

## Modeling the rainfall-runoff of Ajichai Basin using intelligent models

Pouya Allahverdipour<sup>1\*</sup> and Yagob Dinpashoh<sup>2</sup>

<sup>1</sup> PhD Student in Water Resources Engineering, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

<sup>2</sup> Professor, Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

Received: 21 May 2024

Accepted: 25 September 2024

### Extended abstract

#### Introduction

In recent years and decades, due to minor changes and challenges in sufficient and high-quality water resources, sustainable water resources have been the subject of various studies and research. The lack of safe and high-quality water resources is a major obstacle to sustainable development. For this reason, understanding the processes of the water cycle is very important and requires accurate information about hydrological phenomena. Runoff from water transfer is one of the main sources meeting various water demands, including agriculture, industry, and domestic use. The allocation of water resources to these sectors is planned based on runoff data at different times. A significant portion of precipitation in the hydrologic cycle is converted into runoff due to watershed characteristics. Considering the issue that the Lake Urmia Basin is shrinking, identifying the water resources of this basin and its sub-basins is crucial.

#### Materials and methods

The Ajichai Basin is one of the sub-basins of Lake Urmia. In this study, rainfall data from the Tabriz synoptic station and runoff data from the Nahand hydrometric station were used. The aim of this research is to model the daily rainfall-runoff of the Ajichai Basin using intelligent machine learning models, including Artificial Neural Network (ANN), Support Vector Machine (SVM), Gene Expression Programming (GEP), and Random Forest (RF). Seventy percent of the data was used for training, and 30% was used for testing the models. Statistical measures such as the Coefficient of Determination ( $R^2$ ), Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE), and Willmott Index (WI) were used to evaluate the performance of the models.

#### Results and discussion

The results of this research showed that all models performed very well in simulating rainfall-runoff in the Ajichai Basin. According to the obtained results, the GEP model, with  $R^2$  equal 0.84, RMSE equal 0.024  $m^3/s$ , NSE equal 0.864, and WI equal 0.968, was the most accurate in modeling rainfall-runoff in the Ajichai Basin. Based on scatter plots and time series analysis, the GEP model demonstrated higher accuracy than other models in predicting rainfall-runoff values with a high correlation.

#### Conclusions

According to the results, all the investigated models showed good capability in modeling daily rainfall-runoff in the Ajichai Basin. The findings of this research highlight the strong performance of machine learning models in rainfall-runoff modeling. In general, due to the high accuracy of intelligent models, particularly the GEP model, in predicting daily rainfall-runoff, it is recommended to apply these methods to hydrological problems. Additionally, for future research, it is suggested that intelligent methods and data mining techniques be used to model the precipitation-runoff process in different basins separately for drought-affected and wet years.

**Keywords:** Ajichai basin, Artificial neural network, Gene Expression Programming, GEP model, Rainfall-Runoff, Random Forest

Cite this article: Allahverdipour, P., Dinpashoh, Y., 2025. Modeling the rainfall-runoff of Ajichai Basin using intelligent models. *Water. Eng. Manag.* 17(1), 14-28.

\* Corresponding author: a.pouya@tabrizu.ac.ir

© 2025, The Author(s). Published by Soil Conservation and Watershed Management Research Institute (SCWMRI). This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)



## مدل سازی فرایند بارش-رواناب حوزه آبخیز آجی چای با استفاده از روش های هوشمند

پویا اللهویردی پور<sup>۱\*</sup> و یعقوب دین پژوه<sup>۲</sup>

<sup>۱</sup> دانشجوی دکتری مهندسی منابع آب، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

<sup>۲</sup> استاد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران

تاریخ پذیرش: ۱۴۰۳/۰۷/۰۴

تاریخ دریافت: ۱۴۰۳/۰۳/۰۱

### چکیده مبسوط

#### مقدمه

در سال ها و دهه های اخیر به دلیل تغییرات اقلیمی و مشکلات ایجاد شده در دسترسی به منابع آب کافی و با کیفیت، مسأله تأمین منابع آب پایدار مورد توجه محققان و پژوهشگران مختلف قرار گرفته است. کمبود منابع آب سالم و با کیفیت مانعی بزرگ در جهت توسعه پایدار محسوب می شود. به همین علت شناخت فرایندهای چرخه آب بسیار مهم بوده و نیازمند اطلاعات دقیق از پدیده های هیدرولوژیکی است. رواناب های حاصل از بارش یکی از منابع اصلی تأمین نیازهای آبی مختلف از جمله کشاورزی، صنعت و مصارف خانگی هستند. تخصیص منابع آب به این بخش ها با استفاده از داده های مربوط به رواناب در زمان های مختلف، برنامه ریزی می شود. برآورد و پیش بینی دقیق رواناب، اصلی ترین مسأله در استحصال منابع آب های سطحی و استفاده بهینه از پتانسیل حوضه های مختلف آبخیز است. حوزه آبخیز دریاچه ارومیه، به عنوان یکی از مهم ترین حوضه های آبی ایران، در سال های اخیر رو به خشکی نهاده است. با توجه به این موضوع، اطلاع و آگاهی در زمینه منابع آبی این حوضه و زیرحوضه های آن و نیز شبیه سازی و پیش بینی منابع مختلف ورودی به آن از جمله رواناب ها بسیار مهم است.

### مواد و روش ها

حوزه آبخیز آجی چای، یکی از زیرحوضه های دریاچه ارومیه است. این حوضه در ناحیه شمال غربی ایران و در استان آذربایجان شرقی قرار دارد. در این پژوهش از داده های بارش ایستگاه همدیدی تبریز و رواناب ایستگاه هیدرومتری نهند واقع در این حوضه از سال آبی ۷۷-۱۳۷۶ تا ۹۷-۱۳۹۶ استفاده شد. جهت مدل سازی بارش-رواناب از روش های شبکه عصبی مصنوعی (ANN)، برنامه نویسی بیان ژن (GEP)، ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) استفاده شد. ۷۰ درصد از داده ها برای آموزش و ۳۰ درصد از داده ها برای صحت سنجی مدل ها استفاده شد. برای ارزیابی عملکرد مدل ها از معیارهای آماری ضریب تعیین ( $R^2$ )، ریشه میانگین مربعات خطا (RMSE)، ضریب نش-ساتکلیف (NSE) و شاخص ویلموت (WI) استفاده شد.

### نتایج و بحث

نتایج این پژوهش نشان داد که تمامی مدل های مورد بررسی در این پژوهش عملکرد بسیار خوبی در شبیه سازی بارش-رواناب روزانه در حوضه آجی چای دارند. بر اساس نمودارهای پراکنش و سری های زمانی، مدل GEP در مدل سازی مقادیر بارش-رواناب این حوضه با همبستگی بالا نسبت به سایر مدل ها بیشترین دقت را داشت. با توجه به نتایج و معیارهای

\* مسئول مکاتبات: a.pouya@tabrizu.ac.ir

آماری، مدل GEP با مقادیر  $R^2$ ، RMSE، NSE و WI به ترتیب برابر ۰/۸۴، ۰/۰۲۴، ۰/۸۶۴ و ۰/۹۶۸، دقیق‌ترین مدل در شبیه‌سازی فرایند بارش-رواناب روزانه حوضه آبی‌چای بود. مدل‌های مورد بررسی در شبیه‌سازی رواناب‌های بیشینه خطا داشته و این مقادیر را کمتر برآورد کرده‌اند. این نکته می‌تواند به دلیل خاصیت مدل‌های هوشمند باشد که به داده‌های آموزش حساسیت بسیار زیادی دارند. از عوامل مؤثر در عملکرد مدل‌ها در برآورد رواناب‌های بیشینه می‌توان تعیین نوع و ساختار مدل‌ها را نام برد.

### نتیجه‌گیری

به‌طور کلی نتایج این پژوهش نشان داد که مدل GEP نسبت به سایر مدل‌های مورد بررسی، در شبیه‌سازی بارش-رواناب دقت بیشتر و عملکرد بهتری دارد. نتایج این پژوهش نشان‌دهنده عملکرد مناسب مدل‌های یادگیری ماشین در شبیه‌سازی فرایند بارش-رواناب است. به‌طور کلی با توجه به دقت بالای مدل‌های هوشمند به‌ویژه مدل GEP در پیش‌بینی فرایند بارش-رواناب روزانه، استفاده از این روش‌ها در مسائل هیدرولوژیک توصیه می‌شود. همچنین، پیشنهاد می‌شود، در تحقیقات آینده از روش‌های هوشمند و داده‌کاوی در مدل‌سازی فرایند بارش-رواناب در حوضه‌های مختلف به‌طور جداگانه برای سال‌های درگیر خشکسالی و ترسالی استفاده شود.

**واژه‌های کلیدی:** برنامه‌نویسی بیان ژن، حوضه آبی‌چای، جنگل تصادفی، شبکه عصبی مصنوعی، مدل GEP

### مقدمه

در سال‌ها و دهه‌های اخیر به دلیل تغییرات اقلیمی و مشکلات ایجاد شده در دسترسی به منابع آب کافی و با کیفیت، مسأله تأمین منابع آب پایدار مورد توجه محققان و پژوهشگران مختلف قرار گرفته است. کمبود منابع آب سالم و با کیفیت مانعی بزرگ در جهت توسعه پایدار محسوب می‌شود. به همین علت شناخت فرایندهای چرخه آب بسیار مهم بوده و نیازمند اطلاعات دقیق از پدیده‌های هیدرولوژیکی است. رواناب‌های حاصل از بارش یکی از منابع اصلی تأمین نیازهای آبی مختلف از جمله کشاورزی، صنعت و مصارف خانگی هستند. تخصیص منابع آب به این بخش‌ها با استفاده از داده‌های مربوط به رواناب در زمان‌های مختلف، برنامه‌ریزی می‌شود (Worland et al., 2018).

با توجه به وقوع خشکسالی‌های مداوم و تغییرات اقلیمی به‌وقوع پیوسته و در نتیجه آن کاهش منابع آبی در دهه‌های اخیر، پیش‌بینی میزان جریان رودخانه‌ها در حوضه‌های مختلف جهت مدیریت منابع آب مورد توجه قرار گرفته است. زیرا رواناب‌ها اصلی‌ترین منبع تغذیه آب مخازن و سدها هستند و مدیران منابع آب برای برنامه‌ریزی و مدیریت علمی منابع آب به مدل‌های

پیش‌بینی رواناب نیاز دارند و پیش‌بینی دقیق رواناب، اصلی‌ترین مسأله در استحصال منابع آب‌های سطحی و استفاده بهینه از پتانسیل حوزه‌های آبخیز است (Hussain and Khan, 2020).

تاکنون مدل‌های تجربی و فیزیکی متعددی برای برآورد رواناب ناشی از بارش توسعه یافته‌اند. این نوع مدل‌ها با توجه به تجربی بودن آنها، محدودیت‌های فراوانی دارند و کاربرد آنها نیز نیازمند کالیبراسیون (یا تصحیح ضرایب آنها) و صرف هزینه و زمان زیاد است. با توجه به پیچیده بودن مدل‌های فیزیکی و تجربی، عدم قطعیت این مدل‌ها، خاصیت غیرخطی روابط بارش-رواناب، نیاز به اطلاعات و صرف هزینه و زمان زیاد، محققان را به استفاده از مدل‌های هوشمند ترغیب کرده است. این روش‌ها، مبتنی بر داده هستند و نیازمند دانش و اطلاعات اولیه از پدیده مورد مطالعه نیستند (Ahmadian AhmadAbad and Zakeri Niri, 2021).

در دهه‌های اخیر استفاده از روش‌های هوشمند همچون شبکه‌های عصبی مصنوعی<sup>۱</sup>، ماشین بردار پشتیبان<sup>۲</sup>، برنامه‌نویسی بیان ژن<sup>۳</sup> و انواع مدل‌های

<sup>3</sup> Gene Expression Programming (GEP)

<sup>1</sup> Artificial Neural Network (ANN)

<sup>2</sup> Support Vector Machine (SVM)

برابر با ۰/۹۴ و RMSE برابر ۰/۸۵۳ از دقت بیشتری برخوردار بود.

Yoon et al., (2011) دو مدل ANN و SVM را برای مدل‌سازی تغییرات سطح آب زیرزمینی در یک آبخیز ساحلی در کشور کره جنوبی به کار گرفتند. نتایج این پژوهش حاکی از برتری مدل SVM نسبت به مدل ANN بود. Botsis et al., (2011) در پژوهشی فرایند بارش-رواناب روزانه در کالیفرنیا را با استفاده از مدل‌های ANN و SVM شبیه‌سازی کردند. مطابق نتایج، مدل SVM شبیه‌سازی دقیق‌تر از مدل ANN بود. Adamowski (2013) با هدف شبیه‌سازی بارش-رواناب حوضه دهرادان در یوتارانچال هندوستان، با مقایسه مدل‌های ANN و SVM، توانایی بالای مدل SVM را در پیش‌بینی رواناب این منطقه نشان داد.

Dehghani et al., (2016) با استفاده از مدل‌های ANN و ANFIS به مدل‌سازی بارش-رواناب روزانه حوزه آبخیز کسلیان در دوره آماری ۱۳۸۸-۱۳۸۲ پرداختند. با توجه به نتایج این پژوهش، مدل ANFIS با Rmod برابر ۰/۸۱، MAE برابر ۰/۱۱، RMSE برابر ۰/۰۵۶ و NSE برابر ۰/۸۰ نسبت به مدل ANN با Rmod برابر ۰/۸۷، MAE برابر ۰/۱۴، RMSE برابر ۰/۰۵۶ و NSE برابر ۰/۵۴ دقت بیشتر و خطای کمتری داشت. Vyas et al., (2016) به مدل‌سازی فرایند بارش-رواناب زیرحوضه آبخیز راجسمند واقع در حوضه رودخانه باناس راجستان در هند با استفاده از روش ANN و رگرسیون مرسوم<sup>۶</sup> و داده‌های بارش ماهانه این حوضه و رواناب ورودی سد راجسمند در دوره آماری ۲۰ ساله پرداختند. نتایج این پژوهش حاکی از برتری مدل ANN در پیش‌بینی مقادیر رواناب بود.

Patel and Joshi, (2017) با استفاده از مدل ANN به مدل‌سازی بارش-رواناب حوضه رودخانه سابارماتی هند با استفاده از داده‌های بارش و رواناب به مدت ۲۹ سال پرداختند. نتایج حاکی از آن بود که مدل ANN خطای بسیار پایینی داشته و از دقت بسیار بالا (با ضریب همبستگی برابر با ۰/۹۹) برخوردار است. Nourani (2017) عملکرد مدل FFNN را با مدل EANN<sup>۷</sup> در

درخت تصمیم<sup>۱</sup> مثل جنگل تصادفی<sup>۲</sup> و درخت تصادفی<sup>۳</sup> برای شبیه‌سازی فرایندهای مختلف نظیر برآورد دقیق رواناب‌ها مورد توجه قرار گرفته است. از جمله این تحقیقات می‌توان به کارهای علمی انجام شده توسط Shafeizadeh et al., Hassanzadeh et al., (2015) Abrahart et al., Antar et al., (2006)، (2019) Adnan et al., (2018)، (2012) اشاره نمود.

کمیته بررسی کاربرد شبکه‌های عصبی مصنوعی در هیدرولوژی در انجمن مهندسين عمران آمریکا<sup>۴</sup> با بررسی کاربردهای ANN در مباحث مختلف هیدرولوژی، روش ANN را به‌عنوان ابزاری مؤثر برای شبیه‌سازی بسیاری از فرایندهای غیرخطی هیدرولوژیکی مثل بارش-رواناب، آب زیرزمینی، کیفیت آب و بارش معرفی کرده است (ASCE, 2000).

در پژوهشی Nilsson et al., (2005) با استفاده از روش ANN و مدل مفهومی<sup>۵</sup> و ترکیب آنها، رواناب ماهانه خروجی از دو حوزه آبخیز در نروژ را شبیه‌سازی کرده و نتیجه گرفتند که مدل ANN با مقدار  $R^2$  برابر ۰/۸۲ نسبت به مدل مفهومی با مقدار  $R^2$  برابر ۰/۷۱ در هر دو حوضه، رواناب ماهانه را دقیق‌تر شبیه‌سازی می‌کند. Aytek et al., (2008) با استفاده از مدل‌های ANN و GEP بارش-رواناب روزانه حوضه جونیاتا در ایالت پنسیلوانیای کشور آمریکا را شبیه‌سازی کرده و نتیجه گرفتند که مدل GEP نسبت به ANN از دقت بیشتری برخوردار است.

Misra et al., (2009) با هدف پیش‌بینی رواناب و بار رسوبی با استفاده از داده‌های روزانه، هفتگی و ماهانه در یک حوزه آبخیز در هند، به مقایسه مدل‌های SVM و ANN پرداختند. نتایج نشان داد، در هر دو زمینه پیش‌بینی رواناب و بار رسوب، مدل SVM نسبت به ANN عملکرد بهتری دارد. در پژوهشی Eskandarinia et al., (2010) به پیش‌بینی جریان روزانه رودخانه بختیاری با استفاده از مدل‌های ANN و KNN و داده‌های بارش و رواناب در یک دوره ۲۱ ساله پرداختند. نتایج نشان داد که مدل ANN با مقدار R برابر با ۰/۹۵۷ و RMSE برابر ۱/۰۳ نسبت به مدل KNN با مقدار R

<sup>5</sup> Conceptual

<sup>6</sup> Conventional Regression

<sup>7</sup> Emotional ANN

<sup>1</sup> Decision Tree (DT)

<sup>2</sup> Random Forest (RF)

<sup>3</sup> Random Tree (RT)

<sup>4</sup> American Society of Civil Engineers (ASCE)

تحقیق، اختلاف دقت مدل‌ها در شبیه‌سازی رواناب این حوضه ناچیز بود و هر سه مدل عملکرد قابل قبول و نزدیک به هم داشتند. به‌طور کلی، مدل‌های ANN و ANFIS با NS برابر ۰/۶۶ توانایی بیشتری نسبت به مدل SVM داشتند.

Bigdeli et al., (2023) در پژوهشی به مدل‌سازی بارش-رواناب ایستگاه هیدرومتری آخولا در حوضه آجی چای با استفاده از مدل‌های RF و ANN و داده‌های بارش، دبی و دمای ماهانه ۲۰ ساله پرداختند. نتایج نشان داد که مدل ANN با CC برابر ۰/۸۶ نسبت به مدل RF با CC برابر ۰/۸۲ دقت بیشتری داشت. Allahverdi-pour and Sattari, (2023) به مدل‌سازی بارش سالانه ایستگاه همدیدی اهواز با استفاده از روش‌های رگرسیون خطی چندگانه<sup>۲</sup>، تحلیل مولفه‌های اصلی<sup>۳</sup>، GEP و SVM پرداختند. نتایج این پژوهش نشان داد که روش‌های PCA و GEP با  $R^2$  برابر ۰/۸۵، NSE برابر ۰/۸۵ و WI برابر ۰/۹۶ و اختلاف بسیار ناچیز در مقادیر RMSE به‌ترتیب برابر با ۳۵/۴۹ و ۳۵/۷۰ میلی‌متر نسبت به سایر مدل‌ها عملکرد بهتر و دقت بیشتری دارند.

حوزه آبخیز دریاچه ارومیه، به‌عنوان یکی از مهم‌ترین حوضه‌های آبی ایران، در سال‌های اخیر رو به خشکی نهاده است. با توجه به این موضوع به‌نظر می‌رسد، اطلاع و آگاهی در زمینه منابع آبی این حوضه و شبیه‌سازی و پیش‌بینی منابع مختلف ورودی به آن از جمله رواناب‌ها بسیار مهم است. نوآوری این پژوهش، استفاده از چند مدل هوشمند برای مدل‌سازی فرایند بارش-رواناب، آن‌هم در مقیاس زمانی روزانه است. بنابراین، در این پژوهش فرایند بارش-رواناب حوضه آجی‌چای، یکی از زیرحوضه‌های دریاچه ارومیه با استفاده از روش‌های هوشمند شامل ANN، GEP، SVM و RF مدل‌سازی شده است.

### مواد و روش‌ها

منطقه مورد مطالعه و داده‌های مورد استفاده: حوزه آبخیز آجی‌چای یکی از زیرحوضه‌های دریاچه ارومیه است. این حوضه در ناحیه شمال‌غربی ایران و در استان

مدل‌سازی فرایند بارش-رواناب دو زیرحوضه (شامل لایز هول کریک در حوضه رودخانه مورومبیچی در کشور استرالیا و موزل در حوضه رودخانه راین در قاره اروپا) با دو شرایط آب و هوایی متفاوت مقایسه کرد. نتایج این پژوهش نشان داد که مدل EANN در زیرحوضه لایز هول کریک به‌ترتیب با  $R^2$  برابر ۰/۹۲ و RMSE برابر ۰/۰۲۴ نسبت به مدل FFNN با  $R^2$  برابر ۰/۸۵ و RMSE برابر ۰/۰۳۸ و در زیرحوضه موزل با  $R^2$  برابر ۰/۷۴ و RMSE برابر ۰/۱۸۲ نسبت به مدل FFNN با  $R^2$  برابر ۰/۶۴۸ و RMSE برابر ۰/۲۰۳ عملکرد بهتری داشته است.

Sepelri et al., (2016) برای برآورد دبی سیلاب ایستگاه هیدرومتری یلفان واقع در مسیر رودخانه یلفان از مدل ANN و شبکه عصبی بهینه‌سازی شده با الگوریتم ژنتیک استفاده کردند. با توجه به نتایج این پژوهش مدل GANN با  $R^2$  برابر ۰/۸۷ و RMSE برابر ۱۰/۱۲ نسبت به مدل ANN با  $R^2$  برابر ۰/۵۸ و RMSE برابر ۱۳/۴۳ عملکرد بهتری داشت. Zeinalie et al., (2019) به ارزیابی مدل‌های برنامه‌نویسی بیان ژن، شبکه بیزین<sup>۱</sup> و ماشین بردار پشتیبان در مدل‌سازی جریان رودخانه گاماسیاب نه‌اوند با استفاده از داده‌های بارش و جریان روزانه در یک دوره ۱۰ ساله (۱۳۹۱-۱۳۸۱) پرداختند. با توجه به نتایج، مدل GEP با  $R^2$  برابر ۰/۹۲۳ و NSE برابر ۰/۹۰۲ به‌عنوان مدل برتر انتخاب شد.

Hussain and Khan, (2020) با هدف پیش‌بینی جریان رودخانه هنزه در پاکستان از مدل‌های ANN از نوع پرسپترون چندلایه (MLP)، SVM و RT در دوره آماری ۲۰۰۸-۱۹۶۲ استفاده کردند. نتایج نشان داد که مدل RF نسبت به مدل‌های MPL و SVM با توجه به آماره‌های  $R^2$  به‌ترتیب برابر با ۰/۹۹۳، ۰/۹۱۰ و ۰/۸۳۱، RMSE برابر با ۰/۰۶۹، ۰/۰۸۴ و ۰/۱۰۴، مترمکعب بر ثانیه و MAE برابر با ۰/۰۴۰، ۰/۰۵۸ و ۰/۰۶۲ مترمکعب بر ثانیه برتری داشت. Ahmadian AhmadAbad and Zakeri Niri, (2021) به شبیه‌سازی و برآورد رواناب حوزه آبخیز دز با استفاده از روش‌های ANN، SVM و ANFIS پرداختند. با توجه به نتایج این

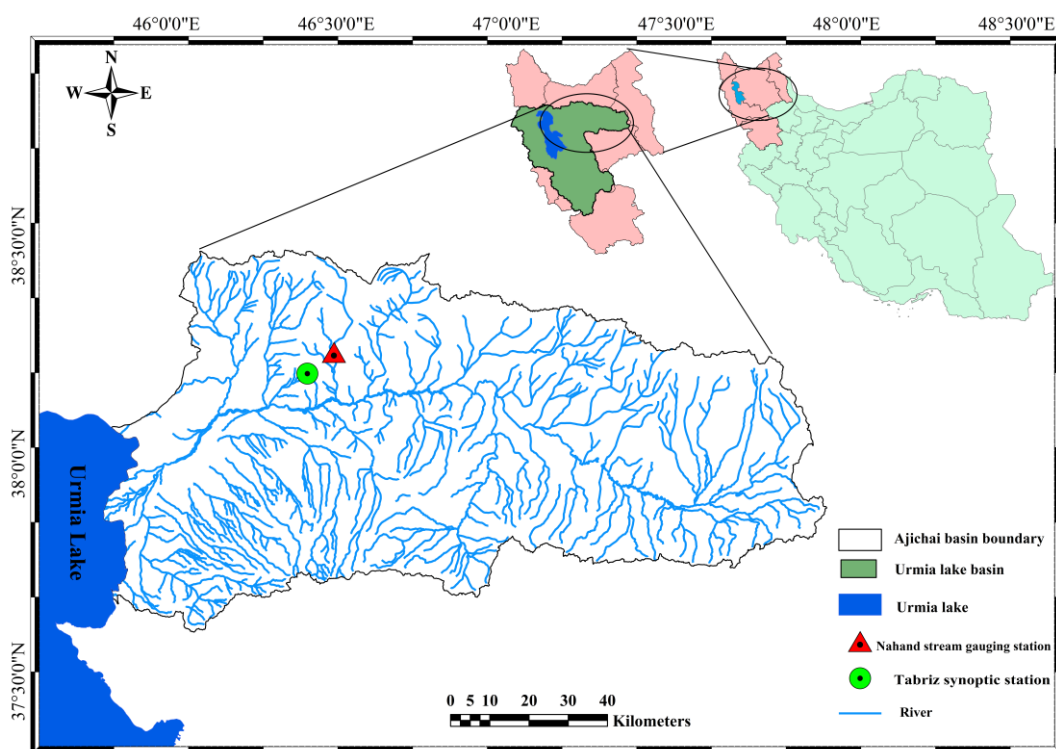
<sup>3</sup> Principal Component Analysis (PCA)

<sup>1</sup> Bayesian Network (BN)

<sup>2</sup> Multiple Linear Regression (MLR)

ایستگاه در طول جغرافیایی ۴۶ درجه و ۲۴ دقیقه شرقی و عرض جغرافیایی ۳۸ درجه و ۱۲ دقیقه شمالی و ارتفاع ۱۳۶۱ متر از سطح دریا واقع است. همچنین، از داده‌های رواناب ایستگاه هیدرومتری نهند واقع در طول جغرافیایی ۴۶ درجه و ۲۸ دقیقه شرقی و عرض جغرافیایی ۳۸ درجه و ۱۴ دقیقه شمالی و ارتفاع ۱۶۲۴ متر از سطح دریا در حوزه آبخیز آجی‌چای استفاده شده است. میانگین بارش و دمای ایستگاه همدیدی تبریز به ترتیب ۲۵۵/۵۸ میلی‌متر و ۱۳/۲۹ درجه سانتی‌گراد و میانگین رواناب ایستگاه هیدرومتری نهند ۰/۹۳۵ مترمکعب بر ثانیه است. موقعیت منطقه مورد مطالعه در شکل ۱ نشان داده شده است.

آذربایجان شرقی و در موقعیت جغرافیایی ۴۵ درجه و ۳۰ دقیقه تا ۴۷ درجه و ۴۵ دقیقه طول شرقی و ۳۷ درجه و ۴۵ دقیقه تا ۳۸ درجه و ۳۰ دقیقه عرض شمالی قرار دارد. این حوضه از شمال به حوزه آبخیز اهرچای، از جنوب به حوضه آبخیز قزل‌اوزن و صوفی‌چای، از شرق به حوزه آبخیز بالخلوچای و از غرب به دریاچه ارومیه محدود می‌شود. مساحت این حوضه حدود ۱۱۴۵۷ کیلومترمربع و شیب متوسط آن ۸/۷۲ درصد است. مرتفع‌ترین نقطه حوزه آبخیز آجی‌چای ۳۸۵۰ متر و پست‌ترین نقطه آن ۱۲۷۴ متر بالاتر از سطح دریا قرار دارد (Sanikhani et al., 2014). براساس طبقه‌بندی دومارتن، اقلیم حوضه آجی‌چای نیمه‌خشک است. در این مطالعه از داده‌های بارش ایستگاه همدیدی تبریز بهره گرفته شد. این



شکل ۱- موقعیت منطقه مورد مطالعه

Fig. 1. The location of the study area

شود، صحت‌سنجی داده‌ها است. واضح است که بدون اطمینان از صحت و کیفیت داده‌ها، نمی‌توان از آنها برای استخراج نتایج بعدی استفاده کرد (Allahverdi-pour and Sattari, 2023). داده‌های مورد استفاده در این پژوهش قبل از انجام فرایند مدل‌سازی مورد بررسی آماری و صحت‌سنجی قرار گرفت و همچنین نسبت به

در این پژوهش، با هدف مدل‌سازی فرایند بارش-رواناب حوزه آبخیز آجی‌چای در مقیاس زمانی روزانه، از داده‌های بارش و رواناب در یک دوره ۲۱ ساله از سال آبی ۱۳۷۶-۷۷ تا سال آبی ۱۳۹۶-۹۷ استفاده شده است. لازمه هر مطالعه هیدرولوژیکی و منابع آب که در آن از داده‌های هواشناسی و هیدرولوژی استفاده می

برای بخش صحت‌سنجی استفاده شد. در جدول ۱، مشخصات آماری متغیرهای مورد استفاده آمده است.

تکمیل داده‌های گم‌شده و حذف داده‌های پرت اقدام شد. ۷۰ درصد از داده‌ها برای بخش آموزش و ۳۰ درصد

جدول ۱- خلاصه مشخصات آماری متغیرهای مورد استفاده در مقیاس روزانه

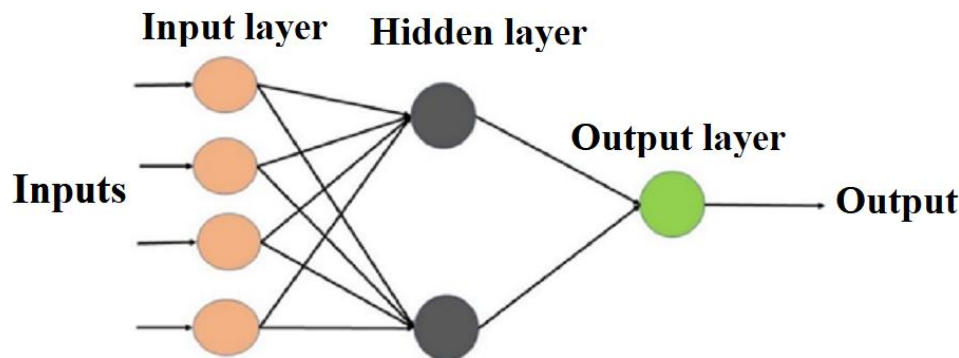
Table 1. Summary statistics of used variables in daily scale

| Variable | Unit              | Minimum | Maximum | Mean | Standard Deviation |
|----------|-------------------|---------|---------|------|--------------------|
| Rainfall | mm                | 0       | 58      | 0.68 | 2.52               |
| Runoff   | m <sup>3</sup> /s | 0       | 25.60   | 0.94 | 1.85               |

لایه یا لایه‌های ورودی، یک یا چند لایه پنهان و تنها یک لایه خروجی هستند. سیگنال مربوط به داده‌های ورودی مساله، در طول شبکه و رو به جلو به صورت لایه به لایه از لایه ورودی به لایه پنهان و از آنجا به لایه خروجی انتشار می‌یابند که با توجه به این نحوه رفتار و عملکرد، به این نوع شبکه‌ها، پرسپترون چندلایه (MLP) گفته می‌شود (Jain et al., 1999).

در شکل ۲ ساختار یک شبکه عصبی مصنوعی از نوع پرسپترون چندلایه (MLP) نشان داده شده است.

شبکه‌های عصبی مصنوعی از نوع پرسپترون چندلایه (MLP): شبکه‌های عصبی مصنوعی (ANN) از روش‌های هوشمند است که با توجه به سیستم عصبی انسان توسعه یافته‌اند. روش ANN با پردازش و بررسی داده‌های مربوط به مشاهدات، نحوه ارتباط و به هم پیوستگی پدیده مورد نظر را آموزش دیده و یاد می‌گیرد و این مدل آموزش دیده شده، برای پیش‌بینی رویدادهای آینده استفاده می‌شود. این شبکه‌ها از مجموعه‌ای از واحدهای حسی که نورون‌های پایه نامیده می‌شوند، تشکیل شده و شامل



شکل ۲- ساختار یک شبکه عصبی مصنوعی از نوع پرسپترون پیشرو سه لایه  
Fig. 2. The structure of a three-layer feed forward perceptron artificial neural network

نگاشت غیرخطی برای آموزش دادن داده‌های هدف استفاده می‌کند. در حالت پیش‌بینی، الگوهای ورودی به فضایی با ابعاد بالا نگاشت می‌شوند، به طوری که می‌توان آنها را به صورت خطی در فضای نگاشت شده پسرفت کرد. دقت رگرسیون بردار پشتیبان با توجه به کمینه کردن تابع خطای بدست آمده کنترل می‌شود. ماشین‌های بردار پشتیبان ابعاد مسأله را از طریق توابع هسته برای حل مسائل غیرخطی تغییر می‌دهند. انتخاب هسته در روش SVM به اندازه و طول داده‌های آموزشی و همچنین ابعاد بردار ویژگی داده‌ها بستگی

ماشین بردار پشتیبان (SVM): ماشین بردار پشتیبان مجموعه‌ای از روش‌های یادگیری تحت نظارت است که توسط Vapnik and Chervonenkis, (2015) بر اساس نظریه یادگیری آماری معرفی شده است. این روش معمولاً برای طبقه‌بندی دوگانه در فضای ویژگی‌های دلخواه اجرا می‌شود و بنابراین برای مسائل پیش‌بینی کاربردی است.

روش‌های SVM به دو گروه اصلی شامل الف) طبقه‌بندی و ب) پیش‌بینی تقسیم می‌شوند. روش دوم یک روش جدید مبتنی بر تئوری آماری است که از



$$g(x) = \sum_{i=1}^{nl} (a_i - \bar{a}_i) K < \phi(x_i), \phi(x_j) > + b \quad (5)$$

در رابطه (۵)، nl تعداد بردارهای پشتیبان است. در این رابطه ضرب داخلی محاسبه شده در فضای ویژگی با استفاده از تابع کرنل انجام می‌شود. یعنی اگر دو بردار  $\phi(x_i)$  و  $\phi(x_j)$  وجود داشته باشد، ضرب داخلی آنها به وسیله تابع کرنل مورد نظر صورت می‌گیرد (Vapnik, 1995).

**برنامه‌نویسی بیان ژن (GEP):** برنامه‌نویسی بیان ژن نوع خاصی از برنامه‌نویسی ژنتیک<sup>۱</sup> (GP) چندشاخه‌ای است که با ایجاد درخت‌های (کروموزوم‌ها) حاوی یک یا چندین ژن، که هر کدام یک درخت فرعی را ایجاد می‌کنند، به حل مسأله مورد نظر می‌پردازد. در روش GEP، متغیر خروجی با پیوند دادن زیر درخت‌های مربوطه با استفاده از توابع جبری یا بولی (Or، And، Not) محاسبه می‌شود. در GP برنامه‌های کامپیوتری (ژن‌ها) از زبان LISP پیروی می‌کنند و به صورت درختان تجزیه با اندازه‌ها و شکل‌های مختلف بیان می‌شوند.

اما در روش GEP برنامه‌های کامپیوتری به‌عنوان رشته‌های خطی با طول ثابت متشکل از یک یا چند ژن در نظر گرفته می‌شوند. هر ژن شامل یک سر و دم است، بنابراین، برای تعداد معین ژن و طول سر، جمعیت اولیه کروموزوم‌های GEP از طریق پرکردن دامنه‌های سر و دم با توجه به توالی کدگذاری ژن‌ها به نام چارچوب خوانش باز ایجاد می‌شوند.

وقتی که الگوریتم GEP راه‌حل کامل مربوط به مسأله مورد نظر را پیدا کند، به این معنی که برازش بهترین کروموزوم به مقدار مشخص شده برسد، یا تعداد ژن‌های تولید شده بدون در نظر گرفتن دقت و کیفیت جواب به‌دست آمده به تعداد معینی رسیده باشد، فرایند تکامل و نتیجه‌گیری در روش GEP به اتمام می‌رسد. در این روش مجموعه توابع، معمولاً شامل توابع حسابی، مثلثاتی و یا توابع تعریف شده توسط کاربر هستند که به‌طور دلخواه و با توجه به ماهیت مسأله تعیین می‌شوند (Ferreira, 2002).

**جنگل تصادفی (RF):** روش جنگل تصادفی (RF) یک روش یادگیری یکپارچه است که بر روی درخت تصمیم

دارد. انتخاب تابع هسته باید با توجه به این موارد باشد تا از قابلیت آموزش ورودی‌های مسأله برخوردار شود (Pai and Hong, 2007).

در روش SVM معادله رگرسیون خطی که در فضای ویژگی با بعد بالا به داده‌ها برازش داده می‌شود، با فرض اینکه به تعداد N سری داده آموزشی به صورت  $\{(x_i, d_i)\}_i^N$  موجود است، طبق رابطه ۱، محاسبه می‌شود.

$$y_i = f(x) = \langle W_i, \phi_i(x) \rangle + b \quad (1)$$

در رابطه بالا، پارامترهای  $W_i$  و  $b$  ضرایب رابطه در فضای ویژگی در بعد بالا بوده و  $\phi_i(x)$  تابعی است که داده‌ها را از فضای ورودی به فضای ویژگی افزایش می‌دهد. در این روش نیز همانند سایر روش‌ها، جهت تخمین پارامترهای این مدل باید یک معیار خطا در نظر گرفته شود، در این رابطه توابع خطای مختلف مثل خطی، چندجمله‌ای، توانی و غیره استفاده می‌شود. از معیار  $\epsilon$  برای بهینه‌سازی پارامترهای مدل SVM استفاده می‌شود (رابطه ۲).

$$L_\epsilon(v, g(u)) = \begin{cases} 0 & \text{for } |v - g(u)| \leq \epsilon \\ |d - g(u)| - \epsilon & \text{otherwise} \end{cases} \quad (2)$$

با در نظر گرفتن این معیار خطا، تخمین پارامترهای رابطه رگرسیونی به مسأله بهینه‌سازی به صورت رابطه‌های (۳) و (۴) تبدیل می‌شود.

$$\min \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \xi_i + \xi_i^* \quad (3)$$

$$\text{subject to } \begin{cases} V_i - (w \cdot \phi(u_i) + b) \leq \epsilon + \xi_i \\ (w \cdot \phi(u_i) + b) - V_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i=1, 2, \dots, n \end{cases} \quad (4)$$

در رابطه‌های (۳) و (۴) پارامترهای  $\xi_i$  و  $\xi_i^*$  ضرایب کمبود هستند که به ترتیب حد بالا و پایین خطای آموزش مدل را تعیین می‌کنند، پارامترهای  $C$  و  $\epsilon$  نیز به ترتیب، پارامتر تنظیم‌کننده و حد بالای خطای خروجی هستند. با در نظر گرفتن ضرایب لاگرانژ و شرایط کان-تاکر، فرم دوگانه مسأله بهینه‌سازی بالا تبدیل به مسأله بهینه‌سازی برنامه‌ریزی چندجمله‌ای درجه دو می‌شود که با بهینه‌سازی آن ضرایب لاگرانژ  $\bar{a}_i$  و بردار وزن‌های بهینه ابر صفحه رگرسیون محاسبه شده و در آخر برای تعیین بردارهای پشتیبان، رابطه رگرسیون به صورت رابطه (۵) بیان می‌شود.

<sup>1</sup> Genetic Programming

اجرای موازی این مدل‌ها اشاره کرد (Cutler et al., 2012).

در این پژوهش، ۱۶ سناریو مختلف با لحاظ کردن ترکیب‌های مختلف ورودی و با استفاده از تأخیر زمانی در پارامترهای بارش و رواناب مورد استفاده قرار گرفته است. جدول ۲، سناریوهای مختلف ورودی به مدل‌ها را در پژوهش حاضر نشان می‌دهد. در این جدول، نمادهای  $Q(t)$ ،  $Q(t-1)$ ،  $Q(t-2)$  و  $Q(t-3)$  سری زمانی رواناب‌های روز  $t$ ، یک روز قبل، دو روز قبل و سه روز قبل از آن و پارامترهای  $P(t)$ ،  $P(t-1)$ ،  $P(t-2)$  و  $P(t-3)$  نیز بارش‌های روز  $t$ ، یک روز قبل، دو روز قبل و سه روز قبل از آن را نشان می‌دهند.

برای بهبود عملکرد پیش‌بینی و تعمیم، با ترکیب درخت‌های تصمیم‌گیری چندگانه ساخته شده است. در یک مسأله طبقه‌بندی، RF طبقه‌بندی نهایی را از طریق مکانیزم رای‌گیری تعیین می‌کند. در مسائل رگرسیونی، مدل RF مقدار نهایی را با در نظر گرفتن میانگین مقادیر پیش‌بینی شده توسط همه درخت‌های تصمیم، تعیین می‌کند (Yin, 2023). از نظر محاسباتی، مدل‌های RF بسیار کاربردی هستند، زیرا هم برای رگرسیون و هم برای طبقه‌بندی (چند طبقه مختلف) می‌توانند به کار گرفته شوند.

همچنین از دیگر ویژگی‌های روش RF می‌توان به آموزش و پیش‌بینی نسبتاً سریع، بستگی به یک یا دو پارامتر تنظیم، برآورد داخلی از خطای تعمیم و قابلیت استفاده برای مسائل با ابعاد بالا و پیچیده و قابلیت

جدول ۲- شانزده سناریوی مختلف برای ترکیب‌های ورودی در پژوهش حاضر

Table 2. Sixteen scenarios of input combinations used in the present study

| Scenario | Inputs combination                                     |
|----------|--|
| 1        | $P(t)$   |
| 2        | $P(t), P(t-1)$   |
| 3        | $P(t), P(t-1), P(t-2)$                                 |
| 4        | $P(t), P(t-1), P(t-2), P(t-3)$                         |
| 5        | $Q(t-1), P(t)$   |
| 6        | $Q(t-1), P(t), P(t-1)$                                 |
| 7        | $Q(t-1), P(t), P(t-1), P(t-2)$                         |
| 8        | $Q(t-1), P(t), P(t-1), P(t-2), P(t-3)$                 |
| 9        | $Q(t-1), Q(t-2), P(t)$                                 |
| 10       | $Q(t-1), Q(t-2), P(t), P(t-1)$                         |
| 11       | $Q(t-1), Q(t-2), P(t), P(t-1), P(t-2)$                 |
| 12       | $Q(t-1), Q(t-2), P(t), P(t-1), P(t-2), P(t-3)$         |
| 13       | $Q(t-1), Q(t-2), Q(t-3), P(t)$                         |
| 14       | $Q(t-1), Q(t-2), Q(t-3), P(t), P(t-1)$                 |
| 15       | $Q(t-1), Q(t-2), Q(t-3), P(t), P(t-1), P(t-2)$         |
| 16       | $Q(t-1), Q(t-2), Q(t-3), P(t), P(t-1), P(t-2), P(t-3)$ |

در رابطه‌های (۶) تا (۹) پارامتر  $O(i)$  مقادیر مشاهداتی (واقعی)،  $P(i)$  مقادیر شبیه‌سازی شده،  $\bar{O}$  میانگین مقادیر مشاهداتی،  $\bar{P}$  میانگین مقادیر شبیه‌سازی شده و  $N$  تعداد مشاهدات است. هرچه مقادیر  $R^2$ ،  $NSE$  و  $WI$  بیشتر و  $RMSE$  کمتر باشد، دقت و عملکرد مدل بهتر تلقی می‌شود.

### نتایج و بحث

با در نظر گرفتن سناریوهای مختلف ورودی به مدل‌ها که در جدول ۲ شرح داده شد، نتایج ارزیابی بهترین ساختار ورودی مربوط به هر یک از مدل‌های مورد بررسی در جدول ۳ آمده است.

معیارهای ارزیابی: برای بررسی عملکرد مدل‌های مورد استفاده در این پژوهش از معیارهای آماری ضریب تعیین ( $R^2$ )، جذر میانگین مربعات خطاها ( $RMSE$ )، ضریب نش-ساتکلیف ( $NSE$ ) و شاخص ویلموت ( $WI$ ) استفاده شد. این معیارها به ترتیب، با رابطه‌های (۶) تا (۹) ارائه شده‌اند.

$$R^2 = \frac{[\sum_{i=1}^N (O_i - \bar{O})(P_i - \bar{P})]^2}{\sum_{i=1}^N (O_i - \bar{O})^2 \sum_{i=1}^N (P_i - \bar{P})^2} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (8)$$

$$WI = 1 - \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (|P_i - \bar{O}_i| + |O_i - \bar{P}_i|)^2}} \quad (9)$$

جدول ۳- نتایج ارزیابی مدل‌های مورد استفاده در پژوهش حاضر در مراحل آموزش و صحت‌سنجی  
**Table 3.** Evaluation results of the used models in the present study in training and testing stages

| Model | Best scenario | Stage | WI    | NSE   | RMSE (m <sup>3</sup> /s) | R <sup>2</sup> |
|-------|---------------|-------|-------|-------|--------------------------|----------------|
| ANN   | 5             | Train | 0.999 | 0.999 | 0.002                    | 0.85           |
|       |               | Test  | 0.967 | 0.865 | 0.024                    | 0.66           |
| SVM   | 10            | Train | 0.999 | 0.999 | 0.002                    | 0.85           |
|       |               | Test  | 0.968 | 0.869 | 0.024                    | 0.81           |
| GEP   | 10            | Train | 0.999 | 0.999 | 0.002                    | 0.85           |
|       |               | Test  | 0.968 | 0.864 | 0.024                    | 0.84           |
| RF    | 8             | Train | 0.999 | 0.999 | 0.002                    | 0.83           |
|       |               | Test  | 0.968 | 0.865 | 0.024                    | 0.74           |

آجی‌چای برخوردار است. پس از آن، به ترتیب مدل SVM با مقادیر R<sup>2</sup> برابر ۰/۸۱، RMSE برابر ۰/۰۲۴، NSE برابر ۰/۸۶۹ و WI برابر ۰/۹۶۸، مدل RF با مقادیر R<sup>2</sup> برابر ۰/۷۴، RMSE برابر ۰/۰۲۴، NSE برابر ۰/۸۶۵ و WI برابر ۰/۹۶۷، مدل ANN با مقادیر R<sup>2</sup> برابر ۰/۶۶، RMSE برابر ۰/۰۲۴، NSE برابر ۰/۸۶۵ و WI برابر ۰/۹۶۷، در جایگاه بعدی قرار دارند.

برای مقایسه بهتر نتایج حاصل از مدل‌سازی، در جدول ۴ مقادیر آماره‌های توصیفی شامل گرایش به مرکز و پراکندگی رواناب مشاهداتی و مدل‌سازی شده ارائه شده است.

با توجه به جدول ۳، می‌توان نتیجه گرفت که در مرحله صحت‌سنجی همه مدل‌ها ضریب تعیین (R<sup>2</sup>) بیش از ۰/۶۶ دارند. بنابراین، می‌توان به دقت مناسب همه مدل‌های هوشمند مورد بررسی در این پژوهش پی برد. نتایج حاکی از قابلیت قابل قبول این مدل‌ها در شبیه‌سازی فرایند بارش-رواناب حوضه مورد مطالعه است.

مطابق جدول ۳ می‌توان نتیجه گرفت که مدل GEP در مرحله صحت‌سنجی با معیارهای آماری R<sup>2</sup> برابر ۰/۸۴، RMSE برابر ۰/۰۲۴، NSE برابر ۰/۸۶۴ و WI برابر ۰/۹۶۸ نسبت به مدل‌های دیگر از بیشترین دقت و کمترین خطا در مدل‌سازی بارش-رواناب حوضه

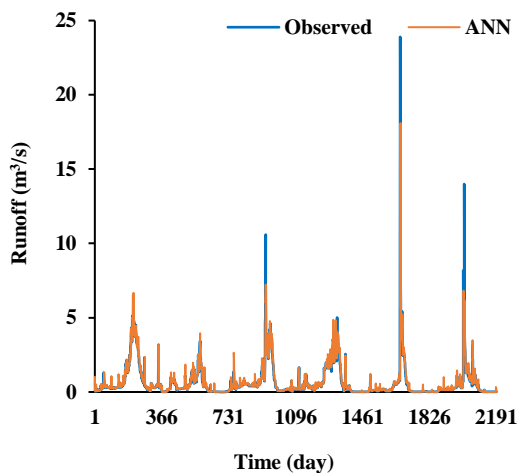
جدول ۴- آماره‌های توصیفی رواناب مشاهداتی و مدل‌سازی شده  
**Table 4.** Descriptive statistics of observed and simulated runoff

| Model    | Stage    | Mean (m <sup>3</sup> /s) | Range (m <sup>3</sup> /s) | Standard Deviation (m <sup>3</sup> /s) |
|----------|----------|--------------------------|---------------------------|--|
| Observed | Training | 1.05                     | 25.60                     | 2.03                                   |
|          | Testing  | 0.65                     | 23.90                     | 1.23                                   |
| ANN      | Training | 1.05                     | 17.80                     | 1.86                                   |
|          | Testing  | 0.68                     | 18.10                     | 1.08                                   |
| SVM      | Training | 1.05                     | 16.06                     | 1.86                                   |
|          | Testing  | 0.69                     | 16.02                     | 1.14                                   |
| GEP      | Training | 1.05                     | 17.80                     | 1.86                                   |
|          | Testing  | 0.70                     | 31.21                     | 1.32                                   |
| RF       | Training | 1.05                     | 17.44                     | 1.86                                   |
|          | Testing  | 0.69                     | 16.86                     | 1.15                                   |

و این امر به دلیل عدم توانایی مدل‌ها در برآورد دبی‌های بیشینه است. به طوری که همه مدل‌ها دبی‌های بیشینه را کمتر برآورد کرده‌اند. این مسأله از ضعف‌های اساسی این مدل‌های هوشمند به شمار می‌رود.

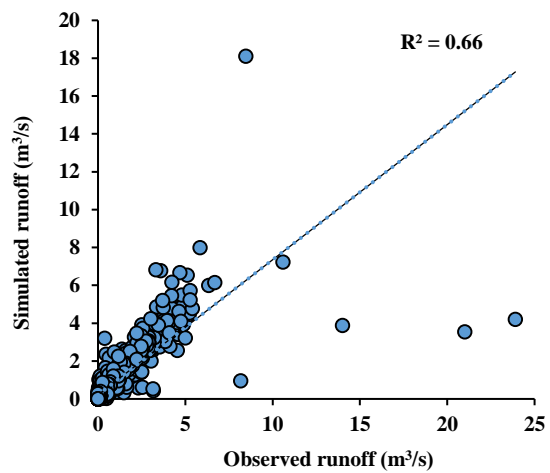
شکل‌های ۳ تا ۱۰ نمودار سری زمانی و پراکنش رواناب مشاهداتی و مدل‌سازی شده با استفاده از مدل‌های مورد استفاده را در مرحله صحت‌سنجی نشان می‌دهند.

مطابق جدول ۴، مشاهده می‌شود که در مرحله صحت‌سنجی مقدار بسیار اندکی تفاوت در میانگین رواناب حاصل از خروجی مدل‌ها نسبت به داده‌های مشاهداتی وجود دارد. ولی در کل در هر دو مرحله آموزش و صحت‌سنجی، میانگین رواناب مشاهداتی و مدل‌سازی شده نزدیک به هم هستند. مشاهده می‌شود که انحراف معیار و همچنین دامنه تغییرات رواناب مدل‌سازی شده تفاوت زیادی نسبت به مقادیر مشاهداتی در هر دو مرحله آموزش و صحت‌سنجی دارند



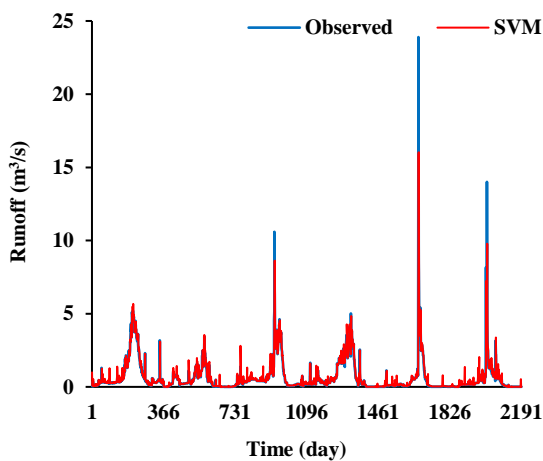
شکل ۴- سری زمانی رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (ANN)

Fig. 4. Time series of observed and simulated runoff in the test phase (ANN)



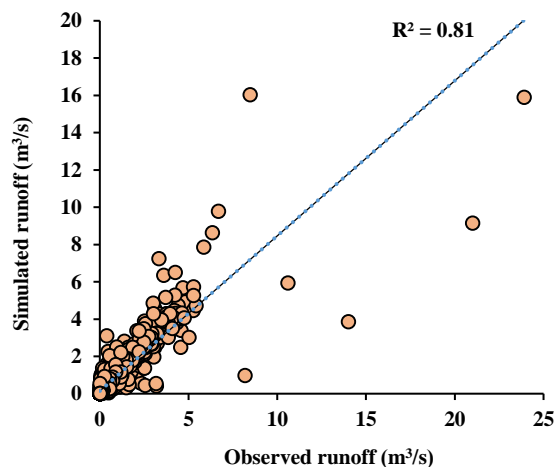
شکل ۳- نمودار پراکنش رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (ANN)

Fig. 3. Scatter Plot of observed and simulated runoff in the test phase (ANN)



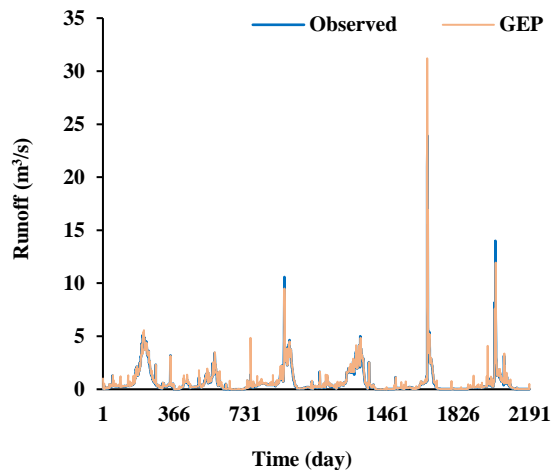
شکل ۶- سری زمانی رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (SVM)

Fig. 6. Time series of observed and simulated runoff in the test phase (ANN)



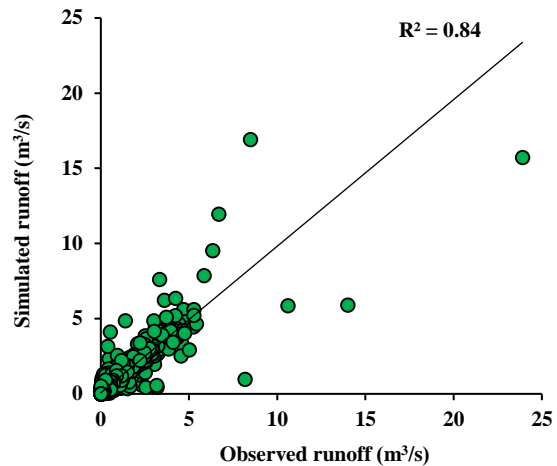
شکل ۵- نمودار پراکنش رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (SVM)

Fig. 5. Scatter Plot of observed and simulated runoff in the test phase (SVM)



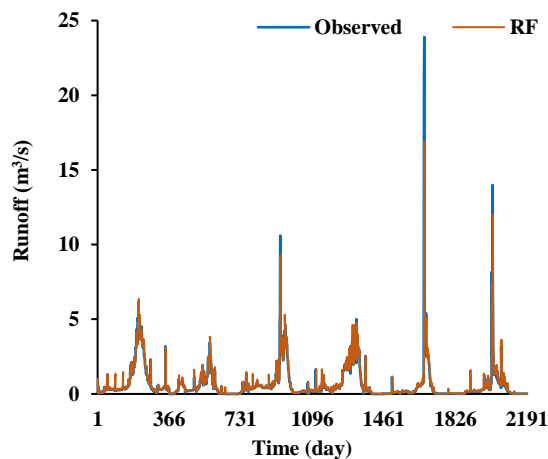
شکل ۸- سری زمانی رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (GEP)

Fig. 8. Time series of observed and simulated runoff in the test phase (GEP)



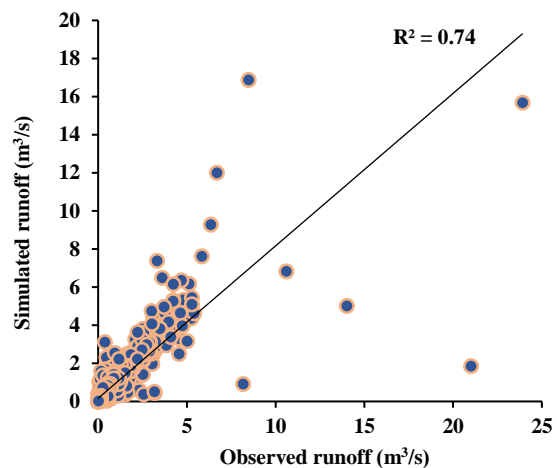
شکل ۷- نمودار پراکنش رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (GEP)

Fig. 7. Scatter Plot of observed and simulated runoff in the test phase (GEP)



شکل ۱۰- سری زمانی رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (RF)

Fig. 10. Time series of observed and simulated runoff in the test phase (RF)



شکل ۹- نمودار پراکنش رواناب مشاهداتی و شبیه‌سازی شده در مرحله صحت‌سنجی (RF)

Fig. 9. Scatter Plot of observed and simulated runoff in the test phase (RF)

همچنین، عملکرد مناسب مدل‌های هوشمند در پژوهش Ahmadian AhmadAbad and Zakeri Niri, (2021) نیز حاصل شده بود.

مشاهده می‌شود که مدل‌های مورد بررسی در شبیه‌سازی رواناب‌های بیشینه (بالا تر از ده مترمکعب بر ثانیه) خطا دارند و این مقادیر را کمتر از مقدار واقعی برآورد کرده‌اند. این نکته می‌تواند به دلیل خاصیت مدل‌های هوشمند یادگیری ماشین باشد که به داده‌های آموزش حساسیت بسیار زیادی دارند و تعداد داده‌های کمتر با این مقادیر، موجب دقت کمتر این مدل‌ها در بخش صحت‌سنجی می‌شود. از عوامل مؤثر در عملکرد مدل‌ها در برآورد رواناب‌های بیشینه می‌توان تعیین نوع

با توجه به شکل‌های ۳ تا ۱۰ مشاهده می‌شود که همه مدل‌های مورد بررسی در این تحقیق عملکرد بسیار مناسبی در شبیه‌سازی بارش-رواناب حوزه آجی‌چای داشته‌اند. این نتایج با نتایج حاصل از پژوهش Hussain and Khan, (2020) که دقت بالای مدل‌های هوشمند RF، MPL و SVM را در پیش‌بینی جریان رودخانه هنزه در کشور پاکستان با  $R^2$  بیش از ۰/۸۳۱ نشان دادند، همخوانی دارد.

Bigdeli et al., (2023) نیز دقت بالای مدل‌های ANN و RF (با ضریب همبستگی بیش از ۰/۸۲) در مدل‌سازی بارش-رواناب ماهانه ایستگاه هیدرومتری آخولا در حوزه آجی‌چای را نتیجه گرفته بودند.

و رواناب متأثر از بارش کل حوضه است. بنابراین، پیشنهاد می‌شود در تحقیقات آتی، به جای یک ایستگاه از چند ایستگاه هم‌دیدگی استفاده شود. با توجه به نتایج این پژوهش پیشنهاد می‌شود در تحقیقات آینده از روش‌های هوشمند و داده‌کاوی در مدل‌سازی فرایند بارش-رواناب در حوضه‌های مختلف به‌طور جداگانه برای سال‌های درگیر خشکسالی و ترسالی استفاده شود. با توجه به بحران آب در نقاط مختلف کشور و خصوصاً حوزه آبخیز دریاچه ارومیه، پیشنهاد می‌شود با استفاده از روش‌های معرفی شده در این پژوهش نسبت به پیش‌بینی بارش‌ها و رواناب‌های ناشی از آن اقدام شود تا مدیریت مناسبی در زمینه منابع آب اعمال شود.

### تشکر و قدردانی

نویسندگان این پژوهش از شرکت مدیریت منابع آب ایران برای ارائه داده‌های مورد نیاز سپاسگزاری می‌نمایند.

### تعارض منافع

در این مقاله تضاد منافی وجود ندارد و این مسأله مورد تأیید همه نویسندگان است.

و ساختار مدل‌ها را نام‌برد. نوع توابع فعال‌ساز و انتقال در مدل ANN، توابع کرنل در مدل SVM، تعداد و نوع کروموزوم‌ها و نوع تابع پیوند بین زیردرخت‌ها در مدل GEP و پارامترهای تنظیم در مدل RF عواملی هستند که ساختار مدل‌ها را تعیین کرده و عملکرد مدل‌ها را تحت تأثیر قرار می‌دهند. در کل دقت همه مدل‌های مورد استفاده در پژوهش حاضر قابل قبول ارزیابی می‌شود. با این حال، لازم به ذکر است که اگر داده‌های صحت‌سنجی از بخش‌های مختلف دوره آماری به‌طور تصادفی انتخاب شود، نتایج ممکن است متفاوت با یافته‌های فعلی باشد.

### نتیجه‌گیری

در این پژوهش، نتایج حاکی از قابلیت بالای این مدل‌های هوشمند در شبیه‌سازی فرایند بارش-رواناب حوضه مورد مطالعه است. به‌طوری که همه مدل‌ها در مرحله صحت‌سنجی  $R^2$  بیش از ۰/۶۶ دارند. مدل GEP نسبت به سایر مدل‌های مورد بررسی، در شبیه‌سازی بارش-رواناب عملکرد بهتری دارد.

از نقاط ضعف این پژوهش، می‌توان به استفاده از داده‌های یک ایستگاه هم‌دیدگی برای کل حوضه اشاره کرد. در صورتی که بارش در کل حوضه یکنواخت نبوده

### منابع مورد استفاده

- Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E., Wilby, R.L., 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Prog. Phys. Geogr.* 36, 480-513.
- Adamowski, J., 2013. Using support vector regression to predict direct runoff, base flow and total flow in a mountainous watershed with limited data in Uttaranchal, India. *Annals of Warsaw University of Life Sciences-SGGW. Land Reclamation.* 45, 71-83.
- Adnan, R. M., Yuan, X., Kisi, O., Adnan, M., Mehmood, A., 2018. Stream flow forecasting of poorly gauged mountainous watershed by least square support vector machine, fuzzy genetic algorithm and M5 model tree using climatic data from nearby station. *Water Resour. Manage.* 32, 4469-4486.
- Ahmadian AhmadAbad, G., Zakeri Niri, M., 2021. Simulation and comparison of runoff estimation by neural network methods, regression and fuzzy inference (A case study:Dez catchment). *Nivar.* 45, 84-96 (in Persian).
- Allahverdipour, P., Sattari, M.T., 2023. Comparing the performance of the multiple linear regression classic method and modern data mining methods in annual rainfall modeling, case study: Ahvaz city. *Water Soil. Manage. Model.* 3, 125-142 (in Persian).
- Antar, M.A., Ellassiouti, I., Allam, M.N., 2006. Rainfall- runoff modelling using artificial neural networks technique: a Blue Nile catchment case study. *Hydrol. Process.* 20, 1201-1216.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology., 2000. Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol. Eng.* 5, 115-123.
- Aytek, A., Asce, M., Alp, M., 2008. An application of artificial intelligence for rainfall-runoff modeling. *J. Earth Syst. Sci.* 117, 145-155.
- Bigdeli, Z., Majnooni Heris, A., Delirhasannia, R., Karimi, S., 2023. Rainfall-Runoff Modeling of Aji Chai Basin Using Random Forest and Artificial Neural Network Models. *New Research in Sustainable Water Engineering. New Res. Sustainable Water Eng.* 1, 27-42 (in Persian).

- Botsis, D., Latinopoulos, P., Diamantaras, K., 2011. Rainfall-runoff modeling using support vector regression and artificial neural networks. In 12th International Conference on Environmental Science and Technology (CEST2011). 8-10.
- Cutler, A., Cutler, D.R., Stevens, J.R., 2012. Random forests. Ensemble machine learning: Methods and applications 1st ed. NY: Springer New York. 157-175.
- Dehghani, N., Vafakhah, M., Bahremand, A., 2016. Rainfall-Runoff Modeling using Artificial Neural Network and Neuro-Fuzzy Inference System in Kasilian Watershed. *J. Watershed Manage. Res.* 7, 137-128 (in Persian).
- Eskandarinia, A., Nazarpour, H., Teimouri, M., Ahmadi, M.Z., 2010. Comparison of neural network and K-nearest neighbor methods in daily flow forecasting. *J. Appl. Sci.* 10, 1006-1010.
- Ferreira, C., 2002. Gene expression programming in problem solving. *Soft computing and industry: recent applications*. London: Springer. 635-653.
- Hassanzadeh, Y., Abdi Kordani, A., Shafiei Najd, M., Khoshtinat, S., 2015. Daily Streamflow Forecasting of Nooranchay River Using the Hybrid Model of Artificial Neural Networks- Principal Component Analysis. *Water Soil Sci.* 25, 53-63 (in Persian).
- Hussain, D., Khan, A.A., 2020. Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *Earth Sci. Inf.* 13, 939-949.
- Jain, S.K., Das, A., Srivastava, D.K., 1999. Application of ANN for reservoir inflow prediction and operation. *J. Water Resour. Plann. Manage.* 125, 263-271.
- Misra, D., Oommen, T., Agarwal, A., Mishra, S.K., Thompson, A.M., 2009. Application and analysis of support vector machine based simulation for runoff and sediment yield. *Biosyst. Eng.* 103, 527-535.
- Nilsson, P., Uvo, C.B., Berndtsson, R., 2006. Monthly runoff simulation: Comparing and combining conceptual and neural network models. *J. Hydrol.* 321, 344-363.
- Nourani, V., 2017. An emotional ANN (EANN) approach to modeling rainfall-runoff process. *J. Hydrol.* 544, 267-277.
- Pai, P.F., Hong, W.C., 2007. A recurrent support vector regression model in rainfall forecasting. *Hydrol. Process.* 21, 819-827.
- Patel, A.B., Joshi, G.S., 2017. Modeling of rainfall-runoff correlations using artificial neural network-A case study of Dharoi Watershed of a Sabarmati River basin, India. *Civ. Eng. J.* 3, 78-87.
- Sanikhani, H., Dinpazhoh, Y., Zamanzad Ghavidel, S. 2014. The Impacts of Climate Change on Runoff in Watersheds, case study: Ajichay Watershed in East Azerbaijan Province, Iran. *Water Soil* 27, 1225-1234 (in Persian).
- Sepehri, M., Ildoromi, A.R., Hosseini, S.Z., Nori, H., Mohammadzade, F., Artimani, M.M., 2018. The combination of neural networks and genetic algorithms is a way to estimate the Peak flood. *Iran. J. Watershed Manage. Sci.* 11, 23-28 (in Persian).
- Shafeizadeh, M., Fathian, H., Nikbakht Shahbazi, A., 2019. Continuous rainfall-runoff simulation by artificial neural networks based on efficient input variables selection using partial mutual information (PMI) algorithm. *Iran Water Resour. Res.* 15, 144-161 (in Persian).
- Vapnik, V., 1999. *The nature of statistical learning theory*. NY: Springer New York.
- Vapnik, V.N., Chervonenkis, A.Y., 2015. On the uniform convergence of relative frequencies of events to their probabilities. In *Measures of complexity: festschrift for alexey chervonenkis*. Switzerland: Springer Cham. 11-30.
- Vyas, S.K., Mathur, Y.P., Sharma, G., Chandwani, V., 2016. Rainfall-Runoff Modelling: Conventional regression and Artificial Neural Networks approach. *International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*. India: IEEE. 1-6.
- Worland, S.C., Farmer, W.H., Kiang, J.E., 2018. Improving predictions of hydrological low-flow indices in ungaged basins using machine learning. *Environ. Modell. Software.* 101, 169-182.
- Yin, L., 2023. Forecasting Sector Rotation of A-share Market Using LSTM and Random Forest. *Advances in Economics, Management and Political Sciences. Adv. Econ. Manage. Polit. Sci.* 49, 109-123.
- Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., Lee, K.K., 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* 396, 128-138.
- Zeinalie, M., Golabi, M.R., Sharifi, M.R., Hafezparast, M., 2019. Evaluation of artificial intelligence models in river flow modeling, case study: Gamasiab River. *Watershed Eng. Manage.* 11, 941-954 (in Persian).