

Estimation of Suspended Sediments in the Coastal Areas of the Caspian Sea Using Machine Learning Techniques Under Drought Conditions

Ebrahim Nohani¹, Hamidreza Babaali²*, Reza Dehghani³

1–Assistant Professor, Department of Civil Engineering, Materials and Energy Research Center, Dezful Branch, Islamic Azad University, Dezful, Iran.

2 –Associate Professor, Department of Civil Engineering, Islamic Azad University, Khorramabad branch, Khorramabad, Iran

3 –PhD in Water Sciences and Engineering, Department of Soil Conservation and Watershed Management, Lorestan Province Agriculture and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization, Khorramabad, Iran.

***Corresponding Author: hr.babaali91@gmail.com**

Extended abstract

Introduction

Accurate modeling of suspended sediment transported by rivers is of great importance in environmental engineering and water resource management, as it directly influences the design, planning, operation, and management of water systems. Moreover, suspended sediment modeling is critical because it significantly affects reservoir capacity, dam operation, water quality, and pollutant transport. However, estimating suspended sediment remains a challenging task for hydrology engineers due to the complex and non-linear interactions between sediment dynamics, watershed geomorphology, and water flow. Suspended sediment transport in rivers is governed by a combination of meteorological and hydrological parameters, making it a complex process to model accurately. In recent years, the non-linear nature of hydrological phenomena, including sediment transport, has led to the increased use of artificial intelligence models. In this research, hybrid models—including Support Vector Regression–Wavelet, Support Vector Regression–Whale Optimization, and Support Vector Regression–Particle Swarm Optimization—were employed to estimate suspended sediment in the Babolrood River, located in Mazandaran Province along the Caspian Sea coast.

Materials and Methods

In this study, the required data were obtained from the Regional Water Company of Mazandaran Province. The Babolrood hydrometric station was selected because it is located in the Caspian Sea coastal area, has no missing data, and offers a long-term statistical record. The parameters used include river discharge and sediment levels from 2002 to 2022 (Iranian calendar). Of this data, 80 percent (2002-2018) was used for training, and 20 percent (2018-2022) was used for testing. To model the suspended sediment levels of the Caspian Sea coasts, the data were first normalized and then input into the model. To construct the hybrid model, the tuning parameters of the kernel functions of the Support Vector Regression (SVR) model were estimated using meta-heuristic algorithms, including Wavelet, Whale, and Particle Swarm Optimization. These parameters were then entered into the SVR model, forming the structure of the hybrid model. The input data were subsequently entered into the hybrid model, and the optimization process of the tuning parameters continued until the model error reached the lowest possible level.

Results

To model the suspended sediment load of the Babolrood River, a Support Vector Regression (SVR) model was used with Wavelet, Whale, and Particle Swarm Optimization algorithms. The results showed that the hybrid models, particularly in the scenario that included all input parameters, performed better than other scenarios. Therefore, increasing the number of effective parameters in artificial neural network-based hybrid models leads to enhanced model performance. Additionally, all models showed better accuracy when using the radial basis kernel function.

According to the evaluation criteria, the Support Vector Regression-Wavelet model, in the combined scenario that included all input parameters, demonstrated the best performance during the validation stage. It achieved the highest correlation coefficient of 0.962, the lowest root mean square error of 0.344 (ton/day), the lowest mean absolute error of 0.158 (ton/day), and the highest Nash-Sutcliffe efficiency coefficient of 0.970.

Conclusion

The research results, based on the evaluation of scenarios consisting of input parameters, showed that in all the models studied, increasing the number of effective parameters led to better performance in estimating river sediment levels. Additionally, the evaluation criteria indicated that the Support Vector Regression-Wavelet model exhibited high accuracy and negligible error. The investigated graphs also showed that the Support Vector Regression-Wavelet model estimated the river's suspended sediment values close to their actual values, as evident in the box plot and Taylor diagrams. Overall, the results of this research suggest that artificial intelligence models based on the Support Vector Regression approach can be used to estimate sediment discharge over 30 years of statistical data for other regions of the country, serving as a step toward making informed management decisions.

Keywords: Babolroud, Support vector regression, Suspended sediment, Modeling

برآورد رسوبات معلق نواحی ساحلی دریای خزر با استفاده از تکنیک یادگیری ماشین در شرایط خشک‌سالی

ابراهیم نوحانی^۱، حمیدرضا باباعلی^{۲*}، رضا دهقانی^۳

۱- استادیار گروه عمران، مرکز تحقیقات مواد و انرژی، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران.

۲- دانشیار، گروه مهندسی عمران، دانشگاه آزاد اسلامی واحد خرم‌آباد، خرم‌آباد، ایران.

۳- دکترای علوم و مهندسی آب، بخش تحقیقات حفاظت خاک و آبخیزداری، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان لرستان، سازمان تحقیقات، آموزش و ترویج کشاورزی، خرم‌آباد، ایران.

*نویسنده مسئول: hr.babaali91@gmail.com

چکیده

مدل‌سازی رسوبات معلق یکی از موضوعات کلیدی در مدیریت منابع آب و حفاظت از محیط‌زیست به‌شمار می‌آید. این فرآیند به تصمیم‌گیرندگان در سطح حوضه آبخیز کمک می‌کند تا درک بهتری از رفتار رسوبات و تأثیرات آن بر کیفیت آب و اکوسیستم‌های آبی داشته باشند. در این تحقیق، یک رویکرد هوشمند ترکیبی جدید مبتنی بر رویکرد مدل رگرسیون بردار پشتیبان برای برآورد رسوبات رودخانه توسعه داده شده است. بدین‌منظور در این پژوهش از دو الگوریتم بهینه‌سازی شامل نهنگ و ازدحام ذرات برای مدل‌سازی میزان رسوبات معلق رودخانه به کار برده شد. جهت مدل‌سازی از آمار و اطلاعات ایستگاه هیدرومتری رودخانه بابلرود واقع در استان مازندران به‌عنوان مطالعه موردی طی ۵ سناریو ترکیبی از پارامترهای ورودی در سال‌های ۱۳۸۲ تا ۱۴۰۲ استفاده شد. به‌منظور ارزیابی عملکرد مدل‌ها از معیارهای ارزیابی ضریب همبستگی، ریشه میانگین مربعات خطا، میانگین قدر مطلق خطا و ضریب نش ساتکلیف استفاده شد. نتایج نشان داد سناریوهای ترکیبی در مدل‌های موردبررسی باعث بهبود عملکرد مدل می‌شود. همچنین نتایج حاصل از معیار ارزیابی نشان داد مدل رگرسیون بردار پشتیبان-موجک دارای ضریب همبستگی ۰/۹۶۲، ریشه میانگین مربعات خطا (ton/day) ۰/۳۴۴، میانگین قدر مطلق خطا (ton/day) ۰/۱۵۸ و ضریب نش ساتکلیف ۰/۹۷۰ در مرحله صحت سنجی برخوردار است. درمجموع نتایج نشان داد استفاده از مدل‌های هوشمند مبتنی بر رویکرد رگرسیون بردار پشتیبان می‌تواند رویکردی مؤثر در پایداری مهندسی رودخانه باشد.

واژگان کلیدی: بابلرود، رگرسیون بردار پشتیبان، رسوبات معلق، مدل‌سازی.

۱- مقدمه

مدل‌سازی دقیق رسوبات معلق منتقل‌شده توسط یک رودخانه از اهمیت بالایی در مهندسی محیط‌زیست و منابع آب برخوردار است، زیرا به‌طور مستقیم بر طراحی، برنامه‌ریزی، بهره‌برداری و مدیریت منابع آب تأثیر می‌گذارد (Kisi et al., 2012). علاوه بر این، مدل‌سازی رسوبات معلق بسیار مهم است زیرا تأثیر عمده‌ای بر ظرفیت مخزن و بهره‌برداری از سد، کیفیت آب و انتقال آلاینده‌ها دارد (Cimen, 2008). با این حال، تخمین رسوبات معلق یک کار چالش‌برانگیز برای مهندسين هیدرولوژی است زیرا تعامل آن با ویژگی‌های ژئومورفولوژیکی حوضه آبریز و جریان آب بسیار پیچیده و غیرخطی است. انتقال رسوبات معلق در رودخانه تابعی از پارامترهای هواشناسی و هیدرولوژیکی به‌عنوان یک فرآیند پیچیده است.

با توجه به مشکلات بسیاری که رسوبات در رودخانه‌ها ایجاد می‌کنند، تلاش‌های زیادی توسط محققان علم رسوب برای دستیابی به روابط حمل‌ونقل رسوب در مطالعات آزمایشگاهی و میدانی صورت گرفته است (Nayak et al., 2013). شرایط مرزی و ساده‌سازی برخی از شاخص‌های مهم منجر به محدودیت‌هایی در معادلات تجربی رسوب شده است که بر اساس ویژگی‌های جریان و رسوب طراحی شده‌اند (Goyal, 2014). به‌تازگی، مدل‌های محاسبات نرم^۱ به‌منظور بهبود دقت برآورد رسوبات رودخانه به کار گرفته شده‌اند. در سال‌های اخیر، این مدل‌ها در زمینه‌های متنوعی از جمله مشکلات هیدرولوژی و منابع آب به‌ویژه در برآورد بار رسوب استفاده شدند (Wu and Chau, 2020). این مدل‌ها از طبیعت و رفتار موجودات زنده الهام گرفته و قادرند مسائل با پیچیدگی و گستردگی فراوان را حل نمایند. این مدل‌ها در زمینه پیش‌بینی رسوبات معلق رودخانه‌ها مورد توجه پژوهشگران قرار گرفته است که می‌توان به موارد زیر اشاره نمود.

حسن‌پور و همکاران (Hassanpour et al., 2019) در پژوهش خود از مدل‌های رگرسیون بردار خوشه‌بندی-پشتیبانی فازی^۲، منحنی رتبه‌بندی رسوب^۳، شبکه عصبی مصنوعی^۴، شبکه عصبی تطبیقی^۵ برای پیش‌بینی بار رسوبی رودخانه سیستان واقع در کشور ایران استفاده کردند در این پژوهش از داده‌های روزانه دبی و رسوب طی سال‌های ۱۹۹۶ تا ۲۰۱۲ با تأخیرهای زمانی متفاوت بهره گرفته شد. نتایج نشان داد مدل ترکیبی مبتنی بر رگرسیون بردار پشتیبان عملکرد بهتری نسبت به سایر مدل‌های هوشمند دارند.

نورانی و همکاران (Nourani et al., 2021) در پژوهش خود برای برآورد رسوبات معلق رودخانه کاتار واقع در کشور اتیوپی از مدل‌های مبتنی بر هوش مصنوعی شامل ماشین بردار پشتیبان^۶، سامانه استنتاج عصبی فازی تطبیقی^۷، شبکه عصبی پیش‌خور^۸ و یک حالت رگرسیون چند خطی معمولی^۹ استفاده کردند. در این پژوهش از داده‌های روزانه دبی و رسوب طی سال‌های ۲۰۰۷ تا ۲۰۱۷ استفاده شد. نتایج نشان داد مدل ماشین بردار پشتیبان نسبت به سایر مدل‌های مورد بررسی از دقت بهتری برخوردار است.

دروودی و همکاران (Doroudi et al., 2021) در پژوهش خود به‌منظور تخمین رسوبات معلق رودخانه چم سیاه در کشور ایران از مدل هیبریدی جدید رگرسیون بردار پشتیبان-الگوریتم معلم یادگیرنده^{۱۰} استفاده کردند در این پژوهش از داده‌های روزانه دبی و بارش طی سال‌های ۱۹۸۶ تا ۲۰۱۵ استفاده شد. نتایج نشان داد مدل هیبریدی جدید رگرسیون بردار پشتیبان با دقت ۰/۹۹۸ و

¹ Soft computing

² Fuzzy clustering-support vector regression

³ Sediment rating curve

⁴ Artificial neural network

⁵ Adaptive neural network

⁶ Support vector machine

⁷ Adaptive neural fuzzy

⁸ Feedforward neural network

⁹ Multilinear regression

¹⁰ Teacher-learner algorithm

میانگین قدر مطلق خطای ۰/۰۲۱ نسبت به مدل‌های هیبریدی مرسوم همچون رگرسیون بردار پشتیبان-ژنتیک از دقت بیشتری برخوردار است.

ایسام و همکاران (Essam et al., 2022) برای پیش‌بینی میزان رسوبات معلق رودخانه سوگای واقع در کشور مالزی از مدل‌های یادگیری ماشین بر اساس الگوریتم‌های ماشین بردار پشتیبان، شبکه عصبی مصنوعی و حافظه کوتاه‌مدت^{۱۱} استفاده کردند در این پژوهش از داده‌های روزانه شامل دبی و رسوب با تأخیرهای زمانی طی دوره آماری ۲۰۱۵ تا ۲۰۲۰ استفاده شد. نتایج نشان داد مدل ماشین بردار پشتیبان ابزاری مفید به‌منظور پیش‌بینی رسوبات معلق رودخانه می‌باشد.

در یک جمع‌بندی می‌توان اشاره داشت، با توجه به پژوهش‌های انجام‌شده، مدل رگرسیون بردار پشتیبان می‌تواند به‌عنوان ابزاری کارآمد در برآورد رسوبات معلق رودخانه و مسائل هیدرولوژیکی قرار گیرد. همچنین طبق پژوهش‌های ذکرشده مدل منفرد رگرسیون بردار پشتیبان در برآورد رسوبات معلق دارای خطا بوده و به‌منظور بهبود عملکرد این مدل از ترکیب الگوریتم‌های فرا ابتکاری جهت بهینه نمودن پارامترهای تنظیم استفاده می‌شود و نتایج این الگوریتم‌ها با مدل هیبریدی ترکیب‌شده با تبدیل موجک به دلیل عملکرد بالای آن سنجیده می‌شود. امروزه به‌منظور افزایش کارایی و بهبود عملکرد مدل رگرسیون بردار پشتیبان، از ترکیب این مدل با الگوریتم‌های فرا ابتکاری به‌عنوان یک راهکاری مناسب برای برآورد رسوبات معلق رودخانه‌ها استفاده می‌شود. در این پژوهش نیز از مدل‌های هیبریدی رگرسیون بردار پشتیبان-موجک، رگرسیون بردار پشتیبان-نهنک و رگرسیون بردار پشتیبان-ازدحام ذرات برای برآورد رسوبات معلق رودخانه بابلرود واقع در استان مازندران و سواحل دریای خزر استفاده شد.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

سواحل جنوبی دریای خزر به‌ویژه منطقه بابلسر آب‌وهوای معتدل و مرطوب که نواری بین رشته‌کوه البرز و دریای خزر است، توسط دشت‌های پست احاطه شده است. به‌طور کلی این منطقه کوچک‌ترین منطقه آب و هوایی ایران است و از دو منطقه جلگه‌ای و کوهستانی تشکیل شده است. بابلسر بین ۵۲ درجه و ۳۹ دقیقه و ۳۰ ثانیه طول جغرافیایی و ۳۶ درجه و ۴۳ دقیقه عرض جغرافیایی واقع شده و در شمال ایران و از نظر رژیم هیدرولوژیکی به بخش‌های شمالی، میانی و جنوبی تقسیم می‌شود. در شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه نشان داده شده است. در سال‌های اخیر به دلیل کاهش بارش و دبی رودخانه بابلرود بستر این رودخانه با حجم قابل توجهی از رسوبات مواجه شده است.

¹¹ Short-term memory



شکل ۱- منطقه مورد مطالعه

Figure 1. Study area

۲-۲- داده‌های مورد استفاده و روش تحقیق

در این پژوهش داده‌های مورد نیاز از شرکت آب منطقه‌ای استان مازندران اخذ شد. ایستگاه هیدرومتری بابلرود به دلیل آنکه در نواحی سواحل دریای خزر واقع شده و فاقد داده گمشده و آمار طولانی مدت است، انتخاب شد. پارامترهای مورد استفاده شامل دبی و میزان رسوبات رودخانه طی سال‌های ۱۳۸۲ تا ۱۴۰۲ می‌باشد که ۸۰ درصد آن‌ها (۱۳۸۲-۱۴۰۰) جهت آموزش و ۲۰ درصد (۱۴۰۱-۱۴۰۲) جهت آزمون بررسی شدند. به منظور مدل‌سازی میزان رسوبات معلق سواحل دریای خزر ابتدا داده‌های مورد استفاده مطابق رابطه ۱ نرمال‌سازی شده و سپس وارد مدل می‌گردد. به منظور ساخت مدل هیبریدی پارامترهای تنظیم توابع کرنل مدل رگرسیون بردار پشتیبان به وسیله الگوریتم‌های فرا ابتکاری شامل مویک، نهنگ و ازدحام ذرات برآورد شده و سپس وارد مدل رگرسیون بردار پشتیبان می‌شود و ساختار مدل هیبریدی شکل می‌گیرد. سپس داده‌های ورودی وارد مدل هیبریدی شده و منجر به خروجی مدل می‌گردد و فرایند بهینه‌سازی پارامترهای تنظیم تا زمانی که خطای مدل به کمترین میزان ممکن برسد انجام گرفت. در شکل ۲ فلوچارت پژوهش نشان داده شده است.

$$X_n = 0.1 + 0.8 \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$



شکل ۲-فلوچارت تحقیق

Figure 2. Research flowchart

۲-۲- رگرسیون بردار پشتیبان^{۱۲}

رگرسیون بردار پشتیبان یکی از روش های هوش مصنوعی می باشد که بر مبنای تئوری بهینه سازی و از قانون کمینه سازی خطا پیروی می نماید که این امر سبب می گردد به یک جواب بهینه کلی منجر شود (Vapnik, 1995). در مدل رگرسیون SVR که شامل تابعی با متغیرهای وابسته Y می باشد که این متغیر وابسته از چند متغیر مستقل X و مقداری خطا تشکیل شده است. همانطور که در مسائل رگرسیون مشاهده می شود میان متغیر وابسته و مستقل رابطه جبری روابط ۲ و ۳ وجود دارد (Vapnik, 1998)

$$f(x) = W^T \cdot \phi(x) + b \quad (2)$$

$$y = f(x) + \text{noise} \quad (3)$$

در معادله بالا، W بردار ضرایب، b یک عدد ثابت است که نشان دهنده ویژگی های تابع رگرسیون است و $\phi(x)$ تابع کرنل است. هدف سپس یافتن شکل تابعی برای $f(x)$ است که از رابطه ۲ حاصل می شود. در رابطه ۲، noise نشان دهنده خطای مجاز است که

¹² Support Vector Regression

فرآیند تلاش می‌نماید تا در هر تکرار آن را به صفر برساند تا تابع به‌ترین پاسخ را ارائه دهد کرنل‌های مختلف می‌توان برای ایجاد انواع مختلف SVM-ε اعمال کرد. مدل SVM از توابع کرنل مختلف از جمله کرنل‌های چندجمله‌ای، تابع پایه شعاعی و خطی پشتیبانی می‌کند آن‌ها را از طریق روابط ۴، ۵ و ۶ محاسبه می‌شوند. (vapnik and Chervoneskis, 1991; Basak et al., 2007). در این پژوهش نیز از این سه تابع کرنل استفاده شد. همچنین مدل رگرسیون بردار پشتیبان در نرم‌افزار متلب ۲۰۱۷ کدنویسی شد.

$$k(x, x_j) = (t + x_i \cdot x_j)^d \quad (۴)$$

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (۵)$$

$$k(x, x_j) = x_i \cdot x_j \quad (۶)$$

۳-۲- تبدیل موجک^{۱۳}

تبدیل موجک به‌عنوان روشی جایگزین برای تبدیل فوری‌ی زمان کوتاه ارائه‌شده است و هدف از ارائه‌ی آن، غلبه بر مشکلات مربوط به قدرت تفکیک‌پذیری فرکانس در تبدیل فوری‌ی زمان کوتاه است. در تبدیل موجک همانند تبدیل فوری‌ی زمان کوتاه، سیگنال موردنظر به پنجره‌هایی تقسیم‌شده و تبدیل موجک بر روی هرکدام از این پنجره‌ها به‌صورت جداگانه انجام می‌گیرد (Vapnik, 1998). اما مهم‌ترین تفاوت آن‌ها در این است که در تبدیل موجک علاوه بر اینکه قدرت تفکیک فرکانس‌های یک سیگنال یا طول پنجره، متناسب با نوع فرکانس تغییر می‌کند، هم‌زمان عرض پنجره یا مقیاس فرکانس نیز متناسب با نوع فرکانس تغییر می‌کند. به‌عبارت‌دیگر، در تبدیل موجک به‌جای فرکانس، مقیاس وجود دارد. یعنی تبدیل موجک، نوعی تبدیل زمان - مقیاس است. بر همین اساس با استفاده از تبدیل موجک، در مقیاس‌های بالا سیگنال منبسط‌شده و جزئیات سیگنال قابل تجزیه و تحلیل است و در مقیاس‌های پایین سیگنال منقبض‌شده و کلیات سیگنال قابل بررسی می‌باشد (Wang et al, 2000). یک موجک به‌معنای موج کوچک، بخشی یا پنجره‌ای از سیگنال اصلی است که انرژی آن در زمان متمرکز شده است. با استفاده از تبدیل یا آنالیز موجک می‌توان یک سیگنال یا سری زمانی مادر را به موجک‌هایی با سطح تفکیک و مقیاس‌های مختلف تجزیه کرد. بنابراین موجک‌ها نمونه‌های انتقال‌یافته و تفکیک‌شده سیگنال مادر هستند که نوساناتی در یک طول متناهی داشته و شدیداً میرا هستند. بر اساس این ویژگی مهم تبدیل موجک، می‌توان سری‌های زمانی نامانا و گذرا را به‌صورت موضعی مورد تجزیه و تحلیل قرار داد (Shin et al, 2005). در این پژوهش از توابع مختلف موجک شامل هار، کلاه مکزیکی، مورلت استفاده شد. تابع موجک مورلت به دلیل آنکه مشتق دوم تابع گوسین می‌باشد از عملکرد بهتری برخوردار بود، انتخاب شد. تبدیل موجک مورلت با ارائه تحلیل زمان-فرکانس دقیق، قابلیت تنظیم فرکانس مرکزی، شکل مناسب برای تحلیل سیگنال‌های ارتعاشی، و امکان استخراج اطلاعات فاز، یک ابزار قدرتمند برای تحلیل سیگنال‌ها در حوزه‌های مختلف است که حتی می‌تواند نویز را از سیگنال حذف نماید. روش‌های حذف نویز مبتنی بر موجک معمولاً عملکرد بهتری نسبت به روش‌های سنتی دارند، زیرا می‌توانند نویز را بدون از دست دادن جزئیات مهم سیگنال حذف کنند.

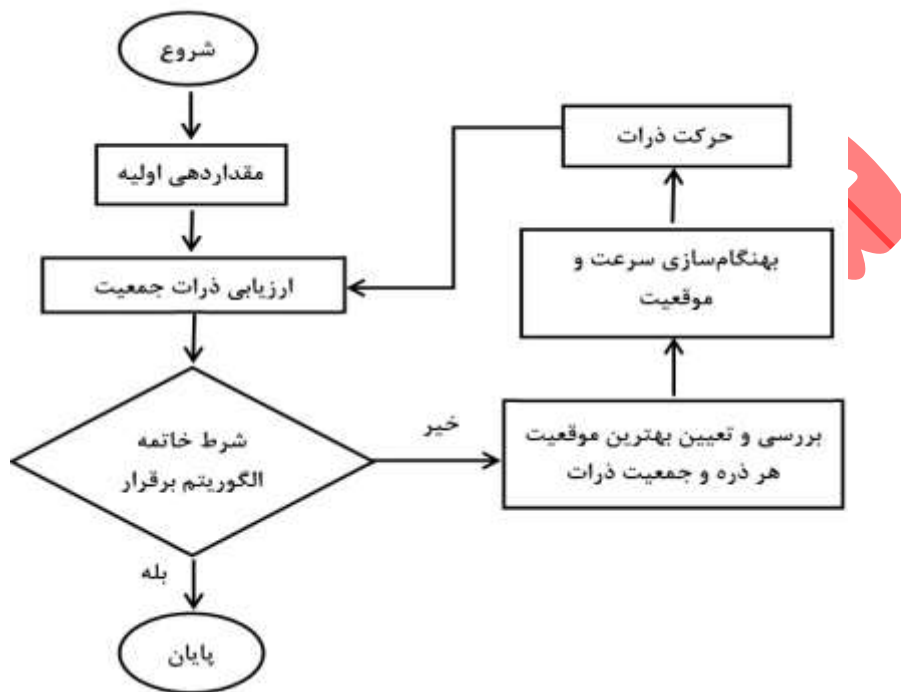
۴-۲- الگوریتم بهینه‌سازی ازدحام ذرات^{۱۴}

الگوریتم بهینه‌سازی ازدحام ذرات یک الگوریتم فرا ابتکاری است که اولین بار توسط ابره‌ه‌ارت و کندی معرفی شد (Eberhart and Kennedy, 1995). این الگوریتم از ماهیت رفتار پرندگان و ماهی‌های الهام گرفته است. به‌گونه‌ای که گروهی از پرندگان و ماهی‌ها مناسب‌ترین مسیر جهت رسیدن به لانه و غذا بدون ممانعت در حرکت ذرات دیگر می‌یابند. مراحل این الگوریتم در این پژوهش

¹³ Wavelet

¹⁴ Particle Swarm Optimization

به گونه ای است که ابتدا جمعیت اولیه تولید شده و بردارهای سرعت ذره در ابتدا صفر و بردار مکان به صورت تصادفی انتخاب می شود. در مرحله بعد، ارزیابی ارزش ذره صورت گرفته و سپس بهترین موقعیت فردی و سرعت ذره به روز می شود (Shrivatava et al., 2015). فلوچارت این الگوریتم در شکل ۳ قابل مشاهده است.



شکل ۳- فلوچارت الگوریتم ازدحام ذرات

Figure 3. Flowchart of the particle swarm algorithm

۲-۵- الگوریتم نهنگ^{۱۵}

الگوریتم بهینه سازی نهنگ از جمله الگوریتم های فرا ابتکاری بوده که از طبیعت و رفتار موجودات زنده الهام گرفته و در زمینه های مختلف مورد استفاده قرار می گیرد و اولین بار توسط میر جلیلی و لویز معرفی شده است (Mirjalili and Lewis, 2016). این الگوریتم از رفتار نهنگ هنگام شکار نشئت گرفته شده است به گونه ای که نهنگ محل شکار را شناسایی و آن را محاصره می کند. در این الگوریتم هدف، مناسب ترین راه حل شکار است. پس از آن که بهترین راه حل شکار را جستجو نمود عوامل دیگر سعی بر آن دارند تا مکان خود را نسبت به بهترین شکار به روزرسانی کنند (Reddy and Saha, 2022). رفتار این الگوریتم مطابق روابط ۷ و ۸ است.

$$\vec{D} = |\vec{C} \cdot \vec{X} - \vec{X}(t)| \quad (7)$$

$$\vec{X}(+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (8)$$

¹⁵ Whale Optimization Algorithm

که در آن A و C بردارهای ضرایب، X^* بردار مکان بهترین راه حل به دست آمده در حال حاضر و X بردار مکان است. بردارهای C و A به صورت روابط ۹ و ۱۰ محاسبه می شوند.

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (9)$$

$$\vec{C} = 2\vec{r} \quad (10)$$

در روابط بالا بازه a بین صفر تا دو به صورت خطی در هر تکرار کاهش می یابد و r نیز یک بردار تصادفی در بازه صفر تا یک است.

۲-۶- معیار ارزیابی

در این پژوهش به منظور ارزیابی مدل های برآورد رسوبات معلق رودخانه از شاخص های ارزیابی در روابط ۱۱ تا ۱۴ استفاده شد.

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad 1 \leq R \leq 1 \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (12)$$

$$MAE = \frac{\sum_{i=1}^N |x_i - y_i|}{n} \quad (13)$$

$$NS = 1 - \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{y})^2} \quad 0 \leq NS \leq 1 \quad (14)$$

در روابط ذکر شده، R ضریب همبستگی، RMSE ریشه میانگین مربعات خطا بر حسب ton/day، MAE میانگین قدر مطلق خطا بر حسب ton/day و NS معیار نش ساتکلیف، x_i و y_i به ترتیب مقادیر مشاهداتی و محاسباتی در گام زمانی t ام، N تعداد گام های زمانی، $\bar{x}$ و $\bar{y}$ نیز به ترتیب میانگین مقادیر مشاهداتی و محاسباتی می باشد. علاوه بر معیارهای فوق از نمودارهای پراکنش و سری زمانی مقادیر مشاهداتی-محاسباتی نسبت به زمان نیز برای مقایسه و تحلیل بیشتر استفاده شد.

۳. نتایج و بحث

در این پژوهش به منظور مدل سازی میزان رسوبات معلق رودخانه بابلرود واقع در استان مازندران از مدل رگرسیون بردار پشتیبان با الگوریتم های موجک، نهنگ و ازدحام ذرات استفاده شد. پارامتر دبی (Q) و رسوبات معلق رودخانه (SS) با تأخیرهای زمانی متفاوت در مقیاس زمانی ماهانه طی سال های ۱۳۸۲ تا ۱۴۰۲ استفاده شد. پارامترهای دبی با تأخیر زمانی $Q(t)$ ، $Q(t-1)$ و $Q(t-2)$ و پارامتر میزان رسوبات معلق $SS(t-1)$ و $SS(t-2)$ به عنوان ورودی مدل در نظر گرفته شد. قابل ذکر است تأخیر زمانی پارامترهای دبی و میزان رسوبات معلق رودخانه بر اساس ضریب همبستگی این پارامترها با خروجی مدل می باشد به گونه ای که تأخیر زمانی تا مرحله ای صورت پذیرفت که همبستگی معنی داری بین پارامترها وجود داشته باشد (Rajaei et al., 2011). در جدول ۱ ویژگی های آماری پارامتر استفاده شده، ارائه شده است. لازم به ذکر است برای مدل سازی ۸۰ درصد داده ها برای آموزش و ۲۰ درصد باقی مانده برای آزمون، به صورت تصادفی، که گستره وسیعی از انواع داده ها را پوشش دهد، انتخاب شد (Nagy et al., 2002; Kisi and Karhan, 2006). یکی از مهم ترین مراحل در مدل سازی، انتخاب ترکیب مناسبی از متغیرهای ورودی است. در مدل های هوشمند

انتخاب ورودی‌های اولیه مناسب و تأثیرگذار در پدیده به‌منظور آموزش ماهیت سازوکار حاکم بر پدیده باعث بهبود عملکرد خواهد شد بنابراین در سناریوسازی میزان رسوبات نیز سعی گردید تا مؤثرترین داده‌های مشاهداتی به‌عنوان داده‌های آموزشی انتخاب شود (Dehghani et al., 2020). بدین‌منظور ترکیب‌های مختلفی از پارامتر ورودی به‌منظور دستیابی به مدل بهینه برای تخمین دبی رسوبی استفاده شد که در جدول ۲ آمده است.

جدول ۱- محدوده پارامتر مورد استفاده برای آموزش و صحت سنجی داده‌ها

Table 1. The used parameter range for training and validation of data

پارامتر Parameters	آموزش Training		آزمون Testing	
	میانگین Mean	مینیمم Minimum	میانگین Mean	ماکزیمم Maximum
دبی (m ³ /s) Discharge	5.128	0.070	2.350	9.459
رسوب (ton/day) Sediment	146.722	0.717	37.678	431.211

جدول ۲- ترکیب‌های منتخب پارامتر ورودی مدل‌های هیبریدی مورد بررسی

Table 2. Selected combinations of input parameters in the investigated models

شماره Number	ورودی Input	خروجی Output
1	Q(t)	SS (t)
2	Q(t), Q (t-1)	SS (t)
3	Q(t), Q (t-1), Q (t-2)	SS(t)
4	Q(t), Q (t-1), Q (t-2), SS (t-1)	SS (t)
5	Q(t), Q (t-1), Q (t-2), SS (t-1), SS (t-2)	SS (t)

به‌منظور مدل‌سازی میزان رسوبات معلق رودخانه بابلرود از مدل رگرسیون بردار پشتیبان با الگوریتم‌های موجک، نهنگ و ازدحام ذرات استفاده شد. همان‌طور که در جدول ۳ مشخص است مدل‌های هیبریدی در سناریو شماره ۵ که شامل کلیه پارامترهای ورودی به مدل می‌باشد از عملکرد بهتری نسبت به سایر سناریوها برخوردار است. بنابراین افزایش تعداد پارامترهای مؤثر در مدل‌های هیبریدی مبتنی بر شبکه عصبی مصنوعی منجر به افزایش عملکرد مدل می‌شود. همچنین کلیه مدل‌ها در تابع کرنل پایه شعاعی از دقت بهتری برخوردار هستند که نتایج مدل‌ها طبق سناریوهای ترکیبی در جدول ۳ با کرنل تابع پایه شعاعی نشان داده شد. همان‌طور که در جدول ۳ مشخص است مدل رگرسیون بردار پشتیبان-موجک در سناریو ترکیبی شماره ۵ با بیشترین ضریب همبستگی ۰/۹۶۲، کمترین ریشه میانگین مربعات (ton/day) ۰/۳۴۴، کمترین میانگین قدر مطلق خطا (ton/day) ۰/۱۵۸ و بیشترین ضریب نش ساتکلیف ۰/۹۷۰ در مرحله صحت سنجی عملکرد بهتری از خود نشان داده است.

در شکل ۴ نمودار سری زمانی مقادیر مشاهداتی و محاسباتی نشان داده شده است همان‌طور که مشاهده می‌گردد مدل رگرسیون بردار پشتیبان-موجک نسبت به مدل‌های هیبریدی رگرسیون بردار پشتیبان- نهنگ و ازدحام ذرات در تخمین اکثر نقاط از جمله کمینه، بیشینه و میانی دقت قابل قبولی از خود نشان داده است همچنین مدل‌ها رگرسیون بردار پشتیبان- نهنگ در تخمین مقادیر میانی عملکرد نسبتاً مطلوبی داشته است و در تخمین مقادیر بیشینه ضعیف عمل نموده است.

در شکل ۵ نمودار جعبه‌ای مدل‌های مورد بررسی نشان داده شد همان‌طور که مشاهده می‌گردد مدل رگرسیون بردار پشتیبان-موجک در تخمین مقادیر چارک اول و میانه نسبت به داده‌های مشاهداتی عملکرد بهتری از خود نشان داده است در صورتی که مدل رگرسیون بردار پشتیبان -ازدحام ذرات ضعیف عمل نموده است و مدل رگرسیون بردار پشتیبان- نهنگ تاب نیز از دقت

مطلوبی برخوردار است و در رتبه دوم می‌باشد. همان‌طور که در شکل ۵ مشخص است کلیه مدل‌ها در برآورد مقادیر مینیمم، چارک اول و میانه عملکرد مطلوبی دارند اما در تخمین مقادیر بیشینه مدل رگرسیون بردار پشتیبان-موجک عملکرد بهتری از خود نشان داده است.

در شکل ۶ نمودار تیلور مدل‌های موردبررسی قابل‌مشاهده است مدل رگرسیون بردار پشتیبان-موجک از عملکرد بهتری برخوردار است زیرا انحراف معیار پیش‌بینی‌شده میزان رسوبات معلق رودخانه نزدیک‌ترین فاصله به انحراف استاندارد داده‌های مشاهداتی را دارد و ضریب همبستگی نیز بالاترین میزان را نشان می‌دهد. همان‌طور که در شکل ۶ مشخص است مدل‌های موردبررسی میزان انحراف معیار نزدیکی نسبت به یکدیگر دارند اما دقت مدل رگرسیون بردار پشتیبان-موجک نسبت به سایر مدل‌ها بیشتر است. بنابراین مدل رگرسیون بردار پشتیبان-موجک نسبت به سایر مدل‌های موردبررسی از عملکرد بهتری برخوردار است که این نتایج با پژوهش‌های مالیک و همکاران (Malik et al., 2021)، هو و همکاران (Hou et al., 2023) و دهقانی و زیدعلی نژاد (Zeidalinejad and Dehghani, 2023) مطابقت دارد. در تحلیل این نتایج می‌توان بیان داشت برتری این مدل ناشی از تبدیل موجک می‌باشد که سیگنال‌های دریافتی را به دودسته بالا گذر و پایین گذر تقسیم نموده و در دسته بالا گذر قدرت تفکیک افزایش‌یافته که سبب می‌گردد مقادیر بیشینه سیگنال با دقت مطلوبی تجزیه و تحلیل گردد. همچنین، در تبدیل موجک میزان نویز شبکه را از بین برده و سبب افزایش دقت مدل می‌گردد.

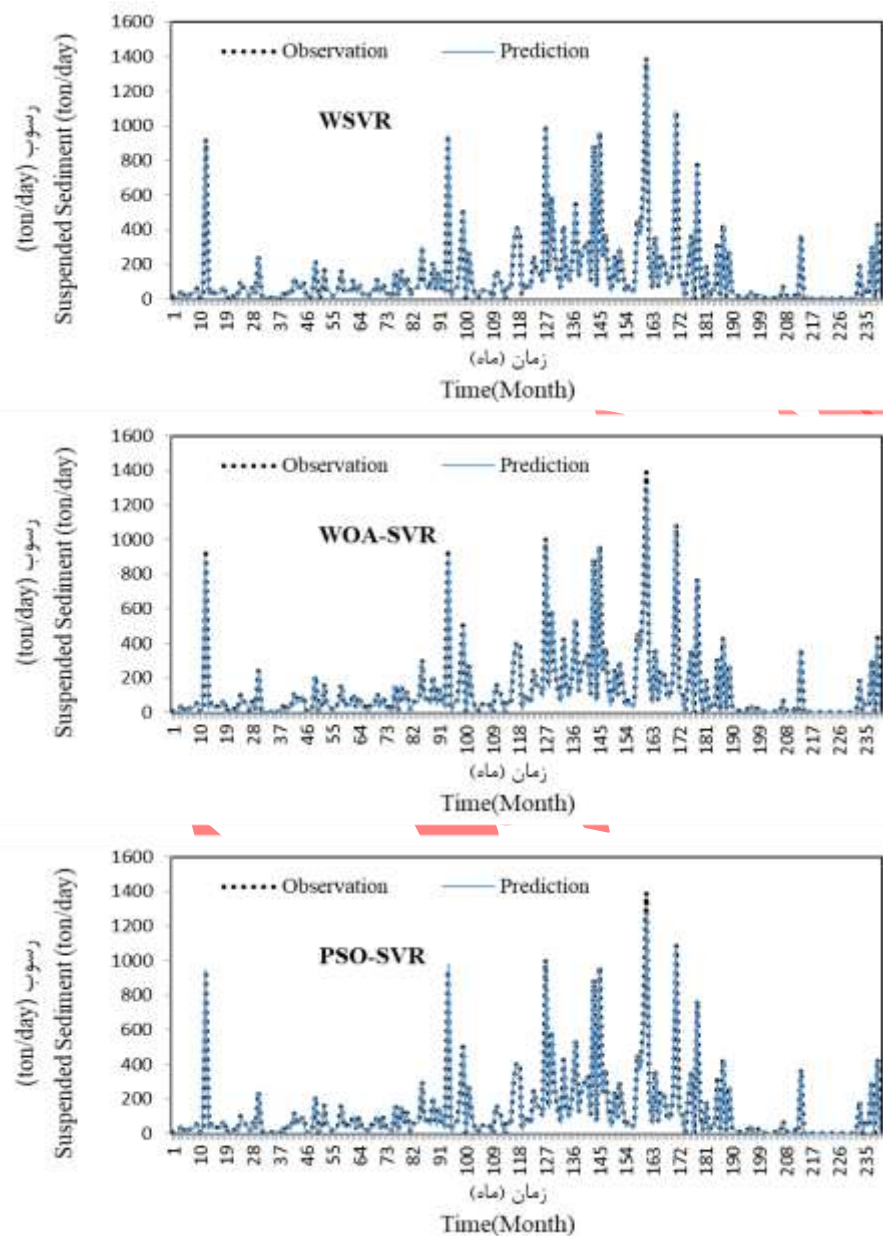
مدل رگرسیون بردار پشتیبان-الگوریتم نهنگ ترکیبی از بهینه‌سازی پیوسته و گسسته است که زمان رسیدن به یک راه‌حل بهینه را در یک منطقه جستجوی وسیع کاهش می‌دهد زیرا از راه‌حل‌های بهینه محلی اجتناب می‌کند. این امر باعث می‌شود که الگوریتم برای حل مسائل غیرخطی با ابعاد بزرگ با سرعت مناسب در همگرایی به سمت یک جواب بهینه قابل قبول باشد. که این امر سبب گردیده این مدل از دقت بالایی نسبت به سایر مدل‌ها برخوردار باشد.

به‌طورکلی پیشنهاد می‌شود از مدل هیبریدی رگرسیون بردار پشتیبان-موجک به‌عنوان مدلی با خطای ناچیز برای حل مسائل غیرخطی با ابعاد بزرگ با سرعت مناسب در همگرایی به سمت یک جواب بهینه استفاده شود. همچنین می‌توان به‌عنوان راهکاری نوین در پیش‌بینی میزان رسوبات معلق رودخانه به‌منظور اتخاذ تصمیمات مدیریتی مناسب برای بهبود منابع آبی، آماده‌سازی زمین و سرمایه‌گذاری اقتصادی، تولید محصولات آبرزی دانست.

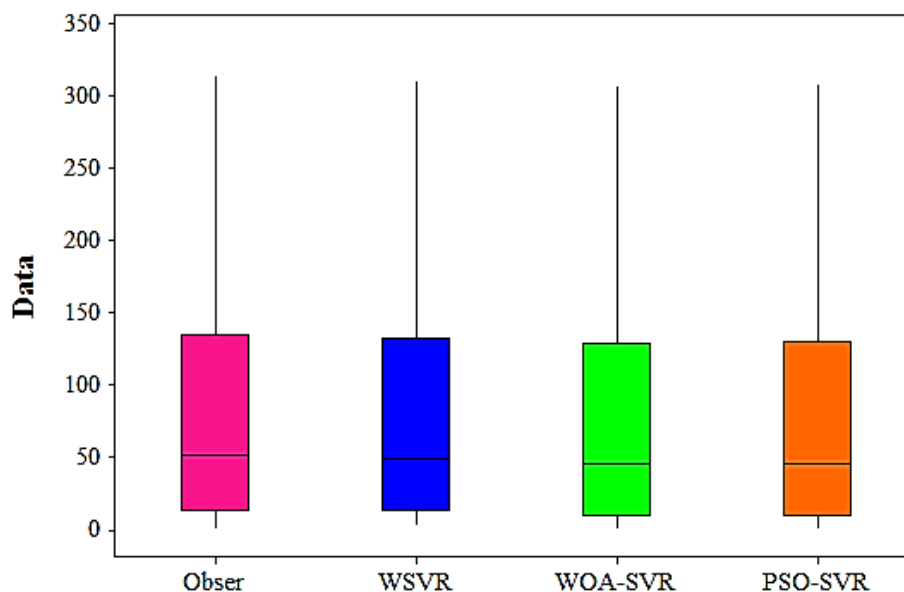
جدول ۳- عملکرد مدل‌های موردبررسی در برآورد سطح آب زیرزمینی

Table 3. Performance of the investigated models in estimating groundwater level

آزمون Testing				آموزش Training				سناریو Senario	مدل Model
NS	MAE (ton/day)	RMSE (ton/day)	R	NS	MAE (ton/day)	RMSE (ton/day)	R		
0.925	0.211	0.401	0.917	0.910	0.298	0.532	0.900	1	WSVR
0.935	0.195	0.388	0.931	0.925	0.277	0.511	0.920	2	
0.950	0.178	0.369	0.942	0.940	0.252	0.486	0.936	3	
0.960	0.166	0.357	0.955	0.950	0.241	0.477	0.940	4	
0.970	0.158	0.344	0.962	0.955	0.234	0.463	0.945	5	
0.910	0.229	0.431	0.904	0.900	0.324	0.547	0.893	1	WOA-SVR
0.920	0.218	0.417	0.910	0.905	0.307	0.532	0.900	2	
0.930	0.208	0.405	0.925	0.915	0.287	0.510	0.908	3	
0.945	0.195	0.392	0.940	0.920	0.266	0.497	0.914	4	
0.950	0.180	0.377	0.947	0.935	0.252	0.487	0.920	5	
0.880	0.317	0.574	0.877	0.870	0.342	0.637	0.850	1	PSO-SVR
0.890	0.304	0.556	0.886	0.880	0.321	0.625	0.864	2	
0.900	0.293	0.532	0.895	0.890	0.307	0.614	0.886	3	
0.910	0.274	0.518	0.902	0.900	0.292	0.592	0.891	4	
0.937	0.258	0.492	0.925	0.910	0.278	0.573	0.900	5	

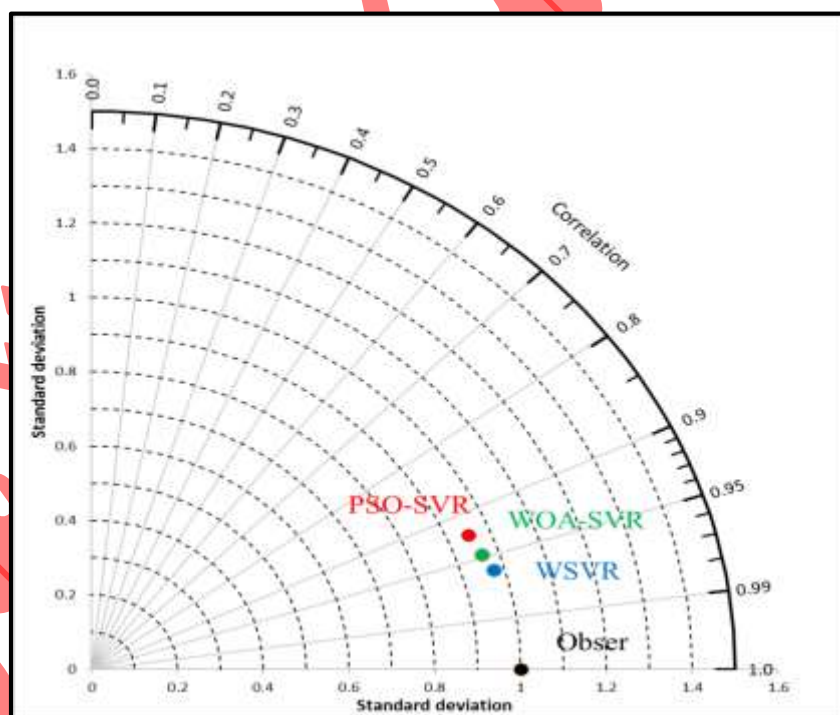


شکل ۴- نمودار سری زمانی مدل های مورد بررسی
Figure 4. Time series graph of the models under the investigation



شکل ۵- نمودار باکس پلات مدل‌های مورد بررسی

Figure 5. Box plot diagram of the investigated models



شکل ۶- نمودار تیلور مدل‌های مورد بررسی

Figure 6. Taylor diagram of the investigated models

۴- نتیجه‌گیری

در پژوهش حاضر مطالعه موردی جهت ارزیابی عملکرد مدل فرا ابتکاری هیبریدی رگرسیون بردار پشتیبان با الگوریتم‌های موجک، نهنگ و ازدحام ذرات به‌منظور برآورد میزان رسوبات معلق رودخانه بابلرود واقع در استان مازندران صورت گرفت. جهت مدل‌سازی

از پارامترهای دبی جریان (Q) و رسوبات معلق (SS) طی گام‌های زمانی مختلف به‌عنوان ورودی و میزان رسوبات معلق رودخانه (SS) به‌عنوان خروجی مدل استفاده شد. جهت ارزیابی مدل‌های موردبررسی از شاخص‌های آماری ضریب همبستگی، ریشه میانگین مربعات خطا، میانگین قدر مطلق خطا و ضریب نش سائکلیف استفاده شد همچنین جهت تحلیل نتایج از نمودارهای سری زمانی، باکس پلات و تیلور استفاده گردید. نتایج تحقیق طبق ارزیابی سناریوهایی متشکل از پارامترهای ورودی نشان داد که در کلیه مدل‌های موردبررسی افزایش تعداد پارامترهای مؤثر در مدل‌های مختلف مدل‌سازی منجر به عملکرد بهتر در تخمین میزان رسوبات رودخانه می‌شود. علاوه بر این، نتایج حاصل از معیارهای ارزیابی نشان داد مدل رگرسیون بردار پشتیبان-موجک از دقت بالا و خطای ناچیزی برخوردار است. همچنین مطابق نمودارهای موردبررسی مدل رگرسیون بردار پشتیبان-موجک مقادیر رسوبات معلق رودخانه را نزدیک به مقدار واقعی‌شان برآورد نموده است که در اشکال باکس پلات و تیلور مشهود می‌باشد. درمجموع، نتایج این تحقیق نشان می‌دهد که استفاده از مدل‌های هوش مصنوعی مبتنی بر رویکرد مدل رگرسیون بردار پشتیبان می‌تواند در زمینه تخمین میزان دبی رسوبی طی ۳۰ سال آماری برای سایر مناطق کشور و گامی در جهت اتخاذ تصمیمات مدیریتی مناسب مورد استفاده قرار گیرد.

منابع

- Cimen, M. (2008). Estimation of daily suspended sediments using support vector machines. *Hydrological Sciences Journal*, 53 (3), 656-666. [10.1623/hysj.53.3.656](https://doi.org/10.1623/hysj.53.3.656)
- Dehghani, R., Babaali, H. (2023). Evaluation of Statistical Models and Modern Hybrid Artificial Intelligence in Simulation of Runoff Precipitation Process. *Sustain. Water Resour. Manag.*, 8, 154-172. <https://doi.org/10.1007/s40899-022-00743-9>.
- Dehghani, R., Torabi, H. (2021). Dissolved oxygen concentration predictions for running waters using hybrid machine learning techniques. *Modeling Earth Systems and Environment*, 6(2), 64-78. <https://doi.org/10.1007/s40808-021-01253-x>.
- Doroudi, S., Sharafati, A., Mohajeri, H. (2021). Estimation of Daily Suspended Sediment Load Using a Novel Hybrid Support Vector Regression Model Incorporated with Observer-Teacher-Learner-Based Optimization Method. *Complexity*, 9(3), 532-545. <https://doi.org/10.1155/2021/5540284>
- Eberhart R., Kennedy J. (1995). A New Optimizer Using Particle Swarm Theory Proc. Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, Piscataway, NJ: IEEE Service Center, 39-43. [10.1109/MHS.1995.494215](https://doi.org/10.1109/MHS.1995.494215)
- Essam, Y., Huang, Y., Birima, A., Ahmed, A., El-Shafie, A. (2022). Predicting suspended sediment load in Peninsular Malaysia using support vector machine and deep learning algorithms. *scientific reports*, 12(2), 344-357. <https://doi.org/10.1038/s41598-021-04419-w>.
- Goyal, M. K. (2014). Modeling of sediment yield prediction using the M5 model tree algorithm and wavelet regression. *Water Resources Management*, 28(7), 1991-2003. <https://doi.org/10.1007/s11269-014-0590-6>
- Hassanpour, F., Sharifazari, S., Ahmadaali, K., Mohammadi, S., Sheikhalipour, Z. (2019). Development of the FCM-SVR Hybrid Model for Estimating the Suspended Sediment Load. *KSCE Journal of Civil Engineering* 23(6), 2514-2523. <https://doi.org/10.1007/s12205-019-1693-7>.
- Hou, W., Yin, G., Gu, J., Ma, N. (2023). Estimation of Spring Maize Evapotranspiration in Semi-Arid Regions of Northeast China Using Machine Learning: An Improved SVR Model Based on PSO and RF Algorithms. *Water*, 15(8), 558-568. <https://doi.org/10.3390/w15081503>
- Kisi, O., Dailr, A. H., Cimen, M., Shiri, J. (2012). Suspended sediment modeling using genetic programming and soft computing techniques, *Journal of Hydrology*, 450(3), 48-58. <https://doi.org/10.1016/j.jhydrol.2012.05.031>
- Malik, A., Tikhmarine, Y., Al-Ansari, N., Shahid, S., Sekhon, H.S., Pal, R., Rai, R., Pandey, K., Singh, P., Elbeltagi, A., Sammen, S. (2021). Daily pan-evaporation estimation in different agro-climatic zones using novel hybrid support vector regression optimized by Salp swarm algorithm in conjunction with gamma test. *Engineering Applications of Computational Fluid Mechanics*, 15(1), 1075-1094. <https://doi.org/10.1080/19942060.2021.1942990>
- Mirjalili, S., and Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software*, 95(6), 51-67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>.

- Nayak, P., Venkatesh, B., Krishna, B., and Jain, S. K. (2013). Rainfall- runoff modeling using a conceptual, data-driven, and wavelet-based computing approach. *Journal of Hydrology*, 493(6), 57-67. <https://doi.org/10.1016/j.jhydrol.2013.04.016>
- Nourani, V., Gokcekus, H., Gelete, G. (2021). Estimation of Suspended Sediment Load Using Artificial Intelligence-Based Ensemble Model. *Complexity*, 8(4), 122-136. 10.1155/2021/6633760
- Rajae, T., Nourani, V., Zounemat-Kermani, M., Kisi, O. (2011). River suspended sediment load prediction: application of ANN and wavelet conjunction model. *J Hydrol Eng*, 16(2), 613-627. 10.1061/(ASCE)HE.1943-5584.0000347
- Reddy, K., Saha, A.K. (2022). A modified Whale Optimization Algorithm for exploitation capability and stability enhancement. *Heliyon*, 8(10), 425-441. <https://doi.org/10.1016/j.heliyon.2022.e11027>
- Shin, S., Kyung, D., Lee, S., Taik, & Kim, J., and Hyun, J. (2005). An application of support vector machines in a bankruptcy prediction model, *Expert Systems with Applications*, 28(4), 127-135. <https://doi.org/10.1016/j.eswa.2004.08.009>
- Shrivatava, M., Prasad, V., Khare, R. (2015). Multi-objective optimization of water distribution system using particle swarm optimization. *IOSR J. Mech. Civ. Eng*, 12(1), 21-28. <https://doi.org/10.5004/dwt.2021.26944>
- Vapnik, V.N. (1995). The Nature of Statistical Learning Theory. *Springer, New York*. <https://doi.org/10.1007/978-1-4757-3264-1>
- Vapnik, V.N. (1998). Statistical learning theory. *Wiley, New York*. https://doi.org/10.1007/978-3-540-28650-9_8
- Wang, D., Safavi, A.A., and Romagnoli, J.A. (2000). Wavelet-based adaptive robust M-estimator for non-linear system identification, *AIChE Journal*, 46(4), 1607-1615. <https://doi.org/10.1002/aic.690460812>
- Wu, C. L., Chau, K. W. (2020). Rainfall-runoff modeling using an artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology*, 399(3), 394-409. <https://doi.org/10.1016/j.jhydrol.2011.01.017>