Title
Estimating the riverine environmental water demand under climate change with data mining models

Authors
Zanjani, Masoud
Bozorg-Haddad, Omid
Zanjani, Mustafa
et al.

Publication Date
2024

DOI
10.1007/s11069-024-06656-4

Peer reviewed

Natural Hazards
https://doi.org/10.1007/s11069-024-06656-4

ORIGINAL PAPER

Estimating the riverine environmental water demand under climate change with data mining models

Masoud Zanjani[1] · Omid Bozorg-Haddad[2] · Mustafa Zanjani[1] · Ali Arefinia[2] · Masoud Pourgholam-Amiji[3] · Hugo A. Loáiciga[4]

Received: 5 June 2023 / Accepted: 30 April 2024

Abstract
This paper presents a statistical approach based on data mining to estimate the riverine environmental water demand (EWD). A river's environmental water demand defines the quantity, timing, and quality of streamflow that are required to sustain riverine ecosystems and human activities. Genetic programming (GP), artificial neural network (ANN), and support vector regression (SVR) are herein applied to model the environmental demand. Input and output data for the use of GP, ANN, and SVR are the average monthly temperature and precipitation in 1995–2005 plus climate projections by the Canadian Land System Model (CanESM2) under the recommended concentration pathways RCPs 2.6, 4.5 and 8.5 in 2025–2035. A case study illustrates this paper's methodology using temperature and precipitation data and monthly discharge of the Karaj River, Iran. The applied data mining models were evaluated with R^2, *RMSE*, and the *NSE* criteria. This work's results show that the largest values of R^2 and the *NSE* equal respectively 0.94 and 0.95, and the smallest value of the *RMSE* equals 0.07, which correspond to the SVR predictions. These results establish that SVR is a suitable model for the purpose of estimating the environmental water demand in comparison to GP and ANN in the study area. The SVR projections indicate that by 2035 and under the RCPs 2.6, 4.5, and 8.5 projected changes of the environmental water demand with respect to baseline conditions would be respectively 63, 118, and 126 m^3/s. It is demonstrated in this work that under climate change conditions the correlation between the EWD index and temperature was 83%, while the said value for rainfall was estimated to be 76%.

Keywords Water demand · Support vector regression · Genetic programming · Artificial neural network · SDSM · CanEMS2 · Karaj River

1 Introduction

Greenhouse gas emissions since the Industrial Revolution have raised the earth's surface air temperature and have had multiple adverse social and environmental impacts (Abbass et al. 2022; Mozaffari 2022; Malla et al. 2022). Climate data demonstrate that the Earth's average temperature increased by 0.6 °C in the last century due to a rise in greenhouse gases

Extended author information available on the last page of the article

Published online: 18 May 2024

concentrations in its atmosphere (Bozorg-Haddad et al. 2020a, b). The atmospheric warming has multiple effects on the hydrologic cycle, including modification of river runoff (Sarzaeim et al. 2017a, b).

Rivers are a valuable water resource of primary importance for water supply and ecosystem conservation (Hailu et al. 2018; Than et al. 2021). Therefore, determining the environmental water demand of rivers downstream of reservoirs whose operation affects river streamflow is of central importance to river management (Mengistu et al. 2021; Sedighkia et al. 2021; Ghoreishi et al. 2021). Four primary methods to estimate the environmental water demand were identified by Tharme's 2003 worldwide assessment of 207 environmental demand studies conducted in 44 countries, which involved the use of hydrological, hydraulic rating, habitat modeling, and holistic methods. Each method has unique data challenges. The straightforward hydrological methods depend on flow time series, and they are the most cost-effective. Comparatively, holistic methods, which are more data intensive, do not share widely accepted procedures (Sarzaeim et al. 2017a, b). Hydraulic rating methods are intended for single-channel rivers (Linnansaari et al. 2012) and they require field research data. Habitat modeling methods are relatively accurate; they require detailed biological data and are costly and time-consuming to acquire. The holistic methods cover a variety of river flow factors; they are typically expensive and time-consuming because they require detailed data and expert judgment (Linnansaari et al. 2012). The Tennant method (Tennant 1976), the Tessman method (Tessman 1980), RVA (Richter et al. 1997), BBM (King and Louw 1998), and DRIFT (King et al. 2003) are examples of commonly used methods to estimate the riverine environmental water demand (EWD). Estimating the riverine environmental water demand in the context of climate change is a complex process that is computationally burdensome and requires the analysis of mega data (Khelifa et al. 2021; Tian et al. 2021; Wang et al. 2021; Ahmadzadeh et al. 2022). The data and modeling challenges involved in the estimation of the environmental water demand (EWD) under climate change dfsinvolve the use of computationally efficient methods with advanced pattern recognition capabilities (Oliazadeh et al. 2021, 2022; Arefinia et al. 2022a).

Data mining is the application of computerized algorithms and process to discover non-obvious and useful information from large datasets (Kelleher and Tierney 2018; Li et al. 2019; Tan et al. 2019; Abualigah and Dulaimi 2021; Arefinia et al. 2021a, b), which is used in many studies across various scientific fields (Akbari-Alashti et al. 2014; Beygi et al. 2014; Bozorg-Haddad and Mariño 2011, 2007; Bozorg-Haddad et al. 2009a, b, 2010a, 2013, 2015, 2016, 2017a; b; Fallah-Mehdipour et al. 2011, 2013a, b; Karimi-Hosseini et al. 2011; Orouji et al. 2014; Sabbaghpour et al. 2012; Soltanjalili et al. 2011). Applications of data mining in water resources include water-quality simulation (Bozorg-Haddad et al. 2017a, b; Arefinia et al. 2020), runoff simulation (Sarzaeim et al. 2017a, b; Al-Juboori 2022), and virtual water forecasting (Arefinia et al. 2022b). This study is a novel application because data mining models have not been used to evaluate the riverine EWD before. This work applies three types of data mining algorithms (i.e., GP, ANN and SVR) to estimate the riverine EWD under climate change conditions. The performances of the three algorithms are herein evaluated.

2 Methodology

The method presented in this paper consists of four parts: (1) temperature and rainfall projections in the period 2025–2035 under three RCPs, (2) projection of runoff scenarios in the same period, (3) estimation of the EWD in the baseline period (1995–2005), and (4) application of data mining models to estimate the riverine EWD under climate-change conditions. The method's flow diagram is depicted in Fig. 1. Figure 1 shows that the estimation of the EWD begins with the prediction of surface air temperature and rainfall in the baseline period, followed by the application of data-mining models to estimate the EWD under climate-change conditions.

2.1 Projecting climate-change temperature and rainfall

The projection of future surface air temperature and rainfall over a study region relies on the application of the CanESM2 general circulation model (GCM), which has been evaluated and used in numerous studies in the Middle East and found to have accurate predictive skill (Sarzaeim et al. 2017a, b; Zolghadr-Asli et al. 2019; Arefinia et al. 2022a, b). The CanESM2 is implemented under the RCP2.6, 4.5, and 8.5 (RCP6's data did not exist for the study region). The GCM outputs are made at coarse spatial scale and are downscaled to the regional scale for the purpose of simulating hydrologic processes in the study region. The Statistical Downscaling Model (SDSM, Wilby and Dawson 2013) is a decision support tool that is applied in this work to downscale the CanESM2 outputs. The SDSM uses public-domain climate projections for downscaling purposes. The downscaling of GCM outputs proceeds in two steps. First, regression functions are obtained between model-projected and observed surface air temperature and rainfall temperature corresponding to the baseline period. Second, these regressions are applied to obtain downscaled surface air temperature and rainfall in a future period (2025–2035) representative of climate-change conditions based on the CanESM2 projections of these variables.

The CanESM2 model's climate projections were obtained from the model site (http://climate-scenarios.canada.ca). The downscaled surface air temperature and rainfall projections required to implement the data-mining models under the RCPs 2.6, 4.5 and 8.5 climate change conditions were downscaled for the period 2025–2035 using the SDSM 4.2.9. The monthly surface air temperature and rainfall were applied to calculate the riverine EWD.

2.2 Data mining models

This work applies artificial neural networks (ANNs), genetic programming (GP), and support vector regression (SVR) to estimate the EWD under climate change. These data-mining models are briefly described next.

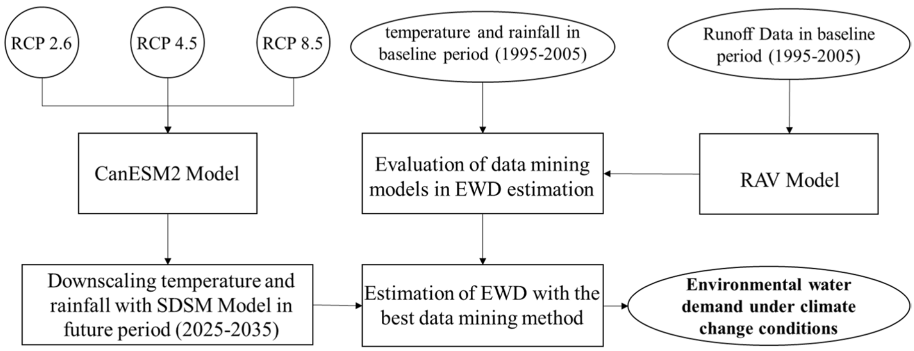

Fig. 1 Flow diagram of this paper's methodology

2.2.1 Genetic programming (GP)

GP was introduced by Cramer (1985). Koza (1992, 1994) developed GP into a practical tool. It is a member of the family of evolutionary algorithms (Arefinia et al. 2021a). GP is applied to find functional relations between variables that are then used to predict dependent variables from independent variables. This is accomplished by generating set of numbers and operators that are treated as decision variables and are improved iteratively until convergence to desired functional relations is achieved. The reader is referred to Mohammad-Azari et al. (2020) for a review of applications of GP in water resources systems analysis.

2.2.2 Artificial neural networks (ANNs)

An ANN is a data-intensive model that discovers functional relations between output and input variables through an algorithmic process called machine learning or data mining. The best functional relations that the ANN can find relating input variables to output variables might be complex and non-linear. An ANN is implemented by first training its input and output data. This is followed by a testing phase whereby the trained ANN's predictive skill is tested with new input and output data. The ANN is used in prediction if it performs well during the testing phase. Several goodness-of-fit criteria are employed to train and test ANNs. The reader is referred to Kelleher and Tierney (2018) and Tan et al. (2019) for a review of ANNs, data mining, and data science, in general.

2.2.3 Support vector regression (SVR)

Vapnik (1995) introduced support vector machine (SVM). Support vector regression (SVR) is a regression variant of SVM. Two functions define SVR. The first function calculates the errors of the values calculated by SVR [see Eq. (1)]:

$$|y - f(x)| = \begin{cases} 0 if |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon = \xi \text{otherwise} \end{cases} \tag{1}$$

in which y= the vector of observed (output values), $f(x)$= the SVR prediction function for the input x; ε = the sensitivity vector of the function $f(x)$, ξ = the penalty vector, w= the weight vector.

The second function predicts the SVR output [see Eq. (2)]:

$$f(x) = w^{\mathrm{T}} \cdot x + b \tag{2}$$

in which b= the deviation of the function $f(x)$, and T denotes the transpose operation.

SVR constitutes an optimization model in the form of Eqs. (3), (4) and (5) (Arefinia et al. 2020; 2021a):

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{m} (\xi_i^- + \xi_i^+) \tag{3}$$

Subjected to:

$$(w^{\mathrm{T}} \cdot x + b) - y_i < \varepsilon + \xi_i^+ \quad i = 1, 2, 3, \ldots, m \tag{4}$$

$$qy_i - (w^{\mathrm{T}} \cdot x + b) \le \varepsilon + \xi_i^- \quad i = 1, 2, 3, \dots, m \tag{5}$$

where C = the penalty coefficient, m = the number of training data, ξ_i^+ and ξ_i^- = the magnitude of the penalties above and below the band $(-\varepsilon, +\varepsilon)$, respectively, and y_i = the i-th observed data value. Problem (3), (4) and (5) is solved by Lagrange optimization for b and w. The output value of the SVR model is then obtained with Eq. (2).

SVR can simulate data exhibiting nonlinear behavior. In this case transfer functions, also referred to as kernel functions, are applied to linearize the data (see Bozorg-Haddad et al. 2020a, b).

2.3 Estimation of the EWD using data mining models

The first step in using data mining tools to estimate the EWD under climate change is to build a database of average surface air temperature and monthly rainfall in a baseline period (1995–2005) to train a model for estimating the EWD.

The range of variability approach (RVA) (Richter et al. 1996, 1997, 1998) is herein applied for the estimation of the EWD. The RVA sets streamflow-based river ecosystem management targets. It uses as its starting point either measured or synthesized daily streamflow values from a period during which human perturbations to the hydrological regime were negligible. This streamflow record is then characterized using thirty-two different hydrologic parameters (Richter et al. 1997). The hydrologic parameters serve as flow targets that guide riverine management. According to the RVA method, the specific environmental water demand range for each month in year k has been obtained by using the first quartile of the river flow in the same month during the 10 years before year k.

Therefore, the EWD in the base period is given by Eq. (6):

$$\mathrm{EWD} = Q_1\left(Q_{k-10}, Q_{k-1}\right) k = 1, 2, \dots \tag{6}$$

where EWD = environmental water demand in year n, $Q_1\left(Q_{k-10}, Q_{k-1}\right)$ = the first quartile of the river discharge for a particular month in the years $k-10$ through $k-1$.

The analysis of the correlation between the EWD and independent variable is essential to obtaining a proper model structure (Gil et al. 2007). The correlation between the EWD and the average monthly surface air temperature and rainfall was found to be significant. The results indicate a strong statistical association (above 70%) between the EWD at time t and the average monthly surface air temperature and rainfall in periods t through t − 3.

Equation (7) describes the function between the output parameter at time t ($(\mathrm{EWD})_{\mathrm{t}}$) and the input parameters at time t through t − 3.

$$(\mathrm{EWD})_{\mathrm{t}} = f\left((T, R)_{\mathrm{t}}, \dots, (T, R)_{\mathrm{t}-3}\right) \tag{7}$$

where $(\mathrm{EWD})_{\mathrm{t}}$ = the EWD at time t, $(T)_{\mathrm{t}}$ = temperature at time t, $(R)_{\mathrm{t}}$ = rainfall at time t.

In the second step all information was normalized according to Eq. (8) to bring the input variables to a common range of values:

$$x_{\mathrm{norm}} = \frac{x - x_{\mathrm{min}}}{x_{\mathrm{max}} - x_{\mathrm{min}}} \tag{8}$$

where x_{norm} = normalized x value, x = value of input variable before normalization, x_{min} and x_{max} = the minimum and maximum values of the parameter before normalization, respectively.

In the last step The GP, ANN, and SVM were trained after creating the database and normalizing the data. 75% of the data were used for training, and 25% of the data were used to test the data-mining model. All models are programmed using MATLAB.

The correlation coefficient (R^2), Nash–Sutcliffe efficiency (*NSE*), and the root mean square error (*RMSE*) were used as performance indicators for each model. The closer the R^2 and *NSE* value to 1, the better the fit between predictions and observations, and the closer the *RMSE* value to zero, the better the fit between predictions and observations. The equations for the R^2, *NSE*, and *RMSE* are as follows:

$$R^2 = \frac{\sum_{t=1}^{n}\left(\mathrm{EWD_{sim}} - \overline{\mathrm{EWD}}_{\mathrm{sim}}\right)^2 \times \left(\mathrm{EWD_{obs}} - \overline{\mathrm{EWD}}_{\mathrm{obs}}\right)^2}{\sum_{t=1}^{n}\left(\mathrm{EWD_{sim}} - \overline{\mathrm{EWD}}_{\mathrm{sim}}\right)^2 \times \sum_{t=1}^{n}\left(\mathrm{EWD_{obs}} - \overline{\mathrm{EWD}}_{\mathrm{obs}}\right)^2} \tag{9}$$

$$NSE = 1 - \frac{\sum_{t=1}^{n}\left(\mathrm{EWD_{obs}} - \mathrm{EWD_{sim}}\right)^2}{\sum_{t=1}^{n}\left(\mathrm{EWD_{obs}} - \mathrm{EWD_{avg}}\right)^2} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\mathrm{EWD_{obs}} - \mathrm{EWD_{sim}}\right)^2}{n}} \tag{11}$$

where n = number of data, $\mathrm{EDW_{obs}}$ = observed EWD, $\overline{\mathrm{EWD}}_{\mathrm{obs}}$ = mean value of observed EWD, $\mathrm{EWD_{sim}}$ = calculated EWD.

2.4 Case study

The Karaj River Basin in northern Iran is located between longitudes 51.05 and 51.35 degrees east and between latitudes 35.85 and 36.24 degrees north. The general topographic slope decreases from north to south. The total area of the basin equals to 845 square kilometers with a minimum and maximum average temperature equal to 9.86 and 21.96 °C, respectively. Table 1 lists the calculated average minimum monthly temperature, the average monthly temperature, the average maximum monthly temperature, and the average monthly rainfall.

The Karaj River provides water for agriculture, to recharge aquifers, and it is a source of water for Tehran, Iran's capital (Arefinia et al. 2020). Meeting Karaj River's EWD is imperative to sustain tourism, and the economic and environmental sectors in its drainage basin.

3 Results and discussion

Three climatic scenarios for the average monthly surface air temperature (*T*) and rainfall (*R*) in the study area were projected with the CanESM2 GCM under the RCPs 2.6, 4.5, and 8.5, and those projections were downscaled with the SDSM for the period 2025–2035 (120 months). The results for *R* and *T* are displayed in Figs. 2 and 3, respectively.

Table 1 Average Monthly temperature and rainfall in the Karaj basin corresponding to the baseline period (1995–2005)

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Min temperature (°C)	−2.5	−0.37	4.8	9.61	15.14	19.63	22.05	20.68	16.26	10.31	3.76	−1.03
Mean temperature (°C)	1.46	4.6	10.9	15.85	21.41	27.05	28.89	29.07	23.92	17.33	9.31	3.72
Max temperature (°C)	7.34	8.31	14.5	22.3	27.4	34.02	36.33	34.51	30.35	25.05	14.95	8.49
R (mm)	28.33	25.36	33.85	32.56	18.14	3.83	5.29	2.58	3.67	9.85	20.81	26.92

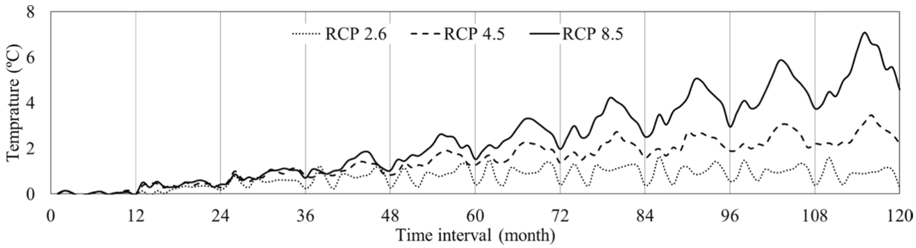

Fig. 2 The anomaly of the average monthly surface air temperature (*T*) downscaled with the SDSM method in the period 2025–2035 under the RCPs 2.6, 4.5 and 8.5

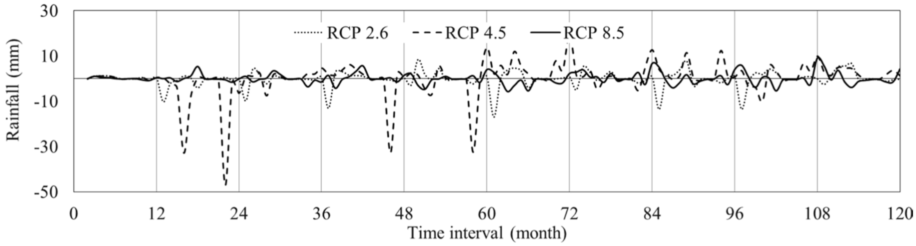

Fig. 3 The anomaly of the average monthly rainfall *(R)* downscaled with the SDSM method in the period 2025–2035 under RCPs 2.6, 4.5 and 8.5

It is seen in Fig. 2 that the average monthly *T* based on RCPs 2.6, 4.5, and 8.5 increases steadily; the results show that the rise in average monthly *T* based on the RCPs 8.5 is larger than the RCP 4.5's, and the RCP4.5's temperature rise is larger than the RCP 2.6's. The maximum *T* rise based on the RCPs 2.6, 4.5, and 8.5 equal 1.60, 3.47, and 7.08 °C, respectively. The largest *T* increase is for May, June, and July, and the lowest *T* increase is for October–November–December, which establishes that the climate change on temperature in the study area is more pronounced in the warmer months.

Figure 3 shows that the projected changes in monthly *R* in the study area do not follow a specific trend. The monthly *R* increases in some years and declines in others, whereas the results show that the extreme daily rainfall events increase.

The EWD was projected with GP, ANN, and SVR using monthly *T* and *R* data in the baseline period (1995–2005 The primary objective of this study is to estimate the environmental water demand of the river in the coming years. Models trained on historical data are used to project potential scenarios for the future based on learned patterns. The results are presented in Figs. 4 and 5. The SVR model's projections are closer to the observational values by (Figs. 4, 5c) than those of the GP model (Figs. 4, 5a), and ANN (Figs. 4, 5b), which establishes the better performance of SVR in estimating the EWD in the study area. A summary of the performance results for GP, ANN, and SVR based on the R^2, *RMSE*, and *NSE* criteria is presented in Table 2.

The data show listed in Table 2 show the GP, ANN, and SVR models' performance metrics (R^2, *RMSE*, and *NSE*) during training and testing. The results validate the conclusion that SVR performs consistently better across a range of assessment criteria, showing improved accuracy when compared to GP and ANN.

The superiority of SVR over the other data mining models used in simulating the EWD in the baseline period was proved. SVR was applied next to project the EWD under climate

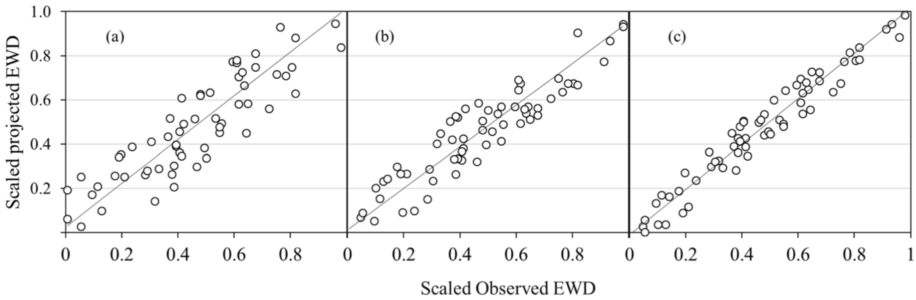

Fig. 4 Scaled EWD scatter diagrams for **a** GP, **b** ANN and **c** SVR

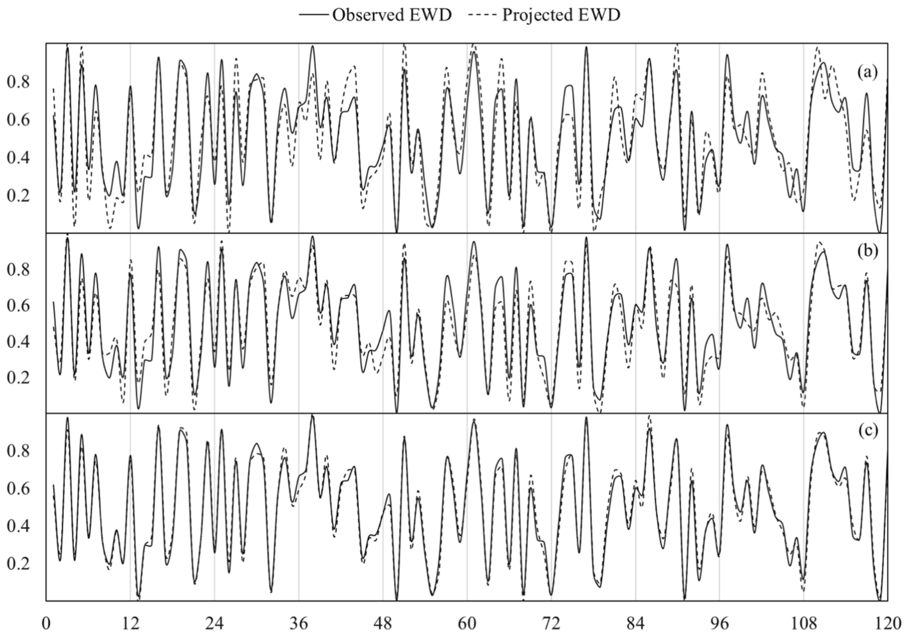

Fig. 5 Comparison of the scaled observed and projected EWD with **a** GP, **b** ANN, and **c** SVR in the period 1995–2005

Table 2 Performance results for *EWD* prediction with data-mining algorithms to the baseline period (1995–2005)

Model	R^2		*RMSE* (m^3/s)		*NSE*	
	Testing	Training	Testing	Training	Testing	Training
GP	0.74	0.83	0.19	0.14	0.62	0.66
ANN	0.86	0.89	0.09	0.08	0.90	0.93
SVR	0.91	0.94	0.08	0.07	0.91	0.95

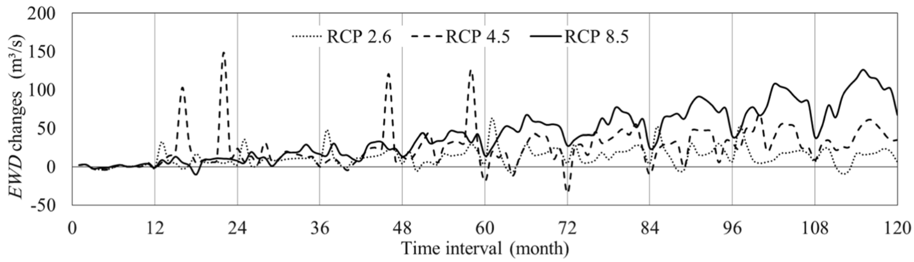

Fig. 6 Comparison of the EWD changes with respect to the baseline period projected with SVR in the period 2025–2035 under RCPs 2.6, 4.5 and 8.5

change conditions (RCPs 2.6, 4.5, and 8.5). The results are displayed in Fig. 6. It is seen in Fig. 6 that the projected change in the EWD with respect to the baseline period exhibits a steady upward trend that is most pronounced under the RCP 8.5. The RCP 8.5-based projections indicate the maximum and minimum changes in values of the EWD equal 126 and −10 m^3/s, respectively. The maximum and minimum changes in values of the EWD according to the RCP4.5 equal 118 and −34 m^3/s, respectively. The projected maximum and minimum changes in the EWD associated with the RCP 2.6 equal respectively 63 and −9 m^3/s. By comparing the results depicted in Fig. 6 with those displayed in Fig. 2 and 3 reveals patterns of statistical association between the EWD and the temperature and rainfall. Specifically, this work's results show that, on average, under climate change conditions the correlation between the EWD index and variable temperature was 83%, while the said value for variable rainfall was estimated to be 76%.

4 Conclusion

This work estimated the Environmental water demand (EWD) changes in the Karaj River under climate-change conditions based on temperature and rainfall projections for the period 2025 through 2035 corresponding to the RCPs 2.6, 4.5, and 8.5.

GP, ANN, and SVR were implemented to predict the EWD in the study area, and their performances were compared based on the R^2, *RMSE*, and *NSE*. It was concluded that SVR ($R^2 = 0.94$, $RMSE = 0.07$(m^3/s), and $NSE = 0.95$) performed better than GP and ANN.

SVR was applied to project the EWD under climate-change conditions. The results establish that the change in the EWD in 2025–2035 exhibits a projected steady upward trend that is more pronounced under the RCP 8.5. The SVR projections by 2035 corresponding to the RCPs 2.6, 4.5, and 8.5 show changes equal to 63, 118, and 126 m^3/s, respectively, in the riverine environmental water demand relative to the baseline period. Also, the results indicate a strong correlation (83%) between the Environmental Water Demand (EWD) index and temperature, emphasizing the significant impact of temperature on EWD. Additionally, the correlation between the EWD index and rainfall is estimated to be 76%, highlighting the influence of rainfall on environmental water demand. These findings underscore the sensitivity of EWD to both temperature and rainfall under climate change conditions.

Acknowledgements The authors thank Iran's National Science Foundation (INSF) for its support for this research.

Author contributions Masoud Zanjani: software, formal analysis, writing—original draft. Omid Bozorg-Haddad: conceptualization, supervision, project administration. Mustafa Zanjani: software, formal analysis, writing—original draft. Ali Arefinia: software, formal analysis, writing—original draft. Masoud Pourgholam-Amiji: software, formal analysis, writing—original draft. Hugo A. Loáiciga: validation, writing—review and editing.

Funding No funding was received for conducting this study specifically.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability The codes that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest There is no conflict of interest.

Consent to participate All authors consent to participate.

Consent for publish All authors consent to publish.

Ethical approval All authors accept all ethical approvals.

References

Abbass K, Qasim MZ, Song H, Murshed M, Mahmood H, Younis I (2022) A review of the global climate change impacts, adaptation, and sustainable mitigation measures. Environ Sci Pollut Res 29:1–21

Abualigah L, Dulaimi AJ (2021) A novel feature selection method for data mining tasks using hybrid sine cosine algorithm and genetic algorithm. Clust Comput 24(3):2161–2176

Ahmadzadeh H, Mansouri B, Fathian F, Vaheddoost B (2022) Assessment of water demand reliability using SWAT and RIBASIM models with respect to climate change and operational water projects. Agric Water Manag 261:107377

Akbari-Alashti H, Bozorg-Haddad O, Fallah-Mehdipour E, Mariño MA (2014) Multi-reservoir real-time operation rules: a new genetic programming approach. Proc Inst Civ Eng Water Manag 167(10):561–576. https://doi.org/10.1680/wama.13.00021

Al-Juboori AM (2022) Solving complex *r*-runoff processes in semi-arid regions using hybrid heuristic model. Water Resour Manage 36(2):717–728

Arefinia A, Bozorg-Haddad O, Oliazadeh A, Loáiciga HA (2020) Reservoir water quality simulation with data mining models. Environ Monit Assess 192(7):1–13

Arefinia A, Bozorg-Haddad O, Keller AA (2021b) "Risk, uncertainty, and reliability in water resources management" essential tools for water resources analysis, planning, and management. Springer Nature Singapore, Singapore, pp 323–343

Arefinia A, Bozorg-Haddad O, Chang H (2021a) "The role of data mining in water resources management" essential tools for water resources analysis, planning, and management. Springer Nature Singapore, Singapore, pp 85–99

Arefinia A, Bozorg-Haddad O, Ahmadaali K, Bazrafshan J, Zolghadr-Asli B, Chu X (2022a) Estimation of geographical variations in virtual water content and crop yield under climate change: comparison of three data mining approaches. Environ Dev Sustain 24(6):8378–8396

Arefinia A, Bozorg-Haddad O, Akhavan M, Baghbani R, Heidary A, Zolghadr-Asli B, Chang H (2022a) Using support vector machine (SVM) in modeling water resources systems. In: Computational intelligence for water and environmental sciences, pp 179-199

Beygi S, Bozorg-Haddad O, Fallah-Mehdipour E, Mariño MA (2014) Bargaining models for optimal design of water distribution networks. J Water Resour Plan Manag 140(1):92–99. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000324

Bozorg-Haddad O, Mariño MA (2007) Dynamic penalty function as a strategy in solving water resources combinatorial optimization problems with honey-bee mating optimization (HBMO) algorithm. J Hydroinf 9(3):233–250. https://doi.org/10.2166/hydro.2007.025

Bozorg-Haddad O, Mariño MA (2011) Optimum operation of wells in coastal aquifers. Proc Inst Civ Eng Water Manag 164(3):135–146. https://doi.org/10.1680/wama.1000037

Bozorg-Haddad O, Moradi-Jalal M, Mirmomeni M, Kholghi MKH, Mariño MA (2009a) Optimal cultivation rules in multi-crop irrigation areas. Irrig Drain 58(1):38–49. https://doi.org/10.1002/ird.381

Bozorg-Haddad O, Afshar A, Mariño MA (2009b) Optimization of non-convex water resource problems by honey-bee mating optimization (HBMO) algorithm. Eng Comput 26(3):267–280. https://doi.org/10.1108/02644400910943617

Bozorg-Haddad O, Mirmomeni M, Mariño MA (2010a) Optimal design of stepped spillways using the HBMO algorithm. Civ Eng Environ Syst 27(1):81–94. https://doi.org/10.1080/10286600802542465

Bozorg-Haddad O, Mirmomeni M, Zarezadeh Mehrizi M, Mariño MA (2010b) Finding the shortest path with honey-bee mating optimization algorithm in project management problems with constrained/unconstrained resources. Comput Optim Appl 47(1):97–128. https://doi.org/10.1007/s10589-008-9210-9

Bozorg-Haddad O, Rezapour-Tabari MM, Fallah-Mehdipour E, Mariño MA (2013) Groundwater model calibration by meta-heuristic algorithms. Water Resour Manag 27(7):2515–2529. https://doi.org/10.1007/s11269-013-0300-9

Bozorg-Haddad O, Ashofteh PS, Mariño MA (2015) Levee layouts and design optimization in protection of flood areas. J Irrig Drain Eng 141(8):04015004. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000864

Bozorg-Haddad O, Janbaz M, Loáiciga HA (2016) Application of the gravity search algorithm to multi-reservoir operation optimization. Adv Water Resour 98:173–185. https://doi.org/10.1016/j.advwatres.2016.11.001

Bozorg-Haddad O, Soleimani S, Loáiciga HA (2017a) Modeling water-quality parameters using a genetic algorithm–least squares support vector regression and genetic programming. J Environ Eng 143(7):04017021

Bozorg-Haddad O, Soleimani S, Loáiciga HA (2017b) Modeling water-quality parameters using genetic algorithm-least squares support vector regression and genetic programming. J Environ Eng 143(7):04017021. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001217

Bozorg-Haddad O, Delpasand M, Loáiciga HA (2020a) Self-optimizer data-mining method for aquifer level prediction. Water Sci Technol Water Supply 20(2):724–736

Bozorg-Haddad O, Zolghadr-Asli B, Sarzaeim P, Aboutalebi M, Chu X, Loáiciga HA (2020b) Evaluation of water shortage crisis in the middle east and possible remedies. J Water Supply Res Technol AQUA 69(1):85–98

Cramer NL (1985) A representation for the adaptive generation of simple sequential programs. In: Proceedings of an international conference on genetic algorithms and their applications

Fallah-Mehdipour E, Bozorg-Haddad O, Beygi S, Mariño MA (2011) Effect of utility function curvature of Young's bargaining method on the design of WDNs. Water Resour Manag 25(9):2197–2218. https://doi.org/10.1007/s11269-011-9802-5

Fallah-Mehdipour E, Bozorg-Haddad O, Mariño MA (2013a) Extraction of multicrop planning rules in a reservoir system: application of evolutionary algorithms. J Irrig Drain Eng 139(6):490–498. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000572

Fallah-Mehdipour E, Bozorg-Haddad O, Mariño MA (2013b) Extraction of optimal operation rules in an aquifer-dam system: genetic programming approach. J Irrig Drain Eng 139(10):872–879. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000628

Ghoreishi M, Razavi S, Elshorbagy A (2021) Understanding human adaptation to drought: agent-based agricultural water demand modeling in the Bow river basin. Can Hydrol Sci J 66(3):389–407

Gill MK, Asefa T, Kaheil Y, McKee M (2007) Effect of missing data on the performance of learning algorithms for hydrologic predictions: implications to an imputation technique. Water Resour Res. https://doi.org/10.1029/2006WR005298

Hailu R, Tolossa D, Alemu G (2018) Integrated water resources management as a system approach for water security: evidence from the Awash river basin of Ethiopia. Ethiop J Soc Sci Humanit 14(1):53–82

Karimi-Hosseini A, Bozorg-Haddad O, Mariño MA (2011) Site selection of raingauges using entropy methodologies. Proc Inst Civ Eng Water Manag 164(7):321–333. https://doi.org/10.1680/wama.2011.164.7.321

Kelleher JD, Tierney B (2018) Data science. MIT Press, Cambridge

Khelifa R, Mahdjoub H, Baaloudj A, Cannings RA, Samways MJ (2021) Effects of both climate change and human water demand on a highly threatened damselfly. Sci Rep 11(1):1–11
King J, Louw D (1998) Instream flow assessments for regulated rivers in South Africa using the building block methodology. Aquat Ecosyst Health Manag 1(2):109–124
King J, Brown C, Sabet H (2003) A scenario-based holistic approach to environmental flow assessments for rivers. River Res Appl 19(5–6):619–639
Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection. MIT Press, Cambridge
Koza JR (1994) Genetic programming II: automatic discovery of reusable programs. MIT Press, Cambridge
Li X, Wang Y, Liu G (2019) Structured medical pathology data hiding information association mining algorithm based on optimized convolutional neural network. IEEE Access 8:1443–1452
Linnansaari T, Monk WA, Baird DJ, Curry RA (2012) Review of approaches and methods to assess environmental flows across Canada and internationally. Fish Oceans Can Sci
Malla FA, Mushtaq A, Bandh SA, Qayoom I, Hoang AT (2022) "Understanding climate change: scientific opinion and public perspective" climate change. Springer, Cham, pp 1–20
Mengistu D, Bewket W, Dosio A, Panitz HJ (2021) Climate change impacts on water resources in the upper blue Nile (Abay) River Basin, Ethiopia. J Hydrol 592:125614
Mohammad-Azari SM, Bozorg-Haddad O, Loáiciga HA (2020) State-of-art of genetic programming applications in water-resources systems analysis. Environ Monit Assess 192(2):1–17
Mozaffari GA (2022) Climate change and its co*nse*quences in agriculture. In: The nature, causes, effects, and mitigation of climate change on the environment, 83
Oliazadeh A, Bozorg-Haddad O, Mani M, Chu X (2021) Developing an urban runoff management model by using satellite precipitation datasets to allocate low impact development systems under climate change conditions. Theor Appl Climatol 146:675–687
Oliazadeh A, Bozorg-Haddad O, Pakdaman M, Baghbani R, Loáiciga HA (2022) Optimal merging of multi-satellite precipitation data in urban areas. Theoret Appl Climatol 147(3–4):1697–1712
Orouji H, Bozorg-Haddad O, Fallah-Mehdipour E, Mariño MA (2014) Extraction of decision alternatives in project management: application of hybrid PSO-SFLA. J Manag Eng 30(1):50–59. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000186
Richter BD, Baumgartner JV, Powell J et al (1996) A method for assessing hydrologic alteration within ecosystems. Conserv Biol 10(4):1163–1174
Richter BD, Baumgartner JV, Wigginton R et al (1997) How much water does a river need? Freshw Biol 37(1):231–249
Richter BD, Baumgartner JV, Braun DP et al (1998) A spatial assessment of hydrologic alteration within a river network. River Res Appl 14(4):329–340
Sabbaghpour S, Naghashzadehgan M, Javaherdeh K, Bozorg-Haddad O (2012) HBMO algorithm for calibrating water distribution network of Langarud city. Water Sci Technol 65(9):1564–1569. https://doi.org/10.2166/wst.2012.045
Sarzaeim P, Bozorg-Haddad O, Bozorgi A, Loáiciga HA (2017a) Runoff projection under climate change conditions with data-mining methods. J Irrig Drain Eng 143(8):04017026
Sarzaeim P, Bozorg-Haddad O, Fallah-Mehdipour E, Loáiciga HA (2017b) Environmental water demand assessment under climate change conditions. Environ Monit Assess 189(7):1–18
Sedighkia M, Abdoli A, Datta B (2021) Optimizing monthly ecological flow regime by a coupled fuzzy physical habitat simulation–genetic algorithm method. Environ Syst Decis 41(3):425–436
Soltanjalili M, Bozorg-Haddad O, Mariño MA (2011) Effect of breakage level one in design of water distribution networks. Water Resour Manag 25(1):311–337. https://doi.org/10.1007/s11269-010-9701-1
Tan P-N, Steinbach M, Karpatne A, Kumar V (2019) Introduction to data mining. Person Educ Inc., New York
Tennant DL (1976) Instream flow regimens for fish, wildlife, recreation and related environmental resources. J Fish 1(4):6–10
Than NH, Ly CD, Van Tat P (2021) The performance of classification and forecasting Dong Nai River water quality for sustainable water resources management using neural network techniques. J Hydrol 596:126099
Tharme RE (2003) A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. River Res Appl 19(5–6):397–441
Tian J, Guo S, Deng L, Yin J, Pan Z, He S, Li Q (2021) Adaptive optimal allocation of water resources respo*NSE* to future water availability and water demand in the Han river basin, China. Sci Rep 11(1):1–18
Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Wang Y, Liu P, Dou M, Li H, Ming B, Gong Y, Yang Z (2021) Reservoir ecological operation considering outflow variations across different time scales. Ecol Ind 125:107582

Wilby R, Dawson CW (2013) The Statistical downscaling model: insights from one decade of application. Int J Climatol 33(7):1707–1719

Zolghadr-Asli B, Bozorg-Haddad O, Chu X (2019) Effects of the uncertainties of climate change on the performance of hydropower systems. J Water Clim Chang 10(3):591–609

Authors and Affiliations

Masoud Zanjani[1] · Omid Bozorg-Haddad[2] · Mustafa Zanjani[1] · Ali Arefinia[2] · Masoud Pourgholam-Amiji[3] · Hugo A. Loáiciga[4]

✉ Omid Bozorg-Haddad
OBHaddad@ut.ac.ir

Masoud Zanjani
Maseod_zanjani@ut.ac.ir

Mustafa Zanjani
Zanjani.mustafa@ut.ac.ir

Ali Arefinia
Ali.Arefinia@ut.ac.ir

Masoud Pourgholam-Amiji
Mpourgholam6@ut.ac.ir

Hugo A. Loáiciga
hloaiciga@ucsb.edu

1 Department of Water Engineering, College of Aburaihan, University of Tehran, Tehran, Iran

2 Department of Irrigation and Reclamation Engineering, Faculty of Agricultural Engineering and Technology, College of Agriculture and Natural Resources, University of Tehran, Karaj, Tehran, Iran

3 Department of Irrigation and Reclamation Engineering, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

4 Department of Geography, University of California, Santa Barbara, CA 93016-4060, USA