

Research Paper



Intelligent Streamflow Prediction Using a Hybrid Metaheuristic Approach: Tasmanian Devil and Red-Tailed Hawk Optimization Algorithms in the Dehgolan Kurdistan Basin



Edris Merufinia ¹, Ahmad Sharafati ², Hiran Abghari ³, and Yousef Hassanzadeh ⁴

1. Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. edris.merufinia@iaui.ac.ir
2. Associate Professor, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. asharafati@srbiiau.ac.ir
3. Associate Professor, Department of Range and Watershed Management, Urmia University, Urmia, Iran. h.abghari@urmia.ac.ir
4. Professor, Department of Water Engineering, Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, and Farazab Co. (Consulting Engineers), Research and Writing Capacity Enhancement Affairs, Tabriz, Iran. yhassanzadeh@tabrizu.ac.ir

Keywords

Pearson correlation coefficient, Streamflow prediction, Tasmanian devil optimization, LSTM, Dehgolan Basin, Iran.

Received: 2025/07/12

Accepted: 2025/12/01

Published: 2026/00/00

A B S T R A C T

Introduction

With the increasing complexity and dynamics of hydrological systems, accurate and reliable river stream flow prediction is necessary for sustainable water resource management. Streamflow estimation plays a significant role in water resources management, especially for flood mitigation, drought warning, and reservoir operation. Streamflow predictions play a crucial role in decision-making for water managers and policymakers, influencing water allocation, reservoir operations, flood control measures, and drought mitigation strategies. The reason for focusing the prediction model on stream flow data is that hydrological processes are complex, requiring a paradigm based entirely in hard facts of data.

Methodology

This research utilized 20 years (from 2001 to 2021) of daily precipitation, river discharge, and mean air temperature data from the Dehgolan basin in Kurdistan Province. To select the optimal combination and model scenarios, Pearson's correlation coefficient was employed using precipitation (Pt), mean temperature (Tt), and river discharge with one to three days of lag (Qt-1 to Qt-3). The Pearson correlation coefficient (PCC) was used to select optimal scenarios and model combinations, establish the relationship between input and output variables, and subsequently choose the model and scenario combinations. For streamflow prediction, we utilized hybrid models including the Artificial Neural Network-Tasmanian Devil Optimizer (ANN-TDO), Support Vector Regression-Red Tailed Hawk (SVR-RTH), and the deep learning model Long Short-Term Memory-Marine Predators Algorithm (LSTM-MPA). Model evaluation involved the following metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R²), and Kling-Gupta Efficiency (KGE). After completing the forecasting and modeling processes, the results were stored and compared with each other. Various objectives were considered for this comparison.

*Correspondin Author: Ahmad Sharafati: asharafati@srbiiau.ac.ir

How to cite this article: Merufinia, Edris; Sharafati, Ahmad; Abghari, Hiran and Hassanzadeh, Yousef. (2025). Intelligent Streamflow Prediction Using a Hybrid Metaheuristic Approach: Tasmanian Devil and Red-Tailed Hawk Optimization Algorithms in the Dehgolan Kurdistan Basin. *Hydrogeomorphology*, 13(46): startpage – endpage. DOI: [10.22034/hyd.2025.68113.1799](https://doi.org/10.22034/hyd.2025.68113.1799)



These included: examining the dominant combined scenario, evaluating the model in both training and testing phases, assessing the increase in accuracy of the hybrid model compared to individual models, comparing different hybrid models with each other, and comparing the results of NFE models with each other in terms of convergence time and accuracy, as well as the extent of getting trapped in local optimal solutions. Finally, time series plots, scatter plots, Taylor diagrams, and cumulative error plots were drawn and examined for visual representation of the results.

Results and Discussion

The ANN-TDO hybrid model demonstrated significantly better performance compared to the standalone ANN model in both the training and testing phases. For all scenarios (M1 to M5), the hybrid model recorded higher R2 and KGE values, and lower MAE and RMSE values. This indicates the higher accuracy and lower error of the hybrid model. The overall best performance for the hybrid model was observed in scenario M1 during the testing phase (with R2=0.881, MAE=0.021, RMSE=0.054, and KGE=0.915). Additionally, in the training phase, the hybrid model in scenario M5 showed more desirable performance (with R2=0.852 and KGE=0.898). The overall worst performance belonged to the standalone ANN model in scenario M4 (4-20-1) during the testing phase (with R2=0.349 and KGE=0.588), indicating an insufficient ability to generalize to new data. In the training phase, by using R2 as a criterion, the results showed that models M1 to M5 improved the performance of the standalone model by 10.66%, 19.25%, 39.40%, 79.45%, and 82.44% respectively. Similarly, in the testing phase, they improved the standalone model's performance by 4.51%, 0.27%, 25.22%, 133.81%, and 90.48% respectively. The LSTM-MPA hybrid model improved the error of the standalone LSTM by 75.11%, 87.53%, 72.27%, 67.86%, and 113.45% in scenarios M1 to M5 respectively during the training phase, and by 54.90%, 92.12%, 68.45%, 75.65%, and 118.87% respectively during the testing phase. Regarding the standalone LSTM model and the LSTM-MPA hybrid model, the results showed that for the standalone model in the training phase, scenario M4 exhibited relatively better performance with the highest R2 value (0.504) and the lowest RMSE (0.109) and MAE (0.074). In the testing phase, scenario M1 had the highest R2 (0.510), but overall, M4 was more dominant in the training phase. This indicates that the LSTM model in scenario M4 had a better ability to learn patterns from the training data. For the LSTM-MPA hybrid model in the training phase, scenario M3 is identified as the dominant scenario with the highest R2 (0.851) and the lowest RMSE (0.060) and MAE (0.022).

Conclusions

The aim of this research is to predict the Dehgolan Kurdistan river flow using artificial intelligence models. Initially, the research data was acquired, and after performing data management tests, the data was divided into training and testing phases at a 70:30 ratios. ANN, SVR, and LSTM models were utilized with meta-heuristic algorithms (TDO, RTH, and MPA) to optimize the models' hyperparameters. Pearson correlation coefficient was employed to create model scenarios, and scenarios with the highest correlation values were selected. Among the hybrid models, ANN-TDO consistently demonstrated the best performance. Its R2 values for the testing phase were frequently above 0.8, and KGE values reached up to 0.915 in some scenarios, indicating a very high correlation with observed data. The LSTM-MPA model also delivered very good performance. Although it performed slightly below ANN-TDO in some scenarios, its R2 and KGE values (often above 0.7 and 0.8), along with low MAE and RMSE values, demonstrated this model's high capability in time series modeling. LSTM's nature in preserving long-term information and MPA's optimization mechanism contributed to LSTM-MPA maintaining high accuracy. In comparison, SVR-RTH, while improved over the standalone SVR, generally exhibited lower performance than ANN-TDO and LSTM-MPA.

مقاله پژوهشی



پیش‌بینی هوشمند جریان رودخانه با ترکیب الگوریتم‌های فراکاوشی شیطان تاسمانی و شاهین دم‌قرمز در حوضه‌ی دهگلان کردستان



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. edris.merufinia@iau.ac.ir

۲- دانشیار، گروه مدیریت ساخت و آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران. asharafati@srbiau.ac.ir

۳- دانشیار، گروه مرتع و آبخیزداری، دانشکده منابع طبیعی، دانشگاه ارومیه، ارومیه، ایران. h.abghari@urmia.ac.ir

۴- استاد، گروه مهندسی آب، قطب علمی هیدروانفورماتیک، دانشکده مهندسی عمران، دانشگاه تبریز و شرکت فرازآب (مهندسین مشاور)، امور ارتقای توانمندی تحقیقات و تألیفات، تبریز، ایران. yhassanzadeh@tabrizu.ac.ir

چکیده

کلیدواژه‌ها

پیش‌بینی دقیق جریان رودخانه نقش مهمی در مدیریت منابع آب، به ویژه برای کاهش مخاطرات ناشی از سیل، هشدار خشکسالی و بهره‌برداری از مخازن سدها ایفا می‌کند. در این پژوهش از آمار ۲۰ ساله (از سال ۱۳۸۰ لغایت ۱۴۰۰) بارش، دبی رودخانه و دمای میانگین در مقیاس روزانه در حوضه آبریز دهگلان استان کردستان استفاده گردید. برای انتخاب ترکیب بهینه و سناریوهای مدل از مقادیر بارش (P_t)، دمای میانگین (T_t) و دبی از یک تا سه روز تاخیر (Q_{t-1} تا Q_{t-3}) با استفاده از ضریب همبستگی پیرسون استفاده شد. جهت انجام مدل‌سازی و پیش‌بینی جریان رودخانه نیز از مدل‌های هیبریدی شبکه عصبی مصنوعی-شیطان تاسمانی (ANN-TDO)، ماشین بردار پشتیبان رگرسیونی-شاهین دم‌قرمز (SVR-RTH) و مدل یادگیری عمیق حافظه طولانی کوتاه‌مدت - شکارچیان دریایی (LSTM-MPA) استفاده گردید. سپس نتایج مدل‌سازی بر اساس معیارهای ارزیابی ضریب تبیین (R^2)، میانگین قدر مطلق خطا (MAE)، ریشه میانگین مربعات خطا (RMSE) و ضریب بهره‌وری کلینگ-گوپتا (KGE) مورد سنجش قرار گرفت. یافته‌های این پژوهش نشان داد که مدل‌های هیبریدی تا حد بسیار خوبی دقت مدل‌های منفرد را بهبود بخشیدند. همچنین نتایج نشان داد که عملکرد مدل‌ها بسیار نزدیک است اما مدل ANN-TDO نسبت به سایر مدل‌ها عملکرد بهتری داشته است. همچنین این مدل در فاز آموزش توانست در سناریوهای یک تا پنج (M_1 تا M_5) بر اساس معیار RMSE به ترتیب ۱۰/۶۶، ۴۰/۲۵، ۳۹/۱۹، ۷۹/۴۵ و ۸۲/۴۴ درصد عملکرد مدل منفرد ANN را بهبود بخشد. همچنین نتایج نمودار همگرایی (NFE) نیز نشان داد که مدل ANN-TDO تا حد بسیار قابل قبولی قادر به خطای کاهش مدل است.

ضریب همبستگی پیرسون،
پیش‌بینی جریان رودخانه،
الگوریتم شیطان تاسمانی،
LSTM، حوضه دهگلان، ایران.

تاریخ دریافت: ۱۴۰۴/۰۴/۲۱

تاریخ پذیرش: ۱۴۰۴/۰۹/۱۰

تاریخ انتشار: ۱۴۰۵/۰۱/۰۰

*نویسنده مسئول: احمد شرافتی asharafati@srbiau.ac.ir **ابجاع به این مقاله:** معروفی نیا، ادریس؛ شرافتی، احمد؛ عبقری، هیراد و حسن زاده، یوسف. (۱۴۰۵). پیش‌بینی هوشمند جریان رودخانه با ترکیب الگوریتم‌های فراکاوشی شیطان تاسمانی و شاهین دم‌قرمز در حوضه‌ی دهگلان کردستان. هیدروژنومورفولوژی، ۱۳(۴۶): ۱-۱۳

ص-ص

شناسه دیجیتال مقاله: 10.22034/hyd.2025.68113.1799



Copyright: ©2025 by the authors

Publisher: University of Tabriz

منابع آب یکی از مهمترین منابع طبیعی و یک مولفه مهم از یک اکوسیستم، که نقش در توسعه پایدار حوضه های آبخیزدار، کشاورزی و مدیریت منابع آب دارد (شریفی و همکاران، ۱۴۰۲: ۵۱). شبیه‌سازی فرآیند بارش- رواناب در یک حوضه از نقطه نظر درک بهتر مسائل هیدرولوژیکی، مدیریت منابع آب، مهندسی رودخانه، سازه‌های کنترل سیل و ذخیره سیلاب اهمیت ویژه‌ای دارد (شاهدی و فروتن دانش، ۱۴۰۱: ۴۲). پیش‌بینی جریان رودخانه به دلیل پیچیدگی سیستم‌های هیدرولوژیکی و تعامل پویای آن‌ها با عوامل اقلیمی، یک چالش بزرگ است. (وینوکیچ^۱ و همکاران، ۲۰۲۵: ۲۷). جریان رودخانه یک متغیر کلیدی برای پایش مؤثر سیلاب و توسعه سیستم‌های هشدار سریع محسوب می‌شود. با این وجود، پیش‌بینی دقیق جریان رودخانه به دلیل محدودیت‌های مدل‌های هیدرولوژیکی موجود خطاهای پیش‌بینی هواشناسی و شرایط اولیه هیدرولوژیکی، امری چالش‌برانگیز است (سولانکی^۲ و همکاران، ۲۰۲۵: ۶۱). مدل‌های هیدرولوژیکی سنتی که مبتنی بر رویکردهای قطعی و تصادفی هستند، به دلیل رابطه پیچیده بین بارش و رواناب، همواره در پیش‌بینی جریان رودخانه با چالش‌هایی روبرو بوده‌اند. این پیچیدگی تحت تأثیر عوامل گوناگونی از جمله الگوهای کاربری اراضی، ویژگی‌های حوضه، رطوبت خاک، پوشش گیاهی و فعالیت‌های انسانی قرار دارد، که این امر، محدودیت‌های قابلیت پیش‌بینی مدل‌ها را برجسته می‌سازد (نصیر^۳ و همکاران، ۲۰۲۵: ۶۰). برای مقابله با پیچیدگی‌های هیدرولوژیکی، مدل‌های یادگیری ماشین به عنوان گزینه‌ای کارآمد ظهور کرده‌اند. این مدل‌ها، برخلاف رویکردهای فرآیند-محور سنتی، مستقیماً از داده‌ها می‌آموزند و قادرند روابط پیچیده و غیرخطی پنهان را کشف کنند (دانش^۴ و همکاران، ۲۰۲۵: ۳۹). یادگیری ماشین به طور فزاینده‌ای برای پیش‌بینی جریان رودخانه به کار گرفته می‌شود؛ این کاربرد هم شامل تخمین‌های مستقیم جریان و هم برآورد پارامترها برای مدل‌های مبتنی بر فرآیندهای فیزیکی می‌شود. مدل‌های حافظه طولانی کوتاه مدت (LSTM) نوعی الگوریتم یادگیری ماشین از دسته شبکه‌های عصبی بازگشتی اصلاح‌شده هستند که برای یادگیری الگوها و انجام پیش‌بینی از داده‌های متوالی، مانند سری‌های زمانی، مناسب می‌باشند (گری^۵ و همکاران، ۲۰۲۵: ۱۳۳). به دلیل همین ویژگی‌ها، مدل‌های LSTM کارایی بالایی برای پیش‌بینی جریان روزانه و پیوسته رودخانه در حوضه‌های فاقد ایستگاه سنجش از خود نشان داده‌اند (احمد^۶ و همکاران، ۲۰۲۴: ۴۴۱). برای دستیابی به پیش‌بینی‌های دقیق، مدل‌های LSTM نیازمند مجموعه داده‌های ورودی بزرگی هستند. این مجموعه داده‌ها باید شامل توالی‌های طولانی از دبی جریان، متغیرهای هواشناسی مانند بارش، و توصیف‌گرهای حوضه آبریز باشد تا الگوریتم بتواند الگوهای زمانی و مکانی را بیاموزد. با این وجود، اگرچه مدل‌های LSTM می‌توانند بسیار دقیق باشند، اما دارای محدودیت‌هایی نیز هستند. به دلیل نیاز به مجموعه داده‌های ورودی بزرگ و ساختار پیچیده‌شان، این مدل‌ها برای آموزش مؤثر به توان محاسباتی بالایی نیاز دارند، مستعد بیش‌برازش بوده و در تشخیص رفتارهای حادی که به ندرت رخ می‌دهند، با چالش مواجه هستند (تریپاثی و میشر^۷، ۲۰۲۴: ۶۲۸). شبکه‌های عصبی مصنوعی (ANN) و ماشین‌های بردار پشتیبان رگرسیونی (SVR) دو مورد از پرکاربردترین و مؤثرترین الگوریتم‌های یادگیری ماشین برای پیش‌بینی جریان رودخانه محسوب می‌شوند. مزیت اصلی هر دو رویکرد در توانایی آن‌ها برای مدل‌سازی روابط پیچیده و غیرخطی بین متغیرهای هیدرومتئورولوژیکی (مانند بارش) و جریان رودخانه، بدون نیاز به تعریف صریح فرآیندهای فیزیکی، نهفته است (بیکیکی^۸ و همکاران، ۲۰۲۵: ۱۸). شبکه‌های عصبی مصنوعی، با الهام از ساختار مغز انسان، قادرند الگوهای پنهان در داده‌های تاریخی را با دقت بالایی بیاموزند و خود را با دینامیک‌های متغیر سیستم تطبیق دهند. از سوی دیگر، ماشین‌های بردار پشتیبان رگرسیون با استفاده از تابع کرنل و تمرکز بر یافتن یک حاشیه بهینه، توانایی بالایی در تعمیم‌پذیری و

1 Vinokić
2 Solanki
3 Nasir
4 Danesh

5 Grey
6 Ahmed
7 Tripathy and Mishra
8 Bickici

جلوگیری از بیش‌برازش از خود نشان می‌دهد (اسماعیل^۱ و همکاران، ۲۰۲۵: ۱۷). دانش^۲ و همکاران (۲۰۲۵) به ارزیابی مقایسه‌ای مدل‌های یادگیری ماشین و یادگیری عمیق برای پیش‌بینی روزانه جریان رودخانه پرداختند. شش مدل جنگل تصادفی (RF)، درخت تصمیم (DT) و نزدیکترین همسایگان (KNN) شبکه‌های عصبی کانولوشنی (CNN)، حافظه طولانی کوتاه‌مدت (LSTM) و مدل هیبریدی CNN-LSTM در دو ایستگاه سنجش بر روی رودخانه مک‌کنزی در ایالات متحده توسعه و آزمایش شد. نتایج نشان داد که مدل‌های یادگیری عمیق به طور مداوم عملکرد بهتری نسبت به روش‌های یادگیری ماشین در هر دو منطقه داشتند. این مطالعه نتیجه‌گیری کرد که مدل هیبریدی CNN-LSTM دقیق‌ترین پیش‌بینی‌ها را به دست آورد. لی^۳ و همکاران (۲۰۲۵) به بررسی پیش‌بینی جریان برای حوضه‌های فاقد ایستگاه سنجش پرداختند. با استفاده از مدل‌های جنگل تصادفی و شبکه عصبی مصنوعی نشان دادند که عملکرد مدل‌ها در این شرایط، قوی و پایدار نیست. نتایج مشخص کرد که دقت پیش‌بینی به شدت تحت تأثیر تفاوت در ویژگی‌های حوضه مانند مساحت، تراکم سدها و سطح شهرنشینی قرار می‌گیرد. ژانگ^۴ و همکاران (۲۰۲۵) کاربرد یادگیری ماشین را برای پیش‌پردازش و پس‌پردازش جهت بهبود شبیه‌سازی روزانه جریان رودخانه در ۵۲۲ زیرحوضه بررسی کردند. در پژوهش مذکور از RF برای پیش‌پردازش ورودی‌ها (بارش ماهواره‌ای) و از RF و LSTM برای پس‌پردازش خروجی‌ها (جریان شبیه‌سازی‌شده) استفاده گردید. یافته‌های آنها نشان داد که اگرچه هر دو روش به تنهایی عملکرد را بهبود بخشیدند، اما پیش‌پردازش به طور قابل توجهی مؤثرتر از پس‌پردازش بود. با این حال، بهترین و پایدارترین نتایج زمانی به دست آمد که هر دو رویکرد پیش‌پردازش و پس‌پردازش با هم ترکیب شدند. صالح^۵ و همکاران (۲۰۲۵) پنج مدل یادگیری ماشین (شامل M5، RF، SVR و MLP) را با یک روش سنتی هیدرولوژیکی (GR2M) برای پیش‌بینی جریان ماهانه مقایسه کردند. نتایج نشان داد که مدل RF نه تنها بهترین عملکرد را در میان مدل‌های یادگیری ماشین داشت، بلکه به طور قابل توجهی از روش سنتی GR2M نیز دقیق‌تر بود. توسان^۶ و همکاران (۲۰۲۵) در زمینه یادگیری ماشین گروهی در هیدرولوژی، حدود ۲۰۰ مقاله را برای شناسایی روندها و متدولوژی‌های کلیدی تحلیل کردند. یافته‌های آنها نشان‌دهنده افزایش چشمگیر در به‌کارگیری تکنیک‌های EML است که در آن مدل‌های تقویتی (مانند XGBoost) و انباشت (مانند ترکیب RF و LSTM) برجسته‌ترین و مؤثرترین استراتژی‌ها برای وظایفی مانند پیش‌بینی جریان رودخانه بوده‌اند. کاتسکپور^۷ و همکاران (۲۰۲۵) به منظور مدیریت سیلاب در حوضه ولتای سفید غنا از مدل‌های یادگیری ماشین (LSTM و RF) جهت پیش‌بینی جریان رودخانه استفاده نمود. نتایج نشان داد که هر دو مدل عملکرد خوبی داشته و LSTM در پیش‌بینی جریان‌های اوج (سیلاب) برتر بود، در حالی که مدل RF پیش‌بینی‌های پایدارتری برای بازه زمانی طولانی‌تر (تا ۱۰ روز) ارائه داد. پیش‌بینی‌های آینده نیز حاکی از افزایش ریسک سیل و خشکسالی در این حوضه بوده است. اگرچه الگوریتم‌های فراکاوشی متعددی با موفقیت برای بهینه‌سازی مدل‌های یادگیری ماشین در پیش‌بینی جریان رودخانه به کار گرفته شده‌اند، اما پتانسیل الگوریتم‌های جدیدتر الهام‌گرفته از طبیعت، مانند بهینه‌ساز شیطان تاسمانی و الگوریتم شاهین دم‌قرمز، در حوزه هیدرولوژی تا حد زیادی ناشناخته باقی مانده است. مهم‌تر از آن، اکثر مطالعات موجود بر به‌کارگیری یک الگوریتم بهینه‌ساز واحد تمرکز دارند و یک شکاف تحقیقاتی کلیدی در توسعه و ارزیابی چارچوب‌های فراکاوشی هیبریدی که نقاط قوت اکتشافی و استخراجی دو الگوریتم متفاوت را با هم ترکیب می‌کنند وجود دارد. بنابراین، این پژوهش با هدف پر کردن این شکاف‌ها، برای اولین بار یک رویکرد ترکیبی نوین را برای پیش‌بینی هوشمند جریان در یک حوضه نیمه‌خشک و پیچیده مانند دهگلان کردستان ارائه می‌دهد. نوآوری این تحقیق چندوجهی بوده و شکاف‌های کلیدی در ادبیات موجود را برطرف می‌کند. نخست، این پژوهش برای اولین بار دو الگوریتم فراکاوشی جدید و الهام‌گرفته از طبیعت—بهینه‌ساز شیطان تاسمانی و الگوریتم شاهین دم‌قرمز را به حوزه مدل‌سازی هیدرولوژیکی معرفی می‌کند. مهم‌تر از آن، نوآوری اصلی در توسعه یک چارچوب فراکاوشی هیبریدی نهفته است. این رویکرد، با ترکیب هم‌افزای دو الگوریتم متفاوت برای بهبود جستجوی هایپرپارامترهای بهینه، فراتر از رویه رایج استفاده از یک بهینه‌ساز واحد

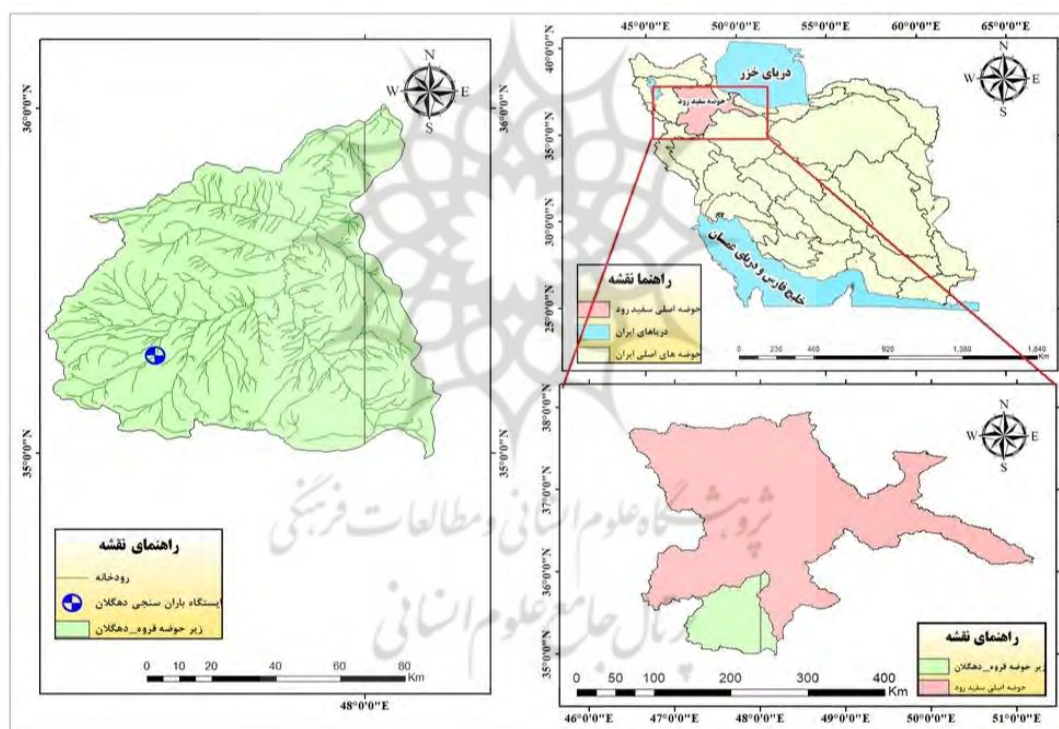
¹ Ismail² Danesh³ Li⁴ Zhang⁵ Salaeh⁶ Tosan⁷ Katsekpor

می‌رود. در نهایت، ارزیابی این متدولوژی پیشرفته در حوضه آبریز پیچیده و نیمه‌خشک دهگلان، اعتبار مهمی برای کارایی و قابلیت تعمیم‌پذیری آن در مسائل چالش‌برانگیز و واقعی مدیریت منابع آب فراهم می‌کند.

مواد و روش‌ها

منطقه مورد مطالعه

شهرستان دهگلان در استان کردستان واقع شده و از شمال به شهرستان بیجار و دیواندره، از شرق به شهرستان قروه، از جنوب به شهرستان کامیاران و سنقر و از غرب به شهرستان سنندج محدود می‌شود. مساحت آن ۲۵۰۰ کیلومتر مربع می‌باشد. حوضه آبریز دهگلان بخشی از حوضه آبریز قزل اوزن است و شامل رودخانه‌های دهگلان، چرخه بیان، تروال و میرکی می‌باشد که آب مورد نیاز منطقه را تامین می‌کنند. دشت دهگلان یکی از بزرگترین دشت‌های کشاورزی استان است و به دلیل اهمیت کشاورزی و توسعه شهری، برداشت از منابع آب زیرزمینی در این منطقه افزایش یافته است. حوضه آبی دهگلان با مساحت ۲۸۷۰ کیلومتر مربع در جنوب شرقی استان کردستان قرار دارد و جزء زیر حوضه آبریز سفیدرود می‌باشد که آب آن به دریای خزر می‌ریزد. این حوضه در طول جغرافیایی ۴۷ درجه و ۳۰ دقیقه تا ۴۷ درجه و ۵۰ دقیقه شرقی و عرض جغرافیایی ۳۵ درجه و ۱۵ دقیقه تا ۳۵ درجه و ۳۰ دقیقه شمالی قرار دارد. ارتفاع از سطح دریا برابر حدود ۱۸۰۰ متر می‌باشد. این حوضه از دو قسمت کوهستان و دشت تشکیل شده است و شکل (۱) موقعیت جغرافیایی منطقه مورد مطالعه را نشان می‌دهد.



شکل (۱): موقعیت جغرافیایی حوضه دهگلان کردستان

Figure (1): Location map of the Dehgholan watershed, Kurdistan

روش‌شناسی تحقیق

گام نخست در این زمینه جمع‌آوری و تحلیل داده‌های میدانی (ایستگاه‌های باران‌سنجی، هیدرومتری و مطالعه خصوصیات فیزیکی حوضه) می‌باشد. گام بعدی شامل تحلیل‌های آماری (تحلیل‌های همبستگی، تحلیل رگرسیون و تحلیل سری‌های زمانی) است. سپس قبل از شروع فرایند مدلسازی، فرایند پیش‌پردازش داده‌ها و مدیریت اطلاعات (شامل حذف یا بازسازی داده‌های مفقوده، حذف داده‌های پرت، نرمال‌سازی داده‌ها) انجام گرفت. بعد از انجام پاکسازی اطلاعات، فرایند تقسیم‌بندی داده‌ها برای مراحل آموزش و آزمون (نسبت ۷۰ به ۳۰) صورت گرفت. گام بعدی تعیین سناریوهای مدل بر اساس روش‌های رایج (روش همبستگی پیرسون، روش MRMR و ...)

می‌باشد. برای این منظور همبستگی بین متغیرهای ورودی و خروجی (از متغیر بارش، دمای متوسط و دبی با تاخیرهای مختلف بعنوان ورودی و دبی رودخانه بعنوان متغیر خروجی) انجام گرفت. بعد از تعیین ارتباط بین متغیرها و تعیین سناریوها، از مدل‌های یادگیری ماشین شبکه عصبی مصنوعی (ANN)، ماشین بردار پشتیبان رگرسیون (SVR) و مدل یادگیری عمیق حافظه طولانی کوتاه مدت (LSTM) جهت پیش بینی استفاده شد. برای بهبود نتایج پیش بینی از مدل‌های فرا ابتکاری شیطان تاسمانی (TDO)، شاهین دم‌قرمز (RTH) و بهینه‌سازی شکارچیان دریایی (MPA) استفاده شد که منجر به تشکیل مدل‌های ANN-TDO، SVR-RTH و LSTM-MPA گردید. در نهایت بر اساس معیارهای ارزیابی مختلف نتایج پیش‌بینی در دو مرحله آموزش و آزمون مورد بررسی قرار گرفته و مدل‌های برتر انتخاب گردید. سپس نتایج مدل بصورت کمی و کیفی مورد ارزیابی قرار گرفته و نمودارهای بصری (نمودار پراکندگی، سری زمانی، دیاگرام تیلور، ویلئون پلات و...) جهت بیان نتایج ترسیم گردید.

الگوریتم‌های فراکاوشی مورد استفاده

الگوریتم بهینه‌سازی شیطان تاسمانی

الگوریتم بهینه‌سازی شیطان تاسمانی (TDO) که توسط دهقانی^۱ و همکاران (۲۰۲۲) برای شبیه‌سازی رفتار طبیعی شیطان تاسمانی در هنگام جستجو برای منبع غذا ارائه شده است. این الگوریتم دو استراتژی حمله و خوردن شکار زنده یا لاشه حیوانات مرده را ارائه می‌دهد (وانگ^۲ و همکاران، ۲۰۲۴: ۱۲). جمعیت اولیه این عوامل به صورت تصادفی بر اساس محدودیت‌های مسئله تولید می‌شود. اعضای جمعیت TDO که جستجوگران فضای حل مسئله هستند، مقادیر منتخب را برای متغیرهای مسئله بر اساس موقعیت خود در فضای جستجو پیشنهاد می‌دهند. بنابراین، از نظر ریاضی، هر عضو از یک جمعیت، برداری است که تعداد عناصر آن برابر با تعداد متغیرهای مسئله است. در نتیجه، مجموعه اعضای TDO را می‌توان طبق رابطه (۱) محاسبه نمود:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N \times m} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix}_{N \times m} \quad (1)$$

در اینجا، X : نشان‌دهنده جمعیت شیطان‌های تاسمانی، X_i : i امین راه‌حل منتخب، $x_{i,j}$ مقدار منتخب آن برای i امین متغیر است. همچنین N : تعداد شیطان‌های تاسمانی جستجوگر و m : تعداد متغیرهای مسائل داده شده است. تابع هدف مسئله را می‌توان با قرار دادن هر یک از راه‌حل‌های منتخب در مقادیر متغیرهای تابع هدف مطابق رابطه (۲) محاسبه کرد:

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N \times 1} \quad (2)$$

در اینجا، F بردار مقادیر تابع هدف است و F_i : مقدار تابع هدف است که توسط i امین راه‌حل منتخب به دست می‌آید. رفتار شیطان تاسمانی در جستجوی منطقه زیستگاه برای یافتن لاشه، مشابه فرایند جستجوی الگوریتم در فضای حل مسئله است. این استراتژی شیطان تاسمانی در واقع قدرت اکتشافی TDO را در اسکن کردن مناطق مختلف فضای جستجو برای شناسایی ناحیه بهینه اصلی نشان می‌دهد. مفاهیم بیان شده در استراتژی خوردن لاشه شیطان تاسمانی با استفاده از روابط (۳) تا (۵) به صورت زیر مدل‌سازی می‌شوند:

$$C_i = X_k, i = 1, 2, \dots, N, k \in \{1, 2, \dots, N \mid k \neq i\} \quad (3)$$

$$x_{i,j}^{new,S1} = \begin{cases} x_{i,j} + r \cdot (c_{i,j} - I \cdot x_{i,j}), & F_{Ci} < F_i \\ x_{i,j} + r \cdot (x_{i,j} - c_{i,j}), & \text{otherwise} \end{cases} \quad (4)$$

$$X_i = \begin{cases} X_i^{new,S1}, & F_i^{new,S1} < F_i \\ X_i, & \text{otherwise} \end{cases} \quad (5)$$

که در این روابط به ترتیب $X_i^{new,S1}$: نشان‌دهنده وضعیت جدید i امین شیطان تاسمانی بر اساس استراتژی اول، $x_{i,j}^{new,S1}$: مقدار آن برای i امین متغیر، $F_i^{new,S1}$: مقدار تابع هدف آن، F_{Ci} : مقدار تابع هدف لاشه انتخاب‌شده است؛ r یک عدد تصادفی در بازه $[0, 1]$ است؛ و I یک عدد تصادفی است که می‌تواند ۱ یا ۲ باشد. مرحله اول انتخاب طعمه و حمله به آن با استفاده از روابط (۶) تا (۸) مدل‌سازی می‌شوند:

$$P_i = X_k, i = 1, 2, \dots, N, k \in \{1, 2, \dots, N \mid k \neq i\} \quad (6)$$

$$x_{i,j}^{new,S2} = \begin{cases} x_{i,j} + r \cdot (p_{i,j} - I \cdot x_{i,j}), & F_{Pi} < F_i \\ x_{i,j} + r \cdot (x_{i,j} - p_{i,j}), & \text{otherwise} \end{cases} \quad (7)$$

$$X_i = \begin{cases} X_i^{new,S2}, & F_i^{new,S2} < F_i; \\ X_i, & \text{otherwise}, \end{cases} \quad (8)$$

که در این روابط به ترتیب P_i : طعمه انتخاب‌شده توسط i امین شیطان تاسمانی، $x_{i,j}^{new,S2}$: نشان‌دهنده وضعیت جدید i -امین شیطان تاسمانی بر اساس استراتژی دوم، $x_{i,j}^{new,S2}$: مقدار آن برای i -امین متغیر، $F_i^{new,S2}$: مقدار تابع هدف و F_{Pi} : مقدار تابع هدف طعمه انتخاب‌شده است. تعقیب طعمه در نزدیکی محل حمله، مشابه جستجوی محلی در فضای جستجو است. این رفتار شیطان تاسمانی در واقع توانایی TDO را در بهره‌برداری برای همگرایی به راه‌حل‌های کاندیدای بهتر نشان می‌دهد.

$$R = 0.01 \left(1 - \frac{t}{T}\right) \quad (9)$$

$$x_{i,j}^{new} = x_{i,j} + (2r - 1) \cdot R \cdot x_{i,j} \quad (10)$$

$$X_i = \begin{cases} X_i^{new}, & F_i^{new} < F_i \\ X_i, & \text{otherwise} \end{cases} \quad (11)$$

الگوریتم بهینه‌سازی شاهین دم‌قرمز

الگوریتم بهینه‌سازی شاهین دم‌قرمز (RTH) که توسط فراهتیا^۱ و همکاران (۲۰۲۳) ارائه شده است، رفتار شکار شاهین دم‌قرمز را در فرایند شکار تقلید می‌کند. این الگوریتم شامل سه مرحله اوج‌گیری بلند، اوج‌گیری کوتاه و شیرجه و غافلگیری می‌باشد. در مرحله اوج‌گیری بلند، شاهین دم‌قرمز به ارتفاعات بالا اوج می‌گیرد و به دنبال بهترین مکان از نظر فراوانی غذا می‌گردد. رابطه (۱۲) مدل ریاضی این مرحله را بیان می‌کند:

$$X(t) = X_{best} + (X_{mean} - X(t-1)) \cdot \text{Levy}(\text{dim}) \cdot TF(t) \quad (12)$$

در این رابطه در اینجا، $X(t)$: نشان‌دهنده موقعیت شاهین دم‌قرمز در تکرار t است. X_{best} : بهترین موقعیت به دست آمده و X_{mean} : میانگین موقعیت‌ها است. Levy نیز تابع توزیع پرواز لوی را نشان می‌دهد که طبق رابطه (۱۳) قابل محاسبه است.

¹ Ferahtia

$$\text{Levy}(\text{dim}) = s \frac{\mu \cdot \sigma}{|v|^{\beta-1}} \quad (13)$$

$$\sigma = \left(\frac{\Gamma(1+\beta) \cdot \sin(\pi\beta/2)}{\Gamma(1+\beta/2) \cdot \beta \cdot 2} \right) (1 - \beta/2)$$

در نهایت، $TF(t)$: تابع فاکتور گذار است که می‌توان آن را طبق رابطه (۱۴) محاسبه کرد:

$$TF(t) = 1 + \sin(2.5 + (t/T_{max})) \quad (14)$$

که در آن S : یک ثابت (۰.۱)، dim : بعد مسئله، β : یک ثابت (۵.۱) است، u و v اعداد تصادفی [۰ تا ۱] هستند. که در آن T_{max} نشان‌دهنده حداکثر تعداد تکرارها است. پرواز در ارتفاع گام بعدی می‌باشد. در این مرحله، شاهین با پرواز در ارتفاع بسیار پایین‌تر از سطح زمین و به صورت خطی ماریپیچ، طعمه را محاصره نموده و مدل آن را می‌توان از معادله (۱۵) بصورت زیر بدست آورد:

$$\begin{aligned} X(t) &= X_{best} + (x(t) + y(t)) \cdot \text{StepSize}(t) \\ \text{StepSize}(t) &= X(t) - X_{mean} \end{aligned} \quad (15)$$

که در آن x و y مختصات جهت را نشان می‌دهند که می‌توانند به صورت زیر از طریق رابطه (۱۶) محاسبه شوند:

$$\begin{cases} x(t) = R(t) \cdot \sin(\theta(t)) \\ y(t) = R(t) \cdot \cos(\theta(t)) \end{cases} \begin{cases} R(t) = R_0 \cdot (r - t/T_{max}) \cdot \text{rand} \\ \theta(t) = A \cdot (1 - t/T_{max}) \cdot \text{rand} \end{cases} \begin{cases} x(t) = x(t)/\max|x(t)| \\ y(t) = y(t)/\max|y(t)| \end{cases} \quad (16)$$

که در آن R_0 : نشان‌دهنده مقدار اولیه شعاع [۰.۵-۳]، A : نشان‌دهنده بهره زاویه [۵-۱۵]، rand : یک بهره تصادفی [۰-۱] و r یک بهره کنترلی [۱، ۲] است. این پارامترها به شاهین کمک می‌کنند تا با حرکات ماریپیچی به دور طعمه پرواز کند.

$$X(t) = \alpha(t) \cdot X_{best} + x(t) \cdot \text{StepSize} 1(t) + y(t) \cdot \text{StepSize} 2(t) \quad (17)$$

که در آن اندازه هر گام را می‌توان به صورت زیر محاسبه از طریق رابطه (۱۸) محاسبه کرد:

$$\begin{aligned} \text{StepSize} 1(t) &= X(t) - TF(t) \cdot X_{mean} \\ \text{StepSize} 2(t) &= G(t) \cdot X(t) - TF(t) \cdot X_{best} \end{aligned} \quad (18)$$

که در آن α و G به ترتیب عامل‌های شتاب و گرانش هستند که می‌توانند به صورت زیر از طریق رابطه (۱۹) تعریف گردد:

$$\begin{aligned} \alpha(t) &= \sin^2(2.5 - t/T_{max}) \\ G(t) &= 2 \cdot (1 - t/T_{max}) \end{aligned} \quad (19)$$

که در آن α : نشان‌دهنده شتاب شاهین است که با افزایش t (زمان یا تکرار) افزایش می‌یابد تا سرعت همگرایی را بهبود بخشد، G اثر گرانش است که کاهش می‌یابد تا تنوع بهره‌برداری را زمانی که شاهین بسیار به طعمه نزدیک است، کاهش دهد.

الگوریتم بهینه‌سازی شکارچیان دریایی

الگوریتم شکارچیان دریایی یک الگوریتم بهینه‌سازی فراابتکاری است که بر اساس قانون جستجوی غذای شکارچیان توسعه داده شده است (فامرزی^۱ و همکاران، ۲۰۲۰: ۱۵۲). این الگوریتم از رفتارهای شکارچی و استراتژی‌های جستجوی غذای موجودات دریایی الهام گرفته است و با شبیه‌سازی قوانین این موجودات دریایی در شکار طعمه‌هایشان، مسئله بهینه‌سازی را حل می‌کند (شین^۲ و همکاران، ۲۰۲۳: ۱۲۳). روش‌های جستجوی غذا عمدتاً شامل دو دسته پرواز لوی و حرکت براونی هستند. در فرآیند بهینه‌سازی، استراتژی‌های جستجوی غذا بر اساس سرعت‌های حرکتی متفاوت بین شکارچی و طعمه انتخاب می‌شوند. احتمال برخورد بین شکارچی و طعمه با

¹ Faramarzi

² Shen

سنجش انتخاب دو استراتژی جستجوی غذا، پرواز لوی و حرکت براونی، در مراحل مختلف به حداکثر می‌رسد (مرادی فر^۱ و همکاران، ۲۰۲۴: ۱۲۳). راه حل‌های اولیه در الگوریتم شکارچیان دریایی که به صورت یکنواخت توزیع شده است از رابطه (۲۰) بصورت زیر بدست می‌آید:

$$X_0 = X_{lb} + rand(X_{ub} - X_{lb}) \quad (20)$$

که در این رابطه به ترتیب که در آن rand: یک بردار تصادفی است که به صورت یکنواخت بین ۰ و ۱ توزیع شده است، X_{lb} : حداقل مرز (کران پایین) متغیرها و X_{ub} : حداکثر مرز (کران بالا) متغیرها است. ماتریس طعمه طبق رابطه (۲۱) بصورت زیر بدست می‌آید:

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,Dim} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,Dim} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,2} & X_{n,2} & X_{n,Dim} \end{bmatrix}_{n \times Dim} \quad (21)$$

که در آن X_j^i : نشان‌دهنده بُعد j-ام از طعمه i-ام، Dim: بُعد عوامل جستجو و n: تعداد عوامل جستجو است. موقعیت به روش زیر طبق رابطه (۲۲) به‌روزرسانی می‌شود:

$$\begin{aligned} \overrightarrow{step}_i &= \overrightarrow{M_B} \otimes (\overrightarrow{El}_i - \overrightarrow{M_B} \otimes \overrightarrow{Pre}_i), i = 1, \dots, n \\ \overrightarrow{Pre}_i &= \overrightarrow{Pre}_i + C * \overrightarrow{rand} \otimes \overrightarrow{step}_i \end{aligned} \quad (22)$$

که در این رابطه به ترتیب $\overrightarrow{M_B}$: برداری است که حرکت براونی را نشان می‌دهد. C: یک عدد ثابت است. همچنین $\otimes$ ضرب جزء به جزء است. rand نیز برداری است که به صورت تصادفی و یکنواخت بین ۰ و ۱ توزیع شده است. به‌روزرسانی دقیق موقعیت طبق رابطه (۲۳) بصورت زیر انجام خواهد شد:

$$\begin{aligned} \overrightarrow{step}_i &= \overrightarrow{M_L} \otimes (\overrightarrow{El}_i - \overrightarrow{M_L} \otimes \overrightarrow{Pre}_i), i = 1, \dots, n/2 \\ \overrightarrow{Pre}_i &= \overrightarrow{Pre}_i + C * \overrightarrow{rand} \otimes \overrightarrow{step}_i \end{aligned} \quad (23)$$

که در آن $\overrightarrow{M_L}$: برداری است که حرکات لوی را نشان می‌دهد. برای نیمه دیگر عوامل جستجو از روابط (۲۴) و (۲۵) استفاده خواهیم نمود.

$$\begin{aligned} \overrightarrow{step}_i &= \overrightarrow{M_B} \otimes (\overrightarrow{M_B} \otimes \overrightarrow{El}_i - \overrightarrow{Pre}_i), i = n/2 + 1, \dots, n \\ \overrightarrow{Pre}_i &= \overrightarrow{El}_i + C * CF \otimes \overrightarrow{step}_i \end{aligned} \quad (24)$$

$$CF = (1 - t/t_{max})(2t/t_{max}) \quad (25)$$

که در آن CF: پارامتری است که در فرآیند بهینه‌سازی برای کنترل اندازه گام استفاده می‌شود.

$$\begin{aligned} \overrightarrow{step}_i &= \overrightarrow{M_L} \otimes (\overrightarrow{M_L} \otimes \overrightarrow{El}_i - \overrightarrow{Pre}_i), i = 1, \dots, n \\ \overrightarrow{Pre}_i &= \overrightarrow{El}_i + C * CF \otimes \overrightarrow{step}_i \end{aligned} \quad (26)$$

$$\overrightarrow{Pre}_i = \begin{cases} \overrightarrow{Pre}_i + CF[\overrightarrow{X}_{lb} + \overrightarrow{rand} \otimes (\overrightarrow{X}_{ub} - \overrightarrow{X}_{lb})] \otimes \vec{U}, & \text{if } r \leq FADs \\ \overrightarrow{Pre}_i + [FADs(1 - r) + r](\overrightarrow{Pre}_{r1} - \overrightarrow{Pre}_{r2}), & \text{if } r > FADs \end{cases} \quad (27)$$

جمع آوری داده و متغیرهای آماری

در این پژوهش، ابتدا فرایند جمع آوری و اخذ داده‌ها از شرکت آب منطقه‌ای استان کردستان و سازمان هواشناسی کشور آغاز گردید. سری زمانی داده‌ها در مقیاس روزانه می‌باشد. بدین منظور آمار ۲۰ ساله بارش (P)، دبی رودخانه (Q) و دمای میانگین (T_{ave}) از سال ۱۳۸۰ تا ۱۴۰۰ اخذ گردیده و مورد استفاده قرار گرفت. جدول (۱) مقادیر پارامترهای آماری تحقیق (کل داده‌ها) را نشان می‌دهد.

جدول (۱): پارامترهای آماری منطقه مورد مطالعه

Table (1): Statistical parameters of the study area

پارامتر	کمینه	بیشینه	میانگین	انحراف معیار	چولگی	کشیدگی
بارش (mm)	۶۷/۵۰	۰	۰/۹۹	۳/۸۷	۶/۷۵	۶۲/۹۴
دبی (m^3/s)	-۲۴	۵۵	۱۱/۹۳	۹/۵۴	-۰/۰۱	-۰/۸۶
دمای میانگین ($^{\circ}C$)	۰	۳۳/۴۷	۰/۷۵	۱/۸۰	۵/۹۴	۵۵/۶۲

ترکیب مدل و انتخاب سناریوها

پیش‌بینی جریان رودخانه فرآیندی پیچیده است که تحت تأثیر عوامل متعددی از قبیل بارش، دمای هوا، ویژگی‌های حوضه آبریز (از قبیل مساحت، شیب، نوع خاک، پوشش گیاهی و کاربری اراضی) دبی گذشته رودخانه (برای ثبت حافظه سیستم و اثرات پویایی حوضه)، تبخیر و تعرق بالقوه، رطوبت خاک قرار دارد. تعیین سناریوهای مدل برای پیش‌بینی جریان رودخانه شامل تعریف ورودی‌های مختلف و حالات ممکن برای پارامترهای مؤثر است. گام نخست جهت تعیین سناریو و ترکیب مدل، انتخاب متغیرها و یافتن ارتباط و همبستگی بین متغیرهای ورودی و خروجی می‌باشد. با توجه به این که بسیاری از متغیرها همبستگی بسیار کمی با متغیر خروجی داشته و عملاً اثرات آنها در مدلسازی بسیار ناچیز است از اثرات آنها صرف‌نظر می‌شود. همچنین محاسبه و تقریب بسیاری از پارامترهای مربوط به فیزیک خاک و دستیابی به داده‌های آن دشوار می‌باشد. لذا در این پژوهش جهت پیش‌بینی جریان رودخانه از متغیرهای بارش (P_t)، دمای میانگین (T_{ave}) و دبی جریان از یک تا سه روز تاخیر (Q_{t-1} تا Q_{t-3}) استفاده گردید. ارتباط بین متغیرهای ورودی و خروجی در جدول (۲) نشان داده شد. بر این اساس انتخاب سناریو و ترکیب مدل‌ها بر اساس ضریب همبستگی پیرسون انجام گرفت. در این راستا اولین سناریو با بیشترین همبستگی می‌باشد. جدول (۳) ترکیب و سناریوهای مدل برای منطقه مورد مطالعه را نشان می‌دهد.

جدول (۲): ارتباط بین متغیرهای ورودی و خروجی پژوهش

Table (2): Correlation between the input and output variables of the study

Q(t)	Q(t-3)	Q(t-2)	Q(t-1)	T_{ave}	P (t)
۱	۰/۷۳۱	۰/۷۹۵	۰/۸۹۹	۰/۰۲۳	۰/۲۳۳

جدول (۳): انتخاب ترکیب و سناریوهای مدل بر اساس ضریب پیرسون

Table (3): Selection of model combinations and scenarios based on the Pearson coefficient

خروجی	ترکیب ورودی	نام سناریو
Q(t)	Q (t-1)	M ₁
Q(t)	Q (t-1), Q (t-2)	M ₂
Q(t)	Q (t-1), Q (t-2), Q (t-3)	M ₃
Q(t)	Q (t-1), Q (t-2), Q (t-3), P (t)	M ₄
Q(t)	Q (t-1), Q (t-2), Q (t-3), P (t), T_{ave}	M ₅

معیارهای ارزیابی مدل

جهت ارزیابی مدل و بررسی دقت پیش‌بینی از معیارهای ارزیابی میانگین مربعات خطا (MAE)، ریشه میانگین مربعات خطا (RMSE)، ضریب تبیین (R^2) و ضریب بهره‌وری کلینگ-گوپتا استفاده گردید. معادلات (۲۸) تا (۳۱) روابط ریاضی و محدوده آنها را نشان می‌دهد:

$$MAE = \frac{1}{n} \sum_{j=1}^n |Q_{prd} - Q_{obs}| \quad (0 < MAE < \infty) \quad (28)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{prd} - Q_{obs})^2} \quad (0 < RMSE < \infty) \quad (29)$$

$$R^2 = 1 - \frac{\sum (\sqrt{Q_{prd}} - \sqrt{Q_{obs}})^2}{\sum (\sqrt{Q_{obs}} - \sqrt{\bar{Q}})^2} \quad (0 < R^2 < 1) \quad (30)$$

$$KGE = 1 - \sqrt{(CC - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (0 < KGE < 1) \quad (31)$$

که در روابط بالا به ترتیب Q_{prd} : دبی پیش‌بینی یا شبیه‌سازی شده جریان رودخانه (حسب مترمکعب بر ثانیه)، Q_{obs} : دبی مشاهداتی یا واقعی، $\bar{Q}$: میانگین دبی مشاهداتی، CC : ضریب همبستگی خطی بین دبی شبیه‌سازی و مشاهداتی، ضریب α : نسبت انحراف استاندارد (انحراف معیار) دبی شبیه‌سازی شده بر دبی مشاهداتی، ضریب β : نسبت میانگین دبی شبیه‌سازی شده بر دبی مشاهداتی می‌باشد.

بحث و نتایج

در این تحقیق برای پیش‌بینی شبیه‌سازی جریان رودخانه دهگلان از مدل‌های هوش مصنوعی ترکیب شده با مدل‌های فرا ابتکاری استفاده گردید. بر این اساس بعد جمع‌آوری اطلاعات و مدیریت داده‌ها و انجام آزمون‌های اولیه تقسیم‌بندی داده‌ها در فرایند آموزش، آزمون و اعتبارسنجی انجام گرفت. سناریوهای ترکیب مدل‌ها بر اساس ضریب همبستگی پیرسون تعیین شد و سپس هر سناریو به صورت جداگانه اجرا گردید. کلیه مدل‌ها در دو فاز منفرد و هیبریدی مورد بررسی قرار گرفت. بر این اساس ابتدا مدل‌های منفرد اجرا گردید. سپس برای بهینه کردن ضرایب درونی و هایپرپارامترهای آن از مدل‌های فرا ابتکاری استفاده گردید. مقایسه نتایج عملکرد مدل منفرد و هیبریدی تحقیق در جدول (۴) نشان داده شد. یافته‌های پژوهش نشان داد که در فاز آموزش در هر دو مدل منفرد ANN و هیبریدی ANN-TDO، سناریوی M_1 بهترین عملکرد را به ثبت رسانده است. در مدل ANN منفرد، این سناریو $R^2=0.843$ و مقدار $RMSE=0.061 m^3/s$ است. در مدل هیبریدی ANN-TDO نیز سناریوی M_1 با $R^2=0.881$ و $RMSE=0.054 m^3/s$ برترین عملکرد را داشته است. این نشان می‌دهد که ترکیب ورودی در سناریوی M_1 بهترین قابلیت را برای آموزش شبکه داشته است. در فاز آزمون در

مدل ANN منفرد، مجدداً سناریوی M_1 با $R^2=0.769$ بهترین نتیجه را نشان می‌دهد. اما در مدل هیبریدی ANN-TDO، سناریوی M_5 با کسب $R^2=0.852$ و $RMSE=0.059 m^3/s$ ، دقیق‌ترین پیش‌بینی را در فاز ارزیابی نهایی ارائه کرده است. این موضوع نشانگر توانایی بالای الگوریتم TDO در بهینه‌سازی پارامترهای شبکه برای تعمیم‌بخشی بهتر به داده‌های جدید در سناریوی M_5 است. در هر دو مدل ANN و ANN-TDO، عملکرد مدل در فاز آموزش به طور کلی بهتر از فاز آزمون است. در بهترین سناریوی مدل ANN-TDO (سناریوی M_1) در آموزش و M_5 در آزمون، مقدار R^2 از ۰/۸۸۱ در فاز آموزش به ۰/۸۵۲ در فاز آزمون کاهش یافته و $RMSE$ از ۰/۰۵۴ به ۰/۰۵۹ افزایش می‌یابد. این افت عملکرد جزئی، امری طبیعی در مدل‌سازی است و نشان می‌دهد که مدل‌ها دچار بیش‌برازش شدید نشده‌اند و قابلیت تعمیم‌بخشی قابل قبولی به داده‌های مشاهده نشده دارند. با این حال، نزدیکی بیشتر نتایج دو فاز در مدل هیبریدی ANN-TDO بیانگر پایداری و قدرت تعمیم بالاتر آن نسبت به مدل منفرد ANN است. در فاز آموزش بهترین عملکرد مدل ANN-TDO در سناریوی M_1 با ثبت R^2 برابر با ۰/۸۸۱ و ضعیف‌ترین عملکرد مدل ANN در سناریوی M_4 با R^2 برابر ۰/۳۴۹ دارای کمترین توانایی را در یادگیری الگوهای حاکم بر داده‌های آموزشی از خود نشان داده است. مقایسه کلی مدل‌ها با یکدیگر نیز نشان داد که بهترین عملکرد مدل-SVR در RTH در سناریوی M_5 با کسب R^2 برابر ۰/۸۸۲ به عنوان بهترین مدل و ضعیف‌ترین عملکرد مدل LSTM منفرد در سناریوی M_5 با

R^2 برابر با ۰/۳۹۴ دارای ضعیف‌ترین نتیجه را در مواجهه با داده‌های جدید کسب کرده است. همچنین مقایسه نتایج مدل منفرد و هیبریدی در میزان افزایش دقت مورد ارزیابی قرار گرفت. نتایج نشان دادند که استفاده از الگوریتم‌های بهینه‌سازی به طور چشمگیری عملکرد مدل‌های پایه را بهبود بخشیده است. بر این اساس مدل ANN-TDO در برابر مدل منفرد ANN در فاز آزمون، با مقایسه سناریوی M_5 ، مدل هیبریدی توانست R^2 را از ۰/۵۶۷ به ۰/۸۵۲ برساند که نشان‌دهنده بهبود عملکردی معادل ۵۰/۲ درصدی می‌باشد. همچنین مدل SVR-RTH در برابر SVR در فاز آزمون، با مقایسه سناریوی M_5 ، مدل هیبریدی R^2 را از ۰/۷۹۹ به ۰/۸۸۲ افزایش داد که به معنای بهبود ۱۰/۴ درصدی در دقت پیش‌بینی است. مدل LSTM-MPA در برابر LSTM بهترین بهبود در این مدل مشاهده شد. در فاز آزمون، با مقایسه سناریوی M_3 ، مدل هیبریدی مقدار R^2 را از ۰/۴۹۴ به ۰/۸۵۱ ارتقا داد که نشان‌دهنده بهبود فوق‌العاده ۷۲/۲ درصدی است. این نتایج به وضوح نشان می‌دهد که ترکیب مدل‌های هوش مصنوعی با الگوریتم‌های فراابتکاری (مانند RTH، TDO و MPA) یک استراتژی بسیار مؤثر برای افزایش دقت و قابلیت اطمینان مدل‌های پیش‌بینی جریان رودخانه است، به ویژه در مدل LSTM که بهبود قابل توجهی را تجربه کرده است. شکل (۲) نمودار توزیع فراوانی و هیستوگرام خطای مدل‌های هیبریدی را نشان می‌دهد. تحلیل نمودار مذکور نشان داد که مدل ANN-TDO با میانگین خطای ($\mu = 0.1252$) و انحراف معیار ($\sigma = 0.1424$)، بهترین عملکرد را در میان سه مدل از خود نشان می‌دهد. مقدار پایین میانگین (μ) بیانگر آن است که به طور متوسط، خطاهای تولید شده توسط این مدل به صفر نزدیک‌تر هستند. به عبارت دیگر، این مدل دارای کمترین میزان بایاس بوده و پیش‌بینی‌های آن به مقادیر واقعی نزدیک‌تر است. انحراف معیار (σ) پایین، نشان‌دهنده پراکندگی کم خطاها حول میانگین است. این یعنی مدل ANN-TDO عملکردی بسیار پایدار و قابل اعتماد دارد و در اکثر پیش‌بینی‌ها، خطایی نزدیک به میانگین پایین خود تولید می‌کند. تمرکز بالای داده‌ها در بازه خطای نزدیک به صفر موید همین موضوع می‌باشد. مدل SVR-RTH نیز با ثبت بالاترین میانگین خطا ($\mu = 0.22015$) و بیشترین انحراف معیار ($\sigma = 0.26558$)، ضعیف‌ترین عملکرد را در این مقایسه دارد. میانگین خطای بالاتر نشان می‌دهد که پیش‌بینی‌های این مدل به طور متوسط فاصله بیشتری از مقادیر واقعی دارند. همچنین انحراف معیار بالا بیانگر پراکندگی زیاد خطاهاست. این بدان معناست که مدل SVR-RTH در برخی موارد پیش‌بینی‌های خوبی ارائه می‌دهد، اما در موارد دیگر خطاهای بزرگی تولید می‌کند و عملکردی ناپایدارتر و غیرقابل اعتمادتر نسبت به دو مدل دیگر دارد. تحلیل توزیع فراوانی خطا نشان داد که اگرچه هر سه مدل هیبریدی توانایی بالایی در پیش‌بینی جریان رودخانه دارند، اما مدل ANN-TDO به دلیل داشتن کمترین میانگین و انحراف معیار خطا، به عنوان دقیق‌ترین، پایدارترین و قابل اعتمادترین مدل در این پژوهش معرفی می‌شود. نمودار پراکندگی مدل‌های هیبریدی در اشکال (۳) تا (۵) نشان داده شد. شکل (۶) نیز نمودار سری زمانی داده‌های مشاهداتی و پیش‌بینی شده با مدل‌های هیبریدی را نشان می‌دهد. همچنین اشکال (۷ و ۸) نیز نمودارهای همگرایی، فرآیند بهینه‌سازی مدل‌های هیبریدی را با نشان دادن کاهش تابع هزینه (RMSE) در طول تکرارهای الگوریتم نشان می‌دهد. نتایج نشان دادند که از نظر سرعت همگرایی و کارایی محاسباتی، مدل SVR-RTH با اختلاف زیاد، برترین عملکرد را دارد. از نظر دستیابی به بهترین راه‌حل نهایی و کمترین خطای ممکن، مدل LSTM-MPA با توانایی جستجوی هوشمندانه خود، موفق‌ترین مدل است. مدل ANN-TDO نیز کارایی خود را در بهینه‌سازی یک مسئله پیچیده نشان داد، اما از نظر سرعت و دقت نهایی در جایگاه پایین‌تری نسبت به دو رقیب خود قرار گرفت. شکل (۹) نیز نمودار ویلئون پلات داده‌های مشاهداتی (واقعی) و داده‌های پیش‌بینی شده توسط مدل‌های هیبریدی را نشان می‌دهد. تحلیل نمودار ویلئون به وضوح برتری مدل‌های ANN-TDO و LSTM-MPA را تأیید می‌کند. این دو مدل نه تنها از نظر دقت، بلکه از نظر توانایی بازتولید ساختار آماری و توزیع احتمالاتی جریان رودخانه نیز بسیار موفق عمل کرده‌اند. در مقابل، مدل SVR-RTH علی‌رغم عملکرد مناسب در برخی معیارهای دیگر، به دلیل تمایل به بیش‌برآورد و به خصوص تولید مقادیر پرت شدید، به عنوان یک مدل کمتر قابل اعتماد ارزیابی می‌شود. بنابراین، برای کاربردهای عملی که در آن پیش‌بینی صحیح مقادیر حدی و همچنین بازتولید رفتار طبیعی سیستم اهمیت دارد، مدل‌های ANN-TDO و LSTM-MPA گزینه‌های

بسیار برتر و قابل اتکایی هستند. در ادامه به ارزیابی مطالعات انجام گرفته با نتایج این پژوهش پرداخته شد. گومان^۱ و همکاران (۲۰۱۸) به مقایسه انواع مختلف مدل‌های شبکه عصبی مصنوعی (ANN) و رگرسیون بردار پشتیبان (SVR) برای پیش‌بینی جریان ماهانه در رود سند علیا پرداختند. نتایج نشان داد که مدل ANN بهینه‌سازی شده با الگوریتم‌های شانون و فلچر و مدل SVR با کرنل تابع پایه شعاعی بهترین عملکرد را در میان مدل‌های مورد آزمایش داشته که با نتایج این مدل در برتری شبکه عصبی منطبق می‌باشد. گیمیر^۲ و همکاران (۲۰۲۱) برای پیش‌بینی ساعتی جریان رودخانه در استرالیا مدل‌های یادگیری عمیق LSTM، CNN و DNN و مدل هیبریدی CNN-LSTM را مورد ارزیابی قرار دادند. نتایج نشان داد که این مدل ترکیبی به‌طور قابل توجهی از مدل‌های منفرد عملکرد بهتری داشته است که با نتایج این تحقیق در برتری مدل‌های هیبریدی بر منفرد انطباق دارد. عصام^۳ و همکاران (۲۰۲۲) مدل‌های SVM، ANN و LSTM را در ۱۱ رودخانه در شبه جزیره مالزی مقایسه کردند. نتایج نشان داد که مدل ANN به دلیل قابلیت اطمینان و سازگاری بالا، بهترین عملکرد کلی را داشته و به عنوان یک مدل جامع برای منطقه پیشنهاد شد که با نتایج این پژوهش همسو می‌باشد. آکینر^۴ و همکاران (۲۰۲۴) به مقایسه مدل‌های ANN و LSTM برای پیش‌بینی جریان رودخانه در حوضه یشیلیرماک ترکیه پرداختند. نتایج نشان داد که یک مدل ANN که به روش آماری و با طراحی آزمایش‌ها بهینه‌سازی شده بود، عملکردی برتر از مدل‌های LSTM داشته و به بالاترین دقت پیش‌بینی دست یافته است که با نتایج این پژوهش در برتری ANN بر LSTM منطبق می‌باشد.

جدول (۴): مقایسه نتایج عملکرد مدل منفرد ANN و هیبریدی ANN-TDO در دو فاز آموزش و آزمون

Table (4): Performance comparison of the standalone ANN and hybrid ANN-TDO models during the training and testing phases

فرایند آموزش (۷۰ درصد)				فرایند آزمون (۳۰ درصد)				ساختار شبکه	سناریوها
KGE	RMSE(m ³ /s)	MAE (m ³ /s)	R ²	KGE	RMSE(m ³ /s)	MAE (m ³ /s)	R ²		
مدل منفرد ANN									
۰/۸۶۸	۰/۰۶۱	۰/۰۲۵	۰/۸۴۳	۰/۸۴۴	۰/۰۷۵	۰/۰۲۸	۰/۷۶۹	۱-۲۰-۱	M ₁
۰/۸۳۸	۰/۰۷۷	۰/۰۳۴	۰/۷۵۰	۰/۸۱۸	۰/۰۸۷	۰/۰۳۵	۰/۶۹۱	۲-۲۰-۱	M ₂
۰/۶۹۶	۰/۰۹۲	۰/۰۹۲	۰/۶۷۰	۰/۶۹۶	۰/۰۹۷	۰/۰۴۹	۰/۵۹۹	۳-۲۰-۱	M ₃
۰/۵۸۸	۱۲۷.۰	۰/۰۷۰	۰/۳۴۹	۰/۶۳۷	۰/۱۱۳	۰/۰۶۵	۰/۴۷۲	۴-۲۰-۱	M ₄
۰/۶۰۴	۰/۱۱۱	۰/۰۷۹	۰/۴۴۱	۰/۶۰۳	۰/۱۱۵	۰/۰۸۱	۰/۴۶۷	۵-۲۰-۱	M ₅
مدل هیبریدی ANN-TDO									
۰/۹۱۵	۰/۰۵۴	۰/۰۲۱	۰/۸۸۱	۰/۸۸۵	۰/۰۶۳	۰/۰۲۲	۰/۸۵۱	۱-۲۰-۱	M ₁
۰/۸۵۳	۰/۰۷۷	۰/۰۲۷	۰/۷۴۸	۰/۸۴۷	۰/۰۶۵	۰/۰۲۴	۰/۸۲۴	۲-۲۰-۱	M ₂
۰/۸۷۲	۰/۰۶۴	۰/۰۲۵	۰/۸۳۹	۰/۸۷۹	۰/۰۶۲	۰/۰۲۴	۰/۸۳۵	۳-۲۰-۱	M ₃
۰/۸۸۹	۰/۰۶۷	۰/۰۲۴	۰/۸۱۶	۰/۸۸۷	۰/۰۶۰	۰/۰۲۳	۰/۸۴۷	۴-۲۰-۱	M ₄
۰/۹۰۳	۰/۰۵۹	۰/۰۲۳	۰/۸۴۰	۰/۸۹۸	۰/۰۶۰	۰/۰۲۳	۰/۸۵۲	۵-۲۰-۱	M ₅
مدل منفرد SVR									
۰/۶۱۱	۰/۰۹۳	۰/۰۶۱	۰/۷۸۹	۰/۶۲۷	۰/۰۹۳	۰/۰۶۰	۰/۷۶۵		M ₁
۰/۵۷۳	۰/۰۷۵	۰/۰۶۴	۰/۶۶۷	۰/۶۷۳	۰/۰۷۵	۰/۰۵۸	۰/۷۵۹		M ₂
۰/۵۸۷	۰/۰۷۹	۰/۰۶۳	۰/۷۰۹	۰/۶۳۵	۰/۰۷۹	۰/۰۶۳	۰/۷۳۶		M ₃
۰/۴۹۲	۰/۰۷۸	۰/۰۶۴	۰/۶۹۸	۰/۵۴۹	۰/۰۷۸	۰/۰۶۲	۰/۷۴۶		M ₄
۰/۴۳۰	۰/۷۱۰	۰/۰۶۶	۰/۶۳۲	۰/۵۸۲	۰/۷۱۰	۰/۰۵۸	۰/۷۹۹		M ₅
مدل هیبریدی SVR-RTH									
۰/۸۶۱	۰/۰۸۳	۰/۰۴۲	۰/۸۱۶	۰/۸۶۶	۰/۰۸۵	۰/۰۴۰	۰/۸۰۹		M ₁
۰/۷۸۹	۰/۰۷۸	۰/۰۴۸	۰/۷۵۵	۰/۸۴۰	۰/۰۵۶	۰/۰۳۹	۰/۸۶۵		M ₂
۰/۷۴۹	۰/۰۸۰	۰/۰۸۰	۰/۰۵۲	۰/۸۰۱	۰/۰۵۴	۰/۰۴۰	۰/۸۷۸		M ₃

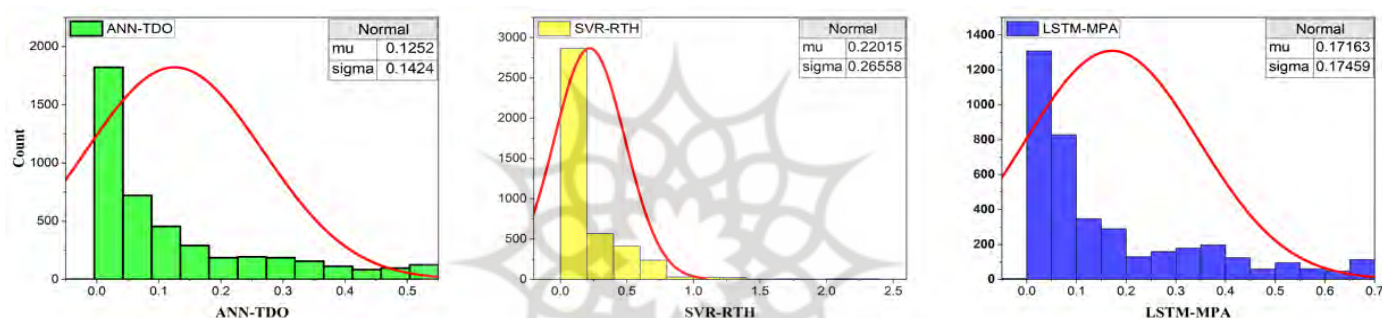
¹ Ghumman

² Ghimire

³ Essam

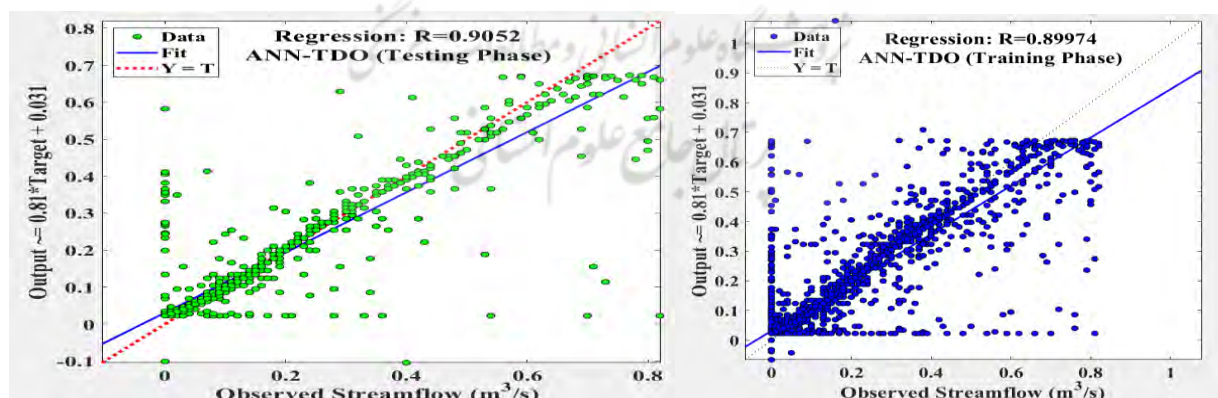
⁴ Akiner

۰/۸۰۴	۰/۰۶۵	۰/۰۴۰	۰/۸۲۳	۰/۰۵۷	۰/۰۵۷	۰/۰۳۶	۰/۸۶۷		M ₄
۰/۷۶۱	۰/۰۸۶	۰/۰۵۴	۰/۶۹۸	۰/۸۳۹	۰/۰۵۳	۰/۰۳۹	۰/۸۸۲		M ₅
مدل منفرد LSTM									
۰/۴۵۱	۰/۱۳۹	۰/۰۹۸	۰/۵۱۰	۰/۴۴۴	۰/۱۴۱	۰/۹۷۰	۰/۴۷۰		M ₁
۰/۴۲۰	۰/۱۱۶	۰/۰۷۹	۰/۴۱۱	۰/۴۲۲	۰/۱۱۶	۰/۰۸۱	۰/۴۴۹		M ₂
۰/۴۴۰	۰/۱۱۱	۰/۰۷۸	۰/۴۸۵	۰/۴۶۳	۰/۱۱۱	۰/۰۷۷	۰/۴۹۴		M ₃
۰/۴۸۳	۰/۱۱۵	۰/۰۷۴	۰/۴۶۴	۰/۴۸۸	۰/۱۰۹	۰/۰۷۴	۰/۵۰۴		M ₄
۰/۳۷۰	۰/۱۲۴	۰/۰۸۹	۰/۳۷۱	۰/۳۶۴	۰/۱۲۱	۰/۰۸۹	۰/۳۹۴		M ₅
مدل هیبریدی LSTM-MPA									
۰/۸۳۸	۰/۰۹۱	۰/۰۳۷	۰/۷۹۰	۰/۸۶۰	۰/۰۸۱	۰/۰۳۳	۰/۸۲۳		M ₁
۰/۸۸۱	۰/۰۶۹	۰/۰۲۳	۰/۷۸۹	۰/۸۸۵	۰/۰۶۲	۰/۰۲۳	۰/۸۴۲		M ₂
۰/۸۶۳	۰/۰۶۶	۰/۰۲۴	۰/۸۱۷	۰/۸۹۴	۰/۰۶۰	۰/۰۲۲	۰/۸۵۱		M ₃
۰/۸۹۱	۰/۰۶۷	۰/۰۲۵	۰/۸۱۵	۰/۸۸۴	۰/۰۶۱	۰/۰۲۴	۰/۸۴۶		M ₄
۰/۸۸۸	۰/۰۶۸	۰/۰۲۶	۰/۸۱۲	۰/۸۹۰	۰/۰۶۲	۰/۰۲۴	۰/۸۴۱		M ₅



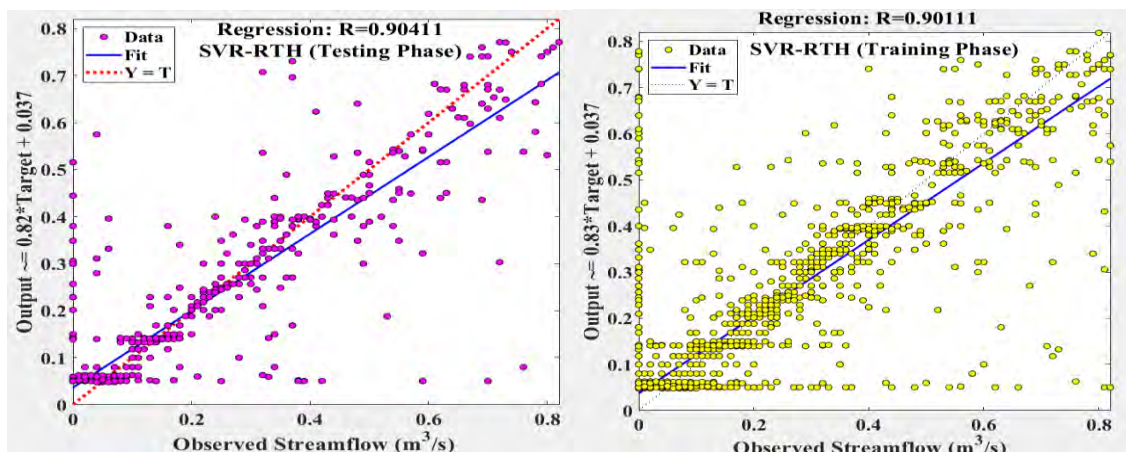
شکل (۲) نمودار توزیع فراوانی و هیستوگرام خطای مدل‌های هیبریدی

Figure (2): Frequency distribution plot and error histogram of the hybrid models.

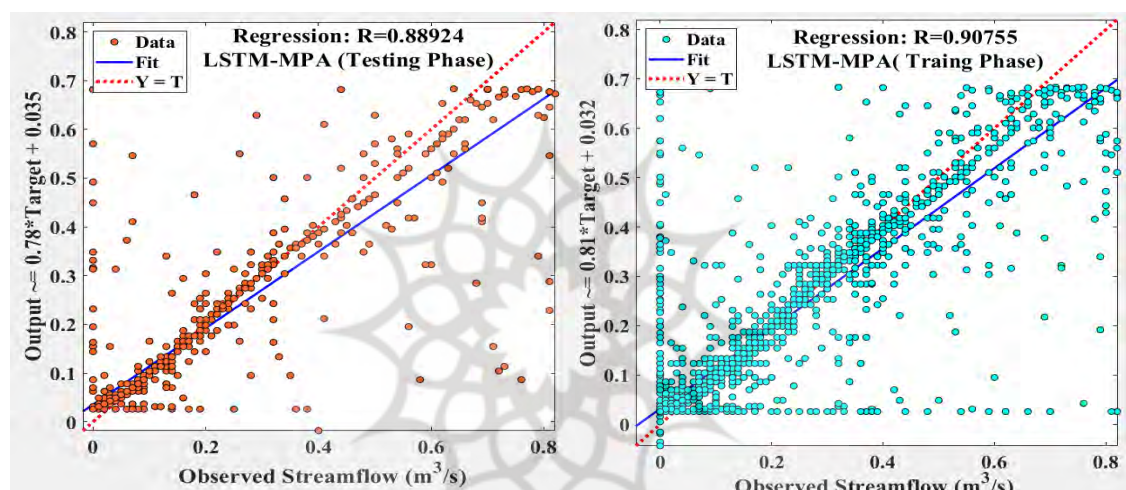


شکل (۳): نمودار پراکندگی مدل هیبریدی ANN-TDO در دو فاز آموزش و آزمون

Figure (3): Scatter plot of the hybrid ANN-TDO model in the training and testing phases

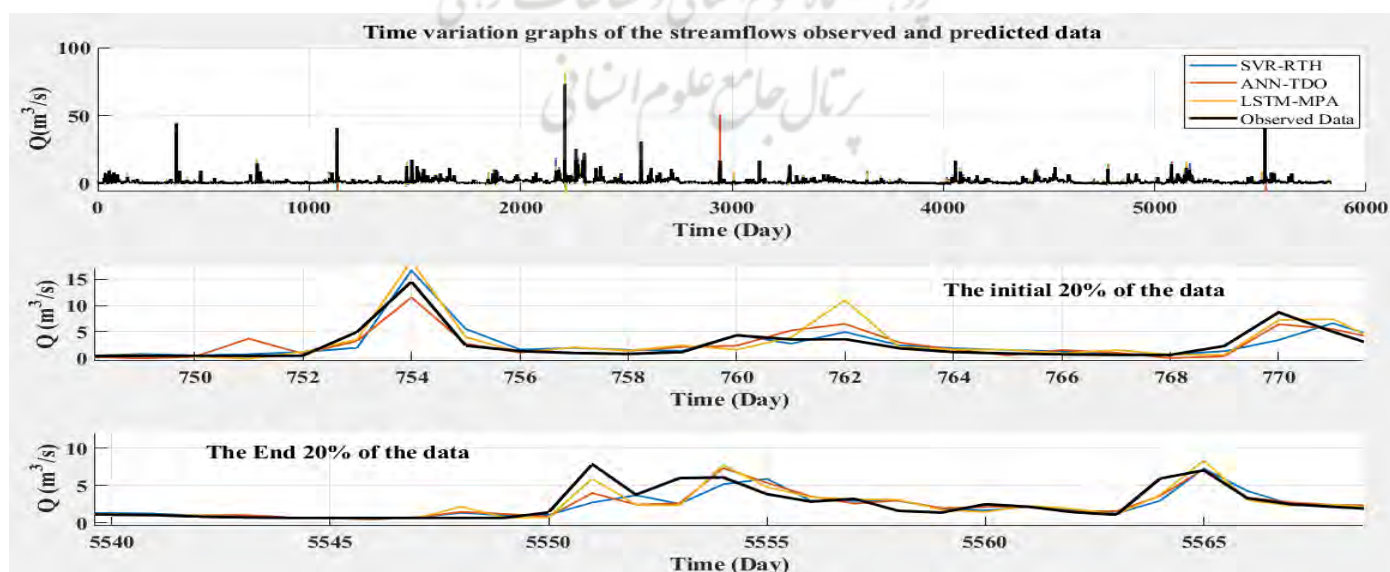


شکل (۴): نمودار پراکندگی مدل هیبریدی SVR-RTH در دو فاز آموزش و آزمون
Figure (4): Scatter plot of the hybrid SVR RTH model in the training and testing phases



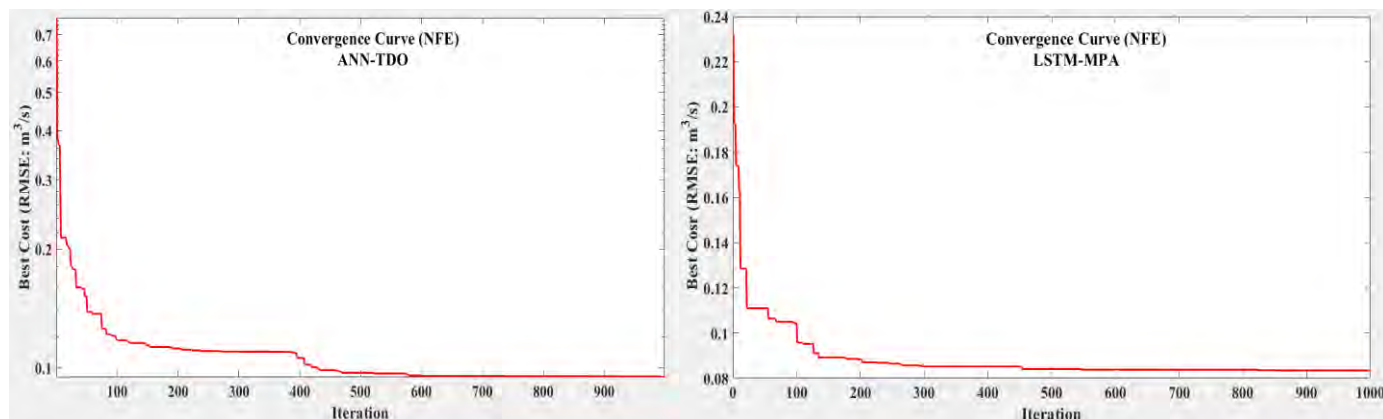
شکل (۵): نمودار پراکندگی مدل هیبریدی LSTM-MPA در دو فاز آموزش و آزمون

Figure (5): Scatter plot of the hybrid LSTM MPA model in the training and testing phases



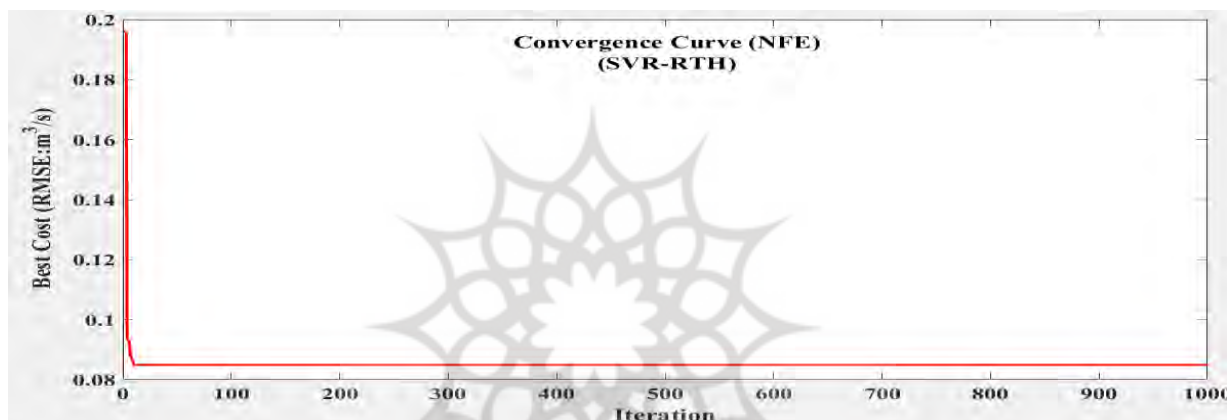
شکل (۶): نمودار سری زمانی داده‌های مشاهداتی و پیش‌بینی شده با مدل‌های هیبریدی

Figure (6): Time series plot of observed and predicted data for the hybrid models.



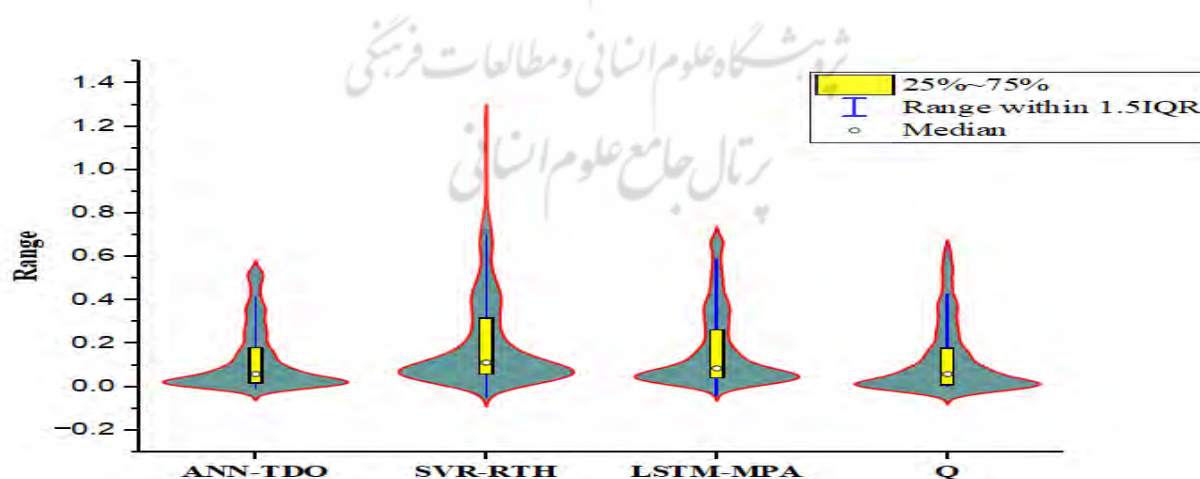
شکل (۷): نمودار منحنی همگرایی (NFE) مدل های هیبریدی ANN-TDO و LSTM-MPA

Figure (7): Convergence curves (NFE) for the hybrid ANN-TDO and LSTM-MPA models



شکل (۸): نمودار منحنی همگرایی (NFE) مدل هیبریدی SVR-RTH

Figure (8): Convergence curve (NFE) for the hybrid SVR-RTH model



شکل (۹): نمودار ویلئون پلات داده های مشاهداتی (واقعی) و داده های پیش بینی شده توسط مدل های هیبریدی

Figure (9): Violin plots of observed versus predicted data for the hybrid models

جمع‌بندی و نتیجه‌گیری

این مطالعه با هدف ارزیابی و مقایسه عملکرد سه مدل هیبریدی در پیش‌بینی سری‌های زمانی، در مقایسه با نسخه‌های منفرد آنها در دو فاز آموزش و آزمون انجام شد. نتایج به‌دست‌آمده از سناریوهای مختلف (M_1 تا M_5) و معیارهای ارزیابی بینش‌های مهمی را در مورد توانایی‌های این مدل‌ها ارائه می‌دهد. بررسی نتایج نشان می‌دهد که مدل‌های هیبریدی به‌طور کلی عملکردی برتر نسبت به نسخه‌های منفرد خود دارند. این امر به‌وضوح در مقادیر بالاتر R^2 و KGE، و مقادیر پایین‌تر MAE و RMSE برای تمامی سناریوها در فازهای آموزش و آزمون قابل مشاهده است. این بهبود عملکرد نشان‌دهنده توانایی مدل‌های هیبریدی در جذب و مدل‌سازی پیچیدگی‌های بیشتر داده‌ها از طریق ترکیب نقاط قوت مدل پایه با مکانیزم‌های بهینه‌سازی یا ویژگی‌افزایی خاص هر مدل هیبریدی است. در میان مدل‌های هیبریدی، ANN-TDO به‌طور مداوم بهترین عملکرد را از خود نشان داد. مقادیر R^2 در این مدل برای فاز آزمون اغلب بالای ۰/۸ و حتی تا ۰/۹۱۵ برای KGE در برخی سناریوها می‌رسید که نشان‌دهنده تطابق بسیار بالا با داده‌های مشاهده‌شده است. مدل LSTM-MPA نیز عملکرد بسیار خوبی را ارائه داد. اگرچه در برخی سناریوها اندکی پایین‌تر از ANN-TDO قرار گرفت، اما مقادیر R^2 و KGE آن (اغلب بالای ۰/۷ و ۰/۸) و MAE و RMSE پایین، نشان‌دهنده قابلیت بالای این مدل در مدل‌سازی سری‌های زمانی است. ماهیت LSTM در حفظ اطلاعات طولانی‌مدت و مکانیزم MPA در بهینه‌سازی، به LSTM-MPA کمک کرده تا دقت بالایی را حفظ کند. در مقایسه، SVR-RTH اگرچه نسبت به SVR منفرد بهبود یافته بود، اما به‌طور کلی عملکردی پایین‌تر از ANN-TDO و LSTM-MPA ارائه داد. بهترین مدل پیشنهادی بر اساس این تحلیل، مدل هیبریدی ANN-TDO است که به‌طور مداوم بهترین عملکرد را در تمامی معیارهای ارزیابی و سناریوهای مختلف به نمایش گذاشت و می‌تواند به‌عنوان ابزاری مناسبی برای پیش‌بینی در این حوزه مورد استفاده قرار گیرد. با توجه به دقت بالا و تعمیم‌پذیری مناسب مدل ANN-TDO، پیشنهاد می‌شود این مدل در سیستم‌های پیش‌بینی کوتاه‌مدت جریان رودخانه و هشدار سیلاب در حوضه‌های فاقد ایستگاه سنجش پیاده‌سازی شود. این مدل می‌تواند به‌عنوان یک ابزار تصمیم‌گیری در مدیریت بحران سیل مورد استفاده قرار گیرد. نتایج نشان می‌دهد که مدل‌های هیبریدی دقت بالایی در شبیه‌سازی جریان دارند. بنابراین، پیشنهاد می‌شود این مدل‌ها در سیستم‌های تصمیم‌گیری یکپارچه مدیریت منابع آب یکپارچه شوند تا به مدیران آب کمک کند تا در تخصیص آب، مدیریت سدها و برنامه‌ریزی کشاورزی تصمیم‌گیری بهینه داشته باشند. برای بهبود بیشتر عملکرد مدل، پیشنهاد می‌شود در مطالعات آتی از داده‌های ماهواره‌ای (مانند بارش TRMM، داده‌های رطوبت خاک SMAP) به عنوان ورودی مدل استفاده شود. همچنین، انجام تحلیل حساسیت بر روی پارامترهای مدل و متغیرهای ورودی می‌تواند به درک بهتر از مکانیزم‌های کنترل‌کننده جریان و افزایش تفسیرپذیری مدل کمک نماید. همچنین پیشنهاد می‌گردد از مدل‌های یادگیری عمیق تقویت شده با مोजک و مدل‌های درختی نیز استفاده شده و اثرات آنها در اقلیم‌های مختلف مورد ارزیابی قرار گیرد.

References

- Ahmed, A. A., Sayed, S., Abdoulhalik, A., Moutari, S., & Oyedele, L. (2024). Applications of machine learning to water resources management: A review of present status and future opportunities. *Journal of Cleaner Production*, 441, 140715. <https://doi.org/10.1016/j.jclepro.2024.140715>.
- Akiner, M. E., Kartal, V., Guzeler, A. C., & Karakoyun, E. (2024). Exploring the applicability of the experiment-based ANN and LSTM models for streamflow estimation. *Earth Science Informatics*, 17(4), 3111-3135. <https://doi.org/10.1007/s12145-024-01332-4>.
- Bickici, B., Beyaztas, B. H., Yaseen, Z. M., Beyaztas, U., & Kahya, E. (2025). Streamflow Intervals Prediction Using Coupled Autoregressive Conditionally Heteroscedastic With Bootstrap Model. *Journal of Flood Risk Management*, 18(1), e70009. <https://doi.org/10.1111/jfr3.70009>.
- Danesh, M., Gharehbaghi, A., Mehdizadeh, S., & Danesh, A. (2025). A comparative assessment of machine learning and deep learning models for the daily river streamflow forecasting. *Water Resources Management*, 39(4), 1911-1930. <https://doi.org/10.1007/s11269-024-04052-y>.
- Danesh, M., Gharehbaghi, A., Mehdizadeh, S., & Danesh, A. (2025). A comparative assessment of machine learning and deep learning models for the daily river streamflow forecasting. *Water Resources Management*, 39(4), 1911-1930. <https://doi.org/10.1007/s11269-024-04052-y>.

- Dehghani, M., Hubálovský, Š., & Trojovský, P. (2022). Tasmanian devil optimization: a new bio-inspired optimization algorithm for solving optimization algorithm. *IEEE access*, 10, 19599-19620. <https://doi.org/10.1109/ACCESS.2022.3151641>.
- Essam, Y., Huang, Y. F., Ng, J. L., Birima, A. H., Ahmed, A. N., & El-Shafie, A. (2022). Predicting streamflow in Peninsular Malaysia using support vector machine and deep learning algorithms. *Scientific Reports*, 12(1), 3883. <https://doi.org/10.1038/s41598-022-07693-4>.
- Faramarzi, A., Heidarinejad, M., Mirjalili, S., & Gandomi, A. H. (2020). Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert systems with applications*, 152, 113377. <https://doi.org/10.1016/j.eswa.2020.113377>.
- Ferahtia, S., Houari, A., Rezk, H., Djerioui, A., Machmoum, M., Motahhir, S., & Ait-Ahmed, M. (2023). Red-tailed hawk algorithm for numerical optimization and real-world problems. *Scientific Reports*, 13(1), 12950. <https://doi.org/10.1038/s41598-023-38778-3>.
- Ghimire, S., Yaseen, Z. M., Farooque, A. A., Deo, R. C., Zhang, J., & Tao, X. (2021). Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks. *Scientific Reports*, 11(1), 17497. <https://doi.org/10.1038/s41598-021-96751-4>.
- Ghumman, A. R., Ahmad, S., & Hashmi, H. N. (2018). Performance assessment of artificial neural networks and support vector regression models for stream flow predictions. *Environmental monitoring and assessment*, 190(12), 704. <https://doi.org/10.1007/s10661-018-7012-9>.
- Grey, V., Fletcher, T., Smith-Miles, K., Hatt, B., & Coleman, R. (2025). Harnessing the strengths of machine learning and geostatistics to improve streamflow prediction in ungauged basins; the best of both worlds. *Journal of Hydrology*, 133936. <https://doi.org/10.1016/j.jhydrol.2025.133936>.
- Ismail, I. I., Jibril, M. M., Muhammad, U. J., Mahmoud, I. A., Aliyu, U. U., Abdullahi, A., & Malami, S. I. (2025). Ensemble Machine Learning Technique Based on Gaussian Algorithm for Stream Flow Modelling. *Techno-computing Journal*, 1(2), 1-17. <https://doi.org/10.71170/tecoj.2025.1.2.pp1-17>.
- Katsektor, J. T., Greve, K., & Yamba, E. I. (2025). Streamflow Forecasting using Machine Learning for Flood Management and Mitigation in the White Volta Basin of Ghana. *Environmental Challenges*, 101181. <https://doi.org/10.1016/j.envc.2025.101181>.
- Li, P. C., Dey, S., & Merwade, V. (2025). Analyzing the effects of data splitting and covariate shift on machine learning based streamflow prediction in ungauged basins. *Journal of Hydrology*, 653, 132731. <https://doi.org/10.1016/j.jhydrol.2025.132731>.
- Nasir, N., Irwan, D., Ahmed, A. N., Ibrahim, S. L., Ibrahim, I., Sherif, M., & El-Shafie, A. (2025). Harnessing machine learning for streamflow prediction: A comparative study of advanced models in the Upper Klang River Basin, Malaysia. *Journal of Hydrology: Regional Studies*, 60, 102565. <https://doi.org/10.1016/j.ejrh.2025.102565>.
- Salaeh, N., Dittthakit, P., Pinthong, S., Wipulanusat, W., Weesakul, U., Elkhachy, I., ... & Elshahabi, M. (2025). Utilizing machine learning to estimate monthly streamflow in ungauged basins of Thailand's southern basin. *Physics and Chemistry of the Earth, Parts A/B/C*, 138, 103840. <https://doi.org/10.1016/j.pce.2024.103840>.
- Sharifi, Z., Mostafazadeh, R., Esmali Ouri, A., Hazbavi, Z., Golshan, M. (2023). Comparing optimization methods of SIMHYD model parameters to simulate daily flow discharge in the Kouzetopraghi Watershed, Ardabil, *Hydrogeomorphology*, 10(34), 33-51. <https://doi.org/10.22034/hyd.2023.49595.1617>.
- Shahedi, k., Forootan Danesh, M. (2022). River Flow Simulation in the Ghorichay Watershed using the Wetspa Model, *Hydrogeomorphology*, 9(32), 25-42. <https://doi.org/10.22034/hyd.2022.49767.1619>.
- Shen, B., Khishe, M., & Mirjalili, S. (2023). Evolving Marine Predators Algorithm by dynamic foraging strategy for real-world engineering optimization problems. *Engineering Applications of Artificial Intelligence*, 123, 106207. <https://doi.org/10.1016/j.engappai.2023.106207>.
- Solanki, H., Vegad, U., Kushwaha, A., & Mishra, V. (2025). Improving streamflow prediction using multiple hydrological models and machine learning methods. *Water Resources Research*, 61(1), <https://doi.org/10.1029/2024WR038192>.

- Tosan, M., Nourani, V., Kisi, O., & Dastourani, M. (2025). Evolution of ensemble machine learning approaches in water resources management: a review. *Earth Science Informatics*, 18(2), 416. <https://doi.org/10.1007/s12145-025-01911-z>.
- Tripathy, K. P., & Mishra, A. K. (2024). Deep learning in hydrology and water resources disciplines: concepts, methods, applications, and research directions. *Journal of Hydrology*, 628, 130458. <https://doi.org/10.1016/j.jhydrol.2023.130458>.
- Vinokić, L., Dotlić, M., Prodanović, V., Kolaković, S., Simonovic, S. P., & Stojković, M. (2025). Effectiveness of three machine learning models for prediction of daily streamflow and uncertainty assessment. *Water Research X*, 27, 100297. <https://doi.org/10.1016/j.wroa.2024.100297>.
- Wang, W., & Lyu, L. (2024). Adaptive Tasmanian devil optimizer for global optimization and application in wireless sensor network deployment. *IEEE Access*, 12, 72382-72407. <https://doi.org/10.1109/ACCESS.2024.3403089>.
- Zhang, Y., Ye, A., Li, J., Nguyen, P., Analui, B., Hsu, K., & Sorooshian, S. (2025). Improve streamflow simulations by combining machine learning pre-processing and post-processing. *Journal of Hydrology*, 655, 132904. <https://doi.org/10.1016/j.jhydrol.2025.132904>.

