Original Article

Journal of Drought and Climate change Research (JDCR)

Winter 2025, Vol. 2, No. 8, pp 141-160

10.22077/jdcr.2025.8683.1102

Evaluating Hybrid Models and Google Earth Engine for Predicting Climate Change Impacts on Runoff in the Kasilian Catchment, Northern Iran

Farhad Hajian[1*], Elham Yousefi[2], Hossein Monshizadeh Naeen[3]

1- Department of Civil Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.
2- Department of Environment, Faculty of Natural Resources and Environment, University of Birjand, Iran
3- Department of Computer Engineering, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran..
*** corresponding author: F.hajian@yahoo.com**

Keywords:
Remote sensing, Hybrid models, Google earth engine (GEE), Erosion, Sediment yield.

Received:
31 Dec 2024

Revised:
24 Jan 2025

Accepted:
25 Jan 2025

Abstrsct

This study investigates the effect of climate change on annual rainfall and runoff of Kasilian catchment through two distinct approaches. Firstly, it utilizes hybrid models by integrating the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) with Artificial Neural Networks (ANN), Support Vector Machines (SVM), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Gene Expression Programming (GEP) separately, as well as employing the Long Ashton Research Station Weather Generator (LARS-WG). Secondly, it employs Google Earth Engine (GEE) to analyze changes in annual rainfall and runoff for the observed period, compensating for incomplete data from hydrometric and climatological stations. The results demonstrate that under the SSP585 scenario, from various climate models in LARS WG and when employing hybrid models, the median annual rainfall is projected to increase in the future compared to the base period, while the median annual runoff is expected to decrease due to rising temperatures and increased evapotranspiration. Consistent with these projections, GEE data from 1981 to 2023 also indicates an increase in annual rainfall and a decrease in annual runoff. Additionally, there is a reduction in annual erosion and sedimentation rates, attributed to the reduced capacity of runoff to transport sediment. These

How to cite this article:
Hajian, F., Yousefi, E. & Monshizadeh Naeen, H. (2024). Evaluating Hybrid Models and Google Earth Engine for Predicting Climate Change Impacts on Runoff in the Kasilian Catchment, Northern Iran. *Journal of Drought and Climate change Research* (JDCR), 2(8), 141-160. 10.22077/jdcr.2025.8683.1102

findings highlight the potential for more extreme rainfall events, increased annual precipitation, and a subsequent decrease in annual runoff and sediment load in the Kasilian catchment, providing essential perspectives for managing water resources.

Introduction

The primary sources of uncertainty in climate change impact studies stem from the structure of Global Climate Models (GCMs) or climate data, the structure of hydrological models, and the parameters used within hydrological models. Different climate models produce varying rainfall patterns, resulting in different runoff magnitudes (Hajian, 2013). For example, Abbaspour et al. (2009) found increased precipitation in northern Iran up to 40% for future periods, while Babaeian et al. (2007) observed a 2% decrease in rainfall for Mazandaran Province for 2010-2039 compared to 1976-2005. The variability in rainfall predictions highlights the necessity of using multiple climate models in climate change impact studies. Many studies have traditionally relied on a single GCM, underscoring the need for diverse model applications to better understand potential impacts on water resources.

Bae et al. (2011) examined uncertainties in climate change impact studies in the Chungju Dam Catchment, South Korea, using three semi-distributed hydrological models (PRMS, SLURP, and SWAT) with multiple potential evapotranspiration methods, and 39 climate change scenarios across two future periods (2011-2040 and 2071-2100). The models performed identically for calibration and validation with observed data, but showed varying runoff results with GCM outputs. Runoff changes were significant during winter (dry period) and minor in other seasons, influenced by model structure, GCM type, and evapotranspiration methods. This highlights the need for careful selection of hydrological models, GCMs, and parameterization methods in climate impact studies, as these choices significantly affect the results.

Shifteh Some'e et al. (2012) analyzed the percent change of annual and seasonal rainfall data for 28 synoptic stations of Iran over the period 1967-2006. The annual rainfall increased by 0-10%, spring rainfall increased by 10%, summer rainfall increased by 0-30%, autumn rainfall increased by 0-10%, and a noticeable decrease of 10-20% in the winter rainfall was observed over the period 1967-2006, for northern parts (southern Caspian Sea coastal area).

The following studies collectively demonstrate the effectiveness of various advanced modeling approaches in accurately predicting and simulating hydrological processes across different regions. One study compared the performance of Gene Expression

Programming (GEP), Adaptive Neuro-Fuzzy Inference System (ANFIS), and Support Vector Machine (SVM) models in simulating rainfall using historical data from the Kiyav-Chay River basin in Iran. The GEP model outperformed the ANFIS and SVM models, achieving the highest R^2 value and the lowest RMSE and MAE values, concluding that GEP is a reliable and accurate method for rainfall simulation in this region (Tabatabaei et al., 2021). Another study reviewed LS-SVM, GEP, and ANFIS-PSO models for simulating rainfall-runoff in the Halil River, determining that the ANFIS-PSO model outperformed others with the lowest RMSE (0.35) and highest R^2 (0.92) values (Kavoosi & Khozeymehnezhad 2021). Different studies in the Anandpur and Champua catchments of the Baitarani catchment, India, evaluated ANN, ANFIS, and SRC models, with ANN and ANFIS models performing best for sediment load simulation (Kumaret al., 2019).

A study introduced a hybrid model named Ia-LSTM, which combines the physical-based HEC-HMS model with the data-driven LSTM model to optimize the "initial loss" (Ia) and accurately capture the rainfall-runoff relationship. This model was tested in the Yufuhe basin in Jinan City, Shandong province, China. The Ia-LSTM model consistently demonstrated superior performance compared to the individual HEC-HMS and LSTM models. It achieved impressive average Nash-Sutcliffe Efficiency (NSE) values of 0.873 and 0.829, and average R^2 values of 0.916 and 0.870 during calibration and validation, respectively (Zhang et al., 2022).

Asadi and Santos (2022) created a hybrid model that integrates artificial intelligence (ANN) with the SWAT semi-distributed hydrological model, achieving high accuracy with NSE values of 0.85 for calibration and 0.82 for validation. This research was conducted in the Upper Sabarmati River Basin in Gujarat, India. In another study, Gebremichael and Hailu (2024) compared the HEC-HMS model with machine learning models (SVM, ANN, and ANFIS) for rainfall-runoff prediction, discovering that the ANFIS model outperformed the others with NSE values of 0.88 for calibration and 0.85 for validation. This study took place in the Rappahannock River basin near Fredericksburg, Virginia, USA.

This study seeks to improve the performance of the HEC-HMS model by integrating it with Artificial Neural Networks (ANN), Gene Expression Programming (GEP), Support Vector Machines (SVM), and Adaptive Neuro-Fuzzy Inference Systems (ANFIS). These integrated approaches, collectively known as hybrid models, will then be applied to climate change studies. The objectives of this study are:

1. To evaluate the accuracy of these

hybrid models in simulating rainfall-runoff processes in the Kasilian Catchment.

2. To evaluate the effect of climate change on runoff in the Kasilian Catchment by applying these hybrid models. While the integration of conceptual model like HEC-HMS with machine learning models are commonly used to improve model performance, limited research has explored their application in assessing climate change effects on water resources.
3. To validate the future projections generated using the hybrid and LARS-WG models with observational data from Google Earth Engine (GEE), considering the trend of change. This additional validation step confirmed the future projections to a significant extent by analyzing changes in annual rainfall, annual runoff, and annual sediment load based on observed data.

Material and Methods

Study area

The Kasilian Watershed, situated in Mazandaran Province in northern Iran, spans coordinates 53°18′ to 53°30′E and 35°58′ to 36°07′N (Figure 1). Covering an area of 65.7 km² above the Valikbon hydrometric station, the watershed flows northward into the Caspian Sea with its longest flow path extending 17.8 kilometers. Approximately 80% of the area is forested, while the lower regions have been cleared for agriculture. The geology consists of shale, sandstone, marl, and siltstone (Hajian, 2013). The Sangdeh meteorological station records daily rainfall and temperature, and the Valikbon station monitors river discharge (Figure 1). From 1977 to 1996, the average annual rainfall was about 756 mm, and the average annual runoff from 1980 to 1996 was 229 mm (Hajian, 2013).

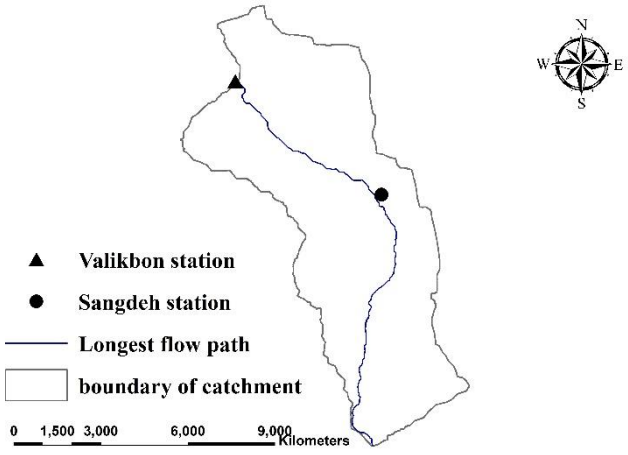

Fig 1. Study area.

Modeling the Rainfall-Runoff Process with HEC-HMS

In this study, the HEC-HMS model, developed by the US Army Corps of Engineers, was used to evaluate hydrological processes within the Kasilian Catchment. Due to limited data availability, the Thornthwaite method was employed to estimate potential evapotranspiration (PET) using temperature data. The study found that uncertainties in PET had a smaller impact on runoff simulations compared to variations arising from different types of General Circulation Models (GCM) or future climate scenarios. The HEC-HMS model, which utilizes 23 parameters, was optimized using the Nelder-Mead method. Initial parameter values were derived from geological, soil, and land use maps, as well as existing literature (Hajian, 2013). The calibration-validation process indicated poor performance, with Nash and Sutcliffe coefficients of -0.293 during the calibration period (September 23, 2007, to March 20, 2018) and 0.060 during the validation period (September 23, 2018, to September 22, 2021). The inadequate performance of HEC-HMS in estimating seasonal runoff volumes is attributed to seasonal variations in parameter values such as maximum infiltration rate and surface storage. To achieve accurate estimations, seasonal optimization might be required; however, the lumped HEC-HMS model cannot accommodate such adjustments, thus limiting its accuracy in simulating seasonal rainfall-runoff dynamics.

Hybrid modelling approach (Integrating machine learning (ML) with HEC-HMS conceptual model)

Improving the precision of runoff predictions can be achieved by combining conceptual hydrological models with machine learning (ML) techniques. Methods such as artificial neural networks (ANN), gene expression programming (GEP), Support Vector Machines (SVM), and the Adaptive Neuro-Fuzzy Inference System (ANFIS) offer powerful tools for refining forecast accuracy. Incorporating streamflow estimates from conceptual models as input for ML algorithms greatly enhances runoff forecasting precision. Research by Farfan et al., (2020) and Hitokoto & Sakuraba (2020) has demonstrated the effectiveness of this method, leading to significantly improved streamflow predictions. These findings highlight the advantages of integrating conceptual models with ML techniques for more accurate and dependable runoff estimations. These studies underscore the potential of combining conceptual models with ML techniques for more precise and reliable runoff predictions. The ML methods use discharge estimates from HEC-HMS (Q_t), along with estimates from one-day (Q_{t-1}) and two-day (Q_{t-2}) lags, representing previous days' discharge. As

highlighted in Table 1, the integration of SVM with HEC-HMS outperformed other hybrid models and the standalone HEC-HMS model. Nonetheless, the impact of climate change on runoff was examined across all hybrid models, regardless of their individual performance.

Table 1. The performance evaluation of different hybrid models.

Integration HEC HMS & ANN (3,3,1) with Tansig activation function		**Integration HEC HMS & GEP**		**Integration HEC HMS & SVM with RBF kernel and gamma=10 and σ^2 = 0.2**		**Integration HEC HMS & ANFIS**	
NSE train	NSE test	NSE train	NSE test	NSE train	NSE test	NSE train	NSE test
0.253	0.11	0.261	0.153	0.514	0.072	0.16	0.146

Artificial Neural Network (ANN)

Complex biological systems, with their intricate networks of interconnected processing units, demonstrate remarkable capabilities in information processing and pattern recognition. Individual processing units, while limited in their computational capacity, collectively engage in sophisticated tasks. These systems learn from their environment and establish connections that facilitate information flow through input pathways, processing stages, and output pathways. Artificial Neural Networks (ANNs) are computational models inspired by the architecture of such biological systems. They mimic the interconnected network of processing units for a range of tasks, including classification, pattern recognition, and regression. In ANNs, inputs are analogous to the input pathways, connection weights to the strength of inter-unit connections, activation functions to the influence of the processing units, and outputs to the output pathways.

ANNs typically comprise interconnected layers of artificial neurons. A crucial feature is the inclusion of hidden layers positioned between the input and output layers. These hidden layers enable ANNs to capture complex non-linear relationships and patterns within the data, functioning through multiple processing stages. Each artificial neuron is a mathematical construct involving inputs, associated weights, a bias term, an activation function, and an output. Mathematically, the output of a single neuron can be expressed as

$$y_j = f\left(\sum_{i=1}^{n} w_{ij}x_i + b_j\right) \quad (1)$$

Where is the output of the th neuron, is the th input to the neuron. is the weight connecting the th input to the th neuron, is the bias term of the th neuron neuron, and is the number of inputs to the neuron. The activation function introduces non-

linearity into the model, allowing ANNs to model complex mappings that are not linearly separable (Glorot et al. 2010).

ANNs are particularly well-suited for regression tasks, which involve the prediction of continuous-valued outputs and the modeling of intricate input-output relationships, rendering them invaluable across various practical domains. Among different ANN architectures, feed-forward backpropagation networks, optimized with the Levenberg-Marquardt algorithm, have demonstrated effective training performance, making them widely applicable.

This study utilizes a feed-forward backpropagation artificial neural network (ANN) model, with weight optimization achieved through the Levenberg-Marquardt algorithm. In many nonlinear problems, a single hidden layer is often sufficient to achieve reliable results. Increasing the number of hidden layers beyond two generally does not lead to significant performance improvements and may introduce issues like overfitting or vanishing gradients (Pashazadeh & Javan, 2019; Kisi et al., 2013). In this study, after evaluating various ANN configurations, it was found that a single hidden layer offered the best performance. Specifically, the Tansig activation function with three neurons in the hidden layer produced optimal results. The final optimal ANN architecture for investigating climate change impacts was determined to be ANN (3,3,1), comprising three input neurons, three hidden layer neurons, and one output neuron (Table 1). The ANN model, combined with a calibrated HEC-HMS model, was trained and tested with different configurations. However, the study does not detail the results for other configurations, such as varying activation functions and different numbers of hidden layer neurons.

Gene Expression Programming (GEP)

Gene Expression Programming (GEP) extends traditional evolutionary algorithms by representing solutions as expression trees. These expression trees offer a transparent and mathematically explicit relationship between inputs and outputs. This characteristic of clarity makes GEP particularly valuable for applications such as runoff prediction, where interpretability of the model is a crucial requirement. The GEP process commences with the definition of a fitness function.

In Gene Expression Programming (GEP) modeling, the initial step involves selecting an appropriate fitness function (Kisi et al., 2013). This study employed the Root Mean Square Error (RMSE) as the fitness function, which guided the evolutionary process of the expression trees by evaluating the differences between predicted and observed values. Various functional and terminal nodes

were utilized to construct the expression trees. The default functions provided by GeneXpro were used, including addition, subtraction, multiplication, division, power, square root, exponential, natural logarithm, absolute value, inverse, cube root, sine, cosine, tangent, cotangent, cosecant, secant, arcsine, arccosine, arctangent, arc cotangent, arc cosecant, arc secant, hyperbolic tangent, and other linking functions. These functions are combined to generate equations that relate input features to the target variables. The general form of a GEP-derived equation can be represented as:

$$y = \sum_{j=1}^{g} c_j f_j(X) \quad (2)$$

Here is the output variable, represents the th functional node, is the coefficient associated with the j-th functional node, is the vector of input variables, and is the number of genes in the model.

This example illustrates how GEP integrates both linear and nonlinear relationships and domain-specific information into a predictive model. The structure of these equations can help to understand the relationships learned between the input and output variables.

A key advantage of combing HEC-HMS with Gene Expression Programming (GEP) is GEP's ability to derive explicit mathematical relationships between input and output variables—something that traditional models like ANN and ANFIS often lack (Kisi et al., 2013). The GEP-derived expression formulated in this study for the model presented in Table 1 is:

$$y = cos\left(cos\left(cos\left(\sqrt[3]{d_2} - d_0 - 1.32938895082628\right)\right)\right) + \left(\frac{(d_2 - d_0) \times d_0 \times acot(-7.12454252754295)}{\cos(\cos(d_2))}\right) + \left(\frac{1}{\tan\left(\sqrt{e^{1.34469875637074} - d_0}\right) - \frac{1}{59.299337086763 - (-7.57133701590014)}}\right) \quad (3)$$

Here, d_0 denotes the discharge estimates from HEC-HMS at time t (Q_t), and d_2 refers to discharge estimates from HEC-HMS two days prior (Q_{t-2}). The genetic operators utilized for the HEC-HMS and GEP integration, as detailed in Table 1, are as follows: chromosome count: 30, head size: 7, gene count: 3, linking function: addition, fitness function: RMSE, Mutation rate: 0.00138, inversion rate: 0.00546.

Support Vector Machine (SVM)

Support Vector Machines (SVMs) are powerful supervised learning algorithms that have been widely applied in hydrological modeling due to their capacity to handle nonlinear relationships and high-dimensional datasets (Mountrakis et al., 2011). The fundamental principle of SVMs involves identifying an optimal hyperplane that separates data points of different classes (or, in the case of regression, accurately models continuous outputs) by maximizing the margin between them. For regression tasks, the SVM algorithm employs the concept of ε-insensitive loss, focusing primarily on predictions that

deviate from actual values by more than a specified tolerance, ε (Smola et al., 2004). This approach allows for more robust and generalized model predictions. For an input vector x and corresponding output y, the SVM regression function can be represented as:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \quad (4)$$

Where represents the support vector coefficients, is the kernel function that defines the similarity between the input vector and the support vectors , is the number of support vectors, and is the bias term.

The kernel function, , is a critical component of SVM, enabling it to map data into higher-dimensional spaces to capture complex nonlinearities. In this study, the Radial Basis Function (RBF) kernel was primarily utilized, alongside linear and polynomial kernels. The RBF kernel, also known as Gaussian kernel, is defined as:

$$K(x, x_i) = \exp(-\gamma \parallel x - x_i \parallel^2) \quad (5)$$

Where is the kernel parameter that controls the width of the Gaussian function, and consequently affects the influence of each support vector.

Parameter optimization is crucial to the performance of SVM models. This includes optimizing the penalty parameter, which controls the balance between minimizing training error and maximizing the margin, and the kernel-specific parameters, such as for the RBF kernel and the degree of the polynomial for the polynomial kernel.

In this study, linear kernel, polynomial kernel, and RBF kernel were used in SVM integration with the HEC-HMS model. The gamma parameter (γ) was optimized through a systematic trial-and-error process for all kernels, with γ values tested over a range of {10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 20, 50, 100}. An additional sigma squared parameter (σ^2) for the RBF kernel was also determined through a similar trial-and-error process, with σ^2 varied in steps including {0.1, 0.2, 0.3, 0.4, ..., 10}. Specifically, the RBF kernel with $\gamma = 10$ and $\sigma^2 = 0.2$ provided the best results in terms of the NSE coefficient (Table 1). The selected values minimized the validation error and achieved a good balance between model complexity and generalization, thereby optimizing the model's performance. The SVM model was configured for function approximation (type f), as shown in Table 1.

Adaptive Neuro-Fuzzy Inference System (ANFIS)

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a hybrid computational model that combines the strengths of fuzzy logic and neural networks. This integration makes ANFIS particularly effective for modeling nonlinear and

complex relationships, such as those commonly found in hydrology (Fahimi et al. 2017). ANFIS utilizes fuzzy IF-THEN rules to mimic human-like reasoning and refines these rules through neural network training techniques to optimize predictive performance. As a result, the model is both interpretable, due to the fuzzy logic component, and capable of learning complex patterns, thanks to the neural network component.

The ANFIS model is based on a Sugeno-type fuzzy inference system, where the output is a linear combination of input variables within each fuzzy rule. A typical fuzzy IF-THEN rule in ANFIS can be expressed as:

(6)

IF is AND is ... AND is THEN

$y = p_0 + p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$

Where represents the input variables, are the membership functions (MFs) associated with each input variable, is the output variable, and are the consequent parameters.

The membership functions (MFs) quantify the degree to which each input variable belongs to a particular fuzzy set. These functions, such as Gaussian or Generalized Bell (GBellMF) functions, mathematically define the degree of membership. The mathematical form of the Generalized Bell membership function (GBellMF) is:

$$\mu_A(x) = \frac{1}{1 + \left|\frac{x - c}{a}\right|^{2b}} \quad (7)$$

Where a, b, and c are parameters that control the shape of the membership function, and represents the degree of membership of in the fuzzy set . Training ANFIS involves optimizing the parameters of the membership functions (a, b, c) and the consequent parameters (). A hybrid optimization method is typically used to refine these parameters, combining gradient descent to adjust the membership function parameters and least squares estimation to determine the consequent parameters.

In this study, an ANFIS model was trained and tested using a hybrid optimization method with three input membership functions (MFs) set to the GBellMF type and a linear output MF type (Table 1). The use of multiple membership functions allows the model to capture a wide range of input variable behaviors. The GBellMF type has been shown to perform well in these contexts.

Future climate scenario projections

LARS-WG was selected for its excellent performance and ability to downscale GCM outputs for the Sangdeh station area. The tool's effectiveness was assessed by replicating observed daily data from 1985 to 2005 and producing 300 years of synthetic data, which mirrored the statistical properties of the observed data. LARS-WG then generated daily datasets for six models (ACCESS-

ESM1-5, CNRM-CM6-1, GFDL-ESM4, HADGEM3-GC31-LL, MPI-ESM1-2-LR, MRI-ESM2-0) under the SSP585 scenario for the periods 2031-2050 and 2051-2070. This scenario is appropriate for Iran due to the nation's high fossil fuel consumption and slow transition to renewable energy, leading to projections of more extreme weather events (Horanyi , 2023). These climate data allow for the comparison of future runoff conditions with data recorded from 2007-2018.

Results and discussion

Future Projections and Managing Uncertainty from Various Climate Model Forecasts

Climate models vary in their predictions of rainfall and runoff patterns. For example, some models within the LARS-WG framework project an increase in mean annual rainfall in certain areas, while others foresee a decrease in those same regions (Semenov et al. 1998). Utilizing a range of climate models is essential for accurate climate change impact assessments. To illustrate the uncertainty in these predictions, two box plots were created using future mean annual values of climatic variables (rainfall) and hydrological variables (runoff) from all General Circulation Models (GCMs). These future values are then compared with historical data. The box plots show the 25th and 75th percentiles, the median (50th percentile), and the minimum and maximum values (Semenov & Shewry, 2011). Figures 2 and 3 display the annual changes in rainfall and runoff for two future periods.

Throughout the calibration period from 2007 to 2018, the mean annual rainfall was 788 mm. When integrating HEC-HMS with ANN, GEP, ANFIS, and SVM, the mean annual runoff values were 174.23 mm, 183.29 mm, 164.62 mm, and 169.79 mm, respectively. The projected annual rainfall shows an increasing trend over the coming decades. From the calibration period, where the mean annual rainfall was 788 mm, there is a projected increase to a median annual rainfall of 807.4 mm in 2031-2051, representing a 2.46% rise. This trend continues, with the median annual rainfall expected to reach 834.9 mm in 2051-2070, marking a 5.95% increase from the calibration period. The projected median annual runoff using the integration of HEC-HMS with SVM shows a decreasing trend over the coming decades. For the period from 2031-2051, the median annual runoff is expected to be 163.81 mm, representing a 3.53% decrease from the calibration period's mean of 169.79 mm. In the period from 2051-2070, the median annual runoff is projected to rise slightly to 167.61 mm, which is a 1.28% decrease from the calibration period. It is crucial to recognize that integrating HEC-HMS with ANFIS, ANN, and GEP

models shows a projected decline in runoff compared to the calibration period. These trends indicate a decrease in annual runoff, highlighting the necessity for adaptive water management strategies to address potential variations in water availability (Figure. 2 and 3). Increased temperatures and higher evapotranspiration rates are anticipated to further lower median annual runoff.

The IPCC Sixth Assessment Report (2023) underscores that anthropogenic global warming of 1.1°C has precipitated unprecedented alterations in Earth's climate, manifesting as more extreme weather events and heightened rainfall in numerous regions. The report stresses that each additional 0.5°C rise in global temperature will trigger more frequent and intense heat extremes, heavy rainfall, and regional droughts. Similarly, Ripple et al. (2024) point out that escalating greenhouse gas emissions are driving more extreme weather events, with human-induced carbon dioxide emissions being the principal contributors to climate change, resulting in increased rainfall and more frequent extreme weather phenomena.

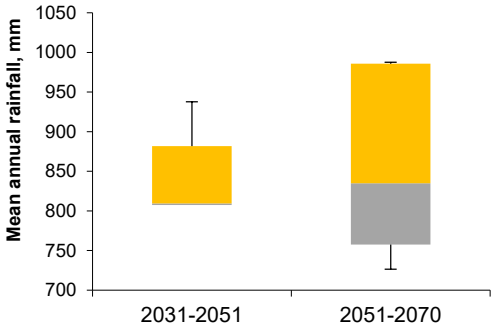

Fig 2. Box plots created from the average annual rainfall data derived from various climate models for the SSP585 scenario and projected future period.

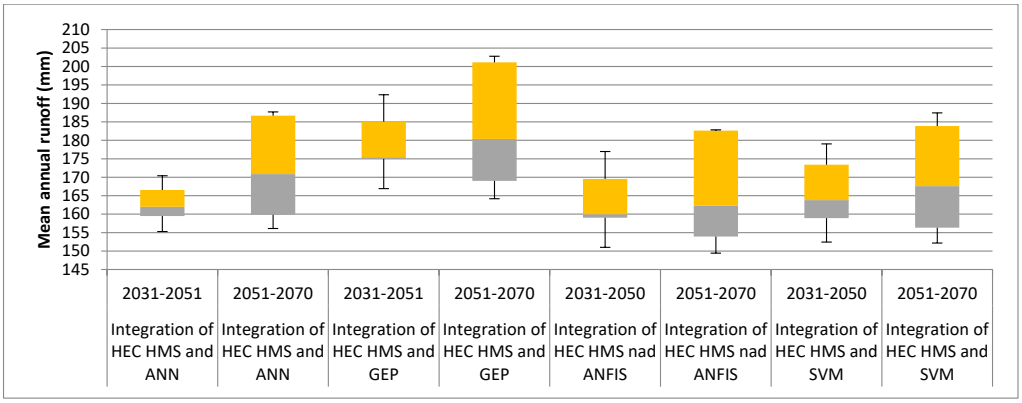

Fig 3. Box plots created using the average annual runoff values derived from various climate models for the SSP585 scenario and projected future period.

Google Earth Engine analysis to investigate the impact of climate change on rainfall and runoff

Over the past decade, extensive research has highlighted the effective use of remote sensing and the Google Earth Engine (GEE) platform for runoff and flood risk calculations. Studies by Yousefi et al. (2022 & 2024) and Tiwari et al. (2020) underscore GEE's capability in estimating runoff, sediment, and mapping flood inundation areas.

Additionally, we utilized Google Earth Engine to analyze climate change impacts using observational data, compensating for the incomplete datasets from the Valikbon Hydrometric Station and the Sangdeh Meteorological Station.

This study utilized Google Earth Engine to analyze annual precipitation trends using the CHIRPS dataset. We centered the region of interest (Kasilain catchment) on the map and aggregated daily CHIRPS precipitation data into annual sums for the years 1981 to 2023. The mean annual precipitation for the Kasilain catchment can be calculated by averaging these annual sums. The resulting data was visualized with a chart, depicting year-to-year variations in annual precipitation and providing insights into long-term climate trends and variability (Figure 4).

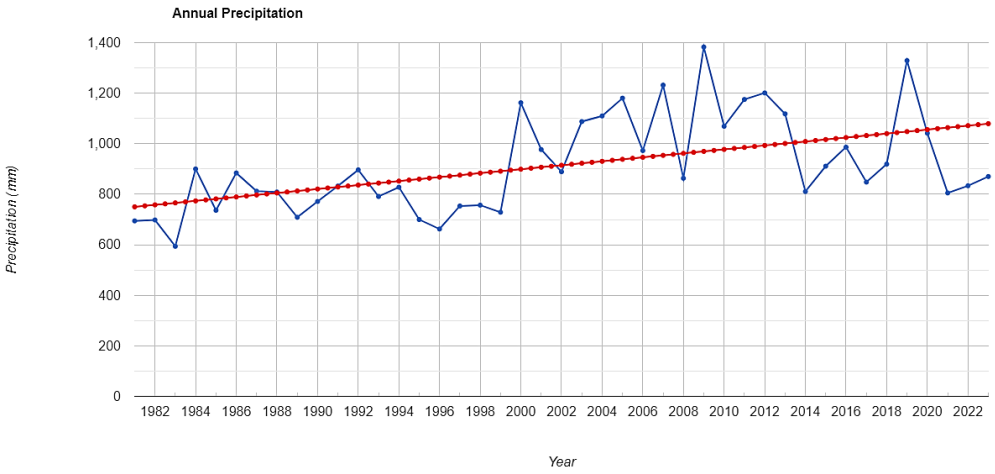

Fig 4. Annual Precipitation in the Kasilian catchment analyzed using Google Earth Engine, blue line: Annual precipitation data points, Red dashed line: Trend line indicating overall precipitation trend

A different code in Google Earth Engine (GEE) was used to analyze and visualize annual runoff data for the Kasilian catchment from 1981 to 2023, utilizing the IDAHO_EPSCOR/TERRACLIMATE dataset. Runoff data is filtered by spatial and temporal bounds, selecting only the runoff ('ro') band. The code then calculates annual runoff by summing the monthly runoff for each year and creates an image collection with each year's total runoff. This data is visualized through a time series

chart, plotting annual runoff over time and including a trendline to highlight long-term patterns. The chart, titled "Annual runoff Kasilian catchment," displays years on the x-axis and runoff values in millimeters on the y-axis, providing valuable insights into the hydrological behavior and trends of the Kasilian catchment over four decades (Figure 5).

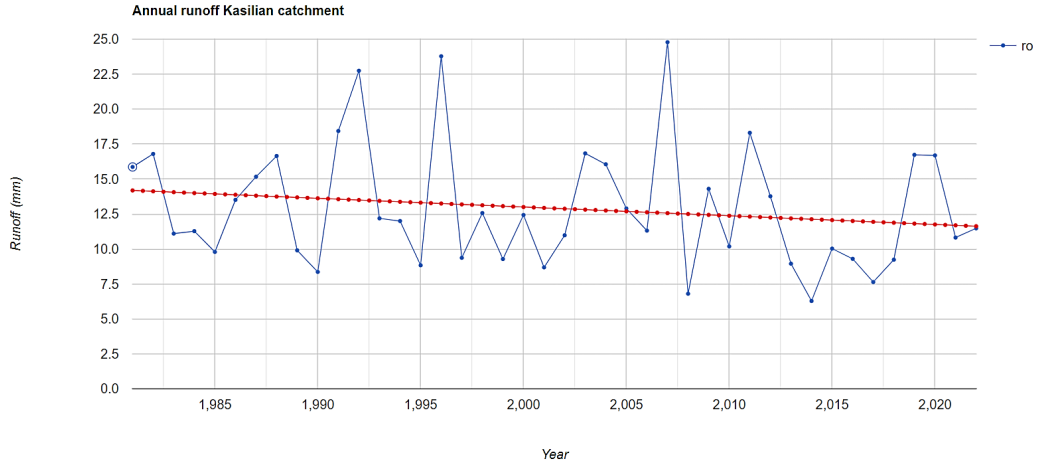

Fig 5. Annual runoff in the Kasilian catchment analyzed using Google Earth Engine, blue line: Annual runoff data points, Red dashed line: Trend line indicating overall runoff trend

A Google Earth Engine code visualizes annual Potential Evapotranspiration (PET) for the Kasilian catchment from 2000 to 2023 (Figure 6). This code centers the map on the catchment, specifies the time range, retrieves the MODIS PET product, and calculates the annual sum of PET. It then generates a time series chart of annual PET values, including axis titles and a trend line. Unfortunately, the MODIS PET data does not extend before 2000 but continues to be updated periodically up to the present. This visualization helps monitor changes in annual PET over time for the Kasilian catchment.

Sediment load and soil loss in Kasilian catchment

The Revised Universal Soil Loss Equation (RUSLE) is a highly used model demonstrating the connection between rainfall and soil erosion. In this equation (Eq. 8), R signifies the rainfall erosivity factor, K represents the soil erodibility factor, LS is the slope length and steepness factor (dimensionless), C denotes the cover management factor, P is the support practice factor (ranging from 0 to 1), and A represents the estimated soil loss ($ton \cdot hectare^{-1} \cdot annum^{-1}$) (Islam 2022, Benavidez et al. 2018, Sakhraoui & Hasbaia 2023).

It is important to note that various formulas from Islam (2022), Benavidez et al. (2018), and Sakhraoui & Hasbaia (2023) were utilized to estimate the C factor, K factor, R

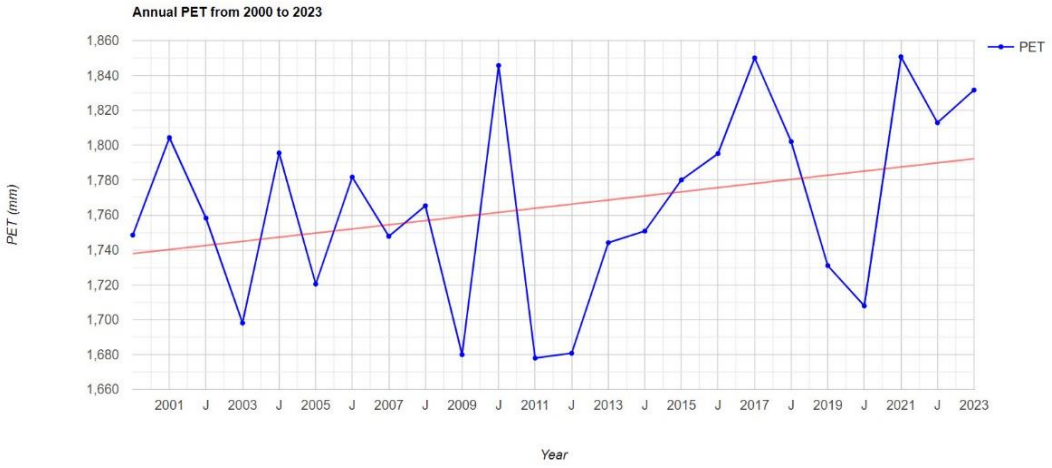

Fig 6. Annual potential evapotranspiration in the Kasilian catchment analyzed using Google Earth Engine, blue line: Annual potential evapotranspiration data points, Red dashed line: Trend line indicating overall potential evapotranspiration trend

factor, and LS factor in order to determine the actual value of mean sediment load for the Kasilian catchment. Ultimately, equations 10 to 15 were identified as the most accurate for estimating the true mean sediment load of the Kasilian catchment.

$$A = R \times K \times LS \times C \times P \quad (8)$$

The sediment delivery ratio (SDR) quantifies the proportion of eroded soil that is carried away from a specific land area over a certain period. This concept is mathematically represented as (Hajian 2013):

$$SDR = sediment\frac{yield}{soil\ erosion} \quad (9)$$

Rainfall factor (R):

$$MFI = \sum_{I=1}^{12} \frac{P_I^2}{p} \quad (10)$$

MFI: the Modified Fournier Index; Pi: the monthly precipitation; P: the annual precipitation;

The R factor is calculated as:

$$R = 0.07397 \times MFI^{1.847} \quad (11)$$

Soil factor (K):

The K factor was determined using Equation 12:

$$K = [0.2 + 0.3 \times exp - 0.0256 \times SAN \times (1 - \frac{SIL}{100}) \times \left[1 - \frac{0.25 \times CLA}{CLA + \exp(3.72 - 2.95 \times CLA)}\right] \quad (12)$$

where SAN represents the percentage of sand, SIL indicates the percentage of silt, and CLA denotes the percentage of clay.

LS factor (LS):

The LS factor was determined using Equation 13 and serves as an accelerating factor for rainfall erosion:

$$LS = \left(\frac{l}{22}\right)^{0.5} \times (0.065 + 0.045s + 0.0065s^2) \quad (13)$$

l: slope length (m)

s: slope steepness (%)

Cover management factor (C):

This relates to cover management, calculated using the Normalized Difference Vegetation Index (NDVI) (Equation 14):

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - RED)}{(\mathrm{NIR} + \mathrm{RED})} \quad (14)$$

NIR denotes the reflectance value in the near-infrared band, while RED indicates the reflectance value in the red band.

Equation 15 was utilized to determine the C factor, which is dependent on NDVI.

$$\mathrm{C\ factor} = \mathrm{e}^{\frac{-2\mathrm{NDVI}}{1-\mathrm{NDVI}}} \quad (15)$$

This methodology utilizes various satellite datasets to derive the factors of the Revised Universal Soil Loss Equation (RUSLE). Specifically, the C factor is obtained from MODIS NDVI data covering the period from January 1, 2000, to December 31, 2023. The LS factor is calculated using SRTM 30-meter Digital Elevation Model (DEM) data. Soil property data from OpenLandMap is used to estimate the K factor. Additionally, the CHIRPS daily precipitation dataset, spanning from January 1, 2000, to December 31, 2023, is employed to determine the R factor. By integrating these datasets with the appropriate formulas, a comprehensive model for estimating soil erosion is developed. The sediment load analysis is limited to the period from 2000 to 2022 due to the availability of NDVI data from the MODIS sensor (MODIS/006/MOD13A2), which was launched on December 18, 1999 (Figure 7). Therefore, prior to 2000, NDVI data necessary for estimating the C factor was unavailable.

Considering the sediment delivery ratio (SDR) of the Kasilian catchment obtained as 0.763 from Hajian (2013), the estimated mean annual sediment load of the Kasilian catchment using GEE (631.667 t km^{-2} yr^{-1}) closely matched the observed value for the Talar catchment (532 t km^{-2} yr^{-1}), of which Kasilian is a subcatchment. Therefore, the sediment load derived from GEE seems to more accurately reflect the sediment load conditions in the Kasilian catchment (Hajian, 2013).

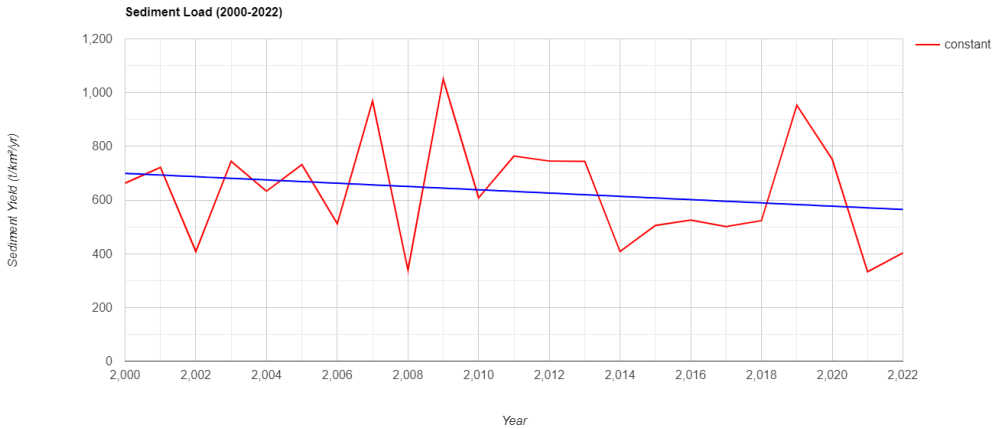

Fig 7. Annual sediment load in the Kasilian catchment analyzed using Google Earth Engine, blue line: Annual sediment load data points, Red dashed line: Trend line indicating overall sediment load trend

The increase in extreme rainfall events has led to higher annual rainfall, while rising temperatures and greater annual potential evapotranspiration have reduced annual runoff. This reduction in runoff has decreased sediment load due to the diminished capacity of streams to transport sediment. From 1981 to 2023, annual rainfall has increased, but annual runoff has decreased, as shown in Figures 4 and 5, consistent with Figures 2 and 3. The increase in rainfall enhances soil moisture, promotes vegetation growth, stabilizes the soil, reduces erosion, and leads to greater infiltration, which decreases erosion and sedimentation rates. Despite the rise in rainfall, the overall impact is a complex interplay of factors affecting hydrological and sedimentation processes in the catchment area.

Research by Cho et al. (2011) and Bae et al. (2008) shows that annual precipitation and mean air temperature significantly impact annual runoff. Higher temperatures increase evapotranspiration, reducing runoff despite higher precipitation. This underscores the necessity to consider temperature effects in water resource management and climate adaptation strategies. Future projections indicate increased rainfall but decreased runoff due to heightened evapotranspiration, stressing the need for integrating these climatic changes into water management strategies.

Conclusions

The integration of Support Vector Machine (SVM) and Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) models has shown superior performance over other hybrid models in simulating rainfall-runoff processes. Despite this, all hybrid models were used to evaluate the impact of climate change on runoff. Combining stream flow forecasts from conceptual models with machine learning techniques like ANN, GEP, SVM, and ANFIS improves runoff prediction accuracy and is highly effective for assessing the impacts of climate change on water resources. In the Kasilian Catchment and northern Iran, climate change is anticipated to result in more extreme rainfall events and increased annual rainfall. However, rising annual evapotranspiration rates are expected to decrease annual runoff. The increase in mean rainfall is likely driven by these extreme events. The annual sediment load likely decreased due to reduced runoff. This reduction in runoff diminished the ability of the streams to transport sediment, leading to lower sediment load. Observed data from the Google Earth Engine platform showed an increase in annual rainfall and a decrease in annual runoff, consistent with future analysis results using hybrid models and LARS-WG.

Author Contributions

The first author was responsible for all aspects of the study, including calculations and writing, as well as other related tasks. The second and third authors provided guidance and support throughout the research process.

References

Abbaspour K.C., Faramarzi M., Ghasemi S.S., & Yang H. (2009). Assessing the impact of climate change on water resources in Iran, *Water Resources*, 45, W10434. Doi:10.1029/2008WR007615.

Asadi, H., & Santos, C.A.G. (2022). A hybrid artificial intelligence and semi-distributed model for runoff prediction, *Water Supply*, 22(7), 6181-6195. doi: https://doi.org/10.2166/ws.2022.123

Babaeian I., Najafinik Z., Zabolabasi F., Habibinokhandan M., Adab H., & Malbosi S. (2007). Climate change investigation of Iran for 2010-2039 using downscaled output of ECHO-G climate model, *Geography and Development Iranian Journal*, 16, 135-152. [in Persian]. https://doi.org/10.22111/gdij.2009.1179.

Bae D.H., Jung W., & Chang H. (2008). Potential changes in Korean water resources estimated by high-resolution climate resolution, *Climate Research*, 35, 213-226. https://doi.org/10.3354/cr00704.

Bae D.H., Jung W., & Lettenmaier D.P. (2011). Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Catchment, Korea, *Journal of Hydrology*, 401, 90–105. https://doi.org/10.1016/j.jhydrol.2011.02.012

Benavidez, R., Jackson, B., Maxwell, D., & Norton, K. (2018). A review of the (Revised) Universal Soil Loss Equation ((R)USLE): with a view to increasing its global applicability and improving soil loss estimates. *Hydrology and Earth System Sciences*, 22(11), 6059-6086. doi:10.5194/hess-22-6059-2018.

Cho, J., Komatsu, H., Pokhrel, Y. et al.(2011). The effects of annual precipitation and mean air temperature on annual runoff in global forest regions. *Climatic Change,* 108, 401–410. https://doi.org/10.1007/s10584-011-0197-3.

Fahimi, F., Yaseen, Z. M., & El-shafie, A. (2017). Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. *Theoretical and applied climatology*, 128, 875-903. doi: https://doi.org/10.1007/s00704-016-1735-8

Farfan, J.F., Palacios, K., Ulloa, J., & Avilés, A. (2020). A hybrid neural network-based technique to improve the flow forecasting of physical and data-driven models: Methodology and case studies in Andean watersheds. *Journal of Hydrology: Regional Studies, 27*, 100652. https://doi.org/10.1016/j.ejrh.2020.100652.

Gebremichael, M., & Hailu, A. (2024). Comparative analysis of HEC-HMS and machine learning models for rainfall-runoff prediction, *Hydrology Research*, 55(3), pp. 456-470. doi: https://doi.org/10.2166/nh.2024.032

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feed forward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics, *JMLR Workshop and Conference Proceedings.*

Hajian, F. (2013). Effects of land cover and climate changes on runoff and sediment yield

from a forested catchment in northern Iran. PhD thesis. Kingston University.

Hitokoto, M., & Sakuraba, M. (2020). Hybrid deep neural network and distributed rainfall-runoff model for real-time river-stage prediction, *Journal of Japan Society of Civil Engineers (JSCE)*, 8(1), 46–58. https://doi.org/10.2208/journalofjsce.8.1_46

Horanyi, A. (2023). CMIP6: Global climate projections. European Centre for Medium-Range Weather Forecasts (ECMWF). https://confluence.ecmwf.int/display/COPSRV/CMIP6%3A+Global+climate+projections.

IPCC. (2023). Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland. doi:10.59327/IPCC/AR6-9789291691647.

Islam, Z. (2022). Soil loss assessment by RUSLE in the cloud-based platform (GEE) in Nigeria. *Modeling Earth Systems and Environment*, 8, 4579-4591. doi:10.1007/s40808-022-01467-7.

Kavoosi, M., & Khozeymehnezhad, H. (2021). Review and compare performance of 4 modeling methods LS-SVM, NN, GEP, and ANFIS-PSO in simulation of rainfall-runoff (Study Area: Halil River - Jiroft Dam). *Iranian Journal of Ecohydrology*, 8(2), 123-135. doi:10.22125/iwe.2021.128115.

Kisi, O., Shiri, J., & Tombul, M. (2013). Modeling rainfall-runoff process using soft computing techniques, *Computers & Geosciences*, 51, 108–117. doi: https://doi.org/10.1016/j.cageo.2012.07.001

Kumar, A., Kumar, P., & Singh, V. K. (2019). Evaluating different machine learning models for runoff and suspended sediment simulation. *Water Resources Management*, 33(4), 1217-1231. doi:10.1007/s11269-018-2178-z.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS journal of photogrammetry and remote sensing*, 66(3), 247-259. doi: https://doi.org/10.1016/j.isprsjprs.2010.11.001

Pashazadeh, A., & Javan, M. (2019). Comparison of the gene expression programming, artificial neural network (ANN), and equivalent Muskingum inflow models in the flood routing of multiple branched rivers, *Theoretical and Applied Climatology*, doi: https://doi.org/10.1007/s00704-019-03032-2.

Ripple, W. J., Wolf, C., Gregg, J. W., Rockström, J., Mann, M. E., Oreskes, N., Lenton, T. M., Rahmstorf, S., Newsome, T. M., Xu, C., Svenning, J.-C., Pereira, C. C., Law, B. E., & Crowther, T. W. (2024). The 2024 state of the climate report: Perilous times on planet Earth. *BioScience, 74*(10), 1-13. doi:10.1093/biosci/biae087.

Sakhraoui, F., & Hasbaia, M. (2023). Evaluation of the sensitivity of the RUSLE erosion model to rainfall erosivity: A case study of the Ksob watershed in central Algeria. *Water Supply, 23*(8), 3262-3284. https://doi.org/10.2166/ws.2023.207.

Semenov M.A., Brooks R.J., Barrow E.M., & Richardson C.W. (1998). Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates, *Climate Research*, 10, 95–107.

Semenov, M., & Shewry, P.R. (2011). Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe, *Scientific Reports*, 1, 66. doi: https://doi.org/10.1038/srep00066

Shifteh Some'e B., Ezani A., & Tabari H. (2012). Spatiotemporal trends and change point of precipitation in Iran, Atmospheric Research, 113, 1-12. https://doi.org/10.1016/j.atmosres.2012.04.016.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14, 199-222. doi: https://doi.org/10.1023/B:STCO.0000035301.49549.88

Tabatabaei, S. M., Nazeri Tahroudi, M., & Hamraz, B. S. (2021). Comparison of the performances of GEP, ANFIS, and SVM artificial intelligence models in rainfall simulation. *IDŐJÁRÁS / Quarterly Journal of the Hungarian Meteorological Service*, 125(2), 195-209. doi:10.28974/idojaras.2021.2.21.

Tiwari V, Kumar V, Matin MA, Thapa A, Ellenburg WL, Gupta N, et al. (2020). Flood inundation mapping- Kerala 2018; Harnessing the power of SAR, automatic threshold detection method and Google Earth Engine. PLoS ONE, 15(8): e0237324. https://doi.org/10.1371/journal.pone.0237324

Yousefi, E., Kaffash, M., & shariati, M. (2024). Spatial analysis of flood risk in Tabas watershed using satellite images and geographic information system (GIS). *Journal of Drought and Climate change Research*, 2(2), 105-118. doi: 10.22077/jdcr.2024.7489.1066

Yousefi, E., Sayadi, M. H., & Chamenhpour, E. (2022). Google Earth Engine platform to calculate the hydrometeorology and hydrological water balance of wetlands in arid areas and predict future changes. *Journal of Applied Research in Water and Wastewater*, 9(1), 52-68. doi: 10.22126/arww.2022.7033.1228

Zhang, Y., Li, X., & Wang, J. (2022). A Hybrid Rainfall-Runoff Model: Integrating Initial Loss and LSTM for Improved Forecasting. *Journal of Hydrology, 610*, 127-139. doi:10.1016/j.jhydrol.2022.127139.