

ارزیابی عملکرد روش‌های مبتنی بر کرنل در تخمین میزان بار رسوبی معلق رودخانه

(مطالعه موردی: رودخانه صوفی‌چای مراغه)

علی‌رضازاده جودی* - باشگاه پژوهشگران جوان و نخبگان، واحد مراغه، دانشگاه آزاد اسلامی، مراغه، ایران

محمدتقی ستاری - عضو هیئت‌علمی گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

پذیرش مقاله: ۱۳۹۵/۰۲/۱۳ تأیید نهایی: ۱۳۹۵/۰۴/۸

چکیده

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. در این پژوهش میزان بار رسوبی معلق رودخانه صوفی‌چای به وسیله روش‌های نوین داده‌کاوی - شامل رگرسیون فرایند گاوسی و رگرسیون بردار پشتیبان - که با بهره‌گیری از توابع کرنل توانایی بسیاری در حل مسائل غیرخطی دارند، تخمین زده شد، سپس، با مقادیر به‌دست‌آمده از روش‌های تجربی منحنی سنجه رسوب و روش فصلی مقایسه شد. روش رگرسیون فرایند گاوسی با ارائه شاخص‌های آماری ضریب همبستگی (R) برابر ۰٫۹۷۷، ضریب همبستگی نش - ساتکلیف (N-S) برابر ۰٫۷۹۴، میانگین خطای مطلق (MAE) برابر ۷۷٫۴۲۷۸ تن در روز، و ریشه میانگین مربعات خطا (RMSE) برابر ۶۹۸٫۷۴۵۵ تن در روز دارای بیشترین دقت و کمترین خطا از میان روش‌های بررسی‌شده در این مطالعه است. نتایج به‌دست‌آمده نشان داد هر دو روش داده‌کاوی بررسی‌شده رگرسیون فرایند گاوسی و رگرسیون بردار پشتیبان به مراتب نتایج بهتری نسبت به منحنی سنجه رسوب و روش فصلی ارائه می‌کنند.

کلیدواژه‌ها: بار رسوبی معلق، رگرسیون بردار پشتیبان، رگرسیون فرایند گاوسی، صوفی‌چای، منحنی سنجه رسوب.

مقدمه

رسوباتی که به وسیله آب رودخانه‌ها حمل می‌شود شامل دو گروه است: ۱. رسوباتی که در آب معلق‌اند و در بالای بستر رودخانه‌ها همراه با آب حرکت می‌کنند؛ به این مواد «بار معلق» گفته می‌شود؛ ۲. رسوباتی که در سطح بستر رودخانه همراه با جریان آب به صورت غلطش، جهش، و لغزش حرکت می‌کنند و «بار بستر» نامیده می‌شوند (نجمایی، ۱۳۶۹). برآورد صحیح حجم رسوبات آبی در یک حوضه آبریز از مهم‌ترین مسائل مطرح‌شده در مهندسی آب، مهندسی رودخانه، منابع آب، تأسیسات و سازه‌های آبی، و محیط زیست برای اجرای بسیاری از طرح‌ها و برنامه‌های توسعه‌ای است. در پژوهش‌های رسوب‌دهی آبریزها به مطالعه غلظت رسوب معلق^۱ توجه خاصی شده است، زیرا بار معلق شاخصی از

* E- mail: alijoudi66@gmail.com

رسوب‌دهی کل سطح حوضه است و، علاوه بر این، بر اثر شست‌وشوی اراضی حاصل‌خیز بالادست شکل می‌گیرد (صادقی و همکاران، ۲۰۰۸: ۱۹۹). بدین منظور، با توجه به اهمیت مسئله، تا کنون محققان بسیاری در سراسر جهان به مطالعه دربارهٔ این موضوع پرداخته و روش‌ها و روابط متعددی را در این زمینه توسعه داده‌اند. شماری از این روابط بر پایهٔ معادلات بار بستر ارائه شده است، لیکن، افزون بر نیاز به شرایط ویژه، مقدار محاسبه‌شدهٔ بار کل رسوب به کمک این روابط معمولاً با مقدار اندازه‌گیری‌شدهٔ بار رسوب کل متفاوت است. دلیل این مسئله می‌تواند تأثیر بار معلق بر مقدار بار کل باشد که معادلات ارائه‌شده قادر به محاسبهٔ آن نیستند (امامی، ۱۳۷۹: ۲۸۵). از سوی دیگر، روابط موجود تحت شرایط خاصی گسترش یافته‌اند و استفاده از آن‌ها در شرایطی دیگر نیازمند بررسی و مقایسهٔ نتایج با ارقام اندازه‌گیری‌شده است. یکی از روش‌های متداول محاسبهٔ بار معلق رودخانه استفاده از منحنی سنجهٔ رسوب^۱ است. این رابطه به دلیل داشتن اریب^۲ در بیشتر مواقع نتوانسته غلظت رسوب در دبی‌های مختلف جریان را به‌خوبی نشان دهد. اریب در واقع باعث شده تا مقدار باقی‌مانده‌ها (اختلاف بین مقادیر مشاهده‌ای و محاسباتی) توزیع نرمالی نداشته باشد (کائو و همکاران، ۲۰۰۵: ۶۵۵). بیشتر روش‌ها و روابط موجود محدودیت‌های خاصی دارند؛ از جمله نیاز به حجم وسیعی از متغیرهای مختلف؛ فقدان دقت کافی و وجود خطا در مقادیر دبی‌های مختلف؛ و محدودبودن به شرایط جغرافیایی خاص. این مسئله موجب شده تا در دههٔ اخیر برای یافتن روشی که محدودیت‌ها و مشکلات موجود در روابط و روش‌های کلاسیک متداول را نداشته باشد بیشتر تلاش شود؛ از جملهٔ این روش‌ها می‌توان به روش‌های نوین داده‌کاوی اشاره کرد. اساس این روش‌ها پیدا کردن الگوها و روابط پنهان بین داده‌هاست. آلپ و جیغیز اوغلو (۲۰۰۷) در پژوهشی از دو الگوریتم شبکهٔ عصبی مصنوعی - شامل توابع پایهٔ شعاعی و الگوریتم پیش‌خور پس‌انتشار خطا - برای مدل‌سازی بار معلق رسوب در حوضهٔ آبریز جونیاتای^۳ آمریکا با استفاده از داده‌های جریان بارندگی و داده‌های بار رسوب معلق استفاده کردند. نتایج نشان‌دهندهٔ برتری روش شبکهٔ عصبی مصنوعی نسبت به روش رگرسیون خطی در این زمینه بود. رجایی و همکاران (۲۰۱۰) با استفاده از مدل ترکیبی موجک و عصبی - فازی به تخمین میزان بار رسوبی معلق روزانه پرداختند و توانایی قابل قبول این روش را در مقابل منحنی سنجه نشان دادند. ایدر و همکاران (۲۰۱۰)، با توجه به اثرهای پسماند، به برآورد بار معلق در رودخانه پرداختند. نتایج نشان داد از بین روش‌های یادشده روش منحنی سنجهٔ رسوب دقت کمتری دارد. اوندراک و همکاران (۲۰۱۲) با استفاده از مدل‌سازی چندمتغیره به تخمین بار معلق در بالادست حوضه اقدام کردند و مشاهده نمودند که مدل درختی ام^۴ ۵ قابلیت خوبی در تخمین بار رسوبی دارد. کیا و همکاران (۲۰۱۳) با استفاده از شبکه‌های عصبی مصنوعی به تخمین میزان بار معلق رودخانهٔ بابل‌رود پرداختند. نتایج نشان داد بهترین نتیجه مربوط به استفاده از شبکهٔ عصبی پرسپترون چندلایه است و در حالت کلی نتایج شبکهٔ عصبی بهتر از نتایج منحنی سنجهٔ رسوب است. کاکایی لخدانی و همکاران (۲۰۱۳) با استفاده از دو مدل شبکهٔ عصبی مصنوعی و ماشین بردار پشتیبان به پیش‌بینی بار معلق رسوب روزانهٔ رودخانه پرداختند و نتایج را با مدل‌های رگرسیونی مقایسه کردند. نتایج نشان داد

1. suspended rating curve

2. bias

3. Juniata

4. m5 model tree

مدل‌های رگرسیونی عملکرد ضعیف‌تری نسبت به دو مدل دیگر در پیش‌بینی رسوب دارند. هنگ و سوئسوگی (۲۰۱۳) کارایی شبکه عصبی مصنوعی در تخمین بار رسوبی حوضه آبخیز رودخانه تونله‌سپ در کامبوج را ارزیابی کردند. نتایج بیانگر عملکرد قابل قبول شبکه عصبی مصنوعی در این زمینه است. کومار گوپال (۲۰۱۴) عملکرد مدل درختی ام ۵ و رگرسیون موجکی^۱ را در مقایسه با شبکه عصبی مصنوعی برای پیش‌بینی آورد رسوبی در حوضه آبریز ناگوا واقع در هند بررسی کرد. نتایج نشان‌دهنده برتری مدل درختی ام ۵ و رگرسیون موجکی است. دوان و همکاران (۲۰۱۵) با استفاده از روش SPARROW به مدل‌سازی منابع رسوب معلق و انتقال آن‌ها در حوضه آیشیکاری^۲ پرداختند و نحوه تولید و انتقال رسوب را در این حوضه مشخص کردند و اعلام نمودند این روش می‌تواند برای مدیریت منابع آب به‌خوبی به‌کار رود. دهقانی و همکاران (۱۳۸۸) با استفاده از شبکه عصبی مصنوعی و منحنی سنجه به تخمین بار معلق رسوبی رودخانه دوغ در استان گلستان پرداختند و اعلام کردند که شبکه عصبی مصنوعی نسبت به منحنی سنجه نتایج بهتری ارائه می‌کند. ولی و همکاران (۱۳۹۰) به پیش‌بینی بار رسوب جریان در حوضه آبریز سمندگان با استفاده از شبکه‌های عصبی مصنوعی و رگرسیون پرداختند. نتایج نشان داد شبکه عصبی ژئومورفولوژی در مقایسه با شبکه عصبی غیرژئومورفولوژی و مدل‌های رگرسیونی ژئومورفولوژی و غیرژئومورفولوژی دقت بیشتری دارد. فلامکی و همکاران (۱۳۹۲) بار رسوب رودخانه را با استفاده از شبکه عصبی مصنوعی مدل‌سازی کردند. نتایج نشان داد مدل‌های ارائه‌شده به وسیله شبکه عصبی مصنوعی در مقایسه با روابط تجربی نتایج دقیق‌تری ارائه می‌کنند. دهقانی و وفاخواه (۱۳۹۲) نتایج به‌دست‌آمده از دو نوع شبکه عصبی مصنوعی و منحنی‌های سنجه مختلف در حوضه آبریز گرگان‌رود را با یکدیگر مقایسه کردند. نتایج این پژوهش نیز نشان داد شبکه‌های پرسپترون چندلایه با ورودی‌های دبی همان روز و دبی یک و دو روز قبل نسبت به سایر روش‌ها از دقت بیشتری در تخمین میزان بار رسوب معلق برخوردار است. طباطبایی و همکاران (۱۳۹۳) با استفاده از شبکه عصبی مصنوعی و خوشه‌بندی داده‌ها به روش نگاشت خودسازمان‌ده^۳ به برآورد غلظت رسوب معلق روزانه در رودخانه کرج پرداختند و دریافتند که استفاده از متغیرهای بارش و دمای روزانه، دبی جریان، و تفکیک زمانی داده‌ها نقش مهمی در افزایش دقت برآورد رسوب رودخانه داشته است. رجبی و همکاران (۱۳۹۴) با استفاده از مدل تبرید تدریجی عصبی^۴ (NDE) و مدل‌های ANFIS و RBF^۵ به تخمین بار معلق رسوبی رودخانه گیوی‌چای پرداختند و دریافتند که مدل تبرید تدریجی قادر به ارائه نتایج دقیق‌تری است. رضازاده جودی و ستاری (۱۳۹۴) با استفاده از روش رگرسیون فرایند گاوسی به پیش‌بینی عمق چاله آب‌شستگی اطراف پایه پل پرداختند و نتایج را با هشت رابطه تجربی مرسوم در این زمینه مقایسه کردند. نتایج نشان‌دهنده کارایی بالای روش رگرسیون فرایند گاوسی در مقایسه با روابط تجربی بررسی شده است. بررسی منابع انجام‌یافته تا کنون نشان می‌دهد گرچه تحقیقات بسیاری درباره کارایی و عملکرد شبکه‌های عصبی مصنوعی در زمینه برآورد میزان بار معلق رودخانه‌ها انجام پذیرفته است، در ایران در زمینه کاربرد سایر

1. wavelet regression

2. Ishikari

3. self organized map

4. neural differential evolution

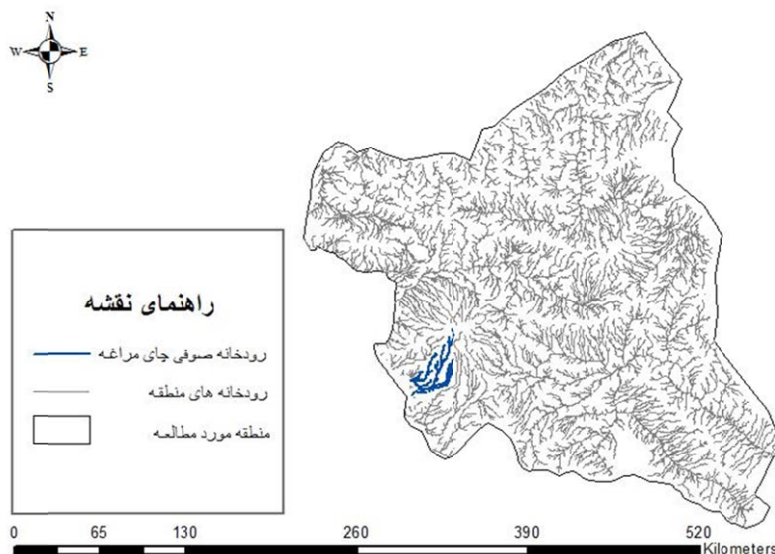
5. radial basis function

روش‌های نوین داده‌کاوی مبتنی بر تابع کرنل- از جمله مدل‌های رگرسیون فرایند گاوسی^۱ و رگرسیون بردار پشتیبان^۲- مطالعه‌ی چندانی صورت نپذیرفته است. بنابراین، هدف از این تحقیق بررسی کارایی و عملکرد روش‌های رگرسیون فرایند گاوسی و رگرسیون بردار پشتیبان در برآورد میزان دبی رسوب معلق رودخانه صوفی‌چای واقع در استان آذربایجان شرقی و مقایسه آن‌ها با دو روش منحنی‌سنج رسوب و روش فصلی به عنوان روش‌های متداول کلاسیک در برآورد میزان دبی رسوب معلق است.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌های مورد استفاده

حوضه آبریز صوفی‌چای به مساحت ۳۱۱ کیلومتر مربع در جنوب استان آذربایجان شرقی و شمال شهر مراغه قرار دارد. این رودخانه از دامنه‌های جنوبی کوه‌های سه‌سند سرچشمه می‌گیرد و از شمال به حوضه آبریز لیقوان و کندچای، از شرق به حوضه آبریز مردق‌چای، از غرب به حوضه آبریز قلعه‌چای، و از جنوب پس از عبور از محدوده سد علویان به دریاچه ارومیه منتهی می‌شود. رودخانه صوفی‌چای در محدوده مختصات جغرافیایی $37^{\circ} 15' 2''$ تا $37^{\circ} 45' 3''$ عرض شمالی و $45^{\circ} 56' 31''$ تا $45^{\circ} 25' 5''$ طول شرقی قرار گرفته است. شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه را نشان می‌دهد.



شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه

1. gaussian process regression
2. support vector regression

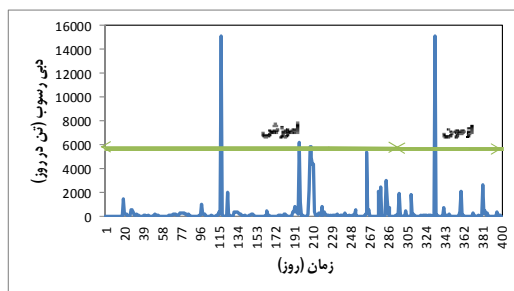
در این تحقیق از داده‌های مربوط به پارامترهای دبی جریان (متر مکعب در ثانیه)، غلظت بار معلق (میلی‌گرم در لیتر)، دبی وزنی رسوب (تن در روز) رودخانه صوفی‌چای مربوط به سال‌های ۱۳۶۳ تا ۱۳۹۳، که در مجموع شامل ۴۰۲ مجموعه داده است، استفاده شد. این داده‌ها از شرکت آب منطقه‌ای استان آذربایجان شرقی اخذ شده است. از این تعداد داده، ۷۵ درصد داده‌ها برای بخش آموزش (از ابتدای مجموعه داده‌ها یعنی از سال ۱۳۶۳ تا ۱۳۸۵) و ۲۵ درصد داده‌ها برای بخش آزمون مدل (از انتهای مجموعه داده یعنی از سال ۱۳۸۵ تا ۱۳۹۳) در نظر گرفته شد. جدول ۱ مشخصات آماری داده‌های مورد استفاده را نشان می‌دهد.

جدول ۱. مشخصات آماری داده‌های مورد استفاده

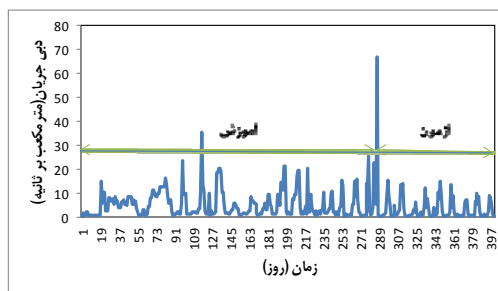
داده‌های بخش آزمون	داده‌های بخش آموزش			داده‌های بخش آزمون		
	دبی رسوب (ton/day)	غلظت مواد معلق (Mg/litr)	دبی جریان (m ³ /s)	دبی رسوب (ton/day)	غلظت مواد معلق (Mg/litr)	دبی جریان (m ³ /s)
حد اقل	۰٫۰۳	۰٫۳۳	۰٫۱۴	۰٫۰۱۸	۰٫۲۵	۰٫۰۶
حداکثر	۱۵۱۴۵٫۳۹	۱۴۳۵۶٫۵۸	۱۵٫۱۴	۱۵۱۳۸٫۲۱	۷۱۰۲٫۶۷	۶۷
میانگین	۲۷۰٫۰۴	۴۵۳٫۴	۳٫۱۵	۳۱۴٫۴۳	۳۲۹٫۷۶	۵٫۶۲
انحراف معیار	۱۵۵۰٫۰۶	۱۶۰۹٫۵۲۱	۳٫۷۱	۱۲۱۳٫۵۲	۸۲۸٫۹۱	۶٫۵۶

با توجه به اینکه به کار بردن داده‌های استاندارد و همگن در میزان دقت و کارایی مدل‌های داده‌محور تأثیر منفی می‌گذارد و موجب کاهش قدرت تعمیم‌دهی آن‌ها می‌شود، در تحقیق حاضر، پیش از اقدام به مدل‌سازی میزان دبی رسوب رودخانه صوفی‌چای، همه داده‌ها از نظر همگن بودن و استاندارد بودن بررسی شد. بر این اساس، همگنی داده‌های موجود با استفاده از نرم‌افزار XLSTAT به وسیله آزمون‌های همگنی نرمال استاندارد^۱، دامنه بیش‌ند^۲، پیتیت^۳، و نسبت فون نیومن^۴ آزمون شد. شکل‌های ۲ و ۳ تغییرات زمانی میزان دبی جریان و دبی رسوبی معلق در رودخانه صوفی‌چای مراغه را نشان می‌دهند. با توجه به اینکه سری زمانی داده‌های انتخاب‌شده برای آموزش مربوط به دوره‌های نسبتاً تر و برعکس داده‌های انتخاب‌شده برای دوره آزمون مربوط به دوره‌های اخیر و نسبتاً خشک است، به نظر می‌رسد اختلاف بین ویژگی‌های آماری از این منظر قابل ارزیابی باشد. از سویی دیگر، تغییرات کاربری اراضی و توسعه سطح زیر کشت در طول بازه زمانی مورد مطالعه می‌تواند بر این موضوع تأثیر بگذارد.

1. standard normal homogeneity test
2. buishand range homogeneity test
3. pettitt homogeneity test
4. von neumann ratio test



شکل ۳. نمودار تغییرات زمانی دبی رسوب معلق



شکل ۴. نمودار تغییرات زمانی دبی جریان

روش منحنی سنجه رسوب یک‌خطی یا روش (USBR¹)

به سبب در دسترس نبودن آمار دقیق میزان فرسایش و رسوب در حوضه‌های آبریز، در بیشتر مواقع از منحنی سنجه رسوب تهیه‌شده با داده‌های دبی و غلظت رسوب یا دبی رسوب معلق استفاده می‌شود. شکل کلی منحنی سنجه رسوب به شکل رابطه ۱ است:

$$Q_s = a Q_w^b$$

رابطه ۱

که در آن Q_w برابر دبی آبراهه (m^3/s) و Q_s برابر دبی وزنی مواد معلق (تن بر روز) است و a و b ضرایب ثابت‌اند. منحنی‌های سنجه رسوب بر اساس اندازه‌گیری‌های انجام‌شده در رودخانه‌ها به‌دست می‌آیند؛ بدین صورت که متناظر با هر دبی خاص در رودخانه غلظت مواد رسوبی حمل‌شده به وسیله آب اندازه‌گیری می‌شود و منحنی مربوط رسم می‌شود (دهقانی و وفاخواه، ۱۳۹۲: ۲۲۳).

روش فصلی

در این روش داده‌های متناظر دبی جریان و دبی رسوب معلق با توجه به تاریخ نمونه‌برداری به چهار سری بهار، تابستانه، پاییزه، و زمستانه تقسیم می‌شود. سپس، منحنی سنجه آن‌ها در نظر گرفته می‌شود (مساعدی و همکاران، ۱۳۸۸: ۳).

رگرسیون فرایند گاوسی

فرایند گاوسی روشی مفید به منظور تعریف توزیع‌های مقدم برای مدل‌های انعطاف پذیر رگرسیون و طبقه‌بندی است که در آن رگرسیون یا توابع احتمال طبقه به اشکال ساده پارامتری محدود نشده باشند. یکی از جذابیت‌های فرایند گاوسی تنوع بسیار توابع کورایانس آن است که به ایجاد توابعی با درجات مختلف همواری یا انواع مختلف ساختارهای پیوستی منجر می‌شود و به محقق این امکان را می‌دهد که از میان آن‌ها گزینش مناسبی انجام دهد. این مدل‌ها می‌توانند توزیع‌هایی را بین توابعی با تعداد یک یا چندین متغیر ورودی مشخص کنند. هنگامی که چنین تابعی میانگین پاسخ را در

یک مدل رگرسیونی با خطاهای گاوسی تعریف می‌کند، می‌توان از محاسبات ماتریس برای استنتاج بهره‌گیری کرد؛ این مسئله برای مجموعه داده‌هایی با بیش از هزار نمونه امکان‌پذیر است. فرایندهای گاوسی در مدل‌سازی آماری بسیار پُراهمیت‌اند؛ زیرا خصوصیات نرمالی دارند (نیل، ۱۹۹۷: ۱).

می‌توان n مشاهده را همیشه در مجموعه داده دلخواه $Y = \{Y_1, \dots, Y_n\}$ به صورت یک نقطه تک‌نمونه‌گیری شده از توزیع گاوسی چندمتغیری (n متغیری) تصور کرد. از این رو، مجموعه داده‌ها را می‌توان با یک فرایند گاوسی متناظر کرد. بنابراین، فرایندهای گاوسی همان قدر که ساده‌اند، فراگیر نیز هستند. اغلب فرض می‌شود که میانگین فرایند گاوسی متناظر در هر جا صفر است. آنچه در چنین حالت‌هایی یک مشاهده را به مشاهده دیگر ارتباط می‌دهد تابع کواریانس $k(x, x')$ است. هر مشاهده y را می‌توان از طریق مدل نویزی گاوسی به یک تابع اصلی ارتباط داد:

$$y = f(x) + N(0, \sigma_f^2) \quad \text{رابطه ۲}$$

که در آن $N(0, \sigma_f^2)$ به عنوان نویز تابع توزیع نرمال با میانگین صفر و واریانس σ_f^2 است. رگرسیون در واقع جست‌وجو برای $f(x)$ است. برای سادگی کار در مرحله بعدی با نوشتن عبارت زیر از روش جدید آمیختن نویز در $k(x, x')$ استفاده می‌شود:

$$k(x, x') = \sigma_f^2 \exp\left[-\frac{(x - x')^2}{2\tau^2}\right] + \sigma_n^2 \delta(x, x') \quad \text{رابطه ۳}$$

که در آن $\delta(x, x')$ تابع دلتای کرونکر است. بنابراین، n مشاهده Y در نظر گرفته شده است؛ هدف پیش‌بینی Y^* است نه f^* واقعی. مقادیر پیش‌بینی‌شده مشاهده‌ها طبق رابطه ۲ یکسان‌اند، اما واریانس‌ها به سبب فرایند نویز مشاهده‌ای با یکدیگر تفاوت دارند. برای آماده‌سازی رگرسیون فرایند گاوسی تابع کواریانس، رابطه ۳ در میان همه ترکیبات ممکن این نقاط محاسبه می‌شود و یافته‌ها در سه ماتریس به طور خلاصه بیان می‌شود:

$$k = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix} \quad \text{رابطه ۴}$$

$$K^* = [k(x_*, x_1) \ k(x_*, x_2) \ k(x_*, x_n)] \quad k_{ij} = k(x_i, x_j) \quad \text{رابطه ۵}$$

ذکر این نکته لازم است که عناصر قطری K به صورت $\sigma_f^2 + \sigma_n^2$ هستند. هرگاه x دامنه تقریباً بزرگی را اختیار کند آن‌گاه عناصر غیرقطری به صفر میل می‌کنند (ابدن، ۲۰۰۸: ۲).

رگرسیون بردار پشتیبان

ماشین‌های بردار پشتیبان شامل دو دسته‌اند: ۱. طبقه‌بندی‌کننده بردار پشتیبانی؛ ۲. رگرسیون بردار پشتیبان. ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که در سال ۱۹۹۵ واپنیک آن را بر پایه تئوری یادگیری آماری معرفی کرد (واپنیک، ۱۹۹۵). در برخی مواقع، به ساختارهای پیچیده و غیرخطی برای جداسازی داده‌ها نیاز است. در این صورت، ماشین بردار پشتیبان داده‌های اصلی را با به‌کارگیری مجموعه‌ای از توابع ریاضی- که کرنل نام دارند- در فضای جدیدی نگاشت و بازآرایی می‌کند؛ به این کار «تبدیل نگاشت» گفته می‌شود (شهرابی و ذوالقدر شجاعی، ۱۳۹۰). یک الگوریتم ماشین بردار پشتیبان به دنبال ابرصفحه‌ای با حداکثر حاشیه می‌گردد. از نقطه‌نظر هندسی حاشیه از فاصله موجود بین ابرصفحه و نزدیک‌ترین نمونه‌های آموزشی محاسبه می‌شود. کوتاه‌ترین فاصله از یک ابرصفحه تا نمونه‌هایی با برچسب +۱ برابر با کوتاه‌ترین فاصله از آن ابرصفحه تا نمونه‌هایی با برچسب -۱ است. در واقع، حاشیه از دو برابر این فاصله به دست می‌آید. یک ابرصفحه جداکننده را می‌توان به صورت زیر تعریف کرد:

$$W.X + b = 0 \quad \text{رابطه ۶}$$

که در آن برداری است که تعداد عضوهای موجود در آن برابر با صفات خاصه است و b مقدار ثابتی فرض می‌شود. در فضای دو بُعدی، که مجموعه داده‌ها با دو صفت خاصه و یک برچسب کلاس تعریف می‌شوند، با فرض $b = w_0$ رابطه ۶ به صورت رابطه ۷ بازنویسی می‌شود:

$$w_0 + w_1 x_1 + w_r x_r = 0 \quad \text{رابطه ۷}$$

بدین ترتیب، نمونه‌هایی (نقاطی) که در فضای بالای این ابرصفحه قرار می‌گیرند نامعادله ۸ و نمونه‌هایی که در فضای پایینی این ابرصفحه قرار می‌گیرند نامعادله ۹ را تکمیل می‌کنند.

$$w_0 + w_1 x_1 + w_r x_r > 0 \quad \text{رابطه ۸}$$

$$w_0 + w_1 x_1 + w_r x_r < 0 \quad \text{رابطه ۹}$$

با تنظیم مقدار W و b خواهیم داشت:

$$w_0 + w_1 x_1 + w_r x_r \geq 1 \quad \text{if } y_i = +1 \quad \text{رابطه ۱۰}$$

$$w_0 + w_1 x_1 + w_r x_r \leq -1 \quad \text{if } y_i = -1 \quad \text{رابطه ۱۱}$$

این بدین معنی است که هر نمونه‌ای بالا یا روی ابرصفحه H_1 قرار گیرد متعلق به کلاس +۱ است و هر نمونه‌ای که پایین و روی ابرصفحه H_2 قرار گیرد متعلق به کلاس -۱ خواهد بود. هر یک از نمونه‌هایی را که دقیقاً بر روی ابرصفحه H_1 و H_2 قرار دارند «بردار پشتیبان» می‌نامیم (اسماعیلی، ۱۳۹۳: ۲۵۰).

توابع کرنل

یکی از روش‌های متداول برای حل مسائل غیرخطی استفاده از توابع کرنل است؛ این توابع بر اساس ضرب داخلی داده‌های مفروض تعریف می‌شود. طراحی روش‌های رگرسیون مبتنی بر فرایند گاوسی نیز شامل استفاده از مفهوم تابع کرنل است. در واقع، با یک تبدیل غیرخطی از فضای ورودی به فضای خصیصه با ابعاد بیشتر (حتی نامتناهی) می‌توان

مسائل را به صورت خطی تفکیک‌پذیر ساخت. با تبدیل نمونه‌ها از فضای ورودی به فضای ویژگی، تفکیک‌کننده غیرخطی به حالت خطی تبدیل خواهد شد.

$$W^T \Phi(x_j) + b = 0 \quad \text{رابطه ۱۲}$$

مسئله اولیه و دوگان مسئله مربوط به فضای خصیصه را تشکیل می‌دهد، با این تفاوت که به جای عبارت $(x_i^T \cdot x_j)$ از $\Phi^T(x_i) \cdot \Phi(x_j)$ استفاده شده است که $k(x_i, x_j)$ تابع کرنل به منظور خطی‌سازی مسائل غیرخطی است. از مهم‌ترین کرنل‌ها می‌توان به کرنل خطی، چندجمله‌ای، چندجمله‌ای نرمال‌شده، کرنل تابع شعاع‌محور، و تابع کرنل پیرسون اشاره کرد.

$$k(x_i, x_j) = (x_i^T x_j + 1)^p \quad \text{رابطه ۱۳} \quad \text{تابع کرنل چندجمله‌ای ساده}$$

$$k(x_i, x_j) = \frac{(x_i^T x_j + 1)^p}{\sqrt{(x_i^T x_i)(x_j^T x_j)}} \quad \text{رابطه ۱۴} \quad \text{تابع کرنل چندجمله‌ای نرمال‌شده}$$

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad \text{رابطه ۱۵} \quad \text{تابع پایه شعاعی}$$

$$k(x_i, x_j) = \frac{1}{\left[1 + \left(\sqrt{\frac{2}{\sigma}} \|x_i - x_j\|\right)^w\right]^w} \quad \text{رابطه ۱۶} \quad \text{تابع کرنل پیرسون}$$

هر یک از این توابع کرنل در ساختار خود دارای پارامترهایی هستند که با نام هایپر پارامتر^۱ شناخته می‌شوند. مثلاً، تابع پایه شعاعی دارای هایپر پارامتر گاما و تابع پیرسون دارای هایپر پارامترهای سیگما و امگا است. در کاربرد روش‌های مبتنی بر تابع کرنل، علاوه بر انتخاب تابع کرنل بهینه، تعیین مقدار بهینه هایپر پارامترهای مربوط به هر تابع بسیار حائز اهمیت است. روش‌های مدل‌سازی مبتنی بر کرنل به ایجاد پارامترهای مناسب تعریف‌شده توسط کاربر نیاز دارد، چون دقت این روش رگرسیونی وابستگی زیادی به انتخاب این پارامترها دارد. علاوه بر انتخاب پارامترهای خاص کرنل، رگرسیون فرایند گاوسی به مشخص کردن مقادیر بهینه سطح نویز گاوسی نیاز دارد. برای انتخاب پارامترهای انتخاب‌شده توسط کاربر (یعنی $C, \gamma, \sigma, \omega, \varepsilon$ و نویز گاوسی) روش‌های متعددی ارائه شده است، مثل روش دستی (سعی و خطا)، روش جست‌وجوی شبکه، الگوریتم ژنتیک، و بهینه‌سازی اجتماع ذره. در مطالعه حاضر، برای انتخاب پارامترهای انتخاب‌شده توسط کاربر از روش سعی و خطا استفاده شد. مقادیر بهینه پارامترهای مختلف تعریف‌شده توسط کاربر طوری انتخاب شده است که ریشه میانگین مربع خطا را کمینه و ضریب همبستگی را بیشینه کند (پال و دسوال، ۲۰۱۰: ۹۴۴).

معیارهای ارزیابی مدل‌ها

عملکرد مدل‌های منحنی سنجه رسوب، روش فصلی، رگرسیون فرایند گاوسی، و رگرسیون ماشین بردار پشتیبان در این

تحقیق بر پایه محاسبه ضریب همبستگی^۱، نش-ساتکلیف^۲، جذر میانگین مربعات خطا^۳، و میانگین خطای مطلق^۴ ارزیابی شد. فرمول‌های محاسبه آماره‌های فوق در رابطه‌های ۱۷ تا ۲۰ ارائه شده است. در این رابطه‌ها مقادیر X شامل مقادیر مشاهداتی و مقادیر Y شامل مقادیر محاسباتی و N تعداد داده‌هاست.

$$r_{\text{pearson}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{رابطه ۱۸}$$

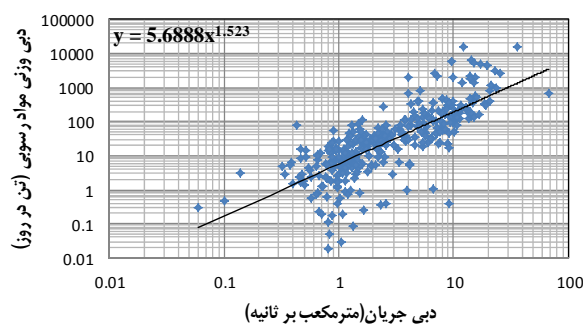
$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{N}} \quad \text{رابطه ۱۷}$$

$$E = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad \text{رابطه ۲۰}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \quad \text{رابطه ۱۹}$$

یافته‌های پژوهش

پیش از مدل‌سازی میزان رسوب معلق رودخانه صوفی‌چای به وسیله روش‌های مختلف، تحلیل‌های آماری روی داده‌های موجود از قبیل آزمون‌های همگنی نرمال استاندارد، دامنه بیش‌ند، پتیت، نسبت فون نیومن، و پالایش داده‌ها انجام شد، زیرا استفاده از داده‌های صحیح در آغاز کار نقش بسزایی در میزان درجه اطمینان نتایج مدل‌ها و تحقیق دارد. برای مقایسه نتایج حاصل از روش‌های رگرسیون فرایند گاوسی و رگرسیون ماشین بردار پشتیبان با سایر روش‌ها، نخست منحنی سنجه رسوب رسم شد. شکل ۴ منحنی سنجه رسوب به دست آمده از داده‌های موجود را نشان می‌دهد.



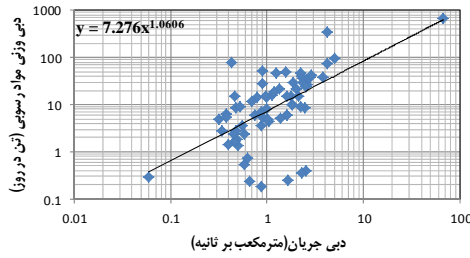
شکل ۴. منحنی سنجه رسوب به دست آمده مربوط به رودخانه صوفی‌چای

معادله به دست آمده از منحنی سنجه رسوب به صورت رابطه ۲۱ است. برای ساخت مدل به کمک روش منحنی سنجه رسوب از داده‌های مربوط به بخش آزمون مدل‌ها استفاده شد و در رابطه ۲۱ جای‌گذاری شد تا مقادیر رسوب

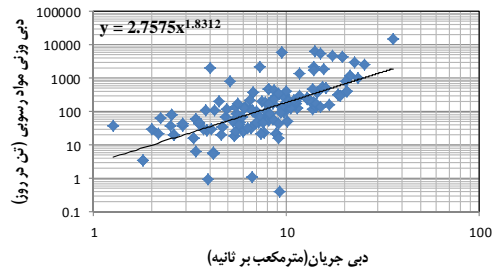
1. correlation coefficient
2. nash-sutcliffe coefficient (E)
3. root mean square error
4. mean absolute error

برآورد شود. پس از برآورد، مقادیر پیش‌بینی شده با مقادیر مشاهده‌ای به وسیله شاخص‌های آماری مطرح‌شده بررسی شد. شکل ۵ روابط خطی و نمودارهای به‌دست‌آمده از روش فصلی را نشان می‌دهد.

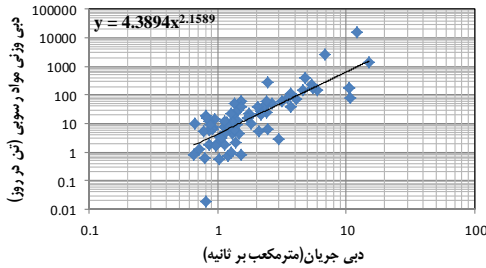
$$Q_s = ۵,۶۸۸۸ Q_w^{۱.۵۲۳} \quad \text{رابطه ۲۱}$$



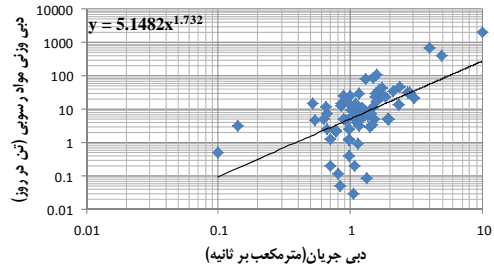
(ب)



(الف)



(د)



(ج)

شکل ۵. روابط خطی و نمودارهای به‌دست‌آمده از روش فصلی: (الف) بهار؛ (ب) تابستان؛ (ج) پاییز؛ (د) زمستان

هنگام مدل‌سازی و برآورد میزان دبی رسوب معلق به وسیله روش‌های رگرسیون فرایند گاوسی و ماشین بردار پشتیبان بعد از سعی و خطاهای متعدد و بررسی هر چهار تابع کرنل معرفی‌شده و هایپر پارامترهای مربوط، مشخص شد در هر دو روش نتیجه بهینه هنگامی حاصل می‌شود که از تابع پایه شعاعی به عنوان تابع کرنل استفاده شود. نتیجه تحقیق حاضر با مطالعات اسکندری و همکاران (۱۳۹۱) - برای تخمین کیفیت آب با استفاده از ماشین بردار پشتیبان و استفاده از تابع کرنل پایه شعاعی - همخوانی دارد. جدول ۲ مقدار بهینه هایپر پارامترهای مربوط را نشان می‌دهد. جدول ۳ نتایج آماری مربوط به چهار روش بررسی‌شده در این تحقیق را نشان می‌دهد.

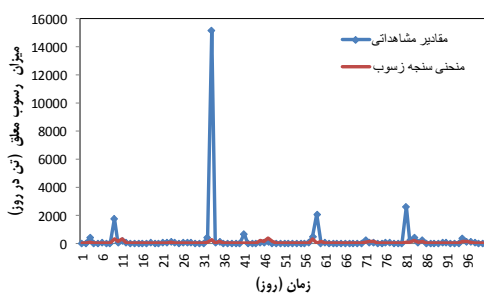
جدول ۲. مقدار بهینه هایپر پارامترهای مربوط به روش‌های داده‌کاوی مبتنی بر کرنل

روش داده‌کاوی مبتنی بر کرنل	تابع کرنل	گاما (γ)	نویز (ε)	C
رگرسیون فرایند گاوسی	پایه شعاعی	۰/۵	۰/۰۱	-
رگرسیون بردار پشتیبان	پایه شعاعی	۰/۰۱	-	۱

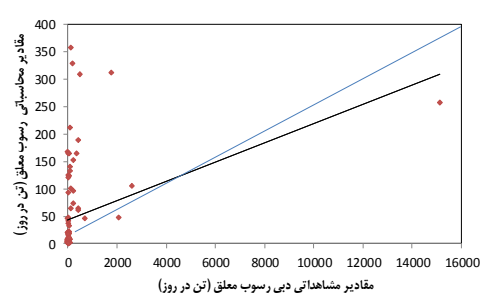
جدول ۳. نتایج آماری به‌دست‌آمده از روش‌های گوناگون در برآورد دبی رسوب معلق رودخانه صوفی‌چای مراغه

آماره‌های ارزیابی	ضریب همبستگی	ضریب نش	ریشه میانگین مربعات خطا (تن در روز)	میانگین خطای مطلق (تن در روز)
روش استفاده‌شده				
منحنی سنجه رسوب	۰/۳۳۹	۰/۰۱۱	۱۵۳۳/۲۸۹۹	۲۵۳/۹۶۶۷
روش فصلی	۰/۶۹۵	۰/۱۰۱	۱۴۶۲/۰۵۳۱	۲۴۷/۹۵۵۲
رگرسیون فرایند گاوسی	۰/۹۷۷	۰/۷۹۴	۶۹۸/۷۴۵۵	۷۷/۴۲۷۸
رگرسیون بردار پشتیبان	۰/۹۴۹	۰/۵۱۷	۱۰۷۱/۰۴۹۳	۱۶۷/۱۴۸۲

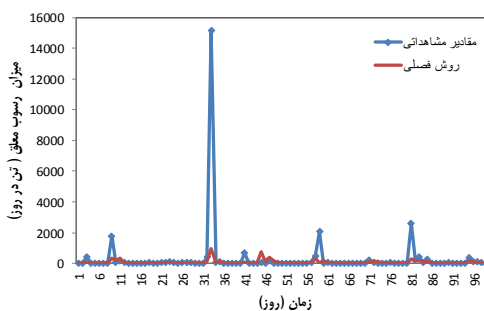
همچنان که در جدول ۳ مشاهده می‌شود، مدل به‌دست‌آمده از روش رگرسیون فرایند گاوسی با ارائه مقادیر بالاتر ضریب همبستگی و نش و مقادیر کمتر آماره‌های مربوط به خطا دارای دقت بیشتری در زمینه برآورد میزان دبی رسوب معلق رودخانه صوفی‌چای است. همچنین، نتیجه دیگری که قابل استنباط است این است که هر دو روش داده‌کاوی بررسی‌شده به مراتب نتایج بهتری نسبت به منحنی سنجه رسوب و روش فصلی ارائه کرده‌اند. برای درک بهتر نتایج، نمودارهای پراکنش و سری زمانی داده‌های محاسباتی به‌دست‌آمده از روش‌های مختلف نسبت به داده‌های مشاهداتی موجود در شکل‌های ۶ تا ۱۳ نشان داده شده است.



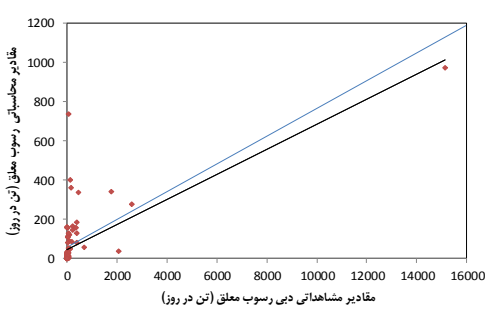
شکل ۵. سری زمانی رسوب به‌دست‌آمده از منحنی سنجه رسوب



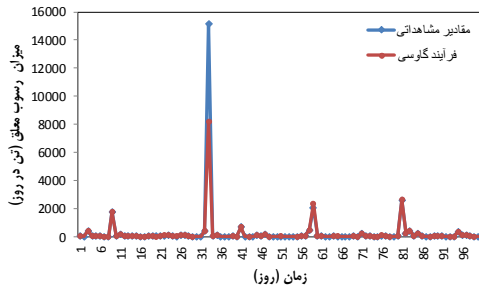
شکل ۶. مقادیر محاسباتی و مشاهداتی رسوب معلق به‌دست‌آمده از منحنی سنجه رسوب



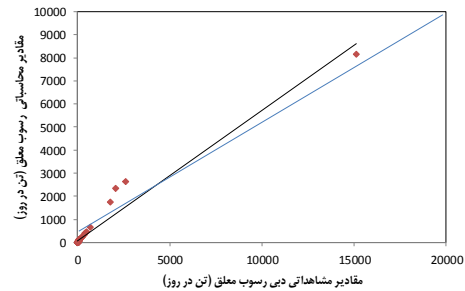
شکل ۷. سری زمانی رسوب معلق به‌دست‌آمده از روش فصلی



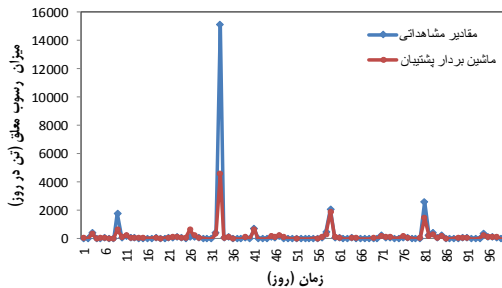
شکل ۸. مقادیر محاسباتی و مشاهداتی دبی رسوب معلق به‌دست‌آمده از روش فصلی



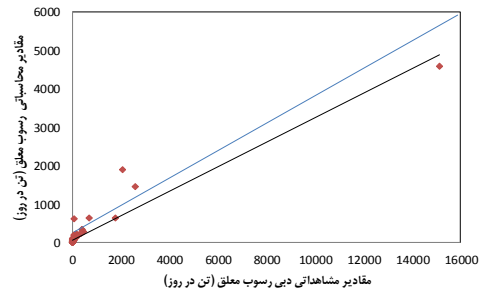
شکل ۱۱. سری زمانی رسوب معلق به‌دست‌آمده از روش رگرسیون فرایند گاوسی



شکل ۱۰. مقادیر محاسباتی و مشاهده‌ای رسوب معلق به‌دست‌آمده از روش رگرسیون فرایند گاوسی



شکل ۱۳. سری زمانی رسوب معلق به‌دست‌آمده از رگرسیون بردار پشتیبان



شکل ۱۲. مقادیر محاسباتی و مشاهده‌ای رسوب معلق به‌دست‌آمده از رگرسیون بردار پشتیبان

همچنان که در شکل‌های ۶ تا ۱۳ مشاهده می‌شود، بیشترین تطابق بین مقادیر محاسباتی و مشاهده‌ای مربوط به نتایج به‌دست‌آمده از روش رگرسیون فرایند گاوسی است و نشان‌دهنده دقت بالای این مدل در زمینه برآورد رسوب رودخانه صوفی‌چای و همخوانی بیشتر مقادیر پیش‌بینی‌شده با مقادیر واقعی میزان رسوب معلق است. هر دو روش رگرسیون فرایند گاوسی و رگرسیون بردار پشتیبان عملکرد بهتری نسبت به روش منحنی سنجه رسوب و روش فصلی داشته‌اند، ولیکن با توجه به بررسی مقادیر ضریب همبستگی و نش-سانکلیف مربوط به این روش‌ها و همچنین بررسی مقادیر خطای ریشه میانگین مربعات خطا و میانگین خطای مطلق مشاهده می‌شود که روش رگرسیون فرایند گاوسی برتری اندکی نسبت به روش رگرسیون بردار پشتیبان دارد؛ این برتری اندک در نمودارهای پراکنش و سری زمانی به‌دست‌آمده از هر دو روش نیز مشهود است. از بین دو روش تجربی مرسوم بررسی‌شده روش فصلی، با توجه به اینکه تغییرات فصول را در نظر می‌گیرد، دقت بیشتری دارد. همچنین، نتیجه دیگری که از بررسی دقیق‌تر این نمودارها قابل استنباط است این است که روش منحنی سنجه رسوب و روش فصلی در تخمین میزان رسوب رودخانه صوفی‌چای مراغه در مقادیر بالای دبی و بار معلق دقت بسیار اندکی دارند و استفاده از این دو روش برای محاسبه رسوب معلق مناسب نیست. این مسئله می‌تواند نشئت‌گرفته از اریب موجود در این روش‌ها باشد. اریب در واقع باعث می‌شود تا مقدار باقی‌مانده‌ها (اختلاف بین مقادیر مشاهده‌ای و محاسباتی) توزیع نرمالی نداشته باشند. اما روش رگرسیون فرایند گاوسی و ماشین بردار پشتیبان توانسته‌اند انطباق نسبتاً خوبی بین مقادیر محاسباتی و مشاهده‌ای میزان رسوب معلق رودخانه

صوفی‌چای مراغه ایجاد کنند؛ این موضوع نشان‌دهنده توانایی نسبتاً خوب این روش در حل مسائل غیرخطی است؛ این مسئله همچنین به سبب استفاده این دو روش از توابع کرنل است. از آنجا که میزان رسوب کل در یک بازه زمانی مشخص در مطالعات منابع آب و طراحی سدها اهمیت زیادی دارد، شکل ۱۴ نحوه عملکرد روش‌های نوین داده‌کاوی و مرسوم کلاسیک در تخمین میزان رسوب کل در بازه مورد مطالعه را نشان می‌دهد.



شکل ۱۴. مقادیر واقعی و محاسبه‌شده رسوب کل

همچنان‌که مشاهده می‌شود، روش‌های سنتی تفاوت چشم‌گیری با یکدیگر نداشتند و در حالت برآورد میزان رسوب کل عملکرد نسبتاً مشابهی داشتند و میزان رسوب کل را با تفاوت بسیار زیادی کمتر از میزان واقعی آن برآورد کردند. این در حالی است که هر دو روش داده‌کاوی مبتنی بر توابع کرنل عملکرد بهتر و دقیق‌تری در زمینه تخمین میزان رسوب سالانه از خود نشان دادند. روش رگرسیون فرایند گاوسی به سبب انعطاف‌پذیری بیشتر مقادیر رسوب را با دقت بیشتر و خطای کمتر تخمین زده است. از آنجا که پیش‌بینی کمتر از حد معمول میزان رسوب کل مشکلات عدیده‌ای در بهره‌برداری و استفاده بهینه از منابع آبی موجود ایجاد می‌کند، استفاده از روش‌های داده‌کاوی مبتنی بر توابع کرنل، که توانایی زیادی در این زمینه از خود نشان داده‌اند، برای تخمین میزان رسوب سالانه توصیه می‌شود.

بحث و نتیجه‌گیری

در این تحقیق، پس از جمع‌آوری داده‌های مورد نیاز مربوط به رودخانه صوفی‌چای مراغه، این داده‌ها به وسیله آزمون‌های همگنی شامل نرمال استاندارد، دامنه بیش‌ند، پتیت، و نسبت فون نیومن بررسی شد. سپس، میزان دبی رسوبی معلق رودخانه صوفی‌چای با استفاده از روش‌های رگرسیون فرایند گاوسی، رگرسیون بردار پشتیبان، منحنی سنجه رسوب، و روش فصلی برآورد شد. برای دستیابی به نتایج بهینه به وسیله روش‌های داده‌کاوی استفاده‌شده سناریوهای متعددی شامل انواع توابع کرنل و بازه‌های متفاوتی از هاپیر پارامترهای توابع کرنل تعریف شد. هنگامی که برای برآورد رسوب معلق رودخانه صوفی‌چای از روش رگرسیون فرایند گاوسی با تابع کرنل پایه شعاعی دارای مقادیر (نویز گاوسی) (E) برابر

۰/۰۱ و گاما (۱) برابر ۰/۵ استفاده شد، مشاهده گردید که، با توجه به معیارهای ارزیابی، این روش دارای بیشترین دقت و کمترین خطا از میان روش‌های بررسی‌شده در این مطالعه است. در تحقیق مشابهی خزایی-پول و طالبی (۱۳۹۲) رسوبات معلق رودخانه قطورچای را با استفاده از شبکه عصبی مصنوعی و منحنی سنجه رسوب پیش‌بینی کردند و هنگام مدل‌سازی رسوبات معلق این رودخانه به وسیله شبکه عصبی مصنوعی ضریب همبستگی را برابر با ۰/۸۷ و مقدار ریشه میانگین مربعات خطا را برابر ۶۲۲ (تن در روز) به‌دست آوردند. در صورتی که در تحقیق حاضر هر دو روش داده‌کاوی نوین بررسی‌شده نتایج دقیق‌تری ارائه کردند؛ این موضوع نشان‌دهنده کارایی و عملکرد دقیق این روش‌هاست. همچنین، در تحقیق دیگری ولی و همکاران (۱۳۹۰) بار رسوب جریان را در حوضه آبریز سمندگان با استفاده از شبکه عصبی مصنوعی پیش‌بینی کردند؛ در بهترین حالت ضریب همبستگی برابر با ۰/۹۳ و ریشه میانگین مربعات خطا برابر با ۱/۸۱۵ (تن در روز) به‌دست آمد؛ در حالی که نتایج تحقیق حاضر نشان‌دهنده برتری نسبی روش فرایند گاوسی با ارائه ضریب همبستگی ۰/۹۷۷ و ریشه میانگین مربعات خطا برابر ۶۹۸/۷۴۶ (تن در روز) است. اختلاف در میزان ریشه میانگین مربعات خطا می‌تواند به تفاوت در نوع و بازه داده‌های مورد استفاده در تحقیق حاضر مربوط باشد. بنابراین، مشخص شد که اگرچه نمی‌توان نتایج به‌دست‌آمده را برای سایر حوضه‌ها تعمیم داد، در این تحقیق هر دو روش داده‌کاوی بررسی‌شده به‌مراتب کارایی و دقت عمل نسبتاً بیشتری در زمینه تخمین بار معلق رودخانه صوفی‌چای داشتند و می‌توانند به عنوان روش‌های جای‌گزین برای تخمین بار رسوبی معلق به‌کار گرفته شوند. ذکر این نکته ضروری است که، با توجه به تغییر مستمر کاربری اراضی و توسعه سطح زیر کشت، شرایط فیزیکی و ساختار مورفولوژیکی حوضه نیز دائماً در حال تغییر است و نمی‌توان نتایج به‌دست‌آمده را برای همه شرایط زمانی بسط داد.

منابع

- اسکندری، ع.؛ نوری، ر.؛ معراجی، ح. و کیاقدری، ا. (۱۳۹۱). توسعه مدلی مناسب بر مبنای شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای پیش‌بینی بهنگام اکسیژن‌خواهی بیوشیمیایی ۵روزه، *مجله محیط‌شناسی*، ۱۶(۱): ۷۱ - ۸۲.
- امامی، س.ا. (۱۳۷۹). *انتقال رسوب*، تهران: انتشارات جهاد دانشگاهی صنعتی امیرکبیر.
- خزایی‌پول، ا. و طالبی، ع. (۱۳۹۲). بررسی امکان پیش‌بینی رسوبات معلق با استفاده از ترکیب منحنی سنجه رسوب و شبکه عصبی مصنوعی (مطالعه موردی: رودخانه قطورچای، پل یزدکان)، *مجله پژوهش‌های فرسایش محیطی*، ۹: ۷۳ - ۸۲.
- دهقانی، ا.؛ زنگانه، م.؛ مساعدی، ا. و کوهستانی، ن. (۱۳۸۸). مقایسه تخمین بار معلق به دو روش منحنی سنجه رسوب و شبکه عصبی مصنوعی، *مجله علوم کشاورزی و منابع طبیعی*، ۱۶: ۳۶ - ۵۱.
- دهقانی، ن. و وفاخواه، م. (۱۳۹۲). مقایسه روش‌های تخمین رسوب معلق روزانه با استفاده از روش‌های منحنی سنجه رسوب و شبکه عصبی (مطالعه موردی: ایستگاه قزاقلی، استان گلستان)، *مجله پژوهش‌های حفاظت آب و خاک*، ۲۰(۲): ۲۲۱ - ۲۳۰.
- رجبی، م.؛ فیض‌الله‌پور، م. و روستایی، ش. (۱۳۹۴). استفاده از مدل تبرید تدریجی عصبی (NDE) در تخمین بار معلق رسوبی و مقایسه آن با مدل ANFIS و RBF (مطالعه موردی: رودخانه گیوی‌چای)، *مجله توسعه و جغرافیا*، ۳۹: ۱ - ۱۶.
- رضازاده جودی، ع. و ستاری، م. (۱۳۹۴). تخمین عمق چاله آبستنگی پایه پل در سازه‌های رودخانه‌ای با روش رگرسیون فرایند گاوسی، *نشریه تحقیقات کاربردی مهندسی سازه‌های آبیاری و زهکشی*، ۱۶(۶۵): ۱۹ - ۳۶.

- شهرابی، ج. و ذوالقدر شجاعی، ع. (۱۳۹۰). *داده‌کاوی پیشرفته (مفاهیم و الگوریتم‌ها)*. انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر.
- طباطبایی، م.; سلیمانی، ک.; حبیب‌الله روشن، م. و کاویان، ع. (۱۳۹۳). برآورد غلظت رسوب معلق روزانه با استفاده از شبکه عصبی مصنوعی و خوشه‌بندی داده‌ها به روش نگاشت خودسازمانده (مطالعه موردی: ایستگاه هیدرومتری سیرا - رودخانه کرج)، *پژوهش‌نامه مدیریت حوضه آبریز*، ۱۰: ۹۸ - ۱۱۶.
- فلامکی، ا.; اسکندری، م.; بغلانی، ع. و احمدی، س.ا. (۱۳۹۲). مدل سازی بار رسوب کل رودخانه‌ها با استفاده از شبکه‌های عصبی مصنوعی، *نشریه حفاظت منابع آب و خاک*، ۲(۳): ۱۳ - ۲۵.
- مساعدی، ا.; زنگانه، م.ا.; مفتاح، م.; دهقانی، ا.ا. و خوشروش، م. (۱۳۸۸). ارزیابی روش‌های هیدرولوژیکی برآورد بار معلق (مطالعه موردی: رودخانه اترک استان گلستان)، *دهمین سمینار سراسری آبیاری و کاهش تبخیر ۱۹ تا ۲۱ بهمن، کرمان*.
- نجمایی، م. (۱۳۶۹). *هیدرولوژی مهندسی*، ج ۲، دانشگاه علم و صنعت ایران.
- ولی، ع.; معیری، م.; رامشت، م.ح. و موحدی‌نیا، ن. (۱۳۹۰). تحلیل مقایسه عملکرد شبکه‌های عصبی مصنوعی و مدل‌های رگرسیون پیش‌بینی رسوب معلق (مطالعه موردی: حوضه آبخیز اسکندری واقع در حوضه آبریز زاینده‌رود)، *پژوهش‌های جغرافیای طبیعی*، ۷۱(۱): ۲۱ - ۳۰.
- Alp, M. and Cigizoglu, H.K. (2007). **Suspended sediment load simulation by two artificial neural network methods using hydro meteorological data**, *Environmental Modelling and Software*, 22: 2-13.
- Dehgani, A.A.; Zangane, M.A.; Mosaedi, A. and Kuhestani N. (2010). **Estimation of suspended sediment in the sediment rating curve and artificial neural network**, *Journal of Agricultural Sciences and Natural Resources*, 16: 36-51 (In Persian).
- Dehgani, N. and Vafakhah, M. (2014). **Comparison of methods for prediction of daily sediment using sediment rating curves and neural network (Case Study: Gezagli Station, Golestan Province)**, *Journal of Soil and Water Conservation*, 20(2): 221-230 (In Persian).
- Duan, W.L.; He, B.; Takara, K.; Luo, P.P.; Nover, D. and Hu, M.C. (2015). **Modeling suspended sediment sources and transport in the Ishikari River basin, Japan, using SPARROW**, *Hydraulic Earth Systems Sciences*, 19: 1293-1306.
- Ebden, M. (2008). **Gaussian processes for regression: a quick Introduction**, Available from: <http://www.robots.ox.ac.uk/~mebden/reports/GPtutorial.pdf> [Accessed 14 August 2015]
- Eder, A.P.; Strauss, T.; Krueger, B.; Iand, J.N. and Quinton, B. (2010). **A Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkirchen catchment), Austria**, *Journal of Hydrology*, 389: 168-176.
- Emami, S.A. (2000). *Sediment transportation*, Jahade daneshgahi Press, Amirkabir industrial university, Tehran, First edition, 716p. (In Persian).
- Eskandari, A.; Nouri, R.; Meraji, H. and Kiagaderi, A. (2012). **The development of an appropriate model based on artificial neural network and support vector machine for predicting biochemical oxygen during 5 days**, *Journal of Ecology*, 61(1): 71-82 (In Persian).
- Falamaki, A.; Eskandari, M.; Baghlani, A. and Ahmadi, S.A. (2013). **Modeling total sediment load in rivers using artificial neural networks**, *Journal of water and soil conservation*, 2(3): 13-25 (In Persian).
- Heng, S. and Suetsugi, T. (2013). **Using artificial neural network to estimate sediment load in ungauged catchments of the Tonle Sap River Basin, Cambodia**, *Journal of Water Resource and Protection*, 5: 111-123.
- Kakaei Lafdani, E.; Moghaddam Nia, A. and Ahmadi, A. (2013). **Daily suspended sediment load prediction using artificial neural networks and support vector machines**, *Hydrology*, 478: 50-62.
- Kao, Sh.; Lee, T. and Milliman, J.D. (2005). **Calculating highly fluctuated suspended sediment fluxes from mountainous rivers in Taiwan**, *TAO*, 16(3): 653-675.
- 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*, 2(9): 73-82 (In Persian).

- Kia, E.; Emadi, A.R. and Fazlola, R. (2013). **Investigation and Evaluation of Artificial Neural Networks in Babolroud River Suspended Load Estimation**, *Journal of Civil Engineering and Urbanism*, 3(4): 183-190.
- Kumar Goyal, M. (2014). **Modeling of Sediment Yield Prediction Using M5 Model Tree Algorithm and Wavelet Regression**, *Journal of Water Resources Management*, 28: 1991-2003.
- Mosaedi, A.; Zangane, M.A.; Meftah, M.; Dehgani, A.A. and Khoshravesh, M. (2010). **Evaluation of hydrological methods to estimate the suspended load (case study: Atrak River of Golestan Province)**, *10th Seminar irrigation and evaporation*, 8-10 February, Kerman (In Persian).
- Najmaei, M. (1990). *Engineering Hydrology*, Second edition, Iran University of Science and Technology (In Persian).
- Neal, R.M. (1997). *Monte carlo implementation of gaussian process models for bayesian regression and classification*, University of Toronto, Toronto: Department of Statistics and Department of Computer Science, Technical report no, 9702.
- Onderka, M.; Krein, A. and Wrede, S. (2012). **Dynamics of storm-driven suspended sediments in headwater catchment described by multivariable modeling**, *Journal of Soils Sediments*, 12: 620-635.
- Pal, M. and Deswal, S. (2010). **Modelling pile capacity using Gaussian process regression**, *Computers and Geotechnics*, 37, 942-947.
- Rajabi, M.; Feizollahpour, M. and Roustaei, S. (2015). **Using NDE model for estimation of suspended sediment load in comparison with ANFIS and RBF case study: Givi Chay**, *Geography and Development Iranian Journal*, 39(2): 1-16 (In Persian).
- Rajaei, T.; Mirbagheri, S.A.; Nourani, V. and Alikhani, A. (2010). **Prediction of daily suspended sediment load using wavelet and neuro fuzzy combined model**, *International Journal of Environment Sciences, Tech.*, 7(1): 93-110.
- Rezazadeh Joudi, A. and Sattari, M. (2016). **Estimation of Scour Depth of Piers in Hydraulic Structures using Gaussian Process Regression**, *Applied Research in Irrigation and Drainage Structures Engineering*, 16(65): 19-36 (In Persian).
- Sadeghi, S.H.R.; Mizuyama, T.; Miyata, S.; Gomi, T.; Kosugi, K.; Fukushima, T.; Mizugaki, S. and Onda, Y. (2008). **Development, evaluation and interpretation of sediment rating curves for a Japanese small mountainous reforested watershed**, *Geoderma*, 144: 198-211.
- Shahrabi, J. and Hejazi, T.H. (2011). *Data mining*, Tehran, Industrial University of Amirkabir, Jahad daneshgahi Press (In Persian).
- Tabatabaei, M.; Solaimani, K.; Habibnejad Roshan, M. and Kavian, A. (2014). **Estimation of Daily Suspended Sediment Concentration using Artificial Neural Networks and Data Clustering by Self-Organizing Map (Case Study: Sierra Hydrometry Station- Karaj Dam Watershed)**, *Journal of Watershed Management Research*, 5(10): 98-116 (In Persian).
- Vali, A.; Moayeri, M.; Ramsht, M.H. and Movahedinia, N. (2010). **Analysis and Comparison of artificial neural networks and regression models in suspended sediment Prediction case study: Eskandari Catchment Area located in Zayanderood Basin**, *Journal of Physical Geography Research Quarterly*, 71(1): 21-30 (In Persian).
- Vapnik, V.N. (1995). *The nature of statistical learning theory*, Newyork: springer-verlag.