min}, Radi., Prec.	Evap.

To strengthen the study's robustness, LARS-WG (Long Ashton Research Station Weather Generator) and SDSM (Statistical DownScaling Model) models were implemented to estimate future meteorological data. The LARS-WG model generates synthetic weather data that retain the essential statistical characteristics of historical climatic observations on a daily scale. The SDSM model downscales large-scale atmospheric data to local meteorological variables, thereby increasing the spatial resolution of future climate projections. To ensure the models' accuracy, a rigorous validation process was conducted, comparing the generated data to historical records using performance metrics. This validation demonstrated the models' reliability in capturing key climatic trends. Additionally, an uncertainty analysis was performed to account for potential discrepancies arising from model assumptions, parameter choices, and the inherent variability in climate scenarios. By thoroughly discussing the implementation, accuracy, and uncertainties of the LARS-WG and SDSM models, the robustness of the results is underscored. The predictive meteorological data obtained from these models were subsequently utilized as inputs for the best-performing machine learning model, enabling precise calculations of future evaporation rates under various climate change scenarios. This comprehensive approach bolsters the credibility and reliability of our findings.

2.2 Methods

2.2.1 Random forest (RF)

RF is a technique that utilizes an ensemble of tree-based algorithms to generate repeated forecasts for individual instances. Single decision trees often suffer from overfitting and limited generalizability. Minor changes in training data can significantly alter the structure of a decision tree (Quinlan, 1993). In contrast, RF is adept at learning complex patterns and accounting for the nonlinear relationships between explanatory and dependent variables. Additionally, RF's ability to handle non-normally distributed data allows it to incorporate and integrate diverse data types in the analysis.

The process of combining multiple decision trees is known as ensemble learning. An ensemble consists of several base learners, which are individual models typically generated from training data using base learning algorithms, such as decision trees, neural networks, or other methods. The generalizability of an ensemble model generally surpasses that of its individual base learners. Ensemble methods are particularly valued for their capacity to enhance the performance of weaker models (Schapire 1990).

RF, an advanced method derived from decision tree algorithms, combines the predictions of multiple individual models by employing predefined rules. The fundamental principle of ensemble learning techniques is based on the hypothesis that they are more accurate than single-model training algorithms. This increased accuracy arises from the idea that a collection of predictive models outperforms any single model. Furthermore, ensembles amplify the strengths of individual models while simultaneously mitigating their weaknesses (Kostantis & Pintelas 2004).

2.2.2 Gradient boosted tree (GBT)

The GBT procedure is one of the best learning algorithms and performs categorization with excellent exactitude for many datasets. In this method, the trees are trained one after the other. Each individual tree is mainly instructed with records that were misclassified by the foregoing tree (Malohlava & Candel 2018). This causes the model to concentrate more on intricate cases and less on problems that are straightforward to forecast. Consequently, this approach yields superior outcomes compared to numerous methods, including the linear regression technique and the bagging approach (Breiman 2001). The basic idea of this method was proposed by Breiman (1999) as a method that can interpret the optimization algorithm on a suitable cost function.

2.2.3 Support vector machine (SVM)

Support vector machine is among the directed learning techniques applicable for both categorization and regression tasks. This approach is grounded in the statistical learning framework proposed by Vapnik (1995). SVM is a technique for binary categorization in any given feature space, making it an appropriate technique for forecasting issues (Vapnik 1998). The SVM fundamentally operates as a binary classifier that distinguishes classes with a linear separator. In this technique, the examples nearest to the decision margin are termed support vectors. These vectors define the formula of the decision margin Traditional intelligent modeling techniques, like artificial neural networks, typically aim to reduce the absolute size of the error or the cumulative sum of the squared errors of the training data, whereas SVM models employ the concept of minimal structural error. Using the concept of internal product multiplication, Vapnik demonstrated that the input vector x could first be transferred by a nonlinear conversion to a space with a large dimension, in which space performed the internal product and proved that if a kernel were symmetric, the condition of the theorem could be reviewed. Applying this kernel in a low-dimensional input space can

greatly reduce the product. If the input vector xi is mapped to the specific space by the nonlinear function Φ(x), the decision function will look like Eq (1):

$$f(w.x)=w.\Phi(x)+b \quad (1)$$

Where; w and b are the function parameter vectors. The problem of nonlinear regression can be like the problem of optimization as Eq. (2)

$$\min 1/2 \sum_{ij}^{1} \left[\left(\partial_i - \left[\partial_i\right]^*\right)\left(\partial_j - \left[\partial_j\right]^*\right)\times\langle\Phi(x_i).\Phi(x_j)\rangle\right.$$

$$+\varepsilon\sum_{(i=1)}^{1}\left[\left(\partial_i - \left[\partial_i\right]^*\right)\right] - \sum_{(i=1)}^{1}\left[y_i\left(\partial_i = \left[\partial_i\right]^*\right)\right] \quad (2)$$

Express with the following limitations (Hamel 2009):

$$\sum_{ij}^{1}\left[\left(\partial_i - \left[\partial_i\right]^*\right)=0\right]$$
$$0\leq\partial_i\leq C, \quad i=1,2,\ldots 1$$
$$0\leq \left[\partial_i\right]^*\leq C, \quad i=1,2,\ldots 1 \quad (3)$$

Where; ε is the error tolerances, ∂i and ∂i* are the LaGrange coefficients, C is the cost constant, and <Φ(xi).Φ(xj)> is the internal multiplication of the Φ(xi) and Φ(xj) functions.

2.2.4 Gaussian process regression (GPR)

Gaussian processes represent intricate machine learning algorithms employed for predictive modeling (Williams & Rasmussen 1996) and classification (Williams & Barber 1998). The Gaussian process comprises a collection of random variables, any finite number of which adhere to Gaussian distributions. The Gaussian process is entirely defined by its mean function m(x) and its covariance function k(x, x'). It naturally extends the Gaussian distribution, wherein the mean and covariance are represented by a vector and a matrix, respectively (refer to Eq 4).

$$f\sim GP(m,k) \quad (4)$$

Regression models utilizing Gaussian processes operate on the premise that observations are mutually informative. Gaussian processes specify priors directly within the function space, thereby extending Gaussian distribution principles, where the mean and covariance are represented as vectors and matrices, respectively. Unlike Gaussian distributions which apply to vectors, Gaussian processes apply to functions. Consequently, models based on Gaussian processes inherently comprehend functional and data interdependencies, obviating the need for a separate validation process for generalization. This inherent understanding facilitates the testing of GPR models.

2.2.5 Deep learning (DL)

Deep Belief Networks (DBNs) are recognized as some of the most powerful algorithms in the field of DL. These networks are constructed by combining multiple nonlinear transformations designed to provide a richer and more useful representation of existing data (Bengio 2009). As the number of layers in a neural network increases, the complexity of the optimization problem also escalates (Keyvanrad & Homayounpour 2015). To address this challenge, one of the training methods used for these networks involves unsupervised pre-training algorithms. In this approach, each layer is trained individually at first, followed by a precise and integrated fine-tuning of the entire network (Bengio 2009). In addition to structural differences, Deep Belief Networks and MLPs differ significantly in their training methodologies. For instance, employing backpropagation in training Deep Belief Networks can result in the vanishing gradient problem (Hinton & Osindero 2006).

In Deep Belief Networks, the training process begins with a Restricted Boltzmann Machine (RBM). Once trained, the hidden layer of the RBM becomes the input for the next RBM. This sequential training continues, with each layer being trained based on the output of the previous layer (Glorot & Bengio 2010).

2.2.6 Generalized linear model (GLM)

GLMs are analytical tools for different types of data, and their mathematical relationships were developed by Nelder & Baker (1972). The theory of this model was developed by McCullagh (1984) and applied to hydrology and meteorology by Chandler & Wheater (2002). GLMs include a wide range of statistical models such as linear regression for normally distributed responses, logistic models for binary data, linear models for counting data, and many other useful statistical models through their general method. GLMs can be used when observations of the normal distribution do not exist and when other regression model methods

are not appropriate. This model's effectiveness is satisfactory when compared to conventional modeling techniques (Pregibon & Hastie 2017). These models have been developed with the aim of reducing limiting assumptions in linear regression and can consider dependent variables from any probabilistic distribution of the exponential family (such as binomial distribution, gamma distribution, Poisson distribution, or normal distribution). In GLMs, instead of a response variable, a transformation of the response variable is modeled using a link function.

2.2.7 Statistical down scaling model (SDSM)

The statistical downscaling model employs a dual-step conditional resampling approach for downscaling purposes (Wilby et al. 2007). Initially, this technique downscales predictor variables like temperature and precipitation through integrated regression and stochastic weather generation methods, followed by replication at the station's site (Tatsumi et al. 2013). In fact, SDSM is a combination of the generative method of statistical meteorology and transformed functions. This model was first prepared by Wilby et al. (2002) under the title of version 2.1 in England. This approach relies on employing a mix of regression techniques and the creation of artificial weather data to achieve downscaling. In this model, the local climate is expressed by the macro-scale climate of the region in the form R = F(X), where R represents the small-scaled local climate variable, X is an array of large-scale climatic factors, and F is a subordinate of determination conditional on X, which is obtained based on training and validation of historical data. This model is one of the most widely used downscaling statistical tools that has many applications in meteorological, hydrological, geographical, and environmental studies (Wilby & Harris 2006). Because in this method, large-scale daily circulation templates, such as atmospheric humidity variables, are used on a station scale, it is used when quick and low-cost estimation of climate changes is needed, and in the case of random weather generators and methods of transformed functions, the results can be obtained. This program performs statistical downscaling in five separate steps (Wilby et al. 2002): 1) Screening predictor variables, 2) Model calibration, 3) Feedback of producing observational data artificially, 4) Production of climate change scenarios, and 5) Troubleshooting test and statistical analysis. Figure 2 illustrates the SDSM model's method for downscaling and creating climate scenarios.

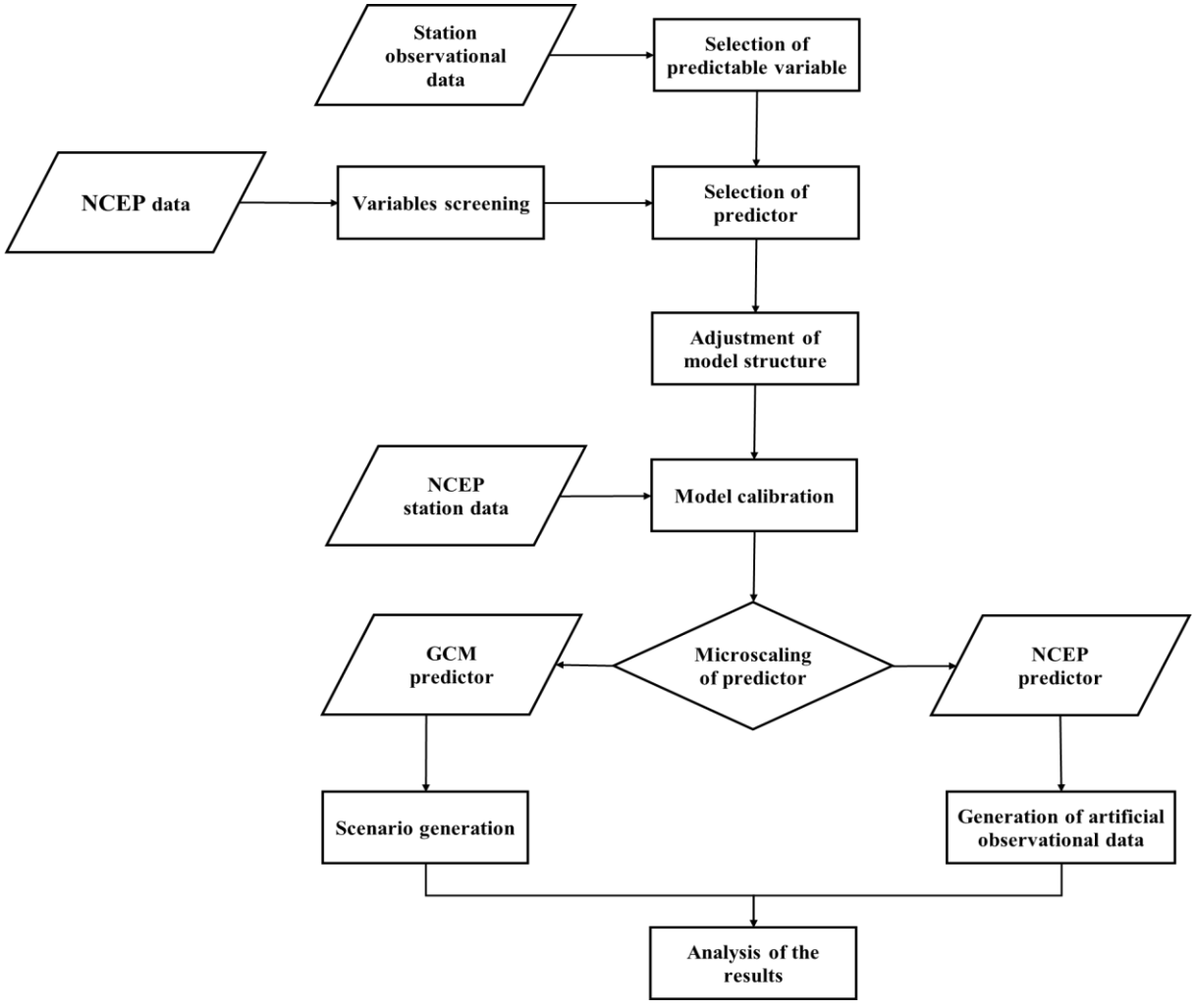

Figure 2- Process of downscaling and climate scenario generation by the SDSM model

2.2.8 Long ashton research station weather generator (LARS-WG)

The Long Ashton Research Station Weather Generator model stands out as a renowned tool for producing stochastic weather data, offering greater utility through iterative computations, minimal data requirements, and straightforward, effective operation (Hay et al. 2000; Dibike & Coulibaly 2005; Kilsby et al. 2007). This model was first presented by Roscoe et al. (1991) and then amended and advanced by Semnoff et al. (1998) at the Langston Research Center. The main reason for producing this model was to surmount the weaknesses of the Markov chain. The 6th version of this model (LARS-WG6) was updated and published in 2018 to microscale the data of the fifth report (CIMP5). Currently, this model is used for two purposes: producing future daily data artificially and generating data for a time frame or stations without statistics.

To generate synthetic data, the model compares long-term daily data (at least 30 years) related to the station (temperature, precipitation, sunshine hours) as inputs. If these two categories of data match, the model is able to produce future time series. For the second function, i.e., generating data for stations without information, the model uses the data of the closest station to the unknown station, checks these data and their statistical characteristics, and due to the same climatic characteristics of these two stations' proximity to each other, uses the data of the known station and their statistical characteristics to generate data for the unknown station. Despite its simpler simulation process and data requirements, this microscaling model possesses a strong predictive capacity for climate change (Semenov & Stratonovitch 2010).

This model's inputs include daily average low and high temperatures, mean precipitation, and hours of sunlight per day. The process of predicting future data is done by this model in four stages:

1. Basic data analysis: Analyzing the statistical characteristics of observational data for the purpose of establishing the statistical characteristics of the data.

2. Initial data generation: Artificially generating data by the model in the base period and determining the statistical attributes of the generated artificial data.

3. Statistical comparison: Matching and comparing the analytical characteristics of observational data and synthetic data.

4. Production of daily data in the future: Using the statistical characteristics of basic data, greenhouse gas emission scenarios, and the output of GCM in the production of daily time series projected into the future with identical statistical characteristics of the basic data.

2.3 Error evaluation indicators

To evaluate the performance of various models, several statistical indicators were employed: the coefficient of determination (R^2), mean absolute error (MAE) to assess the consistency between observed and modeled data, the scatter index (SI) to compare measured values with computed values, Willmott's index (WI) to evaluate the agreement between observed and estimated data, and the Kling-Gupta efficiency (KGE), which provides a comprehensive measure of model performance by accounting for correlation, bias, and variability. The equations for these indicators are presented in Equations (5) to (10).

$$R^2=1-(\sum_{i=1}^{N}〖(O_i-P_i)^2〗)/(\sum_{i=1}^{N}〖(O_i-\bar{O})^2〗) \quad (5)$$

$$MAE=(\sum_{i=1}^{N}(|O_i-P_i|))/N \quad (6)$$

$$SI=\sqrt{((\sum_{i=1}^{N}〖(O_i-P_i)^2〗)/n)}/\bar{O} \quad (7)$$

$$WI=1-[(\sum_{i=1}^{N}〖(O_i-P_i)^2〗)/(\sum_{i=1}^{N}(|P_i-\bar{O}|+|O_i-\bar{O}|)^2)] \quad (8)$$

$$KGE=1-\sqrt{((r-1)^2+(\alpha-1)^2+(\beta-1)^2)} \quad (9)$$

In equations (5) to (9), Oi and Pi represent the observed and predicted (calculated) values, respectively, while $\bar{O}$ denotes the average of the observed values, and N is the total number of observations. For equation (9), which corresponds to the KGE, r is the linear correlation coefficient between the observed and simulated data, measuring the strength of their relationship. α represents the ratio of the standard deviation of the simulated data (σsim) to the standard deviation of the observed data (σobs), indicating variability. Additionally, β denotes the ratio of the mean of the simulated data to the mean of the observed data, representing the bias between the two datasets.

The machine learning models employed in this study were developed using RapidMiner Studio 9.10 software, which provides a robust platform for data analysis and predictive modeling.

Figure 3 shows a schematic structure of methods adopted in this study.

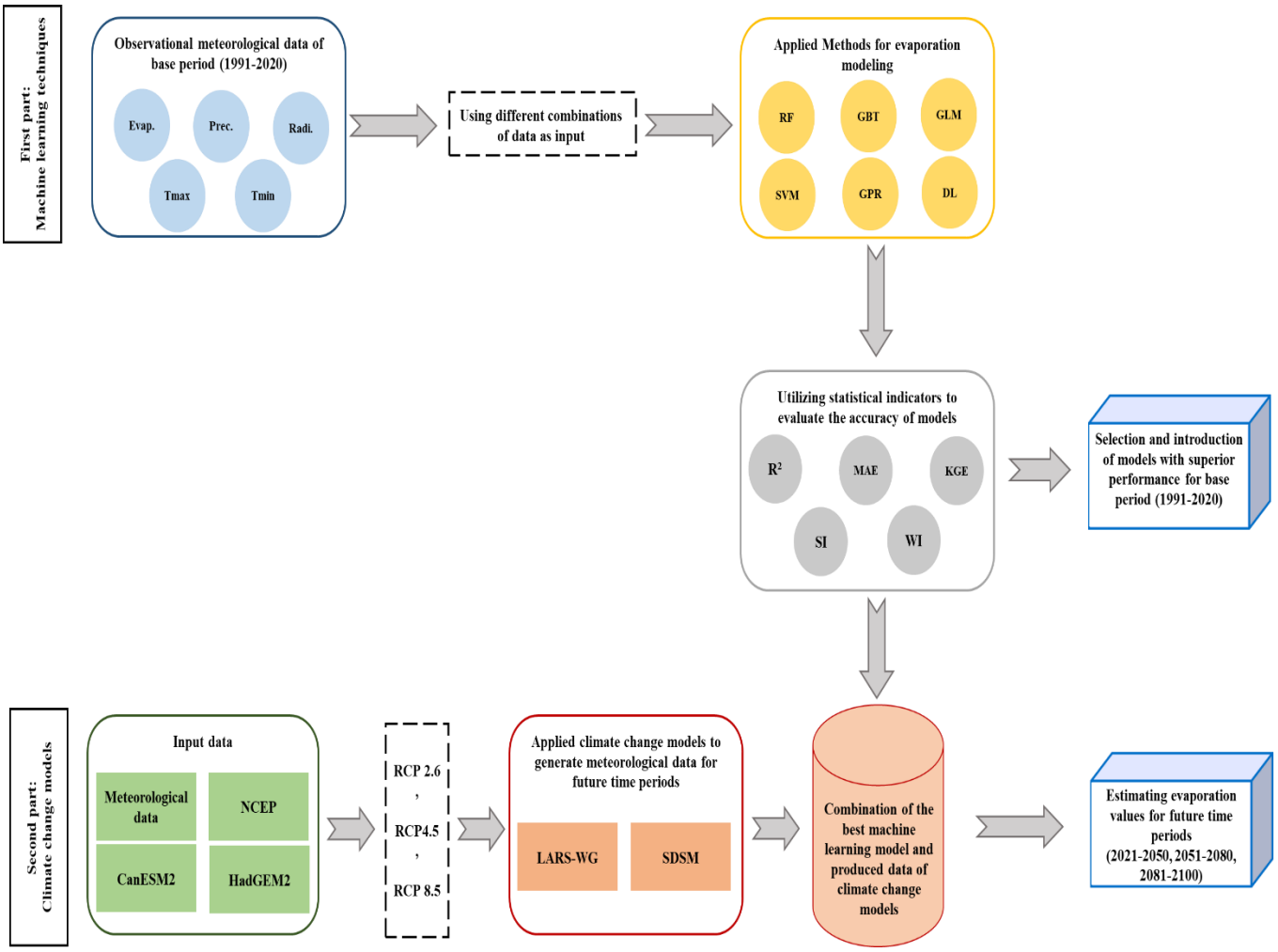

Figure 3- Schematic structure of methods adopted in this study

3. Results

In this study, daily evaporation values for the Quri Gol Wetland were estimated using six machine learning models: RF, GBT, SVM, GPR, GLM, and DL over a 30-year period (1991-2020). The input data included minimum and maximum temperatures, precipitation, and solar radiation, which were combined into three different scenarios. A 70-30% data split was applied, with 70% of the data used for model calibration and 30% for validation. In addition, climate change models were employed to predict future meteorological parameters under three Representative Concentration Pathway scenarios (RCP 2.6, RCP 4.5, and RCP 8.5) for future periods (2021-2050, 2051-2080, and 2081-2100). The estimated meteorological values from these climate change models were then used as inputs to the best-performing machine learning model to predict evaporation for future periods, and the results were compared to the observed evaporation values of the base period.

In this section, the results first evaluate the performance of the machine learning models for the base period, followed by an assessment of the combined machine learning-climate change models.

3.1 Performance of machine learning models

Table 2 presents the statistical parameters, including the coefficient of determination, mean absolute error, scatter index, Willmott's index, and Kling-Gupta efficiency for the machine learning models during the calibration phase.

Table 2- Statistical parameters (R^2, SI, MAE, WI and KGE) values of studied machine learning models in calibration section

Method	*Scenario*	*R^2*	*SI*	*MAE*	*WI*	*KGE*
RF	I	0.811	0.425	0.879	0.946	0.152
	II	0.829	0.405	0.838	0.951	0.143
	III	**0.839**	**0.394**	**0.822**	**0.954**	**0.139**
GBT	I	0.808	0.540	1.391	0.874	0.448
	II	0.833	0.523	1.349	0.883	0.441
	III	0.839	0.519	1.339	0.885	0.439
GLM	I	0.720	0.517	1.221	0.915	0.215
	II	0.729	0.508	1.192	0.919	0.208
	III	0.730	0.507	1.191	0.919	0.207
SVM	I	0.755	0.485	1.005	0.930	0.175
	II	0.759	0.481	0.996	0.931	0.171
	III	0.761	0.479	0.993	0.932	0.170
GPR	I	0.751	0.489	1.019	0.929	0.174
	II	0.736	0.507	1.053	0.924	0.175
	III	0.755	0.489	1.040	0.928	0.176
DL	I	0.771	0.466	0.979	0.934	0.180
	II	0.776	0.462	0.988	0.934	0.186
	III	**0.779**	**0.459**	**0.964**	**0.938**	**0.161**

Table 2 shows the performance of the machine learning models during the calibration phase across three different scenarios. For all methods, the results demonstrate relatively strong performance across the statistical metrics, with RF and GBT generally showing higher R^2 values, indicating better model fit during the calibration phase. Particularly, the RF method achieves the highest R^2 values across all three scenarios, reaching a maximum of 0.839, which suggests a strong correlation between predicted and actual evaporation rates. Additionally, the DL model exhibits competitive performance, with R^2 values ranging from 0.771 to 0.779, showing its ability to capture complex patterns in the data. Moreover, WI values for all methods are consistently above 0.90, indicating high agreement between observed and predicted values. However, models such as the GLM and SVM show slightly lower performance, with higher SI and MAE values. Despite this, the KGE values indicate that all models, while varied in their predictive abilities, perform reasonably well, with RF and GBT models being the most effective overall during calibration. Table 3 presents the statistical parameters (R^2, SI, MAE, WI, and KGE) for the machine learning models in the validation phase, providing an assessment of the models' performance on unseen data. This phase of the analysis is crucial as it reflects the models' ability to generalize beyond the training data, ensuring they are reliable for real-world applications such as evaporation prediction.

Table 3- Statistical parameters (R^2, SI, MAE, WI and KGE) values of studied machine learning models in validation section

Method	*Scenario*	*R^2*	*SI*	*MAE*	*WI*	*KGE*
RF	I	0.809	0.426	0.924	0.943	0.196
	II	0.820	0.417	0.902	0.946	0.200
	III	**0.821**	**0.414**	**0.902**	**0.946**	**0.196**
GBT	I	0.802	0.558	1.489	0.861	0.478
	II	0.810	0.554	1.478	0.863	0.478
	III	0.814	0.550	1.471	0.865	0.475
GLM	I	0.764	0.473	1.190	0.927	0.226
	II	0.775	0.467	1.159	0.928	0.237
	III	0.774	0.467	1.160	0.928	0.237
SVM	I	0.799	0.436	0.956	0.943	0.170
	II	0.807	0.428	0.933	0.945	0.180
	III	0.809	0.426	0.932	0.945	0.179
GPR	I	0.770	0.471	1.029	0.931	0.213
	II	0.746	0.500	1.090	0.922	0.231
	III	0.766	0.477	1.044	0.927	0.233
DL	I	0.816	0.420	0.932	0.944	0.206
	II	0.813	0.433	0.972	0.938	0.239
	III	**0.822**	**0.411**	**0.915**	**0.947**	**0.193**

According to Table 3, in the validation phase, RF and DL emerge as the best-performing models across all statistical parameters. Both models maintain high R^2 values, with RF showing values between 0.809 and 0.821, and DL performing similarly well with R^2 values ranging from 0.813 to 0.822. This indicates that both models exhibit a strong correlation between predicted and observed values, confirming their ability to accurately predict evaporation rates. Additionally, the low SI values for both models, ranging from 0.411 to 0.426, demonstrate minimal error spread, which further supports their precision. These low SI values indicate a tight clustering of predicted data around the observed values, highlighting the high predictive capability of these models in evaporation forecasting.

The MAE values, which measure the average magnitude of the errors between predicted and observed values, also favor RF and DL models. DL models consistently record low MAE values between 0.915 and 0.972, while RF models show slightly lower error margins, ranging from 0.902 to 0.924. These results suggest that DL and RF are particularly adept at minimizing the deviation between predicted and observed evaporation values, making them highly reliable for accurate predictions. Furthermore, the WI values for these models are exceptionally high, with RF scoring between 0.943 and 0.946 and DL between 0.938 and 0.947. The near-perfect WI values indicate strong agreement between the predicted and observed values, reinforcing the reliability of these models.

A significant feature of the analysis is the KGE, which combines aspects of correlation, variability, and bias to offer a comprehensive assessment of model performance. RF and DL again demonstrate balanced performance, with KGE values ranging from 0.193 to 0.239, indicating that these models not only predict evaporation accurately but also capture the overall variability and distribution of the data.

On the other hand, models like GBT and GLM show moderate performance. While the R^2 values for GBT (0.802-0.814) and GLM (0.764-0.775) indicate reasonable correlation, they fall short of the performance seen in RF and DL. This limitation is further reflected in their SI values, where GBT shows higher error spreads (0.550-0.558), signalling reduced precision. Also, the MAE values for GBT (1.471-1.489) are significantly higher than those for RF and DL, indicating that GBT struggles to minimize errors as effectively. Similarly, GLM's MAE values (1.159-1.190) show a lower degree of accuracy compared to the top models. Despite these weaknesses, GLM maintains relatively high WI values (0.927-0.928), indicating that while it may struggle with error minimization, it still demonstrates reasonable agreement between predicted and observed values.

SVM and GPR models occupy an intermediate position in terms of performance. Both models perform reasonably well, with R^2 values close to 0.8, which suggests they can adequately capture the relationship between the inputs and evaporation. SVM models perform particularly well in terms of error minimization, with low MAE values (0.932-0.956) and SI values comparable to those of RF and DL. However, both SVM and GPR models exhibit slightly lower KGE values, indicating potential challenges in balancing bias and variability.

With a general review of the results, RF and DL emerge as the most reliable models for predicting evaporation, demonstrating superior performance across all metrics. The best RF model, particularly in Scenario III, displayed a high R^2 (0.821), low SI (0.414), and a minimal MAE (0.902). Similarly, the DL model, especially in Scenario III, achieved the highest R^2 (0.822), a low SI (0.411), and the lowest MAE (0.915). Both models offer accurate, consistent, and reliable predictions, with low error margins and strong agreement between predicted and observed values. The slightly weaker performance of GBT and GLM suggests that these models may not be as suited for tasks requiring high precision in evaporation prediction, though they still offer a reasonable degree of accuracy. SVM and GPR, while not as robust as RF or DL, offer solid performance but may require further refinement to handle variability and bias more effectively. Overall, the validation results highlight the capability of machine learning models to accurately predict complex hydrological processes like evaporation, with RF and DL proving to be the most effective approaches for this task.

The comparison of calibration and validation results reveals the consistency and reliability of the machine learning models across different phases of evaluation. During the calibration phase, all models demonstrated satisfactory performance, with RF and DL models emerging as the most effective, achieving higher values for metrics such as R^2, WI, and KGE, alongside lower MAE and SI. This strong performance was maintained in the validation phase, where RF and DL models continued to exhibit superior predictive accuracy, though there was a slight decrease in performance compared to the calibration phase, which is common in model testing on unseen data. The consistent performance across both phases indicates that these models can generalize well beyond the training data, offering reliable predictions even under new conditions. Other models, such as GBT and GLM, showed moderate performance, particularly in validation, where a slight increase in MAE and SI was observed, signaling that they may be more prone to overfitting. Overall, as stated above, the results indicate that RF and DL are the most robust and adaptable models for predicting evaporation. Figure 4 provides a comparative visual representation of the simulated and observed evaporation values for the superior scenarios across all six machine learning models. The primary aim of this figure is to evaluate how effectively each model captures the temporal dynamics of evaporation rates over the study period.

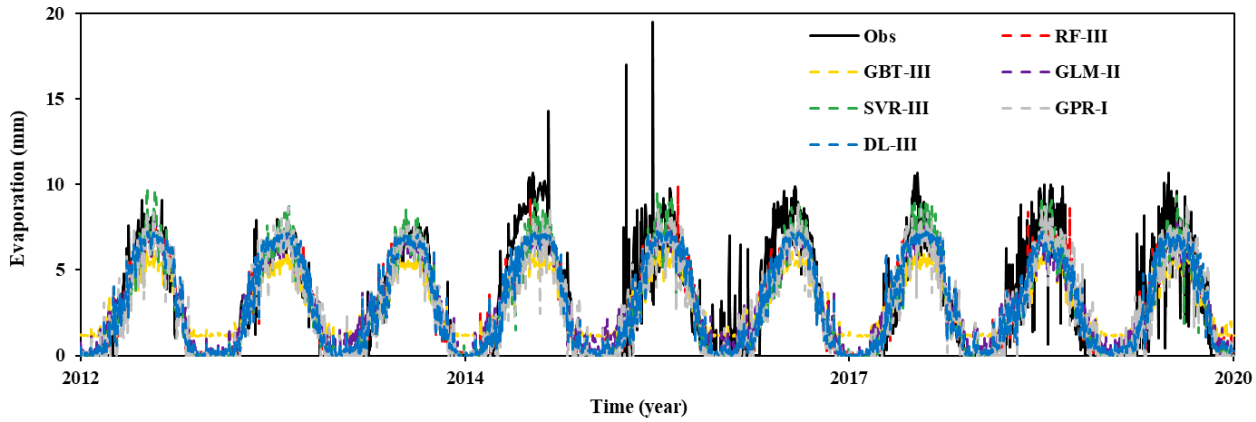

Figure 4- Variation diagrams of simulated and observed evaporation values for the superior scenarios in all six methods

As illustrated in Figure 4, the alignment between the observed and simulated values demonstrates the ability of each model to track real-world variations in evaporation. Both RF and DL show strong agreement with the observed data, reflecting minimal deviations and closely following the seasonal fluctuations of evaporation. This is particularly significant, as it indicates that these models can accurately capture the complex nonlinear interactions between meteorological variables that influence evaporation. The consistent performance of RF and DL, with their ability to maintain accuracy across different time scales and scenarios, underscores their robustness and reliability in modeling evaporation processes. In contrast, models like GLM and SVM exhibit more noticeable discrepancies between observed and simulated values. This suggests that these methods may struggle with capturing the full complexity of evaporation patterns, particularly during peak evaporation periods or abrupt changes in climatic conditions. The larger deviations observed with these models could be attributed to their relatively linear nature (in the case of GLM) or their sensitivity to specific parameter choices (in the case of SVM), which may limit their flexibility in modeling evaporation dynamics. Furthermore, the overall trend illustrated in Figure 4 highlights the seasonal and annual fluctuations in evaporation, with most models effectively capturing the general trend. However, the precision and fit to the data vary depending on the model, with RF and DL demonstrating superior performance in both accuracy and consistency. This visual comparison not only reinforces the quantitative findings from the calibration and validation phases but also provides a clear indication of the predictive strength of RF and DL, especially in handling seasonal variations and sudden shifts in evaporation rates. The figure thus serves as a critical tool for assessing model accuracy in simulating evaporation, offering valuable insights into the strengths and limitations of each approach. The strong performance of RF and DL is particularly noteworthy, as these models demonstrate the capacity to adapt to the inherent variability of evaporation, making them well-suited for hydrological applications in climate-sensitive regions. Figure 5 presents the distribution diagrams comparing the simulated and observed evaporation values for the superior scenarios using the six machine learning models. This figure provides a more detailed statistical view by focusing on how well the models replicate the distributional characteristics of the observed data, particularly in terms of variability, spread, and the prediction of extreme values.

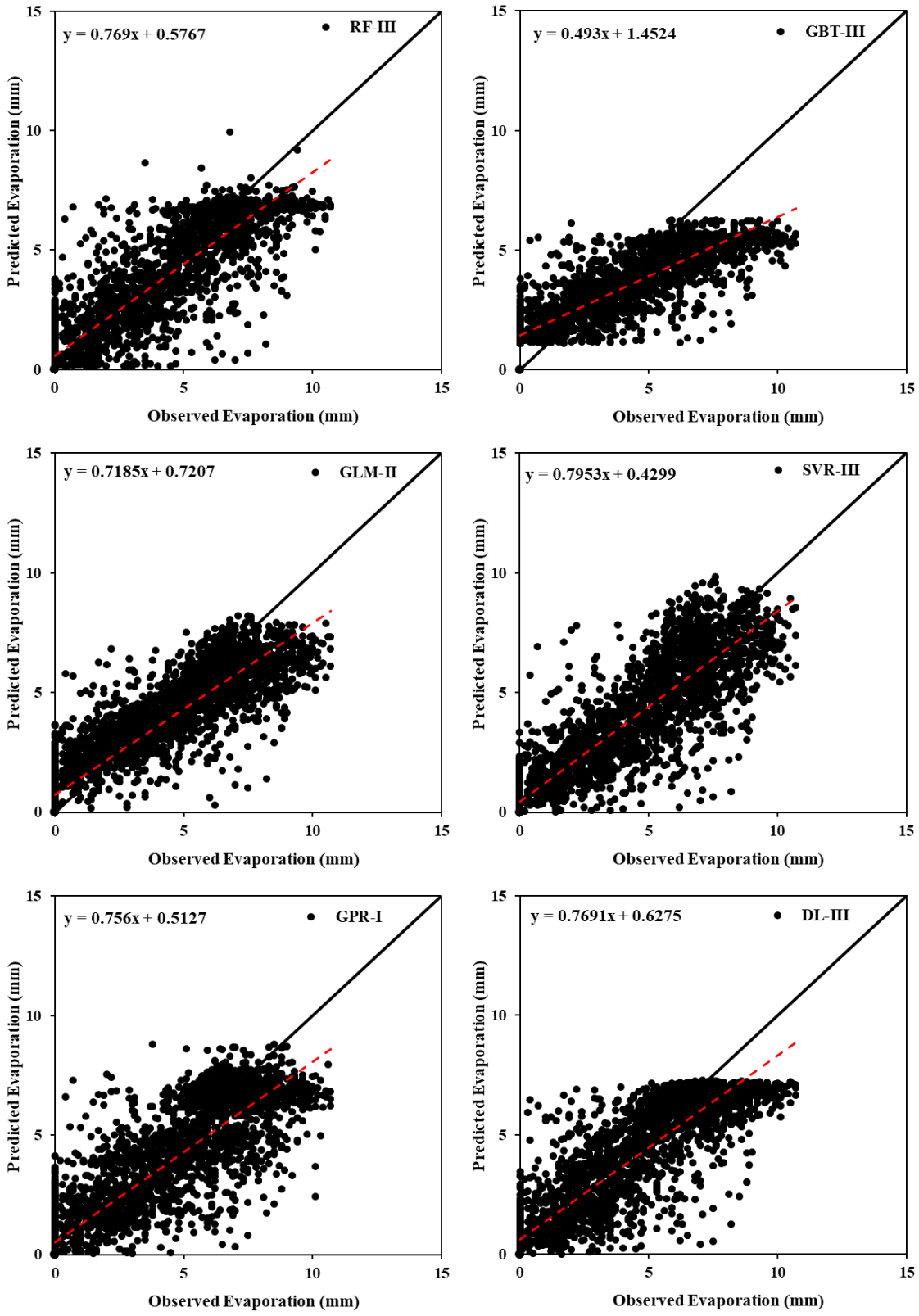

Figure 5- Distribution diagrams of simulated and observed evaporation values for the superior scenarios in all six methods

The distribution plots allow for a deeper examination of the models' ability to capture not just the mean values but also the full range of evaporation rates. As shown in Figure 5, RF and DL once again show a strong alignment with the observed distributions, demonstrating their capability to not only model average evaporation levels but also to accurately reflect the variability and extremes within the dataset. This indicates that both models have a high degree of generalizability and are able to predict not only typical evaporation rates but also more extreme values, which are often critical for understanding hydrological changes under climate stress. The closeness of the simulated distributions to the observed ones for RF and DL suggests that these models effectively handle the inherent variability in evaporation data, including the tail ends of the distribution where extreme

evaporation values may occur. This ability to model variability is crucial in climate change projections, as shifts in extreme values can have significant implications for water resource management. RF and DL's proficiency in capturing both central tendencies and extreme values enhances their applicability for real-world hydrological forecasting and management. Other models, such as GLM and GBT, display greater deviations from the observed distribution shapes. This suggests that while these models may perform adequately for predicting mean evaporation values, they are less adept at capturing the full range of data, particularly at the higher and lower ends of the distribution. The larger spread observed in the GBT model's predictions, for instance, indicates that it may struggle with outlier predictions, making it less reliable in scenarios where accurate modeling of extremes is necessary. Similarly, GLM's inability to capture nonlinear relationships may contribute to its reduced accuracy in reflecting the variability of evaporation values.

By visually comparing the range and distribution of evaporation rates, Figure 5 provides essential insights into the models' ability to handle variability, which is a critical aspect of hydrological forecasting. The clear superiority of RF and DL in these distribution diagrams corroborates earlier quantitative results, reaffirming their role as the most effective models for evaporation prediction in this study. The results also highlight the need for caution when using models such as GLM and GBT in applications where the prediction of extreme or highly variable values is essential. This distribution-based evaluation adds another layer of depth to the overall analysis, providing further evidence that RF and DL models are not only effective in general prediction accuracy but also in handling the distributional complexities of evaporation data, making them suitable for diverse hydrological applications under climate change scenarios.

3.2. Future projections of evaporation and precipitation

Table 4 provides a detailed projection of evaporation values for future periods (2021-2050, 2051-2080, and 2081-2100) using two climate change models, LARS-WG and SDSM, under three different Representative Concentration Pathways (RCP 2.6, RCP 4.5, and RCP 8.5). These models provide critical insight into how climate change is expected to affect evaporation rates in the study area.

Based on the data presented in Table 4, in the period from 2021 to 2050, the LARS-WG model shows a considerable increase in average evaporation values across all scenarios compared to the base period (1991-2020). For example, under the RCP 2.6 scenario, evaporation is projected to increase by 30.40%, while under RCP 8.5, this rise reaches 31.30%. The cumulative evaporation values reflect this trend, showing a significant increase, with the largest being 10565.5 mm under RCP 8.5. On the other hand, the SDSM model also predicts an increase in evaporation, though to a lesser extent than LARS-WG. The SDSM model shows a rise of 5.08% to 11.11% in evaporation across the RCP scenarios, with the RCP 8.5 scenario showing the highest cumulative evaporation increase of 3749.65 mm. Although both models project rising evaporation levels, the LARS-WG model consistently predicts a more substantial increase than the SDSM model. The next period, 2051-2080, follows the same trend, but the rate of increase in evaporation becomes more pronounced. The LARS-WG model shows that cumulative evaporation increases between 33.97% (RCP 2.6) and 41.79% (RCP 8.5). In this period, the average evaporation values rise steadily as well, with a noticeable increase in cumulative evaporation for all scenarios, particularly in the RCP 8.5 scenario, where it reaches 14109.2 mm. Similarly, the SDSM model continues to project an increase in evaporation, though still at a lower magnitude than LARS-WG. In this period, the SDSM model estimates an increase of between 11.08% and 21.49% in average evaporation across the three scenarios, with cumulative values reaching up to 7253.03 mm increase under RCP 8.5. By the final period (2081-2100), the evaporation values predicted by both models have increased significantly compared to the base period, with LARS-WG projecting an even more substantial rise in evaporation. Cumulative evaporation values increase by 32.63% (RCP 2.6) to 50.01% (RCP 8.5), reflecting the growing impact of climate change. Under RCP 8.5, evaporation levels reach 33870.8 mm, demonstrating the severity of expected changes. Meanwhile, the SDSM model also shows continued increases, with cumulative evaporation values rising by 15.95% (RCP 2.6) to 30.14% (RCP 8.5). While the increase in evaporation projected by SDSM remains lower than that of LARS-WG, the rise is still substantial, signaling potential challenges for water management in the region. Overall, both models predict a significant rise in evaporation over the next century, though LARS-WG consistently projects higher rates, especially under the more extreme climate scenarios (RCP 8.5). This projection of substantial evaporation increases highlights the potential for intensified water stress and the need for adaptive strategies in water management.

Table 4- Statistical values of evaporation for future time periods using LARS-WG and SDSM models

Method	*Period*	*Scenario*	*Xmax*	*Xmin*	*Xmean*	***SX***	***SP - SO***
Observations	1991-2020	---	11.500	0.000	3.083	33758.2	---
	2001-2020	---	11.500	0.000	3.093	22579.4	---
LARS-WG	2021-2050	RCP 2.6	8.002	0.114	4.020	44020.2	+10262
		RCP 4.5	7.940	0.114	4.002	43816.5	+10058.3
		RCP 8.5	7.932	0.114	4.048	44323.7	+10565.5
	2051-2080	RCP 2.6	8.000	0.114	4.130	45226.1	+11467.9
		RCP 4.5	8.488	0.114	4.207	46068.3	+12310.1
		RCP 8.5	8.002	0.114	4.371	47867.4	+14109.2
	2081-2100	RCP 2.6	7.907	0.114	4.102	29948.0	+7368.62
		RCP 4.5	8.492	0.114	4.307	31440.6	+8861.21
		RCP 8.5	8.002	0.114	4.640	33870.8	+11291.4
SDSM	2021-2050	RCP 2.6	9.532	0.130	3.240	35474.3	+1716.06
		RCP 4.5	9.250	0.130	3.233	35406.8	+1648.57
		RCP 8.5	9.118	0.130	3.425	37507.9	+3749.65
	2051-2080	RCP 2.6	8.696	0.130	3.424	37498.1	+3739.89
		RCP 4.5	8.758	0.130	3.568	39072.3	+5314.14
		RCP 8.5	9.469	0.130	3.745	41011.2	+7253.03
	2081-2100	RCP 2.6	9.935	0.130	3.586	26180.9	+3601.53
		RCP 4.5	8.842	0.130	4.025	29384.1	+6804.74
		RCP 8.5	8.958	0.130	3.914	28569.0	+5989.61

Table 5 provides a projection of precipitation values for the same future periods (2021-2050, 2051-2080, and 2081-2100) using the LARS-WG and SDSM models under RCP 2.6, RCP 4.5, and RCP 8.5 scenarios.

As detailed in Table 5, unlike the substantial increases in evaporation presented in Table 4, the increase in precipitation is much smaller and varies across the models and scenarios. During the period from 2021 to 2050, the LARS-WG model shows relatively minor changes in cumulative precipitation. Under RCP 2.6, the cumulative precipitation slightly decreases by 106.78 mm, while in the more severe RCP 8.5 scenario, it increases by 530.52 mm, reflecting an increase of about 7%. This is far less than the concurrent rise in evaporation. The SDSM model similarly projects only modest changes in precipitation during this period, with cumulative precipitation increases ranging from 34.12 mm (RCP 2.6) to 321.72 mm (RCP 8.5). Although these projections indicate a slight increase in rainfall, the gap between rising evaporation and precipitation is stark, suggesting potential challenges for water resources in the future. For the 2051-2080 period, the LARS-WG model shows a further increase in precipitation under RCP 2.6 and RCP 8.5, but again, the changes remain modest in comparison to evaporation. Cumulative precipitation increases range from 510.12 mm (RCP 2.6) to 634.72 mm (RCP 8.5), representing an increase of approximately 8%. Similarly, the SDSM model shows a slight increase in cumulative precipitation, with a maximum increase of 330.52 mm under RCP 2.6. These results indicate that while precipitation will rise, the increase is not proportional to the expected rise in evaporation, leading to potential water deficits in the region. By the 2081-2100 period, the discrepancy between evaporation and precipitation becomes more pronounced. The LARS-WG model predicts cumulative precipitation increases of 156.21 mm (RCP 2.6) to 775.61 mm (RCP 8.5), representing a maximum increase of 16% under RCP 8.5. However, this increase is far lower than the corresponding rise in evaporation, which reaches 50.01% under the same scenario. The SDSM model also shows only slight increases in precipitation, with the highest cumulative increase being 234.51 mm under RCP 2.6. In the more severe RCP 8.5 scenario, the model even predicts a slight decrease in cumulative precipitation (-217.59 mm), further highlighting the growing gap between precipitation and evaporation.

Table 5- Statistical values of precipitation for future time periods using LARS-WG and SDSM models

Method	Period	scenario	Xmax	Xmin	Xmean	SX	SP – SO
Observations	1991-2020	---	46.4	0	0.668	7316.68	---
	2001-2020	---	46.4	0	0.662	4833.89	---
LARS-WG	2021-2050	RCP 2.6	46.8	0	0.658	7209.90	-106.78
		RCP 4.5	33	0	0.685	7501.00	+184.32
		RCP 8.5	37	0	0.717	7847.20	+530.52
	2051-2080	RCP 2.6	41.2	0	0.715	7826.80	+510.12
		RCP 4.5	41.2	0	0.654	7161.60	-155.08
		RCP 8.5	38.1	0	0.726	7951.40	+634.72
	2081-2100	RCP 2.6	37.2	0	0.684	4990.10	+156.21
		RCP 4.5	40.2	0	0.701	5117.80	+283.91
		RCP 8.5	42.3	0	0.768	5609.50	+775.61
SDSM	2021-2050	RCP 2.6	38.1	0	0.671	7350.80	+34.12
		RCP 4.5	39.9	0	0.676	7402.80	+86.12
		RCP 8.5	33.5	0	0.698	7638.40	+321.72
	2051-2080	RCP 2.6	38.9	0	0.698	7647.20	+330.52
		RCP 4.5	43.1	0	0.674	7377.20	+60.52
		RCP 8.5	34.6	0	0.686	7516.90	+200.22
	2081-2100	RCP 2.6	38.1	0	0.694	5068.40	+234.51
		RCP 4.5	36.2	0	0.688	5019.60	+185.71
		RCP 8.5	32.3	0	0.632	4616.30	-217.59

The data in Table 5 indicate that while precipitation is expected to increase slightly under most scenarios, these increases are far outweighed by the substantial rise in evaporation. This imbalance between rising evaporation and relatively stable or slightly increasing precipitation suggests that the region may face growing water stress in the future, as the rate of water loss due to evaporation will likely exceed the rate of water replenishment through precipitation. This situation underscores the potential for dehydration in the study area and highlights the need for proactive water management strategies to mitigate the effects of climate change. Additionally, these findings indicate that further research is required to fully understand the complex interactions between evaporation and precipitation in the region and to develop comprehensive plans for managing water resources under future climate conditions.

Figure 6 illustrates the time distribution of observed and calculated evaporation values over the future periods using the LARS-WG and SDSM models, providing a critical comparative analysis between historical data (1991-2020) and model predictions for the future periods of 2021-2050, 2051-2080, and 2081-2100 under various climate scenarios (RCP 2.6, RCP 4.5, and RCP 8.5). The primary aim of this figure is to track how evaporation is expected to change over time in response to climate change, as projected by both models.

Figure 6. Time distribution diagram of observed and calculated evaporation values for future time periods using LARS-WG and SDSM models

Figure 6 highlights a clear upward trend in evaporation rates, with both models consistently showing an increase in evaporation for future time frames, particularly under the RCP 8.5 scenario, which represents a more extreme climate change projection. The calculated evaporation values for both LARS-WG and SDSM align closely with historical observations for the initial period, indicating that both models are well-calibrated for the region. As the future periods progress, however, the models begin to diverge slightly, with the LARS-WG model projecting somewhat higher evaporation values compared to the SDSM model. This discrepancy may be attributed to differences in how each model handles various climatic factors such as temperature, solar radiation, and humidity. The LARS-WG model shows a steeper increase in evaporation, particularly in the second half of the century (2051-2100), which suggests that this model may be more sensitive to extreme climate change scenarios. This is important for understanding long-term shifts in water balance, as rising evaporation could exacerbate water shortages in the region. On the other hand, the SDSM model exhibits a more moderate increase in evaporation values, potentially reflecting a more conservative response to climate variables. Overall, the time distribution diagrams in Figure 6 underscore the robustness of both models in predicting long-term shifts in evaporation patterns. These results highlight the urgent need for adaptive water resource management strategies to mitigate the impacts of rising evaporation in the Quri Gol Wetland region, as projected by both models.

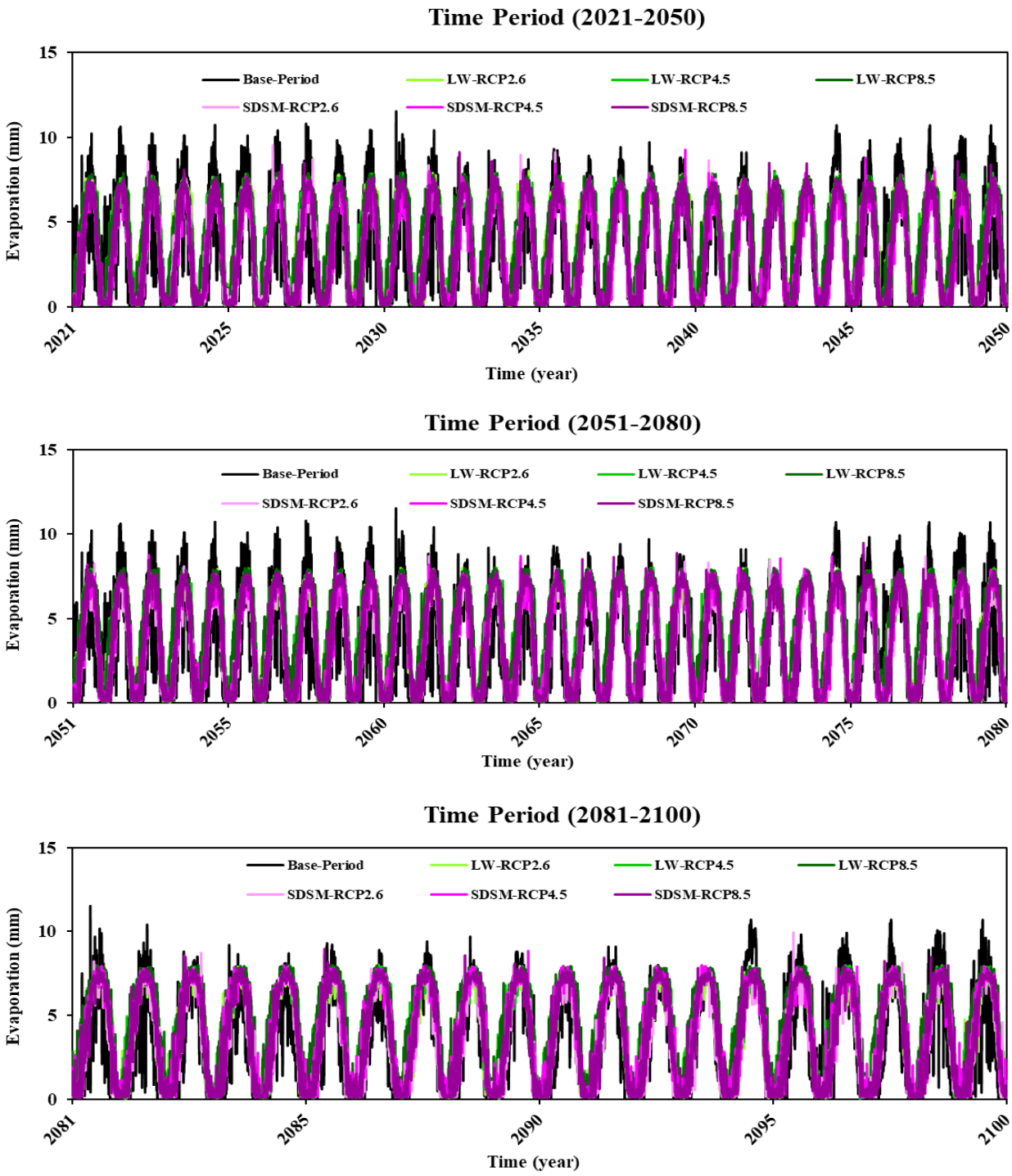

Fig 6- Time distribution diagram of observed and calculated evaporation values for future time

Figures 7 and 8 provide comparative visual analyses of future climate projections using both the LARS-WG and SDSM models. Figure 7 presents violin plots, combining the advantages of box plots and density plots, to offer a statistical comparison of observed and calculated evaporation values. These plots comprehensively visualize the distribution, density, and variability of evaporation data across different time periods and climate scenarios. Figure 8, on the other hand, uses barplots to compare observed and calculated precipitation values, providing a visual representation of precipitation predictions under varying RCP scenarios. Together, these figures highlight the differences in projected evaporation and precipitation trends, providing a clear depiction of their statistical characteristics.

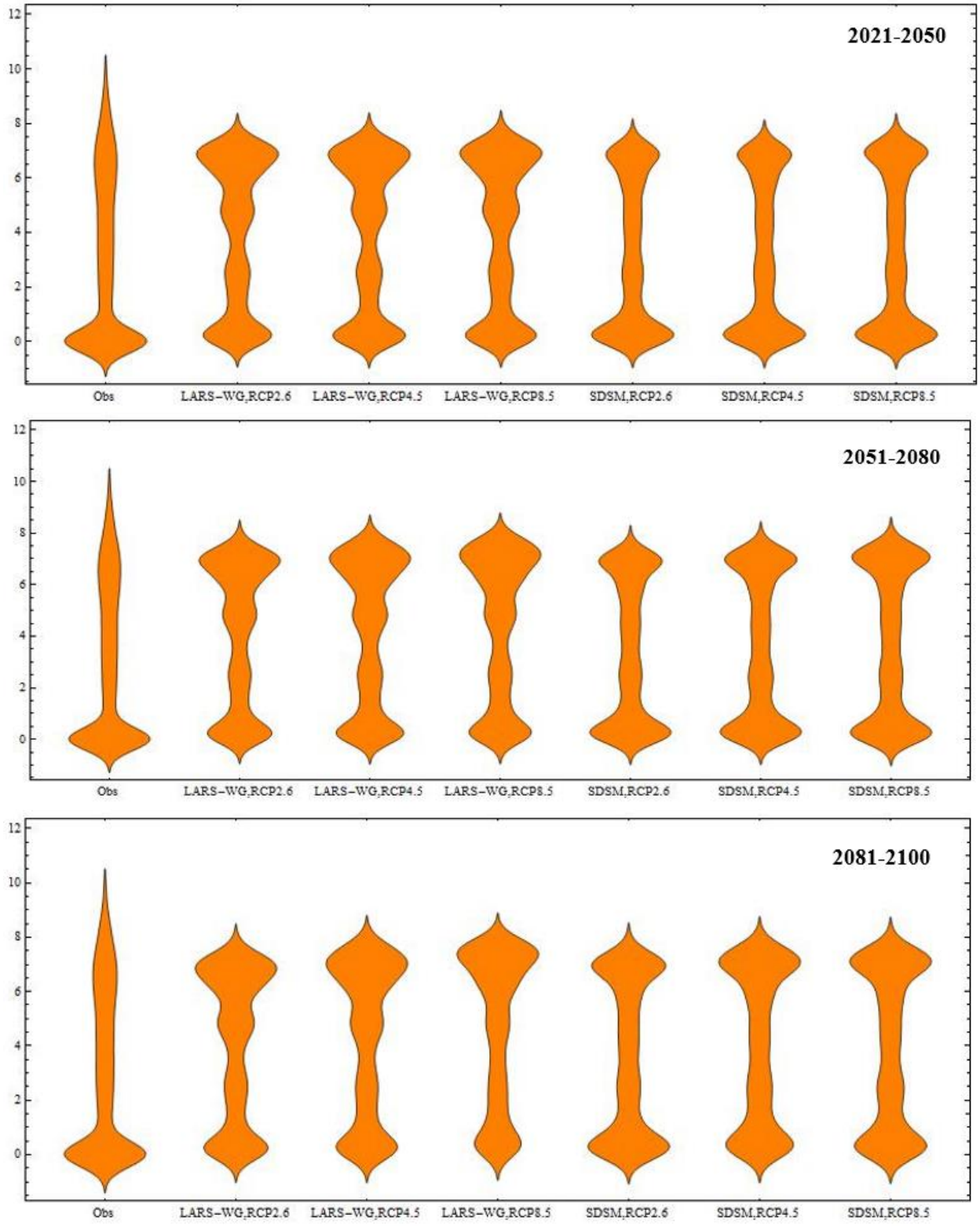

Figure 7- Violin plots of observed and calculated evaporation values for future time periods using LARS-WG and SDSM models

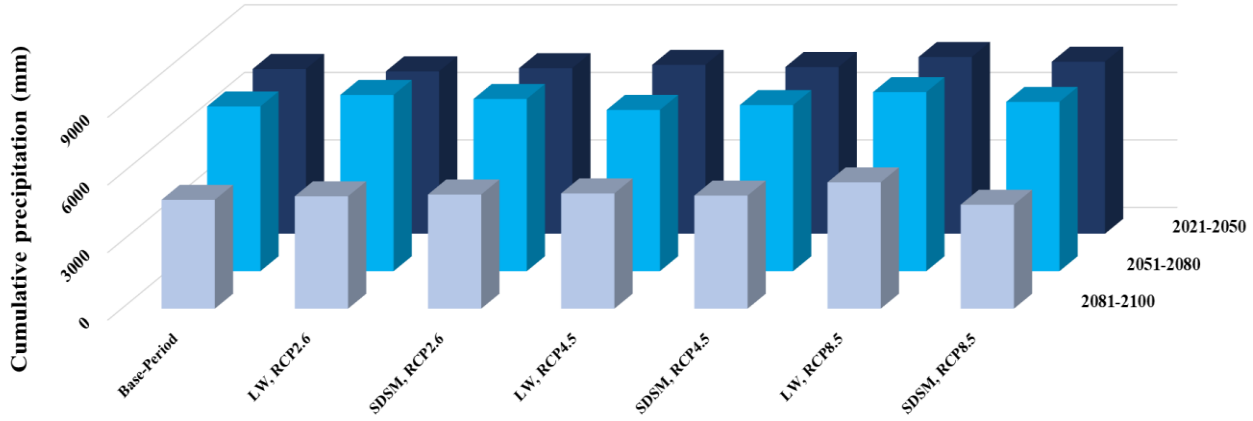

Figure 8- Barplots of observed and calculated precipitation values for future time periods using LARS-WG and SDSM models

According to Figure 7, the width of the violin plots represents the density of data points at various levels of evaporation, with wider sections indicating a higher concentration of values. The observed evaporation values, representing the historical period, display a relatively compact distribution, indicating lower variability in evaporation during the historical period. In contrast, the future predictions show broader distributions, particularly under the more severe climate scenarios (RCP 4.5 and RCP 8.5), indicating greater variability and uncertainty in future evaporation rates. The violin plots reveal that LARS-WG tends to predict slightly higher variability in evaporation values compared to SDSM, especially in the upper quartiles. This suggests that LARS-WG is more sensitive to extreme events and may be better at capturing the tails of the distribution, where extreme evaporation values occur. The SDSM model, while still showing an increase in variability compared to historical data, exhibits a narrower range of predicted values, implying that it projects more conservative changes in evaporation. These plots are critical in understanding the uncertainty associated with future projections, as the broader distribution of values in both models reflects the increased variability in climate-driven evaporation. The higher density of extreme evaporation values in LARS-WG, particularly under RCP 8.5, emphasizes the potential for more frequent and intense evaporation events, which could have significant implications for water availability in the region. The violin plots in Figure 7 serve as a visual confirmation of the increased risk of hydrological extremes due to climate change, reinforcing the need for proactive water management policies to address potential water shortages in the Quri Gol Wetland.

As revealed in Figure 8, the observed data, representing the historical period (1991-2020), show a relatively stable range of precipitation values. This suggests that historical precipitation levels have been relatively consistent, which provides a baseline for comparing future projections. Both models predict an increase in variability for future precipitation values, but the magnitude of the change is far less pronounced compared to the increases in evaporation depicted in Figures 6 and 7. The LARS-WG model predicts a slightly higher median precipitation value compared to SDSM, particularly under the more severe climate scenarios (RCP 4.5 and RCP 8.5). This could indicate that LARS-WG is more optimistic about potential increases in precipitation. However, the overall range of projected precipitation values remains relatively limited, and the barplots for both models indicate that extreme precipitation events are expected to be less prominent compared to extreme evaporation trends. The SDSM model demonstrates a more consistent range, with a smaller spread of predicted values, suggesting a relatively stable projection for future precipitation levels under various scenarios. Despite the slight increases in precipitation shown in both models, the changes are marginal when compared to the substantial increases in evaporation. This imbalance between rising evaporation and relatively stable precipitation levels suggests that the Quri Gol Wetland region could experience increased water deficits in the future. The barplots in Figure 8 reinforce the notion that precipitation alone will not be sufficient to counterbalance the expected rise in evaporation, emphasizing the need for integrated water resource management strategies that account for both increasing evaporation rates and limited precipitation changes.

4. Discussion

This study provides a comprehensive evaluation of how climate change may impact evaporation and precipitation dynamics in the Quri Gol Wetland, employing a combination of machine learning models and climate change projections. The integration of these approaches offers valuable insights into the region's hydrological future, highlighting trends, challenges, and opportunities for sustainable water resource management.

The analysis highlights the effectiveness of machine learning models in capturing the intricate relationships between meteorological variables and evaporation. Advanced techniques, particularly RF and DL, excelled in predicting evaporation patterns. Their ability to model complex nonlinear relationships demonstrates their suitability for hydrological applications. These models reliably reflected observed trends while maintaining accuracy across varying climatic conditions, indicating their

robustness for long-term forecasting. In contrast, other methods like GLM and GBT showed moderate performance. While useful for capturing general trends, these models appeared less adept at handling the variability and extremes that characterize evaporation processes. This underscores the importance of selecting methods tailored to the study's objectives, especially when accurate representation of complex patterns is critical. The comparative analysis of the present study with similar studies is given in Table 6.

Table 6- Comparative analysis of the present study with similar research

Study	*Models Used*	*Location*	*Key Variables*	*Performance Metrics*	*Key Findings*
Present Study	RF, GBT, GLM, SVM, GPR, DL	Quri Gol Wetland (Iran)	maximum and minimum temperature, solar radiation, and precipitation	RF(III) (R^2=0.821, SI= 0.414, and MAE= 0.902) and DL(III) (R^2=0.822, SI= 0.411, and MAE= 0.915) had the highest accuracy	RF and DL models performed better than other models
Tezel & Buyukyildi 2016	MLP, radial basis function network (RBFN), and SVR	Beysehir (Turkey)	temperature, relative humidity, wind speed, and precipitation	the best performance was obtained using MLP (SCG 4,2,2,1) with R^2=0.905.	The comparison of all obtained model performances indicated that all ANN models were more successful than SVR
Wang et al. 2017b	FG, least square support vector regression (LSSVR), MARS, M5Tree and MLR	Dongting Lake Basin (China)	air temperature, surface temperature, wind speed, relative humidity, and sunshine hours	LSSVR and FG models perform better than the MARS, M5Tree and MLR models with respect to statistical indices (MAE, RMSE, and R^2).	the FG, LSSVR and MARS models outperform the M5Tree model, and surface temperature, sunshine hours and air temperature were most important influencing variables
Rezaei-Balf et al. 2019	SVM, model tree (MT), ensemble empirical mode decomposition (EEMD) coupled with support vector machine (EEMD-SVM) and EEMD model tree (EEMD-MT)	Siirt and Diyarbakir stations (Turkey)	wind speed, temperature, relative humidity, and solar radiation	The EEMD-MT model had more accurate performance for both Siirt (NSE=0.89, WI=0.97, and LMI=0.70) and Diyarbakir (NSE=0.92, WI=0.98, and LMI=0.80) stations	It was concluded that the proposed pre-processing technique is very promising for complex time series forecasting
Malik et al. 2022	DL and Gradient Boosting Machine (GBM)	Kiashahr (Iran) and Ranichauri (India)	Monthly maximum temperature and monthly pan evaporation	The best DL models in Kiashahr and Ranichauri stations recorded MAE values of 0.5691 & 0.3693 mm/month, respectively	The DL model was more accurate in both Kiashahr and Ranichauri stations
Bilali et al. 2023	Extra Tree, XGBoost, SVR, and DNN	Bouregreg watershed (Morocco)	air temperature, Relative Humidity, atmospheric pressure, wind speed, and solar radiation	the developed models were accurate in reproducing the daily pan evaporation with NSE ranging from 0.76 to 0.83	the air temperature, solar radiation, followed by relative humidity were the most important climate variables for evaporation estimation in the study area.
Wang et al. 2023	physical model PenPan, MARS, RF, and MLR	China	atmospheric pressure, relative humidity, sunshine hours, air temperature, and wind speed	The RF model had the highest R^2 (0.95± 0.029/ 0.98 ± 0.019) and lowest RMSD (0.62 ± 0.17 mm day^{-1}/9.06 ± 3.45 mm $month^{-1}$) values	It was concluded that the MARS and RF estimates were better than PenPan, and the results of MLR were the worst.

Table 6 offers a detailed comparison between the results of the present study and similar research conducted in various regions, focusing on the use of different machine learning models for predicting evaporation. The table highlights the models employed, key input variables (such as temperature, solar radiation, and precipitation), performance metrics, and key findings. In this study, as highlighted previously, RF and DL models emerged as the most accurate for predicting evaporation in the Quri Gol Wetland (Iran), with the best performance metrics, demonstrating their effectiveness in capturing evaporation patterns. When compared with other studies, models like RF and DL consistently performed better than traditional statistical methods such as MLR and even other machine learning methods like SVM and GBM. The superior accuracy of RF and DL, particularly in handling nonlinear relationships and complex climatic interactions, is evident across different regions and datasets. This comparative analysis underscores the robustness of these models in hydrological applications and highlights their adaptability to various climatic conditions. Table 6 effectively demonstrates that machine learning models, especially RF and DL, provide

reliable and accurate predictions for evaporation, reaffirming their utility in water resource management under diverse climatic conditions.

Climate change projections indicate that evaporation in the Quri Gol Wetland will increase substantially in the coming decades, with the magnitude of the change intensifying under more extreme climate scenarios. This upward trend aligns with rising temperatures and enhanced solar radiation driven by global warming. These changes suggest an urgent need for strategies to address the anticipated water losses from the wetland. In contrast, projected changes in precipitation are comparatively modest. While there are slight increases in precipitation under certain scenarios, these do not compensate for the significant rise in evaporation. This imbalance highlights the potential for water stress in the wetland, which could adversely affect its ecological health and the agricultural activities dependent on it. Such findings underscore the vulnerability of water systems in arid and semi-arid regions under changing climatic conditions.

The study's findings emphasize the need for proactive water resource management strategies to address the anticipated challenges posed by climate change. Adaptation measures should focus on reducing water loss and ensuring sustainable usage. These may include optimizing irrigation systems, enhancing water storage infrastructure, and implementing conservation practices to preserve wetland ecosystems. The application of machine learning models provides a reliable tool for forecasting and monitoring hydrological variables, enabling more informed decision-making. The flexibility and precision of RF and DL models, in particular, make them valuable for real-time applications, offering opportunities to anticipate and mitigate the impacts of climatic variability. Furthermore, incorporating climate change projections into management policies can help stakeholders better prepare for the long-term impacts on water availability.

While the study offers robust insights, several limitations should be acknowledged. The projections depend heavily on the quality of input data and assumptions within the climate change scenarios used. Non-climatic factors, such as human interventions and land use changes, were not accounted for, potentially limiting the scope of the findings. Additionally, the study primarily focused on meteorological variables, omitting other important factors like groundwater interactions and soil moisture dynamics. Future research should aim to address these gaps by integrating additional variables and exploring the interplay between climatic and non-climatic factors. A more holistic approach would provide a deeper understanding of the hydrological processes affecting the wetland. Expanding the analysis to include seasonal and annual variations could also yield more nuanced insights, particularly in capturing short-term dynamics alongside long-term trends.

5. Conclusions

This study evaluated the future impacts of climate change on evaporation and precipitation in the Quri Gol Wetland by integrating machine learning models with climate change projections. Evaporation values for the present period (1991-2020) were estimated using machine learning models (RF, GBT, GLM, SVM, GPR, and DL), and future predictions were made by combining machine learning methods with climate models (LARS-WG and SDSM) under three different RCP scenarios (2.6, 4.5, and 8.5). This comprehensive approach assessed potential hydrological changes in the wetland for the future periods of 2021-2050, 2051-2080, and 2081-2100.

Machine learning models performance:

RF and DL emerged as the most accurate models for predicting evaporation during the calibration and validation phases.

In the validation phase, the best RF model achieved an R^2 of 0.821, an SI of 0.414, and an MAE of 0.902, while the best DL model recorded an R^2 of 0.822, an SI of 0.411, and an MAE of 0.915.

These results indicate the robustness of RF and DL in capturing complex nonlinear relationships between meteorological variables and evaporation.

Future evaporation and precipitation trends:

Evaporation rates are expected to increase significantly across all future periods, with the highest rise projected under the RCP 8.5 scenario.

The LARS-WG model forecasted a maximum increase of 50.01% in evaporation by 2081-2100, while the SDSM model projected a smaller rise of up to 30.14%.

Precipitation increases were comparatively smaller, with a maximum of 16% projected by the LARS-WG model under the RCP 8.5 scenario during the same period.

Implications for hydrological balance:

The imbalance between rapidly increasing evaporation and relatively stable precipitation levels highlights potential water scarcity challenges for the Quri Gol Wetland.

These findings underscore the wetland's vulnerability to climate change and its implications for ecological and agricultural sustainability.

The findings of this study underline the critical need for adaptive strategies to address the challenges posed by climate change on the Quri Gol Wetland. As evaporation rates are projected to rise significantly, outpacing smaller increases in precipitation, effective water resource management becomes imperative to sustain the ecological and agricultural viability of the region. Future research should expand the scope of analysis by considering a broader range of climate scenarios and incorporating additional variables such as land use changes, socio-economic factors, and groundwater interactions. These additions would provide a more holistic understanding of the hydrological impacts of climate change and improve the predictive accuracy of the models. Moreover, it is essential to investigate the implications of these hydrological changes on groundwater recharge and ecosystem health, as these factors are vital for maintaining the wetland's long-term stability. Integrating these variables into future studies will offer deeper insights into the complex interactions between climate, hydrology, and human activity. Such research can guide the development of sustainable resource management practices, ensuring that the Quri Gol Wetland and similar regions can effectively adapt to the pressures of a changing climate.

References

Ahmadaali J, Barani G A, Qaderi K, Hessari B (2018). Analysis of the effects of water management strategies and climate change on the environmental and agricultural sustainability of Urmia Lake basin, Iran. Water 10: 160

Azizi G, Nazif S, Abbasi F (2017). An assessment of the influence of climate change on Urmia Lake water level reduction. Journal of Interdisciplinary Studies in the Humanities 9(4): 1-21

Bilali A E, Abdeslam T, Ayoub N, Lamane H, Ezzaouini M A, Elbeltagi A (2023). An interpretable machine learning approach based on DNN, SVR, Extra Tree, and XGBoost models for predicting daily pan evaporation. J. Environ. Manage. 327: 116890

Bengio Y (2009). Learning Deep Architectures for Artificial Intelligence. Found. Trends Mach. Learn. 2(1): 1-127

Breiman L (1999). Using adaptive bagging to debias regressions. Technical Report 547, Statistics Department, University of California.

Breiman L (2001). Random forests. Mach. Learn. 45(1): 5-32

Chandler R E & Wheater H S (2002). Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. Water Resour. Res. 38(10): 1-10

Dibike Y B & Coulibaly P (2005). Hydrologic impact of climate change in the Saguenay Watershed: Comparison of Ownscaling Methods and Hydrologic Models. *J. Hydrol.* 307: 145-163

Ghaemi A, Rezaie-Balf M, Adamowski J, Kisi O, Quilty J (2019). On the applicability of maximum overlap discrete wavelet transform integrated with MARS and M5 model tree for monthly pan evaporation prediction. Agric. For. Meteorol. 278:107647. https://doi.org/10.1016/j.agrformet.2019.107647.

Ghorbani M A, Deo R C, Karimi V, Yaseen Z M, Terzi O (2018). Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. Stoch. Env. Res. Risk Assess. 32(6): 1683-1697. https://doi.org/10.1007/s00477-017-1474-0.

Glorot X & Bengio Y (2010). Understanding the difficulty of training deep feedforward neural networks. Proceedings of Machine Learning Research 9:249-256

Goyal M K, Bharti B, Quilty J, Adamowski J & Pandey A (2014). Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert Syst. Appl. 41(11): 5267-5276

Goyal M K & Ojha C S P (2014). Evaluation of rule and decision tree induction algorithms for generating climate change scenarios for temperature and pan evaporation on a lake basin. *J. Hydrol. Eng.* 19:828-835

Guven A & Kisi O (2013). Monthly pan evaporation modeling using linear genetic programming. *J. Hydrol.* 503:178-185

Hamel L (2009). Knowledge Discovery with Support Vector Machines. John Wiley, Hoboken, New Jersey.

Hay L E, Wilby R L & Leavesley G H (2000). A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. J. Am. Water Resour. Assoc. 36: 387-397

Helfer F, Lemckert C H & Zhang H (2012). Impacts of climate change on temperature and evaporation from a large reservoir in Australia. *J. Hydrol.* 1-38

Hinton G E, Osindero Y W S (2006). A fast learning algorithm for deep belief nets. Neural Comput. 18(7): 1527-1554

Keshtegar B, Piri J & Kisi O (2016) A nonlinear mathematical modeling of daily pan evaporation based on conjugate gradient method. Comput. Electron. Agric. 127:120-130

Keyvanrad MA, Homayounpour MM (2015) Deep belief network training improvement using elite samples minimizing free energy. Int. J. Pattern Recognit. Artif. 29(5): 155-166

Kilsby C G, Jones P D, Burton A, Ford A C, Fowler H J, Harpham C, James P, Smith A & Wilby R L (2007). A daily weather generator for use in climate change studies. Environ. Model. Softw. 22: 1705-1719

Kim S, Shiri J, Kisi O (2012). Pan evaporation modeling using neural computing approach for different climatic zones. Water Resour. Manage. 26(11):3231-3249.

Kisi O (2015). Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J. Hydrol. 528: 312-320

Kisi O & Heddam S (2019). Evaporation modelling by heuristic regression approaches using only temperature data. *Hydrol. Sci. J.* 64(6):653-672

Kotsiantis S & Pintelas P (2004). Combining bagging and boosting. *Int. J. Comput. Intell. Syst.* 1(4):324-333

Lu X, Ju Y, Wu L, Fan J, Zhang F & Li Z (2018). Daily pan evaporation modeling from local and cross-station data using three tree-based machine learning models. *J. Hydrol.* 566: 668-684

Majhi B, Naidu D, Mishra A P & Satapathy S C (2020). Improved prediction of daily pan evaporation using Deep-LSTM model. Neural Comput. Appl. pp. 1-16

Malik A, Saggi M K, Rehman S, Sajjad H, Inyurt S, Bhatia A S, Farooque A A, Oudah A Y & Yaseen Z M (2022). Deep learning versus gradient boosting machine for pan evaporation prediction, Eng. Appl. Comput. Fluid Mech. 16(1): 570-587

Malohlava M & Candel A (2018). Gradient boosting machine with H2O. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/gbm.html .Accessed 24 FEBRUARY 2020.

McCullagh P (1984). Generalized linear models. *Eur. J. Oper. Res.* 16(3): 285-292

Nelder J A & Baker R J (1972). Generalized Linear Models. Wiley Online Library, New Jersey.

Pregibon D& Hastie T J (2017). Generalized linear models. In: Statistical Models in S. Routledge.

Quinlan J R (1993). C4.5 programs for machine learning. Morgan Kaurmann, San Mateo, CA.

Rezaie-Balf M, Kisi O & Chua L H C (2019). Application of ensemble empirical mode decomposition based on machine learning methodologies in forecasting monthly pan evaporation. Hydrol. res. 50(2): 498-516

Schapire R (1990). The strength of weak learnability. *Journal of Machine learning* 5: 197-227

Seifi A & Soroush F (2020). Pan evaporation estimation and derivation of explicit optimized equations by novel hybrid meta-heuristic ANN based methods in different climates of Iran. Comput. Electron. Agric. 173: 105418

Semenov M A & Stratonovitch P (2010). Use of multi-model ensembles from global climate models for assessment of climate change impacts. Clim. Res. 41:1-14

Shiri J (2019). Evaluation of a neuro-fuzzy technique in estimating pan evaporation values in low-altitude locations. Meteorol. Appl. 26(2):204-212

Shaker Sureh F, Sattari M T, Rostamzadeh H & Kahya E (2024). Meteorological Drought Assessment and Prediction in Association with Combination of Atmospheric Circulations and Meteorological Parameters via Rule Based Models. *J Agr Sci-Tarim Bili.* 30(1):61-78. doi:10.15832/ankutbd.1067486

Tatsumi K, Oizumi T & Yamashiki Y (2013). Introduction of daily minimum and maximum temperature change signals in the Shikoku region using the statistical downscaling method by GCMs. Hydrol. Res. Lett. 7(3): 48-53

Terzi O (2010). Modeling of daily pan evaporation of lake Egirdir using data-driven techniques. International symposium on innovations in Intelligent systems and Applications. Istanbul. Turkey. pp. 320-324

Tezel G & Buyukyildiz M (2016). Monthly evaporation forecasting using artificial neural networks and support vector machines. Theor. Appl. Climatol. 124: 69-80

Vapnik V (1995). The Nature of Statistical Learning Theory. Springer, New York.

Vapnik V (1998). Statistical Learning Theory. Wiley, New York.

Wang L, Kisi O, Zounemat-Kermani M & Li H (2017a). Pan evaporation modeling using six different heuristic computing methods in different climates of China. *J. Hydrol.* 544: 407-427

Wang L, Niu Z, Kisi O, Li C & Yu D (2017b). Pan evaporation modeling using four different heuristic approaches. Comput. Electron. Agric. 140: 203-213

Wang H, Yan H, Zeng W, Lei G, Ao C & Zha Y (2020). A novel nonlinear Arps decline model with salp swarm algorithm for predicting pan evaporation in the arid and semi-arid regions of China. *J. Hydrol.* 582: 124545

Wang H, Sun F, Liu F, Wang T, Liu W & Feng Y (2023). Reconstruction of the pan evaporation based on meteorological factors with machine learning method over China. Agric. Water Manag. 287:108416.

Wilby R L, Dawson C W & Barrow E M (2002). SDSM a decision support tool for the assessment of regional climate change impacts. Environ. Model. Softw. 17:147-159

Wilby RL, Harris I (2006). A frame work for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. Water Resour. Res. https://doi.org/10.1029/2005wr004065

Wilby R L, Tomlinson O J, Dawson C W (2007). Multi-site simulation of precipitation by condition resampling. *J. Clim. Res.* 23: 183-194

Williams C K I, Barber D (1998). Bayesian classification with gaussian processes. IEEE Trans. Pattern Anal. Mach. Intell. 20(12): 1342-1351

Williams C K I, Rasmussen C E (1996). Gaussian processes for regression. In: Advances in Neural Information Processing Systems. MIT Press, pp. 514–520

Zarghami M, Abdi A, Babaeian I, Hassanzadeh Y İ & Kanani R (2011). Impacts of climate change on runoffs in East Azerbaijan, Iran. Glob Planet Change 78: 137-146

Zounemat-Kermani M, Kisi O, Piri J & Mahdavi-Meymand A (2019). Assessment of artificial intelligence–based models and metaheuristic algorithms in modeling evaporation. *J. Hydrol. Eng.* 24(10): 04019033. https://doi.org/10.1061/(ASCE)HE. 1943-5584.0001835.