

مقاله علمی-پژوهشی

ارزیابی عملکرد مدل‌های یادگیری ماشین مبتنی بر روش تجزیه مد متغیر در برآورد جریان رودخانه (مطالعه موردی: رودخانه دز)

نازنین عالی وند^۱، علی محمد آخوند علی^۲، فرشاد احمدی^{۳*}

تاریخ دریافت: ۱۴۰۴/۰۱/۱۱ تاریخ پذیرش: ۱۴۰۴/۰۲/۲۷

چکیده

در مطالعه حاضر جریان ماهانه رودخانه دز در محل ایستگاه تله زنگ در دوره آماری ۱۳۵۰ تا ۱۴۰۱، با رویکرد توسعه مدل تلفیقی مبتنی بر روش تجزیه سیگنال مد متغیر (VMD) مدل سازی شد. داده‌های ورودی به مدل‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) براساس سه سناریو تعرف شد. در سناریوی اول ورودی‌ها به ترتیب از یک تا چهار تأخیر در مقیاس ماهانه در نظر گرفته شدند. در سناریوی دوم علاوه بر داده‌های جریان تأخیر یافته ترم تناوبی خطی و غیرخطی به مدل‌ها وارد شد و در سناریوی سوم داده‌های ورودی توسط روش VMD به زیر سری‌هایی تحت عنوان توابع مد ذاتی تجزیه شده و سپس در اختیار مدل‌ها قرار گرفت. نتایج نشان داد که هرچند مدل‌های منفرد با الگوهای ورودی متفاوتی به حداکثر دقت خود دست می‌یابند اما افزودن ترم تناوبی عملکرد مدل‌ها را تا حدودی بهبود می‌بخشد. در این بین بهترین عملکرد، مربوط به مدل SVM با میانگین شاخص خطای RMSE برابر با $156/83 \text{ (m}^3/\text{s)}$ در سناریوی دوم بود. برای مدل RF نیز میانگین شاخص خطا در سناریوی دوم کمتر از سناریوی اول بوده و برابر با $162/99 \text{ (m}^3/\text{s)}$ به دست آمد. در گام سوم داده‌های مربوطه توسط روش VMD تجزیه و فرآیند مدل‌سازی با روش‌های RF و SVM انجام شد. براساس شاخص‌های ارزیابی، کاهش خطا و افزایش دقت در مدل‌های تلفیقی توسعه داده شده به‌طور قابل ملاحظه‌ای مشهود بود. به طوری که مدل VMD-SVM توانست مقدار شاخص RMSE را به‌طور متوسط حدود $121 \text{ (m}^3/\text{s)}$ (حدود ۷۱ درصد) کاهش داده و نسبت به مدل VMD-RF نیز از دقت بهتری برخوردار بود.

واژه‌های کلیدی: تجزیه سیگنال، توابع مد ذاتی، موج فرانکس مرکزی، مدل جنگل تصادفی، ماشین بردار پشتیبان

مقدمه

پیش‌بینی جریان رودخانه یکی از چالش‌های اساسی در مدیریت منابع آب و برنامه‌ریزی‌های زیست‌محیطی است. تغییرات آب و هوایی، فعالیت‌های انسانی و نوسانات طبیعی می‌توانند به شدت بر جریان‌های آبی تأثیر بگذارند. پیش‌بینی دقیق جریان رودخانه‌ها می‌تواند به بهبود مدیریت سیلاب، تأمین آب شرب و کشاورزی و حفاظت از اکوسیستم‌های آبی کمک کند. در سال‌های اخیر، استفاده

از مدل‌های یادگیری ماشین برای پیش‌بینی جریان رودخانه‌ها مورد توجه قرار گرفته است. از جمله این مدل‌ها می‌توان به ماشین بردار پشتیبان^۴ (SVM) و جنگل تصادفی^۵ (RF) اشاره کرد (Essam et al., 2022; Samantaray et al., 2022; Ahmadi et al., 2022). مدل ماشین بردار پشتیبان به‌عنوان یک روش قدرتمند در یادگیری ماشین و داده‌کاوی، به‌ویژه در مسائل طبقه‌بندی و رگرسیون شناخته می‌شود. این مدل با استفاده از تکنیک‌های ریاضی پیچیده، داده‌ها را در فضای چندبعدی تجزیه و تحلیل کرده و به دنبال بهترین مرز تفکیک بین کلاس‌ها یا پیش‌بینی مقادیر خروجی می‌باشد (Vapnik., 1998). از این روش در مطالعات متعددی برای پیش‌بینی جریان رودخانه استفاده شده است. سمتری و ساهو به‌منظور برآورد دبی جریان از مدل هیبریدی PSR-SVM-FFA که ترکیبی از بازسازی فضای فازی، مدل ماشین بردار پشتیبان و الگوریتم کرم شب

۱- دانشجوی کارشناسی ارشد، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران
۲- استاد، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران
۳- دانشیار، گروه هیدرولوژی و منابع آب، دانشکده مهندسی آب و محیط زیست، دانشگاه شهید چمران اهواز، اهواز، ایران
* - نویسنده مسئول: (Email: F.ahmadi@scu.ac.ir)

4 - Support Vector Machine (SVM)
5 - Random Forest (RF)

از دیرباز، فعالیت‌های انسانی همواره چالش بزرگی برای مسئله پیش‌بینی جریان رودخانه بوده که اختلال در عملکرد مدل‌ها را به همراه داشته است. این فعالیت‌ها به‌همراه اثرات تغییر اقلیم، موجب می‌شود که داده‌های ثبت شده جریان رودخانه متغیر و ناپایدار شود (Meng et al., 2021). از این رو به‌منظور برطرف کردن این مشکل، مدل‌های یادگیری ماشین بر اساس روش‌های تجزیه پایه توسعه داده شده‌اند که می‌توانند سری‌های جریان رودخانه را به زیرسری‌های متعددی تجزیه نمایند. شایان‌ذکر است که وارد کردن زیر سری‌ها به مدل‌ها موجب عملکرد مطلوب‌تر آن‌ها می‌شود (Fang et al., 2019). یکی از روش‌هایی که می‌تواند ناپایستایی داده‌ها را شناسایی کند، تابع موجک است. از مشکلات روش تجزیه موجک وابسته بودن به موجک مادر می‌باشد. تشخیص تابع موجک مادر مناسب با توجه به خروجی‌های متفاوت موجود برای یک سری، زمان‌بر خواهد بود (Ahmadi et al., 2022).

روش تجزیه مد تجزیه (EMD) یکی دیگر از روش‌های تجزیه سیگنال می‌باشد که توسط هانگ و همکاران (Huang et al., 1998) ارائه شده و تاکنون پیشرفت‌های تکاملی فراوانی را طی کرده است. بعد از معرفی روش اولیه، به دلیل مسائل و مشکلات مربوط به ترکیب (اختلاط) مد، وو و هونگ روش تجزیه مد تجزیه گروهی (EEMD) را معرفی نمودند. روش EEMD دارای مشکلاتی از قبیل باقی ماندن مقداری نوفه در داده بازسازی شده از مدهای ذاتی و نیز تولید مدهای ذاتی مختلف در اثر اضافه کردن نوفه‌های گوسی متفاوت است (Wu and Huang, 2009). از این رو، توورس و همکاران با معرفی تجزیه مد تجزیه گروهی کامل (CEEMD) مشکل روش‌های پیشین را برطرف نمودند. در این روش نیز چندین بار نوفه سفید به سیگنال اصلی اضافه می‌گردد، با این تفاوت که در مرحله اول ضمن اضافه کردن نوفه سفید، IMF اول استخراج شده و برای استخراج IMF بعدی با اضافه کردن مجدد نوفه مراحل به همین منوال تکرار می‌گردد (Torres et al., 2011).

دراگومیرسکی و زوسو روش تجزیه مد متغیر (VMD) را برای تجزیه سیگنال ارائه کردند (Dragomiretskiy and Zosso, 2013). VMD می‌تواند سری پیچیده را با قدرت بیشتری نویززدایی کند (Lahmiri, 2015). هی و همکاران برای پیش‌بینی جریان رودخانه از مدل پیشنهادی ترکیبی VMD-DNN که مبتنی بر شبکه عصبی عمیق و تجزیه مد متغیر می‌باشد، استفاده نمودند. همچنین برای ارزیابی بهتر، روش‌های EMD-DNN و EEMD-DNN نیز مورد بررسی قرار گرفتند. نتایج به‌دست‌آمده نشان داد که روش VMD-DNN با $NS = 0.95$ و $RMSE = 9/92$ (m^3/s) و $MAE = 3/82$ مناسبترین عملکرد را به خود اختصاص داده است (He et al., 2019).

با توجه به پیشینه پژوهش ارائه شده، مشاهده می‌شود که برآورد

تاب می‌باشد بهره بردند. این محققان برای پارامترهای ورودی از داده های چهار ایستگاه رودخانه ماهانادی واقع در کشور هند استفاده کردند. برای مقایسه عملکرد مدل مورد نظر از شاخص‌های ارزیابی نش-ساتکلیف (NS)، ویلموت و ریشه میانگین مربعات خطا استفاده شد. نتایج این پژوهش نشان داد که مدل تلفیقی PSR-SVM-FFA از عملکرد مطلوب‌تری نسبت به مدل منفرد SVM برخوردار بوده است (Samantaray & Sahoo, 2024). در پژوهشی دیگر، محمود و همکاران (برای پیش‌بینی دبی رودخانه فرات در بالادست مخزن سد حدیثه در غرب کشور عراق از مدل ماشین بردار پشتیبان بهره بردند. داده‌های جریان روزانه، ماهانه و فصلی طی دوره آماری ۲۰۲۴-۱۹۸۶ به‌عنوان متغیرهای ورودی انتخاب شدند. نتایج حاصل نشان دهنده آن بود که مدل ماشین بردار پشتیبان از عملکرد مطلوبی در برآورد دبی جریان رودخانه فرات برخوردار بوده و با کوتاه‌تر شدن مقیاس زمانی (از فصلی به روزانه) دقت مدل نیز افزایش قابل ملاحظه‌ای می‌یابد (Mahmood et al., 2024). در مطالعه شارما و گوئل جریان ورودی به سد تهری هند را با استفاده از مدل یادگیری ماشین رگرسیون بردار پشتیبان برآورد نموده و عملکرد مناسب آن مورد تایید قرار گرفت (Sharma & Goel, 2024). روش SVM در مطالعات متعددی همچون باقریور و همکاران (۱۳۹۷)، احمدی (۱۳۹۹)، دهقانی و همکاران (۱۴۰۰)، خرمی‌پور و همکاران (۱۴۰۳)، چوبی و همکاران (Choubey et al., 2014) و ساهاو و ساتاپاتی (Sahoo & Satapathy., 2025) به کار گرفته شده و عملکرد آن در پیش‌بینی جریان مطلوب گزارش شده است.

روش جنگل تصادفی توسط بریمان معرفی شد. این روش از یک رویکرد یادگیری ماشینی مجموعه‌ای و تعداد زیادی درخت طبقه‌بندی یا رگرسیون برای حل مسائل مختلف استفاده می‌کند. این روش نیز در مطالعات متعددی به کار گرفته شده است (Breiman, 2001). عبدا و همکاران به‌منظور تخمین جریان روزانه حوضه ساحلی واقعاً در کشور الجزایر مدل‌های یادگیری ماشین شامل مدل جنگل‌های تصادفی، شبکه‌های عصبی مصنوعی و رگرسیون خطی وزنی محلی را به کار گرفتند و دقت و عملکرد مدل‌ها را با شاخص‌های RMSE و R مورد ارزیابی قرار دادند. نتایج نشان داد که مدل جنگل‌های تصادفی بهترین عملکرد را در مرحله آموزش و صحت‌سنجی داشت (Abda et al., 2022). مطالعه لطیف اوغلو حاکی از عملکرد بسیار مطلوب‌تر مدل جنگل تصادفی در پیش‌بینی میان‌مدت جریان رودخانه است (Latifoglu, 2022). در پژوهشی دیگر ایدمودیا و همکاران برای پیش‌بینی جریان رودخانه ایکوپا واقع در جنوب کشور نیجریه، مدل‌های یادگیری ماشین از جمله جنگل تصادفی، درخت تصمیم، الگوریتم تقویت گرادیان و ماشین رگرسیون بردار پشتیبان را به کار گرفتند که نتایج، حاکی از عملکرد بهتر مدل جنگل تصادفی نسبت به سایر روش‌های ذکر شده بود (Idemudia et al., 2024).

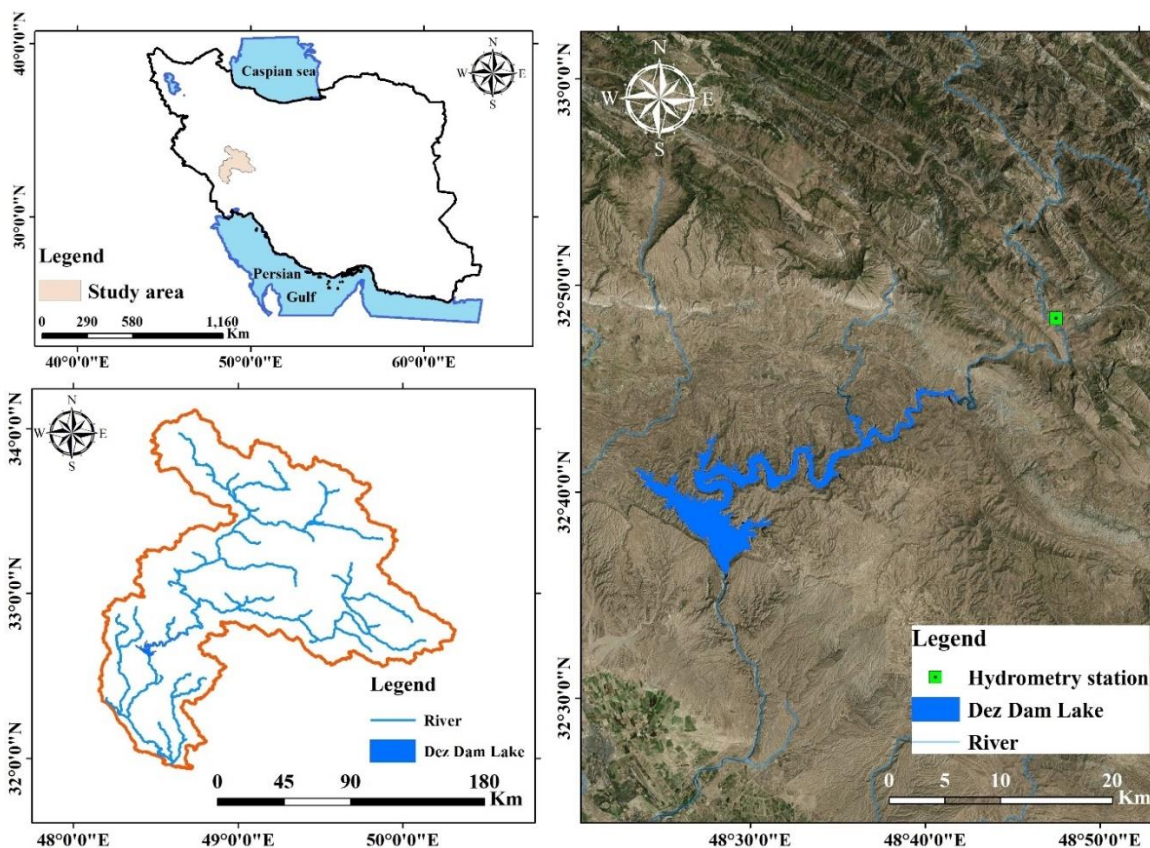
جریان ماهانه رودخانه‌ها همواره مورد توجه پژوهش‌گران بوده و بدین منظور از روش‌های یادگیری ماشین و به‌ویژه RF و SVM استفاده شده‌است. در اکثر مطالعات انجام‌شده روش‌های RF و SVM به صورت منفرد مورد استفاده قرار گرفته و ترکیب آنها با روش‌های پیش پردازش‌کننده نظیر VMD کمتر مورد توجه بوده است. بنابراین هدف از مطالعه حاضر توسعه مدل‌های VMD-SVM و VMD-RF برای پیش‌بینی جریان رودخانه و ارزیابی عملکرد آنها می‌باشد. علاوه بر این در انتخاب الگوهای ورودی مدل‌ها، خاصیت تناوبی نیز با دور رویکرد خطی و غیرخطی در بهبود عملکرد مدل‌ها مورد بحث و بررسی قرار خواهد گرفت.

جریان ماهانه رودخانه‌ها همواره مورد توجه پژوهش‌گران بوده و بدین منظور از روش‌های یادگیری ماشین و به‌ویژه RF و SVM استفاده شده‌است. در اکثر مطالعات انجام‌شده روش‌های RF و SVM به صورت منفرد مورد استفاده قرار گرفته و ترکیب آنها با روش‌های پیش پردازش‌کننده نظیر VMD کمتر مورد توجه بوده است. بنابراین هدف از مطالعه حاضر توسعه مدل‌های VMD-SVM و VMD-RF برای پیش‌بینی جریان رودخانه و ارزیابی عملکرد آنها می‌باشد. علاوه بر این در انتخاب الگوهای ورودی مدل‌ها، خاصیت تناوبی نیز با دور رویکرد خطی و غیرخطی در بهبود عملکرد مدل‌ها مورد بحث و بررسی قرار خواهد گرفت.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌ها

در مطالعه حاضر، از اطلاعات دبی ایستگاه تله‌زنگ واقع در حوضه آبریز دز در طی دوره آماری ۵۱-۱۳۵۰ تا ۴۰۲-۱۴۰۱ استفاده شده است. این حوضه در دسته‌بندی حوضه‌های آبریز کشور، جزء حوضه



شکل ۱- موقعیت حوضه آبریز دز در ایران و موقعیت ایستگاه هیدرومتری تله‌زنگ

جدول ۱- مشخصات آماری دبی ایستگاه هیدرومتری تله زنگ در دوره آماری ۱۳۵۰-۱۴۰۱

مقیاس زمانی	کمینه (m ³ /s)	میانگین (m ³ /s)	بیشینه (m ³ /s)	انحراف معیار (m ³ /s)	ضریب تغییرات (%)
مهر	۲۵/۱	۶۵/۶	۱۳۸/۰	۲۱/۳	۳۲/۴
آبان	۳۲/۴	۱۰۰/۹	۵۰۰/۰	۶۹/۷	۶۹/۱
آذر	۵۰/۹	۱۸۹/۳	۱۰۲۳/۰	۱۴۹/۰	۸۷/۷
دی	۴۵/۰	۲۱۱/۹	۶۳۲/۰	۱۱۸/۴	۵۵/۹
بهمن	۷۱/۹	۲۸۳/۰	۱۰۱۷/۰	۱۷۵/۵	۶۲/۰
اسفند	۱۴۲/۰	۳۹۹/۳	۱۲۸۸/۰	۲۴۲/۲	۶۰/۷
فروردین	۹۵/۱	۵۹۲/۸	۱۸۲۸/۷	۳۳۳/۵	۵۶/۳
اردیبهشت	۱۲۴/۰	۴۶۸/۴	۱۴۱۵/۰	۲۶۶/۰	۵۶/۸
خرداد	۱۰۲/۰	۲۶۰/۴	۷۲۳/۰	۱۲۹/۲	۴۹/۶
تیر	۷۰/۶	۱۵۵/۰	۴۰۶/۰	۶۷/۶	۴۳/۶
مرداد	۴۷/۴	۱۰۶/۶	۲۵۹/۰	۴۵/۴	۴۲/۶
شهریور	۳۰/۲	۷۷/۱	۱۸۰/۰	۲۷/۴	۳۵/۶
سالانه	۸۵/۶	۲۴۲/۵	۵۸۲/۳	۱۰۰/۱	۴۱/۳

مدل ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان در حقیقت یک سیستم یادگیری کاربردی است که بر پایه اصل بهینه‌سازی استوار شده‌است و برای رسیدن به یک جواب بهینه، از کمینه‌سازی خطای ساختاری استفاده می‌کند. به عبارت دیگر، SVM یک الگوریتم یادگیری ماشین است که بدون استفاده از مفاهیم آماری، داده‌ها را به دو گروه تقسیم می‌کند (Mantero et al., 2005). این روش اولین بار توسط (Vapnik, 1995) ارائه شد. در مدل رگرسیون ماشین بردار پشتیبان (SVM) هدف، پیش‌بینی یک متغیر وابسته (Y) بر اساس چندین متغیر مستقل (X) است. این مدل، رابطه بین متغیرها را با استفاده از یک تابع ریاضی مدل می‌کند که شامل ترکیبی خطی از متغیرهای مستقل و یک ضریب ثابت به شرح زیر است:

$$f(x) = W^T \cdot \phi(x) + b \quad (1)$$

چنانچه W (بردار ضرایب) و b (ثابت) مشخصه‌های تابع رگرسیون و ϕ نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با آموزش مدل SVM توسط مجموعه‌ای از نمونه‌ها (مجموعه آموزش) محقق می‌شود. این فرایند شامل بهینه‌سازی یک تابع خطا است. بنابراین برای برآورد W و b بایستی تابع خطا زیر با در نظر گرفتن قیودات آن بهینه شود (Hamel, 2009):

$$\frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \quad (2)$$

$$\text{Subject to: } \begin{cases} W^T \cdot \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ y_i - W^T \cdot \phi(x_i) - b \leq \varepsilon + \xi_i \\ \xi_i^*, \xi_i \geq 0, i = 1, \dots, N \end{cases}$$

در روابط فوق، مقدار مثبت عدد C میزان جریمه‌ای را تعیین می‌کند

که در صورت بروز خطا در فرایند آموزش مدل اعمال می‌شود. N تعداد نمونه‌ها، ϕ تابع کرنل، ξ_i^* و ξ_i متغیرهای کمبود هستند و خطای مجاز ε را مشخص می‌کنند. در مسائل فرض می‌شود که داده‌ها، درون بازه مرزی ε قرار گیرند؛ حال اگر داده‌ای خارج از بازه ε قرار گرفت آنگاه مقدار خطایی معادل با ξ_i^* و ξ_i وجود خواهد داشت. شایان‌ذکر است که مدل SVM مشکلات ناشی از کم تخمینی و فوق‌برازشی را با کمینه‌کردن همزمان عبارت $C \sum_{i=1}^N (\xi_i + \xi_i^*)$ در رابطه فوق حل می‌کند. بنابراین، با معرفی دو ضریب لاگرانژ α_i^* و α_i مسئله بهینه‌سازی با حداکثرسازی عددی تابع درجه دوم به رابطه زیر با شرایط مشخص حل خواهد شد (Fletcher., 2009).

$$\text{Max: } Z = \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^N (\alpha_i - \alpha_i^*) - \quad (3)$$

$$0.5 \sum_{i,j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \phi(x_i)^T \cdot \phi(x_j)$$

$$\text{Subject to: } \begin{cases} \sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C, i = 1, 2, \dots, 1 \\ 0 \leq \alpha_i^* \leq C, i = 1, 2, \dots, 1 \end{cases}$$

تابع هدف در رابطه فوق، تابع محدب است و به همین سبب جواب معادله بهینه خواهد بود و پس از تعریف ضرایب لاگرانژ، مشخصه‌های W و b در مدل SVM رگرسیون با استفاده از شرایط تئوری کرش-کوهن-تاگر محاسبه می‌شوند که در آن SVM رگرسیون خواهیم داشت (Fletcher, 2009):

درختان، رهیافت نهایی فرایند ساخت مدل است.

برای رسیدن به نتیجه نهایی، هر درخت در مدل جنگل تصادفی با بهینه‌سازی یک تابع خطا در هر تقسیم، توسعه داده می‌شود (میانگین مربع خطا برای این مطالعه انتخاب شد). در RF، زیرمجموعه ای تصادفی از مجموعه کامل پیش‌بینی‌کننده‌ها برای هر گره انتخاب می‌شود و از این زیرمجموعه، یک تقسیم (بوت استریپینگ) انتخاب می‌شود تا اطمینان حاصل شود که پارامترهای تأثیرگذار، تمامی درختان را تحت‌الشعاع قرار نمی‌دهند. پارامترهای مختلفی می‌توانند برای تنظیم الگوریتم استفاده شوند، مانند متغیر تعداد درختان. برای کنترل تعداد درختان در جنگل، تعداد متغیرهای موجود می‌تواند تقسیم و اندازه گره عمق هر درخت را کنترل می‌کند. در این مطالعه، پارامترهای تعداد درختان و اندازه گره در تمامی اجراهای مدل ثابت نگه‌داشته شدند. تعداد درختان در مرحله تنظیم مدل بر روی ۲۰۰ و در مرحله آموزش مدل نهایی بر روی ۱۰۰۰ تنظیم شد، در حالی که اندازه گره برای هر دو مرحله عدد پنج در نظر گرفته شد. عملکرد مدل با تعداد درختان بیشتر از ۱۰۰ به‌طور قابل توجهی تحت تأثیر قرار نمی‌گیرد، اما پایداری واریانس با افزایش تعداد درختان بهبود می‌یابد. به همین دلیل، ۱۰۰۰ درخت برای آموزش مدل نهایی انتخاب شد (Doyle et al., 2019).

الگوریتم تجزیه مد متغیر (VMD)

تجزیه مد متغیر (VMD) مدلی نوظهور با توانمندی بالا می‌باشد که به‌منظور تجزیه یک سری زمانی ثابت و غیر ایستا به توابع حالت ذاتی (IMFs) مورد استفاده قرار می‌گیرد. به این سبب مدها در اطراف فرکانس مرکزی دارای خاصیت فشرده‌گی هستند. در نتیجه IMFs می‌تواند داده‌های اولیه را با دقت بالایی بازسازی کند (Dragomiretskiy and Zosso, 2013). برای پی‌بردن به هر حالت و فرکانس مرکزی آن، مسئله تغییرات محدود را می‌توان به‌شکل زیر نوشت (He et al., 2019):

$$\min_{\{u_k(t)\}, \{\omega_k(t)\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (6)$$

$$s.t. \sum_k u_k(t) = f(t)$$

در معادله فوق: t گام زمانی، $\delta(t)$ توزیع دیراک، $u_k(t)$ امین مد، $\omega_k(t)$ فرکانس مرکزی مربوطه و $f(t)$ امین داده سیگنال می‌باشد. علاوه بر این، تبدیل هیلبرت $u_k(t)$ با $\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t)$ بیان شده‌است، که می‌تواند $u_k(t)$ را به داده‌های تحلیلی تبدیل کند تا یک طیف فرکانس یک طرفه با فرکانس‌های مثبت ایجاد کند. در نتیجه، با توجه به اصطلاح شاخص، طیف مدها را می‌توان به یک باند پایه منتقل کرد. به‌منظور تبدیل

$$W = \sum_{i=1}^N (a_i - a_i^*) \phi(x_i)^T \cdot \phi(x) + b \quad (4)$$

ترم‌های لاگرانژ $(\alpha_i - \alpha_i^*)$ می‌توانند صفر یا غیر صفر باشند. در نتیجه فقط مجموعه داده‌هایی که ضرایب α_i آنها غیر صفر است، در معادله رگرسیون نهایی وارد می‌شوند که به‌عنوان بردارهای پشتیبان شناخته می‌شوند. در میان بردارها آنهایی که مقدار $|\alpha_i|$ آنها کمتر از C باشد بردارهای پشتیبان حاشیه‌ای نامیده می‌شوند. هنگامی که مقدار $|\alpha_i|$ بردارهای پشتیبان برابر مقدار C باشد، به‌عنوان بردار پشتیبان خطا، یا بردار پشتیبان کران دار شناخته می‌شود. در نهایت تابع SVM رگرسیون را می‌توان به فرم معادله (δ) بازنویسی کرد:

$$f(x) = \sum_{i=1}^N a_i \cdot \phi(x) + b \quad (5)$$

که در این معادله، محاسبه $\phi(x)$ ممکن است دارای پیچیدگی‌های بسیار باشد. به‌منظور حل این مشکل به‌طور معمول از یک کرنل یا تابع هسته استفاده می‌شود. با توجه به مطالعات انجام‌شده، تابع کرنل شعاعی یا RBF یکی از مناسب‌ترین هسته‌ها برای حل مسائل پیچیده نظیر پیش‌بینی جریان رودخانه می‌باشد (موسی زاده و همکاران، ۱۴۰۱).

مدل جنگل تصادفی (RF)

مدل جنگل تصادفی (RF) یک روش یادگیری جمعی مبتنی بر توسعه درخت تصمیم است که برای مسائل رگرسیون و خوشه‌بندی به کار می‌رود. این روش برای اولین بار توسط بریمن در سال ۲۰۰۱ معرفی شد (Breiman., 2001). جنگل تصادفی در واقع یک دسته بندی‌کننده گروهی بر پایه درخت تصمیم‌گیری است که در مقایسه با سایر دسته‌بندی‌کننده‌های یادگیری ماشین عملکرد بهتری دارد (Cutler et al., 2012). این الگوریتم به‌دلیل پایداری و تعمیم‌دهی خوب آن، به‌طور گسترده‌ای در زمینه‌های مختلف استفاده می‌شود. جنگل تصادفی مجموعه‌ای از درخت‌های تصمیم ساده است که هر درخت با انتخاب نمونه‌ها و ویژگی‌های تصادفی ایجاد می‌شود و نتایج حاصل از این درختان، خروجی نهایی جنگل تصادفی محسوب می‌شود. برای کاهش خطا در این روش، دو پارامتر مهم بهینه‌سازی می‌شود: الف) تعداد درختان و ب) تعداد ویژگی‌هایی که به‌عنوان داده‌های خارج از کیسه شناخته می‌شوند، باید از داده‌های اصلی حذف شوند. هنگامی که RF ایجاد می‌شود، با ورود برداری مانند بردار X به داخل آن، این بردار وارد همه درختان می‌شود و در هر درخت، یک کلاس را انتخاب می‌کند؛ در نتیجه کلاسی که بیشترین آرا را داشته‌باشد، کلاس پیش‌بینی RF برای این بردار ورودی خواهد بود. میانگین نتایج تمام

خطی و غیرخطی بهره گرفته شد. بدین منظور از عبارت‌های $\cos[2\pi \cdot i/12]$ و $\sin[2\pi \cdot i/12]$ ($i = 1, \dots, 12$) و ضرایب ساده ماهانه تحت عنوان α ($\alpha = 1, \dots, 12$) برای بسط اثر خاصیت پربودیک غیرخطی ($NLPT^2$) و خطی (LPT^3) استفاده شده و به الگوهای ورودی افزوده شد. در ساختار سوم، پیش‌پردازش داده‌ها به منظور حذف اثر نوفه‌ها و نیز انتقال بهتر اطلاعات به مدل‌های منفرد، از روش تجزیه مد متغیر (VMD) استفاده شد. به عبارت دیگر در ابتدا الگوهای ورودی LB تجزیه و سپس به مدل‌ها معرفی شده و روش‌های تلفیقی VMD-RF و VMD-SVM توسعه داده شدند. جدول (۲) الگوهای ورودی توسعه یافته برای هر یک از مدل‌ها را نشان می‌دهد.

ارزیابی مدل‌ها

در این پژوهش، به منظور ارزیابی مدل‌های به کار گرفته شده در پیش‌بینی جریان در محل ایستگاه تله‌زنگ از معیارهای جذر میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE)، ضریب کارایی نش-ساتکلیف (NS) و ضریب همبستگی (R) استفاده شده است:

$$RMSE = \left(\frac{\sum_{i=1}^n (Q_i - \bar{Q}_i)^2}{n} \right)^{0.5} \quad (11)$$

$$MAE = \frac{\sum_{i=1}^n |Q_i - \bar{Q}_i|}{n} \quad (12)$$

$$NS = 1 - \frac{\sum_{i=1}^n (Q_i - \bar{Q}_i)^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2} \quad (13)$$

$$R = \frac{\sum_{i=1}^n (Q_i - \bar{Q}_i)(Q_i - \bar{Q})}{\sqrt{\sum_{i=1}^n (Q_i - \bar{Q}_i)^2} \sqrt{\sum_{i=1}^n (Q_i - \bar{Q})^2}} \quad (14)$$

که در روابط فوق، Q_i مقدار مشاهده شده در گام زمانی i ام، \bar{Q}_i مقدار محاسبه شده در همان زمان، n تعداد داده‌ها و \bar{Q} میانگین مقادیر مشاهداتی و \bar{Q}_i میانگین مقادیر محاسباتی می‌باشد. مدلی که دارای کمترین مقدار RMSE و MAE و نیز بیشترین مقدار R و NS باشد، مناسب‌ترین گزینه خواهد بود (احمدی، ۱۴۰۰).

مسئله بهینه‌سازی داده شده به یک عبارت بهینه‌سازی غیر هدف، دو جنبه باید در نظر گرفته شود: ضرب کننده‌های لاگرانژ و پارامتر جریمه درجه دوم. معادله زیر معادله تابع لاگرانژ تقویت شده است (Li et al., 2022):

$$L(u_k, \omega_k, \lambda) = \alpha \sum_{k=1}^K \left\| \hat{\partial}_t \left[(\delta(t) + \frac{j}{\pi t}) \otimes u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \quad (7)$$

$$+ \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t) f(t) - \sum_{k=1}^K u_k(t) \right\rangle$$

در این رابطه، $\left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2$ تابع جریمه درجه دوم

برای کاهش زمان همگرایی می‌باشد. روش جهت متناوب ضرب کننده‌ها (ADMM) یک روش بهینه‌سازی است و در دو نقطه مختلف به روز می‌شود تا اجازه دهد روش VMD تکمیل شود. در نتیجه، معادلات اصلاح شده به شرح زیر قابل بیان می‌باشد (He et al., 2019):

$$u_k^{n+1} = \frac{f(\omega) - \sum_{i \neq k} u_i(\omega) + \frac{\lambda(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)} \quad (8)$$

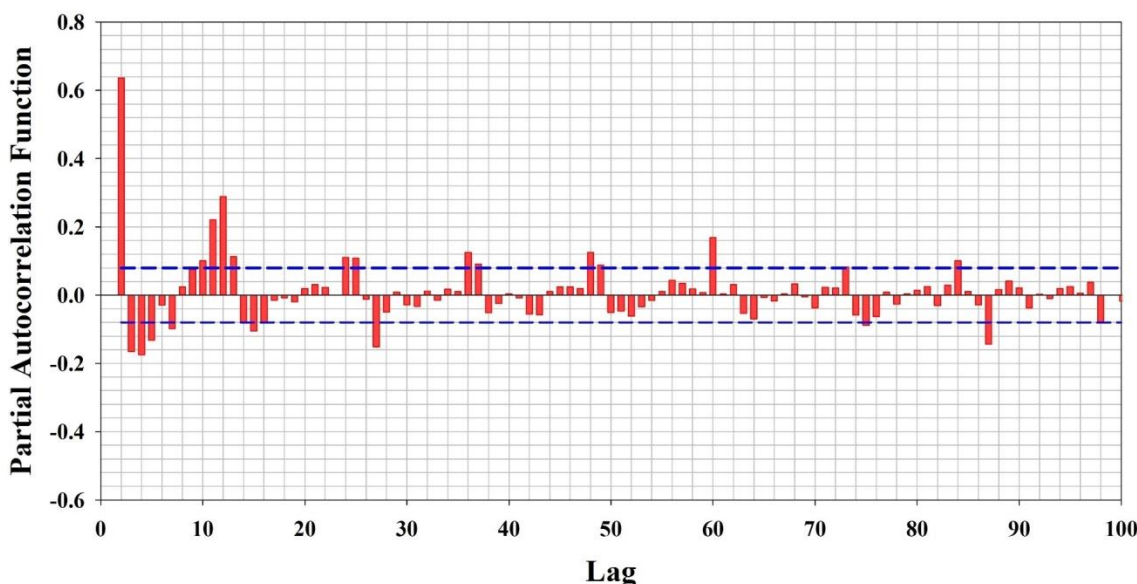
$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |u_k(\omega)|^2 d\omega}{\int_0^\infty |u_k(\omega)|^2 d\omega} \quad (9)$$

$$\lambda^{n+1}(\omega) = \lambda^n(\omega) + \eta \left(f(\omega) - \sum_{k=1}^K u_k^{n+1} \right) \quad (10)$$

که در آن، n تعداد تکرار است. $f(\omega)$ ، $\lambda(\omega)$ و $u_i(\omega)$ پارامترهای تبدیل فوریه هستند و η ضریب تکراری را نشان می‌دهد.

توسعه مدل

یکی از مهم‌ترین جنبه‌های مدل‌سازی جریان رودخانه، انتخاب ورودی‌های مناسب برای مدل‌ها می‌باشد. به منظور ایجاد روابط مناسب بین پدیده‌ها، مدل‌ها به ورودی‌هایی نیاز دارند. بنابراین در مطالعه حاضر سعی گردید الگوهای ورودی با حالت‌های مختلف و براساس انتخاب از طریق تابع خودهمبستگی جزئی توسعه یابد. بدین منظور سه ساختار ورودی تعریف گردید. در ساختار اول، تأخیر مناسب برای مدل‌سازی جریان رودخانه با استفاده از روش تابع خودهمبستگی جزئی وزن‌دهی شده و پس از انتخاب تأخیرهای مؤثر، الگوهای تأخیر پایه (LB^1) ایجاد شدند (شکل ۲). همان گونه که شکل (۲) نشان می‌دهد، تأخیر زمانی چهار ماهه برای انجام مدل‌سازی مناسب می‌باشد. در ساختار دوم، علاوه بر متغیرهای منتخب، از خاصیت فصلی



شکل ۲- نمودار خودهمبستگی جزئی (PACF) جریان در محل ایستگاه هیدرومتری تله زنگ

جدول ۲- الگوهای ورودی تحت سناریوهای مختلف

سناریو	ردیف	الگو	آرایش الگوی ورودی
تأخیر داده‌های جریان (S1)	۱	S1M1	$Q_t = f(Q_{t-1})$
	۲	S1M2	$Q_t = f(Q_{t-1}, Q_{t-2})$
	۳	S1M3	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3})$
	۴	S1M4	$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4})$
تأخیر داده‌های جریان با در نظر گرفتن ترم تناوبی (S2)	۵	S2M1	$Q_t = f(LPT, NLPT, Q_{t-1})$
	۶	S2M2	$Q_t = f(LPT, NLPT, Q_{t-1}, Q_{t-2})$
	۷	S2M3	$Q_t = f(LPT, NLPT, Q_{t-1}, Q_{t-2}, Q_{t-3})$
	۸	S2M4	$Q_t = f(LPT, NLPT, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4})$
تجزیه داده‌های جریان با استفاده از روش VMD (S3)	۱۹	S3M1	$Q_t = f(VMD(Q_{t-1}))$
	۲۰	S3M2	$Q_t = f(VMD(Q_{t-1}, Q_{t-2}))$
	۲۱	S3M3	$Q_t = f(VMD(Q_{t-1}, Q_{t-2}, Q_{t-3}))$
	۲۲	S3M4	$Q_t = f(VMD(Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}))$

نتایج و بحث

پیش‌بینی جریان ماهانه رودخانه با استفاده از مدل‌های مستقل و تناوبی

در این پژوهش، به منظور مدل‌سازی و پیش‌بینی جریان ماهانه رودخانه دز، از دو الگوریتم قدرتمند یادگیری ماشین تحت عنوان ماشین بردار پشتیبان (SVM) و جنگل تصادفی (RF) استفاده گردید. فرآیند مدل‌سازی بر اساس هشت الگوی ورودی مختلف (ارائه شده در جدول ۲) طراحی شد که هر یک از این الگوها با در نظر گرفتن پارامترهای هیدرولوژیکی و خاصیت تناوبی مؤثر بر جریان رودخانه تنظیم گردیدند. داده‌های مورد استفاده در این مطالعه به دو

بخش مجزای آموزش (۸۰ درصد، معادل ۵۰۴ ماه) و آزمون (۲۰ درصد، معادل ۱۲۰ ماه) تقسیم‌بندی شدند تا امکان ارزیابی دقیق عملکرد مدل‌ها فراهم گردد.

یکی از مراحل حیاتی در پیاده‌سازی مدل‌های یادگیری ماشین، انتخاب تابع کرنل مناسب است که تأثیر مستقیمی بر دقت و کارایی پیش‌بینی‌ها دارد. بر اساس پژوهش‌های پیشین، از جمله مطالعه احمدی (۱۳۹۹)، مشخص شده است که تابع کرنل پایه شعاعی (RBF) در مواجهه با مسائل پیچیده هیدرولوژیکی که از روابط غیرخطی و پیچیده برخوردارند، عملکرد بهتری نسبت به سایر توابع کرنل از خود نشان می‌دهد. این تابع با قابلیت بالای خود در شناسایی

مدل RF کاربرد ترم تناوبی کارایی مدل را براساس شاخص RMSE بهبود بخشیده است به طوری که این شاخص بدون در نظر گرفتن ترم تناوبی برای الگوهای ورودی S1M1 تا S1M4 در مرحله آزمون به طور متوسط برابر با ۱۸۷/۹ متر مکعب بر ثانیه بوده و این در حالی است که برای الگوهای S2M1 تا S2M4 این مقدار به ۱۶۲/۹ متر مکعب بر ثانیه کاهش یافته است. بهبود در شاخص NS برای روش RF نیز با در نظر گرفتن خاصیت تناوبی در مرحله آزمون وجود داشته و حدود ۰/۱۹۷ بهبود مشاهده می شود.

الگوهای پیچیده و تعمیم پذیری مناسب، به ویژه در مدل سازی پدیده های طبیعی، مانند جریان رودخانه ها نتایج قابل اطمینان تری ارائه می دهد. بنابراین در مطالعه حاضر، تابع کرنل پایه شعاعی برای ادامه کار انتخاب گردید.

در جدول ۳ نتایج شاخص های ارزیابی RMSE، MAE، R و NS برای مدل های RF و SVM ارائه شده است. با توجه به این جداول مشاهده می شود که الگوهای ورودی منتخب در دقت و عملکرد مدل های منفرد تأثیر مستقیمی داشته اند. به عنوان مثال، برای

جدول ۳ - نتایج شاخص های ارزیابی RMSE، MAE، R و NS برای مدل های RF و SVM

مدل	مرحله آموزش				مرحله آزمون					
	الگو	RMSE (m ³ /s)	MAE (m ³ /s)	R	NS	الگو	RMSE (m ³ /s)	MAE (m ³ /s)	R	NS
RF	S1M1	۱۰۱/۵۴۲	۶۲/۲۵۳	-۰/۹۰۶	-۰/۸۱۴	S1M1	۱۹۰/۹۴۸	۱۰۲/۷۷۴	-۰/۴۹۴	-۰/۱۷۶
	S1M2	۶۹/۵۸۶	۴۰/۰۷۲	-۰/۹۶۳	-۰/۹۱۳	S1M2	۱۸۰/۳۵۲	۸۶/۷۴۵	+۰/۵۴۶	+۰/۲۶۵
	S1M3	۶۶/۲۱۵	۳۸/۰۹۷	-۰/۹۶۷	-۰/۹۲۱	S1M3	۱۸۹/۸۴۸	۹۲/۶۳۲	-۰/۵۰۰	-۰/۱۸۶
	S1M4	۶۵/۷۱۶	۳۸/۱۶۹	-۰/۹۶۷	-۰/۹۲۲	S1M4	۱۹۰/۴۶۱	۹۷/۱۶۹	-۰/۵۱۵	-۰/۱۸۱
	S2M1	۶۶/۴۸۹	۳۶/۶۴۴	-۰/۹۶۲	-۰/۹۲۰	S2M1	۱۷۲/۴۲۴	۸۲/۵۴۴	-۰/۶۱۷	-۰/۳۲۸
	S2M2	۵۵/۸۸۷	۳۰/۸۷۸	-۰/۹۷۵	-۰/۹۴۴	S2M2	۱۶۲/۶۶۷	۸۰/۵۸۴	-۰/۶۵۶	-۰/۴۰۲
	S2M3	۵۶/۵۹۱	۳۰/۸۷۸	-۰/۹۷۵	-۰/۹۴۲	S2M3	۱۵۹/۱۰۰	۷۷/۳۳۹	-۰/۶۶۶	-۰/۴۲۸
	S2M4	۵۵/۸۸۶	۳۱/۵۶۴	-۰/۹۷۶	-۰/۹۴۴	S2M4	۱۵۷/۷۷۶	۸۱/۱۸۹	+۰/۶۸۳	+۰/۴۳۸
SVM	S1M1	۱۶۹/۳۳۳	۱۱۰/۹۰۶	-۰/۶۹۴	-۰/۴۷۹	S1M1	۱۷۵/۵۷۴	۹۴/۶۵۱	+۰/۵۶۳	+۰/۳۰۴
	S1M2	۱۴۴/۷۹۷	۹۰/۸۷۵	-۰/۷۹۲	-۰/۶۲۲	S1M2	۱۷۹/۷۰۹	۸۷/۴۴۲	-۰/۵۳۸	-۰/۲۷۱
	S1M3	۱۵۷/۲۷۷	۹۶/۳۳۴	-۰/۷۴۵	-۰/۵۵۴	S1M3	۱۸۳/۵۶۳	۹۴/۴۵۶	-۰/۵۱۸	-۰/۲۳۹
	S1M4	۱۵۵/۶۱۱	۹۳/۶۱۳	-۰/۷۵۲	-۰/۵۶۴	S1M4	۱۸۱/۳۱۷	۹۴/۴۰۹	-۰/۵۳۸	-۰/۲۵۷
	S2M1	۱۲۰/۵۹۸	۶۸/۵۷۶	-۰/۸۶۰	-۰/۷۳۷	S2M1	۱۵۷/۸۱۰	۷۶/۹۲۳	-۰/۶۷۳	-۰/۴۳۷
	S2M2	۱۱۹/۵۳۵	۶۸/۴۷۲	-۰/۸۶۲	-۰/۷۴۲	S2M2	۱۵۸/۶۷۱	۷۶/۱۳۱	-۰/۶۶۳	-۰/۴۳۱
	S2M3	۱۲۷/۱۴۳	۷۳/۳۱۰	-۰/۸۴۲	-۰/۷۰۹	S2M3	۱۵۴/۸۰۷	۷۶/۳۹۸	+۰/۶۸۲	+۰/۴۵۹
	S2M4	۱۲۵/۰۱۶	۷۲/۳۵۱	-۰/۸۴۸	-۰/۷۱۹	S2M4	۱۵۶/۰۶۸	۷۷/۱۶۵	-۰/۶۷۶	-۰/۴۵۰

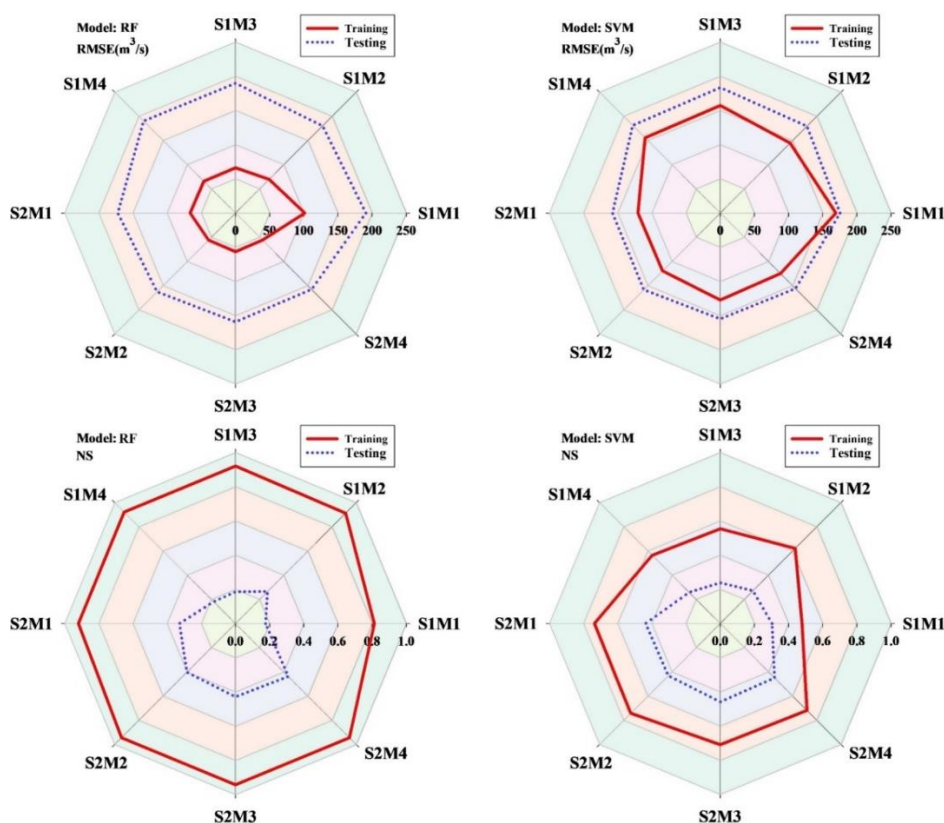
به ورودی های کمتری نیز نیاز دارد به طوری که SVM با الگوی S2M3 کمترین خطا (برابر با ۱۵۴/۸۰۷ متر مکعب بر ثانیه) را نتیجه داده اما مدل RF با الگوی S2M4 خطایی برابر با ۱۵۷/۷۷۶ متر مکعب بر ثانیه را ثبت کرده است. به عبارت دیگر، عملکرد مدل SVM در مقایسه با RF هم از نظر خطا و هم از نظر تعداد داده های مورد نیاز برای انجام مدل سازی جریان ماهانه بهتر بوده است.

نکته مهم دیگری که باید مورد توجه قرار گیرد، تفاوت قابل ملاحظه در معیارهای ارزیابی مراحل آموزش و آزمون در مدل های منفرد می باشد. این اختلاف، رخداد پدیده بیش برازش (اورفیتینگ) را در مدل های SVM و RF نشان می دهد. اورفیتینگ زمانی رخ می دهد که مدل به جای یادگیری الگوهای کلی داده های آموزشی، نویزها و

پژوهش های پیشین به خوبی تأثیر مثبت ترم تناوبی را در بهبود دقت مدل های یادگیری ماشین تأیید کرده اند؛ به طوری که احمدی (۱۳۹۹) در مطالعه خود نشان داد استفاده از خاصیت تناوبی می تواند عملکرد مدل ها را در پیش بینی جریان رودخانه به طور قابل ملاحظه ای ارتقا دهد. این یافته ها با نتایج تحقیقات دیگر پژوهشگران از جمله منتظری و قویدل (۱۳۹۳) که بر کاربرد مؤثر ترم تناوبی در فرآیند مدل سازی تأکید داشته اند، همسو و همخوان می باشد. مجموعه این مطالعات و نتایج به دست آمده از پژوهش حاضر نشان می دهد که در نظر گرفتن ویژگی های تناوبی در اطلاعات ورودی مدل های یادگیری ماشین می تواند به عنوان یک راهکار کارآمد برای افزایش دقت پیش بینی ها مورد توجه قرار گیرد. مقایسه شاخص MAE برای مدل های منفرد RF و SVM نشان می دهد که مدل ماشین بردار پشتیبان در مقایسه با روش جنگل های تصادفی از عملکرد بهتری برخوردار بوده و

(Nasrabadi, 2006). یکی از راهکارهای مؤثر در حل این مشکل افزودن داده‌های آموزشی می‌باشد (Chollet, 2021). در مطالعه حاضر خاصیت تناوبی نیز به ورودی‌های مدل‌ها اضافه شد اما به گواه شکل ۳ نتوانسته است بهبود قابل توجهی در عملکرد مدل داشته باشد. بنابراین، راهکار دیگری تحت عنوان توسعه مدل‌های تلفیقی مبتنی بر تجزیه داده‌ها با روش VMD مورد توجه قرار گرفت.

جزئیات خاص آنها را حفظ می‌کند، که منجر به عملکرد ضعیف در داده‌های جدید می‌شود. برای تشخیص این پدیده، می‌توان از شاخص‌هایی مانند اختلاف زیاد بین دقت آموزش و دقت آزمون یا افزایش ناگهانی خطا (مانند RMSE) در داده‌های آزمون استفاده کرد (James et al., 2021). به عبارت دیگر اگر خطای آموزش بسیار پایین باشد اما خطای آزمون بالا باقی بماند، نشانه اورفیتینگ است (Bishop and



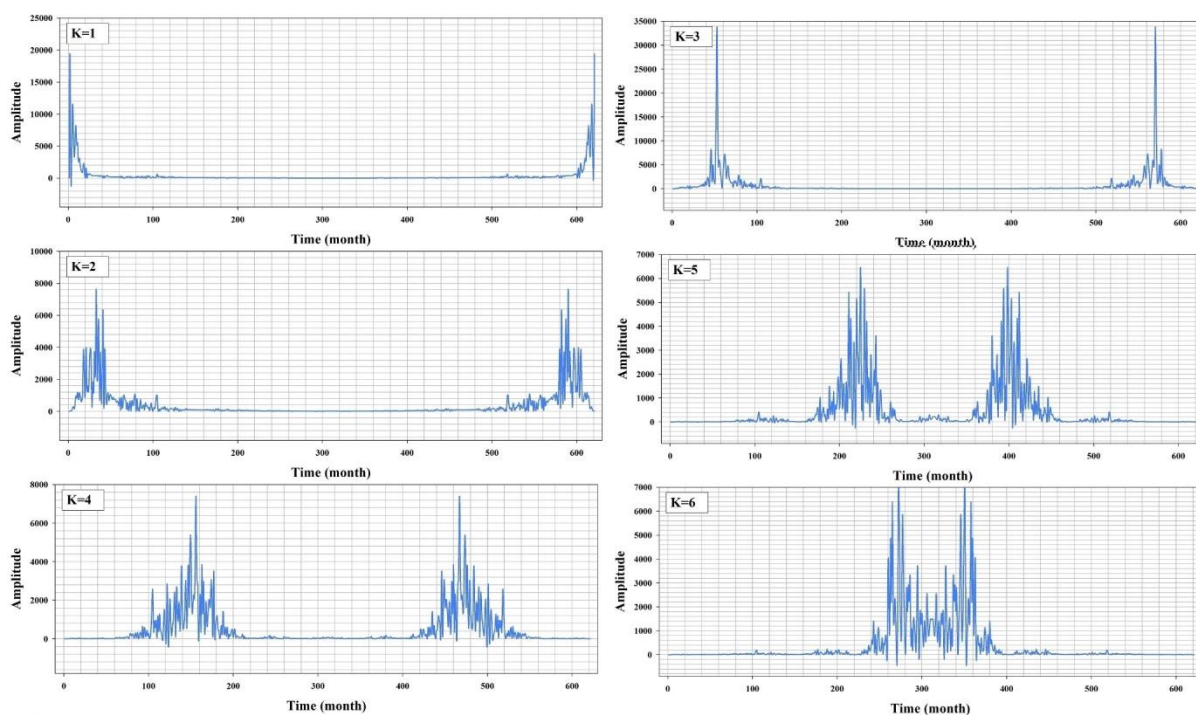
شکل ۳- نمودار راداری معیارهای ارزیابی RMSE و NS برای مدل‌های منفرد و سناریوهای مختلف در دو مرحله آموزش و آزمون

مدل‌ها کاهش می‌یابد. از طرف دیگر، چنانچه سطح تجزیه زیاد انتخاب شود، اختلاط مد فرکانس‌های مرکزی پیش می‌آید و نتایج غیرقابل اعتماد خواهد بود. در این مطالعه برای انتخاب سطح تجزیه مناسب، داده‌های جریان از دو تا نه سطح با استفاده از VMD تجزیه شدند. سپس نمودارهای فرکانس مرکزی برای هر سری ورودی ترسیم شد. نتایج حاصل از تجزیه داده‌ها در سطوح مختلف و فرکانس‌های مرکزی آنها در شکل (۴) برای ورودی Qt-1 ارائه شده است. با توجه به این شکل مشاهده می‌شود که تا سطح تجزیه پنج، فرکانس‌های مرکزی به یکدیگر نزدیک شده و هیچ اختلاط مدی اتفاق نمی‌افتد. اما در سطح تجزیه شش اختلاط مد پیش می‌آید. بنابراین مقدار مناسب سطح تجزیه برای تحلیل داده‌های جریان ماهانه رودخانه با یک تأخیر برابر با پنج می‌باشد.

تجزیه داده‌ها با استفاده از روش تجزیه مد متغیر (VMD)

در این مطالعه داده‌های جریان رودخانه با استفاده از روش VMD تجزیه شد. در روش VMD می‌باید دو پارامتر مهم شامل سطح تجزیه مناسب (K) و مقدار جریمه برای متعادل کردن محدودیت وفاداری داده (alpha) تعیین گردد (He et al., 2020). Alfa پهنای باند IMF ها را مشخص می‌کند. مقدار بزرگ‌تر Alfa پهنای باند گسترده‌تر را نتیجه می‌دهد و از طریق آزمون و خطا به دست می‌آید (Li et al., 2021). در این مطالعه، مقادیر مختلفی از ۱۰ تا ۲۰۰۰ مورد بررسی قرار گرفته و در نهایت Alfa=1000 بهترین نتیجه را در تجزیه داده‌های جریان ماهانه رودخانه به دست داد.

انتخاب سطح تجزیه (K) مناسب تعیین کننده تعداد گام‌های اجرای VMD می‌باشد. اگر تعداد سطح تجزیه کم انتخاب شود در این صورت همه اطلاعات از سری زمانی استخراج نمی‌شود و عملکرد



شکل ۴- فرکانس تجزیه VMD با سطوح تجزیه یک تا شش

نیز منعکس گردد. با بررسی آماره NS مشاهده می‌شود که متوسط مقدار آن در مدل VMD-RF نسبت به مدل مستقل به‌طور میانگین از $0/30$ به $0/722$ افزایش یافته‌است. این امر نشان دهنده تأثیر مثبت پیش‌پردازش داده‌ها در فرآیند مدل‌سازی می‌باشد.

بهبود دقت برآورد جریان ماهانه رودخانه برای مدل VMD-SVM نیز مشهود است (جدول ۴). با توجه به نتایج به‌دست آمده می‌توان چنین استنباط کرد که مدل ترکیبی VMD-SVM توانسته است ضمن کاهش خطای تخمین، کارایی برآوردها را نیز براساس شاخص NS به‌طور متوسط از $0/356$ به $0/911$ ارتقا دهد. بررسی و مقایسه مقادیر شاخص RMSE در مرحله آزمون برای الگوهای ورودی S1M1 تا S2M4 مدل منفرد و ترکیبی SVM نشان می‌دهد که تک تک مقادیر پیش‌بینی شده در مدل تلفیقی تا حدود زیادی با مقادیر اندازه‌گیری شده مطابقت دارد. مقدار این شاخص آماری برای بهترین الگوهای مختلف ورودی در مدل SVM بین $154/807$ تا $183/563$ متر مکعب در ثانیه در حال تغییر می‌باشد، در حالی که این مقدار برای مدل VMD-SVM بین $21/441$ تا $118/699$ متغیر است و بنابراین می‌توان نتیجه گرفت که مدل ترکیبی توسعه داده شده در مقایسه با مدل منفرد SVM توانسته است با ورودی‌های مختلف مقادیر مشاهداتی را به‌صورت نظیر به نظیر بهتر برآورد نماید.

نتایج حاصل از مدل‌های تلفیقی مبتنی بر تجزیه داده‌ها با استفاده از روش VMD

همان‌گونه که اشاره شد، مدل‌های منفرد قادر به استفاده کامل و بهینه از اطلاعات در دسترس نیستند. بنابراین، داده‌های ورودی با استفاده از روش تجزیه‌ی مد متغیر (VMD) به مؤلفه‌های سازنده (IMFها) تفکیک و سپس به مدل‌ها معرفی گردیدند. روش VMD مبتنی بر تجزیه‌ی سری داده‌ها به مؤلفه‌های مد ذاتی (IMF) و یک سری باقیمانده است، به‌طوری‌که جمع این مؤلفه‌ها و باقیمانده باید بتواند سری اصلی را بازسازی کند (موسی زاده و همکاران، ۱۴۰۲). اساس تشکیل IMFها بر عدم تداخل فرکانس‌های مرکزی در سطوح مختلف تجزیه استوار است.

نتایج محاسبه شاخص‌های آماری برای مدل‌های تلفیقی در جدول ۴ ارائه شده‌است. با توجه به این جدول مشاهده می‌شود که با الگوهای ورودی S3M1 تا S3M4 میانگین شاخص RMSE در مرحله آزمون برای مدل RF برابر با $110/124$ متر مکعب بر ثانیه می‌باشد. این در حالی است که با الگوهای ورودی سناریوهای اول و دوم مقدار میانگین این شاخص برای مدل RF به‌ترتیب برابر با $187/902$ و $162/992$ متر مکعب بر ثانیه می‌باشد. به‌عبارت دیگر روش تلفیقی VMD-RF توانسته است مقدار شاخص RMSE را به‌طور متوسط ۳۸ درصد کاهش دهد. این کاهش خطا باید متعاقباً در شاخص دقت

جدول ۴ - نتایج شاخص‌های ارزیابی RMSE، MAE، R و NS برای مدل‌های VMD-SVM و VMD-RF

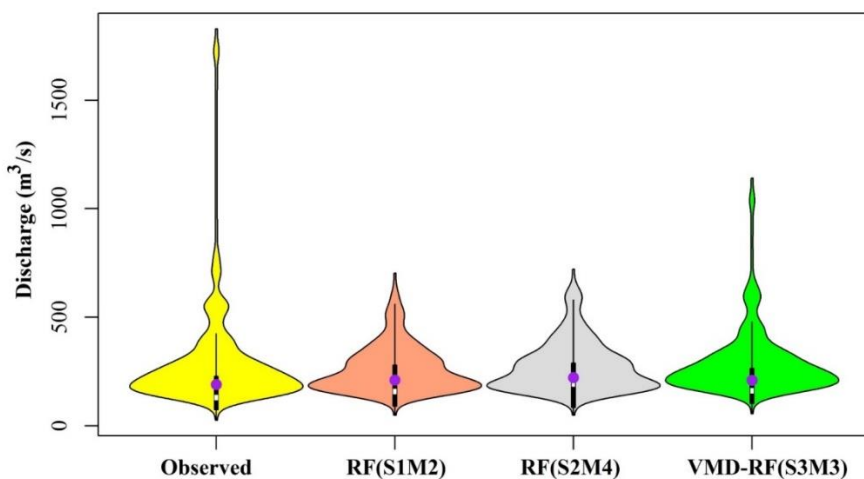
مدل	مرحله آموزش					مرحله آزمون				
	الگو	RMSE (m ³ /s)	MAE (m ³ /s)	R	NS	الگو	RMSE (m ³ /s)	MAE (m ³ /s)	R	NS
RF	S3M1	۵۶/۱۲۱	۳۶/۹۸۲	-۰/۹۷۶	-۰/۹۴۳	S3M1	۱۳۴/۲۵۱	۸۱/۸۸۴	-۰/۷۸۲	-۰/۵۹۳
	S3M2	۴۱/۹۸۳	۲۷/۳۹۱	-۰/۹۹۰	-۰/۹۶۸	S3M2	۱۰۳/۳۶۷	۶۹/۲۲۱	-۰/۸۸۸	-۰/۷۵۹
	S3M3	۳۹/۱۹۴	۲۵/۱۲۰	-۰/۹۹۲	-۰/۹۷۲	S3M3	۹۷/۸۹۹	۶۱/۸۵۷	+۰/۹۰۴	+۰/۷۸۴
	S3M4	۳۷/۹۹۵	۲۳/۸۳۱	-۰/۹۹۳	-۰/۹۷۴	S3M4	۱۰۴/۹۷۸	۶۷/۴۷۹	-۰/۸۹۳	-۰/۷۵۱
SVM	S3M1	۱۳۱/۹۴۲	۹۰/۵۰۳	-۰/۸۲۸	-۰/۶۸۵	S3M1	۱۱۸/۶۹۹	۷۳/۲۱۱	-۰/۸۲۷	-۰/۶۸۲
	S3M2	۲۰/۵۰۸	۱۵/۴۲۴	-۰/۹۹۶	-۰/۹۹۲	S3M2	۲۷/۲۲۶	۱۸/۲۸۵	-۰/۹۹۲	-۰/۹۸۳
	S3M3	۱۷/۲۳۰	۱۲/۵۹۰	-۰/۹۹۷	-۰/۹۹۵	S3M3	۲۱/۴۴۱	۱۵/۲۱۴	+۰/۹۹۵	+۰/۹۹۰
	S3M4	۱۰/۵۶۹	۸/۱۳۸	-۰/۹۹۹	-۰/۹۹۸	S3M4	۲۲/۲۴۱	۱۴/۵۷۰	-۰/۹۹۵	-۰/۹۸۹

است.

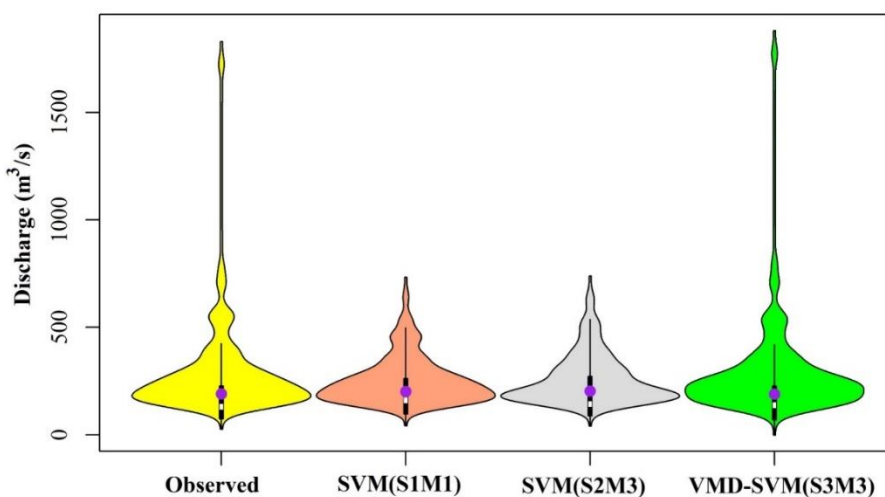
در شکل‌های ۵ و ۶ نمودارهای پراکندگی و ویولونی حاصل از مدل‌های RF، SVM، VMD-RF و VMD-GPR با بهترین الگوهای ورودی در هر سه سناریو ارائه شده است. با مقایسه این شکل‌ها می‌توان مشاهده نمود که مدل RF در همه سناریوها نتوانسته است مقادیر حدی بزرگ را به خوبی برآورد نماید. این مشکل در سناریوهای اول و دوم نسبت به سناریوی سوم حادتر می‌باشد. علاوه بر این توزیع داده‌های کمینه و متوسط نیز در سناریوی سوم که داده‌های وردی پس از تجزیه با روش VMD به مدل RF معرفی شده‌اند بهبود قابل توجهی یافته و شباهت انتهای ویولون مدل VMD-RF(S3M3) با مقادیر مشاهداتی افزایش یافته است. همچنین نکته دیگری نیز که باید مورد توجه قرار گیرد تأثیر پیش پردازش داده‌ها در کاهش بیش برآزش مدل RF می‌باشد. به طوری که در سناریوهای اول و دوم اختلاف شاخص RMSE در مراحل اول و دوم به طور میانین به ترتیب برابر با ۱۱۲/۱۳ و ۱۰۴/۲۸ مترمکعب بر ثانیه بود و این مقدار در سناریوی سوم مدل مذکور با بیش از ۴۸ درصد کاهش به ۶۶/۳۰ مترمکعب بر ثانیه رسیده است.

در شکل ۶ نمودار ویولونی برای مدل SVM در بهترین حالت هر سناریو ارائه شده است. با توجه به این شکل مشاهده می‌شود که بهبود در عملکرد مدل SVM همانند مدل RF بوده و پیش پردازش داده‌ها باعث بهبود عملکرد آن شده است. به طوری که برآورد مقادیر حدی به طور قابل ملاحظه‌ای بهبود یافته اما قدری از مقدار مشاهداتی بیشتر برآورد شده است. همچنین همچون مدل RF بیش برآزش در مدل SVM نیز با پیش پردازش داده‌ها به طور قابل قبولی کاهش داشته است.

جدول ۴ نکته مهم دیگری را نیز نشان می‌دهد و آن عملکرد برتر مدل VMD-SVM در مقابل VMD-RF می‌باشد. مدل تلفیقی SVM در مقایسه با روش تلفیقی RF توانسته است با خطایی کمتر بهینه‌ترین مقادیر را برای جریان ماهانه رودخانه برآورد نماید. روش VMD-RF با الگوی ورودی S3M3 به حداکثر دقت رسیده و براساس شاخص RMSE، ۵۹/۸۷۶ متر مکعب بر ثانیه خطا را کاهش داده است. این در حالی است که مدل VMD-SVM با الگوی S3M3 به بهترین عملکرد خود دست یافته و شاخص RMSE را در الگوی فوق از ۱۵۴/۸۰۷ متر مکعب در ثانیه به ۲۱/۴۴۱ متر مکعب در ثانیه (بیش از ۱۳۳/۳۶۶ متر مکعب بر ثانیه بهبود خطا) کاهش داده است. استفاده از شاخص‌های عددی نمی‌تواند اطلاعات لازم در خصوص توزیع داده‌ها را در اختیار کاربران قرار دهد. از این رو استفاده از نمودارهایی که بتواند توزیع داده‌ها را نشان داده و تفاوت مدل‌ها را نشان دهد، بسیار حائز اهمیت است. از جمله این نمودارها می‌توان به نمودار جعبه‌ای، نمودار پراکندگی و نمودار ویولونی اشاره نمود. نمودار ویولونی یک ابزار مصورسازی داده‌هاست که با ترکیب توزیع چگالی داده‌ها و عناصر باکس پلات (مانند میانه و چارک‌ها)، نمایشی جامع از پراکندگی داده‌ها ارائه می‌دهد. برخلاف نمودارهای رایجی مانند هیستوگرام یا باکس پلات که تنها چارک‌ها یا میانگین را نشان می‌دهند، نمودار ویولونی با استفاده از انتهای متقارن (شبهه ویولون) چگالی داده‌ها در نقاط مختلف را آشکار می‌کند و بینشی از شکل توزیع (نرمال، دوگانه یا کج) ارائه می‌دهد. برتری آن نسبت به باکس پلات در نمایش جزئیات بیشتر است؛ برای مثال، در داده‌های چندوجهی (مثلاً دو قله‌ای) یا مناطق پرتراکم را نشان می‌دهد، در حالی که نمودار جعبه‌ای ممکن است این جزئیات را نادیده بگیرد. در نمودار ویولونی، عرض هر بخش نشان‌دهنده تراکم داده‌ها در آن محدوده است و خطوط داخلی معمولاً نشانگر میانه یا بازه‌های چارکی هستند. این نمودار به طور همزمان خلاصه آماری و الگوی توزیع را در یک نمایش می‌دهد و برای مقایسه گروه‌های مختلف داده بسیار کارآمد



شکل ۵- نمودار ویولونی مدل RF برای سناریوهای مختلف در بهینه‌ترین حالت الگوهای ورودی



شکل ۶- نمودار ویولونی مدل SVM برای سناریوهای مختلف در بهینه‌ترین حالت الگوهای ورودی

نتیجه‌گیری

در این پژوهش، نتایج نشان داد که افزودن ترم تناوبی (پریودیک) به الگوهای ورودی مدل‌های منفرد (همچون SVM و RF) می‌تواند دقت برآورد جریان را نسبت به حالت ساده (بدون این ترم) بهبود بخشد. این ترم تناوبی بر اساس یک رابطه ریاضی ساده یا شماره ماه‌ها محاسبه شد و بدون نیاز به داده‌های اضافی (مانند بارش، دما یا تبخیر)، زمان و هزینه جمع‌آوری اطلاعات را برای پژوهش کاهش داد. در حالی که بسیاری از مطالعات برای افزایش دقت مدل‌ها، به داده‌های هیدرولوژیکی پیچیده‌تر متکی هستند، این پژوهش با بهره‌گیری از ترکیب ترم تناوبی خطی و غیرخطی در ورودی‌ها، همراه با سایر پارامترهای موجود، عملکرد مدل‌ها را ارتقا داد. این رویکرد نه تنها سادگی محاسباتی را حفظ کرد، بلکه نشان داد که ادغام ترم‌های تناوبی خطی و غیرخطی می‌تواند تا حدودی جایگزین مناسبی برای داده‌های مازاد باشد.

در این پژوهش، عملکرد مدل‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) در برآورد جریان ماهانه رودخانه، با تلفیق روش پیش پردازش تجزیه مُد متغیر (VMD) مورد بررسی قرار گرفت. هدف، مقایسه کارایی این مدل‌ها در دو حالت منفرد و هیبریدی بود. نتایج به‌دست آمده نشان دهنده تفاوت‌های معنادار در دقت پیش‌بینی بین دو حالت یادشده است که در ادامه به شرح زیر ارائه می‌گردد:

نتایج به‌دست آمده نشان داد که مدل SVM با استفاده از تابع کرنل RBF نسبت به مدل RF، با سناریوها و الگوهای ورودی مختلف در پیش‌بینی جریان ماهانه رودخانه بهترین عملکرد را به خود اختصاص داده است.

- Abda, Z., Zerouali, B., Chettih, M., Guimaraes Santos, C. A., de Farias, C. A. S., & Elbleltagi, A. 2022. Assessing machine learning models for streamflow estimation: a case study in Oued Sebaou watershed (Northern Algeria). *Hydrological Sciences Journal*. 67(9): 1328-1341.
- Ahmadi, F., Mehdizadeh, S., & Nourani, V. 2022. Improving the performance of random forest for estimating monthly reservoir inflow via complete ensemble empirical mode decomposition and wavelet analysis. *Stochastic Environmental Research and Risk Assessment*. 36(9): 2753-2768.
- Bishop, C. M., & Nasrabadi, N. M. 2006. *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: Springer.
- Breiman, L. 2001. Random forests. *Machine Learning*. 45(1): 5-32.
- Chollet, F. 2021. *Deep Learning with Python* (2nd ed.). Manning Publications.
- Choubey, V., Mishra, S., & Pandey, S. K. 2014. Time series data mining in real time surface runoff forecasting through Support Vector Machine. *International Journal of Computer Applications*. 98(3):274-291.
- Cutler, A., Cutler, D.R., & Stevens, J.R. 2012. Random forests. Pp. 157-175, In: Ensemble.
- Doyle, J. M., Hill, R. A., Leibowitz, S. G., & Ebersole, J. L. 2023. Random forest models to estimate bankfull and low flow channel widths and depths across the conterminous United States. *JAWRA Journal of the American Water Resources Association*. 59(5): 1099-1114.
- Dragomiretskiy, K., & Zosso, D. 2013. Variational mode decomposition. *IEEE transactions on signal processing*. 62(3): 531-544.
- Essam, Y., Huang, Y. F., Ng, J. L., Birima, A. H., Ahmed, A. N., & El-Shafie, A. 2022. Predicting streamflow in Peninsular Malaysia using support vector machine and deep learning algorithms. *Scientific Reports*. 12(1): 38-83.
- Fang, W., Huang, S., Ren, K., Huang, Q., Huang, G., Cheng, G., & Li, K. 2019. Examining the applicability of different sampling techniques in the development of decomposition-based streamflow forecasting models. *Journal of Hydrology*. 568: 534-550.
- Fletcher, T. 2009. Support vector machines explained. *Tutorial paper*. 1118: 1-19.
- Hamel, L. 2009 *Knowledge Discovery with Support Vector Machines*. Hoboken, N.J. John Wiley.
- He, X., Luo, J., Zuo, G., & Xie, J. 2019. Daily runoff forecasting using a hybrid model based on

در این مطالعه، علاوه بر روش‌های ساده، از یک روش ریاضی پیشرفته مبتنی بر تجزیه سیگنال به نام تجزیه مد متغیر (VMD) برای بهبود عملکرد مدل‌های جنگل تصادفی (RF) و ماشین بردار پشتیبان (SVM) استفاده شد. نتایج نشان داد که به کارگیری این روش، خطای مدل‌های منفرد RF و SVM را به ترتیب ۳۸ و ۷۲ درصد کاهش داده و توانایی بالایی در پیش بینی داده‌هایی با میانگین و واریانس داده‌های مشاهده‌ای دارد. در این میان، مدل هیبریدی VMD-SVM با الگوی ورودی S3M3، بهترین عملکرد را نشان داد و مقادیر شاخص‌های ارزیابی RMSE، MAE و NS در مرحله آزمون به ترتیب ۲۱/۴۴۱ متر مکعب در ثانیه، $۱۵/۲۱۴ \text{ (m}^3/\text{s)}$ و ۰/۹۹۰ محاسبه شد. این نتایج، اثربخشی روش VMD در افزایش دقت پیش‌بینی و کاهش خطاهای سیستماتیک مدل‌های هوشمند را تأیید می‌کند.

منابع

- احمدی، ف. ۱۳۹۹. ارزیابی عملکرد روش‌های ماشین بردار پشتیبان و سیستم استنتاج عصبی فازی تطبیقی در پیش‌بینی جریان ماهانه رودخانه‌ها (مطالعه موردی: رودخانه‌های نازلو و سزار). *تحقیقات آب و خاک ایران*. ۵۱(۳): ۶۸۶-۶۷۳.
- احمدی، ف. ۱۴۰۰. بررسی کارایی روش‌های پیش‌پردازش داده‌ها در بهبود عملکرد مدل برنامه ریزی بیان ژن (مطالعه موردی: رودخانه آب زال). *آب و خاک*. ۳۵(۲): ۱۵۳-۱۶۵.
- باقرپور، م، سیدیان، س.م، فتح‌آبادی، ا، محمدی، ا. ۱۳۹۷. پیش‌بینی رواناب با استفاده از مدل‌های جعبه سیاه و خاکستری. *تحقیقات منابع آب ایران*. ۱۴(۵): ۲۱۹-۲۰۴.
- خرمی‌پور، ز، ولیخان انارکی، م، فرزین، س. ۱۴۰۳. رویکردی جدید در روندیابی سیلاب بر مبنای تلفیق تئوری بیض، ماشین بردار پشتیبان و الگوریتم بهینه‌سازی فراابتکاری. *نشریه آبیاری و زهکشی ایران*. ۱۸(۳): ۴۲۰-۴۰۹.
- دهقانی، ر، ترابی پوده، ح، یونسی، ح.ا، شاهی نژاد، ب. ۱۴۰۰. کاربرد مدل هیبریدی ماشین بردار پشتیبان - موجک در تخمین جریان رودخانه‌ها (مطالعه موردی: حوضه دز). *نشریه علمی - پژوهشی مهندسی و مدیریت آب‌خیز*. ۱۳(۱): ۹۸-۱۱۰.
- موسی زاده، ن، آخوندعلی، ع.م، احمدی، ف. ۱۴۰۲. توسعه مدل‌های تلفیقی یادگیری ماشین مبتنی بر روش تجزیه مد تجربی گروهی کامل در برآورد جریان ورودی به سد (مطالعه موردی: سد دز). *نشریه آبیاری و زهکشی ایران*. ۴(۱۷): ۷۰۱-۶۸۵.

- through SVM-based probability density estimation. *IEEE Transactions on Geoscience and Remote Sensing*. 43(3): 559-570.
- Meng, E., Huang, S., Huang, Q., Fang, W., Wang, H., Leng, G., & Liang, H. 2021. A hybrid VMD-SVM model for practical streamflow prediction using an innovative input selection framework. *Water Resources Management*. 35: 1321-1337.
- Sahoo, S. K., & Satapathy, D. P. 2025. An improved support vector machine model for groundwater level prediction: a case study. *Earth Science Informatics*. 18(1): 1-18.
- Samantaray, S., & Sahoo, A. 2024. Prediction of flow discharge in Mahanadi River Basin, India, based on novel hybrid SVM approaches. *Environment, Development and Sustainability*. 26(7): 18699-18723.
- Samantaray, S., Das, S. S., Sahoo, A., & Satapathy, D. P. 2022. Monthly runoff prediction at Baitarani River basin by support vector machine based on Salp swarm algorithm. *Ain Shams Engineering Journal*. 13(5): 101732.
- Sharma, B., & Goel, N. K. 2024. Streamflow prediction using support vector regression machine learning model for Tehri Dam. *Applied Water Science*. 14(5): 1-20.
- Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. 2011. A complete ensemble empirical mode decomposition with adaptive noise. In 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 4144-4147). IEEE.
- Vapnik, V. 1995. *The nature of statistical learning theory* Springer. New York. 10: 978-1.
- Vapnik, V. 1998. *The support vector method of function estimation*. In *Nonlinear modeling: Advanced black-box techniques*. Boston, MA: Springer US.
- Wu, Z., & Huang, N. E. 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*. 1(01): 1-41.
- variational mode decomposition and deep neural networks. *Water resources management*. 33(4): 1571-1590.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.C., Tung, C.C., & Liu, H. H. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*. 454(1971): 903-995.
- Idemudia, O., Ehiorobo, J. O., Izinyon, C. O., & Ilaboya, I. 2024. Evaluating the performance of random forest, decision tree, support vector regression and gradient boosting for streamflow prediction. *CTU Journal of Innovation and Sustainable Development*. 16(2): 116-130.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. 2021. *An Introduction to Statistical Learning*. Springer. 610p.
- Lahmiri, S. 2015. Long memory in international financial markets trends and short movements during 2008 financial crisis based on variational mode decomposition and detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*. 437: 130-138.
- Latifoğlu, L. 2022. A novel approach for prediction of daily streamflow discharge data using correlation based feature selection and random forest method. *International Advanced Researches and Engineering Journal*. 6(1): 1-7.
- Li, B. J., Sun, G. L., Liu, Y., Wang, W. C., & Huang, X. D. 2022. Monthly Runoff Forecasting Using Variational Mode Decomposition Coupled with Gray Wolf Optimizer-Based Long Short-term Memory Neural Networks. *Water Resources Management*. 36(6): 2095-2115.
- Mahmood, O. A., Sulaiman, S. O., & Al-Jumeily, D. 2024. Forecasting for Haditha reservoir inflow in the West of Iraq using Support Vector Machine (SVM). *PloS one*. 19(9): e0308266.
- Mantero, P., Moser, G., & Serpico, S. B. 2005. Partially supervised classification of remote sensing images

Performance Evaluation of Machine Learning Models Based on the Variable Mode Decomposition (VMD) Method in River Flow Estimation (Case Study: Dez River)

N. Alivand¹, A. M. Akhoond-Ali², F. Ahmadi^{3*}
Received: Mar.31, 2025 Accepted: May.17, 2025

Abstract

In the present study, the monthly flow of the Dez River at the Taleh Zang station during the statistical period from 1971 to 2022 was modeled using a hybrid approach based on the Variational Mode Decomposition (VMD) signal decomposition method. The input data for the Random Forest (RF) and Support Vector Machine (SVM) models were defined based on three scenarios. In the first scenario, inputs included lagged flow data with delays of 1 to 4 months. In the second scenario, linear and nonlinear periodic terms were added to the lagged flow data as additional inputs. In the third scenario, the input data were decomposed into sub-series called Intrinsic Mode Functions (IMFs) using the VMD method before being fed into the models. The results revealed that each standalone model achieved maximum accuracy with different input patterns, but adding periodic terms moderately improved their performance. The SVM model in the second scenario showed the best performance, with a mean RMSE error of 156.83 m³/s. For the RF model, the mean error in the second scenario (162.99 m³/s) was lower than in the first scenario. In the third step, the data were decomposed using VMD, and modeling was performed with RF and SVM. Evaluation metrics indicated a significant reduction in error and improved accuracy in the hybrid models. Specifically, the VMD-SVM model reduced the RMSE by an average of 121 m³/s (about 71%) and outperformed the VMD-RF model in terms of precision.

Keywords: Signal Decomposition, Intrinsic Mode Functions (IMFs), Random Forests Model, Support Vector Machine

1 - Msc student in Water Resources Engineering, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

2 - Professor, Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

3 - Associate professor, Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

(* - Corresponding Author Email: f.ahmadi@scu.ac.ir)