



مرکز تحقیقات، آموزش و پژوهش کشاورزی و منابع طبیعی یزد

پژوهش‌های آب‌نخرداری

شاپا: ۲۰۳۸-۲۹۸۱



مجله تحقیقات، آموزش و پژوهش کشاورزی

ارزیابی عملکرد مدل‌های یادگیری ماشین و دستورالعمل‌های بهینه‌سازی برای پیش‌بینی آب‌دهی رود کشکان

فاطمه عوض‌پور^۱، محمدرضا هادیان^{۲*}، علی طالبی^۳

۱- دانشجوی دکتری، دانشکده مهندسی عمران، دانشگاه یزد، یزد، ایران

۲- استادیار دانشکده مهندسی عمران، دانشگاه یزد، یزد، ایران

۳- استاد دانشکده منابع طبیعی و کوبرشناسی، دانشگاه یزد، یزد، ایران

چکیده مبسوط

مقدمه و هدف

شبیه‌سازی آب‌دهی رود در ایستگاه‌های آب‌سنجی به‌منظور آگاهی از آب‌دهی رود در دوره‌های زمانی آینده از مسائل مهمی است که معمولاً به‌وسیله سری‌های زمانی آب‌شناختی مرتبط با همان ایستگاه انجام می‌شود. به‌منظور پیش‌بینی آب‌دهی رود با بیشترین دقت از سه گروه بزرگ شامل روش‌های تجربی و آماری، مفهومی و فرایند-محور استفاده می‌شود. از جمله روش‌های داده-محور، روش‌های مبتنی بر پایه هوش مصنوعی هستند. هدف این پژوهش، بررسی عملکرد مدل‌های یادگیری ماشین شامل: SVM و ANFIS، ANN و بررسی عملکرد مدل آموزش داده‌شده شبکه عصبی با دستورالعمل‌های بهینه‌سازی نهنگ (WOA) و دستورالعمل ازدحام ذرات (PSO) در پیش‌بینی آب‌دهی رود بود. افزون بر ارزیابی مدل‌های هوشمند، تأثیر استفاده از دستورالعمل‌های بهینه‌سازی بر دقت پیش‌بینی آب‌دهی رود بررسی شد. از آنجایی که برای استفاده از مدل‌های داده-محور، داده‌های ورودی تأثیر قابل توجهی بر عملکرد مدل‌ها دارند، از این‌رو سنج‌های مؤثر بر آب‌دهی رود مشخص شد و بهترین ترکیب متغیرهای ورودی برای هر مدل تعیین شد.

مواد و روش‌ها

در این پژوهش، به‌منظور پیش‌بینی آب‌دهی روزانه در ایستگاه آب‌سنجی پل دختر واقع بر رود کشکان، داده‌های آب‌دهی رود و بارش مربوط به سال‌های ۱۳۹۷-۱۳۵۰ تهیه شد.

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: mr_hadian@yazd.ac.ir

استناد: عوض‌پور، ف.، هادیان، م.ر.، طالبی، ع. ۱۴۰۴. ارزیابی عملکرد مدل‌های یادگیری ماشین و دستورالعمل‌های بهینه‌سازی در پیش‌بینی آب‌دهی رود کشکان. پژوهش‌های آب‌نخرداری، ۳۸(۱): ۱۸-۱.

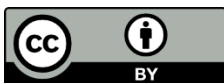
شناسه دیجیتال: 10.22092/WMRJ.2024.365128.1579

تاریخ دریافت: ۱۴۰۲/۱۲/۱۶، تاریخ بازنگری: ۱۴۰۳/۰۲/۲۰، تاریخ پذیرش: ۱۴۰۳/۰۳/۳۱، تاریخ انتشار: ۱۴۰۴/۰۱/۰۱

پژوهش‌های آب‌نخرداری، سال ۱۴۰۴، دوره ۳۸، شماره ۱، شماره پیاپی ۱۴۶، بهار ۱۴۰۴، صفحه‌های ۱ تا ۱۸.

© نویسندگان

ناشر: مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان فارس



از مدل‌های هوشمند ساختار استنتاج فازی عصبی (ANFIS) و ساختار بردار پشتیبان (SVM)، شبکه عصبی مصنوعی (ANN)، مدل ترکیبی شبکه عصبی و دستورالعمل ازدحام ذرات (ANN-PSO) و مدل ترکیبی شبکه عصبی و دستورالعمل نهنگ (ANN-WOA) استفاده شد.

در دو مدل ترکیبی تلاش شد که سنج‌های شبکه عصبی با استفاده از دستورالعمل‌های فراابتکاری تنظیم شوند و تأثیر آن بر عملکرد مدل ANN بررسی شود. همچنین، در این پژوهش، تأثیر داده‌های آب‌دهی رود و بارش همراه با تأخیرهای زمانی (داده‌های مربوط به روزهای گذشته) و ترکیبی از این سنج‌ها به‌عنوان ورودی مدل‌ها بررسی شد. به‌منظور تعیین بهترین ترکیب متغیرهای ورودی، روش‌های آماری تابع همبستگی خودکار (ACF) و تابع همبستگی خودکار جزئی (PACF) و ضریب همبستگی پیرسون (PCC) به‌کارگرفته شد. پس از اعمال ورودی‌های مؤثر و آموزش مدل‌های هوشمند نامبرده، با مقایسه اندازه‌های $RMSE$ ، R^2 و NE ، عملکرد آن‌ها در پیش‌بینی آب‌دهی رود بررسی شد.

نتایج و بحث

نتایج بررسی همه مدل‌ها در این پژوهش نشان داد که آب‌دهی یک (Q-1)، دو (Q-2) و سه (Q-3) روز گذشته و بارش یک روز گذشته (P-1)، بیش‌ترین همبستگی را با آب‌دهی روزانه رود نشان داد. به‌طور کلی، برای مدل کردن آب‌دهی رود در آبخیز کشکان دقت همه مدل‌ها قابل قبول بود. بر اساس نتایج به‌دست آمده، بیشترین دقت برای پیش‌بینی آب‌دهی روزانه جریان مربوط به مدل ANN-WOA با بیش‌ترین اندازه ضریب تبیین ($R^2 = 0/896$) و ضریب نش-ساتکلیف ($NE = 0/803$) و کمترین اندازه خطا ($RMSE = 0/0186$)، بود. پس از آن، مدل SVM با ساختار تابع کرنل پایه شعاعی و اندازه‌های $C=4$ ، $\gamma=1$ و $\epsilon=0/001$ با اندازه ضریب تبیین ($R^2 = 0/895$)، ضریب نش-ساتکلیف ($0/801$) و اندازه خطا ($NE = 0/0187$)، عملکرد بهتری نشان داد و مدل‌های ANN-PSO و ANN نیز به‌ترتیب در رده‌های سوم و چهارم بودند. نتایج بیانگر آن بود که استفاده از دستورالعمل‌های بهینه‌سازی فراابتکاری دقت مدل ANN را افزایش داد و می‌توان از آن برای آموزش شبکه استفاده کرد. بررسی ساختارهای گوناگون ANFIS نشان داد برای مدل‌سازی آب‌دهی رود منطقه مطالعه‌شده عملکرد توابع مثلثی و گوسی بیشتر بود. از سوی دیگر، خطای این مدل با اندازه‌های $RMSE=0/023$ و $NE=0/076$ در مقایسه با دیگر مدل‌ها بیشتر بود.

نتیجه‌گیری و پیشنهادها

نتایج این پژوهش نشان داد برای پیش‌بینی آب‌دهی رود دقت مدل‌های یادگیری ماشین مانند SVM، ANFIS و ANN قابل قبول بود. تنظیم سنج‌های شبکه عصبی با استفاده از دستورالعمل‌های بهینه‌سازی مانند WOA و PSO تأثیر بسزایی در بهبود عملکرد این مدل داشت. سرانجام می‌توان گفت این مدل‌ها می‌توانند جایگزین مناسبی برای مدل‌های مفهومی و آب‌شناختی در حل مسائل آب‌شناختی و آب‌دهی باشند. پیشنهاد می‌شود دو مدل SVM و ANFIS با استفاده از دستورالعمل PSO و WOA آموزش داده شود و سپس نتایج با یافته‌های این پژوهش مقایسه شود.

واژگان کلیدی: آب‌دهی رود، آبخیز کشکان، بارش-رواناب، دستورالعمل ازدحام ذرات، دستورالعمل نهنگ، هوش مصنوعی

مقدمه

رشد روزافزون جمعیت و محدودیت منابع آب سطحی کشور، اهمیت مطالعه، برنامه‌ریزی و مدیریت منابع آبی را دوچندان کرده است. متخصصان همواره به بررسی و پیش‌بینی آب‌دهی رود به‌علت اهمیت آن برای برنامه‌ریزی طرح‌های آبی، ساماندهی رودها، سامانه هشدار سیل و به‌ویژه برنامه‌ریزی برای بهره‌برداری بهینه، توجه داشته‌اند (دیتهاکیت و همکاران ۲۰۲۳). به‌منظور پیش‌بینی آب‌دهی رود با بیشترین دقت روش‌های گوناگونی پیشنهاد شده است که شامل سه گروه بزرگ روش‌های تجربی و آماری، مفهومی و فرایند-محور هستند. از آنجایی که استفاده از مدل‌های مفهومی نیازمند داده و دانش کامل از فرآیندهای مؤثر بر پدیده است، همواره با مشکلات فراوانی همراه بوده است (لطیف ۲۰۲۳). از سوی دیگر، روش‌های پیش‌بینی داده-محور به‌علت صرفه‌جویی زمان، نیاز به کمترین اطلاعات لازم و سهولت اجرا رو به گسترش هستند و پژوهش‌گران به این روش‌ها توجه دارند (اسدآبادی و همکاران ۲۰۱۴). از جمله روش‌های مبتنی بر داده که کاربرد گسترده‌ای در علوم گوناگون دارد، می‌توان به روش‌های مبتنی بر پایه هوش مصنوعی اشاره کرد. در این روش‌ها، با استفاده از هوش مصنوعی، روابط پیچیده میان متغیرها شناسایی می‌شود و سپس مدل‌سازی انجام می‌شود (نیه و همکاران ۲۰۲۳). تاکنون پژوهش‌های گوناگونی در زمینه کاربرد هوش مصنوعی برای پیش‌بینی سری‌های زمانی آب‌شناختی انجام شده است. روش شبکه عصبی (ANN)، ساختار استنتاج فازی عصبی تطبیقی (ANFIS) و ماشین بردار پشتیبان (SVM) از جمله روش‌هایی هستند که در الگوبندی با در نظر گرفتن عامل‌های مؤثر و مدل‌سازی روابط میان ورودی‌های گوناگون، خروجی مناسبی دارند و می‌توان برای پیش‌بینی فرایندهای آب‌شناختی از آنها استفاده کرد. همچنین، استفاده از دستورالعمل‌های بهینه‌سازی متفاوت به‌منظور بهبود عملکرد مدل‌های هوش مصنوعی مورد توجه پژوهش‌گران بوده است. پژوهش‌های پیشین نشان می‌دهد به‌کار بردن دستورالعمل‌های فراابتکاری برای آموزش مدل‌های

هوش مصنوعی مؤثر است و موجب افزایش کارایی و بهبود عملکرد آن‌ها در مقایسه با آموزش سنتی این مدل‌ها می‌شود. این دستورالعمل‌ها، با الهام‌گیری از طبیعت، مسائل پیچیده بهینه‌سازی را به‌شکل تصادفی اما هدفمند و با تقلید از پدیده‌های زیستی یا فیزیکی حل می‌کنند و جواب بهینه مطلق را می‌یابند. پژوهش‌های گسترده‌ای در زمینه استفاده از هوش مصنوعی و دستورالعمل‌های فراابتکاری در بحث‌های آب‌شناختی انجام شده است.

غفاری و وفاخواه (۲۰۱۳) رواناب آبخیز حاجی‌قوشان را با استفاده از شبکه عصبی مصنوعی و شبکه فازی عصبی تطبیقی مدل‌سازی کردند و دریافتند که در فرایند پیش‌بینی بارش-رواناب دقت مدل ANFIS ($MAE=2/18$, $RMSE=7/11$ و $R^2=0/6$) در مقایسه با مدل ANN ($MAE=1/97$, $RMSE=6/03$) و $R^2=0/39$ بیشتر بود. جودی و همکاران (۲۰۱۵) با استفاده از مدل‌های آب‌شناختی SWAT و SVM، رواناب ماهانه رود ليقوان‌چای را مدل‌سازی کردند و نتیجه گرفتند که دقت مدل SWAT ($RMSE=0/41$) در مقایسه با مدل SVM بیشتر بود. عبداله‌پور و ستاری (۲۰۱۵) در استان آذربایجان شرقی آب‌دهی روزانه رود اهرچای را با استفاده از مدل‌های ANN و ANFIS پیش‌بینی کردند. نتایج این پژوهش نشان داد که دقت مدل ساختار استنتاج فازی عصبی تطبیقی ($RMSE=0/318$, $R^2=0/94$) در مقایسه با شبکه‌های عصبی مصنوعی بیشتر بود. در پژوهشی ناظری‌تهرودی و همکاران (۲۰۱۶) با به‌کارگیری مدل ANFIS در شبیه‌سازی آب‌دهی رود نازلوچای، دریافتند که خطای مدل در مقایسه با دو مدل برنامه‌نویسی ژنتیک (GP) و SVM به‌ترتیب ۲۳ و ۳٪ کاهش یافت. صمدی و فتح‌آبادی (۲۰۱۸) با استفاده از داده‌های ایستگاه قزاقلی واقع بر رود گرگان‌رود، اندازه آب‌دهی ورودی به سد گرگان با استفاده از دو مدل ANN و SVM پیش‌بینی کردند. نتایج این پژوهش نشان داد مدل SVM ($RMSE=5/31$) در مقایسه با مدل ANN ($RMSE=9/88$) برتر بود. امامی و

بارش) و آب‌شناختی (مانند آب‌دهی رود) استفاده می‌شود و داده‌های آب‌شناختی در مکان‌ها و زمان‌های گوناگون بسیار متفاوت هستند، بنابراین نمی‌توان این نتایج در یک آبخیز را برای تمام آبخیزها تعمیم داد. کشکان شاخه اصلی آبخیز کشکان است و از نظر تأمین منابع آبی در استان لرستان بسیار حائز اهمیت است. بنابراین، ارزیابی دقیق آن به منظور مدیریت بهتر منابع آبی لازم و ضروری است. از این رو، در این پژوهش افزون بر ارزیابی عملکرد مدل‌های ساختار استنتاج فازی (ANFIS) و ماشین بردار پشتیبان (SVM) برای پیش‌بینی آب‌دهی رود در ایستگاه پل دختر (در خروجی آبخیز کشکان)، عملکرد مدل شبکه عصبی (ANN) همراه با دستورالعمل‌های بهینه‌سازی شامل دستورالعمل ازدحام ذرات (PSO)، و مدل دستورالعمل نهنگ (WOA) بررسی شدند. به این ترتیب ارزیابی دقیقی از عملکرد مدل‌های گوناگون هوش مصنوعی و دستورالعمل‌های بهینه‌سازی انجام شد. سرانجام مدل بهینه بر اساس معیارهای RMSE، R^2 و NE تعیین شد. هدف این پژوهش، تعیین سنجه‌های ورودی بهینه و بررسی عملکرد روش‌های گوناگون محاسبه‌ای نرم و استفاده از دستورالعمل‌های فراابتکاری بهینه‌سازی در پیش‌بینی آب‌دهی رود در آبخیز کشکان بود.

مواد و روش‌ها

معرفی منطقه مطالعه شده

آبخیز کشکان در استان لرستان از زیرمجموعه‌های آبخیز بزرگ کرخه با مساحت کل ۹۵۲۴/۱ کیلومتر مربع، ۱۸/۵٪ از مساحت آبخیز بزرگ کرخه را شامل می‌شود (گزارش منابع آب‌های سطحی وزارت نیرو ۲۰۰۷). که میان مختصات جغرافیایی طول شرقی ۱۱' ۳۴° تا ۱' ۴۹° و طول شمالی ۵' ۳۳° تا ۵۶' ۳۴° است. از پیوستن هررود و آب‌الشر که از بلندی‌های کرین در غرب و جنوب‌غربی بروجرد سرچشمه می‌گیرند رود کشکان تشکیل می‌شود و سپس شاخه‌های فرعی و نسبتاً پر آب دیگری به آن اضافه می‌شود. این رود حدود ۶۰ کیلومتر در جهت غرب به مسیر خود ادامه می‌دهد و پس از ۴۵ کیلومتر به رود

همکاران (۲۰۱۹)، به منظور پیش‌بینی بار رسوب معلق رود زرینه‌رود از سه دستورالعمل فراابتکاری بهینه‌سازی شامل دستورالعمل ژنتیک، دستورالعمل گرگ خاکستری و دستورالعمل انتخابات استفاده کردند. نتایج این پژوهش نشان داد خطای دستورالعمل گرگ خاکستری ($NE=0/74$ ، $RMSE=0/022$) و $R^2=0/96$ ، در مقایسه با دیگر دستورالعمل‌ها کمتر بود. پریسوج و همکاران (۲۰۲۰) با استفاده از مدل‌های SVM، ANN و ماشین بهینه‌یادگیری (ELM) آب‌دهی روزانه و ماهانه رود چهار رود ایالت آمریکا را مدل‌سازی کردند. نتایج نشان‌دهنده دقت بیش‌تر مدل شبکه عصبی و مدل بردار پشتیبان در مقایسه با مدل ELM بود. در هند سامانتاری و ساهو (۲۰۲۳) به منظور توسعه مدل SVM، با استفاده از دستورالعمل‌های بهینه‌سازی نهنگ و ازدحام ذرات، غلظت رسوب رود ماهانادی را پیش‌بینی کردند و نتایج به‌دست آمده را با مدل SVM مقایسه کردند. بررسی نتایج نشان داد، دقت مدل SVM-WOA ($MSE=0/375$)، برای ایستگاه‌های Sundargarh و Salebhata در مقایسه با اندازه‌های مدل SVM-PSO ($MSE=1/005$ ، $MSE=1/014$) بیشتر بود. حکیمی و همکاران (۲۰۲۴)، برای پیش‌بینی آب‌دهی ورودی روزانه سد آبگیر سفیدرود، از ترکیب دستورالعمل‌های فراابتکاری شامل دستورالعمل‌های چرخه آب (WCA)، گرگ خاکستری (GOA)، نهنگ (WOA)، دستورالعمل شبکه عصبی (NNA)، و ملخ (GOA) برای آموزش ساختار عصبی-فازی و به‌روزرسانی سنجه‌های آن استفاده کردند. نتایج این پژوهش نشان داد مدل ترکیبی ANFIS-WCA با کمترین اندازه‌های $MAE=0/0736$ ، $SI=0/5048$ و $NRMSE=0/0736$ بیشترین اندازه $R^2=0/984$ در مقایسه با دیگر مدل‌ها برتر بود.

نتایج پژوهش‌ها بیانگر آن است که عملکرد روش‌های هوش مصنوعی برای پیش‌بینی فرایندهای آب‌شناختی خوب است. اما با توجه به اینکه عملکرد این مدل‌ها، متأثر از داده‌های ورودی است و با توجه به این مهم که در مباحث آب‌شناختی، از داده‌های هواشناسی (مانند

ایستگاه مطالعه‌شده پل‌دختر نام دارد که در انتهای‌ترین نقطهٔ آبخیز کشکان (نقطه تمرکز) است (شکل ۱). از این رو، دلیل بررسی جریان این ایستگاه در مهار و مدیریت جریان به‌ویژه در شرایط سیلابی حائز اهمیت است. مشخصات آبخیز پل‌دختر و ایستگاه پل‌دختر در جدول ۱ آورده شده است.

خرم‌آباد که عموماً جهت جریان آن از شرق به غرب است، می‌پیوندد و در جهت جنوب‌غربی ادامه می‌یابد. پس از ادغام شاخه‌های فرعی چلهول و مادیان‌رود با آن، سرانجام به سیمره پیوسته و رود کرخه را پدید می‌آورد. از میان رودهای استان لرستان، رود کشکان از سال ۱۳۳۴ تاکنون با ۱۷ سیلاب (بیش از ۱۰۰۰ مترمکعب بر ثانیه) سیل‌خیزترین رود این استان به‌شمار می‌آید.



شکل ۱- موقعیت جغرافیایی آبخیز کشکان در آبخیز بزرگ کرخه و ایستگاه پل‌دختر.

Figure 1- Geographical location of the Kashkan Watershed in the large Karkheh basin and Pol Dokhtar Station.

جدول ۱- مشخصات آبخیز پل‌دختر و ایستگاه آب‌سنجی پل‌دختر.

Table 1- Characteristics of the Pol Dokhtar Watershed and station.

| Area (km ²) | slope (degree) | Geographical location | | Q(m ³ /s) | | | P(mm) | | |
|-------------------------|----------------|-----------------------|----------|----------------------|------|------|-------|------|-----|
| | | longitude | latitude | min | mean | max | min | mean | max |
| 202.38 | 23.1 | 47-43-03 | 33-09-42 | 0.011 | 47.8 | 1546 | 0 | 1.06 | 115 |

خاک و (...), پستی‌بلندی آبخیز (شیب آبخیز، مساحت، و ...) کاربری زمین، پوشش گیاهی و ... به شکل مستقیم و غیرمستقیم بر این پدیده تأثیرگذارند. در این پژوهش، با توجه به محدودیت دسترسی به همه داده‌های مؤثر، برای پیش‌بینی جریان یک روز بعد، از داده‌های بارش

روش پژوهش
عامل‌های محیطی گوناگونی بر اندازهٔ آب‌دهی رود در یک آبخیز تأثیرگذار هستند. داده‌های هواشناسی (دمای هوا، رطوبت خاک، تبخیر، تعرق، بارش و ...)، مشخصات زمین‌شناسی (نوع خاک، اندازهٔ نفوذپذیری

سنجه‌های بخش مقدم با استفاده از روش گرادیان کاهشی محاسبه می‌شوند (سامانتاری و ساهو ۲۰۲۲). درباره روش حل ترکیبی اطلاعات کامل‌تر و جامع‌تر در پژوهش‌های جانگ (۱۹۹۳) و مارس (۲۰۰۴) ارائه شده است. برای سادگی کار اگر، فرض شود ساختار فازی دو ورودی x و y دارد و خروجی آن z است، قوانین به شکل زیر تعریف می‌شوند (جانگ ۱۹۹۳).

قانون (۱)
$$\text{if } (x \text{ is } A_1 \text{ and } y \text{ is } B_1) \\ \text{then } (f_1 = p_1x + q_1y + r_1)$$

قانون (۲)
$$\text{if } (x \text{ is } A_2 \text{ and } y \text{ is } B_2) \\ \text{then } (f_2 = p_2x + q_2y + r_2)$$

وایازی بردار پشتیبان (SVR)

ماشین‌های بردار پشتیبان یک ساختار یادگیری است که بر مبنای نظریه بهینه‌سازی مقید عمل می‌کند. در این نظریه از اصل استقرای کمینه‌سازی خطای ساختاری استفاده می‌شود و سرانجام یک جواب بهینه کلی پیشنهاد می‌شود (شاهی‌نژاد و دهقانی ۲۰۱۷). به‌طور کلی ماشین‌های بردار پشتیبان به دو دسته تقسیم می‌شوند. دسته اول که طبقه‌بندی‌کننده نام دارند داده‌ها را دسته‌بندی می‌کند و هدف آن تشخیص طبقه داده جدید است. دسته دوم که عمل وایازی و برازش تابع را انجام می‌دهد، وایازی بردار پشتیبان نامیده می‌شود که هدف آن پیش‌بینی داده است. ماشین‌های بردار پشتیبان با استفاده از فرضیه جداپذیر بودن طبقه‌ها از هم به شکل خطی، ابر صفحاتی را به دست می‌آورد که می‌تواند طبقه‌ها را از هم جدا کند. زمانی که جدا کردن خطی داده‌ها امکان‌پذیر نباشد، ماشین بردار پشتیبان با اعمال توابع (کرنل) داده‌ها را به فضای با ابعاد بیشتر نگاشت می‌دهد تا بتوان آن‌ها را در این فضای جدید به شکل خطی جدا کرد (چوبین و همکاران ۲۰۱۹).

در این مدل، سه سنجه C, γ, ϵ باید بهینه شوند. سنجه γ ممکن است منجر به بیش‌برازش یا کم‌برازش کمتر شود. در شرایطی که مدل نتواند برای داده‌های آزمایش پیش‌بینی دقیقی انجام دهد، برازش کمتر رخ می‌دهد که در این حالت، مدل داده‌های آموزش را به خاطر می‌سپارد و در داده‌های غیر از آموزش دچار خطا

(P) و آبدهی رود (Q) در مقیاس روزانه در بازه زمانی ۴۷ ساله (۱۳۹۷-۱۳۵۰) استفاده شد. این داده‌ها از سازمان‌های مربوطه تهیه شد. با توجه به کامل نبودن داده‌ها در بازه زمانی مطالعه‌شده، ابتدا پیش پردازش داده‌ها انجام شد. از این‌رو، پس از مرتب‌سازی داده‌ها، در شرایط نبودن گزارش داده رسوب یا آبدهی رود و یا وجود داده‌های پرت، ردیف مربوطه حذف شد. در گام بعد، با توجه به اثرپذیری جریان رود از سری زمانی اطلاعات نامبرده، اندازه مؤثر بودن تأخیر زمانی سنجه‌های P و Q مشخص شد و سرانجام به‌عنوان ورودی به مدل تعریف شدند. برای مرحله آموزش ۷۰٪ داده‌ها و برای مرحله آزمایش و صحت‌سنجی ۳۰٪ داده‌ها در نظر گرفته شدند.

شبکه عصبی فازی-استنتاجی تطبیقی (ANFIS)

ساختارهای عصبی-فازی استنتاجی به دست آمده ترکیب ساختارهای فازی و شبکه‌های عصبی مصنوعی هستند و از مزایای هر دو ساختار استفاده می‌کند (اسکندری و همکاران ۲۰۱۹). در این ساختار، سنجه‌های ساختار فازی با استفاده از دستورالعمل‌های آموزش شبکه‌های عصبی مصنوعی به دست می‌آیند. ساختار استنتاج فازی با استفاده از تعدادی قوانین اگر (بخش مقدم) و آنگاه (بخش تالی) یک رابطه جز به جز غیرخطی را از فضای ورودی به فضای خروجی ارائه می‌دهد (عبداله‌پور و همکاران ۲۰۱۵). روش آموزش در ANFIS، روش پس انتشار خطا است. در مقایسه با ANN، مدل ANFIS برای تعیین الگوی موجود میان داده‌ها از یک دستورالعمل ترکیبی استفاده می‌کند. به این شکل که برای بهینه‌سازی سنجه‌های مدل، ترکیب روش‌های گرادیان کاهشی و حداقل مربعات را به کار می‌گیرد. هر تکرار در این مدل شامل دو فاز پیش‌رو و پس‌رو می‌باشد. در فاز پیش‌رو سنجه‌های بخش مقدم ثابت است و سنجه‌های بخش تالی با استفاده از روش حداقل مربعات تعیین می‌شوند. در فاز پس‌رو، سنجه‌های بخش تالی ثابت است و سنجه‌های بخش مقدم با استفاده از روش حداقل مربعات تعیین می‌شوند. در فاز پس‌رو، سنجه‌های بخش تالی ثابت است و

پیرامون طعمه خود و حرکات مارپیچ، روی سطح دریا شکار می‌کنند. روند جستجوی دستوالعمل نهنگ شامل سه مرحله محاصره طعمه، مرحله بهره‌برداری (حمله به حباب تور) و مرحله اکتشاف (جستجو برای طعمه) است. در مرحله محاصره طعمه، از آنجایی که مکان شکار بهینه در فضای جستجو از گذشته مشخص نیست، دستوالعمل فرض می‌کند که بهترین راه حل کاندید حال حاضر، شکار هدف و یا نزدیک به حالت مطلوب است. بعد از اینکه بهترین عامل جستجو شناسایی شد، عامل‌های دیگر جستجو تلاش می‌کنند تا مکان خود را نسبت به بهترین عامل جستجو، به‌روزرسانی کنند. این رفتار نهنگ‌ها با استفاده از معادله‌های ۳ و ۴ تعریف شده است (معماری و همکاران ۲۰۲۲).

$$\vec{D} = |\vec{C} \vec{X}^*(t) - \vec{X}(t)| \quad (3)$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (4)$$

t: شماره تکرار، A و C: بردارهای ضریب‌ها، X^* : بردار مکان بهترین راه حل به دست آمده در حال حاضر و X: بردار مکان است. باید توجه داشت که در شرایط وجود راه حل بهتر، X^* در هر تکرار باید به‌روز شود. بردار A و C با استفاده از معادله‌های ۵ و ۶ محاسبه می‌شوند (معماری و همکاران ۲۰۲۲).

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (5)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (6)$$

a: به شکل خطی از اندازه ۲ به سمت اندازه صفر در تکرارها (در هر دو فاز اکتشاف و استخراج) کاهش می‌یابد و r بردار تصادفی در بازه ۰ تا ۱ است. در مرحله بهره‌برداری (حمله به حباب تور)، برای مدل‌سازی ریاضی رفتار حباب تور نهنگ‌های گوژپشت (حرکت انقباضی و مارپیچی)، دو روش طراحی شده است. روش انقباضی که در آن محدوده نوسان A به وسیله a کاهش می‌یابد. به بیان دیگر، A اندازه تصادفی در فاصله a تا -a است و a در تکرارها، از اندازه ۲ تا ۰ کاهش می‌یابد. با انتخاب اندازه‌های تصادفی A در فاصله ۱ تا -۱، می‌توان مکان جدید عامل جستجو را میان مکان اصلی عامل و مکان بهترین عامل کنونی، تعریف کرد. در مرحله حرکت مارپیچی، ابتدا فاصله میان نهنگ

می‌شود. سنجه C یک مهارکننده اندازه خطای پیش‌بینی است و اندازه بهینه E به نوع نويز موجود در داده‌ها بستگی دارد (یو و همکاران ۲۰۰۴، نورانی و علیزاده ۲۰۱۵). مراحل حل مسئله در دستوالعمل ماشین بردار پشتیبان همانند دیگر مدل‌های یادگیری ماشین به دو مرحله آموزش و آزمایش تقسیم‌بندی می‌شود.

شبکه عصبی مصنوعی پرسپترون چند لایه (ANN-MLP) یکی از معروف‌ترین شبکه‌های عصبی که در این پژوهش استفاده شد، شبکه عصبی MLP است. این شبکه از لایه‌های گوناگون شامل لایه ورودی، لایه پنهان و لایه خروجی تشکیل شده است (اسکندری و همکاران ۲۰۱۹). هر لایه تعدادی نورون دارد که وظایف خاصی دارند. در لایه ورودی، به تعداد سنجه‌های ورودی به مدل، نورون وجود دارد و هر کدام معرف یک سنجه تأثیرگذار بر خروجی است. عملیات پردازش داده‌ها در لایه پنهان انجام می‌شود و لایه خروجی شامل اندازه‌های پیش‌بینی شده به وسیله مدل است. مدل MLP از الگوی پس انتشار (BP) برای آموزش شبکه استفاده می‌کند. داده‌ها به وسیله لایه ورودی به مدل تعریف شده‌اند و پس از عبور از لایه پنهان و لایه خروجی، با مقایسه اندازه‌های پیش‌بینی شده به وسیله مدل با اندازه‌های واقعی، اندازه خطای مدل محاسبه می‌شود. این خطا به شکل پس انتشار در شبکه بازگشت داده می‌شود و سنجه‌های موجود میان نورون‌ها (شامل وزن‌ها و بایاس‌ها) که مهم‌ترین سنجه در فرایند آموزش شبکه هستند، دوباره تنظیم می‌شوند (فلامکی و همکاران ۲۰۱۳).

دستوالعمل بهینه‌سازی نهنگ

دستوالعمل فراابتکاری نهنگ اولین بار به وسیله میرجلیلی و همکاران ارائه شد (میرجلیلی و لویز ۲۰۱۶). این دستوالعمل یکی از دستوالعمل‌های بهینه‌سازی الهام‌گرفته شده از طبیعت و مبتنی بر جمعیت است که در آن از روش شکار نهنگ‌های گوژپشت یا کوهان‌دار به نام روش تغذیه شبکه حبابی استفاده شده است. نهنگ‌ها با ایجاد حباب‌هایی در

تعامل میان افراد یک گروه نیز بستگی دارد (شاه حسینی و همکاران ۲۰۱۱، محمدی و مؤذنزاده ۲۰۱۹). هر ذره با توجه به تجربه و دانش خود و همچنین دانش ذرات دیگر که در همسایگی آن است، کم کم به سمت جواب بهینه حرکت می‌کند. همکاری و حافظه مشترک میان ذرات در این دستورالعمل موجب می‌شود شانس بیشتری برای یافتن راه حل بهینه وجود داشته باشد و سریع‌تر راه حل‌ها با کیفیت مطلوب کشف می‌شوند (وانگ ۲۰۱۸). روش کار دستورالعمل ازدحام ذرات شامل سه مرحله تولید موقعیت و سرعت ذرات، به‌روزرسانی سرعت ذرات و به‌روزرسانی موقعیت ذرات است. به این شکل که ذرات با موقعیت‌ها و سرعت‌های اولیه به شکل تصادفی تولید می‌شوند و سپس ذرات در هر تکرار، بر اساس بهترین موقعیت خود و همسایگانش در گذشته، سرعت و موقعیت خود را به‌روزرسانی می‌کند و به هدف نزدیک‌تری می‌رسد. سرانجام پس از تکرارهای متوالی، به جواب بهینه همگرا می‌شود. با استفاده از رابطه‌های ریاضی ۹ و ۱۰ حرکت ذرات در این دستورالعمل تعیین می‌شود (وانگ و همکاران ۲۰۱۸).

$$V_i(t+1) = \omega V_i(t) + C_1 \times \text{rand}_1(pbest_i(t) - x_i(t)) + C_2 \times \text{rand}_2(gbest_i(t) - x_i(t)) \quad (9)$$

$$x_i(t+1) = x_i(t) + V_i(t+1) \quad (10)$$

بهنجارسازی داده‌ها

وارد کردن داده‌ها به شکل خام بر اندازه سرعت و دقت شبکه مؤثر بوده و موجب کاهش آن‌ها می‌شود. بهینه‌سازی داده‌ها روشی است که برای جلوگیری از بروز چنین شرایطی انجام می‌گیرد. این عمل به‌خصوص زمانی که دامنه تغییرات ورودی‌ها گسترده باشد، کمک قابل توجهی به آموزش بهتر و سریع‌تر مدل می‌کند. بنابراین در این پژوهش داده‌های ورودی با استفاده از رابطه ۱۱ در محدوده ۰ و ۱ بهینه شدند (بویوکیدیز و کامسو ۲۰۱۷).

$$X_{norm} = \frac{X_0 - X_{min}}{X_{max} - X_{min}} \quad (11)$$

قرارگرفته در مختصات فعلی (X, Y) و بهترین عامل جستجو (X^*, Y^*) با استفاده از رابطه ۷ محاسبه می‌شود (D') و سپس با استفاده از رابطه ۸ حرکت نهنگ‌ها به شکل مارپیچی شبیه‌سازی می‌شود (معماری و همکاران ۲۰۲۲).

$$\vec{D}' = |\vec{X}^*(t) - \vec{X}(t)| \quad (7)$$

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (8)$$

b : ثابتی برای تعریف شکل مارپیچ لگاریتمی و l عددی تصادفی میان ۱ تا ۱- است.

دستورالعمل بهینه‌سازی ازدحام ذرات

دستورالعمل فراابتکاری ازدحام ذرات برای بهینه‌سازی مسائل پیچیده و توابع پیوسته غیرخطی استفاده می‌شود. این دستورالعمل، از رفتار دسته جمعی پرندگان و ماهی‌ها در هنگام جستجوی غذا الهام می‌گیرد که در آن هر راه حل (در این پژوهش، وزن‌ها و بایاس‌های شبکه عصبی) که یک ذره نامیده می‌شود، معادل با یک پرند در حرکت دسته جمعی پرندگان فرض شده است و شایستگی آن به وسیله یک تابع هدف مشخص می‌شود. اساس دستورالعمل PSO نه تنها به رفتار فرد در اجتماع وابسته است، بلکه به چگونگی

p_{best} : بهترین موقعیتی که ذره تجربه تاکنون داشته است و g_{best} : بهترین موقعیتی است که به وسیله کل ذرات تشخیص داده شده است. T : نشان دهنده تعداد تکرار، $rand_1$ و $rand_2$: اعداد تصادفی در بازه ۱ و صفر، c : ثابت شتاب و ω : ضریب اینرسی (بازه صفر و ۱) است. ضریب c_1 : سنجه شناختی (تک ذره)، ضریب c_2 : سنجه اجتماعی (همه ذرات) است که نشان دهنده شیب حرکت در جستجوی محلی است. اندازه این ضریب‌ها میان صفر و ۲ است. در این پژوهش تعداد ذرات اولیه ۸ و تعداد تکرار ۱۰۰ در نظر گرفته شد.

به‌کارگرفته شد. این دو روش برای انتخاب ورودی‌ها در پژوهش‌های آب‌شناختی بسیار کاربرد دارند و هدف اصلی آن‌ها توسعه مدلی است که از داده‌های مؤثر ورودی با تأخیر زمانی به‌منظور پیش‌بینی زمان حال استفاده کند (مالیک و همکاران ۲۰۲۱).

معیارهای ارزیابی عملکرد مدل

به‌منظور ارزیابی تأثیر سنج‌های دو مدل ANFIS و SVM در بخش آموزش و صحت‌سنجی، از شاخص‌های ضریب تبیین (R^2)، جذر میانگین مربعات خطا (RMSE) و ضریب نش-سانکلیف (NE) استفاده شد. در این پژوهش با استفاده از معادله‌های ۱۲ تا ۱۴ محاسبه شدند. (موریاسی ۲۰۰۷، کتیپاگلو ۲۰۲۳)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_o - Q_s)^2} \quad (12)$$

$$R^2 = \frac{\sum_{i=1}^n (Q_o - \bar{Q}_o)(Q_s - \bar{Q}_s)}{\sqrt{\sum_{i=1}^n (Q_o - \bar{Q}_o)^2 \sum_{i=1}^n (Q_s - \bar{Q}_s)^2}} \quad (13)$$

$$NE = 1 - \left(\frac{\sum_{i=1}^n (Q_o - Q_s)^2}{\sum_{i=1}^n (Q_o - \bar{Q}_o)^2} \right) \quad (14)$$

این منظور دو سنج بارش روزانه و آب‌دهی رود روزهای گذشته به‌عنوان سنج‌های تأثیرگذار بر آب‌دهی روزانه انتخاب شدند. سپس، برای تعیین سنج‌های ورودی مؤثر در آب‌دهی رود و تأخیر زمانی از روش‌های ACF و PACF استفاده شد (شکل ۲ الف). همچنین، از ضریب همبستگی پیرسون برای تعیین سنج‌های ورودی مؤثر در آب‌دهی رود و بارش در ایستگاه پل‌دختر استفاده شد (شکل ۲ ب). نتایج این پژوهش نشان داد اگرچه میان متغیر هدف (آب‌دهی) و آب‌دهی و بارش روزهای گذشته تا حدودی همبستگی وجود داشت، اما بیش‌ترین همبستگی در پیش‌بینی آب‌دهی روزانه مربوط به آب‌دهی یک (Q-1)، دو (Q-2) و سه (Q-3) روز گذشته و بارش یک روز گذشته (P-1) بود. در نتیجه فضای بردار ورودی به مدل‌ها برای پیش‌بینی روز

X_{norm} : اندازه بهینه‌شده، X_0 : اندازه‌های مشاهده‌ای، X_{min} : کمترین اندازه و X_{max} : بیش‌ترین اندازه از مجموعه داده‌ها است.

انتخاب ورودی‌های مؤثر

از آنجایی که اطلاعات اساسی ساختاری که مدل‌سازی می‌شود به‌وسیله داده‌های ورودی مشخص می‌شوند و داده‌های ورودی مهم‌ترین سنج در تعیین ویژگی‌های یک ساختار هستند، از این‌رو، مسائل مهم در استفاده از روش‌های یادگیری ماشین و داده مینا، انتخاب سنج‌های ورودی بهینه است. در این پژوهش روش‌های آماری تابع همبستگی خودکار (ACF) و تابع همبستگی خودکار جزئی (PACF) و ضریب همبستگی پیرسون برای تعیین ترکیب بهترین متغیرهای ورودی

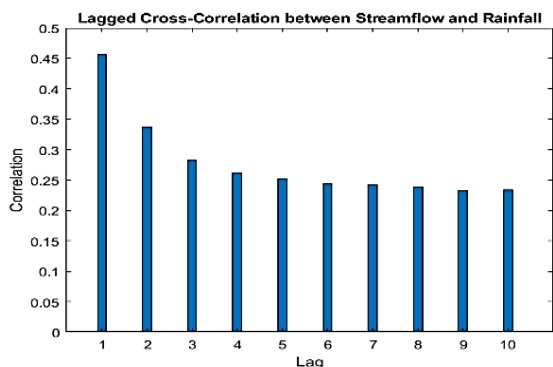
R^2 : ضریب تبیین و بیانگر ارتباط خطی میان داده‌های بزرگ مقیاس و کوچک مقیاس است و بدون بعد است. Q_o : اندازه‌های مشاهده‌ای، Q_s : اندازه‌های محاسبه‌شده به‌وسیله مدل، \bar{Q}_o : اندازه میانگین داده‌های مشاهده‌ای، N : تعداد نمونه‌ها است. هر چه اندازه R^2 به عدد ۱ و اندازه RMSE به عدد صفر نزدیکتر باشد، دقت مدل بیش‌تر است.

نتایج و بحث

انتخاب ورودی‌های اولیه مناسب و تأثیرگذار در پدیده به‌منظور آموزش ماهیت سازوکار حاکم بر پدیده باعث بهبود عملکرد مدل می‌شود. از این‌رو، در الگوسازی آب‌دهی رود کشکان نیز تلاش شد تا مؤثرترین سنج‌های ورودی برای آموزش مدل انتخاب شوند. به

شد که هر کدام شامل ترکیبات گوناگونی از ورودی‌های مؤثر بودند.

بعد حداکثر ۴ متغیر داشت. بر این اساس و بر پایه جدول ۲ شش سناریوی گوناگون برای مدل‌ها تعریف

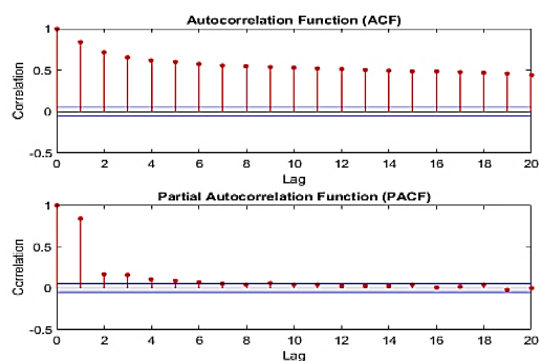


ب- استفاده از ضریب همبستگی پیرسون برای تعیین همبستگی آب‌دهی رود و بارش.

B: Using the Pearson correlation coefficient to determine the correlation between river streamflow and rainfall

شکل ۲- تعیین سنجه‌های ورودی مؤثر با استفاده از روش‌های ACF، PACF و ضریب پیرسون.

Figure 2- Determining effective input parameters using ACF, PACF and Pearson coefficient methods.



الف- استفاده از روش‌های ACF و PACF برای تعیین همبستگی آب‌دهی روزانه با آب‌دهی‌های گذشته.

A: Using ACF and PACF methods to determine the correlation of daily streamflow with past streamflows

جدول ۲- ترکیب ورودی‌های مؤثر به مدل‌ها.

Table 2-combination of effective inputs to models.

| Model number | Effective input structure | Output |
|--------------|---------------------------|--------|
| 1 | Q-1 | Q |
| 2 | Q-1, Q-2 | Q |
| 3 | Q-1, Q-2, Q-3 | Q |
| 4 | Q-1, P-1 | Q |
| 5 | Q-1, Q-2, P-1 | Q |
| 6 | Q-1, Q-2, Q-3, P-1 | Q |

شده است. بر پایه نتایج جدول ۳، عملکرد مدل ANFIS برای پیش‌بینی جریان قابل قبول بود. همچنین، مشخص شد در سناریوی ۱، مدل با چهار تابع عضویت مثلی $R^2=0/88$ و $RMSE=0/025$ در مرحله آموزش و $R^2=0/88$ و $RMSE=0/0208$ در مرحله صحت‌سنجی، کمترین اندازه خطا را در پیش‌بینی آب‌دهی رود داشت. بررسی دیگر ترکیبات ورودی و ساختارهای گوناگون ANFIS نشان داد توابع مثلی و گوسی عملکرد بیشتری در مدل‌سازی آب‌دهی رود منطقه مطالعه شده داشتند. عبدالله پور و ستاری (۲۰۱۵) دریافتند که استفاده از توابع مثلی و گوسی باعث بهبود

نتایج مدل ساختار استنتاج فازی-عصبی تطبیقی

پس از پیش‌پردازش و آماده‌سازی داده‌ها، شش سناریوی تعریف‌شده، با مدل ANFIS مدل‌سازی شدند. به منظور یافتن مدل بهینه در هر سناریو، سنجه‌های مدل شامل نوع و تعداد توابع عضویت بررسی شد. در این پژوهش از سه تابع عضویت مثلی (trimf)، دوزنقه‌ای (trapm) و گوسی (gaussmf) استفاده شد. با اجرای پرتکرار مدل و بر پایه معیارهای ارزیابی، مدلی که بیشترین دقت را داشت به‌عنوان بهترین مدل پیش‌بینی آب‌دهی رود یک روز بعد در ایستگاه پلدختر پیشنهاد شد. نتایج اجرای سناریوهای گوناگون در دو مرحله آموزش و صحت‌سنجی در جدول ۳ نشان‌داده

یک، دو و سه روز تأخیر همراه با بارش یک روز گذشته بود، در پیش‌بینی آب‌دهی رود بیش‌ترین دقت را داشت.

دقت مدل ANFIS شد که با یافته‌های این پژوهش هم‌راستا است. سناریوی ۶ ($RMSE=0.0203$ و $R^2=0.882$) که ترکیبی از آب‌دهی رود با

جدول ۳- نتایج مدل ANFIS در پیش‌بینی آب‌دهی روزانه رود.

Table 3- Results of ANFIS model in predicting daily flow.

| models | Inputs combination | Number and type of membership function | RMSE | | R ² | | NE | |
|--------|-----------------------------------|--|---------------|---------------|----------------|--------------|--------------|--------------|
| | | | Training | Validation | Training | Validation | Training | Validation |
| 1 | Q (t-1) | 4, trimf | 0.0250 | 0.0208 | 0.839 | 0.882 | 0.687 | 0.750 |
| | | 4, trapmf | 0.0251 | 0.0208 | 0.840 | 0.879 | 0.682 | 0.754 |
| | | 3, gaussmf | 0.0250 | 0.0218 | 0.838 | 0.872 | 0.684 | 0.730 |
| 2 | Q (t-1), Q (t-2) | 4, trimf | 0.0251 | 0.0207 | 0.834 | 0.880 | 0.680 | 0.756 |
| | | 2, trapmf | 0.0251 | 0.0219 | 0.837 | 0.866 | 0.683 | 0.727 |
| | | 3, gaussmf | 0.0253 | 0.0218 | 0.834 | 0.870 | 0.677 | 0.728 |
| 3 | Q (t-1), Q (t-2), Q (t-3) | 3, trimf | 0.0254 | 0.0216 | 0.828 | 0.866 | 0.675 | 0.735 |
| | | 4, trapmf | 0.0270 | 0.0228 | 0.803 | 0.848 | 0.632 | 0.705 |
| | | 2, gaussmf | 0.0254 | 0.0218 | 0.833 | 0.864 | 0.673 | 0.730 |
| 4 | Q (t-1), R(t-1) | 3, trimf | 0.0252 | 0.0203 | 0.836 | 0.885 | 0.678 | 0.768 |
| | | 3, trapmf | 0.0260 | 0.0211 | 0.825 | 0.883 | 0.658 | 0.747 |
| | | 3, gaussmf | 0.0250 | 0.0212 | 0.839 | 0.880 | 0.684 | 0.744 |
| 5 | Q (t-1), Q (t-2), R(t-1) | 3, trimf | 0.0256 | 0.0207 | 0.826 | 0.880 | 0.670 | 0.756 |
| | | 3, trapmf | 0.0277 | 0.0233 | 0.808 | 0.868 | 0.613 | 0.717 |
| | | 4, gaussmf | 0.0252 | 0.0206 | 0.831 | 0.881 | 0.679 | 0.758 |
| 6 | Q (t-1), Q (t-2), Q (t-3), R(t-1) | 3, trimf | 0.0264 | 0.0220 | 0.810 | 0.857 | 0.648 | 0.725 |
| | | 3, trapmf | 0.0280 | 0.0240 | 0.790 | 0.834 | 0.605 | 0.673 |
| | | 5, gaussmf | 0.0250 | 0.0203 | 0.835 | 0.882 | 0.682 | 0.761 |

نتایج مدل ماشین بردار پشتیبان

سناریوهای گوناگون در جدول ۴ آورده شده است. با توجه به شاخص‌های آماری به‌دست آمده، مشخص شد در تمام شرایط اندازه ϵ برابر با 0.01 مطلوب بود و با اندازه γ کمتر از ۱، دقت مدل بیش‌تر بود. بر پایه نتایج جدول ۴، ساختار ترکیبی شماره ۶ با تابع کرنل پایه شعاعی $C=4$ و $\gamma=1$ ، $\epsilon=0.01$ کمترین اندازه خطا را در دو مرحله آموزش $RMSE=0.0223$ و صحت‌سنجی $RMSE=0.0187$ داشت و بیش‌ترین اندازه همبستگی را در پیش‌بینی آب‌دهی جریان نشان داد.

در این مرحله نیز مانند مدل ANFIS، داده‌ها پس از آماده‌سازی به مدل SVM تعریف شدند. برای تعیین مدل بهینه، در هر سناریو اندازه‌های گوناگونی برای سه سنجه C, γ, ϵ در نظر گرفته شد. با اعمال اندازه‌های ۱ تا ۲۰ برای مشخصه C ، 0.01 تا 0.1 با گام 0.1 برای ϵ و 0.1 تا 2 با گام 0.1 برای γ تلاش شد، برای عملکرد این سه سنجه بررسی جامعی انجام شد. سرانجام بر اساس اندازه‌های R^2 و $RMSE$ داده‌های پیش‌بینی شده به‌وسیله مدل و اندازه‌های مشاهده‌ای، مدل بهینه انتخاب شد. نتایج به‌دست آمده از مدل SVM برای

جدول ۴- نتایج مدل SVM در پیش‌بینی آب‌دهی روزانه رود.

Table 4 - The results of the SVM model in predicting the daily flow rate.

| model | Inputs combination | C, ϵ, γ parameters | RMSE | | R ² | | NE | |
|----------|--|----------------------------------|---------------|---------------|----------------|--------------|--------------|--------------|
| | | | Training | Validation | Training | Validation | Training | Validation |
| 1 | Q (t-1) | 5,0.001,0.5 | 0.0240 | 0.0203 | 0.844 | 0.874 | 0.706 | 0.765 |
| 2 | Q (t-1), Q (t-2) | 15,0.001,0.8 | 0.0239 | 0.0200 | 0.846 | 0.878 | 0.711 | 0.771 |
| 3 | Q (t-1), Q (t-2), Q (t-3) | 15,0.001,0.8 | 0.0238 | 0.0195 | 0.848 | 0.885 | 0.714 | 0.782 |
| 4 | Q (t-1), R(t-1) | 1,0.001,0.6 | 0.0233 | 0.0202 | 0.854 | 0.876 | 0.725 | 0.766 |
| 5 | Q (t-1), Q (t-2), R(t-1) | 3,0.001,0.4 | 0.0223 | 0.0194 | 0.867 | 0.887 | 0.750 | 0.786 |
| 6 | Q (t-1), Q (t-2), Q (t-3), R(t-1) | 4,0.001,1 | 0.0223 | 0.0187 | 0.865 | 0.895 | 0.737 | 0.801 |

مبتنی بر گرادیان است که با محاسبه و اعمال به‌روزرسانی‌های وزن بر اساس گرادیان تابع خطا، وزن‌های ANN را به‌طور پرتکرار تنظیم می‌کند. در این پژوهش نتایج مربوط به مدل شبکه عصبی نشان داد همانند مدل‌های SVM و ANFIS، سناریوی ۶ ($R^2=0/199$ و $NE=0/793$, $RMSE=0/888$) بیش‌تری در مقایسه با دیگر ترکیبات ورودی داشت. شناسایی مناسب‌ترین دستورالعمل بهینه‌سازی می‌تواند منجر به بهبود نرخ همگرایی، تعمیم بهتر و پیش‌بینی دقیق‌تر شبکه‌های عصبی مصنوعی شود. در این پژوهش به‌منظور بررسی عملکرد دستورالعمل‌های بهینه‌سازی گوناگون در مقایسه با دستورالعمل LM دو دستورالعمل WOA و PSO در نظر گرفته شد. این دو دستورالعمل مبتنی بر جمعیت بود که تعداد جمعیت اولیه در آن‌ها ۸ و تعداد تکرار ۱۰۰ در نظر گرفته شد. نتایج مربوط به هر دو مدل ترکیبی در جدول ۵ آورده شده است. مقایسه نتایج نشان داد که استفاده از این دو دستورالعمل بهینه‌سازی باعث افزایش دقت مدل شبکه عصبی در پیش‌بینی آب‌دهی رود شد. به‌گونه‌ای که مدل ANN-WOA $RMSE=0/186$ و $R^2=0/896$ در مرحله صحت‌سنجی، کمترین اندازه خطا را در پیش‌بینی آب‌دهی رود داشت.

نتایج مدل‌های شبکه عصبی و دستورالعمل‌های بهینه‌سازی در مدل شبکه عصبی مصنوعی، تعداد لایه‌های پنهان، نورون‌های موجود در هر لایه و نوع تابع فعال‌ساز نقش مهمی در کارایی مدل‌ها دارند. با انتخاب مناسب این سنج‌ها می‌توان بهبود قابل توجهی در دقت و کارایی مدل‌های شبکه عصبی ایجاد کرد. یکی از روش‌های متداول برای آموزش شبکه‌های عصبی، استفاده از روش سعی و خطا است. در این روش، سنج‌های شبکه مانند وزن‌ها و بایاس‌ها با تکرار و به‌شکل تجربی بهینه می‌شوند. به این ترتیب که ابتدا سنج‌های اولیه انتخاب می‌شود و سپس مدل آموزش داده می‌شود و عملکرد آن روی داده‌های اعتبارسنجی ارزیابی می‌شود. سپس بر اساس نتایج، سنج‌ها تنظیم می‌شود و دوباره فرآیند آموزش و ارزیابی تکرار می‌شود. در مدل ترکیبی، این فرآیند معمولاً با استفاده از دستورالعمل‌های بهینه‌سازی انجام می‌شود، که به شبکه عصبی کمک می‌کند تا با دقت بیش‌تری به داده‌های آموزشی تطبیق یابد و عملکرد بهتری داشته باشد. از معروف‌ترین دستورالعمل‌های بهینه‌سازی که شبکه عصبی برای حل مسئله از آن استفاده می‌کند، دستورالعمل لوببرگ مارکوارت (LM) است که در این پژوهش نیز برای مدل ANN در نظر گرفته شده است. این دستورالعمل بهینه‌سازی

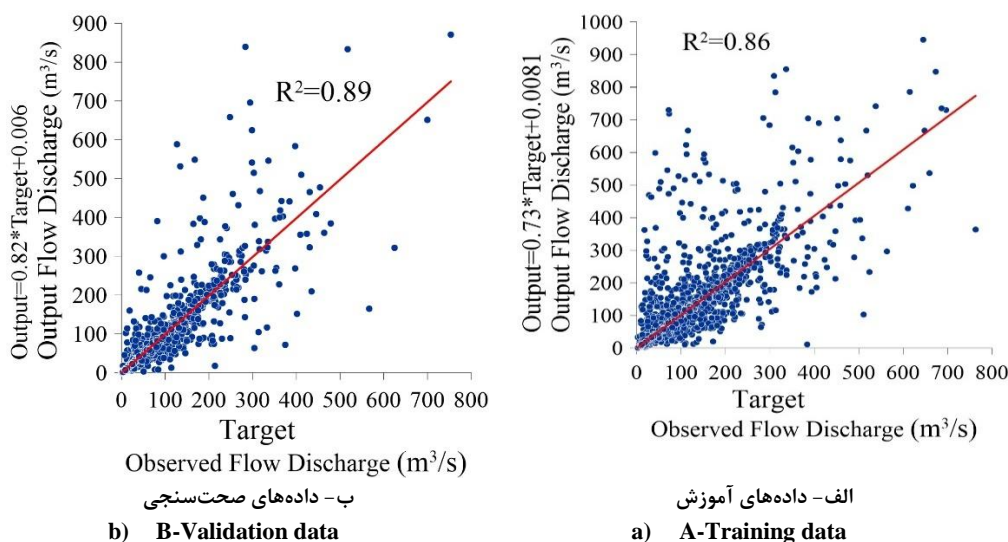
جدول ۵- نتایج مدل ANN، ANN-PSO و ANN-WOA در پیش‌بینی آب‌دهی روزانه رود.

Table 5 - The results of the SVM model in predicting the daily flow rate.

| models | Inputs combination | RMSE | | R2 | | NE | |
|---------|--|---------------|---------------|--------------|--------------|--------------|--------------|
| | | Training | Validation | Training | Validation | Training | Validation |
| ANN | Q (t-1) | 0.0244 | 0.022 | 0.837 | 0.855 | 0.700 | 0.720 |
| | Q (t-1), Q (t-2) | 0.0239 | 0.0200 | 0.843 | 0.879 | 0.711 | 0.771 |
| | Q (t-1), Q (t-2), Q (t-3) | 0.0234 | 0.0210 | 0.850 | 0.867 | 0.723 | 0.749 |
| | Q (t-1), R(t-1) | 0.0221 | 0.0253 | 0.868 | 0.827 | 0.754 | 0.637 |
| | Q (t-1), Q (t-2), R(t-1) | 0.0235 | 0.0204 | 0.850 | 0.876 | 0.722 | 0.764 |
| | Q (t-1), Q (t-2), Q (t-3), R(t-1) | 0.0228 | 0.0199 | 0.895 | 0.888 | 0.730 | 0.793 |
| ANN-PSO | Q (t-1) | 0.0239 | 0.0204 | 0.844 | 0.874 | 0.713 | 0.762 |
| | Q (t-1), Q (t-2) | 0.0237 | 0.0207 | 0.846 | 0.870 | 0.716 | 0.755 |
| | Q (t-1), Q (t-2), Q (t-3) | 0.0235 | 0.0201 | 0.849 | 0.878 | 0.721 | 0.77 |
| | Q (t-1), R(t-1) | 0.0228 | 0.0197 | 0.858 | 0.883 | 0.737 | 0.778 |
| | Q (t-1), Q (t-2), R(t-1) | 0.0229 | 0.0201 | 0.857 | 0.879 | 0.735 | 0.769 |
| | Q (t-1), Q (t-2), Q (t-3), R(t-1) | 0.0222 | 0.0191 | 0.867 | 0.892 | 0.734 | 0.791 |
| ANN-WOA | Q (t-1) | 0.0238 | 0.0205 | 0.844 | 0.873 | 0.713 | 0.760 |
| | Q (t-1), Q (t-2) | 0.0240 | 0.0215 | 0.842 | 0.860 | 0.709 | 0.737 |
| | Q (t-1), Q (t-2), Q (t-3) | 0.0237 | 0.0205 | 0.846 | 0.873 | 0.717 | 0.759 |
| | Q (t-1), R(t-1) | 0.0225 | 0.0203 | 0.862 | 0.877 | 0.744 | 0.765 |
| | Q (t-1), Q (t-2), R(t-1) | 0.0220 | 0.0258 | 0.869 | 0.823 | 0.755 | 0.622 |
| | Q (t-1), Q (t-2), Q (t-3), R(t-1) | 0.0222 | 0.0186 | 0.868 | 0.896 | 0.755 | 0.803 |

نکته قابل توجه این است که هر چه خط برازش داده شده میان داده‌ها نزدیکتر به زاویه ۴۵ درجه با محورهای مختصات باشد، دقت مدل بیشتر است (شکل ۳-الف و ب).

نمودارهای پراکنش اندازه‌های مشاهده‌ای در مقابل اندازه‌های پیش‌بینی شده به وسیله ANN-WOA در شکل ۳ آورده شده است. این نتایج نشان‌دهنده همبستگی زیاد میان نتایج مدل با اندازه‌های واقعی بود.

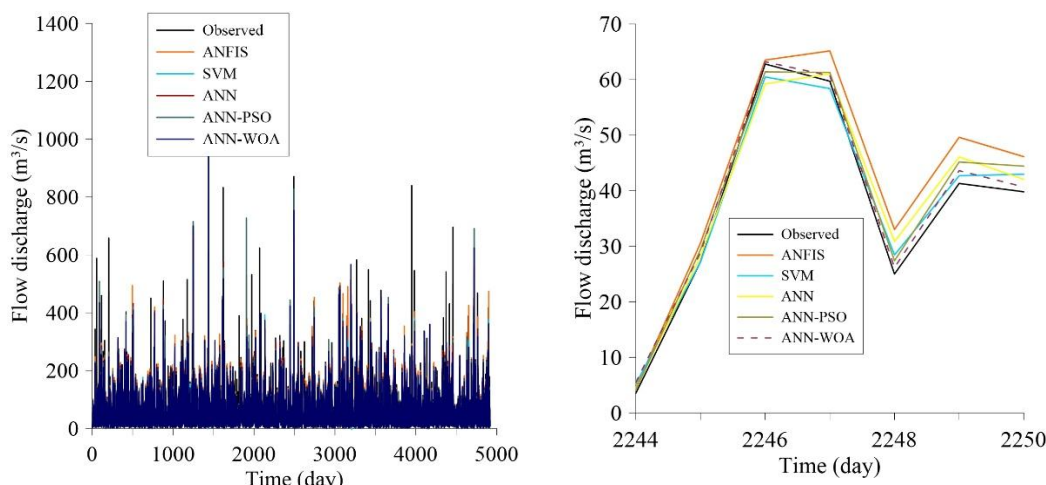


شکل ۳- نمودارهای پراکنش اندازه‌های مشاهده‌ای آب‌دهی در مقابل اندازه‌های پیش‌بینی شده به وسیله مدل بهینه ANN-WOA در مراحل آموزش و صحت‌سنجی.

Figure 3- Distribution charts of observed values of daily flow versus predicted values by the optimal ANN-WOA model in the training and validation stages.

برخی از داده‌ها دشوار باشد، این مدل‌ها می‌توانند با در اختیار داشتن کمترین سنج‌های ورودی، در پیش‌بینی آب‌دهی رود عملکرد قابل قبولی داشته باشند. برای درک بهتر، در شکل ۴ تغییر اندازه‌های مشاهده‌ای و محاسبه‌ای آب‌دهی روزانه رود به وسیله مدل‌ها نسبت به زمان برای بخشی از بازه مطالعه شده آورده شده است. بر پایه شکل ۴ نتایج مدل ترکیبی شبکه عصبی و دستورالعمل نهنگ با داده‌های واقعی هم‌خوانی بیشتر و خطای کمتری داشت. نمودارهای سری زمانی شکل ۴ نشان داد که اگرچه عملکرد مدل‌ها در پیش‌بینی آب‌دهی رود قابل قبول بود اما، در آب‌دهی‌های زیاد ضعیف عمل کردند و اندازه پیش‌بینی شده اختلاف قابل توجهی با اندازه‌های واقعی داشت.

نتایج این پژوهش نشان داد برای پیش‌بینی آب‌دهی روزانه در ایستگاه پل دختر عملکرد مدل‌های هوش مصنوعی قابل قبول بودند. با این حال، عملکرد مدل ANN-WOA با کمترین اندازه‌های RMSE در مراحل آزمایش و صحت‌سنجی، بهتر بود. عملکرد مدل SVM نیز با داشتن ضریب همبستگی زیاد و خطای کمتر در مقایسه با دیگر مدل‌ها، خوب بود. این یافته‌ها با نتایج پژوهش‌های صمدی و فتح‌آبادی (۲۰۱۸) و پریسوج و همکاران (۲۰۲۰) هم‌راستا است. می‌توان گفت این مدل، با استفاده از اصل استقرای کمینه‌سازی خطای ساختاری، به دنبال یک جواب بهینه کلی است و مطلوب‌ترین جواب ممکن را ارائه می‌دهد. از این‌رو، زمانی که کمبود داده‌های آماری وجود دارد یا ثبت



شکل ۴- نمودار سری زمانی اندازه‌های مشاهده‌ای و محاسبه‌ای آب‌دهی روزانه به‌وسیله دو مدل ANFIS و SVM
Figure 4- Time series graph of observed and calculated values of daily flow by ANFIS and SVM model.

همراه است. باید توجه داشت که کیفیت و ماهیت داده‌ها نقش مهمی در عملکرد این مدل‌ها دارد. معمولاً داده‌هایی که برای مدل‌سازی استفاده می‌شوند، محدود و ناقص هستند و ممکن است تأثیر منفی بر دقت پیش‌بینی داشته باشند. در نتیجه عملکرد مدل مطلوب نخواهد بود. همچنین، تغییرات ناگهانی در آبخیز مانند سیلاب و بارش‌های سیل آسا می‌تواند موجب کاهش دقت مدل در پیش‌بینی آب‌دهی رود شود. در این پژوهش تلاش شد با پیش‌پردازش و پالایش داده‌ها، اندازه خطای ناشی از ماهیت داده، کاهش یابد. با این حال نتایج نشان دادند، عملکرد مدل‌ها در پیش‌بینی آب‌دهی‌های زیاد مطلوب نبود. پیشنهاد می‌شود از دیگر مدل‌های یادگیری ماشین به‌ویژه مدل‌های جدید یادگیری عمیق و دستورالعمل‌های بهینه‌سازی دیگر استفاده شود و خروجی‌ها با نتایج این پژوهش مقایسه شود. سرانجام می‌توان گفت استفاده از مدل‌های هوشمند در تحلیل مباحث آب‌شناختی و مدیریت منابع آبی مفید و کم هزینه بودند و می‌توان آنها را جایگزین مناسبی برای مدل‌های مفهومی و آب‌شناختی که معمولاً پرهزینه هستند، دانست.

سپاس‌گزاری

نویسندگان به این وسیله از سازمان هواشناسی کشور و هم‌چنین سازمان آب و برق استان خوزستان که ما را

نتیجه‌گیری و پیشنهادها

در این پژوهش، عملکرد مدل‌های هوشمند در پیش‌بینی آب‌دهی روزانه آبخیز کشکان در استان لرستان بررسی شد. از این رو، داده‌های بارش و آب‌دهی رود ثبت‌شده در ایستگاه پل‌دختر، به‌عنوان انتهایی‌ترین نقطه آبخیز، جمع‌آوری شدند. اندازه‌های مشاهده‌ای آب‌دهی رود با اندازه‌های برآوردشده به‌وسیله مدل‌های ساختار استنتاج فازی عصبی تطبیقی، مدل ماشین بردار پشتیبان، شبکه عصبی، مدل‌های ترکیبی شبکه عصبی-دستورالعمل نهنگ و مدل ترکیبی شبکه عصبی-دستورالعمل ازدحام ذرات مقایسه شدند. بررسی معیارهای ارزیابی آماری نشان داد دقت مدل‌ها با ورودی‌های گوناگون، متفاوتی بود. نتایج بررسی همه مدل‌ها در این پژوهش نشان داد که آب‌دهی رود در یک (Q-1)، دو (Q-2) و سه (Q-3) روز گذشته و بارش یک روز گذشته (P-1)، بیش‌ترین همبستگی را با آب‌دهی روزانه نشان داد. نتایج این پژوهش نشان داد استفاده از دستورالعمل‌های بهینه‌سازی نهنگ و ازدحام ذرات برای آموزش شبکه عصبی، سبب شد تا حد زیادی دقت مدل افزایش یابد. سرانجام، مدل ANN-WOA ($RMSE=0.1186$ ، $NE=0.803$ و $R^2=0.896$) به‌عنوان برترین مدل انتخاب شد. اگرچه استفاده از مدل‌های هوش مصنوعی برای پیش‌بینی آب‌دهی رود می‌تواند مفید باشد، اما با چالش‌ها و محدودیت‌هایی

مشارکت نویسندگان

فاطمه عوض‌پور: مفهوم‌سازی، انجام تحلیل‌های نرم‌افزاری/آماری، نگارش نسخه اولیه مقاله.
محمدرضا هادیان: راهنمایی و مشاوره، بازبینی متن مقاله، بررسی نتایج.
علی طالبی: راهنمایی و مشاوره، بازبینی متن مقاله، بازبینی نتایج.

در انجام این پژوهش یاری کردند، صمیمانه تشکر می‌کنم.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش ندارند.

دسترسی به داده‌ها

داده‌ها و نتایج استفاده‌شده در این پژوهش، با مکاتبه با نویسنده مسئول در اختیار مخاطب قرار خواهد گرفت.

فهرست منابع

- Abdollah Pour Azad MR, Sattari MT. 2015. Forecasting daily river flow of Ahar Chay River using artificial neural networks (ANN) and comparison with Adaptive neuro fuzzy inference system (ANFIS). *Journal of Water and Soil Conservation*. 22(1):287-298. (In Persian). [https://doi : 20.1001.1.23222069.1394.22.1.15.7](https://doi.org/10.1001.1.23222069.1394.22.1.15.7)
- Asadabadi AS, Dinpazhuh Y, Mirabasi Najafabadi R. 2014. Forecasting the average daily discharge of the Beheshtabad River using wavelet analysis. *Journal of Water and Soil Science*. 28(3): 534-45. (In Persian). [https://doi: 20.1001.1.23222069.1394.22.1.15.7](https://doi.org/10.1001.1.23222069.1394.22.1.15.7)
- Buyukyildiz M, Kumcu SY. 2017. An estimation of the suspended sediment load using adaptive network based fuzzy inference system, support vector machine and artificial neural network models. *Journal of Water Resources Management*. 31(5): 1343-1359. <https://doi.org/10.1007/s11269-017-1581-1>
- Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi Hosseini F, Mousavi A. 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Journal of Science of the Total Environment*. 651(2/15): 2087-2096. <https://doi.org/10.1016/j.scitotenv.2018.10.064>
- Ditthakit P, Sirimon P, Nureehan S, Jakkarin W, Thai Th, and Quoc BP. 2023. Comparative study of machine learning methods and GR2M model for monthly runoff prediction. *Journal of Ain Shams Engineering*. 14(4) 101941. <https://doi.org/10.1016/j.asej.2022.101941>
- Emami H, Emami S, Heidari Tashe Kabud SH. 2019. Prediction suspended sediment load of river using meta-heuristic algorithms. *Journal of Irrigation and Drainage*. 13(5): 1426-1438. (In Persian). [https://doi: 20.1001.1.20087942.1398.13.5.21.5](https://doi.org/10.1001.1.20087942.1398.13.5.21.5)
- Eskandari A, Nouri R, Vesali Naseh MR, Saeedi F. 2019. Uncertainty evaluation of ANN and ANFIS models in estimating the inflow to Rais Ali Delwari dam. *Journal of Environmental Science and Technology*. 21(7):34-47. (In Persian). [https://doi: 10.22034/jest.2020.20068.2909](https://doi.org/10.22034/jest.2020.20068.2909)
- Falamaki A, Eskandari M, Baghlani A, Ahmadi SA. 2012. Modeling total sediment load in rivers using artificial neural networks. *Journal of Water and Soil Resources Conservation*. 2(3):13-25. (In Persian). <https://sid.ir/paper/403921/fa>
- Ghafari GH, Vafakhah M. 2013. Simulation of rainfall-runoff process using artificial neural network and fuzzy-adaptive neural system (case study: Hajighoshan watershed). *Journal of Watershed Management Research*. 4(8):120-136. (In Persian). <https://jwmr.sanru.ac.ir/article-1-321-fa.html>
- Hakimi Khansar H, Parsa J, Momeni Keleshteri O, Karami N, Khoshdel Sangdeh M. 2024. Extended estimation of daily inflow of Sefidroud dam using meta-heuristic algorithms combined with fuzzy neural inference system. *Journal of Civil Engineering, Amirkabir*. 56(1): 03-22. (In Persian). [https://doi: 10.22060/ceej.2024.21634.7784](https://doi.org/10.22060/ceej.2024.21634.7784)
- Jang, J-SR. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on systems, man, and cybernetics*. 23(3): 665-85.
- Joudi HA, Kadkhoda Hosseini M, Akhavan S, Nozari H. 2015. Evaluation of AWAT and SVM models in the simulation of Ligvanchai river runoff. *Journal of Water and Soil Science*. 26(4/1): 137-150. (In Persian). <https://civilica.com/doc/1587524>
- Katipoğlu OM. 2023. Modeling the effect of meteorological variables on streamflow estimation: application of data mining techniques in mixed rainfall-snowmelt regime Munzur River.

- Türkiye. *Journal of Environmental Science and Pollution Research*. 30(4): 96312–96328. <https://doi.org/10.1007/s11356-023-29220-2>
- Latif SD, Chong KL, Ahmed AN, Huang YF, Sherif M, Shafie A EL. 2023. Sediment load prediction in Johor River: deep learning versus machine learning models. *Journal of Applied Water Science*. 13(3): 79-93. <https://doi.org/10.1007/s13201-023-01874-w>
- Malik A, Tikhamarine Y, Souag-Gamane D, Rai P, Sammen SS, Kisi O. 2021. Support vector regression integrated with novel meta-heuristic algorithms for meteorological drought prediction. *Journal of Meteorology and Atmospheric Physics*. 133(2): 891–909. <https://doi.org/10.1007/s00703-021-00787-0>
- Marcé R, Comerma M, García JC. Armengol J. 2004. A neuro-fuzzy modeling tool to estimate fluvial nutrient loads in watersheds under time-varying human impact. *Limnology and Oceanography Methods*, 2(11): 342-355. <https://doi.org/10.4319/lom.2004.2.342>
- Memari M, Harifi M, Khalili A. 2020. Improving the classification performance of polynomial neural network using whale algorithm. *Journal of Electrical and Computer Engineering*. 20(2):145-154. (In Persian). <https://sid.ir/paper/959517/fa>
- Mirjalili S, Lewis A. 2016. The whale optimization algorithm. *Journal of Advances in Engineering Software*. 95(6):51-67 <https://doi.org/10.1016/j.advengsoft.2016.01.008>
- Mohammadi B, Moazenzadeh R. 2019. Prediction of stream flow using intelligent hybrid models in monthly scale (case study: Zarrin roud River). *Journal of Environmental Science and Technology*. 21(9):71-81. (In Persian). <https://doi.org/10.22034/jest.2020.24315.3331>.
- Moriasi DN, Jeffrey GA, Michael WVL, Ronald LB, Daren H, Tamie LV. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers (ASABE)*. 50(3): 885–900. <https://doi.org/10.13031/2013.23153>
- Nazeri Tahroudi M, Hashemi SR, Ahmadi F, Nazeri Tahroudi Z. 2016. Examining the accuracy of ANFIS, SVM and GP models in modeling the river flow discharge. *Journal of Echohydrology*. 3(3):347-361. (In Persian). <https://doi.org/10.22059/ije.2016.60024>
- Nie Y, Sun J, Jiehua M. 2023. Seasonal prediction of summer extreme precipitation frequencies over southwest China based on machine learning. *Journal of Atmospheric Research*. 294 (5): 106947. <https://doi.org/10.1016/j.atmosres.2023.106947>
- Nourani V, Alizadeh F, Roushangar K. 2015. Evaluation of a two-Stage SVM and spatial statistics methods for modeling monthly river suspended sediment load. *Journal of Water Resour Management*. 30(9):393-407. <https://doi.org/10.1007/s11269-015-1168-7>
- Parisouj P, Mohebzadeh H, Lee T. 2020. Employing machine learning algorithms for streamflow prediction: A case study of four River basins with different climatic zones in the United States. *Journal of Water Resource Management*. 34(3): 4113-4131. <https://doi.org/10.1007/s11269-020-02659-5>
- Samadi M, Fath Abadi A. 2018. Application of Time Series, ANN, and SVM Models in forecasting the Gorgan Dam inflow rate. *Journal of Environment and Water Engineering*. 4(4): 299-309. (In Persian). <https://doi.org/10.22034/jewe.2018.128256.1256>
- Samantaray S, Sahoo A. 2022. Prediction of suspended sediment concentration using hybrid SVM-WOA approaches. *Journal of Geocarto Internationa*. 37(19): 5609-5635. <https://doi.org/10.1080/10106049.2021.1920638>
- Shahhoseini Sh, Moosavi M, Mollajafari M. 2011. Evolutionary algorithms- fundamentals, applications, implementation. Tehran Press Center, Iran University of Science and Technology, Tehran, 590 p. (In Persian).
- Shahi Nejad B, Dehghani R. 2017. Evaluation and performance of support vector machine model in estimation of suspended sediment. *Journal of Irrigation and Water Engineering*. 8(29):25-34. (In Persian). https://www.waterjournal.ir/article_74207.html
- Studies on the modernization of the country's water comprehensive plan in the border watersheds of the West, Karkheh, Karun and Jarahi-Zahre. 2006. Ministry of Energy and Water Research Institute. pp. 52-58 (In Persian).
- Yu X, Liang SY, Babovic V. 2004. EC-SVM approach for real-time hydrologic forecasting. *Journal of Hydroinformatics*, 6 (3): 209–223. <https://doi.org/10.2166/hydro.2004.0016>



The performance evaluation of machine learning models and optimization algorithms for predicting the River Discharge of Kashkan River

Fatemeh Avazpour¹, Mohammad Reza Hadian², Ali Talebi³

1- Ph.D. Candidate, Department of Civil Engineering, Yazd University, Yazd, Iran

2- Assistant Professor, Department of Civil Engineering, Yazd University, Yazd, Iran

3- Professor, Department of Natural Resources and Desert Studies, Yazd University, Yazd, Iran

Extended Abstract

Introduction and Goal

The simulation of river discharge at hydrometric stations to predict future flow discharge over specific time periods is an important issue typically addressed using hydrological time series associated with the respective station. To predict river discharge with the highest accuracy, three major groups of methods are commonly utilized: empirical and statistical methods, conceptual methods, and process-based approaches. Among data-driven methods, those based on artificial intelligence-based are prominent. The aim of this study was evaluating the performance of machine learning models, including SVM, ANFIS, and ANN, and assessing the performance of a neural network model trained with Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO) to predict flow discharge. In addition to evaluating the intelligent models, the impact of using optimization algorithms on the accuracy of river discharge predictions was also examined. Since input data have a significant impact on the performance of data-driven models, the criteria influencing the river discharge were identified, and the best combination of input variables for each model was determined.

Materials and Methods

In this study, to predict the daily discharge at the Poldokhtar hydrometric station located on the Kashkan River, discharge and precipitation data from 1971 to 2018 were collected, and Intelligent models, including Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machine (SVM), Artificial Neural Network (ANN), and the hybrid model of Artificial Neural Network with Particle Swarm Optimization (ANN-PSO) and the hybrid model of Artificial Neural Network with Whale Optimization Algorithm (ANN-WOA) were employed. In the two hybrid models, efforts were made to adjust the criteria of the artificial neural network using metaheuristic algorithms, and their impact on the performance of the ANN model was examined.

Article Type: Research Article

***Corresponding Author E-mail:** mr_hadian@yazd.ac.ir

Citation: Avazpour, F., Hadian, M.R., Talebi, A. 2025. Evaluation of the performance of machine learning models and optimization algorithms in predicting flow discharge of Kashkan River. *Watershed Management Research*. 38(1): 1-18.

DOI: 10.22092/WMRJ.2024.365128.1579

Received: 06 March 2024, **Received in revised form:** 09 May 2024, **Accepted:** 20 June 2024

Published online: 21 March 2025

Watershed Management Research, Vol.38, No.1, Ser. No:146, Spring 2025, pp. 1-18.

Publisher: Fars Agricultural and Natural Resources Research and Education Center ©Author(s)



Additionally, this study investigated the impact of river discharge and precipitation data, along with their time lags (data from previous days), and combinations of these metrics as input variables for the models. To determine the best combination of input variables, statistical methods such as the Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), and Pearson Correlation Coefficient (PCC) were employed. After applying the effective inputs and training the mentioned intelligent models, their performance in predicting river discharge was evaluated by comparing RMSE, R^2 , and NE metrics.

Results and Discussion

The evaluation of all models in this study showed that the river discharge of one day (Q-1), two days (Q-2), and three days (Q-3) ago, along with the precipitation of one day ago (P-1), exhibited the highest correlation with the river's daily discharge. Overall, all models demonstrated acceptable accuracy in modeling the river discharge in the Kashkan watershed. According to the results, the highest accuracy in predicting daily discharge was achieved by the ANN-WOA model, with the highest coefficient of determination ($R^2 = 0.896$), Nash-Sutcliffe efficiency (NE = 0.803), and the lowest error (RMSE = 0.0186).

Subsequently, the SVM model, using a radial basis kernel function with parameters $C=4$, $\gamma=1$, and $\epsilon=0.001$ demonstrated superior performance, with a coefficient of determination ($R^2 = 0.895$), Nash-Sutcliffe efficiency (NE = 0.801), and an error (RMSE = 0.0187). Then, the ANN-PSO and ANN models ranked third and fourth, respectively. The results indicated that using metaheuristic optimization algorithms significantly improved the accuracy of the ANN model, making it a suitable tool for neural network training. The evaluation of different ANFIS structures revealed that triangular and Gaussian functions performed better for modeling river discharge in the study area. On the other hand, the error of this model, with values of RMSE=0.023 and NE=0.76 was higher compared to the other models.

Conclusion and Suggestions

This study demonstrated that machine learning models, such as SVM, ANFIS, and ANN, exhibited acceptable accuracy in predicting river discharge. Adjusting neural network parameters using optimization algorithms like WOA and PSO significantly enhanced the performance of the ANN model. Finally, it can be concluded that these models can serve as suitable alternatives to conceptual and hydrological models for addressing hydrological and discharge-related issues. It is recommended to train the SVM and ANFIS models using the PSO and WOA algorithms and then compare the results with the findings of this study.

Keywords: Kashkan Watershed, Rainfall-Runoff, Particle Swarm Optimization, Whale Optimization Algorithm, Artificial Intelligence

Article Type: Research Article

Acknowledgement

The authors sincerely thank the Meteorological Organization and the Khuzestan Water & Power Authority for their collaboration in data collection.

Conflicts of interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Data Availability Statement

The datasets are available upon a reasonable request to the corresponding author.

Authors' Contribution

Fatemeh Avazpour: Conceptualization, performing software/statistical analyses, drafting the initial manuscript.

Mohammadreza Hadian: Supervision and guidance, reviewing the manuscript, verifying results.

Ali Talebi: Supervision and guidance, reviewing the manuscript, verifying results.