Archive of SID.ir

Iranian Water Research Journal

Shahrekord University

Vol. 19/ No. 2/ Serial No. 57/ Summer 2025
https://dx.doi.org/10.22034/iwrj.2025.15105.2665

Evaluation of a hybrid model for monthly runoff estimation using support vector machine based on variational mode decomposition with whale optimization algorithm (Mashhad basin)

M. Fardi rad[1], Kh.Hoseini [2*]

Extended Abstract

Introduction:
The water crisis has become one of the fundamental challenges of the modern era, exacerbated by factors such as population growth, climate change, and uneven rainfall distribution. Iran, with its arid and semi-arid climate, faces significant limitations in water resource management, which impacts agriculture and the economy. Accurate runoff prediction is essential for sustainable water resource management and the design of water infrastructures. However, the complexity of nonlinear relationships and irregular flow distributions makes runoff prediction challenging. Various approaches, including physical, conceptual, and data-driven models, have been proposed. Data-driven models, leveraging machine learning, are particularly notable for their ability to analyze large datasets and uncover hidden patterns. Despite challenges (such as selecting suitable algorithms, determining effective inputs, and evaluating estimation accuracy), these models have been widely applied in hydrology. The aim of this research was to improve monthly runoff prediction in Mashhad basin by utilizing advanced machine learning approaches. In this regard, Variable Mode Decomposition (VMD) was employed as a data preprocessing method, and the Whale Optimization Algorithm (WOA) was used to enhance model accuracy. These methods were applied with the goal of achieving more accurate and efficient predictions in this field.

Methods:
The study area of Mashhad basin, located in Khorasan Razavi Province, is bordered by Hezar Masjid heights to the north and the Binalod heights to the south, with the longitudes of 59˚30' to 59˚42' East and the latitudes of 36˚12' to 36˚24' North. Meteorological and hydrometric data were provided for four stations (Ardak Band, Zoshk, Alang asadi and Kartian), from 1968 to 2022. The stationarity and stability of the time series for rainfall, evaporation, temperature, and discharge were first analyzed. Two distinct modeling approaches were employed to estimate runoff for the future month. In the first approach, three different combinations of parameters were used as model inputs. In the first scenario, all parameters with a three-month lag, determined based on the correlation coefficient of rainfall, temperature, evaporation, and discharge, were utilized. For the second scenario, after conducting sensitivity analysis on the first scenario, seven of the most influential parameters were selected. In the third scenario, the effect of using discharge with a three-month lag as the sole model input was evaluated. The Support Vector Machine (SVM) model and Long Short-Term Memory (LSTM) neural network, both in their basic and optimized forms, were applied to the data from each of the four studied stations. The accuracy of the developed models was assessed using statistical indices, including the Nash–Sutcliffe Efficiency (NSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE). In the second approach, the Variational Mode Decomposition (VMD) technique and the Whale Optimization Algorithm (WOA) were employed. Using this approach, the discharge time series was decomposed into several sub-series, each of which, with a three-month lag, was used as an input to both the basic and optimized SVM and LSTM models. Finally, the performance of the developed models was evaluated during the training and testing phases.

1-MSc Student, Department of Water Engineering, Civil Engineering Faculty, Semnan University, Semnan,Iran
2- Professor, Department of Water Engineering, Civil Engineering Faculty, Semnan University, Semnan, Iran.
*** Corresponding Author: khhoseini@semnan.ac.ir**
Received: 2025/02/06 Accepted: 2025/03/22

Archive of SID.ir

Archive of SID.ir

Iranian Water Research Journal
Shahrekord University

Results

This study aimed to evaluate the performance of VMD-WOA-SVM-WOA hybrid model for runoff estimation at four stations within Mashhad study basin. In this study, two different approaches for modeling runoff were investigated. In the first method, data preprocessing was performed using normalization, the optimized LSTM and SVR models performed better than the base models; however, due to the structural complexity of the hydrological data, the models' ability to identify accurate patterns and estimate runoff was limited. In the second method, data preprocessing was performed using Variable Mode Decomposition. This technique allowed for the extraction of more accurate features by separating hydrological data into components with different frequencies. The primary approach involved using the Variational Mode Decomposition (VMD) method to decompose the runoff time series into modes and optimize their frequencies with the Whale Optimization Algorithm (WOA). These modes were then used as inputs to the Support Vector Machine (SVM) model to estimate runoff for the future month. To assess the proposed model, its performance was compared with a simple SVM model using different input configurations. The results demonstrated that the VMD-WOA-SVM-WOA model achieved mean RMSE and Nash–Sutcliffe 0.185 and 0.979, respectively, with the minimum input variables. These findings align with the principles of regression modeling, emphasizing simplicity and avoiding unnecessary complexity. Furthermore, the results confirmed that increasing the number of input variables does not necessarily enhance estimation accuracy and can even introduce noise and reduce the model's generalizability. This research showed that by employing data preprocessing (VMD) and optimizing model parameters with WOA, a simple, accurate, and interpretable model can be developed. Additionally, the proposed model is not only effective for runoff estimation but also has the potential to be applied to similar studies of hydrological processes such as evapotranspiration and sediment modeling in different regions.

Conclusion:

This study investigated two approaches for runoff modeling. The first approach used data normalization and application of optimized LSTM and SVR models, which showed some improvement compared to base models. However, due to the complexity of hydrological data, the models demonstrated limited accuracy, and the overall performance based on low Nash–Sutcliffe values was not satisfactory. The second approach employed for processing, was Variational Mode Decomposition (VMD), which allowed better extraction of temporal features. With parameter optimization using the Whale Optimization Algorithm (WOA), the VMD-WOA-SVR-WOA model achieved significantly better accuracy and generalization. While VMD-WOA-LSTM also performed well, the superior results of VMD-WOA-SVR-WOA model highlighted it as an effective tool for runoff prediction and water resource management.

Keywords: Rainfall-Runoff Modeling, Support Vector Regression, Signal Processing, Optimization.

Citation: Fardi rad, M. and Hoseini Kh., 2025. Evaluation of a hybrid model for monthly runoff estimation using support vector machine based on variational mode decomposition with whale optimization algorithm (Mashhad basin). Iranian Water Research Journal 57(2). pp. 35-48. https://dx.doi.org/10.22034/iwrj.2025.15105.2665.

Archive of SID.ir

Archive of SID.ir

مجله پژوهش آب ایران

جلد ۱۹/ شماره ۲/ پیاپی ۵۷/ تابستان ۱۴۰۴
https://dx.doi.org/10.22034/iwrj.2025.15105.2665

ارزیابی مدل ترکیبی برآورد رواناب ماهانه با استفاده از ماشین بردار پشتیبان بر اساس تجزیه مد متغیر به همراه الگوریتم بهینه‌سازی نهنگ (حوضه مطالعاتی مشهد)

محمدرضا فردی‌راد[۱] و خسرو حسینی[۲]*

چکیده

بحران کمبود آب یکی از چالش‌های مهم قرن حاضر است که نیازمند رویکردهای نوین برای مدیریت پایدار منابع آب می‌باشد. پیش‌بینی دقیق رواناب، به عنوان عنصری کلیدی در این مدیریت، نقشی حیاتی در برنامه‌ریزی و تصمیم‌گیری‌های مرتبط ایفا می‌کند. در این پژوهش با هدف بهبود دقت پیش‌بینی رواناب ماهانه در منطقه مشهد، از روش پیش‌پردازش تجزیه مد متغیر VMD برای بهبود کیفیت داده‌ها و از مدل رگرسیون ماشین بردار پشتیبان SVR و شبکه عصبی حافظه طولانی کوتاه مدت LSTM برای مدل‌سازی رابطه بین ورودی‌ها و خروجی استفاده شده‌است. همچنین، الگوریتم بهینه‌ساز نهنگ WOA برای یافتن پارامترهای بهینه مدل‌ها به‌کار رفته‌است. در این مطالعه، مدل ترکیبی VMD-WOA-SVR-WOA معرفی و کارایی آن در برآورد رواناب با مدل‌های ساده و بهینه‌شده SVR و LSTM با سه سناریو داده ورودی مختلف (نوع و تاخیر زمانی متفاوت) مقایسه شد. داده‌های مورد استفاده شامل داده‌های بارش، دما، تبخیر و دبی با دوره آماری ۵۴ ساله (۱۴۰۱-۱۳۴۷) است. نتایج پژوهش بر اساس معیارهای ضریب نش-ساتکلیف (NSE)، میانگین مطلق خطا (MAE) و ریشه میانگین مربعات خطا (RMSE) نشان می‌دهد که مدل VMD-WOA-SVR-WOA عملکرد بسیار بهتری در برآورد رواناب نسبت به مدل‌های VMD-WOA-LSTM و VMD-WOA-SVR داشته است. برای نمونه، در ایستگاه زشک، این مدل در مرحله آموزش مقادیر NSE، RMSE و MAE را به ترتیب ۰/۹۸۳، ۰/۰۹۹ و ۰/۰۷۲ ارائه داده‌است. همچنین در مرحله آزمون نیز به ترتیب مقادیر ۰/۹۷۵، ۰/۰۸۳ و ۰/۱۲۴ به‌دست آمد که نشان‌دهنده دقت مناسب مدل در برآورد رواناب است. علاوه بر این، در ایستگاه‌های ارداک بند، النگ اسدی و کرتیان مدل ترکیبی علاوه بر دستیابی به دقت بالا (NSE بین ۰/۹۷ تا ۰/۹۸) از نظر کاهش خطاهای RMSE و MAE نسبت به مدل‌های مبتنی بر SVR و LSTM عملکرد به مراتب بهتری از خود نشان داده است.

کلمات کلیدی: الگوریتم بهینه‌ساز نهنگ (WOA)، الگوریتم رگرسیون بردار پشتیبان (SVR)، الگوریتم شبکه عصبی حافظه طولانی کوتاه مدت (LSTM)، تجزیه مد متغیر (VMD)، مدل‌سازی بارش-رواناب.

ارجاع: فردی راد، م.ر. و حسینی، خ. ۱۴۰۴. ارزیابی مدل ترکیبی برآورد رواناب ماهانه با استفاده از ماشین بردار پشتیبان بر اساس تجزیه مد متغیر به همراه الگوریتم بهینه‌سازی نهنگ (حوضه مطالعاتی مشهد). مجله پژوهش آب ایران.۵۷(۲)، صص: ۳۵-۴۸.
https://dx.doi.org/10.22034/iwrj.2025.15105.2665.

[۱] - دانشجوی کارشناسی ارشد، گروه مهندسی آب، دانشکده مهندسی عمران، دانشگاه سمنان
[۲] - استاد گروه مهندسی آب، دانشکده مهندسی عمران، دانشگاه سمنان
* نویسنده مسئول: khhoseini@semnan.ac.ir

تاریخ دریافت: ۱۴۰۳/۱۱/۱۸ تاریخ پذیرش: ۱۴۰۴/۰۱/۰۲

Archive of SID.ir

Archive of SID.ir

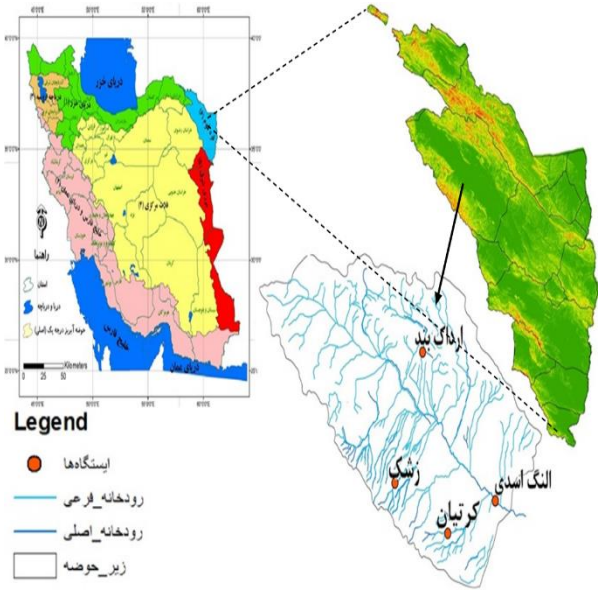
۱-مقدمه

امروزه بحران آب در جوامع بشری به چالشی اساسی تبدیل شده است (Tang et al., 2022) و عواملی مانند افزایش جمعیت، تغییرات اقلیمی و کاهش منابع آب شیرین این بحران را تشدید کرده‌اند (Shafeizadeh et al., 2019). ایران با اقلیم خشک و نیمه‌خشک با چالش‌های جدی در مدیریت منابع آب مواجه است. ناهماهنگی در توزیع زمانی و مکانی بارش با نیازهای کشاورزی و صنعتی، اقتصاد مناطق مختلف کشور را تحت تأثیر قرار می‌دهد. از این‌رو، پیش‌بینی رواناب نقش کلیدی در طراحی زیرساخت‌های پایدار و مدیریت بهینه منابع آب ایفا می‌کند (Merufinia. et al., 2023). با این حال، تخمین دقیق جریان به دلیل پیچیدگی‌های ناشی از روابط غیرخطی بین اجزای حوضه، چالش‌برانگیز است (Xu et al., 2024). علم هیدرولوژی برای مدل‌سازی و پیش‌بینی این فرآیندها سه رویکرد اصلی ارائه می‌دهد: روش‌های فیزیکی، مفهومی و داده محور (Sayed et al., 2023). روش‌های داده محور با استفاده از تکنیک‌های یادگیری ماشین قادرند الگوهای پنهان در داده‌ها را شناسایی کنند (Kavusi and Khozeymeh Nejad, 2020). این ویژگی به‌ویژه در شرایط کمبود داده مؤثر است. با وجود چالش‌هایی مانند انتخاب الگوریتم و تعیین ویژگی‌های مؤثر، یادگیری ماشین به عنوان روشی نوآورانه در مدل‌سازی هیدرولوژیکی مورد توجه قرار گرفته و مطالعات متعددی نیز در این زمینه انجام شده است (Kumar et al., 2023; Yao et al., 2023). برای نمونه در پژوهش Bagherpour et al. (2019)، عملکرد مدل‌های HBV، TOPMODEL، SVM و ANFIS در شبیه‌سازی بارش-رواناب ارزیابی شد. نتایج حاکی از دقت بالاتر مدل ANFIS نسبت به سایر مدل‌ها بود. در پژوهش Guo et al. (2023) به‌منظور پیش‌بینی دقیق رواناب ماهانه در سد میون، مدل ترکیبی VMD-LSTM-Transformer توسعه داده شد. ارزیابی نتایج حاکی از برتری این مدل نسبت به سایر مدل‌هاست. مقدار NSE= ۰/۹۷۶ نیز بیانگر دقت بالای آن در شبیه‌سازی رواناب در منطقه مورد مطالعه است. هدف از تحقیق حاضر، بهبود برآورد رواناب ماهانه در محدوده مطالعاتی مشهد با بهره‌گیری از رویکردهای پیشرفته یادگیری ماشین است. در این راستا، از تجزیه مد متغیر (VMD) به‌عنوان یک روش پیش‌پردازش داده و از الگوریتم بهینه‌سازی نهنگ (WOA) برای افزایش دقت مدل‌ها استفاده شده است. این روش‌ها با هدف دستیابی به برآوردهای دقیق‌تر و کارآمدتر به کار گرفته شده‌اند.

۲-مواد و روش‌ها

۲-۱-منطقه مطالعاتی و داده‌ها

محدوده مطالعاتی این تحقیق مشهد در استان خراسان رضوی واقع شده‌است، از شمال به ارتفاعات هزار مسجد و از جنوب به ارتفاعات بینالود محدود می‌گردد (شکل۱). محدوده مطالعاتی مشهد بین طول جغرافیایی ′۱۱ °۵۹ تا ′۴۸ °۵۹ شرقی و عرض جغرافیایی ′۱۰ °۳۶ تا ′۴۳ °۳۶ شمالی قرار گرفته‌است (Yousefi et al., 2007).



شکل ۱- موقعیت حوضه مشهد و ایستگاه‌های هیدرومتری

داده‌های مورد نیاز این پژوهش شامل داده‌های دبی جریان ایستگاه‌های هیدرومتری (جدول۱) و همچنین داده‌های هواشناسی شامل آمار ماهانه بارش، دما و تبخیر طی سال‌های ۱۳۴۷ تا ۱۴۰۱ است که از سازمان آب منطقه‌ای خراسان رضوی گردآوری شد. همچنین، داده‌های مفقود شده با استفاده از روش رگرسیون خطی ساده و با تکیه بر داده‌های ایستگاه‌های مجاور با بیشترین همبستگی بازسازی گردید. پیش از انجام هرگونه تحلیل، ایستایی و پایداری سری‌های زمانی بررسی شد. ایستایی، که شرط لازم برای مدل‌سازی‌های هیدرولوژیکی است، با آزمون دیکی-فولر افزوده (ADF) مطابق با روش پیشنهادی (Said and Dickey, 1984) ارزیابی شد. همچنین، پایداری سری‌های زمانی با محاسبه نمایه هرست (Hurest Exponent) مورد ارزیابی قرار گرفت. نمایه هرست، که نشان‌دهنده میزان حافظه بلندمدت در داده‌ها است، در بازه صفر تا ۱ متغیر بوده و مقادیر کمتر از ۰/۵، نزدیک به ۰/۵ و بیشتر از ۰/۵ به ترتیب بیانگر رفتار ضدهمبسته، عدم وابستگی بلندمدت و وابستگی بلندمدت مثبت در داده‌ها است (Ghahraman, 2013).

Archive of SID.ir

Archive of SID.ir

جدول ۱- مشخصات ایستگاه‌های هیدرومتری مورد مطالعه

ایستگاه	طول جغرافیایی	عرض جغرافیایی	دبی (m^3/s)		طول دوره آماری
			میانگین	واریانس	
النگ اسدی	۵۹° ۴۸′ ۳۵/۹″	۳۶° ۱۵′ ۳/۱۸″	۰/۹۴	۵/۹۱	۵۴ سال
ارداک بند	۵۹° ۲۳′ ۱/۳۳″	۳۶° ۴۳′ ۹/۴۵″	۱/۰۹	۱/۴۲	۵۴ سال
زشک	۵۹° ۱۱′ ۱/۴۳″	۳۶° ۲۰′ ۱۹″	۰/۵۴	۰/۶۰	۵۴ سال
کرتیان	۵۹° ۳۰′ ۴۹″	۳۶° ۱۰′ ۱۸″	۰/۴۹	۱/۱۴	۵۴ سال

۲-۲-پیش‌پردازش داده‌ها

۲-۲-۱-نرمال‌سازی

نرمال‌سازی به‌منظور جلوگیری از اثر غالب شدن ویژگی‌های با مقیاس بزرگ انجام می‌شود. یکی از روش‌های رایج نرمال‌سازی مقیاس‌گذاری حداقل-حداکثر است (Kavusi and Khozeymeh Nejad, 2020). از رابطه (۱) برای نرمال‌سازی داده‌ها استفاده شد.

$$X_n=\left[\frac{x-x_{min}}{x_{max}-x_{min}}\right] \quad (۱)$$

در رابطه (۱)؛ X_n، x، x_{min} و x_{max} به ترتیب نشان‌دهنده مقادیر نرمال‌شده، واقعی، بیشینه و کمینه سری زمانی موردبررسی می‌باشند.

۲-۲-۲-تجزیه مدی متغیر

تجزیه حالت متغیر (VMD) یک الگوریتم پیشرفته برای تجزیه سیگنال است که در سال ۲۰۱۴ معرفی شد (Feng et al., 2020). این روش برخلاف روش‌های کلاسیک مانند تجزیه مد تجربی، بر پایه یک چارچوب بهینه‌سازی مقید عمل می‌کند و با استفاده از تکرارهای متوالی، اجزای سیگنال را به‌گونه‌ای استخراج می‌کند که هر مؤلفه دارای پهنای باند محدود در حوزه فرکانسی باشد. این ویژگی باعث افزایش پایداری، کاهش حساسیت به نویز و جلوگیری از مشکلاتی مانند اختلاط مد می‌شود. در این روش، تجزیه سیگنال به‌عنوان یک مسئله بهینه‌سازی تعریف می‌شود که هدف آن تجزیه سیگنال به k زیرمد است؛ به‌گونه‌ای که پهنای باند هر مد به حداقل برسد. برای حل این مسئله، از ضرایب لاگرانژ به‌منظور اضافه کردن محدودیت‌ها به تابع هدف استفاده می‌شود (رابطه ۲). سپس، روش بهینه‌سازی ADMM[1] (رابطه ۳) به‌عنوان یک روش عددی برای به‌روزرسانی پارامترها و حل مسئله بهینه‌سازی استفاده می‌شود.

$$L(u_k,w_k,\lambda)=\alpha\sum_{k=1}^{k}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}*u_k(t)\right)\right]e^{-jw_kt}\right\|_2^2+\left\|f(t)-\sum_{k=1}^{k}u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{k}u_k(t)\right\rangle \quad (۲)$$

$$\begin{cases}\hat{\lambda}^{n-1}(w)=\hat{\lambda}^{n}(w)+\tau\left(\hat{f}(w)-\sum_{k=1}^{k}\hat{u}_k^{n+1}(w)\right)\\ \hat{u}_k^{n+1}(w)=\frac{\hat{f}(w)-\sum_{i\neq k}\hat{u}_i(w)+\frac{\lambda(w)}{2}}{1+2\alpha(w-w_k)^2}\\ \hat{w}_k^{n+1}=\frac{\int_0^\infty w\,|\hat{u}_k(w)|^2\,dw}{\int_0^\infty|\hat{u}_k(w)|^2\,dw}\end{cases} \quad (۳)$$

که در این رابطه λ ، u_k و w_k به ترتیب مقادیر ضریب لاگرانژ، مدهای ذاتی و فرکانس مرکزی مدها می‌باشند. همچنین n شماره تکرار فعلی و τ پارامتر تحمل نویز هستند (Dragomiretskiy and Zosso, 2014).

۲-۳-ماشین بردار پشتیبان

ماشین بردار پشتیبان[2] که اولین‌بار در سال ۱۹۶۵ معرفی شد، یکی از الگوریتم‌های محبوب یادگیری ماشین است که برای مسائل طبقه‌بندی و رگرسیون مورداستفاده قرار می‌گیرد. این روش بر اساس نظریه آماری و حداقل‌سازی احتمال زیان طراحی شده و هدف آن یافتن یک جواب بهینه کلی است (Cortes and Vapnik, 1995). در مسائل رگرسیون، نسخه رگرسیونی این مدل SVR به کار گرفته می‌شود که هدف آن یافتن تابعی است که خروجی‌های y_i را با ورودی‌های x_i با حداقل خطای ممکن ε مدل‌سازی کند. فرض می‌شود داده‌های آموزشی به‌صورت D={(),(),…,()} باشند، در این حالت تابع خطی‌سازی مربوطه به‌صورت رابطه (۴) بیان می‌شود.

$$f(x)=(w\,.\,x)+b$$

[1] Alternating Direction Method of Multipliers
[2] Support Vector Machine

Archive of SID.ir

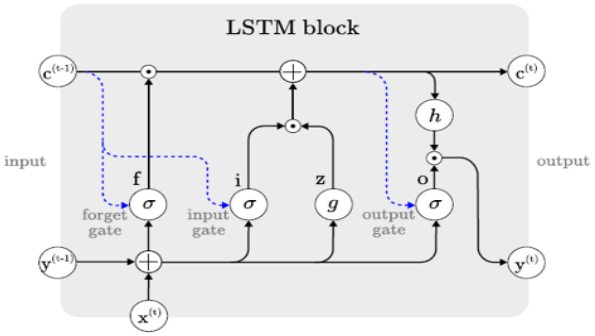
Archive of SID.ir

$$with\ (w, x) \in R^N, b \in R \quad (۴)$$

در رابطه فوق، x فضای ورودی، w معرف بردار وزن و b نیز معرف میزان اختلال (بایاس) می‌باشد. در ماشین بردار پشتیبان هدف پیداکردن بهترین پارامترهای w و b است؛ به‌گونه‌ای که به طور هم‌زمان منجر به کمترین میزان انحراف (ε) شوند.

۲-۴-شبکه عصبی حافظه طولانی کوتاه مدت

شبکه‌های عصبی حافظه طولانی کوتاه‌مدت (LSTM) نوعی شبکه عصبی بازگشتی پیشرفته هستند که با هدف رفع مشکل وابستگی‌های بلندمدت در داده‌های سری‌زمانی توسعه یافته‌اند. این شبکه‌ها با بهره‌گیری از ساختار سلول حافظه و سه دروازه ورودی، فراموشی و خروجی، جریان اطلاعات را به‌صورت هدفمند کنترل کرده و از مشکل کاهش گرادیان که در شبکه‌های بازگشتی سنتی وجود دارد، جلوگیری می‌کنند. این ویژگی‌ها موجب حفظ اطلاعات مهم و حذف داده‌های غیرضروری در طول زمان می‌شود (Hochreiter and Schmidhuber, 1997; Zhang et al., 2021). شکل ۲ ساختار کلی یک شبکه عصبی LSTM را نشان می‌دهد.



شکل ۲ - نمای کلی معماری بلوک LSTM معمولی (Van Houdt et al., 2020)

فرمول‌های ریاضی حاکم بر عملکرد مدل LSTM به صورت روابط (۵) تا (۱۰) است.

$$i^{(t)} = \sigma\left(W_i\, x^{(t)} + R_i\, y^{(t-1)} + p_i \odot c^{(t-1)} + b_i\right) \quad (۵)$$

$$f^{(t)} = \sigma\left(W_f\, x^{(t)} + R_f\, y^{(t-1)} + p_f \odot c^{(t-1)} + b_f\right) \quad (۶)$$

$$o^{(t)} = \sigma\left(W_o x^{(t)} + R_o y^{(t-1)} + p_o \odot c^{(t)} + b_o\right) \quad (۷)$$

$$z^{(t)} = g\left(W_z x^{(t)} + R_z y^{(t-1)} + b_z\right) \quad (۸)$$

$$c^{(t)} = z^{(t)} \odot i^{(t)} + c^{(t-1)} \odot f^{(t)} \quad (۹)$$

$$y^{(t)} = g\left(c^{(t)}\right) \odot o^{(t)} \quad (۱۰)$$

در شبکه LSTM، $x^{(t)}$ ورودی جاری، $y^{(t)}$ خروجی جاری و $c^{(t)}$ وضعیت حافظه جاری است. دروازه‌های ورودی $i^{(t)}$، فراموشی $f^{(t)}$ و خروجی $o^{(t)}$ جریان اطلاعات را کنترل می‌کنند. $z^{(t)}$ مقدار کاندید برای به‌روزرسانی حافظه است. وزن‌های مرتبط با ورودی W، خروجی قبلی R، و حافظه P همراه با بایاس‌ها b، مقادیر را تنظیم می‌کنند. تابع سیگموید σ و فعال‌سازی g عملیات غیرخطی را انجام می‌دهند و ضرب عنصر به عنصر ⊙ تعامل مقادیر را مشخص می‌کند. این عناصر باهم اطلاعات را ذخیره، فراموش یا منتقل می‌کنند.

۲-۵-الگوریتم بهینه‌سازی نهنگ

میرجلیلی و لویس الگوریتم بهینه‌سازی نهنگ (WOA) را بر اساس رفتار نهنگ‌های گوژپشت معرفی کردند (Mirjalili and Lewis, 2016). این الگوریتم شامل سه مرحله است: محاصره طعمه، حمله با حباب و جستجوی طعمه. در محاصره، نهنگ‌ها موقعیت طعمه را شناسایی و آن را محاصره می‌کنند. الگوریتم فرض می‌کند بهترین راه‌حل موجود، طعمه هدف است و سایر عوامل جستجو موقعیت خود را بر اساس آن به‌روزرسانی می‌کنند. این رفتار توسط روابط (۱۱) و (۱۴) نمایش داده می‌شود (Mirjalili and Lewis, 2016; Dunyaei et al., 2020).

$$D = |C.X^*(t) - X(t)| \quad (۱۱)$$

$$X(t+1) = X^*(t) - A.D \quad (۱۲)$$

$$A = 2a\,.r - a \quad (۱۳)$$

$$C = 2\,.\,r \quad (۱۴)$$

در این روابط، X^* موقعیت بهترین راه‌حل، X بردار مکان، A و C بردار ضرایب، t شماره تکرار و r مقدار تصادفی در بازه صفر تا ۱ و a ضریبی خطی است که از ۲ به صفر کاهش می‌یابد و در اکتشاف و استخراج نقش دارد. در مرحله حمله به طعمه با ایجاد حباب (بخش استخراج)، نهنگ‌ها با احتمال برابر، یکی از دو روش کاهش تدریجی محاصره یا حرکت مارپیچی را برای به‌روزرسانی موقعیت خود به‌صورت تصادفی انتخاب می‌کنند. فرمول ریاضی این مرحله به‌صورت رابطه (۱۵) است.

$$X(t+1) = \begin{cases} X^*(t) - A.D & if\ p < 0.5 \\ D'.e^{bl}.\cos(2\pi l) + X^*(t) & if\ p \geq 0.5 \end{cases} \quad (۱۵)$$

که دراین‌رابطه D' موقعیت نهنگ نسبت به طعمه و p متغیر تصادفی بین صفر و ۱ است. در مرحله اکتشاف، نهنگ‌ها به‌صورت تصادفی در فضای جستجو حرکت کرده تا از کمینه‌های محلی اجتناب و نقاط جدید را بررسی کنند. اکتشاف با استفاده از معادلات (۱۶) و (۱۷) انجام می‌شود.

$$\boldsymbol{D = |C\,.\,X_{rand} - X|} \quad (۱۶)$$

$$X(t-1) = X_{rand} - A\,.\,D \quad (۱۷)$$

Archive of SID.ir

Archive of SID.ir

که در این روابط X_{rand} موقعیت تصادفی یک نهنگ و D فاصله بین نهنگ و موقعیت تصادفی X_{rand} است (Mirjalili and Lewis, 2016).

۲-۶- مدل VMD-WOA-SVR-WOA

به‌منظور افزایش دقت برآورد جریان ماهانه، در این مطالعه یک مدل ترکیبی چندمرحله‌ای با استراتژی‌های متنوع پیشنهاد شده است. در این مدل، ابتدا داده‌ها با استفاده از روش تجزیه مد متغیر به مجموعه‌ای از مؤلفه‌های فرکانسی تجزیه می‌شود. برای دستیابی به تجزیه‌ای دقیق‌تر و کاهش نوسانات و پیچیدگی ذاتی داده‌های اولیه، پارامترهای کلیدی VMD (نظیر تعداد مدها و پارامتر کنترل پهنای باند) با بهره‌گیری از الگوریتم بهینه‌سازی نهنگ تنظیم و بهینه شده‌اند. در گام بعد، به‌منظور مدل‌سازی هر یک از مؤلفه‌های حاصل از تجزیه، از مدل رگرسیون بردار پشتیبان استفاده شده است. پارامترهای SVR نیز با استفاده از الگوریتم WOA بهینه‌سازی گردیده‌اند و در نتیجه، مدل ترکیبی WOA-SVR برای برآورد هر مؤلفه به‌کار گرفته شده است. سپس، نتایج برآورد شده مربوط به هر مؤلفه با یکدیگر تلفیق شده و رواناب ماهانه برآورد می‌شود (شکل ۳). مراحل اجرای مدل به‌صورت زیر خلاصه می‌شود:

۱) تجزیه داده‌ها با روش VMD که پارامترهای آن با استفاده از الگوریتم WOA بهینه‌سازی شده‌اند.
۲) با بررسی همبستگی زمانی داده‌ها، تأخیر زمانی مناسب شناسایی و در مدل‌سازی لحاظ می‌گردد.
۳) داده‌ها به دو مجموعه آموزش و آزمون تقسیم می‌شوند.
۴) مدل WOA-SVR برای برآورد هر مد به‌کار گرفته می‌شود.
۵) برآورد نهایی رواناب از جمع نتایج هر مد به‌دست می‌آید.

۲-۷-معیار ارزیابی و سنجش مدل

در این پژوهش، برای انجام مدل‌سازی از زبان برنامه نویسی پایتون استفاده شد. به منظور مدل‌سازی، ۷۰ درصد کل داده‌ها برای آموزش و ۳۰ درصد برای آزمون مدل‌ها استفاده شد. عملکرد مدل‌ها با استفاده از پنج شاخص ضریب تعیین[1] (R^2)، نش-ساتکلیف[2] (NSE) میانگین مطلق خطا[3] (MAE)، ریشه میانگین مربعات خطا[4] (RMSE) و ریشه میانگین مربعات خطا نرمال شده[5] (NRMSE) ارزیابی شد. روابط ریاضی این شاخص‌ها در رابطه‌های ۱۸تا ۲۲ بیان شده‌است.

$$R^2 = \left(\frac{\Sigma(y_o-\bar{y}_o)(y_p-\bar{y}_p)}{\sqrt{\Sigma(y_o-\bar{y}_o)^2}.\sqrt{\Sigma(y_p-\bar{y}_p)^2}} \right)^2 \quad (۱۸)$$

$$NSE = 1 - \frac{\Sigma(y_o-y_p)^2}{\Sigma(y_o-\bar{y}_o)^2} \quad (۱۹)$$

$$MAE = \frac{1}{n}\Sigma_{i=1}^{n}\left|y_{o_i} - y_{p_i}\right| \quad (۲۰)$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}\left(y_{o_i}-y_{p_i}\right)^2}{n}} \quad (۲۱)$$

$$NRMSE = \frac{\sqrt{\frac{\Sigma_{i=1}^{n}\left(y_{o_i}-y_{p_i}\right)^2}{n}}}{\bar{y}_o} \times 100 \quad (۲۲)$$

در این روابط y_o، y_p، $\bar{y}_o$، $\bar{y}_p$ به ترتیب مقادیر داده‌های واقعی (مشاهده شده)، مقادیر پیش‌بینی شده توسط مدل، میانگین مقادیر واقعی، میانگین مقادیر پیش‌بینی شده و n تعداد داده‌های مشاهداتی است.

۳-نتایج و بحث

در فرآیند مدل‌سازی رواناب، شناسایی ویژگی‌های مؤثر مانند بارش، دما و نفوذپذیری به‌عنوان ورودی مدل اهمیت بالایی دارد (Davanlou Tajbakhsh et al., 2019). در این پژوهش، برای تعیین داده‌های ورودی از سه سناریو استفاده شد. در سناریوی نخست، داده‌های بارش (R)، دما (T)، تبخیر (E) و رواناب (Q) با تأخیر زمانی سه‌ماهه (تعیین‌شده بر مبنای ضریب همبستگی) به‌عنوان ورودی مدل‌های SVM و LSTM انتخاب شدند. با توجه به تأثیر تعداد متغیرها بر پیچیدگی و دقت مدل (Araghinejad and Karamouz, 2005; Taormina and Chau, 2015)، در سناریوی دوم، با انجام تحلیل حساسیت روی سناریوی اول، ویژگی‌های کم‌اهمیت حذف و تعداد ویژگی‌های ورودی از ۱۲ به ۷ مورد کاهش یافت. در سناریوی سوم، تنها دبی با تأخیر زمانی سه‌ماهه به‌عنوان ورودی در نظر گرفته شد. در این پژوهش، برای ساخت مدل پایه‌ی رگرسیون بردار پشتیبان، از کرنل پایه شعاعی و مقادیر اولیه‌ی متعارف برای پارامترهای مدل استفاده شده است. بهینه‌سازی هایپرپارامترهای تابع کرنل، نقش مؤثری در بهبود عملکرد مدل دارد (Tikhamarine et al., 2020). در این پژوهش، به‌منظور تنظیم بهینه‌ی هایپرپارامترها و ارتقای دقت مدل، از الگوریتم بهینه‌سازی نهنگ استفاده شد.

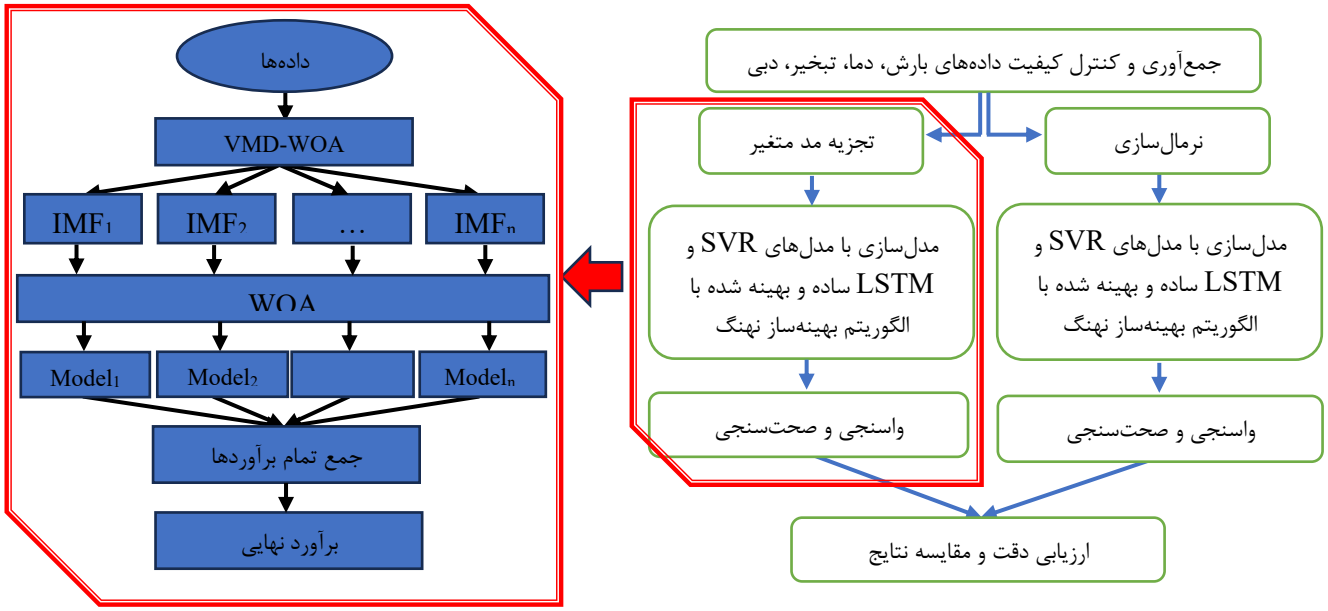
[1] R^2 Score - Coefficient of Determination
[2] Nash–Sutcliffe
[3] Mean Absolute Error
[4] Root Mean Squared Error
[5] Normalized Root Mean Squared Error

Archive of SID.ir

Archive of SID.ir

تابع هدف در فرآیند بهینه‌سازی، بیشینه‌سازی تطابق بین مقادیر پیش‌بینی‌شده و واقعی جریان ماهانه بوده و شاخص نش-ساتکلیف به‌عنوان معیار ارزیابی عملکرد به کار گرفته شد.

جزئیات مربوط به ترکیب داده‌های ورودی مقادیر اولیه و محدوده تغییرات پارامترها در جداول ۲، ۳ و ۴ ارائه شده‌است



شکل ۳ - فلوچارت انجام تحقیق

جدول ۲ - ترکیب داده های ورودی به مدل

ایستگاه	ترکیب داده‌های ورودی به مدل LSTM	ترکیب داده‌های ورودی به مدل SVM	الگو
همه ایستگاه‌ها		$Q_{t+1}=f(R_t, R_{t-1}, R_{t-2}, E_t, E_{t-1}, E_{t-2}, T_t, T_{t-1}, T_{t-2}, Q_t, Q_{t-1}, Q_{t-2})$	۱
ارداک بند	$Q_{t+1}=f(R_t, R_{t-1}, E_t, E_{t-2}, T_{t-1}, Q_t, Q_{t-2})$	$Q_{t+1}=f(R_t, R_{t-1}, R_{t-2}, T_{t-1}, T_{t-2}, Q_t, E_{t-1})$	۲
النگ اسدی	$Q_{t+1}=f(E_t, E_{t-1}, E_{t-2}, T_{t-1}, T_{t-2}, Q_t, Q_{t-1})$	$Q_{t+1}=f(R_t, R_{t-2}, E_t, E_{t-2}, T_t, T_{t-1}, Q_t)$	۲
زشک	$Q_{t+1}=f((R_t, R_{t-2}, T_t, T_{t-1}, T_{t-2}, Q_t)$	$Q_{t+1}=f(R_{t-1}, T_t, T_{t-1}, T_{t-2}, Q_t, Q_{t-1}, Q_{t-2})$	۲
کرتیان	$Q_{t+1}=f(R_{t-1}, R_{t-2}, E_t, E_{t-2}, T_{t-2}, Q_t, Q_{t-1})$	$Q_{t+1}=f(E_t, E_{t-1}, T_t, T_{t-2}, Q_t, Q_{t-1}, Q_{t-2})$	۲
همه ایستگاه‌ها		$Q_{t+1}=f(Q_t, Q_{t-1}, Q_{t-2})$	۳

جدول ۳ - مقادیر پارامترهای مدل رگرسیون بردار پشتیبان

پارامتر	کرنل	ε	γ	C
مقدار	rbf	۰/۱	۲/۵	۱/۰

۳-۱-مدل ساده و بهینه شده رگرسیون بردار پشتیبان

نتایج حاصل از ارزیابی مدل‌های SVR و LSTM در حالت‌های ساده و بهینه‌شده با استفاده از الگوریتم WOA، با سه ترکیب ورودی متفاوت در چهار ایستگاه، که در جدول ۵ ارائه شده است، حاکی از وجود تفاوت‌های معنی‌دار در عملکرد مدل‌ها می‌باشد. بررسی شاخص‌های آماری در مراحل آموزش و آزمون، بیانگر برتری مدل‌های بهینه‌شده نسبت به نسخه‌های ساده است. این بهبود زمانی حائز اهمیت است که منجر به افزایش دقت مدل در مرحله آزمون و ارتقای تعمیم‌پذیری آن شود. به‌عنوان نمونه، در ایستگاه ارداک‌بند مدل SVR-WOA با ترکیب ورودی ۲ به ترتیب به ضرایب NSE برابر با ۰/۵۲ در مرحله آموزش و ۰/۵۸ در مرحله آزمون دست یافت، در حالی‌که این مقادیر برای مدل ساده SVR با همان ترکیب ورودی به‌ترتیب ۰/۳۹ و ۰/۴۴ بود. این بهبود عملکرد ناشی از بهینه‌سازی هایپرپارامترها با الگوریتم WOA است. مقایسه کلی بین مدل‌های بهینه شده در همه ایستگاه‌ها و ترکیب‌های ورودی، نشان‌دهنده برتری معنی‌دار مدل SVR-WOA است.

Archive of SID.ir

Archive of SID.ir

به‌عنوان نمونه، در ایستگاه زشک با ترکیب ورودی ۱، مدل SVR-WOA با شاخص NSE=۰/۵۵، عملکرد بهتری نسبت به مدل LSTM-WOA با NSE=۰/۳۳ از خود نشان داد. این تفاوت عملکرد به قابلیت ذاتی الگوریتم ماشین بردار پشتیبان در پردازش مؤثر داده‌های با ابعاد محدود و مقاومت بالاتر آن در برابر نقاط پرت، نسبت داده می‌شود.

جدول ۴ - پارامترها و دامنه تغییرات الگوریتم بهینه‌ساز نهنگ – رگرسیون بردار پشتیبان

ردیف	پارامتر	دامنه تغییرات
۱	kernel	Linear-poly-rbf-sigmoid
۲	C	۰/۰۱ - ۱۰۰
۳	d	۲ - ۱۰
۴	γ	۰/۱ - ۴
۵	ε	۰/۰۰۱ - ۰/۹
۶	جمعیت اولیه	۲۰۰
۷	تعداد تکرار	۱۰۰۰

ترکیب ورودی ۲ در تمامی ایستگاه‌ها بالاترین دقت پیش‌بینی را ارائه داد. برای نمونه، در ایستگاه کرتیان، مدل SVR-WOA با ترکیب‌های ورودی ۱، ۲ و ۳ به ترتیب به ضرایب NSE برابر ۰/۷۰، ۰/۸۰ و ۰/۵۶ در مرحله آزمون دست یافت. این نتایج بر اهمیت انتخاب بهینه داده‌های ورودی و نقش آن در افزایش دقت مدل تأکید دارد. تحلیل عملکرد مدل‌ها در ایستگاه‌های مختلف نشان داد که ایستگاه النگ اسدی پایین‌ترین دقت پیش‌بینی را دارد. این دقت کم عمدتاً ناشی از موقعیت مکانی و شرایط هیدرولوژیکی ایستگاه است، به‌طوری که ایستگاه النگ اسدی، به‌عنوان ایستگاه خروجی حوضه و تحت تأثیر عوامل بالادستی و تغییرات اقلیمی، به‌دلیل محدود بودن داده‌های ورودی به اطلاعات محلی، دقت کمتری ارائه داده است.

۳-۲-مدل ترکیبی تجزیه مد متغیر، بهینه‌ساز نهنگ و رگرسیون بردار پشتیبان بهینه شده

در این پژوهش، به‌منظور ارزیابی تأثیر روش تجزیه مد متغیر در فرایند مدل‌سازی از این تکنیک استفاده شد. VMD با تجزیه داده‌ها به مدهایی با فرکانس‌های گوناگون، موجب کاهش پیچیدگی داده‌ها شده و تحلیل دقیق‌تر رفتارهای کوتاه‌مدت و بلندمدت هیدرولوژیکی را فراهم می‌سازد. برای افزایش کارایی این روش، پارامترهای کلیدی شامل تعداد مدهای ذاتی (k) و پهنای باند (α) با استفاده از الگوریتم بهینه‌سازی نهنگ WOA تنظیم گردید. در شکل ۴، داده‌های تجزیه‌شده برای ایستگاه‌های مختلف نمایش داده شده است. در این نمودارها، محور (Time) نمایانگر زمان در طول دوره آماری، محور (IMFs) شماره مدهای استخراج شده، و محور (Amplitude) دامنه نوسانات هر مد در طول زمان می‌باشد. مدهای IMF1 تا IMF4 نشان‌دهنده مؤلفه‌های با فرکانس پایین و نوسانات بلندمدت مانند تغییرات اقلیمی هستند؛ در حالی‌که مدهای IMF7 تا IMF10 بیانگر روندهای کوتاه‌مدت نظیر نوسانات فصلی، رخدادهای حدی و نویزهای محلی می‌باشند. در ادامه، داده‌های دبی پردازش‌شده با استفاده از روش‌های VMD-WOA، پس از اعمال تأخیر زمانی سه‌ماهه، به‌صورت مجزا به مدل‌ها وارد شده تا برای پیش‌بینی رواناب ماه آینده مورد استفاده قرار گیرند. نتایج حاصل از این تحلیل در جدول ۶ ارائه شده است.

نتایج ارزیابی مدل‌ها نشان داد که مدل ترکیبی VMD-WOA-SVR-WOA کارآمدترین مدل است. با این حال، عملکرد مدل VMD-WOA-LSTM نیز، علیرغم عدم بهینه‌سازی VMD-WOA-LSTM، قابل‌توجه بود. به‌طور خاص در ایستگاه زشک، مدل VMD-WOA-LSTM با NSE=۰/ در مرحله آزمون، عملکردی بسیار نزدیک به مدل VMD-WOA-SVR-WOA با NSE برابر ۰/۹۷۵ در مرحله ۹۴۹ آزمون ارائه کرد. این نتایج حاکی از پایداری LSTM در محیط‌های غیرخطی با نویز متوسط و کم است. علاوه بر این، یافته‌ها بیانگر توانایی ذاتی شبکه LSTM در یادگیری الگوهای پیچیده زمانی، حتی بدون اعمال بهینه‌سازی پارامترها می‌باشد. با وجود این، مدل ترکیبی VMD-WOA-SVR-WOA در اکثر ایستگاه‌ها مانند اردک‌بند و النگ اسدی، با کسب ضرایب NSE برابر ۰/۹۸۰ و ۰/۹۸۲ در مرحله آزمون، برتری خود را نسبت به مدل‌های SVR و LSTM حفظ کرد.

از سوی دیگر، عملکرد ضعیف‌تر مدل پایه VMD-WOA-SVR در مقایسه با VMD-WOA-LSTM مانند اختلاف NSE برابر ۰/۳۳۶ بین این دو مدل در آزمون ایستگاه کرتیان، نشان می‌دهد که مدل‌های کلاسیک یادگیری ماشین مانند SVR، بدون بهره‌گیری از الگوریتم‌های بهینه‌سازی، نمی‌توانند به‌خوبی با معماری‌های عمیق در کاربردهای پیچیده هیدرولوژیکی رقابت کنند.

درحالی‌که LSTM تنها با ترکیب اولیه VMD-WOA، بدون تنظیمات پیچیده لایه‌ها، به نتایج قابل‌اطمینانی دست یافت. این یافته‌ها اهمیت انتخاب معماری مناسب بر اساس ماهیت داده‌ها را برجسته می‌سازد، به‌گونه‌ای که LSTM به دلیل ساختار بازگشتی خود، برای داده‌های سری زمانی با حافظه بلندمدت (مانند رواناب) عملکرد بهتری نسبت به SVR دارد.

Archive of SID.ir

Archive of SID.ir

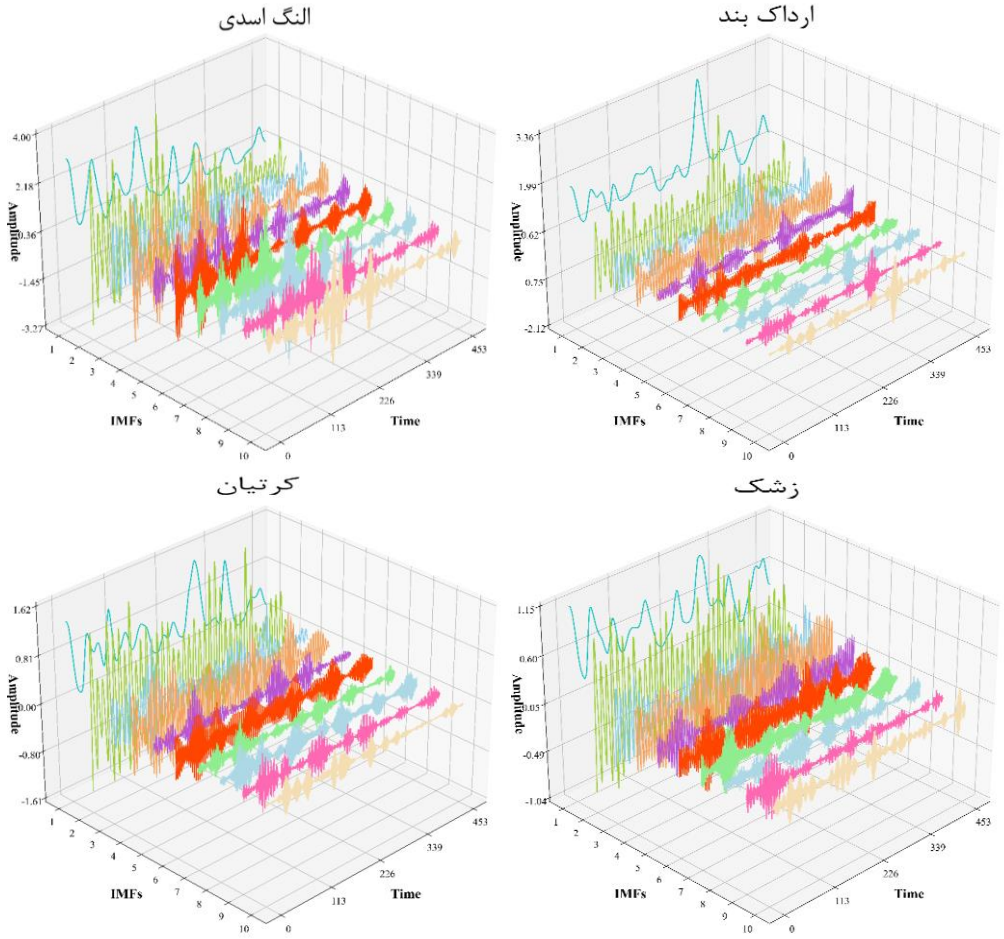
جدول ۵ – ارزیابی دقت مدل‌ها در برآورد رواناب با ترکیب داده‌های ورودی مختلف

ایستگاه	الگوی داده‌ها	مدل	آزمون						آموزش					
			NRMSE	RMSE	MAE	NSE	R^2	NRMSE	RMSE	MAE	NSE	R^2
ارداک بند	۱	SVR	۸۰	۱/۳۸	۰/۸۷	۰/۳۶	۰/۴۷	۹۳	۰/۹۴	۰/۷۶	۰/۱۴	۰/۴۶
		SVR-WOA	۸۷	۱/۱۹	۰/۵۳	۰/۵۲	۰/۵۴	۵۸	۰/۶۰	۰/۲۲	۰/۶۵	۰/۶۶
		LSTM	۶۷	۰/۶۵	۰/۳۳	۰/۴۰	۰/۴۴	۵۳	۰/۶۱	۰/۲۹	۰/۷۶	۰/۷۷
		LSTM_WOA	۶۱	۰/۵۹	۰/۳۰	۰/۴۹	۰/۵۰	۷۸	۰/۹۰	۰/۴۲	۰/۴۹	۰/۵۶
	۲	SVR	۹۹	۱/۳۴	۰/۹۰	۰/۳۹	۰/۵۳	۹۴	۰/۹۷	۰/۸۰	۰/۴۴	۰/۴۷
		SVR-WOA	۸۵	۱/۱۶	۰/۴۷	۰/۵۴	۰/۵۸	۶۸	۰/۶۹	۰/۲۶	۰/۵۲	۰/۵۳
		LSTM	۶۱	۰/۵۹	۰/۲۷	۰/۵۰	۰/۵۱	۶۸	۰/۷۹	۰/۳۳	۰/۶۰	۰/۶۴
		LSTM_WOA	۶۱	۰/۵۹	۰/۲۹	۰/۵۰	۰/۵۲	۷۰	۰/۸۲	۰/۳۵	۰/۵۹	۰/۶۳
	۳	SVR	۱۲۲	۱/۶۵	۱/۰۱	۰/۰۸	۰/۱۰	۹۷	۱/۰۰	۰/۸۱	۰/۰۲	۰/۳۳
		SVR-WOA	۹۹	۱/۳۵	۰/۶۰	۰/۳۸	۰/۴۱	۸۳	۰/۸۵	۰/۴۲	۰/۲۹	۰/۲۹
		LSTM_WOA	۶۶	۰/۶۴	۰/۳۳	۰/۴۱	۰/۴۱	۸۱	۰/۹۴	۰/۴۴	۰/۴۵	۰/۴۷
النگ اسدی	۱	SVR	۱۸۲	۲/۱۷	۱/۵۹	۰/۰۴	۰/۱۹	۲۷۲	۲/۳۶	۱/۵۳	۰/۰۹	۰/۲۲
		SVR-WOA	۱۶۱	۱/۹۲	۱/۰۱	۰/۲۵	۰/۲۶	۲۲۶	۱/۹۷	۰/۷۴	۰/۳۷	۰/۴۶
		LSTM	۱۹۹	۲/۰۹	۰/۸۵	۰/۰۴	۰/۱۵	۲۱۹	۱/۹۷	۰/۶۲	۰/۳۷	۰/۴۰
		LSTM_WOA	۱۹۲	۲/۰۱	۰/۷۶	۰/۱۱	۰/۱۴	۲۲۷	۲/۰۴	.۰/۸۳	۰/۳۲	۰/۳۵
	۲	SVR	۱۷۸	۲/۱۲	۱/۵۱	۰/۰۸	۰/۲۱	۲۷۳	۲/۳۷	۱/۴۹	۰/۰۹	۰/۲۰
		SVR-WOA	۱۵۹	۱/۹۰	۰/۹۸	۰/۲۶	۰/۲۸	۲۳۰	۲/۰۰	۰/۷۷	۰/۳۵	۰/۴۴
		LSTM	۱۸۹	۱/۹۸	۰/۷۵	۰/۱۴	۰/۲۲	۲۳۳	۲/۰۹	۰/۷۰	۰/۲۹	۰/۳۱
		LSTM_WOA	۱۸۷	۱/۹۷	۰/۷۶	۰/۱۶	۰/۲۴	۲۴۵	۲/۲۰	۰/۸۶	۰/۲۱	۰/۲۹
	۳	SVR	۲۲۱	۲/۶۵	۲/۳۷	-۰/۴	۰/۰۶	۳۴۴	۲/۹۹	۲/۵۱	-۰/۴	۰/۱۴
		SVR-WOA	۱۷۲	۲/۰۶	۱/۰۴	۰/۱۴	۰/۳۸	۲۷۱	۲/۳۵	۰/۸۶	۰/۱۰	۰/۱۸
		LSTM_WOA	۱۸۸	۲/۰۱	۱/۰۴	۰/۱۳	۰/۱۹	۲۵۵	۲/۲۹	۱/۰۸	۰/۱۴	۰/۲۵
زشک	۱	SVR	۱۰۷	۰/۷۰	۰/۵۳	۰/۴۰	۰/۴۸	۱۰۹	۰/۵۶	۰/۴۵	۰/۴۲	۰/۵۸
		SVR-WOA	۹۱	۰/۶۱	۰/۳۲	۰/۵۵	۰/۵۹	۹۴	۰/۴۷	۰/۱۸	۰/۶۰	۰/۶۲
		LSTM	۱۶۳	۰/۷۶	۰/۲۶	۰/۲۶	۰/۲۹	۷۸	۰/۴۴	۰/۲۲	۰/۶۶	۰/۷۲
		LSTM_WOA	۱۵۴	۰/۷۲	۰/۳۲	۰/۳۳	۰/۳۴	۹۵	۰/۵۴	۰/۲۸	۰/۵۰	۰/۵۴
	۲	SVR	۱۱۱	۰/۷۲	۰/۵۷	۰/۳۷	۰/۵۵	۱۲۴	۰/۶۴	۰/۵۲	۰/۲۶	۰/۵۵
		SVR-WOA	۸۷	۰/۵۶	۰/۳۰	۰/۶۱	۰/۶۷	۹۴	۰/۴۸	۰/۱۹	۰/۵۷	۰/۵۸
		LSTM	۱۵۲	۰/۷۱	۰/۲۵	۰/۳۵	۰/۳۶	۸۶	۰/۴۹	۰/۲۲	۰/۵۹	۰/۶۲
		LSTM_WOA	۱۵۱	۰/۷۱	۰/۳۳	۰/۳۵	۰/۳۶	۸۴	۰/۴۸	۰/۲۸	۰/۶۱	۰/۶۱
	۳	SVR	۱۱۷	۰/۷۶	۰/۶۰	۰/۲۹	۰/۴۱	۱۲۹	۰/۶۶	۰/۵۴	۰/۲۰	۰/۴۵
		SVR-WOA	۹۸	۰/۶۴	۰/۳۲	۰/۵۰	۰/۵۲	۱۰۲	۰/۵۳	۰/۲۲	۰/۴۹	۰/۵۱
		LSTM_WOA	۱۶۰	۰/۷۶	۰/۳۰	۰/۲۷	۰/۲۷	۹۷	۰/۵۶	۰/۳۰	۰/۴۷	۰/۴۹
کرتیان	۱	SVR	۱۴۹	۰/۸۸	۰/۷۷	۰/۲۱	۰/۵۸	۲۰۶	۰/۹۶	۰/۷۸	/۲۲	۰/۵۲
		SVR-WOA	۹۲	۰/۵۴	۰/۳۴	۰/۷۰	۰/۷۳	۱۶۳	۰/۷۶	۰/۳۳	۰/۵۱	۰/۵۲
		LSTM	۲۱۶	۰/۷۷	۰/۳۰	۰/۲۴	۰/۵۱	۹۸	۰/۵۵	۰/۲۳	۰/۷۳	۰/۷۳
		LSTM_WOA	۱۸۳	۰/۶۶	۰/۲۳	۰/۴۶	۰/۴۸	۱۱۷	۰/۶۶	۰/۲۸	۰/۶۱	۰/۶۶
	۲	SVR	۱۴۶	۰/۸۶	۰/۷۹	۰/۲۴	۰/۶۹	۲۲۱	۱/۰۴	۰/۸۵	۰/۱۰	۰/۴۸
		SVR-WOA	۷۵	۰/۴۴	۰/۲۵	۰/۸۰	۰/۸۴	۱۶۰	۰/۷۵	۰/۲۵	۰/۵۲	۰/۵۴
		LSTM	۱۸۳	۰/۶۵	۰/۲۵	۰/۴۶	۰/۴۹	۱۱۴	۰/۶۴	۰/۲۷	۰/۶۳	۰/۶۴
		LSTM_WOA	۱۷۸	۰/۶۴	۰/۲۷	۰/۴۹	۰/۵۲	۱۱۶	۰/۶۵	۰/۲۸	/۶۲	۰/۶۴
	۳	SVR	۱۸۴	۱/۰۹	۰/۳۲	-۰/۲	۰/۳۵	۲۴۷	۱/۱۶	۱/۰۱	-۰/۱	۰/۴۴
		SVR-WOA	۱۱۲	۰/۶۶	۰/۳۵	۰/۵۶	۰/۶۰	۱۷۵	۰/۸۲	۰/۳۳	۰/۴۳	۰/۴۶
		LSTM_WOA	۱۸۳	۰/۶۷	۰/۳۰	۰/۴۴	۰/۴۶	۱۲۴	۰/۷۰	۰/۳۵	۰/۵۶	۰/۵۸

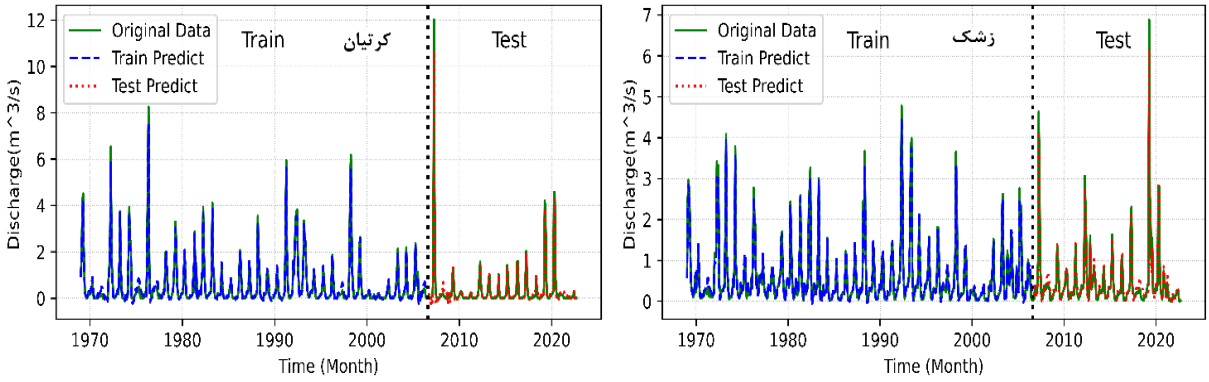
Archive of SID.ir

Archive of SID.ir

به‌طور کلی، مدل ترکیبی ارائه‌شده در این پژوهش، پیشرفت قابل‌توجهی در مدل‌سازی هیدرولوژیکی ایجاد کرده است. این مدل، با کاهش نویز، تفکیک بهتر روندها و نوسانات، و استفاده مؤثر از الگوریتم‌های بهینه‌سازی، ابزار قدرتمندی برای پیش‌بینی رواناب و مدیریت منابع آب محسوب می‌شود. در شکل ۵ داده‌های رواناب مشاهداتی و برآورد شده توسط مدل VMD-WOA-SVR-WOA مقایسه شده‌اند. همان‌گونه که مشاهده می‌گردد، بین داده‌های پیش‌بینی‌شده و داده‌های مشاهداتی برای مراحل آزمون و تست تطابق خوبی برقرار است.



شکل ۴ – مدهای تجزیه شده ایستگاه‌های مورد مطالعه با استفاده از روش VMD-WOA



شکل ۵ – مقایسه داده‌های رواناب برآورد شده مدل VMD-WOA-SVM-WOA و مشاهداتی ایستگاه‌ها

Archive of SID.ir

Archive of SID.ir

جدول ۶ - ارزیابی دقت مدل‌ها در برآورد رواناب با داده‌های ورودی حاصل از تجزیه مد متغیر

ایستگاه	مدل	آزمون					آموزش				
		NRMSE	RMSE	MAE	NSE	R^2	NRMSE	RMSE	MAE	NSE	R^2
ارداک بند	VMD_WOA_LSTM	۱۸/۲	۰/۱۵۳	۰/۱۱۳	۰/۹۶۷	۰/۹۷۵	۲۳/۷	۰/۲۴۸	۰/۱۷۴	۰/۹۶۱	۰/۹۶۸
	VMD-WOA-SVR	۵۵/۳	۰/۵۲۱	۰/۲۵۳	۰/۷۳۹	۰/۷۶۳	۲۵/۵	۰/۲۵۷	۰/۱۷۸	۰/۹۴۵	۰/۹۵۷
	VMD-WOA-SVR-WOA	۱۵/۳	۰/۱۴۵	۰/۰۸۳	۰/۹۸۰	۰/۹۸۶	۱۵/۹	۰/۱۸۴	۰/۱۲۶	۰/۹۷۸	۰/۹۸۴
النگ اسدی	VMD_WOA_LSTM	۳۴/۹	۰/۳۷۲	۰/۲۴۹	۰/۹۷۰	۰/۹۸۱	۴۷/۶	۰/۴۳۴	۰/۲۵۵	۰/۹۶۹	۰/۹۸۱
	VMD-WOA-SVR	۵۵/۳	۰/۵۶۳	۰/۲۶۹	۰/۹۴۲	۰/۹۵۶	۶۱/۶	۰/۵۵۴	۰/۲۶۳	۰/۹۵۰	۰/۹۶۶
	VMD-WOA-SVR-WOA	۳۰/۶	۰/۳۱۴	۰/۱۹۴	۰/۹۸۲	۰/۹۸۸	۳۵/۳	۰/۳۱۹	۰/۲۰۵	۰/۹۸۶	۰/۹۸۸
زشک	VMD_WOA_LSTM	۳۹/۳	۰/۱۹۳	۰/۱۳۰	۰/۹۴۹	۰/۹۶۵	۲۳/۶	۰/۱۹۳	۰/۱۰۲	۰/۹۶۵	۰/۹۷۳
	VMD-WOA-SVR	۷۵	۰/۳۵۱	۰/۱۸۱	۰/۸۰۲	۰/۸۲۸	۳۳/۹	۰/۰۶	۰/۱۳۶	۰/۹۳۶	۰/۹۴۵
	VMD-WOA-SVR-WOA	۲۶/۶	۰/۱۲۴	۰/۰۸۳	۰/۹۷۵	۰/۹۸۲	۱۷/۴	۰/۰۹۹	۰/۰۷۲	۰/۹۸۳	۰/۹۸۶
کرتیان	VMD_WOA_LSTM	۴۰/۶	۰/۱۵۶	۰/۱۰۹	۰/۹۷۱	۰/۹۷۹	۳۵/۲	۰/۲۰۴	۰/۱۴۳	۰/۹۶۲	۰/۹۷۱
	VMD-WOA-SVR	۱۹۴/۴	۰/۶۵۳	۰/۲۱۲	۰/۶۴۳	۰/۶۶۲	۳۹/۷	۰/۲۲۲	۰/۱۵۶	۰/۹۵۵	۰/۹۶۴
	VMD-WOA-SVR-WOA	۴۶/۳	۰/۱۵۶	۰/۰۸۶	۰/۹۷۹	۰/۹۸۷	۴۶/۸	۰/۱۵۰	۰/۱۰۴	۰/۹۷۹	۰/۹۸۴

نتیجه گیری

در این پژوهش، دو رویکرد متفاوت برای مدل‌سازی رواناب بررسی شد. در روش نخست، پیش‌پردازش داده‌ها با استفاده از نرمال‌سازی انجام شد. در این روش، مدل‌های بهینه‌شده LSTM و SVR نسبت به مدل‌های پایه عملکرد بهتری از خود نشان دادند، با این حال به‌دلیل پیچیدگی ساختاری داده‌های هیدرولوژیکی، توانایی مدل‌ها در شناسایی الگوهای دقیق و برآورد رواناب محدود بود. به‌طور کلی، با توجه به مقادیر پایین شاخص نش-ساتکلیف، عملکرد این روش در سطح مطلوبی ارزیابی نمی‌شود. در روش دوم، پیش‌پردازش داده‌ها با استفاده از تجزیه مد متغیر انجام شد. این تکنیک با تفکیک داده‌های هیدرولوژیکی به مؤلفه‌هایی با فرکانس‌های مختلف، امکان استخراج ویژگی‌های دقیق‌تر را فراهم ساخت. در این حالت، پس از بهینه‌سازی پارامترها توسط الگوریتم نهنگ، مدل‌های SVRو LSTM مورد ارزیابی قرار گرفتند. نتایج نشان داد که مدل ترکیبی VMD-WOA-SVR-WOA عملکرد بسیار مطلوبی در برآورد رواناب داشته و نسبت به رویکرد اول، بهبود قابل توجهی در دقت و تعمیم‌پذیری مدل‌ها ایجاد کرده است. اگرچه مدل VMD-WOA-LSTM نیز نتایج خوبی ارائه داد، اما عملکرد برتر مدل VMD-WOA-SVR-WOA آن را به عنوان ابزار کارآمدتری در برآورد رواناب معرفی می‌کند. یافته‌های این پژوهش نشان داد که استفاده از تکنیک‌های پیش‌پردازش داده‌ها، مانند تجزیه مد متغیر همراه با بهینه‌سازی الگوریتمی، دقت مدل‌های پیش‌بینی رواناب را به‌طور قابل‌توجهی افزایش می‌دهد. مدل ترکیبی VMD-WOA-SVR-WOA به دلیل کاهش نویز و تفکیک بهتر روندها و نوسانات، به‌عنوان یک ابزار قدرتمند برای پیش‌بینی و برآورد رواناب و مدیریت منابع آب، به‌ویژه در شرایط بحرانی مانند سیلاب‌ها، پیشنهاد می‌شود.

تقدیر و تشکر

نویسندگان مقاله از شرکت سهامی آب منطقه‌ای خراسان رضوی که اطلاعات را در اختیار قرار دادند، تشکر می‌نمایند.

تضاد منافع نویسندگان

در این مقاله هیچ‌گونه تضاد منافعی وجود ندارد و این مسئله مورد تایید همه نویسندگان است

دسترسی به داده‌ها

داده‌های این تحقیق بخشی از داده‌های مربوط به پایان نامه کارشناسی ارشد نویسنده اول می‌باشد.

مشارکت نویسندگان

نویسنده اول، جمع‌آوری و تحلیل اولیه داده‌ها و انجام مدل-سازی را به‌عهده داشته‌است.

نویسندگان دوم و سوم، تحلیل داده‌ها، راهنمایی، ویرایش و بازبینی مقاله و کنترل نتایج را عهده‌دار بوده‌اند.

اصول اخلاقی

نویسندگان اصول اخلاقی را در انجام و انتشار این اثر عملی رعایت نموده‌اند و این موضوع مورد تایید همه آنها می‌باشد.

Archive of SID.ir

Archive of SID.ir

منابع

1- Araghinejad, S. and Karamouz, M. 2005. Long-Lead Streamflow Forecasting using Artificial Neural Networks and Fuzzy Inference System. *Iran-Water Resources Research*, 1(2), pp. 29–41. (In Persian).

2- Bagherpour, B., Seyedian, S.M., Fatabadi, A. and Mohammadi, A. 2019. Runoff Prediction Using Black and Gray Box Models. *Iran-Water Resources Research*, 14(5), pp. 177–192. (In Persian).

3- Cortes, C. and Vapnik, V. 1995. Support-vector networks. *Machine Learning*, 20(3), pp. 273–297. https://doi.org/10.1007/BF00994018

4- Davanlou Tajbakhsh, A., Nourani, V. and Molajou, A. 2019. Hybrid Wavelet-M5 Model Application in Rainfall-Runoff Process Forecast (Case study: Aji Chay Catchment). *Iran-Water Resources Research*, 15(2), pp. 1–10. (In Persian).

5- Dragomiretskiy, K. and Zosso, D. 2014. Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, 62(3), pp. 531–544. https://doi.org/10.1109/TSP.2013.2288675

6- Dunyaei, A., Saraf, A. and Ahmadi, H. 2020. Evaluation of the algorithms of the whale, magpie and cuckoo whales in optimizing the multi-objective use of the reservoir Golestan dam based on multi-criteria decision making method. *Journal of Water Resources Engineering*, 13(4), pp. 85–100. (In Persian).

7- Feng, Z., Niu, W., Tang, Z., Jiang, Z., Xu, Y., Liu, Y. and Zhang, H. 2020. Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization. *Journal of Hydrology*, 583, p. 124627. https://doi.org/10.1016/j.jhydrol.2020.124627

8- Ghahraman, B. 2013. Performance of some Different Methods of Computing Hurst Exponent for Climatological Time Series. *Journal of Water and Soil*, 27(4), pp. 850–859. (In Persian).

9- Guo, S., Wen, Y., Zhang, X. and Chen, H. 2023. Monthly runoff prediction using the VMD-LSTM-Transformer hybrid model: a case study of the Miyun Reservoir in Beijing. *Journal of Water and Climate Change*, 14(9), pp. 3221–3236. https://doi.org/10.2166/wcc.2023.257

10- Hochreiter, S. and Schmidhuber, J. 1997. Long Short-Term Memory. *Neural Computation*, 9(8), pp. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

11- Kavusi, M. and Khozeymeh Nejad, H. 2020. Review and compare performance of 4 modeling methods LS-SVM, NN, GEP and ANFIS-PSO in Simulation of Rainfall - Runoff (Study Area: Halil River - Jiroft Dam). *Scientific Research Journal of Irrigation and Water Engineering of Iran*, 43. (In Persian).

12- Kumar, V., Kedam, N., Sharma, K.V., Mehta, D.J. and Caloiero, T. 2023. Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. *Water*, 15(14), p. 2572. https://doi.org/10.3390/w15142572

13- Merufinia, E., Sharafati, A., Abghari, H. and Hassanzadeh, Y. 2023. Performance evaluation of ANN-WOA and ANN-BWO simulation-optimization methods in predicting daily runoff (case study: Jelogir station in Karkheh watershed). *Irrigation Sciences and Engineering (JISE)*, 47, pp. 19–36. (In Persian).

14- Mirjalili, S. and Lewis, A. 2016. The Whale Optimization Algorithm. *Advances in Engineering Software*, 95, pp. 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

15- Said, S.E. and Dickey, D.A. 1984. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71(3), p. 599. https://doi.org/10.2307/2336570

16- Sayed, B.T., Al-Mohair, H.K., Alkhayyat, A., Ramírez-Coronel, A.A. and Elsahabi, M. 2023. Comparing machine-learning-based black box techniques and white box models to predict rainfall-runoff in a northern area of Iraq, the Little Khabur River. *Water Science and Technology*, 87(3), pp. 812–822. https://doi.org/10.2166/wst.2023.014

17- Shafeizadeh, M., Fathian, H. and Nikbakht Shahbazi, A.R. 2019. Continuous Rainfall-Runoff Simulation by Artificial Neural Networks Based on Selection of Effective Input Variables Using Partial Mutual Information (PMI) Algorithm. *Iran-Water Resources Research*, 15, pp. 144–161. (In Persian).

18- Tang, W., Pei, Y., Zheng, H., Zhao, Y., Shu, L. and Zhang, H. 2022. Twenty years of China's water pollution control: Experiences and challenges. *Chemosphere*, 295, p. 133875. https://doi.org/10.1016/j.chemosphere.2022.133875

19- Taormina, R. and Chau, K.-W. 2015. Data-driven input variable selection for rainfall–runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines. *Journal of Hydrology*, 529, pp. 1617–1632. https://doi.org/10.1016/j.jhydrol.2015.08.022

20- Tikhamarine, Y., Souag-Gamane, D., Ahmed, A.N., Sammen, S.S., Kisi, O., Huang, Y.F. and El-Shafie, A. 2020. Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. *Journal of Hydrology*, 589, p. 125133. https://doi.org/10.1016/j.jhydrol.2020.125133

21- Van Houdt, G., Mosquera, C. and Nápoles, G. 2020. A review on the long short-term memory model. *Artificial Intelligence Review*, 53(8), pp. 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

22- Xu, D., Li, Z. and Wang, W. 2024. An ensemble model for monthly runoff prediction using least squares support vector machine based on variational modal decomposition with dung beetle optimization algorithm and error correction strategy. *Journal of Hydrology*, 629, p.

Archive of SID.ir

Archive of SID.ir

130558. https://doi.org/10.1016/j.jhydrol.2023.130558

23- Yao, Z., Wang, Z., Wang, D., Wu, J. and Chen, L. 2023. An ensemble CNN-LSTM and GRU adaptive weighting model based improved sparrow search algorithm for predicting runoff using historical meteorological and runoff data as input. *Journal of Hydrology*, 625, p. 129977. https://doi.org/10.1016/j.jhydrol.2023.129977

24- Yousefi, A., Ghafoori, M. and Lashkaripour, G. 2007. Investigation of clay minerals in the subgrade soil of Mashhad city considering their Atterberg limits. Paper presented at the *Iranian Conference on Engineering Geology and the Environment*, Tehran, Iran, 22 July 2007. (In Persian).

25- Zhang, J., Chen, X., Khan, A., Zhang, Y., Kuang, X., Liang, X., Taccari, M.L. and Nuttall, J. 2021. Daily runoff forecasting by deep recursive neural network. *Journal of Hydrology*, 596, p. 126067. https://doi.org/10.1016/j.jhydrol.2021.126067

Archive of SID.ir