

Suspended sediment load modeling by machine learning algorithms in low and high discharge periods (Case study: Kashkan watershed)

Nasrin Beiranvand¹ , Alireza Sepahvand^{2*} , Ali Haghizadeh³ 

¹ M.Sc. Student, Department of Range and Watershed Management Engineering, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

² Assistant Professor, Department of Range and Watershed Management Engineering, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

³ Associate Professor, Department of Range and Watershed Management Engineering, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

Abstract

Introduction

Sediment that moves with water is called suspended sediment, and the amount of suspended sediment material that passes through a river section in a certain period of time is called suspended load. The suspended sediment load (SSL) of a watershed, which passes through a certain section of the river, depends mainly on the climatic characteristics, the characteristics of the watershed, and the capacity to carry sedimentary materials. Actually, suspended sediment transport in the river is a function of meteorological and hydrological parameters as a complicated process. The input suspended load is one of the important and influencing factors on the amount of sediment input to reservoirs of dams and lakes. Determining the amount of sediment carried by rivers is important in many aspects. The calculation of suspended load is very important because of various reasons, one of the most important reasons is the role of suspended sediment load in the quantitative and qualitative management of surface water resources. Therefore, the distribution and transportation of suspended sediment load (SSL) in rivers have a significant effect on water resource management, design of hydraulic structures, river morphology, water quality, and aquatic ecosystems. In fact, accurate and reliable modeling of suspended sediment load (SSL) is very important for planning, managing, and designing of river systems and water resource structures. In addition, the determination of dry and wet periods is very important in studies related to water resources management, especially in arid and semi-arid regions.

Materials and Methods

To compare the result of the proposed models' performance, the Cham Anjir, Bahram Jo, Kaka Reza, and Sarab Syed Ali hydrometry stations in Khorramabad, Biranshahr and Alashtar sub-watersheds (a part of Kashkan watershed) in western of Iran, is used as a case study area. The geographic coordinates of the Cham Anjir, Bahram Jo, Kaka Reza and Sarab Syed Ali are 48° 15' 34" E 33° 26' 55" N, 48° 17' 45" E 33° 34' 8" N, 48° 13' 51" E 33° 43' 39" N and 48° 12' 14" E 33° 44' 55" N, respectively. The studied area has a semiarid climate with a mean annual rainfall of less than 500 mm. The studied area has a maximum elevation of 3578 m in the Alashtar watershed and a minimum elevation of 1158 m in the KhorramAbad watershed. Most parts of the studied sub-watersheds are rangeland, while forest, dry farming, and irrigation lands are in considerable quantities. The surface lithology in the KhorramAbad, Alashtar, and Biranshahr watersheds are covered by the Eocene, Quaternary, Cretaceous, Miocene, Oligocene, Paleocene, and Pliocene geologic formations. Predicting suspended sediment load (SSL) in water resource management requires efficient and reliable predicted models. The present study was carried out for the modeling of Suspended sediment load by learning algorithms in low and high discharge periods. In this study, five soft computing techniques, GP-PUK, GP-RBF, M5P, REEP Tree, and RF were used to predict the SSL in Cham Anjir, Bahram Jo, Kaka Reza, and Sarab Syed Ali hydrometry stations in Khorramabad, Biranshahr and Alashtar sub-watersheds. The total data set consists of rain, discharge, and suspended sediment load (in a period of 18 years from 2000 to 2018) of three sub-watersheds out of which 70% of data was used to train the model and 30% of data were used to test the model. Finally, the model's accuracy was assessed using three performance evaluation parameters, which were Correlation Coefficient

(C.C.), Root Mean Square Error (RMSE), and Maximum Absolute Error (MAE). Finally, a sensitivity investigation was executed to catch the best noteworthy input parameter during the modeling process. This process was carried out by eliminating the one input parameter and noting the output in terms of RMSE and C.C.

Results and Discussion

The obtained results suggest that the Gaussian Process (GP) model with two PUK and RBF kernels is more accurate in estimating the suspended sediment load (SSL) compared to the M5P, ReepTree, and Random Forest (RF) models for the given study area. The results of the test part of the GP-PUK model have given us the best result, which are the correlation coefficient, the root mean square error and the mean absolute error in Bahram Jo station (0.55, 0.42, and 0.27), Cham Anjir station (0.74, 0.18, and 0.08), Sarab Seyed Ali station (0.71, 0.16, and 0.07) and Kakareza station (0.71, 0.24, and 0.15), respectively. In general, the Gaussian Process-PUK model is a powerful model for the prediction of suspended sediment load (SSL) in low and high discharge periods. Therefore, according to the obtained results from this research, these optimal models can be used for costly and time-consuming tasks of the estimation of suspended sediment load from rivers. In addition, these models can be used to estimate the suspended sediment load of nearby rivers by/without a hydrometry station for the management of the quantity and quality of surface water. Also, sensitivity analysis suggests that $AverageQ_t, Q_{t-1}$ in Cham Anjir, Bahram Jo, and Sarab Syed Ali hydrometry stations and rain in Kaka Reza hydrometry station are the most significant parameters in the estimation/prediction of SSL.

Conclusion

The present study focused on the development of GP-PUK, GP-RBF, M5P, REEP Tree, and Random Forest (RF) models to estimate the suspended sediment load. For this purpose, the hydrometry and hydroclimatology data of the Bahram Jo, Cham Anjir, Sarab Syed Ali, and Kaka Reza stations in Khorramabad, Alashtar, and Biranshahr sub-watersheds composed of suspended sediment load (SSL), discharge, and rainfall data were used. In general, the major conclusions of the study are as follows: Among those models with the highest performance, the GP-PUK has the highest performance in both the testing and training phases. The GP-PUK predicted data are closer to observational data compared with the other model's output data. Besides, the GP-PUK is the nearest predicted model with observational data. The GP-PUK model is one of the most extensively used data-driven models in the erosion and sediment literature, while the usages of other data-driven models are comparatively lesser. Also, the structure of the GP-PUK is very simple and much less time-consuming. Thus, the GP-PUK model can be useful in the Suspended Sediment Load (SSL) modeling not only for accuracy but also for its time-saving nature and simple structure compared with other models.

Keywords: Flow Duration Curve (FDC), Gaussian process, Karkhe Watershed, Lorestan province, Random Forest

Article Type: Research Article

*Corresponding Author, E-mail: sepahvand.a@lu.ac.ir

Citation: Beiranvand, N., Sepahvand, A., & Haghi Zadeh, A. (2023). Suspended sediment load modeling by machine learning algorithms in low and high discharge periods (Case study: Kashkan watershed). *Water and Soil Management and Modeling*, 3(2), 50-65.

DOI: 10.22098/mmws.2022.11262.1115

DOR: 20.1001.1.27832546.1402.3.2.4.8

Received: 07 August 2022, Received in revised form: 28 August 2022, Accepted: 28 August 2022, Published online: 28 August 2022

Water and Soil Management and Modeling, Year 2023, Vol. 3, No.2, pp. 50-65

Publisher: University of Mohaghegh Ardabili

© Author(s)





مدل‌سازی رسوب معلق با استفاده از الگوریتم‌های یادگیری ماشین در دوره‌های کم‌آبی و پرابی (مطالعه موردی: حوزه آبخیز کشکان)

نسرین بیرانوند^۱، علیرضا سپهوند^{۲*}، علی حقی‌زاده^۳

^۱ دانشجوی کارشناسی ارشد، گروه مهندسی مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
^۲ استادیار، گروه مهندسی مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
^۳ دانشیار، گروه مهندسی مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران

چکیده

در این پژوهش با استفاده از الگوریتم‌های یادگیری به بررسی کارایی مدل‌های RF، RepTree، GP-PUK، GP-RBF و M5P برای مدل‌سازی بار معلق رودخانه در استان لرستان شامل حوزه‌های آبخیز خرم‌آباد، بیرانشهر و الشتر پرداخته شد. برای انجام این کار از داده‌های ورودی بارش، دبی، دبی یک روز قبل و میانگین دبی و دبی یک روز قبل هم‌چنین داده خروجی رسوب معلق در بازه زمانی ۱۸ ساله (سال‌های ۷۹-۸۰ تا ۹۶-۹۷) استفاده شد. با استفاده از داده‌های در دسترس منحنی تداوم جریان و منحنی سنج رسوب را به دست آورده سپس با استفاده از داده‌های دبی برای هر ایستگاه حد تعیین دوره کم‌آبی و دوره پرابی مشخص شد، سپس رسوب معلق به دو دوره رسوب معلق کم‌آبی و پرابی تقسیم شد، سپس مدل‌سازی داده‌ها (۷۰ درصد داده‌های آموزش و ۳۰ درصد داده‌های آزمایش) با استفاده از مدل‌های ذکر شده انجام شد. نتایج نشان داد باتوجه به معیارهای ارزیابی مدل GP با دو کرنل PUK و RBF در دوره کم‌آبی و پرابی عملکرد بهتری را نسبت به سایر مدل‌ها (RF، RepTree، M5P) داشته است. با توجه به نتایج بخش آزمایش مدل GP-PUK بهترین نتیجه را به ما داده است که به ترتیب ضریب همبستگی، ریشه میانگین مربعات خطا و میانگین خطای مطلق در ایستگاه بهرام‌جو ۰/۵۵، ۰/۴۲ و ۰/۲۷، هم‌چنین ایستگاه چمانجیر ۰/۷۴، ۰/۱۸ و ۰/۰۸، در ایستگاه سراب صیدعلی ۰/۷۱، ۰/۱۶ و ۰/۰۷ و در آخر ایستگاه کاکارضا ۰/۷۳، ۰/۲۴ و ۰/۱۵ به دست آمده است. در مجموع مدل GP-PUK به عنوان مدل برتر، قدرت بالاتری برای مدل‌سازی همه ایستگاه‌ها در رسوب معلق دوره پرابی و کم‌آبی در بخش آزمایش بوده است. لذا با توجه به نتایج به دست آمده از این پژوهش می‌توان از این مدل‌های بهینه برای صرفه‌جویی در هزینه و زمان برای بحث حفاظت آب و خاک و تخمین رسوب معلق خروجی از حوزه‌های آبخیز استفاده کرد. هم‌چنین می‌توان برای اجرای مدیریت بهتر در رابطه با کمیت و کیفیت آب‌های سطحی، این مدل‌ها برای تخمین رسوبات معلق ایستگاه‌های مجاور فاقد آمار دارای شرایط زمین‌ساختی و هیدرولوژیکی یکسان در سطح منطقه مورد استفاده قرار گیرند و نتایج قابل اعتمادی در رابطه با رسوب معلق ارائه دهند.

واژه‌های کلیدی: استان لرستان، حوضه کرخه، منحنی تداوم جریان، جنگل تصادفی، فرآیند گوسی

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: sepahvand.a@lu.ac.ir

استاد: بیرانوند، نسرین، سپهوند، علیرضا، و حقی‌زاده، علی (۱۴۰۲). مدل‌سازی رسوب معلق با استفاده از الگوریتم‌های یادگیری ماشین در دوره‌های کم‌آبی و پرابی (مطالعه موردی: حوزه آبخیز کشکان). *مدل‌سازی و مدیریت آب و خاک*، ۳(۲)، ۵۰-۶۵.
DOI: 10.22098/mmws.2022.11262.1115
DOR: 20.1001.1.27832546.1402.3.2.4.8

تاریخ دریافت: ۱۴۰۱/۰۵/۱۶، تاریخ بازنگری: ۱۴۰۱/۰۶/۰۶، تاریخ پذیرش: ۱۴۰۱/۰۶/۰۶، تاریخ انتشار: ۱۴۰۱/۰۶/۰۶



مدل‌سازی و مدیریت آب و خاک، سال ۱۴۰۲، دوره ۳، شماره ۲، صفحه ۵۰ تا ۶۵
ناشر: دانشگاه محقق اردبیلی
© نویسندهگان

۱- مقدمه

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 (Qobadian and Shokri, 2018). رسوب به دو صورت بار معلق و بار بستر انتقال پیدا می‌کند. رسوباتی که همراه با آب در حرکت باشند را رسوب معلق می‌گویند (Shojaeezadeh et al., 2018). فرآیند انتقال رسوب در رودخانه‌ها از نظر زمانی و مکانی یک فرآیند غیرخطی و پیچیده است، به همین دلیل محاسبه غلظت رسوب معلق امری ضروری در مدیریت منابع آب است (Rahul et al., 2021). بار رسوبی رودخانه از مهم‌ترین پارامترها در مدیریت پروژه‌های هیدرولیکی، شاخص وضعیت خاک و فرسایش حوزه آبخیز است (Asadi and Fathzadeh, 2017). همچنین، مدل‌سازی رسوبات معلق در مهندسی محیط زیست و منابع آب بر ظرفیت مخزن و عملکرد سد اهمیت زیادی دارد. انتقال رسوب معلق به‌عنوان یک فرآیند غیرخطی و پیچیده در رودخانه تابعی از پارامترهای هواشناسی و هیدرولوژیکی است (Nourani et al., 2020).

تعیین توالی دوره‌های خشک و مرطوب دارای اهمیت زیادی در مطالعات مرتبط با مدیریت منابع آب به‌خصوص در مناطق خشک و نیمه‌خشک است. به‌دلیل پیچیدگی‌های موجود در پدیده‌های هیدرواقیمی، درک آن در مقایسه با سایر پدیده‌های هیدرواقیمی دشوارتر بوده (Ansari and Davari, 2008; Mostafazadeh and Zabih, 2016) و مسلماً به اطلاعات بیش‌تر و کامل‌تر به‌منظور مدیریت معضلات و مشکلات منابع آبی ناشی از وقوع دوره‌های خشک در یک منطقه نیاز است. از جمله کاربردهای تعیین دوره‌های مرطوب می‌توان به تولید آب و برنامه‌ریزی کشت، توصیف کمی خشکسالی، سیل‌خیزی و پیش‌بینی و ارزیابی وقوع سیل‌ها اشاره کرد (Bazuhair et al., 1997).

Sepahvand and Azizi Najafkali (2019) به مدل‌سازی رسوب معلق با استفاده از مدل‌های فرآیند گوسی و پرسپترون چند لایه پرداختند. در این پژوهش برای پیش‌بینی رسوب معلق به مقایسه کارایی مدل‌های GP-RBF, GP-PUK, MLP پرداختند. نتایج نشان داد که مدل GP با کرنل RBF در مرحله آموزش و همچنین مرحله آزمایش به‌عنوان بهترین تخمین‌گر برای تخمین رسوب معلق در ایستگاه هیدرومتری مورد مطالعه انتخاب شده است. (Khalighi Sigarodi et al., 2009). در پژوهشی با هدف تحلیل آماری بارش و تعیین سال‌های مرطوب و خشک در ایستگاه‌های استان مازندران به این نتیجه دست یافتند که شاخص SPI به‌دلیل تفکیک دقیق‌تر طبقه‌ها در هر یک از پدیده‌های کم‌آبی و پرآبی و همچنین دقت بالاتر در جداسازی دوره‌های مرطوب و خشک، به تغییرات بارش حساسیت بیش‌تری دارد.

Shojaeezadeh et al. (2018) به مدل‌سازی تصادفی رسوب معلق در رودخانه‌های آبرفتی پرداختند. در این پژوهش یک مدل احتمالی مختصر برای توصیف رابطه بین بار رسوب معلق و حجم تخلیه معرفی شده است. این چارچوب برای هفت رودخانه اصلی در ایالت متحده آزمایش شد و نتایج آن عملکرد امیدوارکننده برای پیش‌بینی SSL و احتمال آن با توجه به سطوح دبی مختلف نشان داده است.

Doroudi et al. (2021) به برآورد بار رسوب معلق روزانه با استفاده از مدل رگرسیون بردار پشتیبان ترکیبی یک‌پارچه با روش بهینه‌سازی مشاهده‌کننده (معلم) یادگیرنده در حوزه آبخیز رودخانه چم‌سیاه واقع در استان کهگیلویه و بویراحمد پرداختند. در این پژوهش چون SVR دارد که پارامترهای ناشناخته روش بهینه‌سازی مبتنی بر مشاهده‌کننده با مدل SVR ادغام شده است و یک مدل ترکیبی جدید ارائه می‌کند. مدل‌های پیش‌بینی با استفاده از شاخص‌های عددی مختلف ارزیابی می‌شوند. نتایج نشان داد که مدل SVR-OTLBO عملکرد بهتری نسبت به سایر مدل‌ها نشان داده است. (Nourani et al., 2020). به برآورد بار رسوب معلق با استفاده از مدل یک‌پارچه مبتنی بر هوش مصنوعی در حوزه آبخیز رودخانه کاتار پرداختند. در این پژوهش از سه مدل مبتنی بر هوش مصنوعی SVM، ANFIS و FFNN و یک حالت رگرسیون چندخطی MLR برای مدل‌سازی SSL استفاده شده است. برای ارزیابی عملکرد مدل‌ها از ضریب تعیین و ریشه میانگین مربعات خطا استفاده شده است. بر اساس نتایج مدل ANFIS کارایی بالاتر نسبت به سایر مدل‌ها از خود نشان می‌دهد. با توجه به بررسی پژوهش‌های انجام شده، مشخص شد که مدل‌های مختلفی مبتنی بر الگوریتم‌های یادگیری توسعه یافته‌اند که در موارد مختلف مدل‌سازی و پیش‌بینی مورد استفاده قرار می‌گیرند. لذا، هدف این پژوهش بررسی کارایی برخی الگوریتم‌های یادگیری برای مدل‌سازی بار معلق رودخانه‌های خرم‌آباد، کاکارضا و الشتر با تاکید بر دوره‌های کم‌آبی و پرآبی است که در صورت برآورد میزان بار معلق رودخانه‌ها، می‌توان با یک مدیریت بهینه فرسایش و رسوب بخشی از خسارت‌ها را کاهش داد. به‌دلیل این‌که تعیین بار معلق رودخانه معمولاً با صرف زمان و هزینه زیاد همراه است، لذا بهتر است با انتخاب یک مدل با دقت زیاد، در رودخانه‌هایی که دارای ایستگاه هیدرومتری نیستند، میزان رسوب معلق را تعیین کرد.

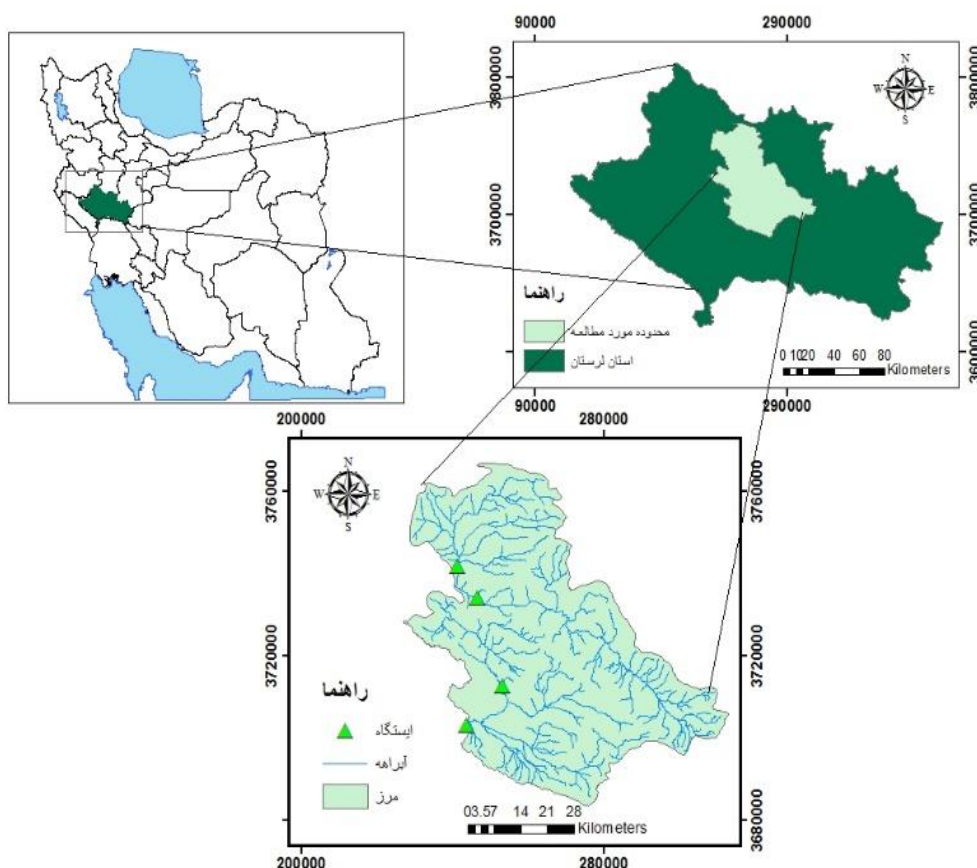
۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

حوزه‌های آبخیز خرم‌آباد، الشتر و بیرانشهر از نظر ژئومورفولوژیکی در بخش مرکزی سلسله جبال زاگرس در استان لرستان قرار

خروجی حوزه آبخیز با ارتفاع ۱۱۵۸ متر از سطح دریا قرار دارد. میانگین بارندگی سالانه منطقه بین ۴۰۰ تا ۵۰۰ میلی‌متر بوده و دارای اقلیم نیمه‌خشک و سرد است (Sepahvand et al., 2019). در جدول ۱ مشخصات جغرافیایی ایستگاه‌های مختلف حوزه آبخیز کرخه نشان داده شده است.

دارند که خود بخشی از حوزه آبخیز کرخه هستند. حوزه‌های آبخیز خرم‌آباد، الشتر و بیرانشهر در مجموع با مساحت ۳۵۶۲/۱۲ کیلومترمربع مطابق شکل ۱ بین طول‌های جغرافیایی ۰۳° ۰۴' تا ۰۷° ۵۹' ۴۸' شرقی و عرض‌های ۰۵' ۱۶' ۳۳" تا ۰۱' ۳۴" شمالی قرار دارند. بلندترین نقطه در بخش شمالی با ارتفاع ۳۵۷۸ متر و پست‌ترین نقطه در بخش غربی، یعنی



شکل ۱- نقشه موقعیت جغرافیایی حوزه آبخیز کرخه، ایستگاه‌های خرم‌آباد، الشتر و بیرانشهر در استان لرستان و ایران

Figure 1- Geographical location map of Karkheh basin, Khoramabad, Alshatar and Biranshahr stations in Lorestan province, Iran

جدول ۱- مشخصات جغرافیایی ایستگاه‌های مورد مطالعه

Table 1 - Geographical characteristics of the studied stations

ردیف	نام ایستگاه	طول جغرافیایی	عرض جغرافیایی	حوزه آبخیز	محل
1	بهرام‌جو	47° 45' تا 48° 36'	33° 10' تا 33° 50'	حوزه آبخیز کرخه	قبل از خروجی
2	چمانجیر	48° 12'	33° 52'	حوزه آبخیز کرخه	خروجی
3	کاکارضا	48° 15'	33° 42'	حوزه آبخیز کرخه	خروجی
4	سراب صیدعلی	48° 15'	33° 30'	حوزه آبخیز کرخه	خروجی

روش‌های ارزنده است با اطلاعاتی مفید که همه دبی‌های رودخانه اعم از کم و جریان سیلابی را نمایش می‌دهد. این منحنی رابطه بین مقادیر دبی و درصد زمانی را (که این دبی مساوی یا بیش‌تر از آن است)، به نمایش می‌گذارد. به عبارت دیگر، رابطه بین بزرگی و فراوانی دبی رودخانه را نشان می‌دهد

۲-۲- روش کار

نخستین مطالعه جامع بر منحنی تداوم جریان را سارسی^۱ در سال ۱۹۵۹ میلادی انجام داد. منحنی تداوم جریان (FDC)^۲ یکی از

¹ Searcy

² Flow Duration Curve

$$\text{IGR}(X, S) = \frac{E(S) = \sum_{i=1}^n x = \frac{E(S) |S_i|}{|S|}}{- \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Log}2 \frac{|s_i|}{|s|}} \quad (1)$$

۲-۳-۲-۳ مدل درختی M5P

یکی از رایج‌ترین الگوریتم‌های مدل درختی، الگوریتم M5 است. این الگوریتم اولین بار توسط کوئینلان در سال ۱۹۹۲ پیشنهاد شد (Quinlan, 1992). سپس الگوریتم M5P که در واقع بازسازی منطقی و توسعه‌یافته از M5 است، توسط ونگ و ویتن در سال ۱۹۹۷ معرفی شد (Wang and Witten, 1997). این مدل‌ها یکی از روش‌های نسبتاً جدید و قدرتمند هوش محاسباتی است که به علت قابلیت مناسب آن، توسط محققان مختلفی مورد استفاده قرار گرفته است. سادگی محاسبه‌ها و دقت نتایج از مهم‌ترین ویژگی‌های این مدل است (Bonakdar et al., 2011). کاهش انحراف معیار مطابق رابطه (۲) محاسبه شده است (Wang and Witten, 1997).

$$\text{SDR} = \text{sd}(T) - \sum \frac{|T_i|}{|T|} \text{sd}(T_i) \quad (2)$$

۲-۲-۲-۴ مدل‌های رگرسیون فرآیند گوسی^۲

مدل‌های رگرسیون فرآیند گوسی بر مبنای این فرض هستند که مشاهدات تنظیم باید دارای اطلاعاتی درباره یکدیگر باشد. این فرآیند راهی برای مشخص کردن اولویتی به صورت مستقیم روی فضای تابع هستند (Yang et al., 2008). مدل‌های فرآیند گوسی به دلیل دانش قبلی درباره وابستگی‌های تابعی و داده‌ها، برای تعمیم به هیچ فرآیند اعتبارسنجی نیاز ندارند و قادر به درک توزیع پیش‌بینی متناظر با ورودی آموزش‌ها هستند (Pal and Deswal, 2010). فرآیند گوسی تعمیمی از توزیع گوسی است. توزیع گوسی در واقع توزیع بین متغیرهای تصادفی بوده، درحالی‌که فرآیند گوسی بیانگر توزیع بین توابع است. فرآیند گوسی $f(x)$ می‌تواند به صورت رابطه (۳) بیان شده است:

$$f(x) \sim \text{GP}(m(x), k(x, x')) \quad (3)$$

۲-۲-۲-۵ مدل جنگل تصادفی^۳

الگوریتم RF نخستین بار توسط Breiman ارائه شد و به‌عنوان روشی دقیق در تهیه نقشه کاربری اراضی و طبقه‌بندی تصاویر ماهواره‌ای استفاده شده است (Pal, 2005; Prasad et al., 2006). رگرسیون RF شامل ترکیبی از متغیرها در هر گره برای رشد یک درخت یا استفاده از متغیر ورودی تصادفی انتخاب شده است (Breiman, 1996). برای رگرسیون RF دو پارامتر لازم

(Smakhtin, 2001). بر این اساس در این پژوهش با توجه به هدف پژوهش مبنی بر مدل‌سازی رسوب معلق در دوره‌های پربابی و کم‌آبی رودخانه‌ها، از منحنی تداوم جریان برای تفکیک دوره‌های پربابی و کم‌آبی استفاده شد. بر این اساس منحنی تداوم جریان مقدار جریانی که در ۹۱ روز از سال دبی مساوی و یا بیش‌تر از آن باشد را به‌عنوان دبی نرمال در حالت پربابی و رسوبات متناظر این دبی‌ها را نیز استخراج نموده و به‌عنوان رسوب معلق دوره پربابی و در حالت کم‌آبی نیز دبی که ۲۷۴ روز از سال دبی مساوی یا آن بیش‌تر از آن باشد را به‌عنوان دبی در حالت کم‌آبی و رسوبات متناظر با این دبی‌ها به‌عنوان رسوبات معلق دوره کم‌آبی تعیین شدند.

۲-۲-۱-۲ مجموعه داده‌های استفاده شده برای مدل‌سازی

داده‌های استفاده شده برای مدل‌سازی در این پژوهش شامل دو گروه داده‌های ورودی و داده‌های خروجی می‌باشند. داده‌های ورودی در این پژوهش شامل، دبی جریان، دبی یک روز قبل، میانگین دو دبی و بارش و داده‌های خروجی شامل مقادیر رسوب معلق در بازه زمانی ۱۸ ساله (از سال ۷۹-۸۰ تا ۹۶-۹۷) رودخانه است. همچنین در این پژوهش از ۷۰ درصد کل داده‌ها برای آموزش و از ۳۰ درصد باقی‌مانده برای آزمایش مدل استفاده شد. در این پژوهش از مدل‌های RF, RepTree, GP-PUK, GP- RBF, M5P برای مدل‌سازی با استفاده از نرم‌افزار Weka استفاده شد. مدل‌های مبتنی بر یادگیری به لحاظ به‌کارگیری الگوریتم‌های تخمین و توزیع، طبیعت داده محور و تکرار بالایی فرآیند مدل‌سازی، توانایی بالایی در شناسایی رفتار وقوع پدیده مورد نظر داشته و در پژوهش‌های علمی زیادی برتری خود را نسبت به دیگر مدل‌های آماری دو متغیره و چند متغیره اثبات نموده‌اند که در ادامه توضیح مختصری در مورد این مدل‌ها آورده شده است.

۲-۲-۲-۲ مدل REP Tree

مدل REPTree، یک مدل از مجموعه درخت تصمیم‌گیری (DT) است که برای رفع مشکلات طبقه‌بندی و رگرسیون کارآمد است (Quinlan, 1992). الگوریتم REPTree یک درخت رگرسیون تصمیم‌گیری است که به‌وسیله تقسیم و هرس کردن درخت رگرسیون، تصمیم‌گیری را براساس بالاترین اطلاعات مقدار سود (IGR)، تشکیل می‌دهد، که مقدار IGR بر اساس تابع آنتروپی (E) و مطابق رابطه (۱) محاسبه شده است.

² Gaussian processes

³ Random Forest

¹ Reduced Error Pruning Tree

هر سال در هر یک از ایستگاه یک دبی حد مطابق منحنی تداوم جریان تعیین شد که بعضی ماه‌های سال کم‌تر و بعضی بیش‌تر از دبی حد بودند که دبی‌های کم‌تر از دبی حد دوره کم‌آبی و دبی‌های بیش‌تر از آن در دوره پربابی قرار می‌گیرند.

۳-۲- مدل‌سازی رسوب معلق دوره پربابی

جدول ۲ خصوصیات آماری داده‌های رسوب معلق دوره پربابی مربوط به مرحله آموزش و آزمایش ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد. مطابق نتایج جدول ۲ مشخص شد که به‌ترتیب بیش‌ترین و کم‌ترین رسوب معلق در مرحله آموزش برابر $۶۷۵۵۹/۹۶۶$ و $۰/۴۸۸$ تن در روز و در مرحله آزمایش به‌ترتیب برابر $۲۷۶۰/۹۲۹$ و $۱/۸۵۴$ تن در روز است. جدول ۳ پارامترهای بهینه برای مدل‌های مختلف در ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد. جدول ۴ نتایج عملکرد برای مدل‌های مختلف در ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد. مطابق جدول ۴ مشخص شد که در تمام ایستگاه‌ها، مدل‌ها در بخش آموزش دارای نتایج مطلوبی بودند، اما در بخش آزمایش نتایج تغییرات زیادی داشتند. بر این اساس در ایستگاه بهرام‌جو نتایج مرحله آزمایش نشان داد که مدل GP-PUK با ضریب همبستگی $۰/۵۵$ ، ریشه میانگین مربعات خطا $۰/۴۲$ و میانگین خطای مطلق $۰/۲۷$ در مرحله دارای بهترین عملکرد است. همچنین در ایستگاه چمانجیر با توجه به نتایج بخش آزمایش، مدل GP-PUK با ضریب کارایی $۰/۷۴$ ، ریشه میانگین مربعات خطا $۰/۱۸$ و میانگین خطای مطلق $۰/۰۸$ به‌عنوان بهترین مدل انتخاب شد. در ایستگاه سراب صیدعلی با توجه به نتایج مرحله آزمایش، مدل GP-PUK با ضریب کارایی $۰/۷۱$ ، ریشه میانگین مربعات خطا $۰/۱۶$ و میانگین خطای مطلق $۰/۰۷$ دارای بهترین کارایی بود و بر اساس نتایج جدول ۴ برای ایستگاه کاکارضا مدل GP-PUK با ضریب کارایی $۰/۷۳$ ، ریشه میانگین مربعات خطا $۰/۲۴$ و میانگین خطای مطلق $۰/۱۵$ در مرحله آزمایش دارای بهترین کارایی است.

شکل ۳ نمودارهای داده‌های مشاهداتی و پیش‌بینی بخش آموزش و آزمایش ایستگاه‌های بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا را نشان می‌دهد. شکل ۴ نمودارهای داده‌های خطاهای بخش آموزش و آزمایش ایستگاه‌های بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا را نشان می‌دهد.

توسط کاربر تعریف شود: تعداد متغیرهای ورودی (متر) استفاده شده در هر گره برای تولید یک درخت و تعداد درختان قابل رشد در هر گره، فقط متغیرهای انتخاب شده برای بهترین تقسیم جستجو می‌شوند (Sepahvand et al., 2020).

۲-۲-۶- معیارهای ارزیابی کارایی مدل

در این پژوهش از پارامترهای ریشه میانگین مربعات خطا^۱ (RMSE)، میانگین خطای مطلق^۲ (MAE) و ضریب همبستگی^۳ (CC)، برای ارزیابی دقت مدل‌ها استفاده شد. از بین مدل‌های مذکور مدلی که از ضریب همبستگی بالاتر، ریشه میانگین مربعات خطا و میانگین خطای مطلق نزدیک به صفر برخوردار باشد، به‌عنوان مناسب‌ترین مدل برای تعیین رسوب معلق رودخانه‌های مورد مطالعه انتخاب شد (رابطه‌های ۴ تا ۶).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (H - F)^2} \quad (4)$$

$$MAE = \frac{\sum_{k=1}^K |H_k - F_k|}{K} \quad (5)$$

$$CC = \frac{n \sum HF - (\sum H)(\sum F)}{\sqrt{n(\sum H^2) - (\sum H)^2} \sqrt{n(\sum F^2) - (\sum F)^2}} \quad (6)$$

در روابط فوق، F مقدار برآورد شده، H مقدار مشاهده شده، \bar{H} میانگین مقادیر مشاهده شده و n تعداد داده‌های اندازه‌گیری رسوب معلق است.

در نهایت باتوجه به کاربرد مدل‌های ذکر شده و معیارهای ارزیابی مدل مشخص می‌شود که در ایستگاه‌های مورد نظر و باتوجه به داده‌های استفاده شده کدام یک از مدل‌ها از عملکرد بهتری نسبت به مدل‌های دیگر برخوردارند. به‌طور کلی شکل ۲ نمودار جریانی مراحل پژوهش را نشان می‌دهد.

۳- نتایج و بحث

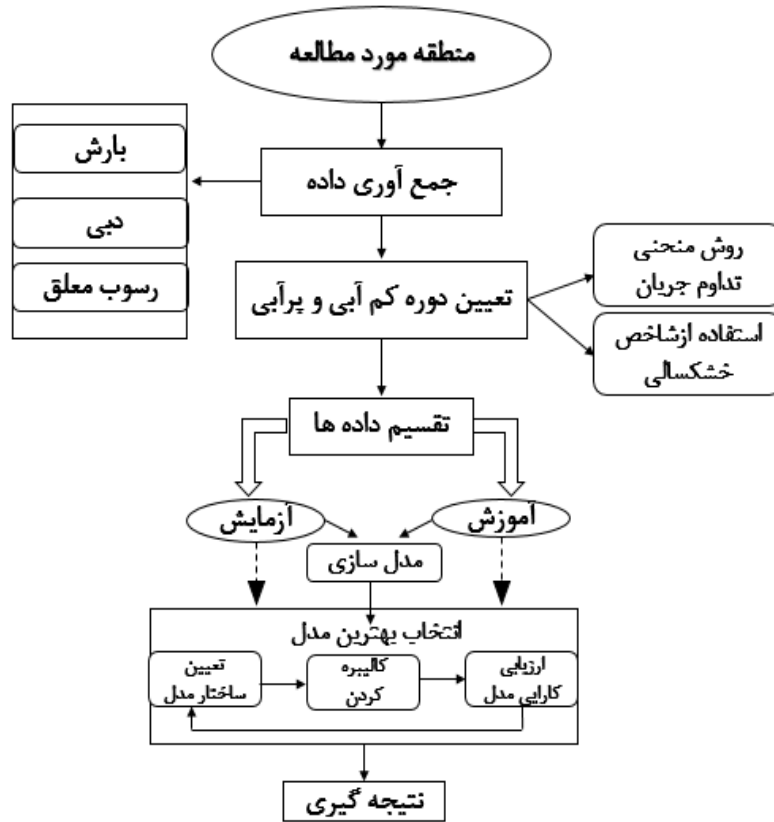
۳-۱- مدل‌سازی رسوب معلق و انحلالی دوره کم‌آبی و پربابی

نتایج حاصل از منحنی تداوم جریان برای تعیین حد کم‌آبی و پربابی هیدرولوژیکی نشان داد که در ایستگاه‌های بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا حدود پربابی و کم‌آبی به‌ترتیب برابر با $۶/۴۴$ ، $۴/۷$ ، $۵/۴۱$ و $۴/۱۳$ مترمکعب بر ثانیه بود که در ادامه تفکیک و تقسیم‌بندی داده‌های دبی جریان برای تعیین رسوب معلق دوره‌های پربابی و کم‌آبی بر اساس این حدود انجام شد. نتایج کلی مدل‌سازی در این پژوهش به دو بخش، نتایج بخش رسوب معلق دوره کم‌آبی و رسوب معلق دوره پربابی تقسیم شدند. تقسیم‌بندی به این صورت است که با توجه به داده‌های دبی برای

¹ Root Mean Square Error

² Mean absolute error

³ Correlation Coefficient



شکل ۲- نمودار جریان‌ی مراحل پژوهش
Figure 2- Flowchart of research steps

جدول ۲- خصوصیات آماری داده‌های آموزش و آزمایش ایستگاه‌های هیدرومتری مورد مطالعه
Table 2- Statistical characteristics of the training and testing data of the studied hydrometric stations

آزمایش				آموزش				داده‌ها
انحراف معیار	میانگین	بیش‌ترین	کم‌ترین	انحراف معیار	میانگین	بیش‌ترین	کم‌ترین	
10.392	42.412	230	0	6.417	31.453	204.5	0	بارش (میلی‌متر)
2.466	11.957	70	2.34	6.479	9.839	144.39	2.91	دبی (مترمکعب در ثانیه)
2.117	9.607	43.658	0.8	6.641	8.863	144.39	0	دبی یک روز قبل (مترمکعب در ثانیه)
1.146	10.782	39.165	2.98	4.401	9.351	78.295	2.18	میانگین هر دو دبی (مترمکعب در ثانیه)
280.651	99.203	2760.929	1.854	3929.406	73.913	67559.966	0.488	رسوب معلق (تن در روز)

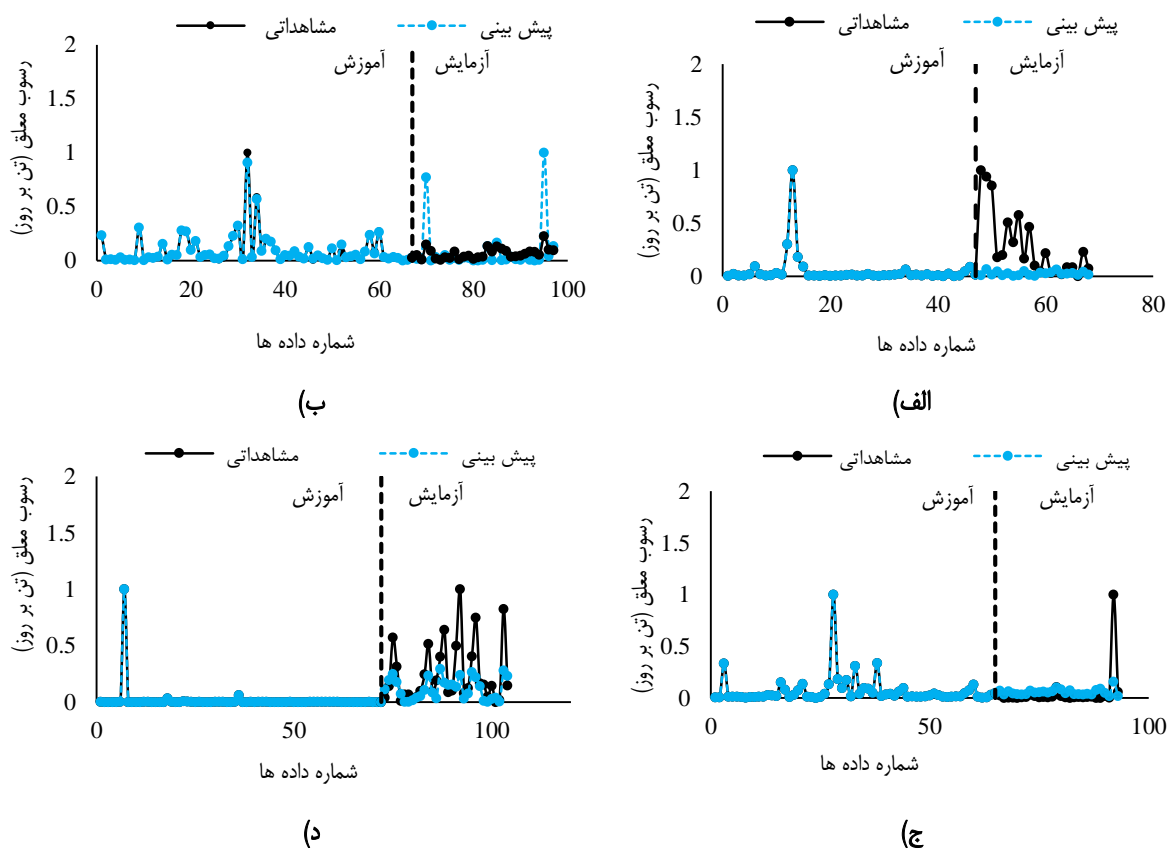
جدول ۳- پارامترهای بهینه مدل‌های مختلف ایستگاه‌های هیدرومتری مورد مطالعه
Table 3- Optimum parameters of different models of studied hydrometric stations

چم‌انجیر	مدل‌ها	بهرام‌جو	مدل‌ها
پارامترهای بهینه	مدل‌ها	پارامترهای بهینه	مدل‌ها
Gaussian noise = 0.1, $\omega=0.1$, $\sigma=5.1$	GP-PUK	Gaussian noise = 0.001, $\omega=0.1$, $\sigma=1.9$	GP-PUK
Gaussian noise = 0.01, $Y=0.5$	GP-RBF	Gaussian noise = 0.001, $Y=0.001$	GP-RBF
num-decimal-places= 0, M=4	M5P	num-decimal-places= 0, M=4	M5P
M=0, V=0, N=3, S=1	REPTREE	M=0, V=0.001, N=3, S=1	REPTREE
K=0, M=1, V=0.001, S=1	RF	K=0, M=1, V=0.001, S=1	RF
کاکارضا		سراب صیدعلی	
Gaussian noise = 0.001, $\omega=0.1$, $\sigma=11.9$	GP-PUK	Gaussian noise = 0.01, $\omega=0.1$, $\sigma=3$	GP-PUK
Gaussian noise = 0.001, $Y=0.001$	GP-RBF	Gaussian noise = 0.01, $Y=0.1$	GP-RBF
num-decimal-places= 0, M=4	M5P	num-decimal-places= 0, M=4	M5P
M=0, V=0.001, N=3, S=1	REPTREE	M=2, V=0.001, N=3, S=1	REPTREE
K=0, M=1, V=0.001, S=1	RF	K=0, M=1, V=0.001, S=1	RF

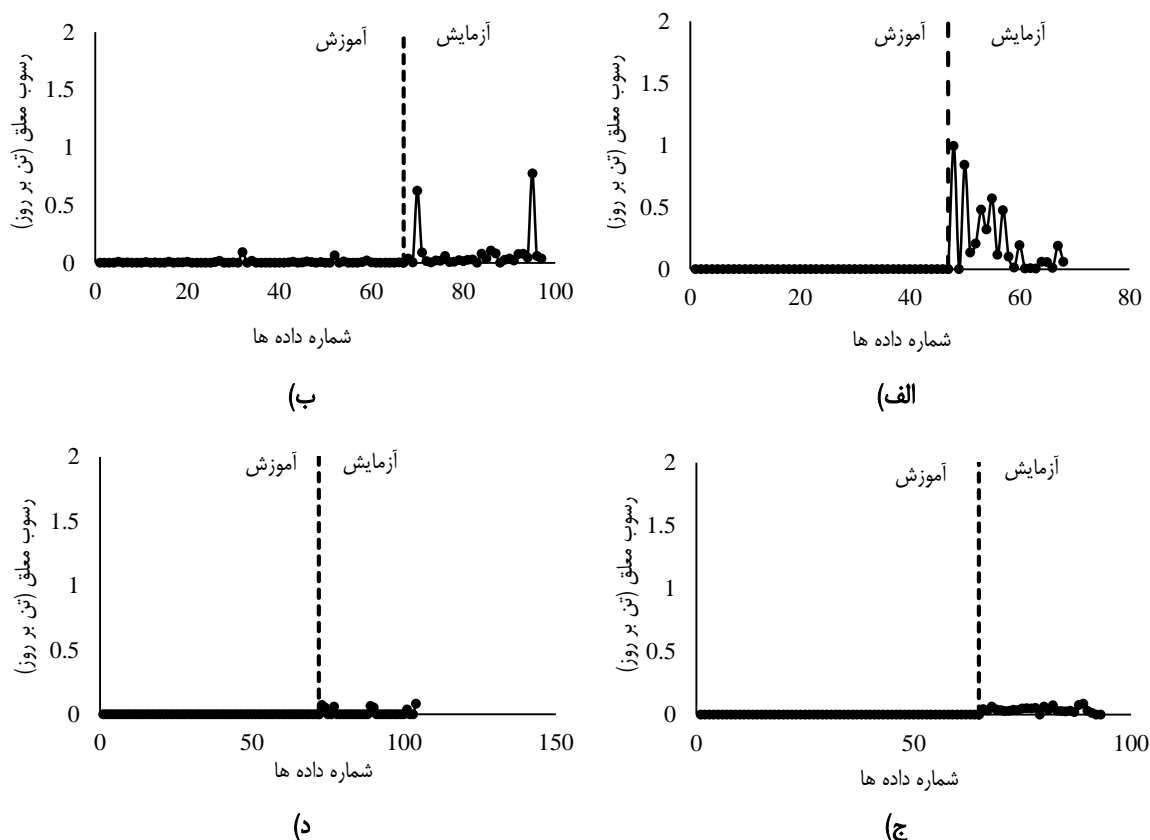
جدول ۴ - پارامترهای ارزیابی عملکرد مدل‌های GP-PUK, GP-RBF, RepTree, M5p, RF ایستگاه‌های هیدرومتری مورد مطالعه

Table 4- Performance evaluation parameters of GP-PUK, GP-RBF, RepTree, M5p, and RF models of studied hydrometric stations

چم‌انجیر						بهرام‌جو						مدل‌ها
آزمایش			آموزش			آزمایش			آموزش			
MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	
0.081	0.187	0.7437	0.005	0.014	0.9972	0.277	0.422	0.553	0.0001	0.0001	0.9999	GP-PUK
0.109	0.204	0.4252	0.057	0.130	0.5291	0.254	0.389	0.2011	0.067	0.145	0.25	GP-RBF
0.094	0.173	0.8632	0.073	0.143	0.3578	0.257	0.401	0.2675	0.057	0.141	0.3522	M5P
0.092	0.192	0.4821	0.052	0.108	0.7089	0.274	0.408	0.221	0.019	0.032	0.9768	REPTREE
0.087	0.184	0.5919	0.022	0.050	0.968	0.261	0.400	0.2554	0.026	0.069	0.9639	RF
کاکارضا						سراب صیدعلی						
0.157	0.241	0.7309	0.0001	0.0001	0.9999	0.07	0.165	0.7158	0.0001	0.0001	0.9999	GP-PUK
0.140	0.198	0.8284	0.002	0.007	0.9981	0.269	0.383	0.4704	0.034	0.089	0.7618	GP-RBF
0.106	0.141	0.8505	0.015	0.038	0.9457	0.046	0.138	0.9951	0.054	0.118	0.5193	M5P
0.183	0.255	0.7542	0.002	0.008	0.9976	0.048	0.151	0.9934	0.047	0.116	0.5381	REPTREE
0.130	0.176	0.8121	0.007	0.046	0.9914	0.061	0.153	0.7655	0.021	0.052	0.9467	RF



شکل ۳- نمودار کارایی مدل GP-PUK داده‌های مشاهداتی و پیش‌بینی در مرحله آموزش و آزمایش، (الف) بهرام‌جو، (ب) چم‌انجیر، (ج) سراب صیدعلی و (د) کاکارضا
Figure 3- Performance chart of GP-PUK model, observation and prediction data in the training and testing phase, (a) Bahram Jo, (b) Cham-Anjir, (c) Sarab Syed Ali, and (d) Kakarza



شکل ۴- نمودار کارایی مدل GP-PUK داده‌های خطا در مرحله آموزش و آزمایش، (الف) بهرام‌جو، (ب) چمانجیر، (ج) سراب صیدعلی و (د) کاکارضا
Figure 4- Performance diagram of GP-PUK model of error data in the training and testing phase, (a) Bahram-Jo, (b) Cham-Anjir, (c) Sarab Syed Ali, and (d) Kakarza

۳-۳- مدل‌سازی رسوب معلق دوره کم‌آبی

جدول ۵ خصوصیات آماری داده‌های رسوب معلق دوره خشک مربوط به مرحله آموزش و آزمایش ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد. مطابق نتایج جدول ۵ مشخص شد که به ترتیب بیش‌ترین و کم‌ترین رسوب معلق در مرحله آموزش برابر ۶۹۸/۷۵۵ و ۰/۳۲۴ تن در روز و در محله آزمایش به ترتیب برابر ۷۰/۰۴ و ۰/۰۱۸ تن در روز است. جدول ۶ پارامترهای بهینه برای مدل‌های مختلف در ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد. جدول ۷ نتایج عملکرد برای مدل‌های مختلف در ایستگاه‌های هیدرومتری مورد مطالعه را نشان می‌دهد.

مطابق جدول ۷ مشخص شد که در تمام ایستگاه‌ها، مدل‌ها در بخش آموزش دارای نتایج مطلوبی بودند، اما در بخش آزمایش نتایج تغییرات زیادی داشتند. بر این اساس در ایستگاه بهرام‌جو نتایج مرحله آموزش نشان داد که مدل GP-PUK با ضریب همبستگی ۰/۷۴، ریشه میانگین مربعات خطا ۰/۳۰ و میانگین خطای مطلق ۰/۱۹ در مرحله دارای بهترین عملکرد است. همچنین، در ایستگاه چمانجیر با توجه به نتایج بخش آزمایش،

مدل GP-PUK با ضریب کارایی ۰/۳۰، ریشه میانگین مربعات خطا ۰/۲۳ و میانگین خطای مطلق ۰/۱۴ از سایر مدل‌ها بهتر شده اما در مجموع هیچ‌یک از مدل‌ها عملکرد خوبی در این ایستگاه نداشته‌اند. در ایستگاه سراب صیدعلی با توجه به نتایج مرحله آموزش، مدل GP-PUK با ضریب کارایی ۰/۶۷، ریشه میانگین مربعات خطا ۰/۱۶ و میانگین خطای مطلق ۰/۱۱ دارای بهترین کارایی بود. بر اساس نتایج جدول ۷ برای ایستگاه کاکارضا مدل GP-PUK با ضریب کارایی ۰/۳۱، ریشه میانگین مربعات خطا ۰/۲۵ و میانگین خطای مطلق ۰/۱۷ در مرحله آموزش از سایر مدل‌ها بهتر شده، اما در مجموع هیچ‌یک از مدل‌ها عملکرد خوبی در این ایستگاه نداشته‌اند.

شکل ۵ نمودارهای داده‌های مشاهداتی و پیش‌بینی بخش آموزش و آزمایش ایستگاه‌های بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا را نشان می‌دهند. شکل ۶ نمودارهای داده‌های خطای بخش آموزش و آزمایش ایستگاه‌های بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا را نشان می‌دهد.

جدول ۵- خصوصیات آماری داده‌های آموزش و آزمایش ایستگاه‌های هیدرومتری مورد مطالعه

Table 5- Statistical characteristics of the training and testing data of the studied hydrometric stations

آزمایش			آموزش			داده‌ها	
انحراف معیار	میانگین	بیش‌ترین	کم‌ترین	انحراف معیار	میانگین	بیش‌ترین	کم‌ترین
28.671	50.674	309	0	21.352	45.007	257	0
0.978	3.189	10.880	0.8	0.421	3.122	8.56	0.33
3.269	5.123	60	0.62	0.928	4.141	27.4	0
1.455	4.156	31.93	0.78	0.567	3.632	17.98	0.75
4.656	9.45	70.04	0.018	42.423	14.139	698.755	0.324

جدول ۶- پارامترهای بهینه مدل‌های مختلف ایستگاه‌های هیدرومتری مورد مطالعه

Table 6- Optimum parameters of different models of studied hydrometric stations

چمانجیر		بهرام‌جو	
پارامترهای بهینه	مدل‌ها	پارامترهای بهینه	مدل‌ها
Gaussian noise = 0.01, $\omega=1.5$, $\sigma=1.9$ Gaussian noise = 2.9, $Y=1.9$ num-decimal-places= 0, M=4 M=0, V=0.001, N=3, S=1 K=0, M=1, V=0.001, S=1	GP-PUK GP-RBF M5P REPTREE RF	Gaussian noise = 0.1, $\omega=0.1$, $\sigma=4.2$ Gaussian noise = 0.2, $Y=0.1$ num-decimal-places= 0, M=4 M=2, V=0.001, N=3, S=1 K=0, M=1, V=0.001, S=1	GP-PUK GP-RBF M5P REPTREE RF
کاکارضا Gaussian noise = 3.1, $\omega=0.01$, $\sigma=0.001$ Gaussian noise = 0.001, $Y=0.01$ num-decimal-places= 0, M=4 M=0, V=0.001, N=3, S=1 K=0, M=1, V=0.001, S=1	GP-PUK GP-RBF M5P REPTREE RF	سرآب صیدعلی Gaussian noise = 0.001, $\omega=0.01$, $\sigma=0.1$ Gaussian noise = 0.001, $Y=0.29$ num-decimal-places= 0, M=4 M=2, V=0.001, N=3, S=1 K=0, M=1, V=0.001, S=1	GP-PUK GP-RBF M5P REPTREE RF

جدول ۷- پارامترهای ارزیابی عملکرد مدل‌های GP-PUK, GP-RBF, RepTree, M5p, RF ایستگاه‌های هیدرومتری مورد مطالعه

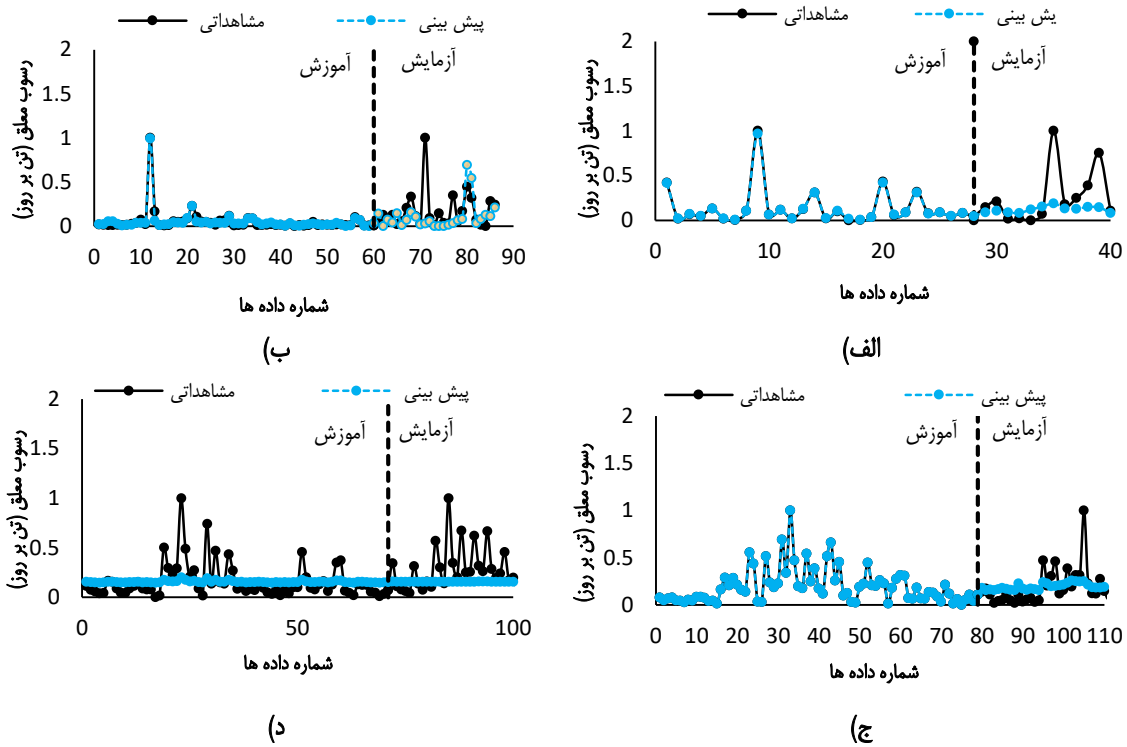
Table 7- Performance evaluation parameters of GP-PUK, GP-RBF, RepTree, M5p, and RF models of studied hydrometric stations

چمانجیر			بهرام‌جو			مدل‌ها						
آزمایش			آموزش			آزمایش			آموزش			
MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	MAE	RMSE	CC	
0.140	0.236	0.301	0.013	0.021	0.9864	0.195	0.309	0.7454	0.004	0.007	0.9999	GP-PUK
0.135	0.236	0.3421	0.044	0.116	0.9471	0.182	0.287	0.8198	0.103	0.182	0.4666	GP-RBF
0.156	0.233	0.142	0.050	0.103	0.6716	0.203	0.324	0	0.128	0.203	0	M5P
0.133	0.238	0.199	0.041	0.114	0.4776	0.219	0.269	0.7369	0.118	0.187	0.3818	REPTREE
0.149	0.235	0.1882	0.021	0.057	0.9605	0.207	0.312	0.3557	0.059	0.084	0.962	RF
کاکارضا						سرآب صیدعلی						
0.170	0.252	0.3103	0.108	0.162	0.9964	0.118	0.169	0.6703	0.0001	0.0001	0.9999	GP-PUK
0.423	0.872	0.4199	0.069	0.121	0.7097	1.978	5.159	0.734	0.017	0.037	0.9789	GP-RBF
0.144	0.215	0.5046	0.097	0.151	0.4863	0.121	0.170	0.6182	0.106	0.142	0.6321	M5P
0.155	0.253	0.1377	0.092	0.146	0.5298	0.135	0.189	0	0.135	0.184	0	REPTREE
0.167	0.242	0.1681	0.01	0.064	0.9576	0.100	0.138	0.7017	0.047	0.065	0.967	RF

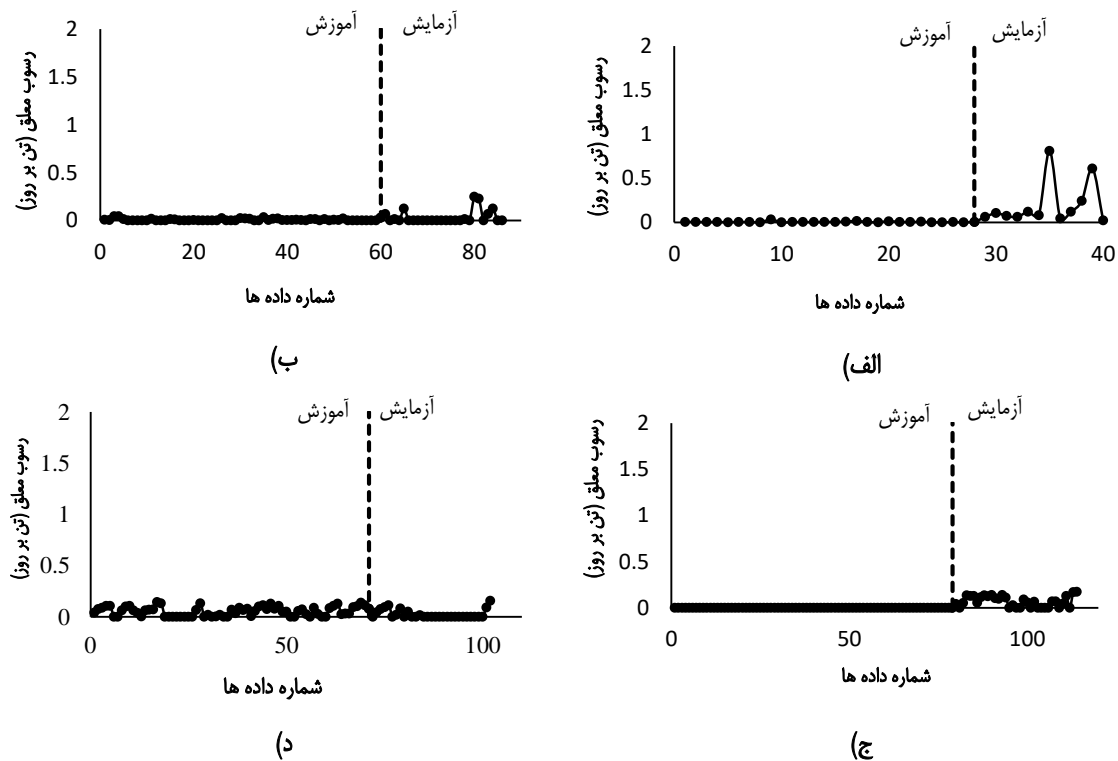
۳-۴- تحلیل حساسیت

کاکارضا با حذف بارش حساسیت نتیجه بهتری داشته است. هم‌چنین در دوره کم‌آبی ایستگاه بهرام‌جو با حذف بارش، ایستگاه چمانجیر با حذف دبی و ایستگاه‌های سرآب صیدعلی و کاکارضا با حذف دبی یک روز قبل حساسیت نتیجه بهتری داشته است.

جدول ۸ تحلیل حساسیت بهترین مدل برای هر ایستگاه را در رسوب معلق دوره کم‌آبی و پرآبی نشان می‌دهد. مطابق این جدول در دوره پرآبی ایستگاه‌های بهرام‌جو، چمانجیر و سرآب صیدعلی با حذف میانگین بین دبی و دبی یک روز قبل و ایستگاه



شکل ۵- نمودار کارایی مدل GP-PUK داده‌های مشاهداتی و پیش‌بینی در مرحله آموزش و آزمایش، (الف) بهرام‌جو، (ب) چمانجیر، (ج) سراب صیدعلی و (د) کاکارضا
 Figure 5- Performance chart of GP-PUK model, observation and prediction data in the training and testing phase, (a) Bahram Jo, (b) Cham-Anjir, (c) Sarab Syed Ali, and (d) Kakarza



شکل ۶- نمودار کارایی مدل GP-PUK داده‌های خطا در مرحله آموزش و آزمایش، (الف) بهرام‌جو، (ب) چمانجیر، (ج) سراب صیدعلی و (د) کاکارضا
 Figure 6- Performance diagram of GP-PUK model of error data in the training and testing phase, (a) Bahram-Jo, (b) Cham-Anjir, (c) Sarab Syed Ali, and (d) Kakarza

جدول ۸- تحلیل حساسیت بهترین مدل ایستگاه‌های هیدرومتری مورد مطالعه
Table 8- Sensitivity analysis of the best model of the studied hydrometric stations

بهرام‌جو							
رسوب معلق خشک (تن بر روز)			رسوب معلق (تن بر روز)			تحلیل حساسیت	
MAE	RMSE	CC	MAE	RMSE	CC	حذف	
0.195	0.309	0.7454	0.277	0.442	0.553	-	$R - Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1816	0.3041	0.8098	0.368	0.7037	0.1693	R	$Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1977	0.3069	0.7287	0.2697	0.4182	0.3849	Q_t	$R - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1947	0.3069	0.6753	0.2823	0.424	0.448	Q_{t-1}	$R - Q_t - (AverageQ_t, Q_{t-1}) - TDS$
0.1976	0.3098	0.7275	0.2781	0.4248	0.5497	$(AverageQ_t, Q_{t-1})$	$R - Q_t - Q_{t-1} - TDS$
چم‌انجیر							
0.140	0.236	0.301	0.081	0.187	0.7437	-	$R - Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1708	0.2748	0.1745	0.1021	0.1848	0.5447	R	$Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1355	0.2245	0.3688	0.0886	0.1928	0.6489	Q_t	$R - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1541	0.2479	0.1898	0.0851	0.1898	0.652	Q_{t-1}	$R - Q_t - (AverageQ_t, Q_{t-1}) - TDS$
0.1541	0.2479	0.1898	0.0861	0.182	0.7038	$(AverageQ_t, Q_{t-1})$	$R - Q_t - Q_{t-1} - TDS$
سراب صیدعلی							
0.118	0.169	0.6703	0.07	0.165	0.7158	-	$R - Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.118	0.1685	0.6225	0.0807	0.1621	0.7133	R	$Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1197	0.1724	0.6424	0.0764	0.1707	0.5157	Q_t	$R - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.118	0.1698	0.6647	0.0682	0.1621	0.6783	Q_{t-1}	$R - Q_t - (AverageQ_t, Q_{t-1}) - TDS$
0.1199	0.1721	0.6583	0.0702	0.1647	0.716	$(AverageQ_t, Q_{t-1})$	$R - Q_t - Q_{t-1} - TDS$
کاکارضا							
0.170	0.252	0.3103	0.157	0.241	0.7309	-	$R - Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1697	0.2526	0.3042	0.1331	0.1755	0.7757	R	$Q_t - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1703	0.253	0.2755	0.1686	0.2375	0.5184	Q_t	$R - Q_{t-1} - (AverageQ_t, Q_{t-1}) - TDS$
0.1701	0.2528	0.3115	0.1446	0.1932	0.7094	Q_{t-1}	$R - Q_t - (AverageQ_t, Q_{t-1}) - TDS$
0.1702	0.253	0.2614	0.1421	0.2089	0.7373	$(AverageQ_t, Q_{t-1})$	$R - Q_t - Q_{t-1} - TDS$

این نتیجه رسیدند که مدل GP با کرنل RBF نسبت به کرنل PUK برتری دارد. همچنین نتیجه به دست آمده با پژوهش Soleimani et al. (2021) نیز همخوانی ندارد؛ زیرا آن‌ها به این نتیجه رسیدند که مدل RF دارای بهترین عملکرد و مدل GP با کرنل PUK دارای کم‌ترین عملکرد در بحث مدل‌سازی نفوذ است. علاوه بر این نتایج با نتایج پژوهش Kohestani et al. (2016) نیز مطابقت ندارد؛ زیرا آن‌ها در پژوهش خود به این نتیجه رسیدند که مدل M5P دارای بیش‌ترین کارایی است. در حالی که، نتایج این پژوهش با نتایج Eshghi et al. (2015) مغایرت داشت.

۴- نتیجه‌گیری

رسوبات معلق، رسوباتی هستند که در یک بازه زمانی مشخص از یک مقطع رودخانه عبور می‌کنند و در اثر فرسایش خاک و جابه‌جایی این مواد فرسایش یافته به آبراهه‌ها ایجاد می‌شوند.

در حالت کلی در رسوب معلق پرآبی و کم‌آبی مدل M5P نتیجه خوبی نداشته و مدل‌های RepTree و RF در بعضی ایستگاه‌ها نتیجه بهتری نسبت به ایستگاه‌های دیگر داشته‌اند و مدل GP با کرنل‌های PUK و RBF نتایج مطلوب‌تری در همه ایستگاه‌ها به دست آوردند. در مجموع مدل GP-PUK به عنوان مدل برتر دارای قدرت بالاتری برای مدل‌سازی همه ایستگاه‌ها در رسوب معلق پرآبی و کم‌آبی در بخش آزمایش بوده است. نتایج این پژوهش با نتایج سایر پژوهش‌گران (Sattari et al., 2015; Malik et al., 2017; Mohammadi, 2018; Mirfalah Nasiri et al., 2019; Adnan Khan et al., 2021; Rahul et al., 2021) مبنی بر توانایی الگوریتم‌های هوش مصنوعی مبتنی بر یادگیری در تخمین رسوب معلق رودخانه همخوانی دارد. نتایج این پژوهش مبنی بر برتری کرنل PUK نسبت به کرنل RBF با نتایج پژوهش Sepahvand and Azizi (2019) همخوانی ندارد؛ زیرا آن‌ها در پژوهش خود به

در ایستگاه سراب صیدعلی نتایج همه مدل‌ها برای بخش آموزش و آزمایش خوب بوده و در نهایت مدل‌ها در ایستگاه کاکارضا در بخش آموزش نتایج بسیار عالی و در بخش آزمایش خیلی خوبی به دست آمده‌اند. در قسمت رسوب معلق کم‌آبی نیز ایستگاه بهرام‌جو در قسمت آموزش فقط برخی مدل‌ها نتایج خوبی داشته و در بخش آزمون در حالت کلی مدل‌ها نتایج بهتری داشته‌اند. در ایستگاه چمانجیر نتایج مدل‌ها در بخش آموزش خوب بوده اما در بخش آزمایش نتایج خوبی به دست نیامده است. در ایستگاه سراب صیدعلی نیز مدل‌ها در بخش آموزش و آزمایش هر دو نتایج خوبی داشته و در آخر ایستگاه کاکارضا در بخش آموزش مدل‌ها نتایج خوبی دارند ولی در بخش آموزش در حالت کلی نتایج خوبی نداشته‌اند.

با توجه به نتایج به دست آمده از این پژوهش می‌توان از این مدل‌های بهینه برای صرفه‌جویی در هزینه و زمان برای بحث حفاظت آب و خاک و تخمین رسوب معلق خروجی از حوضه‌ها استفاده کرد. همچنین می‌توان برای اجرای مدیریت بهتر در رابطه با کمیت و کیفیت آب‌های سطحی، این مدل‌ها برای تخمین رسوبات معلق ایستگاه‌های مجاور فاقد آمار دارای شریط زمین‌ساختی و هیدرولوژیکی یکسان در سطح منطقه مورد استفاده قرار گیرند و نتایج قابل اعتمادی در رابطه با رسوب معلق ارائه دهند.

برآورد رسوب معلق در مدیریت کمی و کیفی آب سطحی نقش اساسی دارد. با اینکه رسوب معلق یک موضوع مهم است، اما اطلاعات دقیقی از میزان رسوب معلق رودخانه‌ها در کشور در دسترس نیست و بین مقدار اندازه‌گیری شده و محاسبه شده اختلافات زیادی وجود دارد. در این پژوهش به بررسی عملکرد برخی مدل‌ها در ارزیابی رسوب معلق رودخانه پرداخته شد. برای ارزیابی عملکرد مدل‌ها از ضریب تعیین و ریشه میانگین مربعات خطا استفاده شده است.

بر اساس هدف انجام این پژوهش مدل‌سازی بار معلق رودخانه با استفاده از الگوریتم‌های یادگیری و به‌کارگیری مدل‌های GP-PUK, GP-RBF, RepTree, M5P, RF در ایستگاه‌های هیدرومتری حوزه آبخیز کرخه شامل بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا انجام شد. نتایج حاصل از مدل‌سازی با استفاده از مدل‌های GP-PUK, GP-RBF, RepTree, M5P, RF برای داده‌های ورودی و خروجی رسوب معلق در دوره کم‌آبی و پرآبی در ایستگاه‌های مختلف (بهرام‌جو، چمانجیر، سراب صیدعلی و کاکارضا) نشان می‌دهد در بخش رسوب معلق پرآبی برای ایستگاه بهرام‌جو در قسمت آموزش بیش‌تر مدل‌ها نتایج خوبی داشته است. در حالی‌که در بخش آموزش نتایج همه مدل‌ها دلخواه نبوده و فقط برخی مدل‌ها نتیجه بهتری داشته‌اند. در ایستگاه چمانجیر هم در بخش آموزش و هم در بخش آزمایش نتیجه مدل‌ها خوب بوده است. همچنین،

منابع

- اسدی، مریم، و فتح‌زاده، علی (۱۳۹۷). بررسی کارایی مدل‌های مبتنی بر هوش محاسباتی در برآورد بار معلق رودخانه (مطالعه موردی: استان گیلان). *مرتع و آبخیزداری، منابع طبیعی ایران*، ۱(۷۱)، ۴۵-۶۰. doi:10.22059/jrwm.2018.222810.1083
- انصاری، حسین، و داوری، کامران (۱۳۸۷). پهنه‌بندی دوره خشک با استفاده از شاخص بارش استاندارد در محیط GIS. *تحقیقات جغرافیایی*، ۳۹(۶۰)، ۹۸-۱۰۸.
- خلیقی سیگارودی، شهرام، صادقی سنگدهی، سیدعلی، اوسطی، خالد، و قویدل رحیمی، یوسف (۱۳۸۸). بررسی نمایه‌های ارزیابی پدیده‌های ترسالی و خشک‌سالی: (SPI, PNPI, Nitzche) (مطالعه موردی، استان مازندران). *تحقیقات مرتع و بیابان ایران*، ۱۱(۱)، ۴۴-۵۴.
- سپه‌وند، علیرضا، نظری سامانی، علی اکبر، محمدیان، هوشنگ، احمدی، حسن، و فیض‌نیا، سادات (۱۳۹۹). تغییرات بار فصلی رسوب انحلالی و تعیین سرعت انحلال‌پذیری سازندهای آهکی. *علوم و مهندسی آبخیزداری ایران*، ۱۴(۴۸)، ۲۱-۳۲. doi:10.1001.1.20089554.1399.14.48.4.1
- سپه‌وند، علیرضا، و عزیزی نجفقلی، زینب (۱۳۹۹). مدل‌سازی رسوب معلق با استفاده از مدل‌های فرآیند گوسی و پرسپترون چند لایه.
- پانزدهمین همایش ملی علوم و مهندسی آبخیزداری ایران، دانشگاه علوم کشاورزی و منابع طبیعی ساری.
- ستاری، محمدتقی، رضازاده جودی، علی، صفدری، فروغ، و قهرمانیان، فراز (۱۳۹۵). ارزیابی عملکرد روش‌های مدل درختی M5 و رگرسیون بردار پشتیبان در مدل‌سازی رسوب معلق رودخانه. *حفاظت منابع آب و خاک*، ۶(۱)، ۱۰۹-۱۲۴.
- سلیمانی، لیلی، میردیکوند، بهرام، و سپه‌وند، علیرضا (۱۴۰۱). مدل‌سازی نفوذپذیری در کلاس‌های مختلف بافت خاک با استفاده از الگوریتم‌های یادگیری. *پژوهش‌های آبخیزداری*، ۳۵(۴)، ۱۵-۱.
- doi:10.22092/wmrj.2022.358213.1461
- عشقی، پریسا، دستورانی، محمدتقی، عرب اسدی، زینب، و فرزاد مهر، جلیل (۱۳۹۵). بررسی کارایی مدل‌های هوشمند در برآورد رسوبات معلق رودخانه‌ای. *پژوهشنامه مدیریت حوزه آبخیز*، ۷(۱۴)، ۸۸-۹۵.
- قبادیان، رسول، و شکری، حامد (۱۳۹۸). بررسی عددی عوامل موثر بر توزیع غلظت رسوب نامتعادل در رودخانه‌های طبیعی. *آب و خاک (علوم و صنایع کشاورزی)*، ۲(۳۴)، ۲۴۱-۲۵۳. doi:10.22067/jsw.v34i2.76326

- مصطفی‌زاده، رئوف، و ذبیحی، محسن (۱۳۹۵). مقایسه شاخص‌های SPI و SPEI در ارزیابی خشکسالی هواشناسی با استفاده از برنامه‌نویسی R. *فیزیک زمین و فضا*، (۴۲)، ۶۳۳-۶۴۳. doi:10.22059/jesphys.2016.57881
- میرفلاح نصیری، سیدسامان، امیری، ابراهیم، و شادابی بجند، محبوبه (۱۳۹۹). مدل‌سازی برآورد میزان رسوب معلق رودخانه پسیخان با استفاده از شبکه مصنوعی درخت تصمیم. *حفاظت منابع آب و خاک*، (۲)۱۰، ۳۱-۴۲. doi:20.1001.1.22517480.1399.10.2.3.6
- کوهستانی، وحیدرضا، حسنلو، محمود، و بازرگان لاری، محمدرضا (۱۳۹۵). پیش‌بینی ظرفیت باربری نهایی پی‌های سطحی واقع بر خاک‌های دانه‌ای با استفاده از مدل درختی M5P. *مهندسی عمران فردوسی*، (۲۷)۲، ۹۹-۱۱۰. doi:10.22067/civil.v27i2.33915
- محمدی، صدیقه (۱۳۹۸). شبیه‌سازی بار رسوب معلق با استفاده از روش‌های شبکه عصبی مصنوعی، عصبی-فازی و منحنی سنج رسوب در حوزه آبخیز هلیل‌رود. *مهندسی و مدیریت آبخیز*، (۱۱)۲، ۴۵۲-۴۶۶. doi:10.22092/ijwmse.2017.108140.1219
- ## References
- Adnan Khan, M., Stamm, J., & Haider, S. (2021). Assessment of soft computing techniques for the prediction of suspended sediment loads in rivers. *Applied Science*, 11, 82-90. doi:10.3390/app11188290
- Ansari, H., & Davari, K. (2008). Classification of dry period using standard precipitation index in GIS environment. *Geographical Research*, 39(60), 98-108. [In Persian]
- Asadi, M., & Fathzadeh, A. (2017). Investigating the effectiveness of computational intelligence-based models in river load estimation (case study: Gilan province). *Range and Watershed Management, Journal of Natural Resources of Iran*, 1(71), 45-60. doi:10.22059/jrwm.2018.222810.1083 [In Persian]
- Bazuhair, S.A., Gohani, A., & Sen, Z. (1997). Determination of monthly wet and dry periods in Saudi Arabia. *International Journal of Climatology*, 17, 303-311.
- Bonakdar, L., & Etemad Shahidi, A. (2011). Predicting wave run-up on rubble-mound structures using M5 model tree. *Ocean Engineering*, (38), 111-118.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. doi:10.1007/BF00058655
- Doroudi, S., Sharafati, A., & Mohajeri, S.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*, 5540284, 1-13.
- Eshghi, P., Farzad Mehr, J., Dasturani, M.T., & Arabasadi, Z. (2015). The effectiveness of intelligent models in estimating the river suspended sediments (Case study: Babaaman Basin, Northern Khorasan). *Journal of Watershed Management Research*, 7(14), 88-95. [In Persian]
- Khalighi, Sh., Sadeghi Sangehi, S.A., Osta, Kh., & Qavidel Rahimi, Y. (2009). The study of drought and wet year assessment models for stations in Mazandaran province. *Iranian Journal of Range and Desert Research*, 16(1), 44-54. [In Persian]
- Kohestani, V.R., Hasanlorad, M., & Bazargan Lari, M. (2016). Prediction of the ultimate bearing capacity of surface foundations located on granular soils using the M5P tree model. *Ferdowsi Civil Engineering*, 2(27), 99-110. doi:10.22067/civil.v27i2.33915 [In Persian]
- Malik, A., Kumar, A., & Piri, J. (2017). Daily suspended sediment concentration simulation using hydrological data of Pranhita River Basin, India. *Computers and Electronics in Agriculture*, 1(138), 20-28. doi:10.1016/j.compag.2017.04.005
- Mirfalah Nasiri, S., Amiri, E., & Shadabi Bejand, M. (2019). Modeling Estimation of Suspended Sediment Rate in Pasikhan River Using Decision Tree Artificial Neural Network. *Journal of Water and Soil Resources Conservation*, 10(2), 31-42. doi:20.1001.1.22517480.1399.10.2.3.6 [In Persian]
- Mohammadi, S. (2018). Simulation of suspended sediment load using artificial neural network, neural-fuzzy and sediment gauge curve methods in Halil-Roud watershed. *Scientific-Research Journal of Watershed Engineering and Management*, 2(11), 452-466. doi:10.22092/ijwmse.2017.108140.1219 [In Persian]
- Mostafazadeh, R., & Zabihi, M. (2016). Comparison of SPI, SPEI indices in meteorological drought assessment using R programming (study area: Kurdistan province). *Journal of Earth and Space Physics*, 42, 633-643. doi:10.22059/jesphys.2016.57881 [In Persian]
- Nourani, V., Gokcekus, H., & Gelete, G. (2020). Estimation of suspended sediment load using artificial intelligence-based ensemble model. *Complexity*, 1-19. doi:10.1155/2021/6633760
- Pal, M., & Deswal, S. (2010). Modelling pile capacity using Gaussian process regression. *Computers and Geotechnics*, 37, 942-947.

- Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217-222. doi:10.1080/01431160412331269698
- Prasad, A.M., Iverson, L.R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, 9(2), 181-199. doi:10.1007/s10021-005-0054-1
- Qobadian, R., & Shokri, H. (2018). Numerical investigation of factors affecting the distribution of unbalanced sediment concentration in natural rivers (case study: Qarasu River, Kermanshah). *Water and Soil (Agricultural Sciences and Industries)*, 2(34), 241-253. doi:10.22067/jsw.v34i2.76326 [In Persian]
- Quinlan, J.R. (1992). Learning with continuous classes. In Proceedings of the 5th Australian joint Conference on Artificial Intelligence. Hobart: Singapore.
- Rahul, A.K., Shivhare, N., Kumar, S., Dwivedi, S.B., & Dikshit, P.K.S. (2021). Modelling of daily suspended sediment concentration using FFBPNN and SVM algorithms. *Journal of Soft Computing in Civil Engineering*, 5(2), 120-134. doi:10.22115/SCCE.2021.283137.1305
- Sattari, M., Rezazade Jodi, A., Safdari, F., & Kahramanzadeh, F. (2015). Performance evaluation of M5 tree model and support vector regression methods in river suspended sediment modeling. *Journal of Water and Soil Resources Protection*, 6(1), 109-124. [In Persian]
- Sepahvand, A., & Azizi Najafkali, Z. (2019). Suspended sediment modeling using Gaussian process and multi-layer perceptron models. 15th National conference on Watershed Management Sciences and Engineering of Iran, Sari, Iran.
- Sepahvand, A., Nazari Samani, A.A., Mohammadian, H., Ahmadi, H., & Feiz Nia, S. (2020). Seasonal variation of the solute and determine the solubility of limestone formations. *Iranian Journal of Watershed Management Science and Engineering*, 14(48), 21-32. dor:20.1001.1.20089554.1399.14.48.4.1
- Sepahvand, A., Singh, B., Sihag, P., Nazari, A., Hasan Ahmadi, S., & Fiz Nia, S. (2019). Assessment of the various soft computing techniques to predict sodium absorption ratio (SAR). *ISH Journal of Hydraulic Engineering*, 27, 124-135. doi:10.1080/09715010.2019.1595185
- Shojaeezadeh, S.A., Nikoo, M.R., McNamara, J.P., AghaKouchak, A., & Sadegh, M. (2018). Stochastic modeling of suspended sediment load in alluvial rivers. *Advances in Water Resources*, 119, 188-196. doi:10.1016/j.advwatres.2018.06.006
- Smakhtin V.U. (2001). Low-flow hydrology: a review. *Journal of Hydrology*, (240), 147-186. doi:10.1016/S0022-1694(00)00340-1
- Soleimani, L., Derikund, B., & Sepehvand, A. (2021). Modelling of infiltration rate in different soil textures using soft computing techniques in Kashkan Watershed, Lorestan Province. *Journal of Watershed Management Research*, 35(4), 1-150. doi:10.22092/wmrj.2022.358213.1461 [In Persian]
- Wang, Y., & Witten, I.H. (1997). Inducing model trees for continuous classes. Proceedings of the Ninth European Conference on Machine Learning, Prague, Czech Republic: Springer.
- Yang, D., Zhang, X., Pan, R., Wang, Y., & Chen, Z. (2018). A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *Journal of Power Sources*, 384, 387-395. doi:10.1016/j.jpowsour.2018.03.015