

Research Article

Modeling Estimation of River Suspended Sediment Amount Using Support Vector Regression and the Group Method of Data Control

A. Moradinejad

Received: December 17, 2022

Accepted: June 13, 2023

Revised: June 2, 2023

Published online: June 21, 2024

Assist. Prof., Soil Conservation and Watershed Management Research Department, Markazi Agricultural and Natural Resources Research and Education Center, Arak, Agricultural Research Education & Extension Organization (AREEO), Arak, Iran.

Corresponding Author, Email: Amir_24619@yahoo.com

Background and Objectives

Erosion and deposition cause irreparable damage to water construction projects. Among these damages, we can mention the accumulation of sediment behind the dams and the reduction of their useful volume, the destruction of structures, the reduction of capacity and the increase in the maintenance cost of irrigation channels. On the other hand, sediment transport affects water quality indicators in terms of drinking and agriculture. Accurate estimation of river-suspended sediment load from various aspects of water resources engineering, environmental issues and water quality is of particular importance. In this regard, the hydrological models of the basin have not shown proper efficiency in estimating the amount of suspended sediment, due to the many influencing factors. Most of the simulation studies of suspended sediment estimation are based only on the flow rate of the basin outlet, and the obtained results are proof of their inefficiency. In the meantime, the development and use of new sediment estimation methods that are easy to use in addition to sufficient accuracy will play an important role. Nowadays, the fuzzy and neural intelligent conductor system has found many applications in various water engineering problems, including sedimentation, due to its ability to solve complex and nonlinear phenomena. Due to the great importance of sediment transport in the optimal use of water resources and the design of dams, it is very necessary to obtain an accurate method for estimating the amount of suspended sediment load in rivers.

Methodology

The purpose of this research is to evaluate and compare the two methods of support vector machine models (SVR) and group data control method (GMDH) for estimating the suspended sediment load of the Pol Doab station of Qarachai River in Markazi Province and comparing it with the results of the measuring curve. For this purpose, the daily data of discharge, sediment, temperature, and rainfall parameters of Shazand Pol Doab station were used. For this work, 13 scenarios and different combinations of parameters were defined. Then the results of the two methods were compared with each other and with the results of the measuring curve. Finally, the best method was suggested. For this purpose, library, and field studies and a review of related sources, statistics and information were collected. The statistics of temperature, rainfall, and daily average discharge of stream and sediment measured daily during a long-term statistical period of 40 years at Pol Doab Hydrometry Station were received from the Meteorology and Regional Water Department of Markazi Province. The received data were categorized and converted into the input format of the models. According to the discharge and corresponding sediment, the curve of the sediment gauge was drawn and its equation was obtained. Appropriate patterns of input variables were selected based on trial and error. Considering that the mentioned parameters have a historical course, therefore, the design of the input patterns of soft computing models should be done based on time delays (like what is discussed in the analysis and forecasting of time series). Then the model was taken for each input and output pattern. In the next step, the most appropriate time delay of the input parameters in the modeling, which had a higher R^2 determination coefficient and a lower root mean square error (RMSE) was selected. In this research, 70% of the research data was used as training and 30% for validation and testing. Finally, two data-mining methods were compared with each other and also with sediment rating curve (SRC) and observational data.

Findings



The results obtained from this research indicated the acceptable performance of the methods used in predicting suspended sediment amounts. Comparing the results of GMDH, SVR, and SRC models shows the superiority of GMDH and SVR models in predicting suspended sediment values compared to Verdi model number 6. The results showed the acceptable performance of the GMDH model with the highest R^2 determination coefficient of 0.99 and the lowest root mean square error of 83 tons per day. According to the obtained results, it can be said that the GMDH model as a powerful and high-speed model can be used to

model suspended sediment. The results of the research showed that both data-mining methods have far better efficiency and accuracy in estimating the suspended load of river sediment than the sediment gauge curve. Data mining-based methods can be used as an alternative to estimate the river's suspended load. It should be noted that due to climate change and droughts, industrial development, colonization of land use, and changes in the morphology of watersheds, the obtained results cannot be used forever at any time, but should be used whenever the models need to be updated. Another weakness of the models is that with the increase in the number of developed layers, the accuracy of the produced answers increases, but the produced relationships between the input and output variables become very complicated.

Conclusion

The results of modeling with (SVR) showed that scenario number 6, which includes discharge at the current time and discharge and sediment with a time delay step, with the highest determination coefficient R^2 with a value of 0.98 and the lowest root mean square error RMSE equal to 185 tons per day. It performs better in other scenarios. In the next step, the best model selection model (SVR) was used as the input of the GMDH model. The results were compared with the model (SVR) and sediment gauge curve. The results show the acceptable performance of the GMDH model with the highest R^2 determination coefficient of 0.99 and the lowest root mean square error of 83 tons per day compared to the other two methods. The obtained results showed that both investigated data mining methods provide much better results than the sediment gauge curve. The coefficients and functions used to calibrate the intelligent models used in this research can be very useful for estimating the suspended sediments of nearby stations without statistics having the same geological and hydrological conditions at the regional level.

Keywords: Modeling, GMDH, PolDoab, Suspended sediment, SVR.

مقاله پژوهشی

مدل‌سازی برآورد میزان رسوب معلق رودخانه با استفاده از رگرسیون بردار پشتیبان

و روش گروهی کنترل داده‌ها

امیر مرادی نژاد

تاریخ دریافت: ۱۴۰۱/۰۹/۲۶

تاریخ پذیرش: ۱۴۰۲/۰۳/۲۳

تاریخ ویرایش: ۱۴۰۲/۰۳/۱۲

تاریخ انتشار آنلاین: ۱۴۰۳/۰۴/۰۱

استادیار مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان مرکزی

*مسئول مکاتبات، پست الکترونیکی: amir_24619@yahoo.com

چکیده

انتقال رسوب از فرآیندهای مهم هیدرودینامیکی است که همواره بسیاری از سازه‌ها و تأسیسات عمرانی حاشیه رودخانه‌ها را تحت تأثیر قرار می‌دهد. انتقال رسوب و رسوبگذاری یکی از بزرگترین مشکلات بهره‌برداری از منابع آبی محسوب می‌گردد. به سبب اهمیت فراوان انتقال رسوب در استفاده بهینه از منابع آبی و طراحی سدها، دستیابی به روشی با دقت مناسب برای تخمین میزان بار رسوبی معلق رودخانه‌ها بسیار ضروری است. هدف از این پژوهش ارزیابی و مقایسه دو روش مدل‌های رگرسیون بردار پشتیبان (SVR) و روش گروهی کنترل داده‌ها (GMDH) در برآورد بار رسوب معلق ایستگاه پل دوآب رودخانه قره‌چای استان مرکزی و مقایسه با نتایج منحنی سنجی می‌باشد. برای این کار از داده‌های روزانه پارامترهای دبی، رسوب، دما و بارندگی ایستگاه پل دوآب شازند استفاده شد. برای این کار ۱۳ سناریو و ترکیب مختلف از پارامترها تعریف شد. نتایج مدل‌سازی با (SVR) نشان داد که سناریوی شماره ۶ که شامل دبی در زمان فعلی و دبی و رسوب با یک گام تأخیر زمانی با بیشترین ضریب تعیین R^2 با مقدار ۰/۹۸ و کمترین ریشه میانگین مربعات خطا RMSE برابر ۱۸۵ کیلوگرم در روز نسبت به سناریوهای دیگر عملکرد بهتری دارد. در مرحله بعد از بهترین الگوی انتخابی مدل (SVR) به‌عنوان ورودی مدل GMDH استفاده شد. نتایج با مدل (SVR) و منحنی سنجی رسوب مقایسه شد. نتایج بیانگر عملکرد قابل قبول مدل GMDH با بیشترین ضریب تعیین R^2 برابر ۰/۹۹ و کمترین ریشه میانگین مربعات خطا برابر ۸۳ کیلوگرم در روز نسبت به دو روش دیگر است. نتایج به‌دست آمده نشان داد هر دو روش داده‌کاوی بررسی شده به‌مراتب نتایج بهتری نسبت به منحنی سنجی رسوب ارائه می‌کنند. ضرایب و توابع به کار رفته برای واسنجی مدل‌های هوشمند مورد استفاده در این پژوهش می‌تواند برای تخمین رسوبات معلق ایستگاه‌های مجاور فاقد آمار دارای شرایط زمین‌ساختی و هیدرولوژیکی یکسان در سطح منطقه بسیار مفید واقع شود.

واژه‌های کلیدی: پل دوآب، رسوب معلق، مدل‌سازی، GMDH، SVR.

مقدمه

برآورد دقیق میزان انتقال بار رسوبات حمل شده از طریق رودخانه در بسیاری از پروژه‌های مهندسی منابع آب از قبیل کانال‌های کشتیرانی، اجرای مخازن سدها، طول عمر تجهیزات هیدروالکتریکی، مسائل زیست محیطی و سایر موضوعات مرتبط مورد نیاز است (کیشی و همکاران ۲۰۱۲). معادلات زیادی بر پایه روش‌های مختلف برای برآورد نرخ انتقال رسوب در رودخانه‌ها توسعه داده شده است. روش‌های متداول که در بسیاری از مدل‌سازی‌ها استفاده شده‌اند غالباً با یک شکل ساده شده از توابع تجربی یا تحلیلی آغاز می‌شوند و با یک تحلیل رگرسیونی یا مشخص کردن منحنی سنج مناسب با استفاده از داده‌های آزمایشگاهی برای تعیین ضرایب مجهول مدل دنبال می‌شوند (نکریا و همکاران ۲۰۱۰). تقریباً همه معادلات تجربی برآورد بار رسوب برای شرایط ساده جریان-های یکنواخت دائمی و تعادل انتقال رسوب به دست آمده‌اند که این شرایط در رودخانه‌های طبیعی خیلی کم اتفاق می‌افتد. بنابراین، دبی رسوب برآورد شده با توجه به فرضیات ساده کننده مذکور دقیق نمی‌باشد (حزروی و شفاعی بجزستانی ۲۰۱۰). برای شبیه‌سازی پدیده‌هایی از این دست، می‌توان از روش‌های داده محور که زیرمجموعه‌ای از روش‌های محاسبات نرم هستند و نیازی به روابط پیچیده حاکم بر پدیده ندارند، استفاده نمود. اجزای اصلی محاسبات نرم شامل منطق فازی، شبکه‌های عصبی و الگوریتم‌های تکامل گرا می‌باشند که توانایی‌های زیادی در شناسائی سیستم‌های غیرخطی و مسائل کنترلی دارند (ایوانکو ۱۹۶۸). جیغیزاوغلو و کیشی (۲۰۰۶) با استفاده از داده‌های دبی جریان و بار معلق به شبیه‌سازی انتقال رسوب با استفاده از شبکه عصبی مصنوعی پرداختند. در مطالعه دیگری ژو و همکاران (۲۰۰۷) با استفاده از شبکه عصبی انتقال رسوب معلق را شبیه‌سازی نمودند. نتایج این تحقیق نشان دهنده قابلیت خوب مدل‌های توسعه داده شده در

پیش‌بینی رسوب است. آلپ و جیغیزاوغلو (۲۰۰۷) در مطالعه خود با استفاده از شبکه عصبی چندلایه و همچنین شبکه عصبی پایه شعاعی به پیش‌بینی میزان بار رسوب معلق و مقایسه آنها با روش رگرسیون خطی پرداختند. نتایج نشان‌دهنده برتری شبکه عصبی در تخمین میزان بار رسوبی بوده است. جی و یو (۲۰۱۱) با استفاده از روش رگرسیون ماشین بردار پشتیبان و شبکه عصبی به شبیه‌سازی میزان انتقال رسوب در حوزه رودخانه کوپینگ در تایوان اقدام نمودند. چن و چاو (۲۰۱۶) با به کارگیری دو نوع از شبکه‌های عصبی میزان رسوب معلق را مورد مطالعه قرار دادند. نتایج این تحقیق نشان‌دهنده قابلیت خوب این مدل‌ها در یادگیری روابط بین پارامترهای ورودی و خروجی است. مرادی نژاد و همکاران (۲۰۱۸) در مطالعه‌ای به بررسی کارایی روش‌های برآورد بار رسوب معلق رودخانه قره‌چای پرداختند. آنها از روش-های مدل‌های شبکه عصبی مصنوعی، نروفازی، منحنی سنج رسوب و همچنین مدل رگرسیونی چند متغیره استفاده کردند. نتایج نشان داد که روش نروفازی براساس ورودی‌های دبی و رسوب و مدل‌های شبکه عصبی مصنوعی براساس ورودی‌های دبی، نسبت به رگرسیون چندمتغیره و منحنی سنج رسوب، دارای دقت مطلوب‌تری می‌باشند. ستاری و همکاران (۲۰۱۶) به ارزیابی عملکرد روش‌های مدل درختی M5 و رگرسیون بردار پشتیبان در مدل‌سازی رسوب معلق رودخانه اهرچای در مقایسه با روش کلاسیک منحنی سنج رسوب پرداختند. برای این کار از داده‌های ماهانه رسوب، دبی و بارندگی استفاده کردند. جهت ارزیابی عملکرد روش‌های استفاده شده از سه آماره شامل ضریب همبستگی، جذر میانگین مربعات خطا و میانگین خطای مطلق استفاده گردید. نتایج به دست آمده بیانگر دقت بسیار بالای روش‌های داده‌کاوی نسبت به منحنی سنج رسوب می‌باشد. اگرچه هر دو روش داده‌کاوی بررسی شده، دقت بیشتر و خطای کم‌تری نسبت به

معیارهای ارزیابی مدل GP با دو کرنل PUK و RBF در دوره کم‌آبی و پرآبی عملکرد بهتری را نسبت به سایر مدل‌ها (RF, RepTree, M5P) داشته است. همچنین با توجه به نتایج بخش آزمایش مدل GP-PUK بهترین نتیجه را داشته است. کشتکار و همکاران (۲۰۲۳) در تحقیقی از روش‌های نرم برای تخمین رسوب معلق رودخانه در پاکستان پرداختند. آنها نتایج مدل‌های RM5Tree با مدل‌های رگرسیون بردار پشتیبان (SVR)، شبکه عصبی مصنوعی (ANN)، اسپلاین رگرسیون تطبیقی چند متغیره (MARS)، منحنی سنجه رسوب (SRC) و روش سطح پاسخ (RSM) مقایسه کردند. از پارامترهای دبی جریان، پوشش برف، دمای هوا، تبخیر و تعرق و بارندگی مؤثر استفاده کردند. دقت حاصل از مدل‌ها با استفاده از ضریب همبستگی پیرسون (R^2)، ریشه میانگین مربع خطا (RMSE) و میانگین درصد مطلق خطا (MAPE) ارزیابی شد. نتایج نشان داد که مدل RM5Tree پیک‌های رسوب معلق را با دقت نسبی ۸۴/۱۰ درصد در مقایسه با مدل‌های MARS، ANN، SVR، M5Tree، RSM و SRC به ترتیب با دقت ۸۰/۶۲، ۷۷/۸۶، ۸۱/۹۰، ۸۰/۵۸، ۸۰/۵۸ و ۷۶ درصد بهتر پیش‌بینی کرده است. تهنشین شدن رسوبات معلق در یک رودخانه همچنین باعث کاهش سطح مقطع رودخانه و تغییر شکل پلان رودخانه و در نتیجه کاهش زیستگاه آبزیان رودخانه می‌شود (ادنان و همکاران ۲۰۲۲). مرادی‌نژاد و همکاران (۲۰۲۳) در تحقیقی به مدل‌سازی مبتنی بر هوش محاسباتی در برآورد بار معلق رودخانه ایستگاه باقرآباد قمرود پرداختند. عملکرد سه نوع مدل رگرسیون ماشین بردار پشتیبان (SVR)، سیستم عصبی فازی-تطبیقی (ANFIS) و برنامه‌ریزی بیان ژن (GEP) در شبیه‌سازی بار رسوبی رودخانه‌ها بررسی و نتایج به دست آمده از این سه روش با یکدیگر و با نتایج منحنی سنجه مقایسه کردند. نتایج تحقیق آنها، برتری مدل (GEP) با بیشترین ضریب تعیین R^2 با مقدار ۰/۹۹ و کمترین ریشه میانگین مربع خطا RMSE

روش متداول منحنی سنجه رسوب داشته‌اند، اما با توجه به روابط خطی ساده و قابل فهم ارائه شده توسط مدل درختی M5 کاربرد این روش کارآمد، را در موارد مشابه توصیه کردند. ایوانی و همکاران (۲۰۱۶) در مطالعه‌ای به برآورد بار معلق در سیستم رودخانه‌ای با استفاده از روش گروهی کنترل داده‌ها (GMDH) پرداختند. مقادیر معیارهای آماری به دست آمده و نیز مقایسه نتایج با سایر الگوریتم‌ها مانند شبکه‌های عصبی، نورفازی و برنامه‌ریزی ژنتیک خطی نشان داد که روش گروهی کنترل داده‌ها نسبت به سایر روش‌ها از قابلیت بالاتری برای پیش‌بینی و شبیه‌سازی نرخ انتقال رسوب در رودخانه‌ها برخوردار است. نیکپور و سانی‌خانی (۲۰۱۷) در مطالعه‌ای از مدل‌های فازی-عصبی تطبیقی (ANFIS)، برنامه‌ریزی بیان ژن (GEP) و رگرسیون بردار پشتیبان (SVR) جهت مدل‌سازی و پیش‌بینی میزان رسوبات معلق حوضه آبریز دره‌رود در استان اردبیل استفاده کردند. نتایج بیانگر عملکرد قابل قبول مدل‌ها و برتری مدل SVR با بیشترین ضریب تعیین ($R^2=0/97$) کمترین ریشه میانگین مربعات خطا ($RMSE=17342 \text{ ton day}^{-1}$) و همچنین شاخص‌های نش-ساتکلیف ($NS=0/97$) و ویلموت ($WI=0/98$) در مرحله صحت‌سنجی بود. یکی از عوامل تأثیرگذار بر میزان رسوب ورودی به مخازن سدها و دریاچه‌ها بار معلق ورودی است. برآورد دقیق میزان بار رسوب معلق از جنبه‌های مختلف دارای اهمیت است (قبادیان و شکری ۲۰۱۸). بیرانوند و همکاران (۲۰۲۲) در پژوهشی به مدل‌سازی رسوب معلق با استفاده از الگوریتم‌های یادگیری ماشین به بررسی کارایی مدل‌های GP-RBF¹، GP-PUK²، RepTree³، RF⁴ و M5P⁵ برای مدل‌سازی بار معلق رودخانه در دوره‌های کم‌آبی و پرآبی در حوزه آبخیز کشکان پرداختند. نتایج نشان داد با توجه به

¹ Gaussian processes radial base core (RBF)

² Gaussian processes Pearson's Kernel VII (PUK)

³ Reduced Error Pruning Tree

⁴ Random Forest

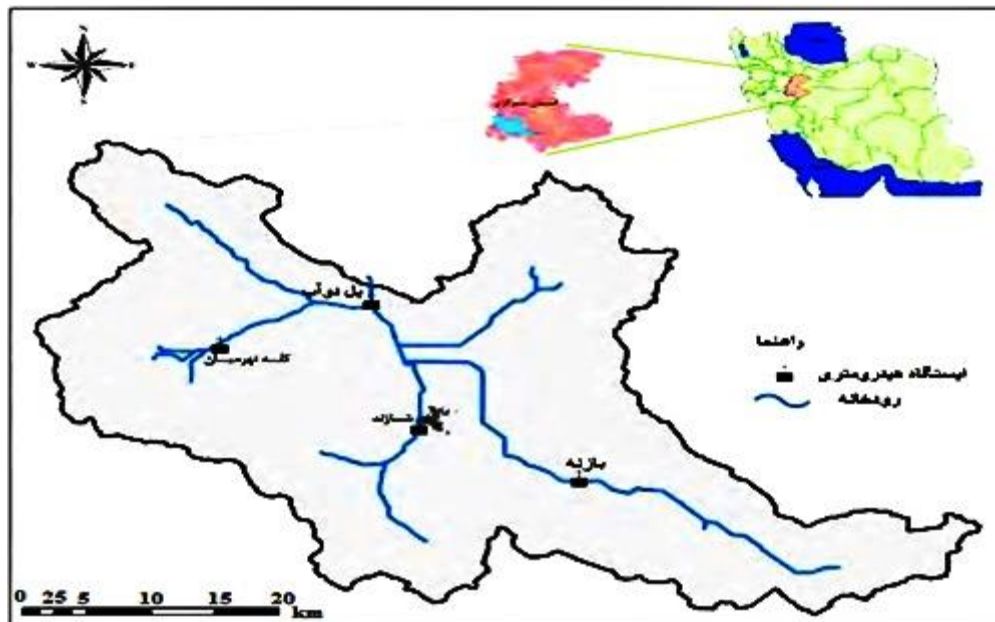
⁵ M5 model trees

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مطالعاتی این تحقیق حوضه آبریز پلدوآب با مختصات طول " ۱۵ ۰۴' ۴۹° تا " ۱۲ ۱۲' ۵۲° شرقی و " ۴۲ ۴۴' ۳۳° تا " ۱۳ ۱۲' ۳۴° شمالی، از زیر حوزه‌های دریاچه نمک است. رودخانه خروجی آبریز که به شفاء معروف است از سه زهکش اصلی به نام- های بازنه، شازند و نهرمیان تغذیه می‌شود (شکل ۱). این رودخانه در پایین دست پس از ادغام با سرشاخه- های اصلی رودخانه قره‌چای که از استان همدان سرچشمه می‌گیرد، در نهایت به سد الغدیر ساوه منتهی می‌شود. متوسط بارندگی سالانه حوزه برابر ۴۲۹ میلی‌متر و متوسط دمای سالانه برابر ۱۱/۵ درجه سانتی‌گراد است.

بر حسب تن در روز با مقدار ۰/۰۱۰ را نشان داد. در این خصوص کارایی مدل SVR با R^2 با مقدار ۰/۹۵ و RMSE برابر ۱/۴۹ تا حدی بهتر از مدل ANFIS با R^2 با مقدار ۰/۹۱ و RMSE برابر ۷/۹۴ بود. همچنین نتایج به دست آمده نشان داد هر سه روش داده‌کاوی بررسی شده نسبت به منحنی سنجه رسوب، نتایج به مراتب بهتری ارائه می‌کنند. به دلیل اینکه تعیین بار معلق رودخانه معمولاً با صرف زمان و هزینه زیاد همراه است، لذا بهتر است با انتخاب یک مدل با دقت زیاد، میزان رسوب معلق را تعیین کرد. با توجه به سابقه تحقیق به سبب متفاوت بودن شرایط آب و هوایی و مشخصات مناطق مطالعه شده با یکدیگر مطالعه و اندازه‌گیری رسوب آن از اهمیت خاصی برخوردار است. لذا هدف از این تحقیق بررسی کارایی روش‌های نرم در برآورد رسوب معلق ایستگاه پلدوآب رودخانه قره‌چای و انتخاب مناسب‌ترین روش می‌باشد.



شکل ۱- موقعیت حوزه آبخیز پلدوآب شازند در استان مرکزی.

به‌عنوان ابزاری با قابلیت بالا در مسیریابی و تشخیص روندهای غیرخطی پیچیده، به‌ویژه با تعداد مشاهدات محدود، برای الگوسازی بار رسوبی

روش تحقیق

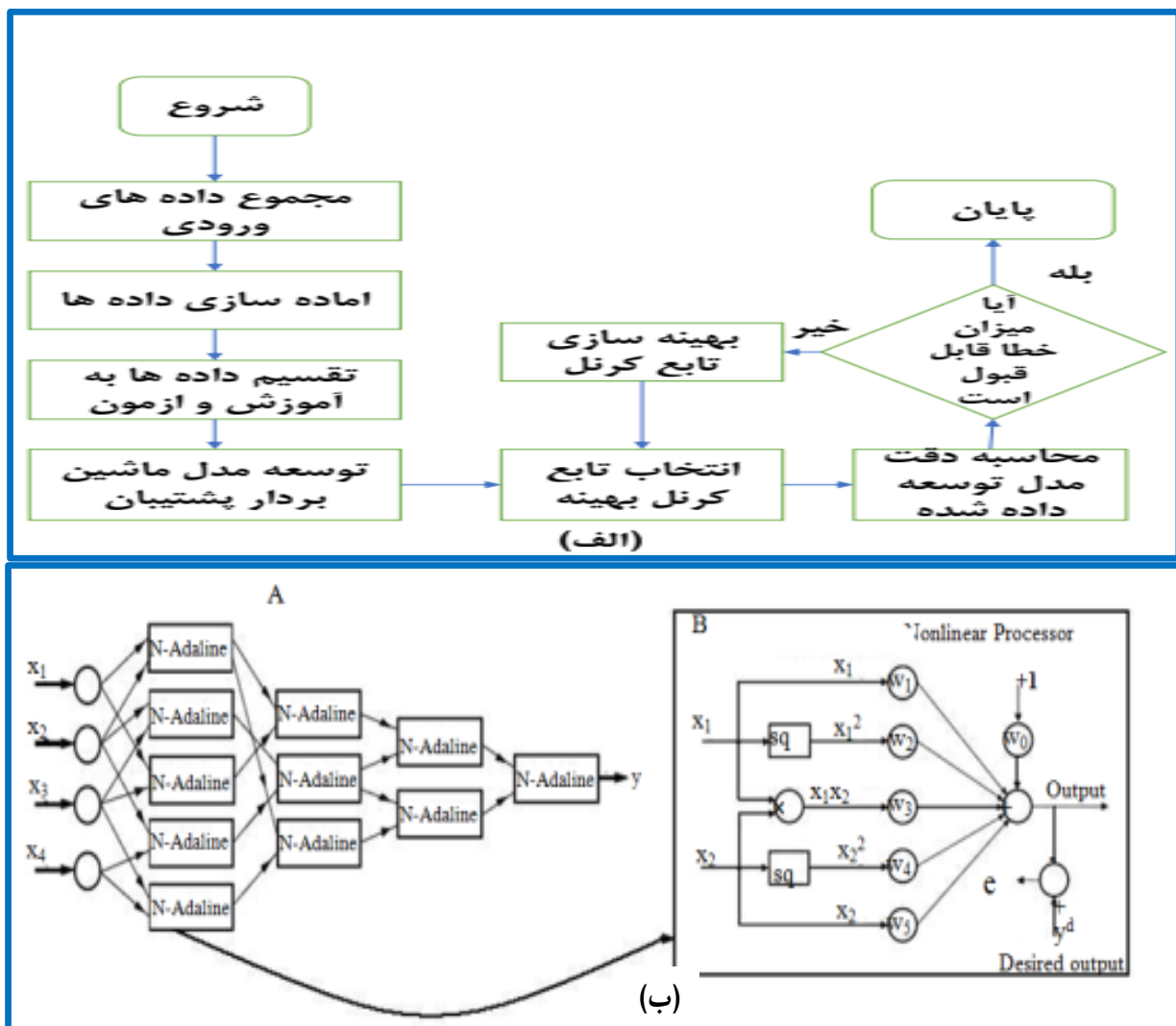
در این پژوهش به‌عملکرد دو نوع مدل ماشین بردار پشتیبان (SVR) به‌دلیل کارایی و سرعت بالا و GMDH

ایستگاه پل دوآب رودخانه قره‌چای استان مرکزی استفاده شد. توجه به این نکته حائز اهمیت است که چون GMDH بر مبنای داده‌های حاصل از سیستم رودخانه عمل می‌کند، ویژگی‌های رودخانه بر برآورد نتایج تأثیرگذار است. سپس نتایج دو روش با یکدیگر و با نتایج منحنی سنجه مورد مقایسه قرار گرفت. در نهایت بهترین روش پیشنهاد شد. برای این کار ابتدا مطالعات کتابخانه‌ای، میدانی و مرور منابع مرتبط، جمع‌آوری آمار و اطلاعات انجام شد. آمار دما، بارندگی، دبی متوسط روزانه جریان و رسوب اندازه-گیری شده روزانه در طول یک دوره آماری بلندمدت ۴۰ ساله ایستگاه هیدرومتری پل دوآب از اداره هواشناسی و آب منطقه‌ای استان مرکزی دریافت شد. داده‌های دریافت شده دسته‌بندی و به‌فرمت ورودی مدل‌ها تبدیل شدند. با توجه به دبی و رسوب متناظر به آن، منحنی سنجه رسوب رسم و معادله آن به‌دست آمد. الگوهای مناسبی از متغیرهای ورودی بر اساس آزمون و خطا انتخاب شد. با توجه به اینکه پارامترهای ذکر شده دارای سیر تاریخی هستند لذا طراحی الگوهای ورودی مدل‌های محاسبات نرم باید بر اساس تأخیرات زمانی انجام گیرد. (نظیر آنچه در تحلیل و پیش‌بینی سری‌های زمانی مطرح است). سپس مدل برای هر الگو ورودی ران و خروجی گرفته شد. در مرحله بعد مناسب‌ترین

تأخیر زمانی پارامترهای ورودی در مدل‌سازی که ضریب تعیین R^2 بیشتری و ریشه میانگین مربعات خطا RMSE کمتری داشت انتخاب شد. در این پژوهش ۷۰ درصد داده‌های تحقیق به‌عنوان آموزش و ۳۰ درصد برای اعتبارسنجی و آزمون مورد استفاده قرار گرفت. در نهایت دو روش داده‌کاوی با یکدیگر و همچنین با منحنی سنجه و داده‌های مشاهداتی مقایسه شدند. محدوده تغییرات و مشخصات آماری پارامترهای دبی جریان، دبی رسوب، بارش و دما روزانه در جدول ۱ ارائه شده‌اند. شکل ۲ نمودار و فلوجارت ماشین بردار پشتیبان و روش گروهی کنترل داده‌ها را نشان می‌دهد. همچنین با توجه به اهمیت پاسخ حوزه به متغیرهای ورودی به مدل‌ها، علاوه بر متغیر دبی و رسوب از بارندگی و دما نیز استفاده شد. از متغیر دینامیک بارندگی به‌دلیل نقش تأثیرگذار در ایجاد فرسایش و تولید رسوب استفاده شد. دلیل تأثیر دما را می‌توان گفت، آنچه که بر میزان رسوب معلق تأثیر فروانی دارد رطوبت خاک اراضی منطقه است. فرض کنید بارندگی با شدت یکسان روی دو خاک با رطوبت متفاوت بیارد. با توجه به اینکه فرسایش خاک خشک بیشتر است مقدار بار معلق رودخانه بیشتر می‌شود. اگر دما بیشتر باشد رطوبت خاک کمتر است.

جدول ۱- آمار خلاصه داده‌های کمی در ایستگاه پل دوآب.

پارامترها	تعداد داده	کمینه	بیشینه	متوسط	انحراف معیار
P(mm)	۳۰۳	۰	۸۳	۲/۵۲	۸
T(°c)	۳۰۳	-۱۵/۲۰	۲۹/۲۵	۱۰/۸۶	۸/۵
Q(m ³ s ⁻¹)	۳۰۳	۰/۰۲۳	۱۱۰	۶/۵	۱۵
S(kg day ⁻¹)	۳۰۳	۰	۳۵۰/۳۳	۱۷/۹۸	۶۰/۳



شکل ۲- مراحل اجرای (فلوچارت) الف: ماشین بردار پشتیبان (SVR)، ب: روش گروهی کنترل داده‌ها (GMDH).

$$R^2 = \frac{\sum_{i=1}^n (x_o - \bar{x}_o)(x_c - \bar{x}_c)}{(\sum_{i=1}^n (x_o - \bar{x}_o)^2)^{0.5} (\sum_{i=1}^n (x_c - \bar{x}_c)^2)^{0.5}} \quad [۱]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_o - x_c)^2}{n}} \quad [۲]$$

$$NS = 1 - \frac{\sum_{i=1}^n (x_o - x_c)^2}{\sum_{i=1}^n (x_o - \bar{x}_o)^2} \quad [۳]$$

$$MBE = \frac{1}{n} \sum_{i=1}^n (x_o - x_c) \quad [۴]$$

معیارهای ارزیابی

برای بررسی دقت نتایج مدل‌ها از چهار معیار آماری شامل میانگین قدرمطلق خطا^۱ (MAE)، ریشه میانگین مربعات خطا^۲ (RMSE)، میانگین خطای اربیبی^۳ (MBE) و ضریب تشخیص^۴ (R²) به ترتیب برای بررسی مقادیر خطای برآورد، کم یا بیش برآورد و همبستگی بر اساس روابط ۱ تا ۴ و همچنین از دیاگرام تیلور استفاده شده است.

که در آن x_o داده مشاهده شده، x_c داده پیش‌بینی شده، n تعداد داده‌های مشاهداتی، \bar{x}_o میانگین داده مشاهده شده و \bar{x}_c میانگین داده های پیش‌بینی شده است.

¹ Mean Absolute Error (MAE)

² Root Mean Squared Error (RMSE)

³ Mean Bias Error (MBE)

⁴ Coefficient of Determination (R²)

نتایج و بحث

برای بررسی اینکه داده‌ها دارای توزیع نرمال هستند یا نه از آزمون‌های شاپیرو ویلک، اندرسون دارلینگ، لیلیه فورس و جارکویرا استفاده شده است. آزمون نرمالیته برای داده‌ها با انجام آزمون‌هایی توسط نرم-افزار آماری XLSTAT مورد بررسی قرار گرفته و نتایج مربوطه در جدول ۲ ارائه گردیده است. در این آزمون‌ها فرض صفر برابر با نرمال بودن داده‌ها و فرض مقابل برابر با غیرنرمال بودن داده‌ها در نظر گرفته شده است. همان‌طور که مشاهده می‌گردد داده‌ها در تمامی آزمون‌های مورد بررسی نرمال نبوده‌اند. داده‌های دما، بارندگی، دبی جریان و بار رسوب برای ایستگاه انحراف قابل توجهی از توزیع نرمال دارند. برای ایستگاه پل دوآب داده‌ها انحراف قابل توجهی را از توزیع نرمال نشان می‌دهند اما به‌طور کلی می‌توان گفت که انحراف برای داده‌های بار رسوب بیشتر از داده‌های دبی جریان است. شایان ذکر است که دلیل استفاده از نمودارهای چندک احتمالی نرمال توانایی آنها در نشان دادن میزان انحراف از توزیع نرمال است. به کمک این نمودارها می‌توان راجع به میزان انحراف داده‌ها از نرمال و اثر آن بر عملکرد رگرسیون‌گیری اظهارنظر کرد. با توجه به اهمیت استفاده از داده‌های صحیح آماری، کلیه داده‌های موجود از نظر همگن بودن توسط آزمون همگنی نرمال استاندارد که یکی از روش‌های متداول برای ارزیابی همگنی داده‌ها می‌باشد، مورد بررسی قرار گرفت. لازم به‌ذکر است که در روش‌های داده‌کاوی لزومی به تبعیت پارامترهای به‌کار رفته در مدل‌سازی از توزیع نرمال نبوده و از این‌روی مشکلی در فرآیند مدل‌سازی ایجاد نمی‌کند. به‌عنوان یک قانون کلی در تکنیک‌های آماری فرض اولیه بر این اساس است که توزیع داده‌ها مشخص است که بیشتر موارد فرض بر این است که توزیع نرمال است و در نهایت درستی یا نادرستی نتایج نهایی به درست بودن فرض اولیه وابسته است. در مقابل روش‌های داده‌کاوی از

هیچ فرض در مورد داده‌ها استفاده نمی‌کنند (حائری مهریزی ۲۰۱۲، ستاری و همکاران ۲۰۱۶). با توجه به اهمیت استفاده از داده‌های صحیح آماری، کلیه داده‌های موجود از نظر همگن بودن توسط آزمون همگنی نرمال استاندارد که یکی از روش‌های متداول برای ارزیابی همگنی داده‌ها می‌باشد، مورد بررسی قرار گرفت. نتایج مربوط به آزمون همگنی نرمال استاندارد در جدول ۳ و شکل ۳ ارائه شده است. در این آزمون فرض صفر، همگنی داده‌ها و فرض یک، ناهمگن بودن داده‌ها می‌باشد. چنانکه مقدار P-Value از مقدار درجه اطمینان بزرگتر باشد، فرض صفر صحیح است، در غیر این صورت فرض یک قابل قبول می‌باشد. پس از استفاده از آزمون همبستگی پیرسون مشخص شد بین پارامترها رابطه معنی‌داری در سطح ۰/۰۵ درصد جود دارد. نتایج جداول ۲ تا ۴ نشان‌دهنده این مطلب است که داده‌های روزانه دما، بارش، دبی جریان و دبی رسوب در ایستگاه پل دوآب و طول آماری مورد مطالعه همگن بوده و می‌توانند با اطمینان به‌کار برده شوند. یکی از مباحثی که در مطالعات هیدرولوژیکی مطرح می‌باشد مسئله کفایت یا عدم کفایت طول آماری داده‌ها برای انجام مطالعات و مدل‌سازی می‌باشد. روش متداول در این زمینه محاسبه ضریب هرست^۱ است (صفوی ۲۰۱۲) اگر این ضریب بیشتر از مقدار ۰/۵ باشد، طول آماری داده‌ها برای انجام مطالعات کافی و اگر کمتر باشد، بدین معنی است که قبل از انجام مطالعات باید طول آماری داده‌ها تطویل گردد. محاسبه و بررسی ضریب هرست نشان داد طول آماری داده‌ها با مقدار ضریب هرست برابر با ۰/۵۸ برای انجام مطالعات کافی و مناسب می‌باشد. نقشه همبستگی و همگنی داده‌ها و تصویر ماتریس همبستگی آنها نمودارهای پراکنش پارامترها در شکل‌های ۳ و ۴ آمده است.

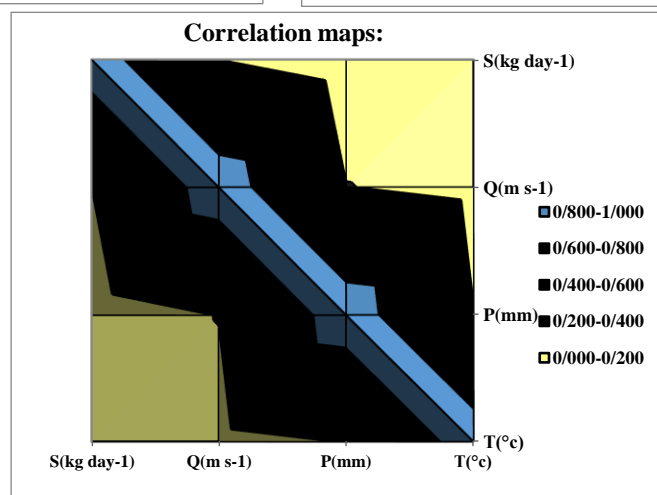
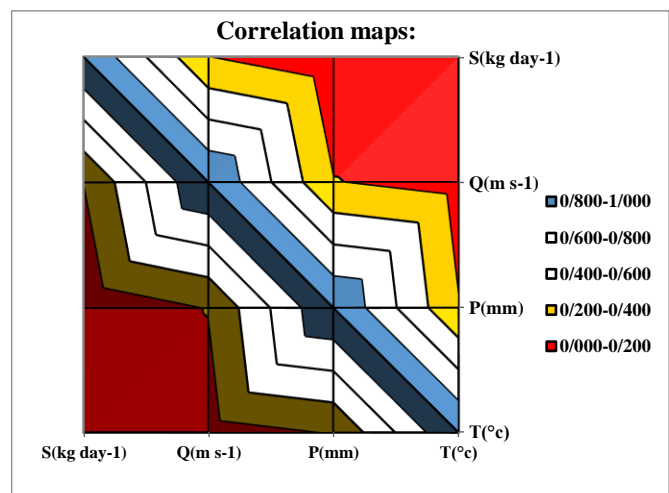
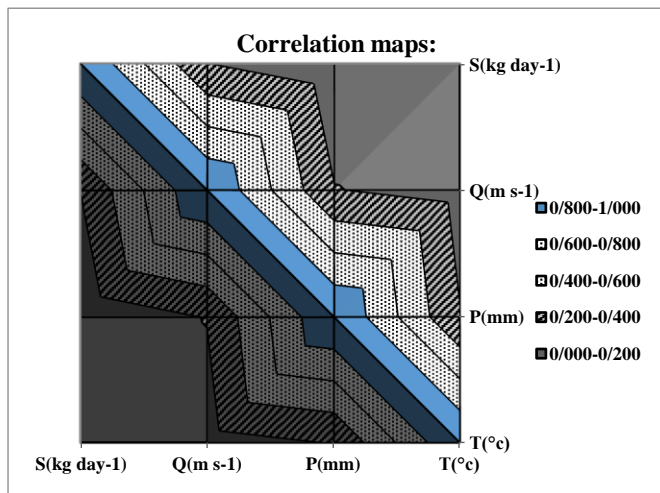
¹ Hurst coefficient

جدول ۲- نتایج آزمون‌های تعیین نرمال بودن داده‌ها ایستگاه پل دوآب.

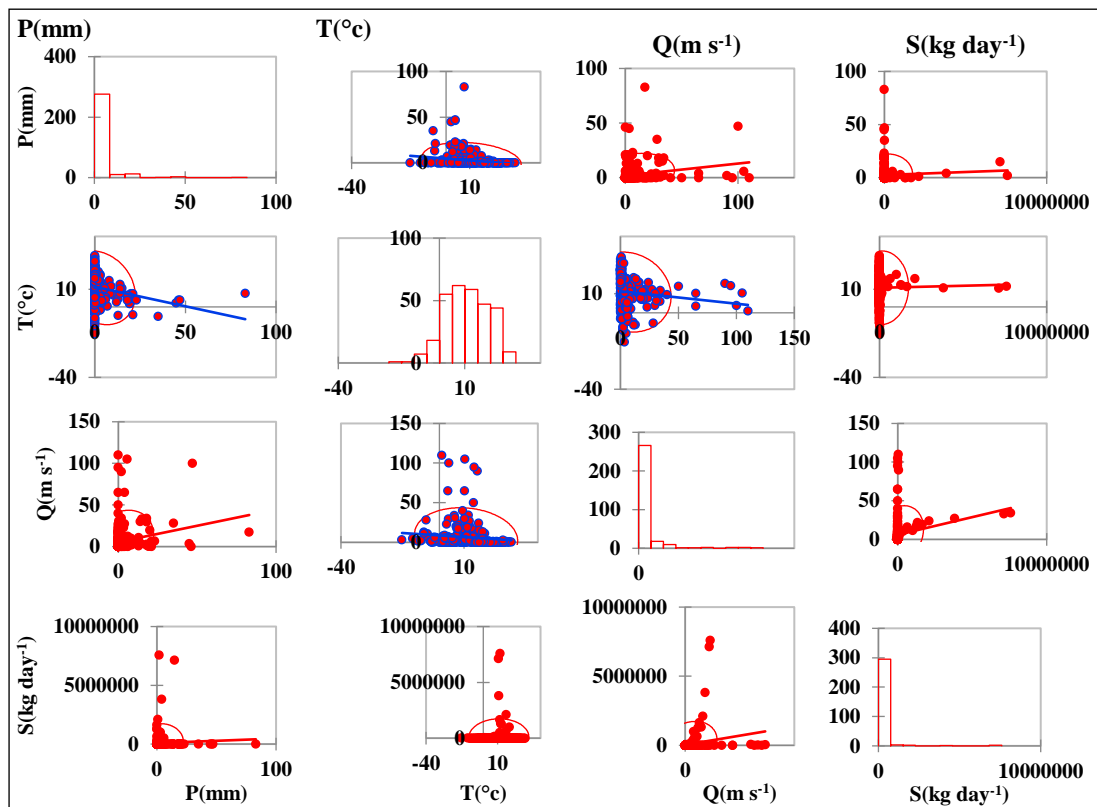
	آزمون شاپیرو ویلک			آزمون اندرسون دارلینک			آزمون لیلیه فورس			آزمون چارکو برا		
	سطح	ریسک	اماره	سطح	ریسک	اماره	سطح	ریسک	اماره	سطح	ریسک	اماره
P	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵
T	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵
Q	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵
S	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵	<۰/۰۰۰۱	۰/۰۱	۰/۰۵

جدول ۳- مقادیر p-values و ضرایب تعیین (پیرسون).

Variables	p-values (Pearson)				Coefficients of determination (Pearson)			
	P(mm)	T(°C)	Q(m ³ s ⁻¹)	S(kg day ⁻¹)	P(mm)	T(°C)	Q(m ³ s ⁻¹)	S(kg day ⁻¹)
P(mm)	.	.	.	۰/۱۸۳	۱	-۰/۴۴	-۰/۴۳	-۰/۰۶
T(°C)	.	.	۰/۰۴۴	۰/۷۲۹	-۰/۴۴	۱	-۰/۱۳	.
Q(m ³ s ⁻¹)	.	۰/۰۴۴	.	<۰/۰۰۰۱	-۰/۴۳	-۰/۱۳	۱	-۰/۶۵۲
S(kg day ⁻¹)	۰/۱۸۳	۰/۷۲۹	<۰/۰۰۰۱	.	-۰/۰۶	.	-۰/۶۵۲	۱



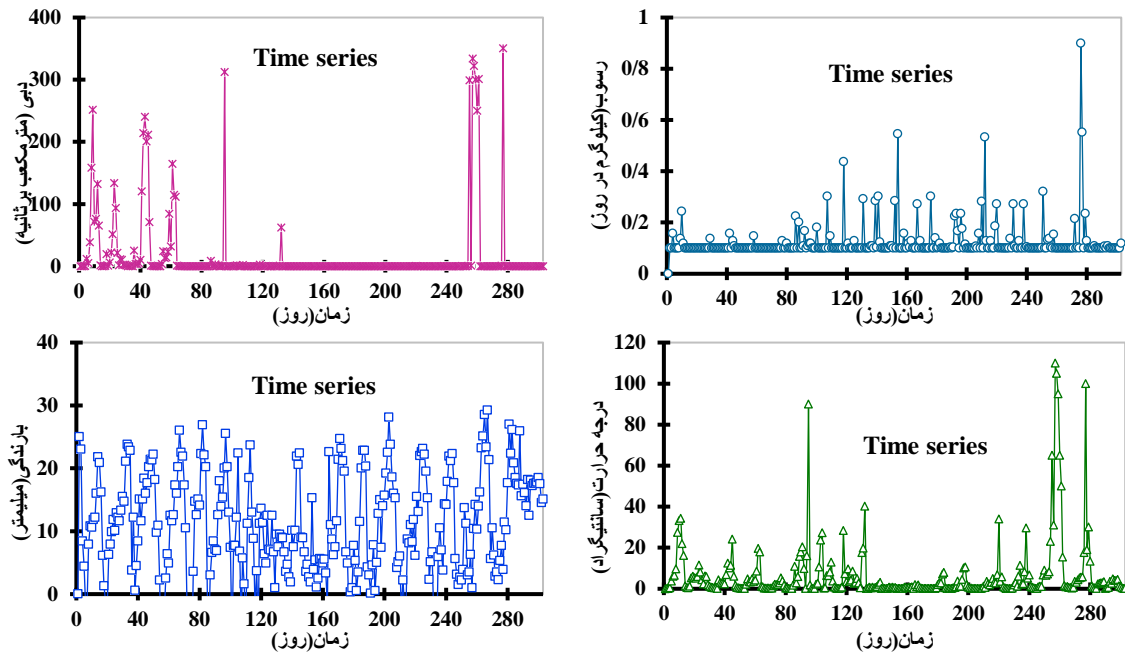
شکل ۳- نقشه‌های همبستگی بین پارامترها ورودی در ایستگاه پل دوآب.



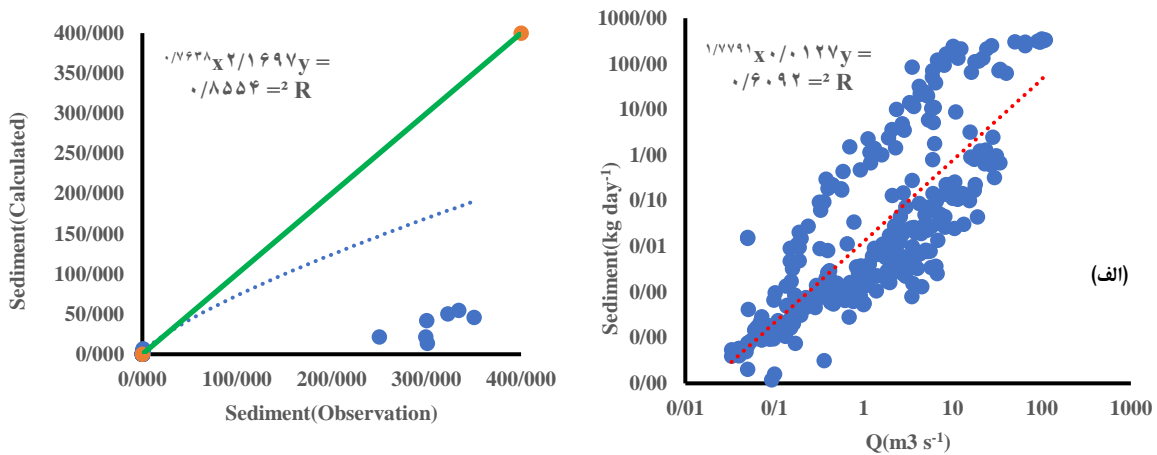
شکل ۴- نمودارهای پراکنندگی در ایستگاه پل دوآب.

اندازه‌گیری و منحنی مربوط رسم گردید. شکل (۶-الف) منحنی سنجۀ رسوب را برای مرحله آموزش در ایستگاه باقرآباد نشان می‌دهد. در شکل (۶-ب) منحنی پلات برای آزمون نرمال بودن داده‌ها استفاده شد. با رسم این منحنی، احتمال تجمعی مشاهدات در برابر احتمالات تجمعی مقادیر محاسبه شده از معادله به دست آمده رسم گردید. در ابتدا داده‌ها به دو قسمت تقسیم شدند، پس از آن ۷۰ درصد از داده‌ها برای آموزش و ۳۰ درصد برای آزمون در نظر گرفته شد.

سری زمانی و فراوانی بارش، رسوب و جریان ماهانه ۴۰ سال آماری ایستگاه باقرآباد رودخانه قمرود در شکل ه ارائه شده است. به دلیل در دسترس نبودن آمار دقیق میزان فرسایش و رسوب در حوضه آبریز، از منحنی سنجۀ رسوب تهیه شده با داده‌های دبی و غلظت رسوب یا دبی رسوب معلق استفاده شد. منحنی‌های سنجۀ رسوب براساس اندازه‌گیری‌ها در رودخانه به- دست آمد. به این صورت که متناظر با هر دبی خاص در رودخانه، غلظت مواد رسوبی حمل شده توسط آب



شکل ۵- سری زمانی دبی جریان، دبی رسوب، بارندگی و دما روزانه ایستگاه پل دوآب.



شکل ۶- الف: منحنی سنجه رسوب قسمت آموزش، ب: مقادیر محاسباتی و مشاهداتی برای قسمت آزمون.

برنامه تدوین شده در محیط نرم افزار MATLAB استفاده شد. در گام نخست، ۱۳ سناریوی مختلف (f1 تا f13) به عنوان الگوهای ورودی در مدل SVR به شرح جدول ۳ استفاده گردید. معیار انتخاب سناریوها بر اساس مرور منابع و استفاده از تجربیات دیگران، روش آزمون و خطا، شرایط آب و هوایی، کوهستانی بودن منطقه و رژیم رودخانه انتخاب شدند.

برای تخمین رسوب معلق از الگوهای ورودی مختلف شامل مقادیر پیشین دما، بارندگی، رسوب و مقادیر دبی جریان در گام زمانی فعلی و قبل استفاده شد تا تأثیر هر یک از این متغیرها در مدل سازی رسوب معلق مشخص گردد. از بین مدل های مورد بررسی، ابتدا مدل SVR به دلیل کارایی و سرعت بالاتر در فرآیند مدل سازی انتخاب گردید. برای به کارگیری این مدل، از

جدول ۴- سناریوهای مختلف استفاده شده برای مدل‌ها

شماره	تابع	شماره	تابع
(f1)	Q_t, S_t, T_t, P_t	(f8)	Q_t, Q_{t-1}, S_t
(f2)	$Q_t, Q_{t-1}, S_{t-1}, S_t, T_t, P_t, P_{t-1}$	(f9)	$Q_t, Q_{t-2}, Q_{t-1}, S_{t-1}, S_t$
(f3)	Q_t, S_{t-1}, S_t	(f10)	$Q_t, Q_{t-1}, S_{t-1}, S_t, P_t$
(f4)	Q_t, S_t, T_t	(f11)	$Q_t, Q_{t-1}, S_{t-1}, S_t, T_t, T_{t-1}$
(f5)	Q_t, S_t, P_t	(f12)	$Q_t, Q_{t-1}, S_{t-1}, S_t, T_t, T_{t-1}, P_{t-1}, P_t$
(f6)	$Q_t, Q_{t-1}, S_{t-1}, S_t$	(f13)	$Q_t, Q_{t-1}, S_{t-1}, S_t, P_{t-1}, P_t$
(f7)	Q_t, S_t		

نتایج حاصل از روش SVR در ایستگاه پل دوآب

پس از مرتب کردن داده‌ها و تعیین پارامترهای مستقل و وابسته برای هر الگو یا سناریو، داده‌ها وارد مدل شدند، و مدل اجرا گردید. در این مرحله برای هر الگو، ۱۵ بار مدل اجرا و خروجی گرفته شد. در نهایت میانگین R^2 و RMSE این ۱۵ مدل انتخاب و در جدول ۵ آورده شد. در این جدول عملکرد مدل SVR به‌ازای ۱۳ الگوی ورودی مختلف در قالب شاخص‌های آماری خطا و همچنین مقادیر بهینه پارامترهای مدل (σ) به‌ازای هر الگو نشان داده شده است. در مرحله بعد ۱۳ الگو (سناریو) با هم مورد مقایسه قرار گرفت و بهترین آنها که ضریب تبیین (R^2) بالاتری و RMSE کمتری داشت انتخاب شد. با توجه به جدول ۵ و مقادیر شاخص‌های آماری خطا، می‌توان دید بهترین عملکرد مدل SVR به‌ازای الگوی شماره ۶ حاصل شده است که در آن از مقادیر دبی جریان در همان روز و دبی جریان و رسوب معلق در یک گام زمانی قبل به‌عنوان ورودی استفاده شده است. همچنین ضعیف‌ترین عملکرد مدل به‌ازای الگوی شماره ۷ بوده است که شامل مقدار دبی جریان در همان روز می‌باشد. دلیل این موضوع می‌تواند عدم وجود متغیر دبی جریان در گام زمانی قبل می‌باشد و نشان می‌دهد که مقادیر رسوب معلق وابستگی زیادی به دبی جریان و رسوب در گام زمانی قبل دارند. به‌ازای بهترین عملکرد مدل مقادیر شاخص‌های آماری R^2 ، RMSE، در مرحله آزمون به ترتیب برابر ۰/۹۸ و ۱۸۵ تن در روز و در مرحله آموزش ۰/۸۳ و ۲۴۸ تن در روز

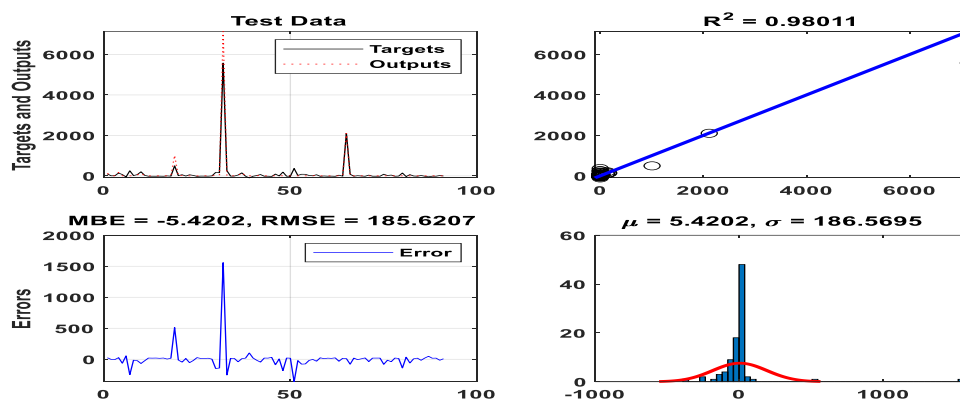
حاصل شده است. نتایج به‌دست آمده در این پژوهش با تحقیق آیتک و کیشی (۲۰۰۸) مطابقت دارد. مقایسه نتایج به‌دست آمده در این تحقیق با مطالعه کیشی و همکاران (۲۰۱۲) حاکی از آن است که در مطالعه آنها نیز بهترین الگوی ورودی مدل‌های هوشمند شامل استفاده از مقادیر رسوب معلق در گام زمانی قبل، دبی جریان در گام زمانی فعلی و قبل بوده است که با نتایج تحقیق حاضر همخوانی دارد. همچنین با توجه به نتایج ارائه شده در جدول ۵، می‌توان گفت در الگوهای ۱ و ۲ نیز نتایج تا حدودی رضایت بخش می‌باشد. در حالتی که از دبی جریان همان روز و دبی رسوب در گام زمانی قبل (الگوی شماره ۳) استفاده شده است، نتایج شاخص‌های آماری خطا نسبتاً قابل قبول می‌باشد که این موضوع اثرات دوگانگی رفتار بین رسوب معلق و دبی جریان را نشان می‌دهد. می‌توان گفت که با در نظر گرفتن دبی جریان فعلی و دما در همان روز به‌عنوان ورودی (الگوی شماره ۴) عملکرد مدل تا حدود ضعیف می‌باشد. نتایج نشان می‌دهد که استفاده از دبی جریان و مقدار بارندگی در همان روز (الگوی شماره ۵) عملکرد مدل تا حدود ضعیف می‌باشد. همچنین استفاده از مقادیر رسوب با تأخیر در یک گام زمانی و دبی با یک گام زمانی تأخیر (الگوی شماره ۶) نتایج بهتر شده است. با به‌کارگیری مقادیر دبی جریان به تنهایی (الگوی شماره ۷) نیز نتایج ضعیفی حاصل شده است. در الگوهای ۸ تا ۱۳ نتایج مدل تا حدودی بهبود یافته است. شکل‌های ۷ تا ۱۰ خروجی برنامه به‌ازای بهترین الگو

رابطه غیرخطی و پیچیده وجود بین مقادیر ورودی و خروجی را نشان دهد. به طور کلی نتایج حاصله در این بخش نشان می‌دهد با حذف مقادیر دما و بارندگی، عملکرد مدل به طور قابل توجهی بهبود می‌یابد.

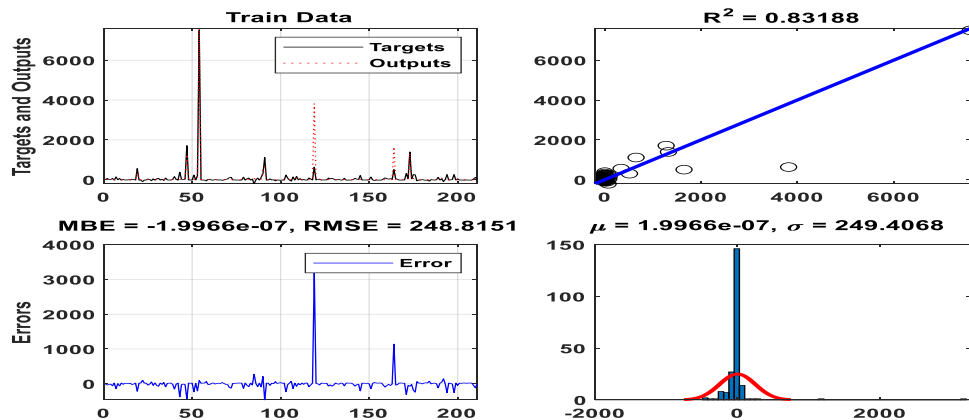
(الگوی ۶) را نشان می‌دهد. نمودار سری زمانی و پراکندگی داده‌های مشاهداتی و شبیه‌سازی شده رسوب معلق به ازای بهترین الگوی ورودی (الگوی شماره ۶) برای مدل SVR نشان داده شده است. با توجه به این شکل می‌توان گفت که مدل SVR به خوبی توانسته است

جدول ۵-مقادیر پارامترها و عملکرد مدل SVR به ازای الگوهای ورودی مختلف در دوره آزمون در ایستگاه پل دوآب.

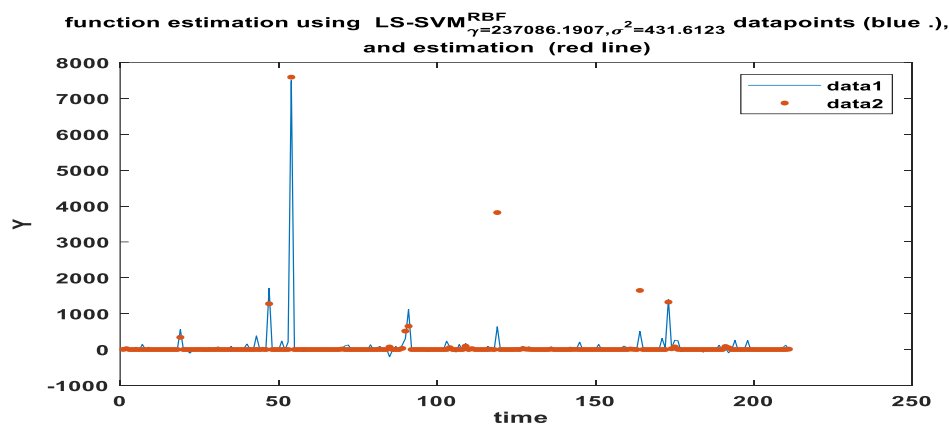
شاخصهای آماری ارزیابی عملکرد در مرحله آموزش			شاخصهای آماری ارزیابی عملکرد در مرحله			مدل
R^2	σ	RMSE (kg day ⁻¹)	R^2	σ	RMSE (kg day ⁻¹)	
۰/۸۱	۲۳/۵۹	۲۳/۶۵	۰/۶۲	۳۸/۸	۳۸/۷	SVR1
۰/۸۵	۳۱۴/۷	۳۱۳/۸۸	۰/۹۹	۶/۶۵	۶/۶۳	SVR2
۰/۹۵	۶۴۰/۶	۶۴۰/۹	۰/۴۶	۴۵۵/۷	۴۵۴/۶	SVR3
۰/۳۲	۲۴۳/۵	۲۶۲/۶	۰/۳۴	۱۳۶۵/۴	۶۶۴/۲	SVR4
۰/۴۸	۱۴۴	۱۵۱/۶۹	۰/۳۵	۶۷۰/۴۴	۶۶۸/۸	SVR5
۰/۹۸	۱۸۶/۵	۱۸۵/۶	۰/۸۳	۲۴۹/۴	۲۴۸/۸	SVR6
۰/۳۱	۷۲۱	۷۱۸/۹	۰/۱۸	۵۵۴/۸	۵۵۳	SVR7
۰/۶۵	۵۰۸/۸	۵۰۶	۰/۹۴	۱۳۷/۷	۱۳۷/۳	SVR8
۰/۹۴	۵۷۵	۵۷۳/۶	۰/۹۰	۱۹۱/۵	۱۹۱	SVR9
۰/۸۶	۲۰۷/۶	۲۰۹/۴۶	۰/۹۹	۱۳/۲۸۶	۱۳/۱۷۲	SVR10
۰/۷۶	۲۹۹/۸	۳۰۵/۲	۰/۹۹	۲۰/۲۷	۲۰/۲۲	SVR11
۰/۹۲	۳۶۸	۴۶۹/۵	۰/۹۹	۳/۷۶	۳/۷۵	SVR12
۰/۷۹	۱۸۴/۸	۱۸۴/۸	۰/۹۹	۷/۱۳	۷/۱۱	SVR13



شکل ۷- نتایج خروجی مدل SVR در محیط متلب در مرحله آزمون ایستگاه پل دوآب.



شکل ۸- نتایج خروجی مدل SVR در محیط متلب در مرحله آموزش ایستگاه پل دوآب.



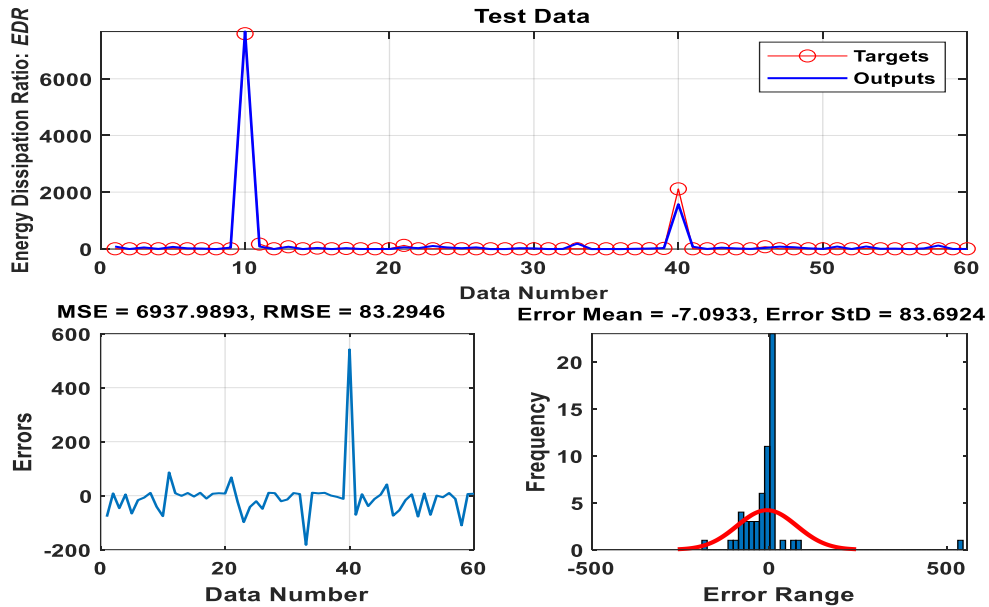
شکل ۱۰- نتایج خروجی مدل SVR و داده‌های واقعی در ایستگاه پل دوآب.

در روز شد. نتایج شاخص‌های آماری برای مدل-های SRC و SVR، GMDH، به‌عنوان ورودی و رودی به‌پینه مربوط به خود در جدول ۶ و در شکل ۱۴ دیگرام تیلور جهت بررسی بصری ارائه شده است. با توجه به نتایج حاصله می‌توان دید که عملکرد مدل GMDH در مقایسه با مدل SRC و SVR (منحنی سنج رسوب) بهتر بوده است. مدل SRC و SVR در رتبه‌های دوم، سوم قرار می‌گیرند. نتایج به‌دست آمده نشان داد هر دو روش داده‌کاوی بررسی شده به مراتب نتایج بهتری نسبت به منحنی سنج رسوب ارائه می‌کنند. با توجه به نتایج حاصله می‌توان گفت، مدل GMDH به‌عنوان مدلی توانمند و با سرعت بالا می‌تواند جهت مدل‌سازی رسوب معلق در حوضه آبریز پل دوآب مورد استفاده

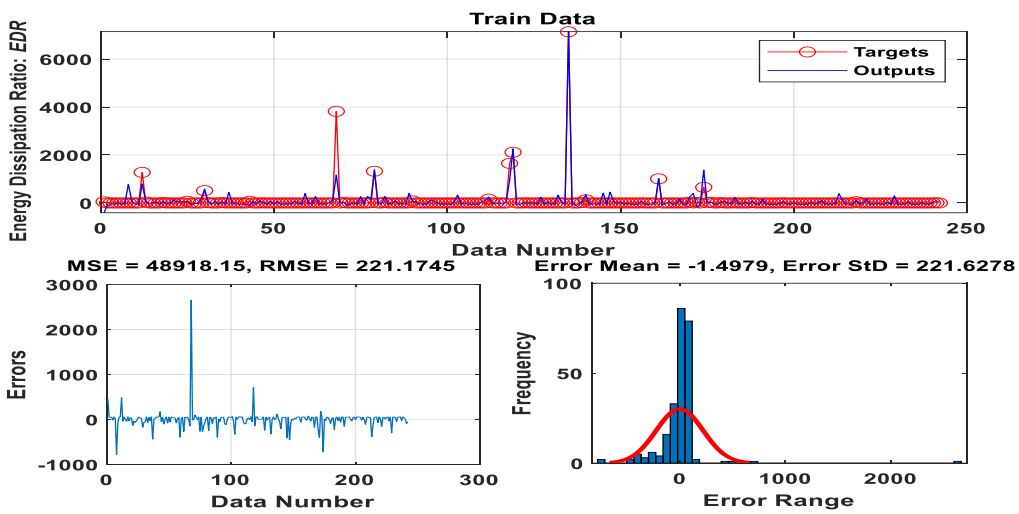
در مرحله بعد، بهترین الگوی انتخابی مدل‌های (SVR) به‌عنوان ورودی مدل GMDH استفاده شد. الگوی ورودی ۶ که بهترین الگو برای مدل‌های (SVM) انتخاب شد به‌عنوان ورودی مدل GMDH معرفی شد. در مرحله آموزش و آزمون، مقادیر شاخص‌های آماری R^2 به ترتیب برابر ۰/۹۹ و ۰/۷۹ مقدار خطای RMSE به ترتیب برابر ۸۳ و ۲۲۱ و مقدار MSE به ترتیب برابر ۶۹۳۷ و ۴۸۹۱۸ و مقادیر StD به ترتیب برابر ۸۳ و ۲۲۱ به‌دست آمد. شکل‌های ۱۱ و ۱۳ نتایج خروجی مدل GMDH در محیط متلب در مرحله آموزش و آزمون در ایستگاه پل دوآب را نشان می‌دهد. نتایج بیانگر عملکرد قابل قبول مدل GMDH با بیشترین ضریب تعیین R^2 برابر ۰/۹۹ و کمترین ریشه میانگین مربعات خطا برابر ۸۳ کیلوگرم

قرار گیرد. از آنجا که نقاط پیک در تعیین مقدار ظرفیت ذخیره و یا آبگذری سازه‌های مختلف اهمیت دارد این مقادیر از جمله مهمترین اطلاعات مورد نیاز در طراحی همه سازه‌ها به شمار می‌روند که این مدل توانسته است به خوبی مقادیر پیک رسوب را پیش‌بینی کند. مقایسه با مطالعات انجام شده نشان داد که نتایج این تحقیق با نتایج تحقیقات شیخعلی‌پور و همکاران (۲۰۱۴) و کیسی و شیرینی (۲۰۱۲) و ایوانی و همکاران (۲۰۱۶) مطابقت دارد. بالا بودن مقدار رسوب در این ایستگاه می‌تواند دلایل مختلفی داشته باشد. من جمله، به خاطر کوهستانی بودن منطقه، تغییر کاربری شدید و تبدیل شدن زمین-های کشاورزی به صنعتی و... می‌باشد. روش‌های ماشین‌یادگیری دارای دقت و صحت بالاتری نسبت به روش منحنی سنج رسوب هستند. دلیل این امر می‌تواند به دلیل خاصیت رگرسیونی منحنی سنج رسوب باشد. بدین معنی که اغلب داده‌های استفاده شده برای مدل-سازی از روش SRC مربوط به دبی‌های کم بوده و به-دلیل اینکه بیشترین مقدار رسوب در دبی بالا حمل می-گردد، این مدل قادر به معرفی دوره‌های پر رسوب نیست و همچنین تبدیل نتایج این مدل از فضای لگاریتمی به فضای حسابی باعث کم‌برآوردی بار رسوب معلق خواهد شد. در صورتیکه مدل‌های هوشمند مانند SVR و GMDH حتی زمانی که داده‌ها از کمیت و کیفیت مناسبی برخوردار نیستند، برآورد دقیقی از مقدار SSL را خواهند داشت. با توجه به نتایج به‌دست آمده، علت برتری نتایج برخی مدل‌ها نسبت به برخی دیگر در این تحقیق را می‌توان به شرح زیر بیان کرد: مدل ماشین‌های بردار پشتیبان از دقت و کارایی خوبی برخوردار است. زیرا مدل ماشین بردار پشتیبان، بهترین مرز تصمیم را برای جداکردن داده‌ها تعیین می-کند، در صورتی که مثلاً شبکه عصبی مصنوعی به اولین مرز تصمیم جداکننده که می‌رسد یادگیری را پایان می‌دهد. در این صورت ممکن است مرز تصمیم

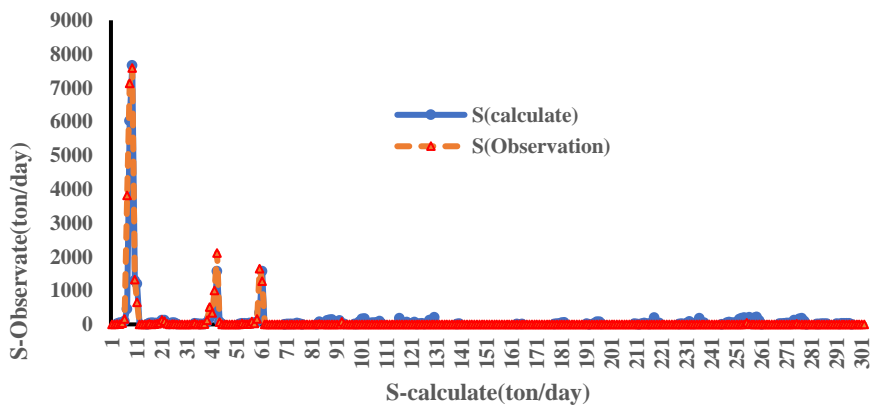
به‌دست آمده، مرز مناسبی نباشد. زمان اجرای ماشین-های بردار پشتیبان به تعداد دسته‌های موجود وابسته است و به مدت زمان کمتری نیاز دارد، در صورتی که زمان اجرای شبکه عصبی مصنوعی علاوه بر تعداد ورودی‌های آموزشی به تعداد نرون‌های لایه پنهان، تابع آموزشی، ضریب یادگیری و مومنتم که به‌صورت آزمون و خطا به‌دست می‌آیند، وابسته است. از نظر سرعت عمل مدل‌ها (سرعت اجرای کد و محاسبات) در برآورد، مدل گروهی کنترل داده‌ها در جایگاه اول، مدل ماشین بردار پشتیبان در جایگاه دوم قرار دارند. روابط تولید شده بین متغیرهای ورودی و خروجی در مدل‌ها بسیار پیچیده است که این یکی از نقاط ضعف می‌باشد. مدل SVR به جهت کارایی و سرعت بالا و GMDH به‌عنوان ابزاری با قابلیت بالا در مسیریابی و تشخیص روندهای غیرخطی پیچیده، به ویژه با تعداد مشاهدات محدود، برای الگوسازی استفاده شد. در این روش احتیاجی به شناخت روابط حاکم بر سیستم و خطی‌سازی روابط وجود ندارد. فقط لازم است که یک سری از داده‌های ورودی و خروجی از سیستم در دسترس باشند. زمان انجام محاسبات در این روش کم است ضمن آن که با افزایش تعداد لایه‌های توسعه داده شده دقت پاسخ‌های تولید شده اضافه می‌گردد، اما روابط تولید شده بین متغیرهای ورودی و خروجی بسیار پیچیده می‌شود که از این مورد می‌توان به یکی از نقاط ضعف روش مورد استفاده یاد کرد. در مطالعات آتی می‌توان پارامترهای ورودی به مدل را تغییر داد و اثر سایر پارامترهای موثر بر انتقال رسوب را بررسی کرد. می‌توان از روش‌های آموزش پارامترها مانند بهینه‌سازی بر اساس مشتق‌گیری یا الگوریتم‌های تکامل‌گرا مانند الگوریتم ژنتیک برای آموزش پارامترهای مورد نظر در GMDH استفاده کرد.



شکل ۱۱- نتایج خروجی مدل GMDH در محیط متلب در مرحله آزمون در ایستگاه پل دوآب.



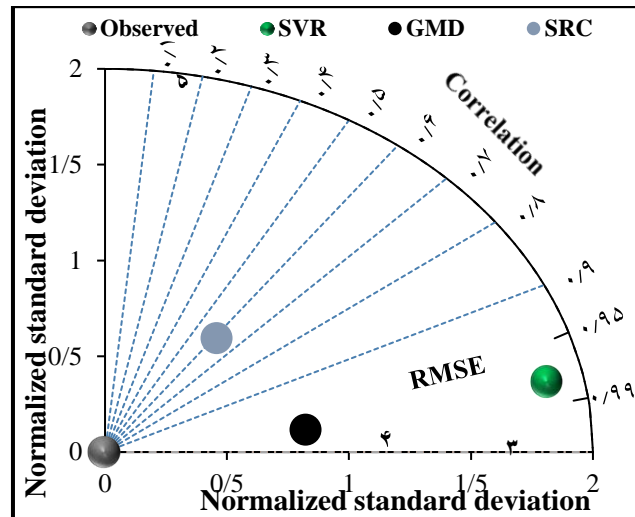
شکل ۱۲- نتایج خروجی مدل GMDH در محیط متلب در مرحله آموزش در ایستگاه پل دوآب.



شکل ۱۳- مقایسه بصری داده های رسوب مشاهداتی و داده های رسوب خروجی از مدل GMDH

جدول ۶- مقایسه عملکرد مدل‌های مورد استفاده جهت تخمین بار معلق به‌ازای بهترین الگوی ورودی ایستگاه پل دوآب.

مدل	الگوی بهینه	R ²	RMSE (Kg day ⁻¹)	MBE	NS
SVR	F6	۰/۹۸	۱۸۵	-۵/۴۲	۰/۸۹
GMDH	F6	۰/۹۹	۸۳	۲/۲	۰/۹۲
منحنی سنجه		۰/۶۱	۷۵	۲۰	۰/۱۰۵



شکل ۱۴- مقایسه عملکرد مدل‌های مورد استفاده جهت تخمین بار معلق با روش بصری از روش منحنی تیلور.

نتیجه‌گیری کلی

نتایج به‌دست آمده از این تحقیق حاکی از عملکرد قابل قبول روش‌های مورد استفاده در پیش‌بینی مقادیر رسوب معلق بود. مقایسه نتایج مدل‌های SVR، GMDH و SRC حاکی از برتری مدل SVR و GMDH در پیش‌بینی مقادیر رسوب معلق نسبت به الگوی ورودی شماره ۶ می‌باشد. نتایج بیانگر عملکرد قابل قبول مدل GMDH با بیشترین ضریب تعیین R^2 برابر ۰/۹۹ و کمترین ریشه میانگین مربعات خطا ۸۳ کیلوگرم در روز شد. با توجه به نتایج حاصله می‌توان گفت، مدل GMDH به‌عنوان مدلی توانمند و با سرعت بالا می‌تواند جهت مدل‌سازی رسوب معلق مورد استفاده قرار گیرد. نتایج تحقیق نشان داد که هر دو روش داده‌کاوی به مراتب کارایی و دقت عمل بهتری در زمینه تخمین بار معلق رسوب رودخانه نسبت به منحنی سنجه رسوب دارند. می‌توان از روش‌های مبتنی بر داده‌کاوی به‌عنوان جایگزین برای تخمین بار معلق رودخانه استفاده کرد.

باید توجه داشت که با توجه به تغییر اقلیم و خشکسالی‌ها، توسعه صنعتی، تغییر مستعمر کاربری اراضی و تغییرات مورفولوژی حوضه‌های آبریز نتایج به‌دست آمده را نمی‌توان برای همیشه در هر زمانی استفاده کرد بلکه بایستی در مواقع استفاده از مدل‌ها شرایط به روزرسانی شود. از دیگر نقاط ضعف مدل‌ها این است که با افزایش تعداد لایه‌های توسعه داده شده دقت پاسخ‌های تولید شده اضافه، ولی روابط تولید شده بین متغیرهای ورودی و خروجی بسیار پیچیده می‌شود. در اکثر مطالعات انجام شده از یک یا دو پارامتر ورودی دبی جریان و بارندگی استفاده شده است. از نقاط قوت این تحقیق استفاده از چندین پارامتر (بارندگی، دما، دبی جریان، دبی رسوب) ورودی می‌باشد. در نهایت استفاده از روش آزمون گاما به‌عنوان یک روش پیش‌پردازش داده‌ها برای انتخاب ترکیباتی از متغیرهای ورودی مناسب به مدل‌ها پیشنهاد می‌شود. همچنین پیشنهاد می‌شود کارایی ورودی‌ها این تحقیق با

شود. همچنین، استفاده از الگوریتم‌های فرا ابتکاری (به-عنوان مثال الگوریتم ژنتیک) در تنظیم پارامترهای مدل-SVR می‌تواند دقت مدل‌سازی برآورد رسوب معلق را افزایش دهند. این بهبود در پیش‌بینی رسوب معلق برای برنامه‌ریزی و مدیریت منابع آب ارزشمند می‌باشد.

استفاده از سایر مدل‌های هوشمند (نظیر مدل‌های نروفازی، درخت تصمیم‌گیری و غیره) و مقایسه نتایج آنها با پژوهش حاضر و نیز کاربرد آن در آبریزهای دیگر مورد بررسی قرار گیرد. در راستای تکمیل پژوهش صورت گرفته بهتر است، علاوه بر متغیرهای هیدرولوژیکی و اقلیمی حوضه (بارندگی و دما)، از متغیرهای دیگر به‌عنوان ورودی‌های مدل‌ها استفاده

منابع مورد استفاده

- Adnan RM, Yaseen ZM, Heddami S, Shahid S, Sadeghi-Niaraki A and Kisi O, 2022. Predictability performance enhancement for suspended sediment in rivers: Inspection of newly developed hybrid adaptive neuro-fuzzy system model. *International Journal of Sediment Research* 37(10): 383-398. <https://doi.org/10.1016/j.ijsrc.2021.10.001>
- Alp M and Cigizoglu HK, 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling and Software* 22: 2-13.
- Aytek A and Kisi O, 2007. A genetic programming approach to suspended sediment modeling. *Journal of Hydrology* 351: 288-298.
- Biranvand N, Sepahvand A and Haqizadeh A, 2022. Modeling of suspended sediment using machine learning algorithms in periods of low and high water (Case study: Kashkan watershed). *Soil and Water Modeling and Management* (3)2: DOI: 10.22098/mmws.2022.11262.1115. (In Persian with English Abstract).
- Chen XY and Chau KW, 2016. A Hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation, *Water Resource Management*. 30(7): 2179-2194. <https://doi.org/10.1007/s11269-016-1281-2>
- Cigizoglu HK and Kisi O, 2006. Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology* 317: 221-238.
- Eivani Z, Ahmadi MM and Qaderi K, 2016. Estimation of suspended sediment load concentration in river system using Group Method of Data Handling (GMDH). *Journal of Watershed Management Research* 7(13):218-229. (In Persian with English Abstract).
- Hesavi M and Shafae-Bajestan M, 2010. Estimate of Sediment Bed Load in Karoon River Ahwaz Station. 8th International River Engineering Conference, Date 26 June 2010, Ahwaz, Iran. (In Persian with English Abstract).
- Ivakhnenko AG, 1968. The Group Method of Data Handling-a Rival of the Method of Stochastic Approximation, *Soviet Automatic Control c/c of Avtomatika* 1: 43-55.
- Jie LC and Yu ST, 2011. Suspended sediment load estimate using support vector machines in Kaoping River basin. *International Conference on Consumer Electronics, Communications and Networks (CECNet)*. 16-18 April .doi: [10.1109/CECNET.2011.5769267](https://doi.org/10.1109/CECNET.2011.5769267)
- Keshtegar B, Piri J, Hussan WU, Ikram K, Yaseen M, Kisi O, Adnan RM, Adnan M and Waseem M, 2023. Prediction of sediment yields using a Data-Driven Radial M5 Tree Model. *Journal of Water*, 15: 1437. <https://doi.org/10.3390/w15071437>
- Kisi O and Shiri J, 2012. River Suspended sediment estimation by climate variables implication: Comparative study among soft computing techniques. *Computer and Geosciences* 43: 73-82.
- Kisi O, Ozkan C and Bahriye A, 2012. Modeling discharge-sediment relationship using neural networks with Artificial Bee Colony Algorithm. *Journal of Hydrology*: 94-103.
- Mehrizi Haeri AA, 2012. Data mining: concepts, methods and applications. Master's thesis in economic and social statistics, Faculty of Economics, Allameh Tabatabai University. (In Persian with English Abstract).

- Moradi-Nejad A, Davoudmaqami D and Moradi M, 2018. Investigating the efficiency of methods for estimating the suspended sediment load of Qarachai river. *Environment and Water Engineering* (4)5: <https://doi.org/10.22034/jewe.2020.211925.1341>. (In Persian with English Abstract).
- Moradi-Nejad A, Khosrobigi S, Akbari M and Hosseini SA, 2023. Evaluation of soft calculation methods in estimation of river suspended sediment (Hassanabad station of Tirah River). *Soil and Water Modeling and Management*. <https://doi.10.22098/mmws.2023.12620.1258>. (In Persian with English Abstract).
- Nikpour MR and Sani Khani H, 2016. Modeling of river suspended sediments using soft calculations (Darah-Rood River). *Irrigation and Water Engineer's Scientific and Research Quarterly*. 8(30):29-44. <https://civilica.com/doc/888603>. (In Persian with English Abstract).
- Qobadian R and Shokri H, 2018. Numerical investigation of factors affecting the distribution of unbalanced sediment concentration in natural rivers (Case study: Qarasu River, Kermanshah). *Journal of Water and Soil* 2(34): 241-253. DOI: 10.22067/jsw.v34i2.7632. (In Persian with English Abstract).
- Safavi H, 1390 *Engineering Hydrology*, Arkan Danesh Publications, 3rd edition, Isfahan. 699P.
- Satari MT, Rezazadeh Jodi A, Safdari F and Kahramanzadeh F, 2016. Performance evaluation of M5 tree model and support vector regression methods in river suspended sediment modeling. *Protection of Water and Soil Resources* 6(1):109-124. <https://civilica.com/doc/1297623>. (In Persian with English Abstract).
- Sheikh Alipour Z, Hasanpour F and Bigi M, 2014. Comparison of artificial intelligence methods in estimating suspended sediment load (Case study: Sistan River). *Water and Soil Conservation Research* 7(2): 41-60. (In Persian with English abstract).
- Zakaria N, Azamathulla HMD, Chang CHK and Ghani AAB, 2010. Gene expression programming for total bed material load estimation-a case study. *Science of the Total Environment*, 408: 5078-5085.
- Zhou Y, Lu XX, Huang Y and Zhu YM, 2007. Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the upper Yangtze catchment, China. *Geomorphology* 84: 111-125.