

The streamflow prediction of Kurkursar river using hybrid artificial intelligence models

Edris Merufinia¹, Ahmad Sharafati^{2*}, Hiran Abghari³, Yousef Hassanzadeh⁴

¹ Ph.D. Student, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

² Associate Professor, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

³ Associate Professor, Department of Range and Watershed Management, Faculty of Natural Resources, Urmia University, Urmia, Iran

⁴ Professor, Department of Water Engineering, Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

Abstract

Introduction

Streamflow prediction is a challenging and critical task for water resource management. Streamflow prediction is one of the essential steps for reliable and robust water resources planning and management. On the other hand, streamflow prediction plays a crucial role in water resources systems planning and mitigating hydrological extremes such as floods and droughts. Since a variety of uncertainties exist in streamflow prediction, it is necessary to enhance our efforts to robustly address uncertainties and their interactions for improving the reliability of streamflow prediction. It is highly vital for hydropower operation, agricultural planning, and flood control. Physical process-based models are developed based on the understanding of the runoff generation processes, transport in channels, and mathematical formulations or parameterization of these physical processes. Data-driven models have the advantages of low demand for model input and ease of use. Due to the influence of various factors, including climate change, human activities, and socio-economic development, the hydrological time series leads to its complicated stochastic characteristics. Statistical models are the typical examples of this category and have been widely used in streamflow predictions over different regions. Different statistical models, machine learning methods are effective in describing the nonlinear characteristics of the observations and offer an alternative approach for streamflow prediction. Over the last two decades, a variety of machine learning techniques, including the artificial neural network (ANN) and the support vector machine (SVM), have been extensively applied in water resource management and hydrological prediction.

Materials and Methods

In this research, from the data of precipitation (P_t), precipitation with a delay of one day (P_{t-1}) until precipitation with a delay of three days (P_{t-3}) and discharge with a delay of one day (Q_{t-1}) until discharge with a delay of 3 days (Q_{t-3}) are used as input variables and discharge (Q_t) is used as output variable to predict the flow of Kurkursar river in Nowshahr. The Kurkursar river basin in the north of Iran is considered a sub-basin of the North Sea, its area is 75.495 km², the average height of the basin is about 860 m, and the overall average slope is about 80.21%. This basin has an oblique shape and is limited to the Caspian Sea from the north, the Meshlak river and part of the Chalus watershed from the south, and the Chalus watershed from the west. Various geomorphological forms such as floodplains, alluvial cones, and sediment dams have been identified in the river basin area. The time series is daily and 70% and 30% of the data are respectively used for the training and test processes. The models used in this research include three individual models (random forest model - artificial neural network and regression support vector machine) and three hybrid models including bagging tree model - random forest (BA-RF), neural network - creative rifleman (ANN-AIG) and Support Vector Machine Regression-Crow Search (SVR-CSA) was used. In order to evaluate the model, the Mean square error (MAE), root mean squared error (RMSE), Nash-Sutcliffe model efficiency coefficient (NSE), and the PSR were used. Pearson's correlation coefficient (PCC) is used for the relationship between input and output variables. Therefore, we calculate the correlation coefficients between input and output variables, then we evaluate the composition of the input model based on different scenarios.

The model combination that has the highest correlation is selected as the selected model. Then, the internal coefficients of each model are optimized with the help of meta-heuristic optimization algorithms. Therefore, the prediction results and observational data of the model are compared with each other. Finally, time series graphs, data dispersion, and box plots will be drawn and we will compare different evaluation indicators quantitatively and qualitatively. Finally, we compare the results of all models with each other and choose the model that has the best result as the best model. In the final step, we will compare the accuracy increase of the hybrid model (error reduction) with the standalone model to determine the error reduction percentage.

Results and Discussion

The Koukursar Nowshahr river flow was predicted using seven model with combinations of predictive variables such as $R(t)$, $Q(t-1)$, $Q(t-2)$, $R(t-1)$, $Q(t-3)$, $R(t-2)$ and $R(t-3)$ as input variables and flow rate (Q_t) as output variable. The results show that the precipitation variable (R_t) with a value of 0.563 has the highest correlation with the output variable, followed by $Q(t-1)$ with 0.463, $Q(t-2)$ with 0.297, $R(t-1)$ with 0.281, $Q(t-3)$ with 0.251, $R(t-2)$ with 0.124, and $R(t-3)$ with 0.072. Variable $R(t)$ with 0.563 has the highest correlation, and variable $R(t-3)$ with 0.072 has the lowest correlation with the output variable. Based on this, the relationship between the input and output variables was shown based on the correlation coefficient using the radar chart. In this regard, in this research, three individual models and three hybrid models were used to predict the river flow. Among the different scenarios and combinations of the input model, model 7 was used as the final model for forecasting. The evaluation results show that the RF model in the training phase has $R^2=0.957$ and in the test, a phase has $R^2=0.717$. The RF model has the highest correlation coefficient among all research models in the training phase. Also, in the test phase, it has the third rank among all models. ANN-AIG model improved the error of the single model by 32.94 %, the single SVR-CSA model by 23.17%, and the BA-RF model by 17.74%.

Conclusion

The qualitative results of the models show that all the models performed very well in predicting the model. Among the research models, the ANN-AIG model has performed best in forecasting.

Keywords: Artificial neural network, Crow search, Hybrid models, Innovative gunner, Kurkursar, Random forest

Article Type: Research Article

*Corresponding Author, E-mail: asharafati@srbiau.ac.ir

Citation: Merufinia, E., Sharafati, A., Abghari, H., & Hassanzadeh, Y. (2023). The streamflow prediction of Kurkursar river using hybrid artificial intelligence models with soft computing approach. *Water and Soil Management and Modeling*, 3(1), 181-199.

DOI: 10.22098/mmws.2022.11657.1150

DOR: 20.1001.1.27832546.1402.3.1.12.0

Received: 11 October 2022, Received in revised form: 30 October 2022, Accepted: 01 November 2022, Published online: 01 November 2022

Water and Soil Management and Modeling, Year 2023, Vol. 3, No. 1, pp. 181-199

Publisher: University of Mohagheh Ardabili

© Author(s)





پیش بینی جریان رودخانه کورکورسر نوشهر با استفاده از مدل های هیبریدی هوش مصنوعی

ادریس معروفی نیا^۱، احمد شرافتی^{۲*}، هیراد عبقری^۳، یوسف حسن زاده^۴

^۱ دانشجوی دکتری، گروه مدیریت ساخت و آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۲ دانشیار، گروه مدیریت ساخت و آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۳ دانشیار، گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه ارومیه، ارومیه، ایران

^۴ استاد گروه مهندسی آب، مرکز آموزش عالی هیدروانفورماتیک، دانشکده مهندسی عمران، دانشگاه تبریز، تبریز، ایران

چکیده

پیش بینی دقیق جریان رودخانه یکی از موضوعات مهم در برنامه ریزی، طراحی، بهره برداری و مدیریت سیستم منابع آب است. هم چنین یک فعالیت ضروری و چالش برانگیز برای شناسایی دوره های خشکسالی هیدرولوژیکی، هشدار و کنترل سیل، بهینه سازی سیستم هیدرولوژیکی یا برنامه ریزی جامع توسعه منابع آب در سند چشم انداز، مدل سازی فعل و انفعالات جریان آب زیرزمینی است. مدل سازی بارش-رواناب یکی از روش های تخمین رواناب و ابزاری مناسب برای مطالعه فرآیندهای هیدرولوژیکی، ارزیابی منابع آبی و مدیریت حوضه آبخیز است. اما پیچیدگی و ماهیت غیر خطی فرآیند بارش-رواناب و ناشناخته بودن تأثیر عوامل روی یکدیگر و نهایتاً روی دبی خروجی حوضه، مدل سازی را مشکل می کند. در این پژوهش از داده های بارش (P_t)، بارش با یک روز تأخیر (P_{t-1}) تا بارش با سه تأخیر (P_{t-3}) و دبی با یک روز تأخیر (Q_{t-1}) تا دبی با سه روز تأخیر (Q_{t-3}) به عنوان متغیرهای ورودی و از دبی (Q_t) به عنوان متغیر خروجی جهت پیش بینی جریان رودخانه کورکورسر نوشهر استفاده شد. سری زمانی، روزانه بوده و از ۷۰ درصد داده ها برای فرآیند آموزش (۱۳۷۶ تا ۱۳۸۷) و ۳۰ درصد داده ها برای آزمون (۱۳۸۷ تا ۱۳۹۱) استفاده شد. مدل های مورد استفاده در این پژوهش، سه مدل منفرد جنگل تصادفی (RF)، شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان رگرسیون (SVR) و سه مدل ترکیبی هیبریدی شامل مدل بگینگ-جنگل تصادفی (BA-RF)، شبکه عصبی-تفنگدار خلاق (ANN-AIG) و ماشین بردار پشتیبان رگرسیون-الگوریتم بهینه سازی جستجوی کلان (SVR-CSA) می باشد. هم چنین جهت ارزیابی مدل های مورد استفاده از شاخص های ارزیابی مجذور میانگین مربعات خطا (RMSE)، میانگین قدر مطلق خطا (MAE)، ضریب بهره وری نش-ساتکلیف (NSE) و ضریب نسبت مجذور میانگین مربعات خطا به انحراف استاندارد مشاهداتی (PSR) استفاده شد. نتایج نشان داد که همه مدل های مورد استفاده (منفرد و هیبریدی) در پیش بینی جریان، عملکرد مطلوبی دارند. هم چنین مدل ANN-AIG به میزان ۳۲/۹۴ درصد، مدل SVR-CSA منفرد ۲۳/۱۷ درصد و مدل BA-RF نیز ۱۷/۷۴ درصد خطای مدل منفرد را بهبود بخشیدند. در بین تمامی مدل های به کار رفته نیز، ANN-AIG دارای بهترین عملکرد در پیش بینی جریان رودخانه کورکورسر نوشهر بوده است.

واژه های کلیدی: تفنگدار خلاق، جنگل تصادفی، جستجوی کلان، شبکه عصبی مصنوعی، کورکورسر، مدل های هیبریدی

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: asharafati@srbiau.ac.ir

استاد: معروفی نیا، ا.، شرافتی، ا.، عبقری، ه.، و حسن زاده، ی. (۱۴۰۲). پیش بینی جریان رودخانه کورکورسر نوشهر با استفاده از مدل های هیبریدی هوش مصنوعی با رویکرد محاسبات نرم. *مدل سازی و مدیریت آب و خاک*، ۳(۱)، ۱۸۱-۱۹۹.

DOI: 10.22098/mmws.2022.11657.1150

DOR: 20.1001.1.27832546.1402.3.1.12.0

تاریخ دریافت: ۱۴۰۱/۰۷/۱۹، تاریخ بازنگری: ۱۴۰۱/۰۸/۰۸، تاریخ پذیرش: ۱۴۰۱/۰۸/۱۰، تاریخ انتشار: ۱۴۰۱/۰۸/۱۰

مدل سازی و مدیریت آب و خاک، سال ۱۴۰۲، دوره ۳، شماره ۱، صفحه ۱۸۱ تا ۱۹۹

© نویسندگان

ناشر: دانشگاه محقق اردبیلی



۱- مقدمه

عصبی، معایی نیز برای این شبکه‌ها وجود دارد. دقت نتایج بستگی زیادی به داده‌های مسأله و نحوه آموزش مدل دارد و پیش‌بینی عملکرد شبکه آن (قابلیت تعمیم) به سادگی امکان‌پذیر نیست. ماشین بردار پشتیبان (SVM) یکی از روش‌های یادگیری با نظارت است که در دو حالت طبقه‌بندی و رگرسیون مورد استفاده قرار می‌گیرند (Yousefi et al., 2020). فرآیند آموزش در این مدل نسبتاً ساده است. همچنین برخلاف شبکه‌های عصبی در بیشینه‌های محلی دچار مشکل نمی‌شود و برای داده‌های با ابعاد بالا جواب نسبتاً مناسبی ارائه می‌نماید (Zhu et al., 2020).

پژوهش‌های مختلفی در خصوص پیش‌بینی جریان رودخانه و مدل‌سازی فرآیند بارش-رواناب انجام گرفته است. Wilby and Dawson (2001) جهت مدل‌سازی بارش-رواناب از شبکه‌های عصبی پرسپترون چندلایه (MLP)، تابع پایه شعاعی (RBF)، رگرسیون خطی چندگانه گام به گام (SWMLP) و پیش‌بینی با مرتبه صفر (ZOF) استفاده نمودند. نتایج پژوهش فوق نشان داد که از بین روش‌های مذکور روش MLP عملکرد بهتری در منطقه مورد مطالعه داشته است. Choy and Chan (2003) از روش SVR برای تعیین تابع شعاعی در مدل‌سازی دبی رودخانه فوجی در کشور ژاپن استفاده نمودند. نتایج پژوهش فوق نشان داد که واریانس خطا با مقدار عددی در پیش‌بینی خیلی کمتر از واریانس خطا در مدل‌سازی بود. Aytek et al. (2008) با استفاده از روش‌های شبکه عصبی (MLP و GRNN) و برنامه‌ریزی ژنتیک اقدام به مدل‌سازی بارش-رواناب رودخانه جونیا آمریکا در مقیاس روزانه نمودند. نتایج به‌طور معناداری نشان‌دهنده برتری برنامه‌ریزی ژنتیک نسبت به روش شبکه عصبی داشته است. Misra et al. (2009) به مقایسه مدل SVR با ANN در پیش‌بینی رواناب در هندوستان پرداختند. نتایج پژوهش آن‌ها نشان داد که مدل SVR دارای عملکرد بهتری نسبت به ANN است. Adamowski (2013) جهت مدل‌سازی بارش-رواناب حوضه کوهستانی دهرادان در یوتا رانچال هندوستان به مقایسه SVR و ANN پرداختند. نتایج برتری مدل SVM بر ANN در پیش‌بینی جریان پایه و جریان کل این منطقه را نشان داد. Sudheer et al. (2014) به پیش‌بینی جریان ماهانه با استفاده از مدل ترکیبی SVM-PSO پرداختند. در پژوهش فوق، عملکرد مدل SVM با مدل میانگین متحرک خود همبسته (ARMA) و شبکه‌های عصبی مصنوعی (ANN) مقایسه شد. نتایج پژوهش آن‌ها نشان داد که، SVM می‌تواند جایگزین بهتری برای پیش‌بینی جریان‌های ماهانه باشد؛ زیرا درجه بالایی از دقت و قابلیت اطمینان را فراهم می‌کند. Ghafari and Vafakhah (2014) جهت شبیه‌سازی فرآیند بارش-رواناب حوضه آبخیز حاجی‌قوشان از مدل‌های ANN و

برآورد دقیق جریان رودخانه، مهم‌ترین عامل در مدیریت سیل و جلوگیری از خسارت‌های اقتصادی ناشی از آن است. بنابراین، تهیه روشی مناسب برای تخمین جریان رودخانه ضروری است (Dehghani et al., 2020; Guven and Kisi, 2011). پیش‌بینی فرآیند جریان رودخانه، منابع مهمی را برای برنامه‌ریزی کنترل سیل، تولید انرژی برق‌آبی و برنامه‌های آب‌رسانی فراهم می‌کند. با این حال، ماهیت دینامیک غیرخطی و چندبعدی بودن این فرآیند، مدل‌سازی آن را دشوار می‌کند (Remesan et al., 2009). یکی از طبقه‌بندی‌های رایج در حوزه یادگیری ماشین، مدل‌های فیزیکی، مفهومی و مدل‌های داده‌منا است (Samsudin et al., 2011; Malik et al., 2020). مدل‌های فیزیکی به پژوهشگران کمک می‌کند تا روند پیش‌بینی جریان رودخانه را بهتر درک نمایند، اما این روش‌ها به دلیل تعدد و همچنین پیچیدگی فرآیندهای دخیل و کمبود داده‌ها با محدودیت همراه خواهد شد (Behzad et al., 2009). در مدل‌های مفهومی، معادلات و ساختارهای فیزیکی به همراه معادلات تجربی به کار می‌رود. با وجود این‌که در مدل‌های مفهومی جزئیات فیزیکی حوضه نیز لحاظ می‌شود، اما جزئیات فیزیکی به اندازه‌ای نیست که کلیه پارامترهای مدل با اندازه‌گیری مستقیم قابل‌حصول باشند (Beven, 1989; Flint et al., 2001). در دهه اخیر الگوریتم‌های هوش مصنوعی، داده‌کاوی و یادگیری عمیق در حال توسعه هستند و به‌خاطر ظرفیت‌های خود برای حل نقاط ضعف الگوریتم‌های سنتی یادگیری ماشین و افزایش کارایی آن مورد توجه قرار می‌گیرند (Ghorbanzadeh et al., 2019). مدل‌های یادگیری ماشین با رویکرد محاسبات نرم، دارای یک ساختار غیرخطی و انعطاف‌پذیر هستند و توانایی پیش‌بینی پدیده‌های پیچیده با دقت بالا را دارند (Ahmed and Sarma, 2007; Afan et al., 2015). در میان مدل‌های یادگیری ماشین، شبکه‌های عصبی مصنوعی (ANNs) یکی از پرکاربردترین مدل‌های مورد استفاده در فرآیند مدل‌سازی فرآیندهای هیدرولوژیکی هستند. این شبکه‌ها به دو دسته نظارت شده و غیرنظارتی تقسیم‌بندی می‌شوند. از نمونه‌های رایج یادگیری نظارت شده می‌توان به شیوه یادگیری پیش‌خور رو به جلو با پس‌انتشار خطا (FFBP)، تابع پایه شعاعی (RBF)، پرسپترون چندلایه (MLP)، شبکه عصبی رگرسیون تعمیم یافته (GRNN) و شبکه بدون نظارت شامل آشنایی با شبکه‌های عصبی خودسازمانده (SOM) اشاره نمود (Danandeh, 2016; Mehr et al., 2015; Deo and Sahin, 2016). ANN دارای مزایای متعددی از قبیل قابلیت یادگیری تطبیقی، خودسازماندهی، تحمل خطا، تعمیم‌دهی و قابلیت پایداری و انعطاف‌پذیری نسبتاً بالا است. در کنار مزایای متعدد شبکه‌های

مدل‌سازی جریان حوضه آبریز رودخانه گاماسیاب پرداختند. نتایج پژوهش فوق‌حاکمی از عملکرد بهتر مدل‌های ANN، RBF و LSTM به‌ویژه در نقاط اوج جریان نسبت به IHACRES در مدل‌سازی جریان برای منطقه مورد مطالعه است. Zeinali and Khashei Siuki (2018) به ارزیابی دقت مدل‌های هم‌زمان سری زمانی (CARMA) و شبکه عصبی (ANN) در مدل‌سازی بارش-رواناب حوضه آبریز نالوچای پرداختند. نتایج پژوهش آن‌ها نشان داد که مدل CARMA به‌مراتب دقت خوبی نسبت به مدل ANN دارد به‌طوری که معیار ریشه میانگین مربعات خطا در مدل CARMA برابر با ۷/۷ و در مدل ANN برابر با ۹/۵۰ مترمکعب بر ثانیه بوده است.

در پژوهش حاضر برای پیش‌بینی جریان رودخانه کورکورسر نوشهر از یک مدل هیبریدی بگینگ-جنگل تصادفی (BA-RF)، مدل شبکه عصبی مصنوعی-الگوریتم بهینه‌سازی تفنگدار خلاق (ANN-AIG) و مدل ماشین بردار رگرسیون-مدل بهینه‌سازی جست و جوی کلاغ (SVR-CSA) استفاده خواهد شد. در ابتدا هدف تعیین مقدار بهینه برای ضرایب درونی مدل و در نهایت نتایج نهایی مدل‌های هیبریدی با مدل‌های منفرد مقایسه خواهد شد تا میزان افزایش دقت و خطای مدل مشخص گردد.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

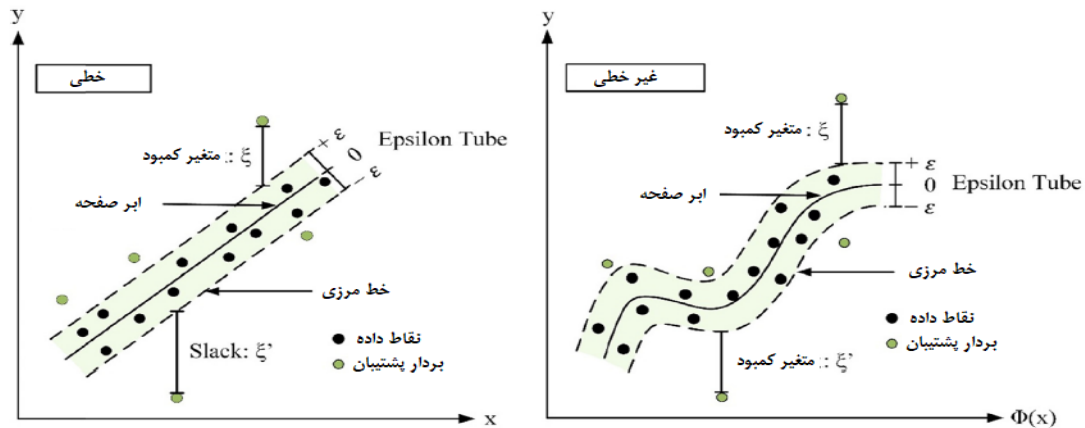
حوضه آبریز کورکورسر در شمال ایران جزء زیرحوضه دریای شمال محسوب می‌شود که بین طول‌های جغرافیایی ۵۱ دقیقه و ۲۹ درجه و ۵۱ دقیقه و ۴۲ درجه شرقی و عرض‌های جغرافیایی ۳۶ دقیقه و ۲۹ درجه و ۳۶ دقیقه و ۴۰ درجه شمالی با مساحت ۷۵/۴۹۵ کیلومتر مربع واقع شده است. میانگین ارتفاع حوضه حدود ۸۶۰ متر و شیب متوسط کلی حدود ۸۰/۲۱ درصد است. از شمال به دریای خزر، از جنوب به رودخانه مشک و بخشی از حوضه آبریز چالوس و از غرب به حوضه آبریز چالوس محدود می‌شود (Khosravi et al., 2021). این منطقه دارای آب و هوای مدیترانه‌ای با میانگین بارندگی و جریان سالانه به‌ترتیب ۹۰۰ میلی‌متر و ۲/۱ مترمکعب بر ثانیه است. حداکثر و حداقل دبی به‌ترتیب ۷۸ و ۰/۰۰۲ مترمکعب بر ثانیه است. شکل ۱ موقعیت رودخانه کورکورسر و ایستگاه هیدرومتری در حوضه آبریز را نشان می‌دهد.

۲-۲- ماشین بردار رگرسیون (SVR)

ماشین بردار پشتیبان، توسط واپنیک و همکارانش (۱۹۹۶) توسعه یافت. SVR یک رابطه بین بردار ورودی بعد $(x \in R)$ و متغیر خروجی $(y \in R)$ را پیش‌بینی می‌کند. هدف از SVR تلاش برای ارائه تمام داده‌های مشاهداتی در داخل مرزها (حداقل خطا)

ANFIS استفاده نمودند. نتایج آن‌ها نشان داد که مدل ANFIS با دقت بالاتری فرآیند بارش-رواناب را پیش‌بینی نموده است. Ghorbani et al. (2016) به مقایسه عملکرد ماشین بردار پشتیبان با سایر مدل‌های هوشمند (توابع کرنل پایه شعاعی، چندجمله‌ای و خطی، مدل شبکه بیزین با الگوریتم یادگیری PC و نیز مدل‌های متداول شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن) جهت شبیه‌سازی فرآیند بارش-رواناب سیمینه‌رود پرداختند. نتایج پژوهش فوق‌حاکمی از عملکرد قابل قبول هر چهار مدل و برتری مدل برنامه‌ریزی بیان ژن با بیش‌ترین ضریب همبستگی بوده است.

Hadi and Tombul (2018) به پیش‌بینی جریان روزانه حوضه‌ها با خصوصیات فیزیکی مختلف از طریق روش‌های داده‌مبنا پرداختند. در پژوهش فوق به ارزیابی دقت و کارایی ANN، ANFIS و SVM در پیش‌بینی جریان روزانه با مدل سنتی (AR) پرداخته شد. نتایج نشان داد که عملکرد ANN و ANFIS برای کل حوضه‌ها از AR و SVM دقیق‌تر است همچنین ANN و ANFIS به‌طور مشابه عمل می‌کنند، ولی ANN از ANFIS بهتر عمل می‌کند. (Pandhiani et al. (2020) به پیش‌بینی سری زمانی جریان‌های رودخانه‌های مالزی با استفاده از تکنیک‌های داده‌مبنا پرداختند. نتایج نهایی پژوهش فوق نشان می‌دهد که عملکرد مدل‌های مبتنی بر RF بالاتر از مدل‌های LS-SVM و M5P برای پیش‌بینی جریان رودخانه است. (Najibzade et al. (2020) از روش‌های SVR و ANN جهت مدل‌سازی بارش-رواناب در حوضه آبریز سد صفارود پرداختند. نتایج حاکمی از عملکرد مناسب روش‌های ANN و SVR در مدل‌سازی در محدوده مورد مطالعه است. (2021) Dehghani et al. به مدل‌سازی فرآیند بارش-رواناب با استفاده از SVR و تبدیل موجک و ANN پرداختند. نتایج نشان داد که مدل ترکیبی ماشین بردار پشتیبان-موجکی در مقایسه با سایر مدل‌های ترکیبی شبکه عصبی مصنوعی موجکی دارای کارایی و دقت بالاتری است. (Kambalimath and Deka (2021) به پیش‌بینی جریان روزانه رودخانه با استفاده از ماشین بردار و تبدیل موجکی پرداختند. نتایج پژوهش آن‌ها نشان داد، زمانی که تبدیل موجک با مدل SVR همراه می‌شود، ضریب تبیین آن (R^2) برای پیش‌بینی‌های یک، سه و پنج روزه به‌ترتیب به ۰/۷۳، ۰/۹۹ و ۰/۶۸ بهبود می‌یابد. (Mirzania et al. (2021) با استفاده از شبکه عصبی و عصبی-موجکی به بررسی تأثیر موجک بر افزایش دقت مدل‌های تخمینی در مدل‌سازی بارش-رواناب پرداختند. نتایج نهایی نشان داد که مدل موجکی عملکرد شبکه منفرد را بهبود بخشیده است. (Momeneh (2022) به مقایسه عملکرد مدل‌های هوش مصنوعی با مدل IHACRES در



شکل ۲- ماشین بردار پشتیبان رگرسیون خطی و غیرخطی (Yu et al., 2006)

Figure 2- Support vector machine for linear and non-linear regression (Yu et al., 2006)

تصمیم می‌گیرد کلاغ ز را دنبال کند تا به مخفیگاه کلاغ ز نزدیک شود. در این جا، دو حالت ممکن است رخ دهد: الف) اطلاع نداشتن از زیر نظر بودن: در این حالت کلاغ ز نمی‌داند کلاغ ز آن را دنبال می‌کند. در نتیجه کلاغ ز به مخفیگاه کلاغ ز نزدیک می‌شود. در این حالت، موقعیت جدید کلاغ ز از طریق رابطه (۷) به صورت زیر به دست می‌آید:

$$X^{i,iter+1} = X^{i,iter} + r_i \times fl^{i,iter} \times (m^{i,iter} - X^{i,iter}) \quad (7)$$

در این رابطه، r_i یک عدد تصادفی با توزیع یکنواخت بین صفر و یک است و $fl^{i,iter}$ طول پرواز کلاغ ز را با تکرار $iter$ است. شکل ۳ شماتیک این حالت و تأثیر fl بر روی قابلیت جست و جو را نشان می‌دهد. مقادیر کوچک fl منجر به جست و جوی محلی می‌شود و مقادیر زیاد منجر به جست و جوی سراسری می‌شود. همان‌طور که در شکل ۳ الف نشان داده شده است، اگر مقدار fl کمتر از یک انتخاب شود، موقعیت بعدی کلاغ ز روی خط تیره بین $X^{i,iter}$ و $m^{i,iter}$ است و همان‌طور که در شکل ۳ نشان داده شده است، اگر مقدار fl بیش از یک انتخاب شود، موقعیت بعدی کلاغ ز روی خط تیره است که ممکن است از m_j فراتر رود.

حالت ب) اطلاع داشتن از زیر نظر بودن: کلاغ ز می‌داند که کلاغ ز آن را دنبال می‌کند. در نتیجه، کلاغ ز برای محافظت از حافظه پنهان خود در برابر سرقت، با رفتن به موقعیت دیگری از فضای جست و جو، کلاغ ز را فریب می‌دهد. در کل، حالت‌های

الف و ب را می‌توان به صورت زیر از طریق رابطه (۸) بیان کرد:

$$X^{i,iter} + r_i \times fl^{i,iter} \times (m^{i,iter} - X^{i,iter}), r_j \quad (8)$$

$= Ap_j^{i,iter}$
a random position, otherwise

در آن r_j یک عدد تصادفی با توزیع یکنواخت بین صفر و یک و $Ap_j^{i,iter}$ بیانگر احتمال آگاهی کلاغ ز در تکرار $iter$ است.

اصول پایه الگوریتم CSA روی چهار محور اصلی استوار است. کلاغ‌ها به صورت گروهی زندگی می‌کنند، موقعیت مخفیگاه‌های خود را حفظ می‌کنند، برای سرقت یکدیگر را تعقیب می‌کنند و از ذخیره‌سازی خود در برابر احتمال سرقت محافظت می‌کنند این الگوریتم فرا ابتکاری این است که کلاغ ز بتواند کلاغ ز را تعقیب کند تا محل اختفای غذای آن را کشف کند. در طی این فرآیند، کلاغ ز باید موقعیت خود را به تدریج به روز کند. علاوه بر این، کلاغ ز باید موقعیت غذا را هنگام دزدیده شدن تغییر دهد. برای این منظور، CSA با یک دسته از کلاغ‌های N و دو ماتریس شروع می‌شود. فرض بر این است که یک محیط d بعدی شامل تعدادی کلاغ وجود دارد. تعداد کلاغ‌ها (اندازه گروه) N است و موقعیت کلاغ ز در زمان تکرار ($iter$) در فضای جست و جو توسط یک بردار $X^{i,iter}$ مشخص می‌شود، که در آن $i=1,2,3,\dots,N$ و $iter=1,2,3,\dots,iter_{max}$ تعریف می‌شود. در حقیقت، d : ابعاد مسأله و $iter_{max}$: حداکثر تعداد تکرار است. $X^{i,iter}$ در فضای d بعدی از طریق رابطه (۶) محاسبه می‌شود:

$$X^{i,iter} = [X_1^{i,iter}, X_2^{i,iter}, \dots, X_d^{i,iter}] \quad (6)$$

در حقیقت هر کلاغ حافظه‌ای دارد که در آن موقعیت مخفیگاه آن حفظ می‌شود. در تکرار $iter$ ، موقعیت مخفیگاه کلاغ توسط $m^{i,iter}$ نشان داده می‌شود. این بهترین موقعیتی است که کلاغ ز تاکنون به دست آورده است. در واقع، به یاد هر کلاغ موقعیت بهترین تجربه آن حفظ شده است. کلاغ‌ها در محیط حرکت می‌کنند و به دنبال منابع غذایی بهتر (مخفیگاه‌ها) می‌گردند. با فرض این که در تکرار $iter$ ، کلاغ ز می‌خواهد از مخفیگاه خود یعنی $m^{j,iter}$ دیدن کند. در این تکرار، کلاغ ز

تعداد متغیرهای تصمیم‌گیری است. موقعیت اولیه کلاغ‌ها از رابطه (۹) به دست می‌آید:

$$Crows = \begin{bmatrix} X_1^1 & X_2^1 & \dots & X_d^1 \\ X_1^2 & X_2^2 & \dots & X_d^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_1^N & X_2^N & \dots & X_d^N \end{bmatrix} \quad (9)$$

حافظه هر کلاغ مقداردهی اولیه می‌شود. از آن‌جا که در تکرار اولیه، کلاغ‌ها هیچ تجربه‌ای ندارند، فرض بر این است که آن‌ها غذاهای خود را در موقعیت اولیه خود پنهان کرده‌اند.

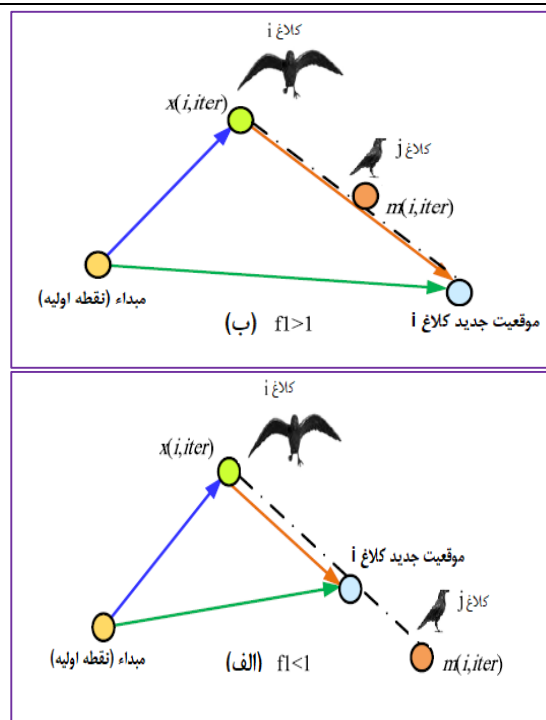
حافظه کلاغ از طریق رابطه (۱۰) به دست می‌آید:

$$Memory = \begin{bmatrix} m_1^1 & m_2^1 & \dots & m_d^1 \\ m_1^2 & m_2^2 & \dots & m_d^2 \\ \vdots & \vdots & \ddots & \vdots \\ m_1^N & m_2^N & \dots & m_d^N \end{bmatrix} \quad (10)$$

مرحله سوم ارزیابی تابع تناسب است. در این حالت برای هر کلاغ، کیفیت موقعیت آن با قرار دادن مقادیر متغیر تصمیم در تابع هدف محاسبه می‌شود. مرحله چهارم ایجاد موقعیت جدید است. در این مرحله کلاغ‌ها موقعیت جدیدی را در فضای جست و جو ایجاد می‌کنند. فرض می‌شود که کلاغ i می‌خواهد موقعیت جدیدی ایجاد کند. برای این منظور، این کلاغ به‌طور تصادفی یکی از کلاغ‌های گروه را انتخاب می‌کند (به‌عنوان مثال کلاغ j) و آن را دنبال می‌کند تا موقعیت غذاهای پنهان شده توسط این کلاغ (m_j) را کشف کند. موقعیت جدید کلاغ i توسط رابطه (۸) به دست می‌آید و این روند برای همه کلاغ‌ها تکرار می‌شود. مرحله پنجم بررسی امکان‌سنجی موقعیت‌های جدید است. در این حالت امکان موقعیت جدید هر کلاغ بررسی می‌شود. اگر موقعیت جدید کلاغ عملی باشد، کلاغ موقعیت خود را به‌روز می‌کند. در غیر این صورت، کلاغ در موقعیت فعلی باقی می‌ماند و به موقعیت جدید ایجاد شده حرکت نمی‌کند. مرحله ششم ارزیابی تابع تناسب برای موقعیت‌های جدید است. در این مرحله مقدار عملکرد تابع تناسب برای موقعیت جدید هر کلاغ محاسبه می‌شود. مرحله هفتم به‌روزرسانی حافظه است که در این حالت کلاغ‌ها حافظه خود را از طریق رابطه (۱۱) به‌روزرسانی می‌کنند:

$$m_i^{i,iter+1} = \begin{cases} X_i^{i,iter+1}, & f(X_i^{i,iter+1}) \leq f(m_i^{i,iter}) \\ m_i^{i,iter}, & \text{Otherwise} \end{cases} \quad (11)$$

اگر مقدار عملکرد تابع تناسب موقعیت جدید کلاغ بهتر از مقدار تابع تناسب موقعیت محافظت شده باشد، کلاغ حافظه خود را با موقعیت جدید به‌روز می‌کند. مرحله هشتم بررسی شرایط خاتمه الگوریتم است. مراحل ۴ تا ۷ برای رسیدن به $iter_{max}$ تکرار می‌شوند. وقتی ملاک خاتمه یافتن، بهترین موقعیت حافظه از نظر مقدار تابع هدف به‌عنوان حل مسأله بهینه‌سازی گزارش می‌شود. جهت انجام مدل‌سازی با SVR از سه پارامتر



شکل ۳- تأثیر طول پرواز (fl) بر قابلیت جست و جوی اکتشاف و بهره‌برداری الگوریتم CSA (Hussien et al., 2020)

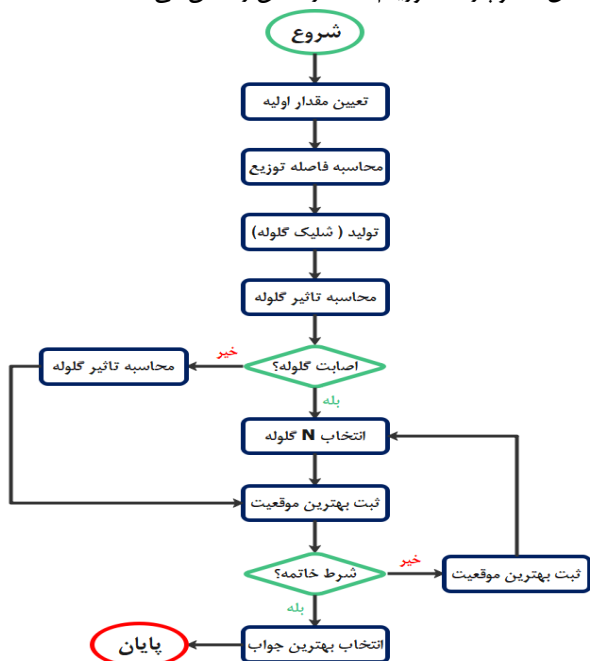
Figure 3- The effect of flight length (fl) on the ability to search, explore and exploit the CSA algorithm (Hussien et al., 2020)

الگوریتم‌های فرا ابتکاری باید تعادل خوبی بین تنوع^۱ و تشدید^۲ ایجاد کند. شدت و تنوع عمدتاً توسط پارامتر احتمال آگاهی (AP) کنترل می‌شود. با کاهش مقدار احتمال آگاهی، الگوریتم تمایل به جست و جو در منطقه محلی دارد که یک راه‌حل خوب در حال حاضر در این منطقه پیدا شده است. در نتیجه، استفاده از مقادیر کم احتمال آگاهی، شدت را افزایش می‌دهد. از طرف دیگر، با افزایش مقدار احتمال آگاهی، احتمال جست و جو در مجاورت راه‌حل‌های خوب فعلی کاهش می‌یابد و CSA تمایل دارد فضای جست و جو را در مقیاس سراسری (تصادفی‌سازی) کشف کند. در نتیجه، استفاده از مقادیر زیاد احتمال آگاهی تنوع را افزایش می‌دهد. در الگوریتم CSA سپس پارامترهای قابل تنظیم اندازه گروه یا جمعیت اولیه (N)، حداکثر تعداد تکرار ($iter_{max}$)، طول پرواز (fl) و احتمال آگاهی (AP) ارزیابی می‌شوند. مرحله اول شامل مقداردهی و تنظیم پارامترهای اولیه است. مرحله دوم مقداردهی موقعیت اولیه و حافظه کلاغ‌ها است. در ابتدا تعداد N کلاغ به‌طور اعضای گروه یا جمعیت به‌طور تصادفی در یک فضای جست و جوی d بعدی قرار می‌گیرند. هر کلاغ بیانگر یک حل عملی از مسأله است و d

¹ Diversification

² Intensification

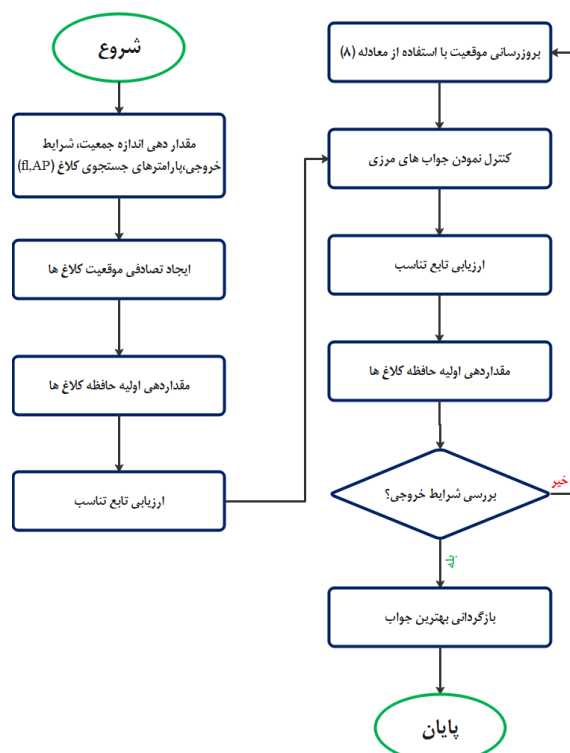
پاسخ بر اساس رویکردهای ازدحامی، این الگوریتم اکتشافی‌های متعددی را در فضای جست و جو انجام می‌دهد و این امر به‌طور قابل‌توجهی به یافتن پاسخ و همگرایی بهینه کمک می‌کند (Sarma, 2021). الگوریتم AIG راه‌حل‌ها و پاسخ‌های مختلف را جست و جو می‌کند و در اجتناب از پاسخ‌های بهینه محلی (گرفتار شدن در پاسخ‌های بهینه محلی) کارایی بالایی دارد. پیش‌بینی می‌شود که نتایج به‌دست آمده توسط این مدل نسبت به سایر روش‌های هوش ازدحامی شناخته شده (مانند الگوریتم ژنتیک، ازدحام ذرات، ملخ، میگو و غیره) کارآمدتر باشد (Dehghani and Torabi Poudeh, 2022). مرحله اول در الگوریتم فوق شروع مدل با نقطه شروع (تعریف مقدار اولیه برای گلوله اول، به‌صورت تصادفی) است. مرحله بعدی تعیین فاصله پخش (فاصله شلیک گلوله تا هدف) است. مرحله سوم محاسبه گلوله تولید شده (گلوله دوم در مرحله سوم از گلوله اول گرفته شده است) است. مرحله چهارم مطالعه احتمال برخورد گلوله با هدف (مطالعه نقاط برخورد احتمالی گلوله و هدف) است. مرحله پنجم انتخاب N گلوله تصادفی به‌عنوان گلوله اصلی (در صورت برخورد مناسب باهدف) است. مرحله ششم ارزیابی و به‌روزرسانی محل و مختصات برخورد گلوله با هدف (در صورت برخورد گلوله با مرکز هدف، ضوابط پایانی محقق می‌شود و همه چیز انجام می‌شود. در صورت عدم اصابت گلوله به هدف، مقدار اولیه باید دوباره تعریف شود) است. مرحله هفتم تعریف بهترین وضعیت ثبت شده و مرحله آخر خاتمه مراحل بهینه‌سازی است (Dehghani and Torabi Poudeh, 2021). شکل ۵ فلوجارت الگوریتم تفنگدار خلاق را نشان می‌دهد.



شکل ۵- نمودار جریان الگوریتم تفنگدار خلاق (Dehghani and Torabi Poudeh, 2021)

Figure 5- Flowchart of the innovative gunner algorithm (Dehghani & Torabi Poudeh, 2021)

ضریب جرمه (C)، ضریب خطا (ϵ) و ضریب تابع کرنل (σ) استفاده می‌شود. بنابراین تعداد متغیرهای تصمیم (nVar) برابر سه متغیر است. مقداردهی اولیه در برنامه MATLAB به‌ترتیب با ضریب $[C, \sigma, \epsilon]$ شروع می‌شود. کران پایینی متغیرها (VarMin) برابر $1e^{-1} * [1 \ 1 \ 1]$ و کران پایینی متغیرها (VarMax) برابر $[50 \ 2 \ 5]$ است. هم‌چنین حداکثر تعداد تکرارها (MaxIt) برابر ۱۰۰ تکرار و تعداد جمعیت اولیه (nPop) برابر ۲۰ کلاغ انتخاب می‌شود. پارامترهای به‌کار رفته در الگوریتم کلاغ شامل ضریب احتمال آگاهی (AP) است که برابر ۰/۵ و ضریب طول پرواز (FL) برابر ۲ انتخاب می‌شود. در نهایت از روش اعتبارسنجی متقابل (K-fold Cross Validation) استفاده می‌شود. شکل ۴ نمودار و مراحل الگوریتم جست و جوی کلاغ را نشان می‌دهد.



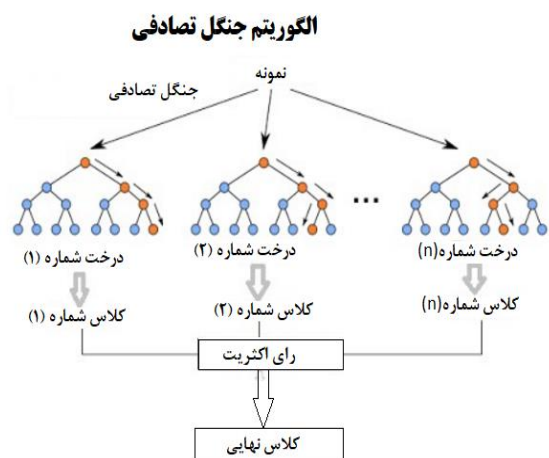
شکل ۴- فلوجارت و مراحل الگوریتم جست و جوی کلاغ (Hussien et al., 2020)

Figure 4- Flowchart and stages of Crow search algorithm (Hussien et al., 2020)

۲-۴ الگوریتم بهینه‌سازی تفنگدار خلاق (AIG)

الگوریتم تفنگدار خلاق (AIG) یکی از روش‌های بهینه‌سازی فرا ابتکاری است که توسط Pijarski and Kacejko (2019) توسعه یافته است. این الگوریتم به‌اندازه کافی سریع و کارآمد در حل مسائل مختلف بهینه‌سازی است. همگرایی بالا و پتانسیل یافتن پاسخ بهینه در کوتاه‌ترین بازه زمانی با هزینه کم و دقت بالا تنها برخی از مزایای این الگوریتم است. به‌دلیل استفاده از بردارهای

بارش و فراوانی وقوع آن‌ها و عوامل ژئومورفولوژیکی و فیزیوگرافیکی از قبیل خصوصیات زمین‌شناسی، خصوصیات خاک، پوشش گیاهی، ارتفاع، شیب، رطوبت نسبی، تبخیر و تعرق اثرگذار هستند. دقت مدل در پیش‌بینی به کیفیت داده‌ها بستگی دارد. لذا یک پیش‌بینی دقیق و توسعه یافته با استفاده از داده‌های محدود زمانی که با کمبود داده مواجه هستیم، کیفیت داده مناسب انجام کار نیست و یا تعداد داده‌های گم شده زیادی داریم که باعث ضعف مدل‌سازی می‌شود، انجام می‌شود.



شکل ۶- الگوریتم جنگل تصادفی (Sharafati et al., 2019)

Figure 6- Random forest algorithm (Sharafati et al., 2019)

در پژوهش حاضر در ابتدا اثرات عوامل مختلف چون دما، تبخیر، رطوبت نسبی، تابش خورشیدی، سرعت باد و بارش مورد بررسی قرار گرفت. اما چون همبستگی مناسبی با متغیر خروجی نداشتند صرفاً از اثر بارش و دبی با تأخیرهای مختلف استفاده شد. مدل‌سازی پیش‌بینی جریان رودخانه با استفاده از متغیرهای $P(t)$ ، $P(t-1)$ ، $P(t-2)$ ، $P(t-3)$ و $Q(t-1)$ ، $Q(t-2)$ ، $Q(t-3)$ که در آن به ترتیب بارش، بارش با یک روز تأخیر تا بارش با سه روز تأخیر و دبی با یک روز تأخیر تا دبی با سه روز تأخیر انجام می‌شود. از ۷۰ درصد داده‌ها برای آموزش (۱۳۷۶ تا ۱۳۸۷ به مدت ۱۰ سال) و ۳۰ درصد برای داده‌ها (۱۳۸۸ تا ۱۳۹۱ به مدت ۴ سال) استفاده می‌شود. برای یافتن ارتباط بین متغیرهای ورودی و خروجی از ضریب همبستگی پیرسون (PCC) استفاده می‌شود. برای این کار همبستگی بین تک‌تک متغیرهای ورودی و خروجی را از طریق ضریب PCC ترکیب مدل ورودی را انتخاب می‌کنیم. در ابتدا بین تک‌تک متغیرهای ورودی و خروجی همبستگی گرفته می‌شود. لذا، متغیر اول دارای بالاترین ضریب همبستگی و متغیر آخر دارای کم‌ترین ضریب همبستگی است. سپس بر اساس ضریب همبستگی پیرسون ترکیب مدل‌های ورودی شکل گرفت. ترکیب اول شامل متغیری است که دارای بالاترین ضریب همبستگی

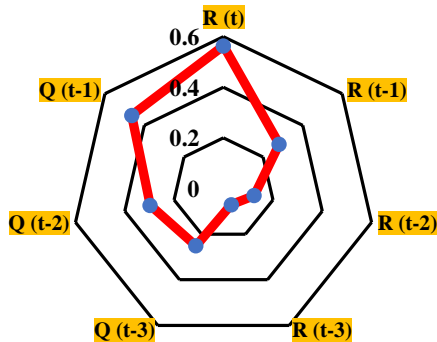
۵-۲- الگوریتم جنگل‌های تصادفی (RF)

الگوریتم RF اولین بار توسط Breiman (2001) معرفی شد. جنگل تصادفی یک روش یادگیری ترکیبی مبتنی بر درخت تصمیم است که برای دو حالت دسته‌بندی و رگرسیون مورد استفاده قرار می‌گیرد. جنگل تصادفی یک طبقه‌بندی کننده گروهی بر اساس درخت تصمیم‌گیری است که در مقایسه با دیگر طبقه‌بندی کننده‌های یادگیری ماشین، عملکرد مطلوب‌تری دارد (Cutler et al., 2012). این الگوریتم به دلیل پایداری و تعمیم خوب آن به‌طور گسترده در زمینه‌های مختلف استفاده شده است (Were et al., 2015). جنگل تصادفی، ترکیبی از درخت‌های تصمیم ساده است که در آن هر درخت با انتخاب نمونه‌ها و ویژگی‌های تصادفی از همه پیش‌بینی‌کننده‌ها تولید می‌شود (Oliveira et al., 2012, Naghibi et al., 2017). اکثر نتایج جمع‌آوری شده از درختان تصمیم به‌عنوان خروجی نهایی RF در نظر گرفته می‌شوند. به‌طور کلی، بخش جداگانه‌ای از نمونه‌ها باید از مجموعه داده اصلی که به آن‌ها نمونه‌های خارج از کیسه (OOB) می‌گویند حذف شود (Abadi et al., 2016). دو پارامتر مهم وجود دارد که باید در طول فرآیند کالیبراسیون مدل RF تعیین شوند. در این روش مقدار بهینه برای دو پارامتر تعداد درخت تصمیم و تعداد متغیرهایی که در ساخت درخت تصمیم به‌کار می‌رود، مورد بررسی قرار می‌گیرد. به عبارت دیگر، حداقل کردن خطا در هر درخت تصمیم، معادل با مقدار بهینه برای تعداد متغیرها (mtree) و تعداد درختان (ntree) است. اگر یک جنگل تصادفی ساخته شده باشد، با ورود یک نمونه بردار X به داخل جنگل تصادفی، این بردار وارد تمامی درخت‌ها می‌شود و در گره‌های پایانی قرار می‌گیرد. سپس بردار ورودی X در هر کدام از درخت‌ها، کلاسی را به خود اختصاص می‌دهد. در نهایت، از بین تمامی کلاس‌ها آن کلاسی که بیش‌ترین تعداد (بیش‌ترین آرا) را داشته باشد، کلاس پیش‌بینی جنگل تصادفی برای بردار ورودی X خواهد بود. در طول فرآیند مدل‌سازی جنگل تصادفی، هر انشعاب درخت با استفاده از یک زیرمجموعه تصادفی از متغیرها (فاکتورها) در هر گره مشخص می‌شود. نتیجه نهایی فرآیند ساخت مدل، میانگین نتایج تمام درختان است. به‌منظور به حداقل رساندن خطای تعمیم، پارامترهای ذکر شده (mtree و ntree) باید بهینه‌سازی شوند (Liaw and Wiener, 2002). شکل ۶ الگوریتم جنگل تصادفی را نشان می‌دهد.

۶-۲- انتخاب سناریو و ترکیب مدل بهینه

در پیش‌بینی جریان رودخانه عوامل مختلف فیزیکی، هیدرولوژیکی، کلیماتولوژی و هواشناسی از قبیل دما، مقدار و نوع

(معیار)، چولگی، کشیدگی و دامنه تغییرات برای بارش (میلی‌متر) و دبی (مترمکعب بر ثانیه) برای مجموعه داده‌های آموزش و آزمون در مجموعه داده رودخانه کورکورس را نشان می‌دهد.



شکل ۷- نمودار راداری برای نمایش ارتباط بین متغیرهای ورودی و خروجی بر اساس ضریب همبستگی

Figure 7- Radar chart to show the relationship between input and output variables based on the correlation coefficient

است. ترکیب دوم، شامل تکرار ترکیب اول و متغیر دوم با بالاترین ضریب همبستگی است و تا آخرین ترکیب و سناریو انتخاب می‌شود. برای نمایش گرافیکی ارتباط بین متغیرهای ورودی و خروجی از نمودار راداری مطابق شکل ۷ استفاده شد. نمودار راداری (نمودار عنکبوتی)، یک نمودار تصویری با چند متغیر است که برای نشان دادن ارتباط چند متغیر در مسائل مختلف استفاده می‌شود. از چند پره با زاویه‌های یکسان تشکیل شده است که در آن هر پره یکی از متغیرهای مسئله را نشان می‌دهد. در این نمودار پس از مشخص کردن مقدار هر متغیر، نقاط نشانگر هر متغیر را به هم متصل کرده و شکلی شبیه به ستاره ساخته می‌شود که در آن ارتباط متغیرهای مختلف بر اساس ضریب همبستگی انجام گرفته است. سپس ترکیب‌های مختلف ورودی که بر اساس ضریب همبستگی پیروان انتخاب شده مطابق جدول ۱ جهت ورود به مدل انتخاب می‌شود. سپس شاخص‌های آماری را برای دو مجموعه داده‌های آموزش و آزمون برای متغیر ورودی (P) و خروجی (Q) به دست آوردیم که جدول ۲ شاخص‌های آماری پایه شامل مقادیر بیشینه، کمینه، میانگین، واریانس، انحراف استاندارد

جدول ۱- ترکیب‌های ورودی مختلف بر اساس ضریب همبستگی پیروان

Table 1- Different input combinations based on Pearson's correlation coefficient

خروجی	پارامترهای ورودی							سناریو
Q (t)	R (t)							1
Q (t)	R (t)	Q (t-1)						2
Q (t)	R (t)	Q (t-1)	Q (t-2)					3
Q (t)	R (t)	Q (t-1)	Q (t-2)	R (t-1)				4
Q (t)	R (t)	Q (t-1)	Q (t-2)	R (t-1)	Q (t-3)			5
Q (t)	R (t)	Q (t-1)	Q (t-2)	R (t-1)	Q (t-3)	R (t-2)		6
Q (t)	R (t)	Q (t-1)	Q (t-2)	R (t-1)	Q (t-3)	R (t-2)	R (t-3)	7

جدول ۲- پارامترهای آماری داده‌های آموزش و آزمون

Table 2- Statistical parameters for the training and testing phase

پارامترهای آماری	مرحله آموزش		مرحله آزمون	
	P (میلی‌متر)	Q (مترمکعب بر ثانیه)	P (میلی‌متر)	Q (مترمکعب بر ثانیه)
بیشینه	149	73.10	147	41.40
کمینه	0	0	0	0
میانگین	3.11	1.29	3.628	1.115
واریانس	124.480	1153.649	131.219	3.585
انحراف استاندارد (معیار)	11.157	4.700	11.455	1.893
چولگی	6.111	13.435	6.029	10.259
کشیدگی	48.631	333.658	53.234	185.232

نمونه‌های فاقد آمار شناسایی شده و با مقدار صفر جایگزین شد. سپس برای شناسایی و حذف داده‌های پرت اقدام شد. یک داده پرت، داده‌ای است که در فاصله غیرعادی از بقیه مقادیر داده در یک نمونه تصادفی از یک جمعیت مشاهده می‌شود؛ تا حدی فاصله این داده با بقیه داده‌ها غیرعادی است که تحلیل‌گر را به این کار وا

۲-۷- پیش‌پردازش داده‌ها

جهت انجام مدل‌سازی از برنامه MATLAB 2020b استفاده شد. قبل از انجام فرآیند مدل‌سازی، ابتدا فرآیند پیش‌پردازش داده‌ها و همچنین آزمون حذف داده‌های گم شده (NaN) انجام گرفت. داده‌های گم شده، مقادیر فاقد آمار است که فرآیند مدل‌سازی را با خطا مواجهه می‌کند. لذا برای حذف یا پر کردن نمونه‌ها، ابتدا

۳- بحث و نتایج

در این پژوهش برای پیش‌بینی جریان رودخانه کورکورس نوشهر از هفت ترکیب مدل از متغیرهای پیش‌بینی کننده از قبیل $R(t)$ ، $Q(t-1)$ ، $Q(t-2)$ ، $Q(t-3)$ ، $R(t-1)$ ، $R(t-2)$ و $R(t-3)$ به‌عنوان متغیرهای ورودی و از دبی (Q_t) به‌عنوان متغیر خروجی استفاده شد. نتایج ارزیابی‌ها (شکل ۷) نشان داد که متغیر بارش (R_t) با مقدار $0/563$ بیش‌ترین همبستگی با متغیر خروجی (Q_t) را دارد و سپس $Q(t-1)$ با $0/463$ ، $Q(t-2)$ با $0/297$ ، $R(t-1)$ با $0/281$ ، $Q(t-3)$ با $0/251$ ، $R(t-2)$ با $0/124$ و $R(t-3)$ با $0/072$ در رتبه‌های بعدی قرار دارند. همان‌گونه که در نتایج دیده می‌شود متغیر $R(t)$ با $0/563$ بیش‌ترین همبستگی و متغیر $R(t-3)$ با $0/072$ کم‌ترین همبستگی با متغیر خروجی را دارند. بر این اساس با استفاده از نمودار راداری ارتباط بین متغیرهای ورودی و خروجی بر اساس ضریب همبستگی نشان داده شد. در این راستا در پژوهش حاضر برای پیش‌بینی جریان رودخانه از سه مدل منفرد و سه مدل هیبریدی استفاده شد. در مرحله اول مدل‌سازی با مدل‌های منفرد RF، ANN و SVR انجام شد. سپس برای افزایش دقت مدل و کاهش خطا و جهت محاسبه ضرایب درونی مدل‌های مذکور از الگوریتم‌های بهینه‌سازی استفاده شد. مدل جنگل تصادفی با مدل بگینگ ترکیب شده و شبکه عصبی مصنوعی با الگوریتم تفنگدار خلاق و ماشین بردار با الگوریتم جست و جوی کلاغ ترکیب شد و سپس شاخص‌های ارزیابی محاسبه شد. ابتدا مدل جنگل تصادفی در حالت منفرد مورد ارزیابی قرار گرفت. نتایج ارزیابی نشان داد که مدل RF در مرحله آموزش دارای ضریب $R^2=0/957$ و در مرحله آزمون دارای ضریب $R^2=0/717$ است. مدل RF در بین تمامی مدل‌های استفاده شده در پژوهش در مرحله آموزش دارای بالاترین ضریب همبستگی قرار دارد. هم‌چنین در مرحله آزمون نیز رتبه سوم را بین تمامی مدل‌ها دارند. هم‌چنین یافته‌ها نشان می‌دهد که در مرحله آزمون، مدل هیبریدی BA-RF توانسته است ۱۷/۷۴ درصد دقت مدل منفرد RF را کاهش دهد. نتیجه مدل منفرد و هیبریدی در مرحله آموزش و آزمون بسیار خوب ارزیابی می‌شود. هم‌چنین سایر نتایج در مرحله آموزش برابر $RMSE=0/5303$ مترمکعب بر ثانیه، $MAE=0/1422$ مترمکعب بر ثانیه، $NSE=0/8056$ و $PSR=0/4408$ است. هم‌چنین بر اساس شاخص خطای $RMSE$ نتایج بررسی‌ها در بین مدل‌های منفرد نشان می‌دهد که در مرحله آموزش به‌ترتیب مدل‌های RF با $0/5305$ مترمکعب بر ثانیه، مدل SVR با $1/4585$ مترمکعب بر ثانیه و مدل ANN با $1/4585$ مترمکعب بر ثانیه رتبه اول تا سوم را دارند. هم‌چنین در مرحله آزمون نیز مدل‌های ANN، SVR و RF به‌ترتیب با $0/8365$ ، $0/8624$ و $1/007$ مترمکعب بر ثانیه رتبه اول تا سوم در پیش‌بینی را دارند. هم‌چنین نتایج مدل‌های هیبریدی نیز نشان داد که بر

می‌دارد که بررسی کند چه چیزی در داده‌ها به نظر غیرعادی می‌آید. در این راستا داده‌های پرت نیز از فرآیند مدل‌سازی حذف شد. سومین مرحله اساسی در فرآیند پیش‌پردازش داده‌ها نرمال کردن داده‌ها است. نرمال‌سازی داده‌ها روشی برای یکنواخت کردن بازه مقادیر مربوط به متغیرهای پژوهش است. اگر واحد سنجش متغیرهای مورد مطالعه متنوع باشد با استفاده از روش‌های نرمال‌سازی می‌توان داده‌ها را بدون مقیاس کرد. از رابطه (۱۲) برای نرمال کردن داده‌ها استفاده شد:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (12)$$

در این رابطه X_n داده نرمال شده، X_{max} داده بیشینه و X_{min} داده کمینه است. رابطه (۱۲) بازه داده‌ها را بین صفر و یک انجام می‌دهد. در این پژوهش برای پیشگیری از مشکلات احتمالی در داده‌ها از بازه یک تا منفی یک استفاده شد. بر این اساس رابطه (۱۳) برای بردن اعداد بین بازه مذکور استفاده شد:

$$X_n = 2 \times \left(\frac{X - X_{min}}{X_{max} - X_{min}} \right) - 1 \quad (13)$$

۲-۸- شاخص‌های ارزیابی مدل

پس از شناسایی بهترین سناریوی ترکیب ورودی و مقدار بهینه برای هر اپراتور، عملکرد هر یک از مدل‌ها توسط شاخص‌های ارزیابی مورد بررسی قرار گرفت. در پژوهش حاضر جهت ارزیابی دقت مدل از شاخص‌های ضریب تبیین (R^2)، ضریب میانگین قدر مطلق خطا (MAE)، ضریب جذر میانگین قدر مطلق خطا (RMSE)، ضریب بهره‌وری نش-سانتکلیف (NSE) و ضریب PSR استفاده شد. ضرایب مذکور به‌ترتیب از روابط (۱۴) تا (۱۸) به‌دست می‌آیند:

$$R^2 = 1 - \left(\frac{\sum_{i=1}^n (Q_t^{obs} - Q_t^{prd})^2}{\sum_{i=1}^n (Q_t^{obs})^2} \right) \quad (14)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_t^{prd} - Q_t^{obs}| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (Q_t^{prd} - Q_t^{obs})^2} \quad (16)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_t^{prd} - Q_t^{obs})^2}{\sum_{i=1}^n (Q_t^{prd} - \overline{Q_t^{prd}})^2} \quad (17)$$

$$PSR = \sqrt{\frac{\sum_{i=1}^N (Q_t^{obs} - Q_t^{pred})^2}{\sum_{i=1}^N (Q_t^{obs} - \overline{Q_t^{pred}})^2}} \quad (18)$$

در روابط بالا به‌ترتیب Q_t^{obs} ؛ دبی مشاهداتی (دبی واقعی)، Q_t^{prd} ؛ دبی محاسباتی (شبیه‌سازی شده)، N : تعداد داده‌ها است. هم‌چنین ضریب PSR نیز نسبت مجذور میانگین مربعات خطا به انحراف استاندارد مشاهداتی را نشان می‌دهد.

کاهش دادند. بر این اساس در مرحله آموزش مدل ANN-AIG با ۳۱/۴۶ درصد بهترین عملکرد را دارد و بیش‌ترین کاهش خطا را احراز نموده و دقت مدل منفرد را بهبود بخشیده است. مدل BA-RF با ۴/۵۲ درصد کم‌ترین کاهش خطا را داشته است. همچنین، در مرحله آزمون نیز مدل ANN-AIG با ۳۲/۹۴ درصد بیش‌ترین کاهش خطا را داشته و لذا دقت مدل منفرد را به بهترین شکل ممکن افزایش داده است و مدل SVR-CSA با ۴/۱۱ درصد کم‌ترین کاهش خطای مدل را به خود اختصاص داده است. جدول ۴ میزان کاهش خطای مدل‌های منفرد و هیبریدی بر اساس شاخص ارزیابی RMSE را نشان می‌دهد. همچنین نتایج کیفی مدل‌ها در جدول ۵ نشان داده شده است.

اساس شاخص RMSE در مرحله آموزش به‌ترتیب مدل‌های BA-RF، ANN-AIG و SVR-CSA به‌ترتیب با ۰/۹۹۹۶، ۰/۷۶۷ و ۱/۱۰۲۲ مترمکعب بر ثانیه و در مرحله آموزش مدل‌های ANN-AIG، SVR-CSA و BA-RF به‌ترتیب با ۰/۸۲۶۹، ۰/۵۶۰۹ و ۰/۸۲۸۳ مترمکعب رتبه‌های اول تا سوم را احراز نمودند. جدول ۳ خلاصه نتایج ارزیابی مدل بر اساس معیارهای مختلف ارزیابی را در دو مرحله آموزش و آزمون را نشان می‌دهد. همچنین نتایج بررسی‌ها نشان می‌دهد که ترکیب BA-RF در مرحله آزمون حدود ۱۷/۷۴ درصد، مدل ANN-AIG در مرحله آموزش و آزمون به‌ترتیب ۳۱/۴۶ و ۳۲/۹۴ درصد و مدل SVR-CSA در مرحله آموزش و آزمون به‌ترتیب ۲۳/۱۷ و ۴/۱۱ درصد خطای مدل را

جدول ۳- نتایج ارزیابی مدل بر اساس معیارهای مختلف ارزیابی در دو مرحله آموزش و آزمون

Table 3- Results of model evaluation based on different evaluation criteria in two phases of training and testing

معیارهای ارزیابی/مدل‌ها	R ²	MAE	RMSE	NSE	PSR
RF-Train	0.957	0.1422	0.5303	0.8056	0.4408
RF-Test	0.717	0.3355	1.007	0.7132	0.5354
BA-RF-Train	0.897	0.1962	0.767	0.5874	0.6423
BA-RF-Test	0.7113	0.3329	0.8283	0.7099	0.5385
ANN-Train	0.5485	0.4754	1.4585	-0.4589	1.2078
ANN-Test	0.7342	0.2849	0.8365	0.7498	0.5011
ANN-AIG-Train	0.791	0.3486	0.9996	0.3146	0.8278
ANN-AIG-Test	0.881	0.2008	0.5609	0.8876	0.3352
SVR-Train	0.563	0.4730	1.4346	-0.4118	1.1881
SVR-Test	0.821	0.0708	0.8624	0.7936	0.4542
SVR-CSA-Train	0.753	0.3636	1.1022	0.1667	0.9128
SVR-CSA-Test	0.831	0.3171	0.8269	0.8102	0.4355

جدول ۴- میزان کاهش خطای مدل‌های منفرد و هیبریدی بر اساس شاخص ارزیابی RMSE

Table 4- Best models in terms of performance

نام مدل	مرحله	درصد کاهش خطای RMSE
BA-RF	آموزش	4.52
BA-RF	آزمون	17.74
ANN-AIG	آموزش	31.46
ANN-AIG	آزمون	32.94
SVR-CSA	آموزش	23.17
SVR-CSA	آزمون	4.11

نیز نمودار ویلثون پلات را نشان می‌دهد. همگرایی بالا و پتانسیل یافتن پاسخ بهینه در کوتاه‌ترین بازه زمانی با هزینه کم و دقت بالا تنها برخی از مزایای الگوریتم بهینه‌ساز تفنگدار خلاق است. به دلیل استفاده از بردارهای پاسخ بر اساس رویکردهای ازدحامی، این الگوریتم اکتشافی‌های متعددی را در فضای جست و جو انجام می‌دهد و این امر به‌طور قابل‌توجهی به یافتن پاسخ و همگرایی بهینه کمک می‌کند.

شکل ۸ نمودار سری زمانی بین داده‌های مشاهداتی (واقعی) و پیش‌بینی (محاسباتی) را در مرحله آزمون برای مدل‌های منفرد و هیبریدی نشان می‌دهد. محور افقی زمان بر حسب روز و محور عمودی دبی بر حسب مترمکعب بر ثانیه را نشان می‌دهد. شکل ۹ نیز نشان‌دهنده نمودار پراکندگی و رگرسیونی را در مرحله آزمون است. محور افقی، دبی مشاهداتی و محور عمودی، دبی محاسباتی بر حسب مترمکعب بر ثانیه را نشان می‌دهد. شکل ۱۰

جدول ۵- ارزیابی کیفی مدل‌های پژوهش در دو مرحله آموزش و آزمون

Table 5- Qualitative evaluation of research models in two phases of training and testing

معیارهای ارزیابی/مدل‌ها	R ²	NSE	PSR
RF-Train	خیلی خوب	خیلی خوب	خیلی خوب
RF-Test	خیلی خوب	خوب	خوب
BA-RF-Train	خیلی خوب	رضایت‌بخش	رضایت‌بخش
BA-RF-Test	خیلی خوب	خوب	خوب
ANN-Train	رضایت‌بخش	عدم رضایت	عدم رضایت
ANN-Test	خیلی خوب	خوب	خوب
ANN-AIG-Train	خیلی خوب	عدم رضایت	عدم رضایت
ANN-AIG-Test	خیلی خوب	خیلی خوب	خیلی خوب
SVR-Train	رضایت‌بخش	رضایت‌بخش	رضایت‌بخش
SVR-Test	خیلی خوب	خیلی خوب	خیلی خوب
SVR-CSA-Train	خیلی خوب	عدم رضایت	عدم رضایت
SVR-CSA-Test	خیلی خوب	عدم رضایت	خیلی خوب

دقت مدل را نسبت به سایر مدل‌ها افزایش دهد که منطبق با یافته‌های پژوهش حاضر است. نتایج نهایی پژوهش حاضر نشان داد که بعد از انجام ۱۰۰ تکرار نتایج پارامترهای درونی SVR بهینه شد که این مقادیر به ترتیب برابر $C=13/6862$ ، $\epsilon=0/1$ و ضریب $\sigma=0/6508$ است. شکل ۱۱ نیز بررسی تعداد مشخصی از پاسخ‌ها (NFE) را نشان می‌دهد که برابر تعداد دفعات ارزیابی تابع هدف در تکرار ۱۰۰ است.

۴- نتیجه‌گیری

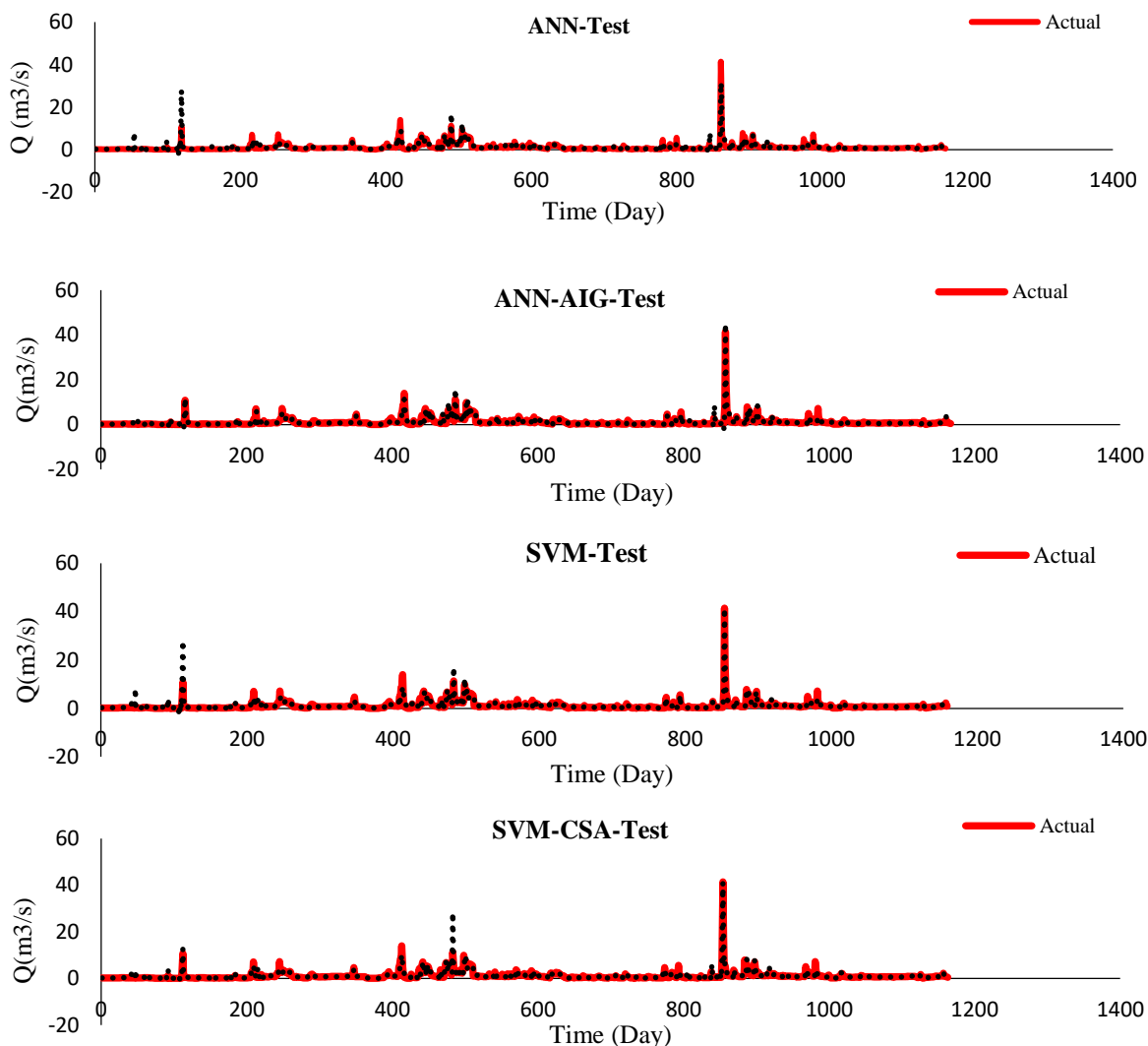
در پژوهش حاضر برای پیش‌بینی جریان رودخانه کورکورسر نوشهر از دو دسته مدل‌های منفرد و ترکیبی استفاده شد. در مرحله نخست پارامتر و متغیرهای ورودی و خروجی پژوهش تعریف شدند. با توجه به دلایل ذکر شده، از ترکیب بارش و دبی به‌عنوان متغیرهای ورودی و خروجی استفاده شد و سپس برای بارش و دبی تأخیرهایی در نظر گرفته شد. از ۷۰ درصد داده‌ها برای آموزش و از ۳۰ درصد برای آزمون استفاده شد. نتایج پژوهش حاضر به‌صورت خلاصه به شرح ذیل است:

متغیر بارش (R_t) با مقدار $0/563$ بیش‌ترین همبستگی با متغیر خروجی (Q_t) را دارد و متغیر بارش با سه روز تأخیر (R_{t-3}) با $0/072$ کم‌ترین همبستگی با متغیر خروجی را دارند.

در بین سناریو و ترکیب‌های مختلف مدل ورودی مدل ۷ به‌عنوان مدل نهایی جهت انجام پیش‌بینی استفاده شد.

نتایج ارزیابی نشان می‌دهد که مدل RF در مرحله آموزش دارای ضریب $R^2=0/957$ و در مرحله آزمون دارای ضریب $R^2=0/717$ است. مدل RF در بین تمامی مدل‌های مورد استفاده در پژوهش حاضر، در مرحله آموزش دارای بالاترین

یافته‌های این پژوهش انطباق مناسبی با پژوهش Dehghani and Torabi Poudeh (2022) دارد. نتایج پژوهش آن‌ها روی شبیه‌سازی تراز آب زیرزمینی ناشی از دقت بالای مدل مذکور بوده است ($R^2=0/964-0/995$). همچنین نتایج با پژوهش Dehghani et al. (2021) برای تخمین جریان رودخانه در حوضه آبریز کرخه داشته است ($R^2=0/985-0/995$). همچنین نتایج با پژوهش Roshni et al. (2022) در پیش‌بینی تراز آب زیرزمینی داشته است ($R^2=0/995$). همچنین نتایج با پژوهش Roustami et al. (2022) در خصوص ارزیابی کارایی برای برآورد دبی سیلابی حوضه کشکان را داشته است ($R^2=0/951$). نتایج نشان می‌دهد که الگوریتم بهینه‌ساز CSA نیز توانایی جست و جو و سرعت همگرایی سریعی از خود نشان می‌دهد و سرعت رسیدن به جواب بدون گیرکردن در بهینه‌های محلی داشته است. نتایج به‌دست آمده با نتایج پژوهش Moldovan et al. (2018) تطابق را دارد. همچنین نتیجه دقت مدل BA-RF با پژوهش Khosravi et al. (2018) در خصوص بار انتقال رسوب هم‌خوانی مناسبی دارد. همچنین بوی و همکاران نیز دقت پیش‌بینی مدل مذکور در پیش‌بینی کیفیت آب رودخانه را مطلوب گزارش نمودند. همچنین نتایج پژوهش Hajian et al. (2022) در پیش‌بینی تراز آب دریاچه ارومیه با استفاده از مدل BA-RF مطلوب گزارش شده است و نتایج حاکی از دقت مدل مذکور بوده است. همچنین Mamun et al. (2022) در خصوص پیش‌بینی رسوب معلق و Shahdad and Saber (2022) در خصوص پیش‌بینی خشکسالی دقت مدل مذکور را مطلوب ارزیابی نمودند. همچنین نتایج پژوهش Khosravi et al. (2020) روی مدل‌سازی فرآیند بارش-رواناب با مدل‌های مبتنی بر درخت تصمیم نشان می‌دهد که مدل RF توانسته است ۲۶/۸۲ درصد



شکل ۸- نمودار سری زمانی بین داده‌های مشاهداتی و محاسباتی در مرحله آزمون برای مدل‌های منفرد و هیبریدی
Figure 8- Time-variation graph for the measured vs. predicted values over the test phase

مدل ANN-AIG به میزان ۳۲/۹۴ درصد، مدل SVR-CSA منفرد ۲۳/۱۷ درصد و مدل BA-RF نیز ۱۷/۷۴ درصد خطای مدل منفرد را بهبود بخشیدند.

منابع

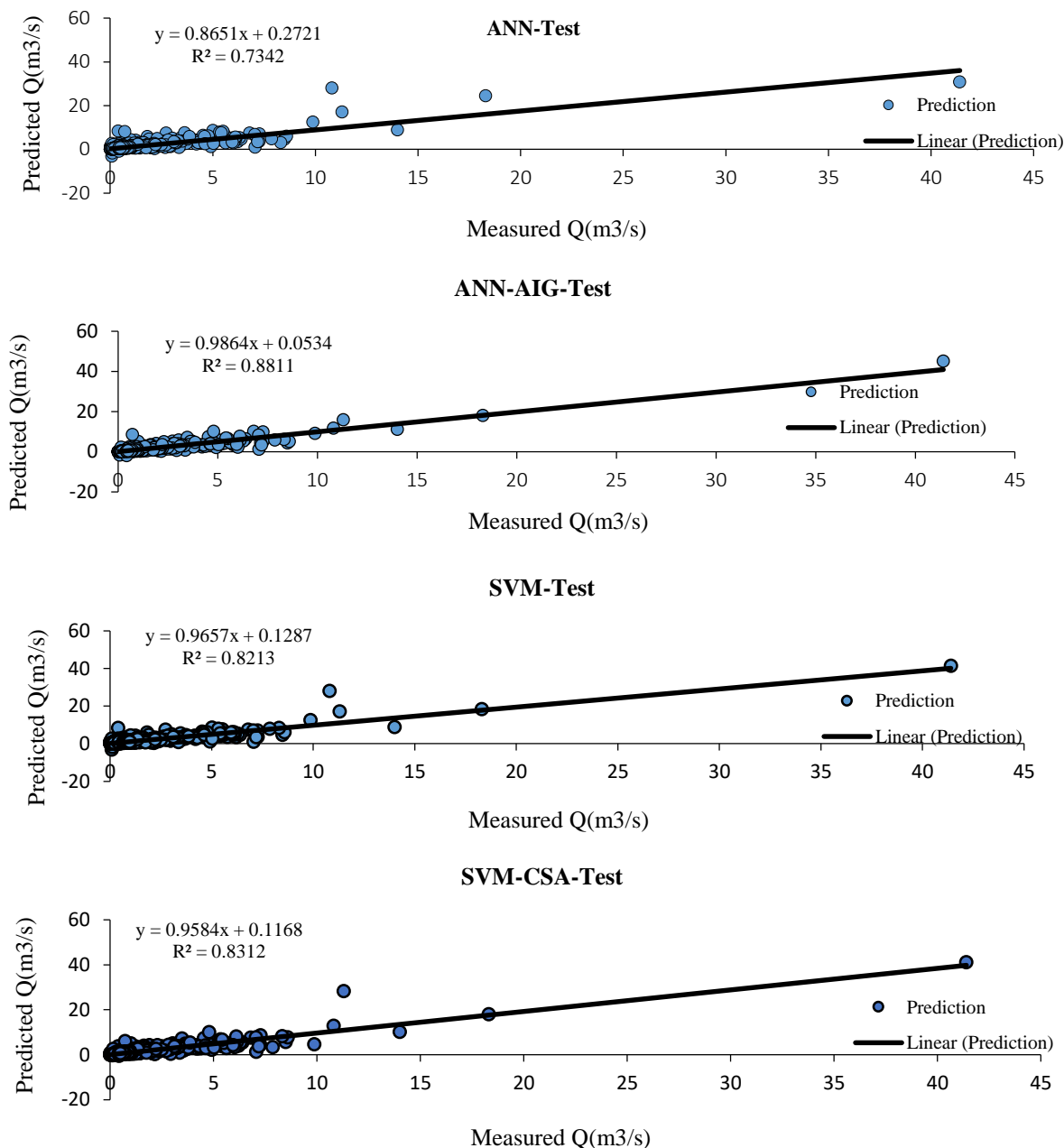
رستمی، س.، شاهی نژاد، ب.، الله یونسی، ح.، ترابی پوده، ح.، و دهقانی، ر. (۱۴۰۰). ارزیابی کارایی مدل‌های هوش مصنوعی هیبریدی نوین در برآورد دبی سیلابی. *هیدروژئومورفولوژی*، ۸(۲۹)، ۱۸۷-۲۰۱.
زینلی، م.ج.، و خاشعی سیوکی، ع. (۱۳۹۷). ارزیابی دقت مدل‌های هم‌زمان سری زمانی و شبکه عصبی در مدل‌سازی بارش-رواناب (مطالعه موردی: حوضه آبریز نازلوچای). *پژوهش‌های حفاظت آب و خاک*، ۲۵(۲)، ۳۱۵-۳۲۱.
غفاری، و وفاخواه، م. (۱۳۹۲). شبیه‌سازی فرآیند بارش-رواناب با استفاده از شبکه عصبی مصنوعی و سیستم عصبی-فازی تطبیقی (مطالعه موردی: حوضه آبخیز حاجی‌قوشان). *پژوهشنامه مدیریت حوزه آبخیز*، ۴(۸)، ۱۲۰-۱۳۶.

ضریب همبستگی قرار دارد. همچنین در مرحله آزمون نیز رتبه سوم را بین تمامی مدل‌ها دارد.

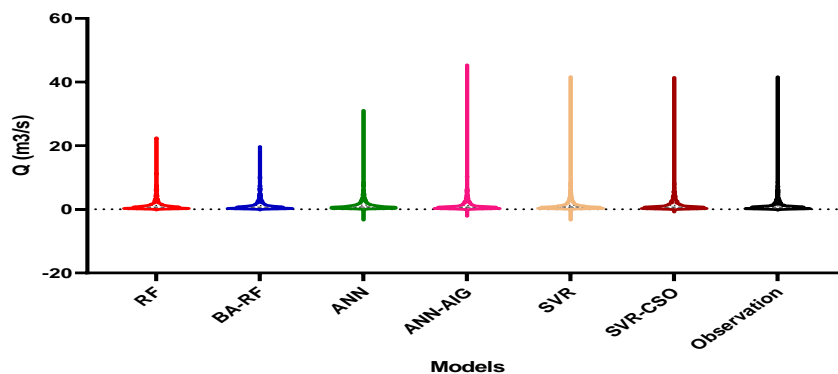
هر چه زمان تأخیر هر متغیر ورودی بیش‌تر باشد، ضریب همبستگی و اثربخشی آن برای اهداف پیش‌بینی کم‌تر است. به عبارت ساده‌تر با افزایش زمان تأخیر، ضریب همبستگی کاهش یافته است.

تحلیل نمودار ویولن پلات نشان می‌دهد که مدل مبتنی بر گروه ANN-AIG بهترین نتایج را برای پیش‌بینی مقادیر جریان با مقادیر بالا (شدید) دارد.

نتایج کیفی مدل‌ها نشان می‌دهد که تمامی مدل‌ها در پیش‌بینی مدل، عملکرد بسیار خوبی از خود نشان دادند. در بین مدل‌های پژوهش، مدل ANN-AIG بهترین عملکرد در پیش‌بینی را داشته است.



شکل ۹- نمودار پراکندگی بین دبی‌های مشاهده‌ای و محاسباتی برای مدل‌های منفرد و هیبریدی در مرحله آزمون
 Figure 9- Scatter plot between the measured vs. predicted values in the test phase



شکل ۱۰- نمودار جعبه‌ای بین داده‌های دبی محاسباتی و دبی مشاهده‌ای در مرحله آزمون برای داده‌های منفرد و هیبریدی
 Figure 10- Box plot diagram between measured vs. predicted values in the test phase

مدل‌سازی بارش-رواناب. *مدل‌سازی و مدیریت آب و خاک*, (۳)۱, ۶۷-۷۹.

نجیب‌زاده، ن.، قادری، ک.، و احمدی، م.م. (۱۳۹۸). بهره‌گیری از روش‌های رگرسیون بردار پشتیبان و شبکه عصبی مصنوعی در مدل‌سازی بارش رواناب (مطالعه موردی: حوضه آبریز سد صفارود). *آبیاری و زهکشی ایران*, (۶)۱۳, ۱۷۰۹-۱۷۲۰.

قربانی، م.ع.، ازانی، ع.، و نقی پور، ل. (۱۳۹۵). مقایسه عملکرد ماشین بردار پشتیبان با سایر مدل‌های هوشمند در شبیه‌سازی فرآیند بارش-رواناب. *پژوهشنامه مدیریت حوزه آبخیز*, (۱۳)۷, ۹۲-۱۰۳.

مومنه، ص. (۱۴۰۱). مقایسه عملکرد مدل‌های هوش مصنوعی با مدل IHACRES در مدل‌سازی جریان حوضه آبریز رودخانه گاماسیاب. *مدل‌سازی و مدیریت آب و خاک*, (۳)۲, ۱-۱۶.

میرزانی، ا.، ملک احمدی، ح.، شاه‌محمدی، ی.، و ابراهیم‌زاده، ع. (۱۴۰۰). تأثیر موجک بر افزایش دقت مدل‌های تخمینی در

References

- 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(1).
- Afan, H.A., El-Shafie, A., Yaseen, Z.M., Hameed, M.M., Wan Mohtar, W.H.M., & Hussain, A. (2015). ANN based sediment prediction model utilizing different input scenarios. *Water Resources Management*, 29(4), 1231-1245.
- Ahmed, J.A., & Sarma, A.K. (2007). Artificial neural network model for synthetic streamflow generation. *Water Resources Management*, 21(6), 1015-1029.
- Al-Abadi, A.M., & Shahid, S. (2016). Spatial mapping of artesian zone at Iraqi southern desert using a GIS-based random forest machine learning model. *Modeling Earth Systems and Environment*, 2(2), 1-17.
- Askarzadeh, A. (2016). A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm. *Computers & Structures*, 169, 1-12.
- Aytek, A., Asce, M., & Alp, M. (2008). An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science*, 117(2), 145-155.
- Barzegari Banadkooki, F., Ehteram, M., Panahi, F., Sammen, S.Sh., Binti Othman, F., & El Shafie, A. (2020). Estimation of total dissolved solids (TDS) using new hybrid machine learning models. *Journal of Hydrology*, 587, 124989.
- Behzad, M., Asghari, K., Eazi, M., & Palhang, M. (2009). Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems With Applications*, 36(4), 7624-7629.
- Beven, K. (1989). Changing ideas in hydrology- The case of physically-based models. *Journal of Hydrology*, 105(1-2), 157-172.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Choy, K.Y., & Chan, C.W. (2003). Modelling of river discharges and rainfall using radial basis function networks based on support vector regression. *International Journal of Systems Science*, 34(14-15), 763-773.
- Cutler, A., Cutler, D.R., & Stevens, J.R. (2012). Random forests. Pp. 157-175, In: *Ensemble machine learning*. Springer, Boston, MA.
- Danandeh Mehr, A., Kahya, E., Şahin, A., & Nazemosadat, M.J. (2015). Successive-station monthly streamflow prediction using different artificial neural network algorithms. *International Journal of Environmental Science and Technology*, 12(7), 2191-2200.
- Dawson, C.W., & Wilby, R.L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography: Earth and Environment*, 25(1), 80-108.
- Dehghani, R., & Poudeh, H.T. (2021). Applying hybrid artificial algorithms to the estimation of river flow: a case study of Karkheh catchment area. *Arabian Journal of Geosciences*, 14(9), 1-19.
- Dehghani, R., & Torabi Poudeh, H. (2022). Application of novel hybrid artificial intelligence algorithms to groundwater simulation. *International Journal of Environmental Science and Technology*, 19(5), 4351-4368.
- Dehghani, R., Poudeh, H.T., Younesi, H., & Shahinejad, B. (2020). Daily streamflow prediction using support vector machine-artificial flora (SVM-AF) hybrid model. *Acta Geophysica*, 68(6), 1763-1778.
- Dehghani, R., Torabi, H., Younesi, H., & Shahinejad, B. (2021). Application of wavelet support vector machine (WSVM) model in predicting river flow (Case study: Dez basin). *Watershed Engineering and Management*, 13(1), 98-110.
- Deo, R.C., & Şahin, M. (2016). An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environmental Monitoring and Assessment*, 188(2), 1-24.
- Flint, A.L., Flint, L.E., Bodvarsson, G.S., Kwicklis, E.M., & Fabryka-Martin, J. (2001). Evolution of the conceptual model of unsaturated zone

- hydrology at Yucca Mountain, Nevada. *Journal of Hydrology*, 247(1-2), 1-30.
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D., & Aryal, J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2), 196.
- Ghafari, G.A., & Vafakhah, M. (2014). Simulation of rainfall-runoff process using artificial neural network and adaptive neuro-fuzzy interface system (case study: Hajighoshan Watershed). *Journal of Watershed Management Research*, 4(8), 120-136 (in Persian).
- Ghorbani, M.A., Azani, A., & Naghipour, L. (2016). Comparison of the performance of support vector machine with other intelligent techniques to simulate rainfall-runoff process. *Journal of Watershed Management Research*, 7(13), 92-103 (in Persian).
- Güven, A., & Kişi, Ö. (2011). Daily pan evaporation modeling using linear genetic programming technique. *Irrigation Science*, 29(2), 135-145.
- Hadi, S.J., & Tombul, M. (2018). Streamflow forecasting using four wavelet transformation combinations approaches with data-driven models: A comparative study. *An International Journal Water Resources Management*, 32(14), 4661-4679.
- Hajian, R., Jalali, M.R., & Mastouri, R. (2022). Multi-step Lake Urmia water level forecasting using ensemble of bagging based tree models. *Earth Science Informatics*, 15, 2515-2543.
- Hassanien, A.E., Rizk-Allah, R.M., & Elhoseny, M. (2018). A hybrid crow search algorithm based on rough searching scheme for solving engineering optimization problems. *Journal of Ambient Intelligence and Humanized Computing*, 1-25.
- Hussien, A.G., Amin, M., Wang, M., Liang, G., Alsanad, A., Gumaei, A., & Chen, H. (2020). Crow search algorithm: theory, recent advances, and applications. *IEEE Access*, 8, 173548-173565.
- Kambalimath, S.S., & Deka, P.C. (2021). Performance enhancement of SVM model using discrete wavelet transform for daily streamflow forecasting. *Environmental Earth Sciences*, 80(3), 1-16.
- Khosravi, K., Cooper, J.R., Daggupati, P., Pham, B.T., & Bui, D.T. (2020). Bedload transport rate prediction: Application of novel hybrid data mining techniques. *Journal of Hydrology*, 585, 124774.
- Khosravi, K., Mao, L., Kisi, O., Yaseen, Z.M., & Shahid, S. (2018). Quantifying hourly suspended sediment load using data mining models: case study of a glacierized Andean catchment in Chile. *Journal of Hydrology*, 567, 165-179.
- Khosravi, K., Miraki, S., Saco, P.M., & Farmani, R. (2021). Short-term river streamflow modeling using ensemble-based additive learner approach. *Journal of Hydro-environment Research*, 39, 81-91.
- Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, 2(3), 18-22.
- Malik, A., Tikhmarine, Y., Souag-Gamane, D., Kisi, O., & Pham, Q.B. (2020). Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. *Stochastic Environmental Research and Risk Assessment*, 34(11), 1755-1773.
- Mamun, A.A., Islam, A.R.M.T., Khosravi, K., & Singh, S.K. (2022). Suspended sediment load prediction using hybrid bagging-based Heuristic Search Algorithm. *Geocarto International*, 37(27), 1-32.
- Merahi, Y., Gabis, A.B., Ramdane-Cherif, A., & Acheli, D. (2021). A comprehensive survey of crow search algorithm and its applications. *Artificial Intelligence Review*, 54(4), 2669-2716.
- Mirzania, E., Malek Ahmadi, H., Shahmohammadi, Y., & Ebrahim Zadeh, A. (2021). Impact of wavelet on accuracy of estimated models in rainfall-runoff modeling (Case study: Sufi Chay). *Water and Soil Management and Modeling (WSMM)*, 1(3), 67-79 (in Persian).
- 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. *Biosystems Engineering*, 103(4), 527-535.
- Moldovan, D., Chifu, V., Pop, C., Cioara, T., Anghel, I., & Salomie, I. (2018). Chicken swarm optimization and deep learning for manufacturing processes. 17th RoEduNet conference: networking in education and research (RoEduNet), Cluj-Napoca, Romania, Pp. 1-6.
- Momeneh, S. (2022). Performance comparison of artificial intelligence models with IHACRES model in streamflow modeling of the Gamasiab River catchment. *Water and Soil Management and Modeling*, 2(3), 1-16 (in Persian).
- Najibzade, N., Qaderi, K., Ahmadi, M.M., (2020). Rainfall-runoff modelling using support vector regression and artificial neural network models (case study: SafaRoud Dam Watershed). *Iranian Journal of Irrigation & Drainage*, 13(6), 1709-1720 (in Persian).
- Naghbi, S.A., Ahmadi, K., & Daneshi, A. (2017). Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential

- mapping. *Water Resources Management*, 31(9), 2761-2775.
- Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., & Pereira, J.M. (2012). Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*, 275, 117-129.
- Panah, P.G., Bornapour, M., Hemmati, R., & Guerrero, J. M. (2021). Charging station stochastic programming for hydrogen/battery electric buses using multi-criteria crow search algorithm. *Renewable and Sustainable Energy Reviews*, 144, 111046.
- Pandhiani, S.M., Sihag, P., Shabri, A.B., Singh, B., & Pham, Q.B. (2020). Time-series prediction of streamflows of Malaysian rivers using data-driven techniques. *Journal of Irrigation and Drainage Engineering*, 146(7), 04020013.
- Pijarski, P., & Kacejko, P. (2019). A new metaheuristic optimization method: the algorithm of the innovative gunner (AIG). *Engineering Optimization*, 51(12), 2049-2068.
- Rashidi, S., Vafakhah, M., Lafdani, E.K., & Javadi, M.R. (2016). Evaluating the support vector machine for suspended sediment load forecasting based on gamma test. *Arabian Journal of Geosciences*, 9(11), 1-15.
- Remesan, R., Shamim, M.A., Han, D., & Mathew, J. (2009). Runoff prediction using an integrated hybrid modelling scheme. *Journal of Hydrology*, 372(1-4), 48-60.
- Roshni, T., Mirzania, E., Hasanpour Kashani, M., Bui, Q. A.T., & Shamshirband, S. (2022). Hybrid support vector regression models with algorithm of innovative gunner for the simulation of groundwater level. *Acta Geophysica*, 70(4), 1885-1898.
- Roustami, S., Shahinejad, B., Younesi, H., Torabi poudeh, H., & Dehghani, R. (2022). Analysis of new hybrid artificial intelligence models in estimating flood flow. *Hydrogeomorphology*, 8(29), 187-201 (in Persian).
- Samantaray, S., Tripathy, O., Sahoo, A., & Ghose, D.K. (2020). Rainfall forecasting through ANN and SVM in Bolangir Watershed, India. *Smart Intelligent Computing and Applications*, 159, 767-774.
- Samsudin, R., Saad, P., & Shabri, A. (2011). River flow time series using least squares support vector machines. *Hydrology and Earth System Sciences*, 15(6), 1835-1852.
- Sarma, S.K. (2021). Optimally configured deep convolutional neural network for attack detection in internet of things: Impact of algorithm of the innovative gunner. *An International Journal Wireless Personal Communications*, 118(1), 239-260.
- Shahdad, M., & Saber, B. (2022). Drought forecasting using new advanced ensemble-based models of reduced error pruning tree. *Acta Geophysica*, 70(2), 697-712.
- Sharafati, A., Khosravi, K., Khosravinia, P., Ahmed, K., Salman, S.A., Yaseen, Z.M., & Shahid, S. (2019). The potential of novel data mining models for global solar radiation prediction. *International Journal of Environmental Science and Technology*, 16(11), 7147-7164.
- Sudheer, C., Maheswaran, R., Panigrahi, B.K., & Mathur, S. (2014). A hybrid SVM-PSO model for forecasting monthly streamflow. *Neural Computing and Applications*, 24(6), 1381-1389.
- Vapnik, V., Golowich, S., & Smola, A. (1996). Support vector method for function approximation, regression estimation and signal processing. *Advances in neural information processing systems*, 9.
- Were, K., Bui, D.T., Dick, B., & Singh, B.R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators*, 52, 394-403.
- Yousefi, V., Kheiri, S., & Rajebi, S. (2020). Evaluation of K-nearest neighbor, bayesian, perceptron, RBF and SVM neural networks in diagnosis of dermatology disease. *International Journal on Technical and Physical Problems of Engineering*, 129(1), 114-120.
- Yu, P.S., Chen, S.T., & Chang, I.F. (2006). Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328(3-4), 704-716.
- Zhu, M., Sun, W., Hahn, A., Wen, Y., Xiao, C., & Tao, W. (2020). Adaptive modeling of maritime autonomous surface ships with uncertainty using a weighted LS-SVR robots to outliers. *Ocean Engineering*, 200(2), 107053.
- Zeinali, M.J., & Khashei Siuki, A. (2018). Assessing the Accuracy of Contemporaneous time series and neural network models in modeling rainfall-runoff (case study: Nazloochei Catchment). *Journal of Water and Soil Conservation*, 25(2), 315-321 (in Persian).