Projection of Groundwater Level Fluctuations Using Different Machine Learning Algorithms under Climate Change in the Mashhad Aquifer, Iran

Ghasem Panahi
Ferdowsi University of Mashhad

Mahya Hassanzadeh Eskafi (mahya.eskafi@mail.um.ac.ir)
Ferdowsi University of Mashhad

Alireza Faridhosseini
Ferdowsi University of Mashhad

Saeed Reza Khodashenas
Ferdowsi University of Mashhad

Abbas Rohani
Ferdowsi University of Mashhad

Research Article

Keywords: Groundwater Level, CMIP6 Model, Climate Change, Machine Learning Algorithms

Posted Date: December 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2319553/v1

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

Additional Declarations: No competing interests reported.
Projection of Groundwater Level Fluctuations Using Different Machine Learning Algorithms under Climate Change in the Mashhad Aquifer, Iran

Ghasem Panahi¹, Mahya Hassanzadeh Eskafi²*, Alireza Faridhosseini³, Saeed Reza Khodashenas⁴, Abbas Rohani⁵

¹ Ph.D. student of Water Science and Engineering Department, Agriculture Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. ghasempanahi71@mail.um.ac.ir

² Master in Water Science and Engineering, Water Science and Engineering Department, Agriculture Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. mahya.eskafi@mail.um.ac.ir

³ Associate Professor of Water Science and Engineering Department, Agriculture Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. farid-h@ferdowsi.um.ac.ir

⁴ Professor of Water Science and Engineering Department, Agriculture Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. khodashenas@ferdowsi.um.ac.ir

⁵ Associate Professor of Biosystems Engineering, Agriculture Faculty, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. arohani@um.ac.ir

Corresponding Author:

Mahya Hassanzadeh Eskafi
Master in Water Science and Engineering
Tel: (+98) 933 802 2298
Email: mahya.eskafi@mail.um.ac.ir
Projection of Groundwater Level Fluctuations Using Different Machine Learning Algorithms under Climate Change in the Mashhad Aquifer, Iran

Ghasem Panahi1, Mahya Hassanzadeh Eskafi2*, Alireza Faridhosseini3, Saeed Reza Khodashenas4, Abbas Rohani5

Abstract

Due to population growth in recent years and climate change in arid and semi-arid regions, the lack of rainfall and the reduction of surface water flows required in various sectors, monitoring and projection of the climate change impact on the Groundwater Level (GWL) in the future is vital in the management and control of these resources. The purpose of this study is the projection of climate change impact on the GWL fluctuations in the Mashhad aquifer during the future period (2022-2064). In the first step, the climatic variables using ACCESS-CM2 under the Shared Socio-economic Pathways (SSPs) 5-8.5 scenario from the CMIP6 model were extracted. We used the CMhyd model to downscale the climatic data from the GCMs model. In the second step, different machine learning algorithms, including Multilayer Perceptron Neural Network (MLP), Adaptive Neuro-fuzzy Inference System Neutral Network (ANFIS), Radial Basis Function Neural Network (RBF), and Support Vector Machine (SVM) were used to predict the GWL fluctuations under climate change in the future period. Our results point out that temperatures and evaporation will increase in the autumn season, and precipitation will decrease by 26% in the future in the Mashhad aquifer. The results showed that the RBF model was an excellent performance in predicting GWL compared to other models. Based on the result of the RBF model, the GWL will decrease by 6.60 meters under the SSP5-8.5 scenario in the future. The findings of this research have a practical role in making helpful groundwater resources management decisions.

Keywords: Groundwater Level, CMIP6 Model, Climate Change, Machine Learning Algorithms

1. Introduction

Due to the increased population and the growing of urban and rural areas, amenities, and agricultural requirements, the amount of demand for water consumption has increased. It is provided through the surface and groundwater resources. Reduced rainfall during the hot seasons (Dai et al., 2020) and surface water quality are some of the reasons for water supply using natural resources. Therefore, using groundwater resources is a solution to reduce water tensions in different regions.

Due to the emission of greenhouse gases such as CO2 caused by human activities, including fossil fuels, the earth is warming (Montzka, Dlugokencky and Butler, 2011). Long-term changes in temperature and weather patterns that are caused by natural or unnatural factors are called climate change (United Nations). According to the IPCC report, the world temperature will increase by 1.5 centigrade by 2050, which will cause poverty and dangers in the lives of
more than 100 million people. Avoiding it requires a significant reduction in carbon dioxide emissions before 2030 (Masson-Delmotte et al., 2018). Based on the United Nations, climate change, directly and indirectly causes extensive changes in different regions. Its most critical direct effects are the global increase in temperature (Lionello and Scarascia, 2018; Arnell et al., 2019; Boukal et al., 2019) and extreme fluctuations in rainfall that cause drought in some regions (Berg and Sheffield, 2018; Cook, Mankin and Anchukaitis, 2018; Dai, Zhao and Chen, 2018; Goodarzi, Abedi-Koupai and Heidarpour, 2019; Kim and Jehanzaib, 2020; Philip et al., 2020). It causes severe floods in some regions (Hirabayashi et al., 2013; Arnell and Gosling, 2016; Hettiarachchi, Wasko and Sharma, 2018; Swain et al., 2020). Among the indirect effects of climate change, we can point out the economic consequences (Ciscar et al., 2011; Park et al., 2018; Allam and Jones, 2019; De Angelis, Di Giacomo and Vannoni, 2019; Piontek et al., 2021), social (Schwindt et al., 2016; Markkanen and Anger-Kraavi, 2019; Austin et al., 2020), changes in marine ecosystems (Blanchard et al., 2012) and eco-geomorphology of wetlands and beaches (Day et al., 2008), creating extensive changes in the physical characteristics of the soil (Trnka et al., 2013), changes in genetic plants (Alsos et al., 2012), changes in the forest ecosystem (Price et al., 2013) and plant species (Thuiller et al., 2011). Also, it changes the GWL by changing the aquifer’s recharge (Guevara-Ochoa, Medina-Sierra and Vives, 2020).

As mentioned, temperature changes and weather patterns play a significant role in the intensity and amount of rainfall. Considering that groundwater resources are formed from surface and subsurface water infiltration into the depths of the earth, changes in the intensity and amount of rainfall have a direct role in them. Since groundwater resources have a vital role in the water supply and increasing global warming, a projection of the climate change impacts on the GWL fluctuations is necessary. Investigating it is done through software and numerical modeling. Some of the studies in this field are mentioned below.

The HadCM3 model was used to investigate climate change impacts on the Umm er Radhuma unconfined aquifer in the western plains in Iraq from 2020 to 2099 (Hassan, 2020). The results showed that the average annual precipitation and temperature would increase under the A2 and B2 scenarios, and the annual groundwater recharge rate would decrease by 16%. The effect of climate change on surface and groundwater resources was investigated in the Yom and Nan basins in northern Thailand using the combination of SWAT and MODFLOW models and the CMIP5 model (Petpongpan, Ekkawatpanit and Kositgittiwong, 2020). Their results showed that under RCP 2.6 and RCP 8.5 scenarios, the average annual temperature will increase by 0.5 to 0.6 and 1 to 0.9 centigrade, respectively. The total surface and groundwater in the Yom basin will decrease. The total surface and groundwater will increase under the RCP 2.6 scenario and decrease under the RCP 8.5 scenario in the Nan basin. The effect of climate change on the groundwater resources in the Lake Tana basin in Ethiopia was studied (Tigabu et al., 2021). Their results indicated that the flow rate of groundwater resources would decrease compared to the base period. The research results regarding the effect of climate change, urbanization development, and sea level change on groundwater resources under A1B, B1, and A2 scenarios showed the GWL will decrease (Akbarpour and Niksokhan, 2018). The effect of climate change on the GWL was studied under RCP2.6, RCP4.5, and RCP8.5 scenarios in Shabestar Plain in Iran (Jeihouni et al., 2019). The study results regarding the effect of climate change on groundwater recharge using the CORDEX model under the RCP scenarios showed the amount of precipitation will decrease by 20%, and the evaporation will increase under the RCP 4.5 scenario by 8.1% in the Tekeze basin (Kahsay, Pingale and Hatiye,
Also, the amount of groundwater supply will decrease. The research results regarding the effect of climate change using the CMIP6 model and the SWAT model in Ho Chi Minh City in Vietnam showed an increase in the surface flow and groundwater recharge (Khoi et al., 2022).

Based on the reviews published by (Rajaee, Ebrahimi and Nourani, 2019; Ahmadi et al., 2022), Artificial Intelligence (AI) methods perform well in simulating and predicting GWL fluctuations in different aquifers. In the lack of hydrogeological information conditions, an Artificial Neural Network (ANN) is a helpful tool to predict the GWL (Lee, Lee and Yoon, 2019). In this regard, various types of research have been conducted to predict the GWL fluctuations using ANN, which can be referred to the research of (Poursaeid, Poursaeid and Shabanlou, 2022; Samani et al., 2022; Seidu et al., 2022; Singh and Panda, 2022). The effects of climate change was investigated under RCP2.6, RCP4.5, and RCP8.5 scenarios in Tasuj plain in Iran using ANN, Least Square Support Vector Machine (LSSVM), and Nonlinear Autoregressive Exogenous model (NARX) (Ghazi, Jeihouni and Kalantari, 2021). The results showed that the temperature would increase, and the amount of precipitation and the GWL would decrease. The effect of climate change on the groundwater resources in the Gaza Strip was investigated (Al-Najjar et al., 2021). Their results showed that the average annual precipitation will decrease by 5.2% and the average annual temperature will increase by 1 centigrade during 2020-2040. Also, the results of ANN modeling showed the GWL, which is currently between 0.38 and 18.5 meters compared to the surface of the earth, will reach 1.13 to 28 meters in 2040. (Jeihouni, Mohammadi and Ghazi, 2021) investigated the effects of climate change on the GWL in the Shabestar Plain under the RCP2.6, RCP4.5, and RCP8.5 scenarios during 2020-2050. Their results showed the average annual temperature will increase and the amount of rainfall will decrease. The deep learning method was used to project GWL in Germany until 2100 under climate change (Wunsch, Liesch and Broda, 2022). The results of using ANN neural network and NARX model indicated decreased in GWL. The results of GWL modeling in Ardabil plain in Iran using ANN, SVM, and ANFIS algorithms indicated the excellent performance of the ANFIS model (Seifi et al., 2020). The results of GWL projection using Linear Regression (LR), SVM, Gaussian Processes Regression (GPR), and Neural Network (NN) models showed the performance of GPR and LR models are the best (Sapitang et al., 2021).

Although according to the conducted research, climate change has a significant effect on the quality and quantity of groundwater resources, no research has been done to predict climate change under the IPCC Sixth Assessment Report (AR6) on the GWL fluctuations in the Mashhad aquifer. Since the Mashhad aquifer is one of the critical basins in Iran due to the passage of important rivers like the Kashafroud river and it has faced severe water resource crises such as floods and droughts in recent years, the projection of GWL fluctuations under climate changes in future periods plays an essential role in creating the attitude of water resources management in this region. The clear purposes of this research are as follows: 1) Using the ACCESS-CM2 model from the AR6 under the SSP5-8.5 scenario to predict climatic variables changes including maximum and minimum temperatures, precipitation and evaporation during the future period (2022-2064) compared to the historical period (1992-2021) 2) Projection of the GWL fluctuations in the Mashhad aquifer using different machine learning algorithms, including MLP, ANFIS, RBF, and SVM. The results of our research will play a significant role in the control and management of groundwater resources in this region.

2. Methods and Materials
2.1 Study area

The study area under investigation is the Mashhad aquifer, located in the northeast of Iran at the longitude of 58° 29' to 59° 56' east, and latitude of 35°58' to 37° 3' north. The Kashafroud river flows from northwest to southeast in the Mashhad plain. This river passes through the northern parts of Mashhad city. In this area, the average annual rainfall is 219.35 mm and has a dry and semi-arid climate. The maximum and minimum temperatures are +44 and -24 centigrade, respectively. Fig. 1 shows the geographic location of the Mashhad aquifer.

![Fig. 1 The Location of Mashhad Aquifer](image)

2.2 Groundwater Resources

Mashhad aquifer is an unconfined aquifer with an area of 3351 square kilometers. The thickness of its layer fluctuates by 130 meters. The morphology of the bedrock in this area is very uneven and has appeared on the alluvial surface. In some places, the thickness of alluvium reaches up to 300 meters. The characteristics of the Mashhad watershed are shown in Table 1.
Table 1 Characteristics of Mashhad Watershed

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watershed Area ($km^2$)</td>
<td>9909.4</td>
</tr>
<tr>
<td>Precipitation Average (mm)</td>
<td>247.5</td>
</tr>
<tr>
<td>Evapotranspiration Average (mm)</td>
<td>2300</td>
</tr>
<tr>
<td>Elevation Average (m)</td>
<td>1214.3</td>
</tr>
<tr>
<td>Minimum Elevation (m)</td>
<td>900</td>
</tr>
<tr>
<td>Maximum Elevation (m)</td>
<td>1600</td>
</tr>
</tbody>
</table>

According to the Khorasan Razavi Regional Water organization (www.khrw.ir), the number of groundwater sources in Mashhad plain is 7433, including springs, aqueducts, deep wells, and semi-deep wells. There are 6 thousand wells in which more than 4 billion cubic meters of water are extracted. The total annual discharge of water resources is 1116.27 million cubic meters, of which 850.30 million cubic meters are used for agriculture, 32.11 million cubic meters for industry, and 233.86 million cubic meters for drinking.

2.3 Water Balance

In the water resources balance cycle, the relationship between different balance factors, input and output sources has been. The inputs of water resources balance cycle include rainfall, surface and groundwater flows, and transfers. The outputs of this cycle include evaporation, exploitation of surface and groundwater resources, evaporation from lakes, surface, underground outflows and transfers. The annual rainfall volume in the Mashhad plain is 2704.97 million cubic meters and the annual evaporation volume from rainfall is 2246.77 million cubic meters. Therefore, effective rainfall is 458.2 million cubic meters. The amount of incoming and transfer surface flows is 99.36 million cubic meters. Therefore, the total volume of water produced in Mashhad plain is 557.56 million cubic meters. The total exploitation of surface and groundwater is 1124.78 million cubic meters, of which the net consumption volume is 607.32 million cubic meters. In this plain, the volume of surface and groundwater flow is 28.24 and 11.68 million cubic meters, respectively. The evaporation from the surface of the lakes is 11.3 million cubic meters. Therefore, the total resources extracted from the Mashhad plain is 658.54 million cubic meters. As a result, the water balance is 100.98 million cubic meters in the Mashhad plain.

2.4 Groundwater Fluctuations

Decreasing groundwater quality, moving salty water towards fresh areas and their salinization, and reversal of the direction of groundwater flow due to the infiltration of effluents and surface pollutants of rivers into groundwater
resources occurs due to the drop in the GWL and the reservoir deficit. The climatic variables and GWL fluctuations in the Mashhad aquifer during the historical period (1992-2021) can be seen in Fig. 2.
2.5 Methodology

According to the decreasing GWL fluctuations in the Mashhad aquifer mentioned in the previous section, the projection of GWL fluctuations under changing climatic patterns during the future period will play an essential role in improving water resources management. Based on the purpose of this research, different machine learning methods will be used to find the relationship between climate changes and the GWL fluctuations in the Mashhad aquifer for the historical period (1992-2021). Finally, the GWL fluctuations under climate change will be predicted for the future (2022-2064). The step of this research is shown in Fig. 3.
Fig. 3 Flow chart of the study

2.5.1 Datasets

The information of the piezometer wells on the Mashhad aquifer from 1992 to 2021, recorded by the Iran Water Resources Management organization (www.wrm.ir), was used. The daily climatic data, including maximum and minimum temperatures, precipitation, and evaporation of Mashhad plain during 1992 to 2021 were used by the Iran Meteorological Organization (www.irimo.ir). To predict the effect of climate change on the GWL in the Mashhad
aquifer in the future, the daily data of maximum and minimum temperatures and precipitation using the CMIP6 model (esgf-node.llnl.gov) were downloaded. In this research, the ACCESS-CM2 model was used under the SSP5-8.5 scenario. ACCESS-CM2 model specifications are shown in Table 2. Since selecting a suitable method for estimating evaporation depends on various factors, including the available meteorological data, and it is impossible to use a lysimeter in all conditions, the Torrent White method was used to calculate it. The monthly potential evaporation was calculated based on Eq. 1.

\[ ET = 16N_m \left( \frac{10T_m}{I} \right)^a \]  

\[ Im = \left( \frac{T_m}{5} \right)^{1.514} \]  

\[ I = \sum_{n=1}^{12} Im \]  

\[ A = (675 \times 10^{-9} \times I_3) \cdot (771 \times 10^{-7} \times I_2) + (179 \times 10^{-4} \times I) + 0.492 \]  

In Eq. 1, \( T_m \) is the monthly average temperature in centigrade and \( N_m \) is the correction factor. Based on Eq. 2, the thermal profile is calculated for each month. The annual thermal profile and coefficient \( A \) are calculated based on Equations 3 and 4, respectively.

### Table 2 Characteristics of the GCM model used

<table>
<thead>
<tr>
<th>Model name</th>
<th>Institution</th>
<th>Country</th>
<th>Grid Size (lon*lat)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS-CM2</td>
<td>Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology</td>
<td>Australia</td>
<td>144 * 192</td>
<td>(Bi et al., 2020)</td>
</tr>
</tbody>
</table>

#### 2.5.2 Downscaling

In this research, the CMhyd model (Climate Model data for hydrologic modeling) was used for downscaling climatic data from the GCMs model and correcting their bias. The CMhyd tool is written in Python 2.7 using NetCDF41, NumPy, and SciPy packages and in the PyQt42 program environment. Bias correction methods include linear scaling, delta change correction, precipitation local intensity scaling, power transformation of precipitation, variance scaling of temperature, and distribution mapping of precipitation and temperature. According to the research results by (Ringard, Seyler and Linguet, 2017; Switanek et al., 2017; Smitha et al., 2018; Enayati et al., 2021), the distribution mapping method has a good performance for bias correction. This technique was used to correct the accuracy of the simulations resulting from the micro scaling of GCMs models and to minimize the difference between the observed and simulated climate variables (Rathjens et al., 2016; Vaittinada Ayar, Vrac and Mailhot, 2021). In this study, after selecting the investigated climate variable in the CMhyd model environment and selecting the type of bias correction method, introducing the input path of the investigated climate variable, which were downloaded in netCDF
(*.nc file) format from the ACCESS-CM2, introducing daily observational climate data (1992-2021) in ASCII format, climate data were extracted in the form of text file for the future period.

2.5.3 Machine learning

Machine learning algorithms are a powerful tool in hydrological studies that can present anomalies that are not considered in conceptually-physical models (Rozos, Dimitriadis and Bellos, 2021). Since they have a good performance in simulating and predicting hydrological and hydrogeological processes (Dehghani, Poudeh and Izadi, 2022), different machine learning methods were used to predict the GWL fluctuations in the Mashhad aquifer under the GCMs model. The purpose of using machine learning methods is to train a process using datasets for data mining, image processing, and projection of time series as automatic. There are different methods to train the machine. Selecting the best method depends on the type of problem and the number of input parameters (Mahesh, 2020). It uses the human brain structure to process the information on the examined data. The appropriate number of neurons and hidden layers can be obtained using the machine’s performance in training and testing. In the first step, the pre-processing stage should be done to normalize the data. In such a way that the data is in the range of 0 and 1. In the present study, Equation 5 was used to normalize data (Ashtiani, Rohani and Aghkhani, 2020).

\[
X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \times (r_{max} - r_{min}) + r_{min}
\]  

In Eq. 5, \(X_n\) is the normalized variable value, \(X\) is the main value, \(X_{max}\) and \(X_{min}\) are the maximum and minimum values, \(r_{max}\) and \(r_{min}\) are the maximum and minimum ranges for normalizing the investigated variable. Two series of data should be created to perform the calibration and validation steps to ensure the efficiency of each machine learning method. The purpose of machine training is finding the most appropriate weight vector, bias vector, and minimize the error function. After the training step, the validation stage is performed. At this stage, the model is executed from another part of the data that was not used in the network training stage. In this step, the model outputs are compared with the observed values. Finally, the accuracy of the model is calculated. The training and validation steps are repeated until finding the best result from the model. Finally, the test stage is performed for the final control of the model. Machine learning algorithms used to use MATLAB software in this research include ANFIS, SVM, MLP, and RBF. The general structure of machine learning neural network methods used in the present study is shown in Fig. 4. In the following, a summary of the performance basis of different machine learning algorithms will be mentioned.
Multilayer Perceptron Neural Network

The structure of the MLP neural network consists of an input layer, an output layer, and at least one hidden layer that consists of hidden neurons. The number of neurons in the hidden layers is done using the optimal trial and error method and the network’s output is done using synaptic weights. Eq. 6 shows the performance of this algorithm. $W_{ji}$ is the weight in the hidden layer, $X_i$ is the input variable, $b_j$ is the bias for the hidden neuron $j$, and $y_i$ is the output variable:

$$y_i = f\left(\sum_{i=1}^{N} W_{ji}X_i + b_j\right)$$

Radial Basis Function Neural Network

The RBF neural network is a two-layer network with a single-layer architecture in which a linear mapping with Gaussian function is performed in the first layer and classification is performed in the second layer. The values of network weights are randomly selected between zero and one. Eq. 7 shows how to calculate the output of this method. In Eq. 7, $W_i$ is the weight of the edges, b is the bias and $f$ is the Gaussian activation function.

$$Y = \sum_{i=1}^{M} W_i f(X) + b$$

Fig. 4 Machine learning methods architecture of the present study
**Support Vector Machine**

Among the supervised learning methods, the SVM method can be mentioned. This method is one of the kernel methods in machine learning. The basis of this method is the linear classification of data. In this method, to categorize data with high complexity, the data is transferred to a high dimensional space by a function and uses Lagrange Duality Theorems. In general, the SVM is a model that fits a curve with the least error and with a certain thickness to the data. Equation 8 shows the basis of using this algorithm (Soltanali et al., 2021):

$$f(X) = \sum_{i=1}^{N} (a_i - a_i^*) K(X_i, X) + b$$  \hspace{1cm} (8)

In the above Equation, $a_i$ and $a_i^*$ are the Lagrange coefficients and $K(X_i, X)$ is the kernel function.

**Adaptive Neuro-fuzzy Inference System Neural Network**

An adaptive neuro-fuzzy inference system (ANFIS) that approximates real continuous functions has good performance in training, generation, and classification. This model was introduced by (Jang, 1993). Using this model, fuzzy rules can be extracted from numerical data. The structure of the ANFIS model consists of five layers including input nodes, rule nodes, intermediate nodes, subsequent nodes, and output nodes.

**Evaluation of Machine Learning Models**

To evaluate the performance of machine learning algorithms used (Li et al., 2021) in the simulation and projection of GWL fluctuations from the GCMs model under the SSP5-8.5 scenario in the Mashhad aquifer, different evaluation criteria including coefficient of determination ($R^2$) and root mean square error (RMSE) were used. The $R^2$ and RMSE values close to 1 and 0 indicate the high accuracy of the model to predict the studied variable. The evaluation criteria equations used in this research are shown in the following:

$$R^2 = 1 - \frac{\sum_{t=1}^{N}(GWL_{t-1} - GWL_t)^2}{\sum_{t=1}^{N}(GWL_{t-1} - \bar{GWL}_{t-1})^2}$$  \hspace{1cm} (9)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(GWL_{t-1} - \bar{GWL}_t)^2}{(N-1)}}$$  \hspace{1cm} (10)

In the above Equations, $GWL_{t-1}$, $GWL_t$, and $\bar{GWL}_t$ show the observed, predicted, and the average of the predicted values, respectively.
3. Results and Discussion

3.1 Results of downscaling

Fig. 4 shows the annual average changes of rainfall variables and minimum and maximum temperatures during 2022-2064 under the SSP5-8.5 scenario. Based on it, the precipitation area will have the highest amount in the northwestern and western regions in Iran. The southeastern regions in Iran will have the lowest amount of precipitation with an annual rainfall of 63 mm per year. The amount of precipitation in the investigated area located in the northeastern regions, is the average amount of it in Iran.

![Maps showing yearly changes in precipitation, minimum and maximum temperatures under the SSP5-8.5 scenario during 2022-2064.](image)

**Fig. 4** Yearly changes precipitation, minimum and maximum temperatures under the SSP5-8.5 scenario during 2022-2064

Based on Fig. 4, the annual average minimum temperature in Iran under the SSP5-8.5 scenario during the future period will vary between 4 and 28 centigrade. The maximum amount of the minimum temperature zone will start from the southern and southeastern regions and will extend to the southwestern and central regions. In the eastern regions, the minimum temperature will be average in Iran. The lowest area of this variable will belong to the northwestern regions of Iran. According to Fig. 4, the maximum temperature zone in the southern and southeastern regions in Iran will have the highest value at 35 centigrade, extending to the central and eastern regions. Its lowest amount will be in the northwestern regions around 12.5 centigrade. The results of climatic changes under the SSP5-8.5 scenario during the future period compared to the base period are shown in Fig. 5. According to Fig. 5a, the highest increasing percentage of changes in the maximum temperature will belong to October and December. Also, the biggest drop of it will occur in February equal to 12%. According to Fig. 5a, the highest and lowest percentage of minimum
temperature changes belong to January and February with 91% and 104%, respectively. Also, the minimum temperature will increase in the summer and autumn seasons.

According to Fig. 5b, the amount of precipitation will decrease in most months in the future compared to the historical period. The largest decrease for the summer in July and August is 73%, for the autumn in September with 37%, for the winter in January with 22%, and for the spring in May with 25%. Based on Fig 5.b, the highest and lowest amount of evaporation changes will happen in October and February equal to 36 and 34%, respectively. The increase

**Fig. 5** Monthly percentage changes climatic variables a) Minimum and Maximum Temperature, b) Rainfall and Evaporation under the SSP5-8.5 scenario during 2022–2064 compared to the historical period
in the percentage of evaporation changes for the future period in the winter is due to an increase in temperature and a decrease in precipitation. Fig 6 shows the monthly time series of precipitation, minimum and maximum temperatures, and evaporation variables in the future period (2022-2064) under the SSP5-8.5 scenario in the study area.
3.2 The performance of different machine learning algorithms

As mentioned, RBF, MLP, ANFIS, and SVM models were used to predict GWL. To use mentioned models to predict GWL, their parameters must be optimized. Fig. 7 shows the changes in some important parameters of the models in the training and testing stages. The results show, increasing the number of neurons in the hidden layer of the MLP and RBF neural networks leads to a decrease in the predicted error in the training and testing stages. The number of neurons in the hidden layer of the MLP and RBF neural networks were determined 7 and 33, respectively. Influence radius (IR) as one of the most important parameters of the ANFIS model is based on the subtractive clustering method to build a fuzzy inference system (FIS). Based on the RMSE results in training and testing steps, the value of the IR parameter was considered equal to 0.6. Also, the kernel function is one of the important parameters to design of the SVM model. Based on RMSE values in training and testing steps, the polynomial kernel degree 2 types was chosen as the best. Other model parameters were determined in the same way.

After calculating the optimal values of the parameters of the models for GWL, the results of their predictions are shown in Fig 8. Although they show a perfect agreement between the actual and predicted GWL values in the training step for all models ($R^2=0.99$), the predicted error is much more critical in the testing step.
Train: $y = 0.99x + 0.11$, $R^2 = 0.99$, RMSE = 0.05
Test: $y = 0.99x + 0.15$, $R^2 = 0.99$, RMSE = 0.06

Train: $y = 0.99x + 3.00$, $R^2 = 0.99$, RMSE = 0.35
Test: $y = 0.97x + 35.0$, $R^2 = 0.94$, RMSE = 1.07

Train: $y = 0.99x + 0.29$, $R^2 = 0.99$, RMSE = 0.13
Test: $y = 0.99x + 4.86$, $R^2 = 0.95$, RMSE = 0.82
Fig. 8 The evaluation of the agreement between the actual and predicted values using a) RBF, b) MLP, c) ANFIS, and d) SVM algorithms.

According to Fig. 8, RBF, ANFIS, MLP, and SVM models have the best performance in the test stage, respectively. The reason for it is the highest value of $R^2$ and the lowest value of the width from the origin. Also, the slope value is close to 1 for the regression line between actual and predicted GWL. Fig. 9 shows the results of GWL fluctuations using different machine learning algorithms under the SSP5-8.5 scenario in the future. According to Fig. 9, the GWL using RBF, ANFIS, MLP, and SVM models will decrease by 6.60, 3.32, 6.59, and 5.18 meters, respectively.

![Fig. 8](image_url)

Fig. 9 Monthly predicted GWL during 2022-2064 using different Machine Learning Algorithms.

Fig. 10 shows the monthly predicted GWL in the future period by RBF, MLP, ANFIS, and SVM compared to the historical period (Observation). Based on the results of the best model to predict GWL (RBF model), it will decrease

---

Figure captions:

- **Fig. 8**: The evaluation of the agreement between the actual and predicted values using different algorithms. The values include training and testing results with their respective $R^2$ and RMSE.
- **Fig. 9**: Monthly predicted GWL during 2022-2064 using different Machine Learning Algorithms. The graph shows the predicted GWL for each algorithm and the historical GWL (Observation) for comparison.
- **Fig. 10**: Monthly predicted GWL in the future period for RBF, MLP, ANFIS, and SVM algorithms compared to the historical period (Observation). The graph includes a linear regression line with the equation $y = 0.99x + 0.29$ for the training phase and $y = 0.85x + 4.86$ for the test phase, along with their respective $R^2$ and RMSE values.
in all seasons in the future under the SSP5-8.5 scenario.

The most significant decrease will occur in the summer and autumn seasons. It is the consequence of the decrease in precipitation and the increase in temperature and evaporation in the summer and autumn seasons under the SSP5-8.5 scenario based on Fig. 10. Therefore, the decrease in the GWL in the future period can be related to the increase in temperature and evaporation. Our result is consistent with the results of (Chang et al., 2015; Arkoç, 2022; Wunsch et al., 2022).

4. Conclusions

According to the results of predicted climate variables in the Mashhad aquifer using the CMIP6 under the SSP5-8.5 scenario, the highest increasing minimum and maximum temperatures changes in the future period (2022-2064) will belong to October by 2.50 and 3.16 degrees Celsius, respectively. Moreover, the amount of precipitation will decrease by 26% in the future. The amount of evaporation will increase in the autumn season. The findings of this research indicated that the most amounts of precipitation will be unavailable due to the increase in temperature and evaporation in the future, which will hurt the feeding rate in the Mashhad aquifer. The results of various machine learning algorithms, including RBF, MLP, ANFIS, and SVM models to predict the effects of climate changes on the GWL of the Mashhad aquifer showed that the RBF method has the best performance in projecting the GWL fluctuations compared to the other algorithms. Also, the result of GWL predicted under climate change showed it will decrease by 6.60 meters in this region. Since the changes in weather patterns in the long term lead to the reduction of
the GWL, it causes an increase in desert areas. Reducing these resources for agricultural, industrial, and domestic purposes will face a severe crisis. Therefore, the improved management of the Mashhad aquifer is vital to control the water demand and supply. It is suggested to further investigation regarding using underground dams to reduce evaporation and the relationship between the drought index and the salinity level of the Mashhad aquifer in the future periods.

References


water demands and its resulting consequences on groundwater using CMIP5 models’, *Groundwater*. Wiley Online Library, 57(2), pp. 259–268.


Statements and Declarations:

Funding:

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

Competing Interests:

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

Author Contributions:

Ghasem Panahi: writing—original draft preparation, conceptualization, visualization, data acquisition

Mahya Hassanzadeh Eskafi: methodology, writing—original draft preparation, data acquisition, supervision

Alireza Faridhosseini: conceptualization, supervision, editing of the manuscript

Saeed Reza Khodashenas: conceptualization, supervision, editing of the manuscript

Abbas Rohani: methodology
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Highlights.docx