

بررسی کارایی مدل‌های مبتنی بر هوش محاسباتی در برآورد بار معلق رودخانه‌ها (مطالعه موردی: استان گیلان)

- ❖ مریم اسدی؛ دانشجوی دکتری آبخیزداری، دانشکده منابع طبیعی، دانشگاه تهران، ایرن
❖ علی فتح‌زاده*؛ دانشیار دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان، ایران

چکیده

آگاهی از میزان رسوب معلق رودخانه‌ها یکی از مسائل اساسی در پروژه‌های آبی است که طراحان تأسیسات آبی همواره با آن روبرو بوده‌اند. با توجه به صرف هزینه و زمان طولانی جهت اندازه‌گیری بار معلق رودخانه‌ها، استفاده از منحنی‌های سنجه رسوب معمول‌ترین روش برآورد بار رسوب معلق رودخانه‌ها محسوب می‌گردد. این در حالی است که روش‌های نوین مبتنی بر هوش مصنوعی و داده‌کاوی در بسیاری از علوم مهندسی رخنه کرده است. بر همین اساس هدف اصلی این تحقیق به چالش کشیدن توانمندی روش کلاسیک برآورد بار معلق در مقایسه با برخی روش‌های نوظهور می‌باشد. ما در این پژوهش شش مدل، K نزدیک‌ترین همسایه، شبکه عصبی پس انتشار خطا، فرآیند گوسی، درخت تصمیم‌گیری M5، ماشین بردار پشتیبان و ماشین بردار پشتیبان تکاملی را انتخاب و به مقایسه آنها با مدل سنجه رسوب در هشت حوزه آبخیز واقع در استان گیلان پرداختیم. طول دوره آماری داده‌های ورودی به مدل‌ها به صورت روزانه و ۳۰ ساله در نظر گرفته شد. ارزیابی نتایج حاصله نشان داد مدل فرآیند گوسی در مقایسه با سایر مدل‌ها، با کمترین مجموع مربعات باقیمانده (RMSE) (متوسط مجموع مربعات باقی مانده = $37/05$ در هشت حوزه) و بیشترین ضریب همبستگی (r) (متوسط ضریب همبستگی $0/72$ در هشت حوزه) و با بهترین ضریب ناش-ساتکلیف (متوسط $0/66$ در هشت حوزه) نسبت به سایر مدل‌ها از کارایی بیشتری برخوردار است. لذا استفاده از مدل‌های مذکور به جای روش‌های معمول برآورد بار معلق می‌تواند دقت این برآوردها را به میزان قابل ملاحظه‌ای بهبود بخشد.

کلید واژگان: بار رسوبی معلق، منحنی سنجه رسوب، فرآیند گوسی، داده‌کاوی، هوش مصنوعی

۱. مقدمه

این رو به نظر می‌رسد استفاده از رهیافت‌های مبتنی بر هوش محاسباتی می‌تواند کمک شایانی در زمینه پیش‌بینی بار رسوبی معلق نماید به‌گونه‌ای که این مدل‌ها با توجه به توانمندی‌شان در مدل‌سازی فرآیندهای غیر خطی قادر به پیش‌بینی بار رسوبی معلق با دقت بیشتری نسبت به مدل منحنی سنجه رسوب می‌باشند.

به‌عنوان مثال در مطالعه‌ای به بررسی امکان پیش‌بینی رسوبات معلق با استفاده از ترکیب منحنی سنجه رسوب و شبکه عصبی مصنوعی پرداخته شد [۱۱]. در این تحقیق کاربرد روش سیستم عصبی مصنوعی در بهینه‌سازی نتایج مدلی مبتنی بر روش منحنی سنجه رسوب برای پیش‌بینی بار معلق جریان مورد بررسی قرار گرفت. پس از ساختن مدل ترکیبی، مدل ANN و منحنی سنجه رسوب به‌تنهایی نیز با داده‌های مشابه مورد استفاده قرار گرفت و نتایج حاصل با مقادیر واقعی مقایسه گردید. نتایج حاصل از کاربرد ترکیبی این مدل‌ها از کیفیت به مراتب بالاتری نسبت به کاربرد هر یک از آنها به‌تنهایی برخوردار بوده است به طوری که میانگین مربعات خطا در مدل منحنی سنجه رسوب و شبکه عصبی به‌ترتیب ۶۹۲/۲۸ و ۶۱۶/۹۶ بوده و در مدل ترکیبی ۶۰۳/۰۹ کاهش پیدا کرده است. برخی محققین در پژوهشی به مدل‌سازی رسوب رودخانه آبخیز جونیاتا در آمریکا پرداختند [۸]. آنها در مطالعه خود روش‌های پس انتشار خطا، منحنی سنجه رسوب و رگرسیون چندمتغیره خطی را در پیش‌بینی بار رسوبی معلق بر مبنای معیارهای ارزیابی (میزان خطا و همبستگی بین مقادیر پیش‌بینی و اندازه‌گیری شده) مقایسه نمودند. نتایج آن‌ها نشان داد که عملکرد شبکه عصبی مصنوعی نسبت به روش‌های دیگر بهتر می‌باشد.

با پیشرفت‌های صورت گرفته در علم هوش محاسباتی استفاده از مدل‌های پیچیده‌تر و تکامل یافته‌تر در مطالعات رسوب مرسوم گردید. به طوری که گروهی از محققین در مطالعه‌ای [۲۳] به مدل‌سازی غلظت رسوب معلق براساس مدل‌های داده‌محور پرداختند. آنان در

رسوب معلق به‌عنوان یکی از اجزای مهم عملکرد هیدرولوژیکی، ژئومورفولوژیکی و اکولوژیکی حوزه‌های آبخیز و عامل اصلی کاهش کیفیت آب رودخانه‌ها شناخته شده است [۱۷]. بار رسوبی رودخانه یکی از مهم‌ترین پارامترها در مدیریت پروژه‌های هیدرولیکی، شاخص وضعیت خاک و فرسایش حوزه می‌باشد [۲۲]. ارزیابی و پیش‌بینی حجم رسوبات معلق به‌منظور اجرای برنامه‌های حفاظت خاک و کاهش رسوب‌زایی و همچنین محاسبه و طراحی دقیق حجم سدهای رسوب‌گیر و مخزنی و به‌طور کلی در زمینه پایداری منابع آبی و مدیریت محیط زیست ضرورت دارد [۱، ۳ و ۲۳]. پدیده فرسایش و انتقال رسوب از فرآیندهای پیچیده هیدرودینامیکی می‌باشد که به دلیل دخالت فاکتورهای مختلف شبیه‌سازی رفتاری آن به آسانی میسر نمی‌باشد. استفاده از منحنی‌های سنجه رسوب یکی از روش‌های ساده بسیار متداول در پیش‌بینی بار رسوبی معلق می‌باشد که در آن با استفاده از ارقام دبی به برآورد بار رسوبی معلق پرداخته می‌شود. در مطالعه‌ای به برآورد بار معلق در منطقه دلتای رود مروارید با استفاده از منحنی سنجه رسوب پرداخته شد [۲۱]. در این مطالعه، با استفاده از داده‌های ماهانه ۵۰ سال اخیر سه رودخانه دلتای رود مروارید و با توجه به روابط بین دبی و غلظت رسوبات معلق مجموعه‌ای از منحنی‌های سنجه رسوب رسم گردید. نتایج آنان نشان داد منحنی سنجه رسوب به‌خوبی قادر به پیش‌بینی بار رسوبی معلق بوده است و پارامترهای سنجه رسوب در منحنی‌های مختلف با هم متفاوت بوده است.

علی‌رغم متداول بودن این روش در بین هیدرولوژیست‌ها، خطی نبودن رابطه دبی و رسوب معلق و از طرفی پیچیده بودن مکانیسم حمل رسوب موجب پایین آمدن ضریب همبستگی در منحنی‌های سنجه رسوب می‌گردد که این امر ناکارآمدی منحنی‌های سنجه رسوب در پیش‌بینی بار رسوبی معلق را به همراه دارد. از

مدل‌های هوش محاسباتی نسبت به روش کلاسیک منحنی‌سنجی رسوب می‌باشد. این در حالی است که مدل‌های هوش محاسباتی نیز متعدد بوده و رفتار هر حوزه نیز در تولید رسوب متفاوت می‌باشد. از این رو لازم است تا با استفاده از یک طیف وسیعی از مدل‌های هوش محاسباتی و در شرایط حوزه‌های متفاوت نسبت به ارزیابی و مقایسه این مدل‌ها با مدل‌های کلاسیک اقدام گردد. لذا در این مطالعه هفت مدل منحنی‌سنجی رسوب، K نزدیک‌ترین همسایه، شبکه عصبی پس انتشار خطا، فرآیند گوسی، درخت تصمیم‌گیری M5، ماشین‌بردار پشتیبان و ماشین‌بردار پشتیبان تکاملی در هشت حوزه متفاوت به منظور پیش‌بینی بار رسوبی معلق مورد مقایسه و ارزیابی قرار گرفتند.

۲. مواد و روش‌ها

۱،۲. منطقه مورد مطالعه

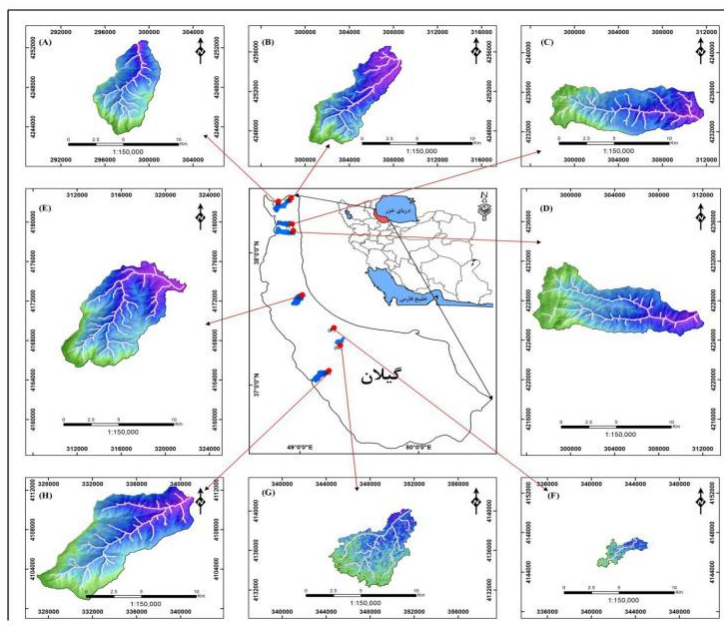
در مطالعات پیش‌بینی پارامترهای هیدرولوژیکی حوزه آبخیز اولین اقدام مهم تهیه یک پایگاه داده با ساختار مناسب می‌باشد. به همین دلیل ابتدا داده‌های دبی و رسوب روزانه هشت ایستگاه هیدرومتری واقع در استان گیلان بر روی رودخانه‌های چوبر، لمیر، گشت رودخان، ملاهادی، کل سر، خاله‌سرا، بهبه‌بر و بهارستان طی سال‌های ۱۳۶۳ تا ۱۳۹۳ تهیه گردید. منطقه مورد نظر از نظر اقلیمی دارای شرایط معتدل خیزی بوده و حداقل و حداکثر ارتفاع در آنها به ترتیب ۸۹- متر و ۳۷۰۲ متر می‌باشد. شکل ۱ پراکنش ایستگاه‌های مطالعاتی را نشان می‌دهد.

با انجام بررسی‌های لازم بر روی داده‌های برداشته شده در هر ایستگاه، طول دوره آماری مشترک مشخص شد. با تعیین تعداد ایستگاه‌های موردنظر پارامترهای آماری دبی و بار رسوبی معلق در هر یک از ایستگاه‌ها محاسبه گردید.

مطالعه خود مدل‌های شبکه عصبی مصنوعی، SVR و منحنی‌سنجی رسوب را مورد مقایسه قرار دادند. نتایج بر اساس معیارهای ارزیابی در این مدل‌ها نشان داد که مدل‌های ANN و SVR به ترتیب ۲۳٪ و ۱۸٪ بهبود در پیش‌بینی و ۱۸٪ و ۱۵٪ بهبود در برآورد بیش از مدل‌های سنتی داشته است و مدل SVR با تابع RBF دارای عملکرد مناسب‌تری به منظور پیش‌بینی بار رسوبی معلق می‌باشد. همچنین در پژوهشی دیگر به مدل‌سازی ارتباط رسوب معلق و دبی با استفاده از حداقل مربعات ماشین‌بردار پشتیبان (LS-SVM) و شبکه عصبی منحنی‌سنجی رسوب پرداخته شد [۱۲]. جریان روزانه رودخانه و داده غلظت رسوبات معلق، در دو ایستگاه بالادست و پایین دست حوزه‌ای در منطقه کالیفرنیا به عنوان ورودی‌های مدل استفاده شد. نتایج نشان داد که در ایستگاه بالادست مدل LS-SVM با RMSE، ۲۸۳، بهترین مدل می‌باشد اما در ایستگاه پایین دست منحنی‌سنجی رسوب با RMSE، ۱۴/۷۴، نتایج بهتری را ارائه داده است.

به کارگیری روش‌ها و تکنیک‌های هوش محاسباتی در پیش‌بینی بار رسوبی معلق باعث افزایش دقت پیش‌بینی‌ها گشته است. به طوری که طی پژوهشی یک شبکه جدید مدل موجک عصبی مصنوعی (WANN) برای پیش‌بینی بار رسوب معلق روزانه ارائه گردید [۱۳]. برای این منظور، از داده‌های دبی روزانه و رسوب معلق رودخانه Yarkin در ایالات متحده آمریکا استفاده گردید. به منظور بررسی دقت مدل ارائه شده، مقایسه مدل با مدل‌های شبکه عصبی مصنوعی، رگرسیون چند خطی (MLR) و مدل منحنی‌سنجی رسوب معمولی صورت گرفت و مقایسه دقت پیش‌بینی از مدل‌ها نشان داد که WANN مدل دقیق‌تری در پیش‌بینی بار رسوبی معلق می‌باشد به طوری که مدل WANN به ترتیب ۷۲،۲٪، ۷۵،۸٪ و ۹۲،۶٪ به ترتیب نسبت به مدل ANN، MLR و SRC میزان RMSE را بهبود بخشید.

جمع‌بندی تحقیقات صورت گرفته بیانگر توانمندی



شکل ۱. پراکنش ایستگاه‌های مطالعاتی در استان گیلان

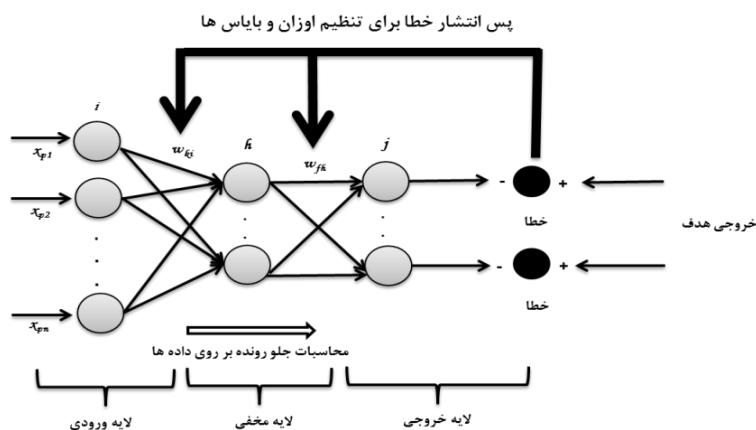
وزن ورودی را از لایه قبلی دریافت کرده آن را به‌عنوان خروجی نرون در لایه بعدی انتقال می‌دهد. مجموع وزن‌های ورودی محاسبه شده و این مجموع به وسیله یک تابع غیرخطی انتقال می‌یابد [۹]. نتایج شبکه با مقادیر واقعی مقایسه شده و خطای شبکه توسط معادله محاسبه می‌شود. روند آموزش تا زمانی که خطا به مقدار قابل قبولی برسد ادامه می‌یابد (شکل ۲).

۲.۲. مدل‌های مورد استفاده

اجرای مدل‌های داده‌کاوی و محاسبه مقادیر معیارهای ارزیابی با استفاده از نرم افزار داده‌کاوی ریپید ماینر (v.5 RapidMiner) انجام شد.

شبکه عصبی پس انتشار خطا

شبکه عصبی پس انتشار خطا شامل سه لایه ورودی، خروجی و لایه مخفی است. هر نرون در یک لایه



شکل ۲. مدل شبکه عصبی پس انتشار خطا [۱۹]

گره می‌رسد. T_i بیانگر نمونه‌هایی است که آامین خروجی سری پتانسیلی را دارند و sd بیانگر انحراف معیار است [۲]. به دلیل فرآیند انشعاب، داده‌های قرار گرفته در گره‌های فرزند، انحراف معیار کمتری نسبت به گره مادر داشته و بنابراین خالص‌تر هستند. پس از حداکثرسازی تمامی انشعاب‌های ممکن M5 صفتی را انتخاب می‌کند که کاهش مورد انتظار را بیشینه نماید [۲۰].

در شکل (۳) نحوه عملکرد مدل درخت تصمیم M5 برای یک مسأله فرضی هر مدل نشان‌دهنده یک معادله رگرسیونی خطی می‌باشد.

- فرآیند گوسی (GP)^۲

فرآیند گوسی یک روش یادگیری ماشینی غیرپارامتری قوی برای ایجاد مدل‌های احتمال گرایانه جامع از مسائل دنیای واقعی است. فرآیندهای گوسی، توزیعات گوسی چندمتغیره را به بعدیت نامتناهی توسعه می‌دهند. به طور رسمی، یک فرآیند گوسی تولید داده‌هایی می‌کند که در طول این طیف قرار دارد به طوری که هر زیر مجموعه متناهی دامنه یک توزیع گوسی چندمتغیره را دنبال می‌کند. شکل ۴ ساختار فیزیکی را برای مدل فرآیند گوسی ارائه می‌دهد.

- ماشین بردار پشتیبان (SVM)^۳

روش ماشین بردار پشتیبان از روش‌های یادگیری ماشینی است که بر مبنای تئوری یادگیری آماری Vapnik در دهه ۹۰ میلادی توسط Vapnik و همکاران ارائه شد. در ماشین بردار پشتیبان از اصول کمینه کردن ریسک ساختاری استفاده شده؛ در حالی که سایر روش‌ها از اصول کمینه کردن تجربی بهره می‌برند. این روش بر پایه نظریه یادگیری محاسباتی توسعه یافته است [۱۰].

تابع کرنلی، برای ایجاد تابع ماشین‌هایی با انواعی مختلف از سطوح تصمیم‌گیری غیرخطی در فضای

- K نزدیک‌ترین همسایه^۱

الگوریتم k نزدیک‌ترین همسایه شامل انتخاب تعداد مشخصی از بردار داده‌ها و سپس باز نمونه‌گیری تصادفی از آن مجموعه برای شبیه‌سازی دوره زمانی متعاقب یک دوره معین می‌باشد. از مزایای استفاده از این الگوریتم در پیش‌بینی‌ها می‌توان به اجرای ساده، عدم نیاز به مرحله تخمین پارامترها، قابلیت مدل‌سازی غیرخطی، مؤثر بودن و عملکرد با بازدهی بالا در برخورد با تعداد دسته‌های زیاد از داده‌ها می‌باشد. در روش K-NN تابع توزیع مقادیر پیش‌بینی با استفاده از توزیع ناپارامتری تابع کرنل به دست می‌آید [۴].

- مدل درختی M5

این مدل زیر مجموعه‌ای از روش‌های یادگیری ماشینی و داده‌کاوی است. ساختار یک مدل درختی شامل ریشه، گره‌های داخلی و برگ می‌باشد. از مدل‌های درخت تصمیم در حل بسیاری از مسائل طبقه‌بندی و رگرسیون استفاده شده است. این مدل، بر خلاف مدل‌های درخت تصمیم معمول که کلاس یا رده‌های گسسته را به عنوان خروجی ارائه می‌کنند، یک مدل خطی چندمتغیره را برای داده‌ها در هر گره از مدل درختی می‌سازد. تشکیل ساختار مدل‌های درخت تصمیم‌گیری شامل مراحل ایجاد درخت و هرس کردن آن است [۱۸]. در مرحله ساختن درخت، از یک الگوریتم استنتاجی یا معیار تقسیم (انشعاب) برای تولید درخت تصمیم استفاده می‌شود. معیار تقسیم برای الگوریتم مدل M5 ارزیابی انحراف معیار مقادیر کلاسی است که به عنوان کمیتی از خطا به یک گره می‌رسد و کاهش مورد انتظار در این خطا را به عنوان نتیجه آزمون هر صفت در آن گره محاسبه می‌نماید.

کاهش انحراف معیار (SDR) از رابطه زیر به دست می‌آید:

$$SDR = sd(T) - \sum \frac{|T_i|}{|T|} sd(T_i) \quad (۱) \text{ رابطه}$$

که در آن T، بیانگر یک سری نمونه‌هایی است که به

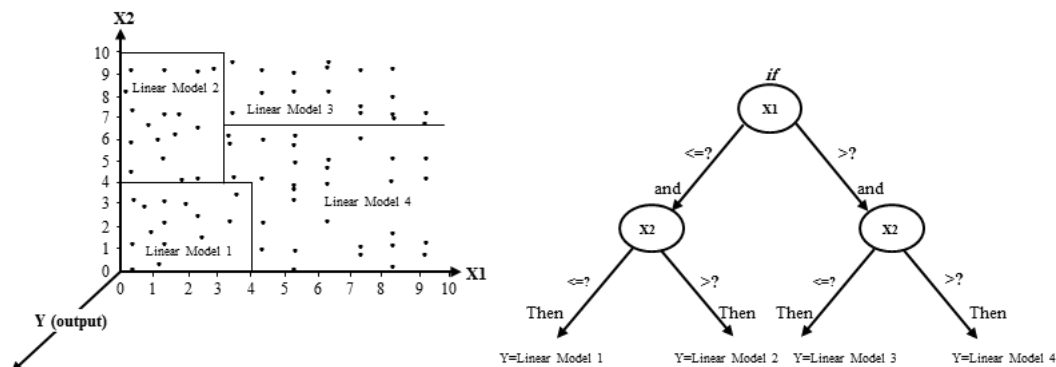
¹ K Nearest Neighbor

² Gaussian Process

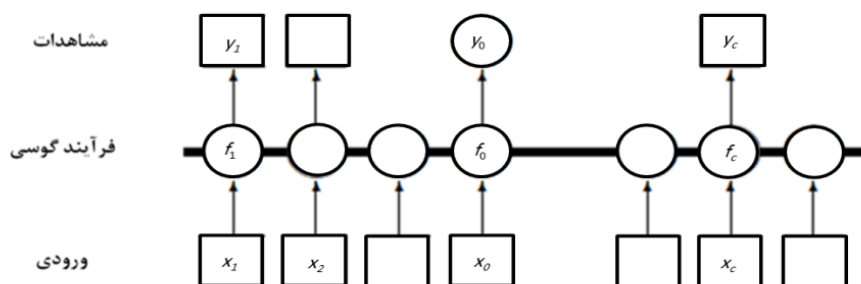
³ Support Vector Machine

ورودی‌ها را برای تبدیل به تابع خطی در فضای ویژگی‌ها فراهم می‌کند (شکل ۵).

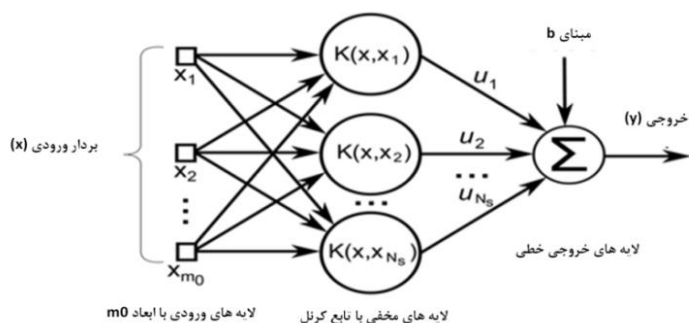
داده‌ها، ضرب‌های داخلی تولید می‌کند. انتخاب مناسب تابع کرنل امکان استفاده از یک تابع غیرخطی در فضای



شکل ۳. نحوه عملکرد مدل درختی M5 بیان معیار تقسیم فضای پارامترهای ورودی به صورت درختی [۱۶]



شکل ۴. ساختار فیزیکی برای مدل فرآیند گوسی [۱۴]

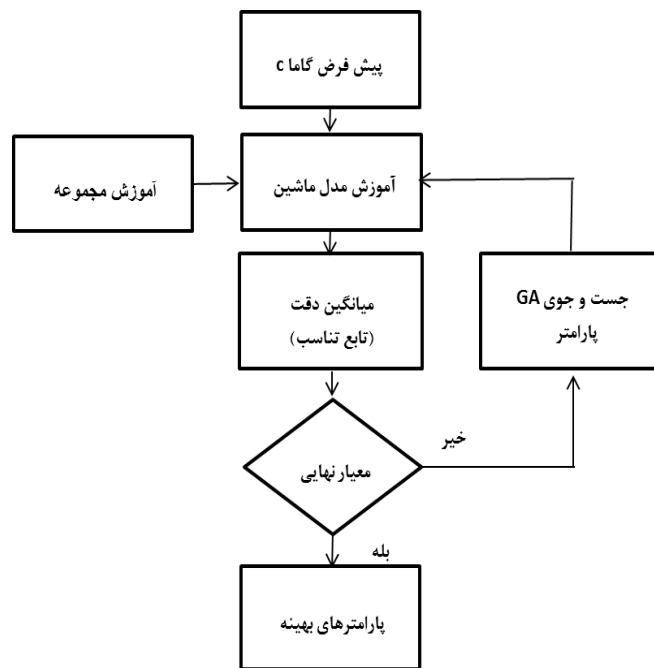


شکل ۵. ساختار مدل ماشین بردار پشتیبان [۲۳]

پارامترهای سیستم با طراحی مدل کروموزومی GA^2 کارآمد دارد و به‌عنوان یک عملگر هوشمند می‌باشد. ESVM با تجزیه و تحلیل فرکانس از ویژگی‌های انتخاب شده، یک مجموعه کوچک از ویژگی‌های مؤثر را شناسایی کرده و از این رو باعث بهبود دقت طبقه‌بندی می‌گردد [۱۰] شکل (۶).

– ماشین بردار پشتیبان تکاملی^۱ (ESVM)

مدل ماشین بردار تکاملی یک استراتژی تکاملی را به‌منظور بهینه‌سازی به کار می‌برد. در واقع این مدل یک الگوریتم تکاملی را برای حل مشکل بهینه‌سازی دوگانه ماشین بردار پشتیبان ارائه می‌دهد. اجرای این الگوریتم در مقایسه ماشین بردار پشتیبان سریع‌تر و آسانتر می‌باشد. EVSM بهترین استفاده را از IGA در بهینه‌سازی



شکل ۶. ساختار مدل ماشین بردار پشتیبان تکاملی [۵]

۳.۲. ارزیابی مدل

ارزیابی مدل‌ها بر اساس روش ارزیابی K_Fold ، $Cross_Validation$ (Leaving one Out) انجام گردید، که در آن مقدار K برابر N در نظر گرفته شد. در این روش در هر آزمون یکی از داده‌ها به‌عنوان داده آزمایشی از سایر داده‌ها جدا شده و فرآیند یادگیری با مابقی داده‌ها ($N-1$) انجام می‌شود [۱۵]. از فواید این شیوه ارزیابی این است که در هر آزمون یکی از داده‌ها به‌منظور

– منحنی سنج‌ه رسوب

در این روش داده‌های ثبت شده دبی و رسوبات معلق در ایستگاه هیدرومتری، در یک گراف رسم گردیده و به آن یک منحنی برازش داده شده و رابطه رگرسیونی به دست آمده از این روش به صورت زیر به دست می‌آید:

$$Q_s = a(Q_w)^b \quad \text{رابطه (۲)}$$

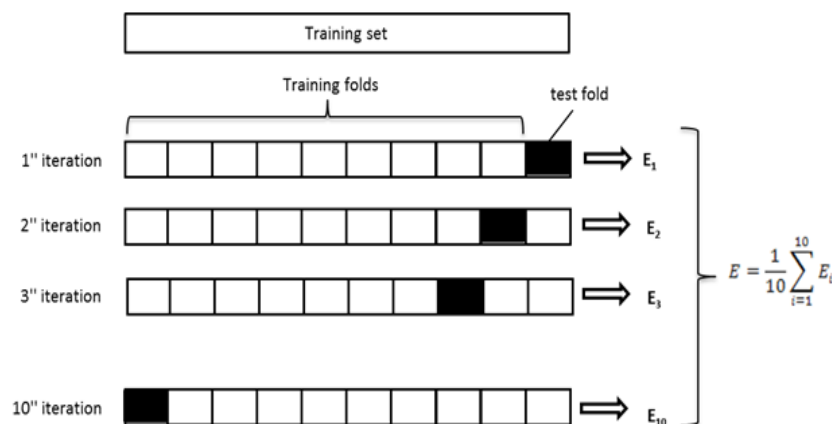
که در این رابطه Q_s دبی رسوبات معلق، Q_w دبی رودخانه و a ، b ضرایب خطی بین مقادیر $\log Q_w$ و $\log Q_s$ می‌باشند.

¹ Evolutionary support vector machine

² Genetic Algorithm

خواهد بود. نحوه فرآیند ارزیابی در شکل ۷ ارائه شده است:

ارزیابی مورد استفاده قرار می‌گیرد و با توجه به دخالت مجموعه داده‌ها در فرآیند ارزیابی نتایج قابل اعتماد



شکل ۷. ساختار فرآیند ارزیابی با روش K_Fold Cross_Validation

معیارهای ارزیابی در هر یک از مدل‌ها با دیگر مدل‌ها مقایسه گردید.

۳. نتایج

ابتدا با آماده‌سازی داده‌ها پارامترهای آماری داده‌های دبی و رسوب در هشت ایستگاه منتخب از استان گیلان مورد بررسی قرار گرفت. نتایج خلاصه آماری هشت ایستگاه در جدول ۱ ارائه شده است.

با اعمال مدل‌های هوش محاسباتی و همچنین تهیه منحنی‌های سنجه رسوب برای هر یک از ایستگاه مطالعاتی نتایج حاصل از پیش‌بینی هر یک از مدل‌ها بر اساس معیارهای ارزیابی مورد بررسی قرار گرفت. نتایج ارزیابی مدل‌ها با استفاده از معیار RMSE در جدول ۲ ارائه شده است. مقادیر RMSE کم نشان‌دهنده میزان خطای کمتر در پیش‌بینی رسوب معلق نسبت به مقادیر واقعی رسوب برداشت شده در ایستگاه‌های هیدرومتری می‌باشد.

نتایج به دست آمده از این معیار حاکی از آن است که

به‌منظور ارزیابی نتایج الگوریتم‌های اعمال شده بر روی داده‌ها از معیارهای ارزیابی مجموع مربعات باقی مانده^۱، ضریب همبستگی^۲ و ضریب ناش-سانتکلیف استفاده شد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (o_i - p)^2}{n}} \quad \text{رابطه (۳)}$$

$$r = \frac{\sum_{i=1}^n (p_i - \bar{p})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2 \sum_{i=1}^n (o_i - \bar{o})^2}} \quad \text{رابطه (۴)}$$

$$NS = 1 - \frac{(p - o_i)^2}{(o_i - \bar{o})^2} \quad \text{رابطه (۵)}$$

که در این معادلات o مقادیر مشاهده‌ای، \bar{o} میانگین داده‌های مشاهده شده، p داده‌های پیش‌بینی شده، \bar{p} میانگین داده‌های پیش‌بینی شده، n تعداد داده‌ها می‌باشد.

با به دست آوردن مقادیر r و $RMSE$ برای کلیه روش‌ها در هر هشت ایستگاه هیدرومتری مقادیر

¹ Root mean squared error (RMSE)

² Correlation coefficient (r)

می‌باشد. البته مقدار خطا در ایستگاه‌های مختلف متفاوت بوده به طوری که مدل فرآیند گوسی در ایستگاه جبرئیل محله دارای حداقل RMSE ۰/۴۲ و در ایستگاه بهارستان دارای حداکثر RMSE ۲/۴۲ می‌باشد.

با توجه به این معیار کلیه مدل‌های هوش محاسباتی دارای مقدار خطای کمتری در پیش‌بینی بار رسوبی معلق نسبت به مدل منحنی سنجه رسوب می‌باشد. از بین مدل‌های هوش محاسباتی ارائه شده نیز مدل فرآیند گوسی دارای نتیجه مناسب‌تری نسبت به سایر مدل‌ها

جدول ۱. خلاصه آماری ایستگاه‌های مورد مطالعه

حوزه	حداقل	حداکثر	میانگین	انحراف معیار	درصد ضریب تغییرات
چوپربالامحله	۰/۰۳۱	۴۶۸۱/۳	۱۰/۴۰	۱۵/۷۲	۱/۵۱
	۰/۰۰۱	۵۱۱۹/۵	۱۱/۳۷	۴۴/۱۲	۳/۸۷
لمیرقربانعلی محله	۰/۲۹۱	۷۸	۱۱/۴۸	۱۵/۰۱	۱/۳۰
	۰/۰۰۱	۴۸۵/۶	۸/۸۲	۳۳/۰۲	۳/۷۴
پیرسرا	۰/۰۷۳	۱۵۵	۳۱/۴۶	۳۶/۶۴	۱/۱۶
	۰/۰۰۵	۳۱۳/۴	۱۴/۶۶	۳۸/۱۷	۲/۶۰
جبرئیل محله	۰/۰۰۵	۱۸۶	۳۸/۵۷	۵۰/۱۲	۱/۲۹
	۰/۰۰۱	۲۸۵/۶	۸/۲۱	۲۶/۲۴	۳/۱۹
کل سر	۰/۰۲۲	۲۴۰	۴۰/۲۳	۴۷/۰۱	۱/۱۶
	۰/۰۱۳	۳۸۶/۳	۳۲/۷۲	۶۲/۷۴	۱/۹۱
کله سرا	۰/۱۰۵	۱۱۰	۲۴/۸۲	۳۱/۵۳	۱/۲۷
	۰/۰۰۱	۵۷۸/۷	۱۳/۷۰	۵۳/۲۳	۳/۸۸
آقامحله	۰/۰۰۷	۲۲۷	۲۷/۵۲	۳۷/۵۵	۱/۳۶
	۰/۰۰۴	۵۰۶/۸	۱۵/۳۸	۴۷/۳۰	۳/۰۷
بهارستان	۰/۰۱	۱۶۰	۵۶/۶۶	۴۶/۵۶	۰/۸۲
	۰/۰۰۱	۲۰۷۴/۹	۱۸/۴۸	۱۱۲/۶۴	۶/۰۹

جدول ۲. معیارهای ارزیابی RMSE (متر مکعب بر ثانیه)

حوزه	ANN	KNN	GP	SVM	ESVM	M5	RRS
چوپربالامحله	۱/۱۵	۰/۹۲	۰/۷۳	۱/۱۸	۲/۲۴	۱/۰۷	۱/۲۴
لمیرقربانعلی محله	۰/۹۰	۰/۷۴	۰/۶۱	۰/۹۳	۱/۸۵	۰/۷۶	۰/۹۲
پیرسرا	۱/۰۳	۰/۸۰	۰/۷۵	۱/۰۸	۱/۸۸	۰/۹۳	۱/۰۸
جبرئیل محله	۰/۷۰	۰/۵۵	۰/۴۲	۰/۷۴	۰/۹۳	۰/۵۸	۰/۷۴
کل سر	۱/۹۷	۱/۲۱	۱/۱۷	۱/۷۲	۲/۴۲	۲/۱۳	۱/۷۴
کله سرا	۱/۴۹	۱/۲۵	۱/۰۶	۱/۵۲	۱/۹۳	۱/۳۴	۱/۵۱
آقامحله	۱/۳۲	۰/۹۸	۰/۸۸	۱/۳۰	۷/۷۵	۱/۰۹	۱/۳۶
بهارستان	۳/۱۷	۲/۵۰	۲/۴۲	۳/۱۸	۶/۷۶	۳/۰۴	۳/۲۳

نیز در ایستگاه‌های مختلف نتایج متفاوتی را نشان داده‌اند. نتایج ارزیابی مدل با استفاده از ضریب همبستگی در جدول ۳ ارائه شده است.

نتایج ارزیابی نتایج با استفاده از ضریب ناش-ساتکلیف در جدول ۴ ارائه شده است.

بررسی ارزیابی مدل‌ها با استفاده از معیار ضریب همبستگی نیز نشان داد که مدل فرآیند گوسی از میان مدل‌های اعمال شده دارای ضریب همبستگی بیشتری بوده است. مدل K نزدیک‌ترین همسایه نیز در اکثر حوزه‌های مطالعاتی دارای ضریب همبستگی بیشتری نسبت به منحنی سنجه رسوب بوده است. سایر مدل‌ها

جدول ۳. معیارهای ارزیابی r

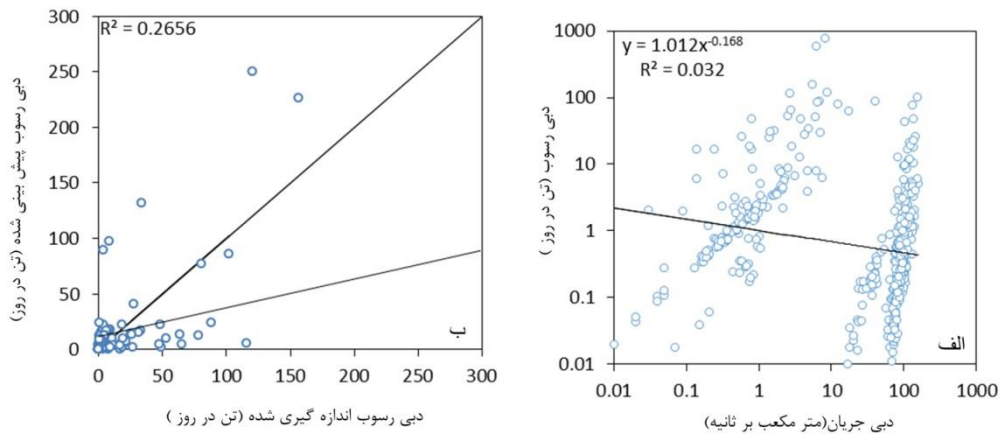
حوزه	ANN	KNN	GP	SVM	ESVM	M5	SRC
چوپربالامحله	۰/۳۸	۰/۶۷	۰/۸۱	۰/۵۱	۰/۵۴	۰/۵۱	۰/۶۵
لمیرقربانعلی محله	۰/۴۰	۰/۶۳	۰/۷۵	۰/۴۳	۰/۵۶	۰/۵۷	۰/۶۱
پیرسرا	۰/۳۰	۰/۶۶	۰/۷۱	۰/۴۳	۰/۲۱	۰/۵۰	۰/۵۶
جبرئیل محله	۰/۲۹	۰/۶۶	۰/۸۲	۰/۱۰	۰/۶۶	۰/۶۴	۰/۷۱
کل سر	۰/۳۳	۰/۷۳	۰/۷۴	۰/۴۴	۰/۵۰	۰/۶۳	۰/۶۶
کله سرا	۰/۲۱	۰/۵۸	۰/۷۰	۰/۲۴	۰/۱۳	۰/۴۶	۰/۶۱
آقامحله	۰/۲۲	۰/۶۷	۰/۷۵	۰/۴۷	۰/۰۹	۰/۵۸	۰/۴۷
بهارستان	۰/۰۸	۰/۶۱	۰/۵۱	۰/۱۰	۰/۶۰	۰/۲۹	۰/۱۷

جدول ۴. معیارهای ارزیابی ناش-ساتکلیف

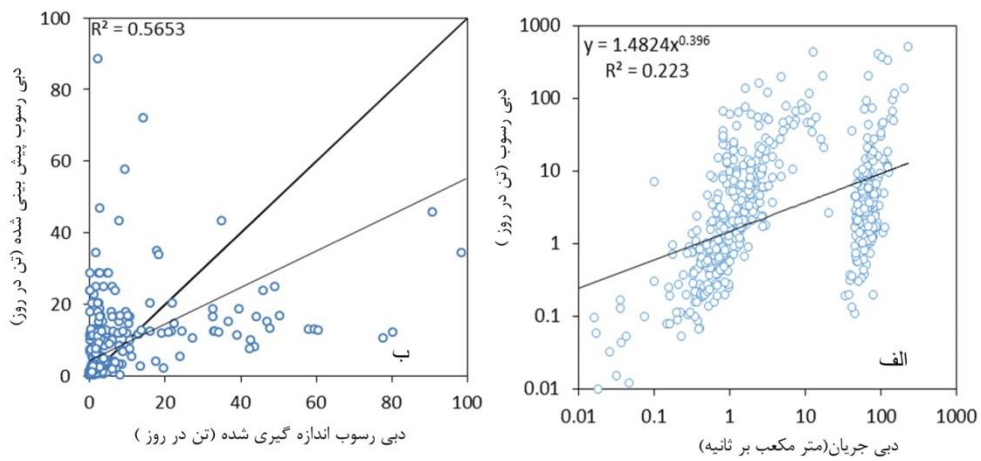
حوزه	ANN	KNN	GP	SVM	ESVM	M5	SRC
چوپربالامحله	۰/۱۷	۰/۴۷	۰/۲۰	۰/۶۶	۰/۱۳	-۲/۱۲	۰/۲۹
لمیرقربانعلی محله	۰/۰۶	۰/۳۶	۰/۱۸	۰/۵۵	۰/۰۰۴	-۲/۹۳	۰/۳۲
پیرسرا	۰/۰۸	۰/۴۴	۰/۱۱	۰/۵۱	-۰/۰۱۵	-۲/۰۵	۰/۲۵
جبرئیل محله	۰/۰۸	۰/۴۴	۰/۰۹	۰/۶۶	۰/۰۰۴	-۰/۵۸	۰/۳۷
کل سر	۰/۰۸	۰/۵۳	۰/۱۳	۰/۵۵	۰/۰۵۵	-۰/۸۶	۰/۳۹
کله سرا	۰/۰۱۹	۰/۳۰	۰/۵۴	۰/۵	-۰/۰۲	-۰/۶۴	۰/۱۹
آقامحله	۰/۹۵	۰/۹۵	۰/۹۷	۰/۹۵	۰/۹۵	۰/۱۴	۰/۹۵
بهارستان	۰/۸۲	۰/۸۹	۰/۸۳	۰/۸۷	۰/۸۳	۰/۲۳	۰/۸۴

با توجه به برتری مدل فرآیند گوسی نسبت به سایر مدل‌های هوش محاسباتی در مقایسه با منحنی سنجه رسوب و منحنی پراکندگی مقادیر مشاهده‌ای رسوب معلق و پیش‌بینی شده آن، در این دو مدل در هر یک از ایستگاه‌ها در شکل‌های ۲ تا ۹ ارائه می‌گردد.

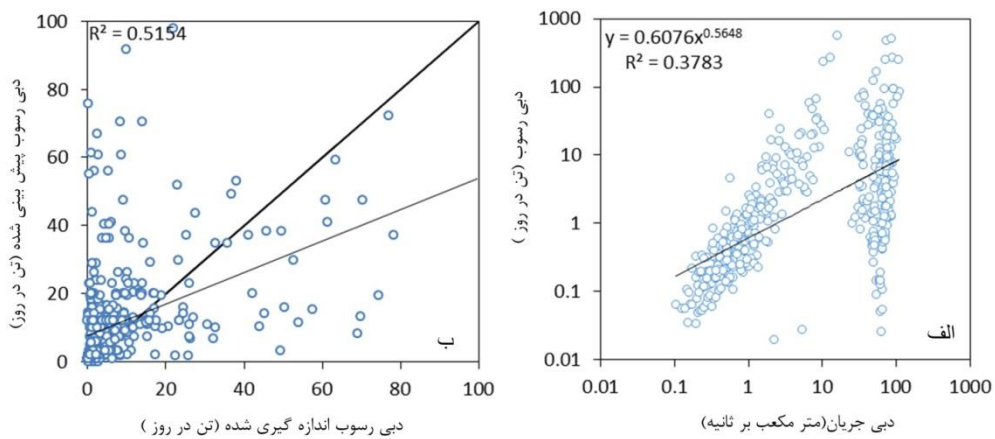
نتایج بررسی این معیار نیز نشان داده است که مدل فرآیند گوسی نسبت به سایر مدل‌ها از دقت بالاتری برخوردار بوده است. البته میزان دقت در ایستگاه‌های مختلف متفاوت بوده است. هرچه مقادیر ضریب ناش-ساتکلیف به مقدار یک نزدیک تر باشد نشان از دقت بالای مدل و خطای کمتر در برآورد را دارد.



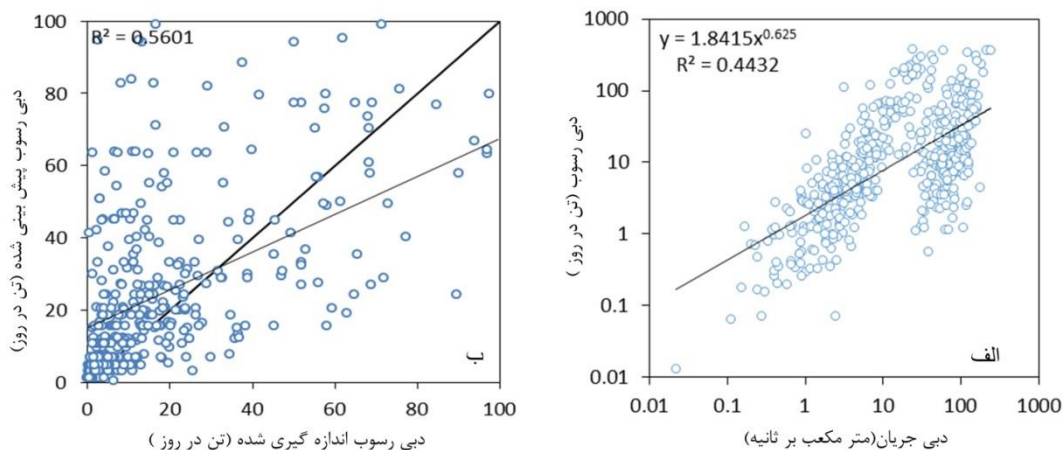
شکل ۸. الف: منحنی سنجه رسوب، ب: منحنی پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه بهارستان



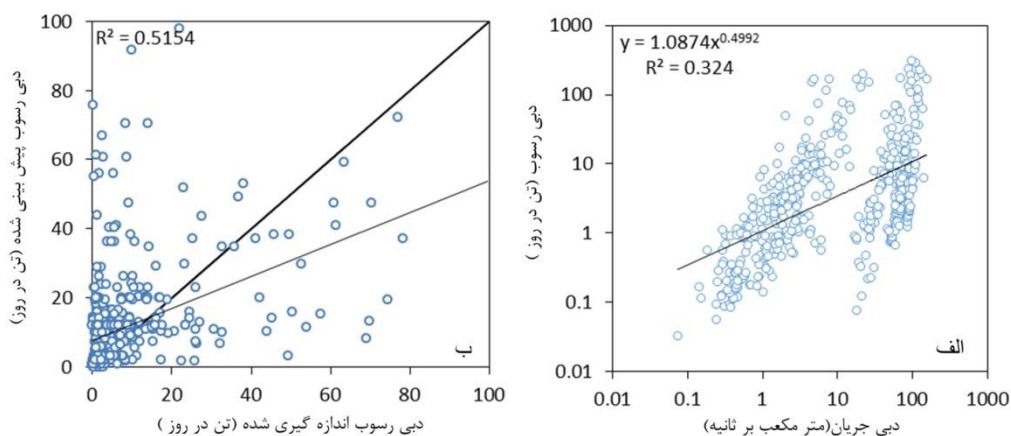
شکل ۹. الف: منحنی سنجه رسوب، ب: منحنی پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه آقامحله



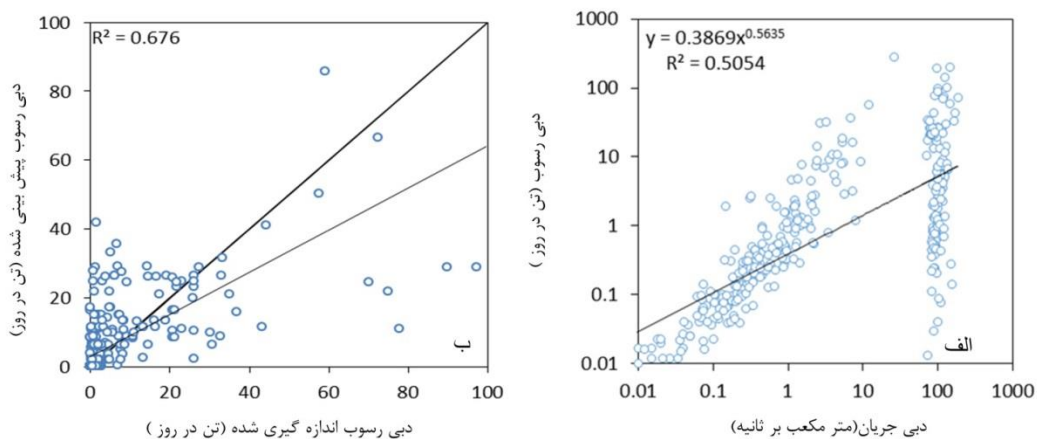
شکل ۱۰. الف: منحنی سنجه رسوب، ب: منحنی پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه کله سرا



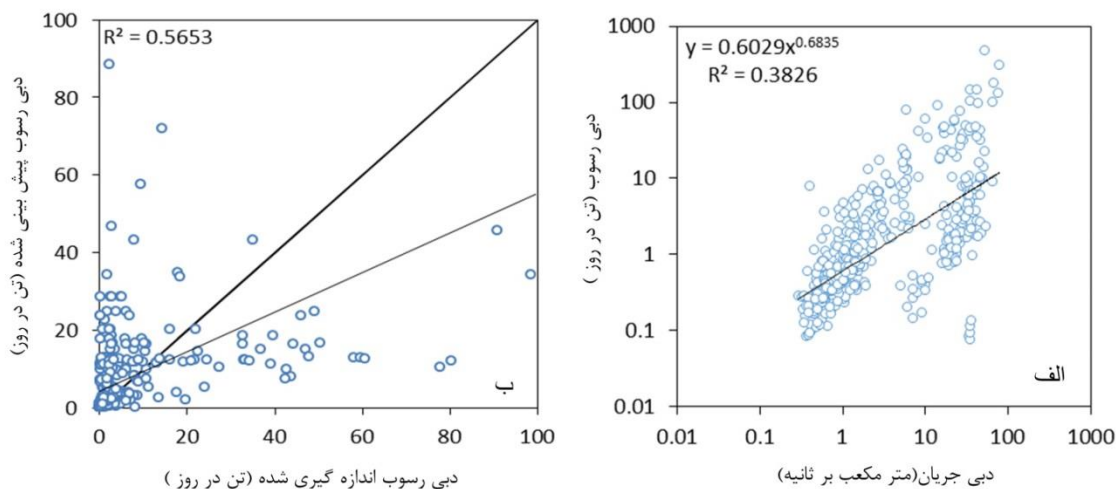
شکل ۱۱. الف: منحنی سنجه رسوب، ب: منحنی پراکنگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه کل سر



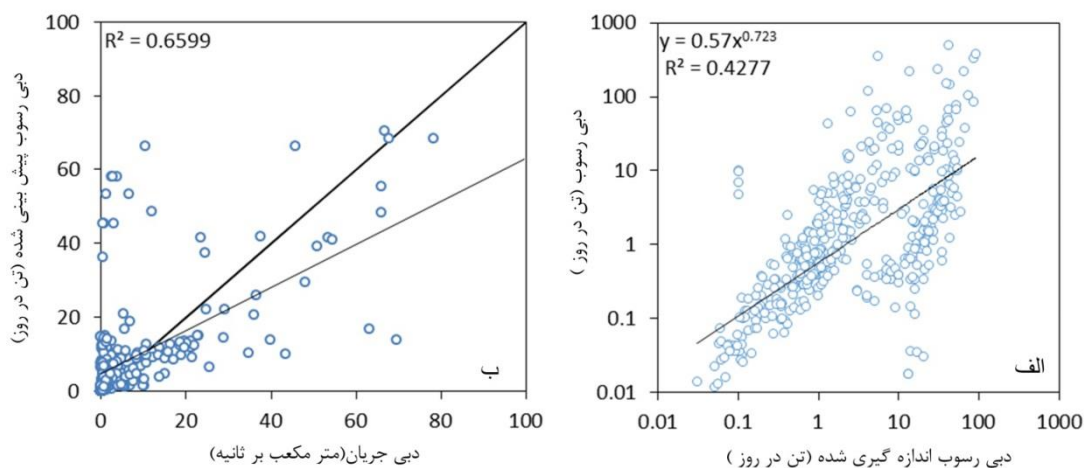
شکل ۱۲. الف: منحنی سنجه رسوب، ب: منحنی پراکنگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه پیرسا



شکل ۱۳. الف: منحنی سنجه رسوب، ب: منحنی پراکنگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه جبرئیل محله



شکل ۱۴. الف: منحنی سنجۀ رسوب، ب: منحنی پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه لمیرقربانعلی محله



شکل ۱۵. الف: منحنی سنجۀ رسوب، ب: منحنی پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده در ایستگاه چوپربالامحله

۴. بحث و نتیجه‌گیری

در این مطالعه به ارزیابی شش مدل مبتنی بر هوش محاسباتی، K نزدیک‌ترین همسایه، شبکه عصبی پس انتشار خطا، فرآیند گوسی، درخت تصمیم‌گیری M5، ماشین بردار پشتیبان و ماشین بردار پشتیبان تکاملی با مدل منحنی سنجۀ رسوب پرداخته شد. مدل‌های هوش محاسباتی با توجه به توانایی‌شان در برقراری ارتباط بین

همانطور که نمودارها پراکندگی مقادیر مشاهده‌ای و پیش‌بینی شده با استفاده از دو روش منحنی سنجه و فرآیند گوسی را نشان می‌دهد نمودارها نیز نشان از همبستگی بیشتر مقادیر مشاهده‌ای و پیش‌بینی شده در مدل فرآیند گوسی دارد. به طور مثال در حوزه مطالعاتی چوپربالامحله میزان همبستگی در مدل فرآیند گوسی ۰/۶۵ درصد و در منحنی سنجۀ رسوب ۰/۴۲ محاسبه گردیده است.

داده مدل‌ها دارای مقدار ضریب همبستگی کمتر یا بیشتر نسبت به مدل منحنی سنجه رسوب می‌باشد. از آنجا که روش‌های هوش مصنوعی برای تحلیل مسائلی به کار می‌روند که شناخت و توصیف صریحی از ماهیت مسأله وجود ندارد، بنابراین بسیاری از مسائل مربوط به رسوبات معلق در رودخانه‌ها را می‌توان با این روش‌ها حل نمود. اما شاید برخی مدل‌ها به‌درستی نتوانند روابط درست بین پدیده‌ها را تعیین نمایند و تعیین مدل مناسب با دقت بالا امری مهم تلقی می‌گردد. بنابراین می‌توان گفت مدل فرآیند گوسی با توجه به داشتن ساختار غیرپارامتری قوی ابزار مناسبی جهت ایجاد مدل‌های احتمال‌گرایانه جامع از مسائل دنیای واقعی می‌باشد و این مدل رسوب را با دقت بالاتری نسبت به سایر مدل‌ها برآورد می‌نماید. از سوی دیگر یکی از مشکلات موجود در منحنی‌های سنجه رسوب که در نمودارهای ارائه شده نیز کامل مشخص می‌باشد این است که کلیه داده‌های دبی و رسوب در منحنی سنجه رسوب از یک روند پیروی نکرده که این امر موجب دو قسمتی بودن منحنی سنجه رسوب گردیده است. در واقع با افزایش مقادیر دبی و رسوب از یک مقدار مشخص روند اولیه داده‌ها برهم خورده و داده‌ها از شیب تندتری پیروی می‌نمایند که این امر موجب افزایش خطا در منحنی‌های سنجه رسوب می‌گردد [۷]. به منظور برطرف‌سازی و یا کاهش این خطا می‌توان برای هر دو مجموعه داده منحنی سنجه جداگانه‌ای تهیه نمود. از طرف دیگر با توجه به تأثیر سایر پارامترهای منطقه‌ای نظیر پوشش گیاهی، بارندگی و مورفولوژی رودخانه بر رسوب‌دهی حوزه‌های آبخیز لازم است تا با بهینه‌سازی مدل‌ها با استفاده از پارامترهای مؤثر در بار رسوب دقت مدل‌ها را افزایش داد.

پارامترهای غیرخطی به خوبی قادر به پیش‌بینی بار رسوبی معلق می‌باشند. به طوری که در مطالعه [۳] نتایج، برتری شبکه عصبی را نسبت به منحنی سنجه رسوب به‌همراه داشته است و در مطالعه [۱۲] شبکه عصبی موجک برتری خود را نسبت به منحنی سنجه رسوب نشان داده است و در مطالعه [۲۳] مدل SVR دارای نتایج بهتری نسبت به منحنی سنجه رسوب بوده است. آنچه در این مطالعات کمتر به آن توجه شده است متفاوت بودن نتایج پیش‌بینی مدل‌ها در حوزه‌های آبخیز مختلف می‌باشد. همچنین در اکثر مطالعات تعداد محدودی از مدل‌های هوش محاسباتی با منحنی رسوب مقایسه گردیده است در صورتی مدل‌های مختلف می‌تواند رفتارهای متفاوتی را در پیش‌بینی بار رسوبی معلق از خود نشان دهند. لذا در این مطالعه شش مدل مذکور در هشت حوزه آبخیز مختلف واقع در استان گیلان مورد ارزیابی و بررسی قرار گرفتند. در این مطالعه ارزیابی مقادیر RMSE نشان داد که تمامی مدل‌های هوش محاسباتی ارائه شده در مطالعه دارای مقدار RMSE کمتری نسبت به مدل سنجه رسوب می‌باشند به طور مثال در حوزه کله‌سرا مقادیر RMSE مدل‌های شبکه عصبی ۱/۴۹، نزدیک‌ترین همسایه ۱/۲۵، فرآیند گوسی ۱/۰۶، ماشین‌بردار پشتیبان ۱/۵۲، ماشین‌بردار پشتیبان تکاملی ۱/۹۳، درخت تصمیم‌گیری ۱/۳۴ و منحنی سنجه ۱/۵۱ به دست آمد. دلیل این امر را می‌توان مربوط به ساختار غیرخطی بار رسوب معلق دانست. همچنین مقادیر همبستگی در مدل‌ها نشان داد که مدل فرآیند گوسی در تمامی ایستگاه‌ها دارای بیشترین مقدار نسبت به سایر مدل‌ها هوش محاسباتی و همچنین منحنی سنجه رسوب می‌باشد. اما در رابطه با سایر مدل‌ها در برخی ایستگاه‌ها بر حسب تعداد نمونه

References

- [1] Akbari, Z. (2010). Performance of the decision tree and regression model to estimate the amount of sediment in the dam area of Ilam, Master's thesis, Department of Natural Resources desert Studies, Yazd University (In Persian).
- [2] Alberg, D., Last, M. and Kandel, A. (2012). Knowledge discovery in data streams with regression tree methods. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1):69-78.
- [3] Alp, M. and Cigizoglu, H.K. (2007). Suspended sediment load simulation by two artificial neural network methodemis using hydrometeorological data. *Environ. Modell. Software* 22 (1): 2–13.
- [4] Azmi, M. and Araghinejad, M. (2012). Development of K-Nearest Neighbour Regression Method in Forecasting River Stream Flow. VO(3).PP.108-119.(In Persian)
- [5] Cheng, M.Y. and Wu, Y.W. (2009). Evolutionary support vector machine inference system for construction management. *Automation in Construction*, 18(5):.597-604.
- [6] Cigizoglu, H. K. and Alp, M. (2006). Generalized regression neural network in modeling river sediment yield. *Advances in Engineering Software*, Vol 37: 63-68.
- [7] Dehghani, A.A., Zanganeh, M.E., Mosaedi, A. and Kohestani, N. (2009). Comparison of Suspended Sediment Estimation by Artificial Neural Network and Sediment Rating Curve Methods (Case Study: Doogh River in Golestan Province). *J. Agric. Sci. Natur. Resour.*, Vol. 16 (Special issue 1-a) (In Persian).
- [8] Harrington, S.T. and Harrington, J.R. (2013). An assessment of the suspended sediment rating curve approach for load estimation on the Rivers Bandon and Owenabue, Ireland. *Geomorphology*, 185, pp.27-38.
- [9] Hornik, K., Stinchcombe, M. and White, H. (1989), "Multilayer feedforward networks are universal approximators". *Neural networks*, 2(5):359-366.
- [10] Huang, H.L. and Chang, F.L. (2007), "ESVM: Evolutionary support vector machine for automatic feature selection and classification of microarray data", *Biosystems*, 90(2):516-528.
- [11] Khazaie Poul . A. and Talebi . A. (2013) Investigation of Possibility of Suspended Sediment Prediction Using The Combination of Sediment Rating Curve and Artificial Neural Network (Case Study: Ghatorchai River, Yazdakan Bridge). *Quarterly Journal of Environmental Erosion Researches* No. 9: 73-82(In Persian).
- [12] Kisi, O. (2012). Modeling discharge-suspended sediment relationship using least square support vector machine. *Journal of hydrology*, 456, pp.110-120.
- [13] Rajae, T. (2011) Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Science of the total environment*, 409(15):2917-2928.
- [14] Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*, the MIT Press, , ISBN 026218253X. c. (2006). Massachusetts Institute of Technology. www.GaussianProcess.org/gpml
- [15] Sani Abade, M., Mahmoudi, S. and Taherparvar, D. (2014). *Data mining applications (second edition)*, Niaz-e-Danesh Pub.Tehran. (In Persian).
- [16] Solomatine, D.P. and Xue, Y. (2004). M5 model trees and neural networks: Application to flood Forecasting in the upper reach of the Huai river in China. *J. Hydr. Engine.* 9: 6. 1-10.
- [17] USEPA. (2000). *The Quality of Our Nation's Waters. A Summary of the National Water Quality Inventory: 1998 Report to Congress*, Office of Water, 841-S-00-001, Washington, DC., 19 p .
- [18] Witten, I.H. and Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- [19] Yousefi, M. and Barzegari, F. (2014). Estimation of suspeded sediment using decision tree method and copmarision the results with rating curve method (Case study: Rivers in Lorestan Province). *Watershed Management Research (Pajouhesh & Sazandegi)* No 104: 187-195.
- [20] Zahiri1, A.R. and Ghorbani, Kh. (2013). Flow discharge prediction in compound channels by using decision model tree M5. *J. of Water and Soil Conservation*, ol. 2V0(3):113-132.

- [21] Zhang, W., Wei. X., Zheng. J., Zhu. Y. and Zhang. Y. (2012). Estimating suspended sediment loads in the Pearl River Delta region using sediment rating curves. *Geomorphology* 185 :27–38
- [22] Zhu, Y.M., Lu, X.X. and Zhou, Y. (2007). Suspended sediment flux modeling with artificial neural network: an example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology*, 84(1):111-125.
- [23] Zounemat-Kermani, M., Kişi, Ö., Adamowski, J. and Ramezani-Charmahineh, A. (2016). Evaluation of data driven models for river suspended sediment concentration modeling. *Journal of Hydrology*, 535:457-472.