



مرکز تحقیقات آموزش کشاورزی و منابع طبیعی فارس

پژوهش‌های آبخیزداری

شاپا: ۲۰۳۸-۲۹۸۱



مركز تحقيقات آموزش و آموزش کشاورزی

پیش‌بینی آب‌دهی رود فریزی با بهره‌گیری از محاسبه‌های نرم

صابر جمالی^۱، فرشته رحیمی‌آغ‌چشمه^۲، محمدجواد امیری^{۳*}

- ۱ - دانشجوی دکتری، گروه علوم و مهندسی آب، دانشکده‌ی کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران
 ۲ - دانشجوی دکتری، گروه مرتع و آبخیزداری، دانشکده‌ی محیط‌زیست و منابع طبیعی، دانشگاه فردوسی مشهد، مشهد، ایران
 ۳ - دانشیار، گروه علوم و مهندسی آب، دانشکده‌ی کشاورزی، دانشگاه فسا، فسا، ایران

چکیده‌ی مبسوط

مقدمه و هدف

پیش‌بینی دقیق جریان رود به‌عنوان یک منبع مهم آب شیرین روی زمین، در مهندسی و مدیریت منابع آب ضروری است؛ از این رو، توسعه‌ی فن‌آوری که آب‌دهی رود را پیش‌بینی کند، ضروری است. در این زمینه مدل‌های مختلفی به‌وسیله‌ی پژوهش‌گران پرشماره ارائه‌شده است. این مدل‌ها به دو دسته‌ی مدل‌های فیزیکی مبتنی بر اصول آب‌شناسی/آبی و مدل‌های مبتنی بر محاسبه‌های نرم تقسیم می‌شوند. تمام مقاله‌های چاپ‌شده شواهدی از اهمیت کاربرد مدل‌های مبتنی بر محاسبه‌های نرم برای مشکلات آب‌شناختی، به‌ویژه آب‌دهی رود هستند.

مواد و روش‌ها

در این پژوهش برای پیش‌بینی آب‌دهی رود فریزی از مدل برنامه‌ریزی بیان ژن (GEP) و ماشین بردار پشتیبان (SVM) استفاده شد. در این پژوهش، از گروه روزانه آب‌دهی در ده سال (۱۳۹۹-۱۳۹۰) مربوط به ایستگاه آب‌سنجی موشنگ استفاده شد. میانگین بلندی منطقه‌ی بررسی‌شده ۲۱۷۱ متر از سطح دریا با طول جغرافیایی $58^{\circ} 4' 30''$ تا $58^{\circ} 49' 30''$ شرقی و عرض جغرافیایی $36^{\circ} 20' 1''$ تا $36^{\circ} 32' 1''$ است. از داده‌های آب‌دهی روزانه رود از ۱ تا ۵ روز قبل به‌عنوان ورودی مدل‌های GEP و SVM استفاده شد. به‌منظور اطمینان از همگنی و تکمیل

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: mj_amiri@fasau.ac.ir

استناد: جمالی، ص، رحیمی‌آغ‌چشمه، ف، امیری، م.ج.، ۱۴۰۳. پیش‌بینی آب‌دهی رود فریزی با بهره‌گیری از محاسبه‌های نرم. پژوهش‌های آبخیزداری، ۳۶ (۱): ۹۷-۸۱.

شناسه‌ی دیجیتال: 10.22092/WMRJ.2023.360554.1504

تاریخ دریافت: ۱۴۰۱/۰۹/۰۸، تاریخ بازنگری: ۱۴۰۱/۱۰/۱۳، تاریخ پذیرش: ۱۴۰۱/۱۲/۲۸، تاریخ انتشار: ۱۴۰۲/۱۰/۰۱

پژوهش‌های آبخیزداری، سال ۱۴۰۲، دوره‌ی ۳۶، شماره‌ی ۴، شماره‌ی پیاپی ۱۴۱، زمستان ۱۴۰۲، صفحه‌های ۸۱ تا ۹۷

© نویسندگان

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



کمبود داده‌های آب‌دهی استفاده‌شده از آزمون ران و ضریب همبستگی میان ایستگاه‌های هم‌جوار استفاده شد. سپس داده‌ها به شکل تصادفی به دو بخش، ۸۰٪ برای آموزش و ۲۰٪ برای آزمون و تعیین خطای مدل‌سازی تفکیک شدند. در مراحل آموزش و اعتبارسنجی براساس ریشه‌ی میانگین مربعات خطا (RSME)، ضریب همبستگی (R)، اریبی مدل، کارایی مدل کلینگ گوپتا (KGE) و نش-ساتکلیف (NSE) عملکرد مدل بررسی شد. در این پژوهش به منظور برآورد جریان ورودی به رود فریزی با کاربرد مدل SVM، سه نوع تابع کرنل رایج در آب‌شناسی شامل تابع‌های پایه‌ی خطی، چند جمله‌ای و شعاعی بررسی شد.

نتایج و بحث

از میان تابع‌های گوناگون، تابع مبنای شعاعی به دلیل داشتن کمترین اندازه‌ی خطا برای متغیرها انتخاب شد. بهترین الگوی ورودی، الگوی شماره‌ی ۵ بود که در آن متغیرهای آب‌دهی پیشین با پنج گام زمانی تأخیر استفاده شد، و در مرحله‌ی آموزش در مدل GEP و SVM بهترین عملکرد را در پیش‌بینی آب‌دهی روزانه‌ی ایستگاه موشنگ داشت. عملکرد مدل اعمال‌شده نشان SVM ($RSME = 1.15$, $R = 0.985$, $NSE = 0.85$ و $KGE = 0.79$) در مرحله‌ی اعتبارسنجی برای پیش‌بینی آب‌دهی روزانه‌ی رود از مدل GEP ($RSME = 1.65$, $R = 0.964$, $NSE = 0.78$ و $KGE = 0.69$) دقیق‌تر است.

نتیجه‌گیری و پیشنهادها

این پژوهش نشان داد که روش محاسبه‌های نرم (مانند SVM و GEP)، ابزار قدرتمندی در پیش‌بینی جریان رود است. با کاربرد این مدل‌ها می‌توان میان سنج‌های ورودی و خروجی رابطه‌ی مطلوب ایجاد کرد و امکان شبیه‌سازی دقیق جریان میانگین و حداکثر روزانه را فراهم ساخت.

واژگان کلیدی: آبخیز فریزی، بیان ژن، رواناب، محاسبه‌های نرم، هوش مصنوعی

مقدمه

به مزیت‌های این مدل، استفاده از آن در پیش‌بینی آب‌دهی جریان، بیش از پیش ضروری به نظر می‌رسد. نتایج پیش‌بینی آب‌دهی رود با استفاده از محاسبه‌های نرم نشان داد که در پیش‌بینی آب‌دهی‌های زیاد، مدل برنامه‌ریزی بیان ژن (GEP) در مقایسه با مدل شبکه‌ی عصبی مصنوعی (ANN) عملکرد بهتری داشت (عبداللهی و همکاران ۲۰۱۷). نتایج شبیه‌سازی آب‌دهی رود با استفاده از معیارهای ارزیابی RMSE و ضریب نش-ساتکلیف (NS) نشان داد که در برآورد جریان رود مدل جدید ترکیبی ANN در مقایسه با GEP عملکرد بهتری داشت (دل‌افروز و همکاران ۲۰۱۸). نتایج پژوهش شافو و همکاران (۲۰۲۱) نشان داد که مدل برنامه‌ریزی بیان ژن در پیش‌بینی آب‌دهی ماهانه از سالانه عملکرد ضعیفی داشت. نتایج بررسی قربانی و همکاران (۲۰۱۶) نشان داد که در برآورد آب‌دهی رود مدل ماشین بردار پشتیبان (SVM) در مقایسه با مدل وایزلی چندگانه عملکرد بهتری داشت. نتایج پیش‌بینی آب‌دهی رود با استفاده از مدل‌های هوشمند نشان داد که مدل ANN در مقایسه با مدل SVM عملکرد بهتری داشت (قربانی و همکاران ۲۰۱۸). در کشور پاکستان نتایج پیش‌بینی آب‌دهی رود با استفاده

در طول زمان با افزایش جمعیت و تغییر کاربری زمین‌ها، اهمیت برنامه‌ریزی برای مدیریت جامع آبخیزها بیشتر شده است. مشکل اساسی برای برنامه‌ریزی کمبود ایستگاه آب‌سنجی است. یک روش برای حل این مشکل، برآورد اندازه‌ی رواناب آبخیز براساس ویژگی‌های طبیعی آن است. به‌منظور تصمیم‌گیری و برآورد سیل در آبخیز می‌توان از مدل‌های آب‌شناختی-توزیعی و مدل‌های داده محور استفاده کرد (بهره‌مند و همکاران ۲۰۰۵؛ صفری و همکاران ۲۰۱۴؛ آذین‌مهر و همکاران ۲۰۱۶). پیش‌بینی آب‌دهی جریان برای تأمین و بهره‌برداری از منابع آب سطحی و همچنین مدیریت سیلاب، ضروری است. بررسی تغییرات آب‌دهی رودها به دلیل خشک‌سالی‌ها و تغییرهای اقلیمی رخ داده و همچنین زیاد بودن متغیرهای مؤثر و قطعی نبودن آن حائز اهمیت است. از این رو، یکی از رویکردهای مناسب در پژوهش برآورد آب‌دهی رود و آب‌دهی سیلاب، استفاده از محاسبه‌های نرم به‌عنوان روشی برای پیش‌بینی در شرایط نبودن قطعیت و حقیقت جزئی است (خاشعی-سیوکی و سربازی ۲۰۱۵). محاسبه‌های نرم برای نشان دادن ابهام در تفکر انسان با نبودن قطعیت زندگی واقعی فرصت جذابی فراهم کرده است. با توجه

به‌وسیله‌ی مناطق مسکونی، باغ‌ها و مراکز تفریحی اشغال شده است. با وجود شرایط طبیعی و دخالت‌های انسانی، این منطقه در معرض خطر شدید سیل است. بنابراین، بررسی شرایط آب‌شناخت این آبخیز اجتناب‌ناپذیر است. از طرف دیگر در خروجی آبخیز یک ایستگاه آب‌سنجی است که آب‌نگار سیل را به‌شکل ساعتی ثبت می‌کند. بنابراین می‌توان افزون بر بررسی شرایط آب‌شناسی، با مقایسه نتایج آن با داده‌های ایستگاه آب‌سنجی، سنج‌های مدل را واسنجی کرد و برای شرایط مشابه استفاده کرد. از این رو، این پژوهش با هدف شبیه‌سازی جریان رود فریزی با گام زمانی کوتاه‌مدت انجام شد تا دقت مدل در شبیه‌سازی جریان‌های سیلابی مشخص شود.

مواد و روش‌ها

منطقه‌ی بررسی‌شده و داده‌های استفاده‌شده

آبخیز فریزی با مساحت تقریبی ۲۷۸ کیلومترمربع در شمال غرب مشهد در استان راسان رضوی است. طول جغرافیایی آن $30^{\circ} 58'$ تا $30^{\circ} 59'$ شرقی و عرض جغرافیایی آن $36^{\circ} 20'$ تا $36^{\circ} 32'$ است. میانگین شیب این آبخیز ۳۶/۷۶٪ است. میانگین بلندی آبخیز ۲۱۷۱ متر است. موقعیت آبخیز فریزی در استان خراسان رضوی و ایران در شکل ۱، نشان‌داده شده است. در این پژوهش داده‌های مربوط به گروه زمانی آبدهی (۹۹-۱۳۹۰) برای ایستگاه آب‌سنجی موشنگ (در خروجی آبخیز فریزی) از آب منطقه‌ای خراسان رضوی تهیه شد. در گام نخست، به‌منظور اطمینان از همگنی و تکمیل کمبود داده‌های آبدهی استفاده‌شده از آزمون ران و ضریب همبستگی بین ایستگاه‌های هم‌جوار استفاده شد. سپس داده‌ها به‌شکل تصادفی به دو قسمت، ۸۰٪ برای آموزش و ۲۰٪ برای آزمون و تعیین خطای مدل‌سازی تفکیک شدند. براساس نقشه‌ی مدل رقومی ارتفاع (DEM) آبخیز فریزی که در شکل ۲ نشان‌داده شده است، بیشترین و کمترین بلندی به‌ترتیب ۳۱۳۳ و ۱۴۱۰ متر از سطح آزاد دریا است. با استفاده از روش سنجش از دور و براساس ترکیب تصویرهای ماهواره‌ی لندست ۸ و سنجنده‌ی مودیس و با کد نویسی در محیط گوگل ارث انجین (GEE) نقشه‌ی کاربری زمین‌های آبخیز فریزی تهیه شد (شکل ۲). اطلاعات آبدهی لحظه‌ای جمع‌آوری شد و سپس برای شبیه‌سازی رواناب، از مدل بیان ژن GEP (با بهره‌گیری از نرم‌افزار GenXPro) و مدل ماشین بردار پشتیبان (با کد نویسی در محیط Matlab) استفاده شد.

از محاسبه‌های نرم نشان داد که مدل جنگل تصادفی (RF) در مقایسه با مدل SVM به اندازه‌ی ۳۳/۶٪ عملکرد بهتری داشت (حسین و خان ۲۰۲۰). همچنین در کشور مالزی نتایج پیش‌بینی آبدهی رود با استفاده از محاسبه‌های نرم نشان داد که مدل‌های مبتنی بر RF در مقایسه با مدل‌های مبتنی بر LS-SVM عملکرد بهتری داشت (پندهیانی و همکاران ۲۰۲۰). نتایج پژوهش ساهو و همکاران (۲۰۲۱) نشان داد که مدل ماشین بردار پشتیبان بهینه‌شده با دستورالعمل کرم شبتاب (SVM-FA) توانست آبدهی رود را با ضریب تبیین ۰/۹۸ پیش‌بینی کند و در مقایسه با مدل SVM عملکرد بهتری داشت. در استان گیلان نتایج پیش‌بینی آبدهی رودهای پسیخان و خالکایی نشان داد که مدل‌های گروه زمانی SARIMA در مقایسه با ماشین بردار پشتیبان حداقل مربعات (LS-SVM) عملکرد بهتری داشت (خداخواه و همکاران ۲۰۲۲). نتایج پیش‌بینی آبدهی رود کسینجا با استفاده از محاسبه‌های نرم نشان داد که مدل درخت تصمیم M5P در مقایسه با مدل بردار پشتیبان براساس ضریب نش-ساتکلیف عملکرد بهتری داشت (نیوش و همکاران ۲۰۲۲). نتایج پیش‌بینی آبدهی رود باراندوزچای با استفاده از محاسبه‌های نرم نشان داد که در برآورد آبدهی براساس شاخص‌های ضریب تبیین و RMSE، مدل GP در مقایسه با مدل SVM عملکرد بهتری داشت (احمدی و همکاران ۲۰۱۶). نتایج پیش‌بینی آبدهی رود ورودی به سد گرگان با بهره‌گیری از مدل‌های ANN و SVM نشان داد که مدل SVM در مقایسه با مدل شبکه عصبی، RMSE کمتری داشت (صمدی و فتح‌آبادی ۲۰۱۹). نتایج پیش‌بینی آبدهی رود گاماسیاب با استفاده از فرا مدل‌های هوشمند نشان داد که مدل GEP در مقایسه با مدل بیزین ارجحیت و برتری نسبی داشت (زینعلی و همکاران ۲۰۲۰). نتایج پژوهش رحیمی و حافظ‌پرست‌مودت (۲۰۲۰) نشان داد که در شبیه‌سازی آبدهی در دوره‌ی آینده (۲۰۲۱-۲۰۵۰) مدل GEP در مقایسه با مدل SVM عملکرد بهتری داشت. نتایج بررسی فاطمی و همکاران (۲۰۲۲) در شبیه‌سازی آبدهی رود نشان داد که براساس ضریب NS به‌دست آمده، مدل SVM در مراحل آموزش و آزمایش به‌ترتیب عملکرد خیلی خوب و خوب داشت و بهره‌گیری از دستورالعمل ژنتیک سبب بهبود عملکرد این مدل شد.

آبخیز فریزی مشهد یک آبخیز با شیب زیاد و تغییرات زیاد کاربری است که بیشتر بخش‌های کناره‌ی رود

یک‌راه حل پیدا شود و یا نسل‌ها به تعداد معینی برسند، تکامل متوقف‌شده و بهترین راه‌حل یافت شده، گزارش می‌شود (سلگی و همکاران ۲۰۱۷؛ زینعلی و همکاران ۲۰۲۰). چکیده‌ی مراحل این روش در شکل ۳-الف نشان‌داده شده است. تابع برازش به‌وسیله‌ی پردازش تعدادی از مسئله هدف، ارزیابی شد (رابطه‌ی ۱).

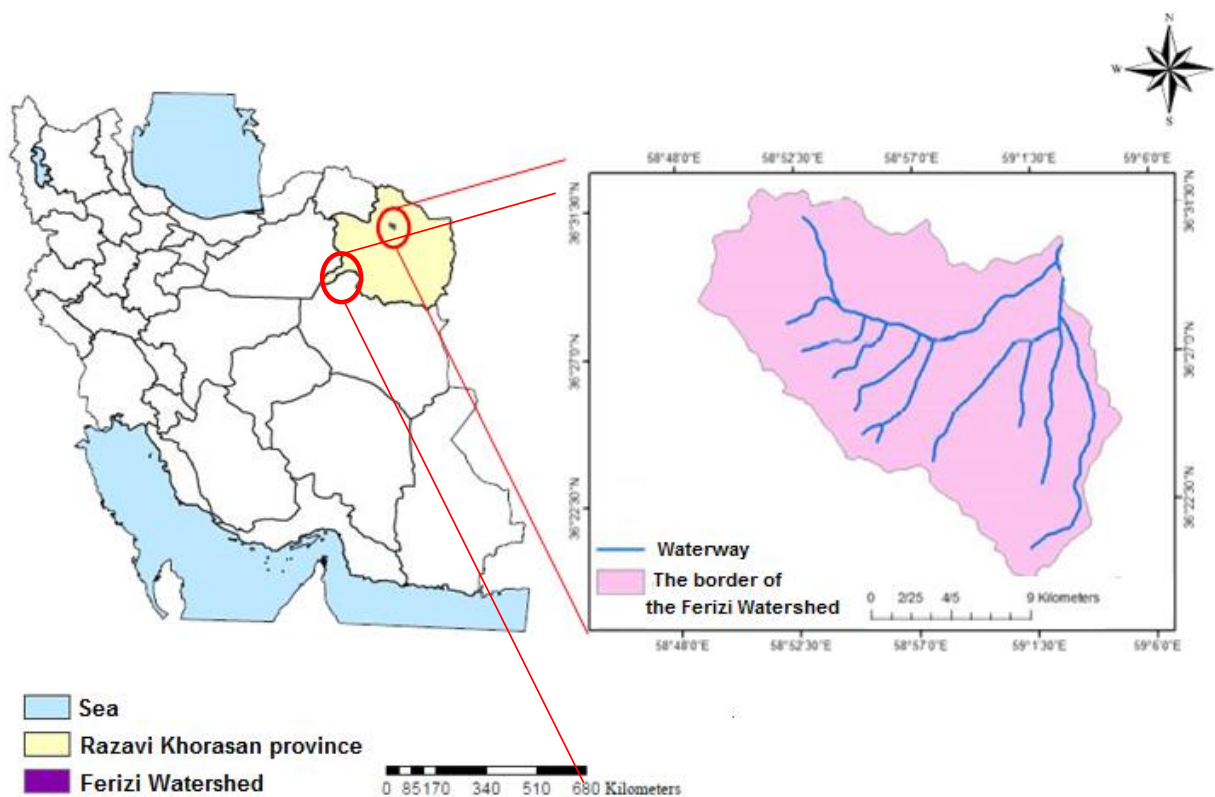
$$f_i = \sum_{j=1}^n \left(R - \left| 100 \times \frac{P_{ij} - T_j}{T_j} \right| \right) \quad (1)$$

R: محدوده‌ی انتخابی، P_{ij}: اندازه‌ی پیش‌بینی‌شده به‌وسیله‌ی برنامه‌ی انفرادی i برای برازش j (از میان n تعداد برازش) و T_j: اندازه‌ی هدف برای برازش j است (قبادیان و همکاران ۲۰۱۳).

اندازه‌های مربوط به تنظیم‌های کلی و عملکردهای ژنتیکی در جدول ۱ نشان‌داده شده است. سنجه‌های ورودی به مدل در الگوهای مختلف در جدول ۲ آورده شده است. براساس نتایج پژوهش‌های پیشین، بهترین نتایج از میان مدل‌های مختلف برای پیش‌بینی جریان روزانه در شرایطی اتفاق می‌افتد که داده‌های ورودی تا پنج تأخیر زمانی استفاده شود (سالاری جزی و همکاران ۲۰۱۶).

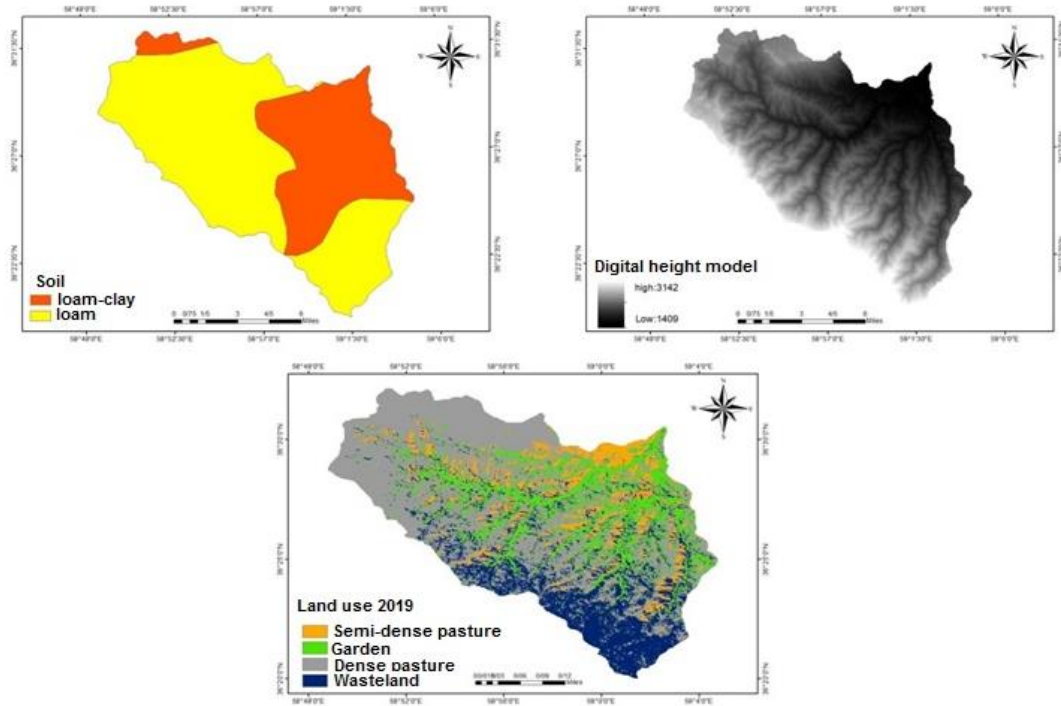
برنامه‌ریزی بیان ژن

مدل برنامه‌ریزی بیان ژن (GEP) تعمیم‌یافته دستورالعمل ژنتیک GA است که براساس نظریه‌ی داروین ارائه و به‌وسیله‌ی فریرا و همکاران (۲۰۰۱) ابداع شد. برنامه‌ریزی GEP نیز همانند GA و برنامه‌ریزی ژنتیک (GP)، یک دستورالعمل ژنتیکی است که جمعیتی از افراد را مطابق برازندگی انتخاب می‌کند و تغییرات ژنتیکی را با استفاده از یک یا چند عملگر ژنتیکی اعمال می‌کند. تفاوت اساسی بین این سه دستورالعمل، مربوط به ماهیت افراد آن‌هاست؛ به‌طوری‌که در GA افراد رشته‌های خطی با طول ثابت (کروموزوم) و در GP نهادهای غیرخطی با اندازه‌ها و شکل‌های متفاوت (درختان تجزیه) می‌باشند، درحالی‌که در GEP افراد به‌شکل رشته‌های خطی با طول ثابت (ژنوم یا کروموزوم) کدگذاری می‌شوند و سپس به‌شکل نهادهای غیرخطی با اندازه‌ها و شکل‌های متفاوت بیان می‌شوند (سالاری جزی و همکاران ۲۰۱۶). اولین مرحله در دستورالعمل GEP، تولید جمعیت اولیه بود که به‌وسیله‌ی فرایند تصادفی و یا با استفاده از بخشی از اطلاعات درباره‌ی مسئله انجام شد. کروموزوم‌ها به‌شکل بیان درختی در شکل ۳ نشان‌داده شده است. اگر کیفیت رضایت‌بخش از

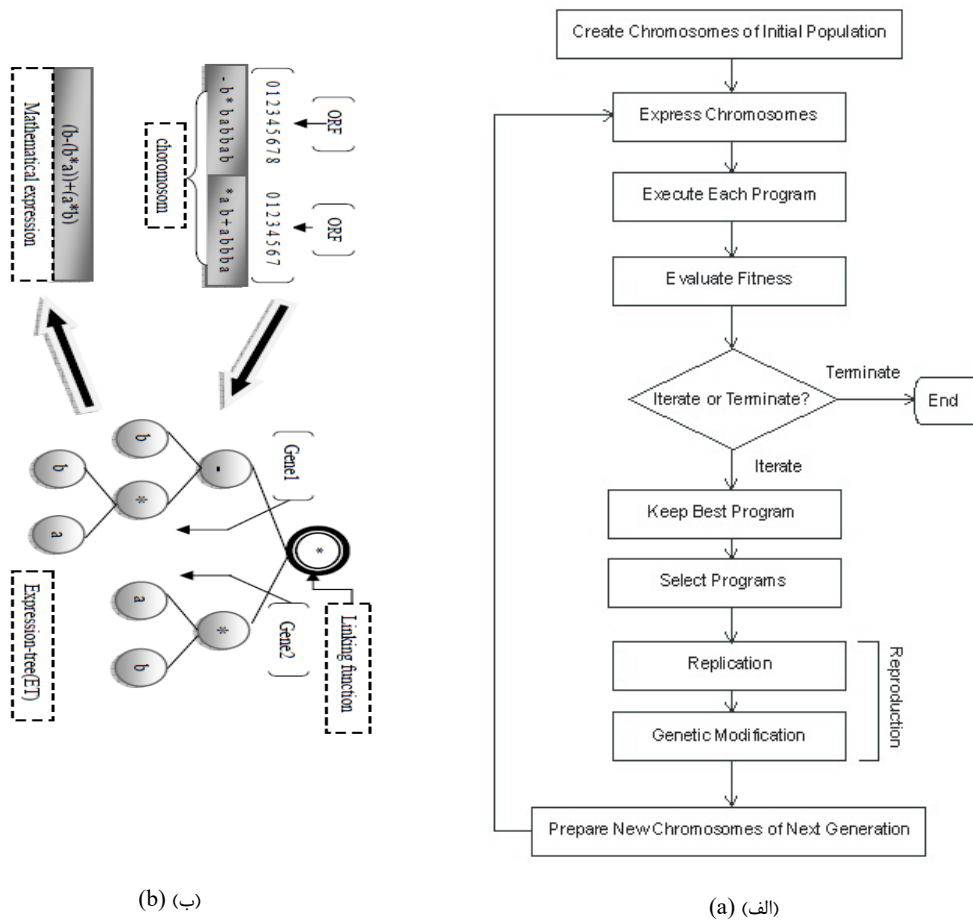


شکل ۱- موقعیت جغرافیایی و شبکه‌ی آبراهه‌های آبخیز فریزی در استان خراسان رضوی.

Figure 1- Geographical Location and Waterway Network of the Ferizi Watershed in the Razavi Khorasan Province.



شکل ۲- نقشه‌ی رقمی ارتفاعی، بافت خاک و کاربری زمین‌های آبخیز فریزی، خراسان رضوی.
 Figure 2- Altitude Digital Map, Soil Texture, and Land Use of the Ferizi Watershed in Razavi Khorasan.



شکل ۳- طرح‌واره‌ی انجام پژوهش با GEP (الف)، کدگذاری و کد برداری به شکل بیان درختی ET در یک کروموزوم با دو ژن (ب).
 Figure 3- The Scheme of Conducting Research with GEP (a) and Coding and Decoding ET as a Tree Expression in a Chromosome with Two Genes (b)

جدول ۱- اندازه‌های سنج‌های استفاده‌شده در مدل برنامه‌ریزی بیان ژن.

Table 1 - Parameter Values Used in the Gene Expression Programming Method.

Parameter	Value	Parameter	Value
Number of chromosomes	30	Mutation rate	0.044
Head size	7	Gene transposition rate	0.1
General settings	Number of genes per chromosome	Genetic functions	Gene recombination rate
	5		0.1
	Error criterion of the fitting function		One-point recombination rate
	R		0.3
	Linking function		Two-point recombination rate
	SUM (+)		0.3

جدول ۲- ترکیب‌های انتخاب‌شده‌ی سنج‌های ورودی مدل برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان.

Table 2- Selected Combinations of Input Parameters for the Gene Expression Programming Model and Support Vector Machine.

Model number	Selected combination
1	$Q(t)=f[Q(t-1)]$
2	$Q(t)=f[Q(t-1) Q(t-2)]$
3	$Q(t)=f[Q(t-1) Q(t-2) Q(t-3)]$
4	$Q(t)=f[Q(t-1) Q(t-2) Q(t-3) Q(t-4)]$
5	$Q(t)=f[Q(t-1) Q(t-2) Q(t-3) Q(t-4) Q(t-5)]$

جبری میان متغیرهای مستقل و وابسته با تابع جبری مانند $f(x)$ به‌علاوه مقداری خطای مجاز است (رابطه‌های ۲ و ۳؛ دهقانی و همکاران ۲۰۱۷).

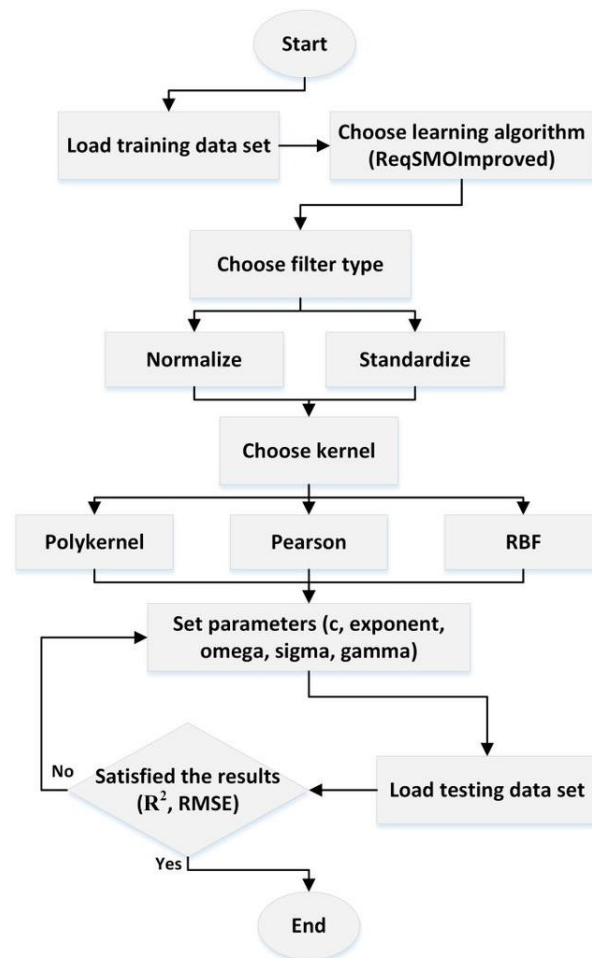
$$f(x) = W^T \cdot \varphi(x) + b \quad (2)$$

$$y = f(x) + noise \quad (3)$$

اگر W بردار ضریب‌ها و b ثابت مشخصه‌های تابع وایزی φ : تابع کرنل باشد، آنگاه هدف پیدا کردن شکل تابعی برای $f(x)$ است. این مهم با آموزش مدل به‌وسیله‌ی مجموعه‌ای از نمونه‌ها محقق می‌شود.

ماشین بردار پشتیبان SVM

ماشین بردار پشتیبان به‌عنوان یک ساختار یادگیری کارآمد بر مبنای نظریه‌ی بهینه‌سازی مقید است و با بهره‌گیری از اصل استقرای کمینه‌سازی خطای ساختاری، توسعه‌یافته است؛ به‌طوری‌که در انتها به جواب بهینه کلی منتج می‌شود (عیسی‌زاده و همکاران ۲۰۱۹). طرح‌واره‌ی انجام پژوهش با مدل ماشین بردار پشتیبان در شکل ۴ برگرفته از عیسی‌زاده و همکاران ۲۰۱۹ است. در این روش تابعی با متغیر وابسته y که خود تابعی از چند متغیر مستقل x است، برآورد شده و مشابه دیگر روش‌های وایزی‌گیری، بیان‌گر رابطه‌ی

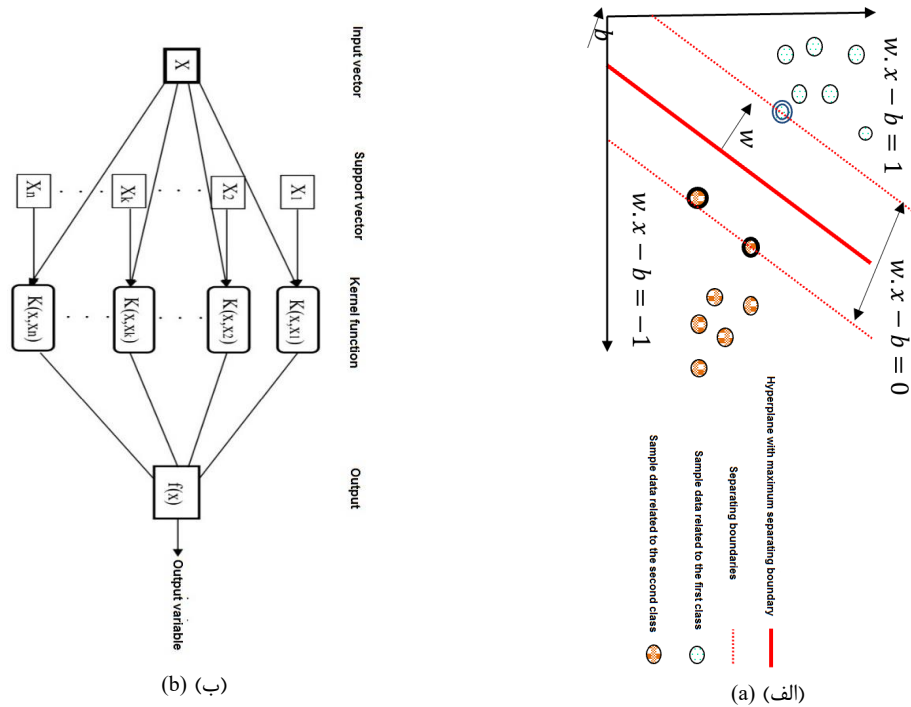


شکل ۴- طرح‌واره‌ی انجام پژوهش با مدل ماشین بردار پشتیبان.

Figure 4- Research Methodology for Conducting Research with the Support Vector Machine Method

می‌کند (ستاری و همکاران ۲۰۱۶). ابر صفحه جداساز و فاصله‌ی نقاط تا خطوط جداکننده و شمای ساده‌ای از مدل‌سازی با ماشین بردار پشتیبان برگرفته از ستاری و همکاران (۲۰۱۶) در شکل ۵ نشان داده شده است.

در این روش هدف، یافتن تابعی با کمترین اندازه‌ی انحراف از Y است. ماشین بردار پشتیبان، داده‌های متغیرهای مستقل (داده‌های اصلی) را با به‌کارگیری مجموعه‌ای از تابع‌های کرنل در فضای جدیدی بازآرایی



شکل ۵- ابر صفحه جداساز و فاصله‌ی نقاط تا خطوط جداکننده (الف) و شمای ساده‌ای از مدل‌سازی با ماشین بردار پشتیبان (ب).
Figure 5- Separator Superplane and Distance of Points to Separator Lines (a), and a Simple Diagram of Modeling with Support Vector Machine (SVM) (b).

O: اندازه‌های مشاهده‌ای، n : تعداد داده‌ها، \bar{P} : میانگین داده‌های شبیه‌سازی‌شده و \bar{O} : میانگین داده‌های مشاهده‌ای است.

معیارهای آماری استفاده‌شده برای ارزیابی مدل
 برای ارزیابی مدل در مراحل واسنجی و اعتبارسنجی از معیارهای آماری نش-ساتکلیف، شاخص کلینگ-گوپتا، ضریب تبیین و RMSE استفاده شد.

ضریب نش-ساتکلیف NS
 ضریب نش-ساتکلیف نشان‌دهنده‌ی درستی شبیه‌سازی آب‌دهی‌های جریان با استفاده از مدل است. اندازه‌های این شاخص میان منفی بی نهایت تا یک متغیر است و اندازه‌های نزدیک به یک بیان‌گر کارایی خوب مدل است. در این پژوهش ضریب نش-ساتکلیف از رابطه‌ی ۶ تعیین شد.

$$NS = 1 - \frac{\sum(Q_o - Q_s)^2}{\sum(Q_o - \bar{Q}_o)^2} \quad (6)$$

Qo: داده مشاهده‌ای، Qs: داده‌ی شبیه‌سازی‌شده و \bar{Q}_o : میانگین داده‌ی مشاهده‌ای است. طبقه‌بندی ضریب نش-ساتکلیف در جدول ۳ آورده شده است.

برای پیش‌بینی با مدل SVM از تابع کرنل RBF استفاده شد و کد نویسی در نرم‌افزار Matlab انجام شد. در این مدل همانند مدل بیان ژن از پنج ورودی جدول ۲، استفاده شد.

نمودار تیلور

برای ارزیابی عملکرد الگوهای مختلف استفاده‌شده در پیش‌بینی آب‌دهی و مقایسه‌ی با آب‌دهی مشاهده‌ای از نمودار تیلور استفاده شد (تیلور ۲۰۰۱). در این نمودار برای ارزیابی از سه آماره‌ی ریشه‌ی میانگین مربعات خطا (RMSE)، ضریب همبستگی و انحراف معیار میان داده‌های آب‌دهی شبیه‌سازی‌شده و مشاهده‌ای استفاده شد. برای محاسبه ریشه‌ی میانگین مربعات خطا و ضریب همبستگی به ترتیب از رابطه‌های ۴ و ۵ استفاده شد. برای رسم این نمودار، کد نویسی با نرم‌افزار Matlab انجام شد.

$$RMSE = \sqrt{\frac{\sum(P-O)^2}{n}} \quad (4)$$

$$r = \frac{\sum((O-\bar{O}) \times (P-\bar{P}))}{\sqrt{\sum(P-\bar{P})^2} \times \sqrt{\sum(O-\bar{O})^2}} \quad (5)$$

P: اندازه‌های شبیه‌سازی‌شده به‌وسیله‌ی مدل‌ها،

جدول ۳- طبقه‌های اندازه‌های معیار نش-ساتکلیف (احمدیوسفی و همکاران ۲۰۱۹).

Table 3- Classes corresponding to Nash-Sutcliffe criterion values (Ahmadyousefi et al. 2019).

Evaluation criteria	Very weak	Weak	Good	Very good	Excellent
Nash-Sutcliffe	<0.2	0.2-0.5	0.5-0.65	0.65-0.85	>0.85

از مدل برنامه‌ریزی بیان ژن و ماشین بردار پشتیبان استفاده شد. برای تعیین بهترین الگوی استفاده‌شده در شبیه‌سازی آبدهی به‌وسیله‌ی مدل‌های مزبور در مرحله‌ی آموزش از نمودار تیلور استفاده شد. نمودار تیلور در شکل ۶ نشان داده شده است. براساس نتایج شبیه‌سازی آبدهی با مدل برنامه‌ریزی بیان ژن و با توجه به شکل ۶-الف، مشخص شد که استفاده از الگوی شماره‌ی ۵، کمترین اندازه‌ی RMSE و بیشترین اندازه‌ی r را داشت. همچنین این الگو بیشترین دقت را در میان الگوهای بررسی شده داشت. در این پژوهش براساس نتایج شبیه‌سازی آبدهی در مرحله‌ی آموزش با استفاده از مدل ماشین بردار پشتیبان و با توجه به شکل نشان داد شکل ۶-ب مشخص شد که الگوی شماره‌ی ۵ در مقایسه با دیگر روش‌ها بیشترین اندازه‌ی r و کمترین اندازه‌ی RMSE را داشت. بخش‌هایی از گروه زمانی آبدهی در سال‌های ۱۳۹۰ تا ۱۳۹۶ در دو مدل GEP و SVM در جدول ۴ آورده شده است.

شاخص کلینگ-گوپتا KGE

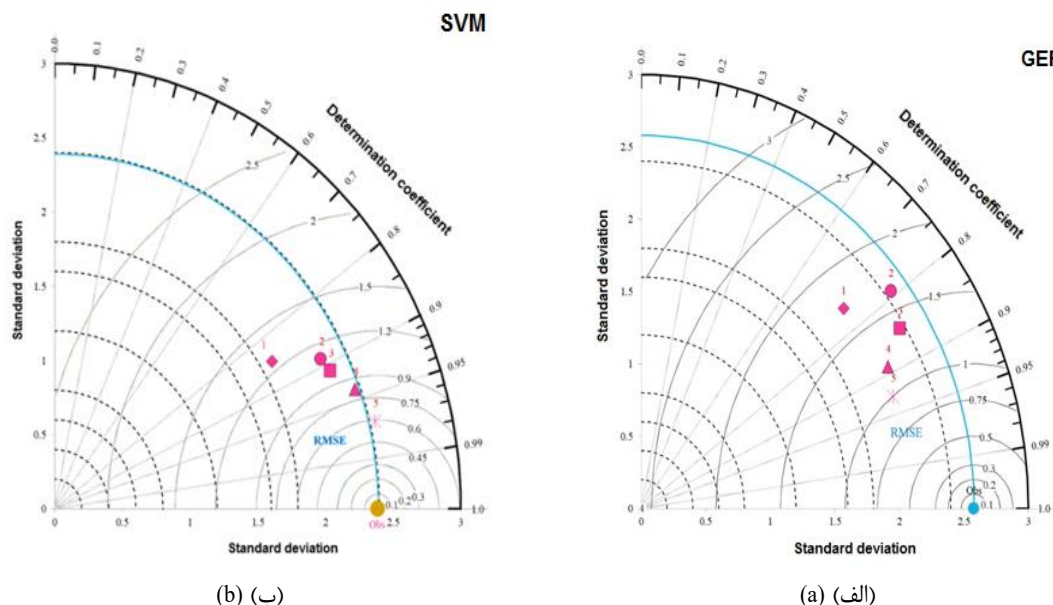
در این پژوهش برای ارزیابی مدل از شاخص کلینگ-گوپتا استفاده شد. بهترین اندازه برای معیار ارزیابی کلینگ-گوپتا عدد یک است که نشان‌دهنده‌ی تطابق کامل آب‌نگارها است. در این پژوهش این شاخص با استفاده از رابطه‌ی ۷ محاسبه شد.

$$1 - \sqrt{(r-1)^2 + (a-1)^2 + (B-1)^2} \quad a = \frac{\sigma_s}{\sigma_o} \quad \beta = \frac{\mu_s}{\mu_o}$$

سنجه‌های r : ضریب همبستگی داده‌های میان آبدهی شبیه‌سازی و مشاهده‌ای، a : نسبت انحراف معیار، β : نسبت میانگین، σ_s : انحراف معیار اندازه‌های شبیه‌سازی، σ_o : انحراف معیار اندازه‌های مشاهده‌ای، μ_s : میانگین اندازه‌های شبیه‌سازی و μ_o : میانگین اندازه‌های مشاهده‌ای است.

نتایج و بحث

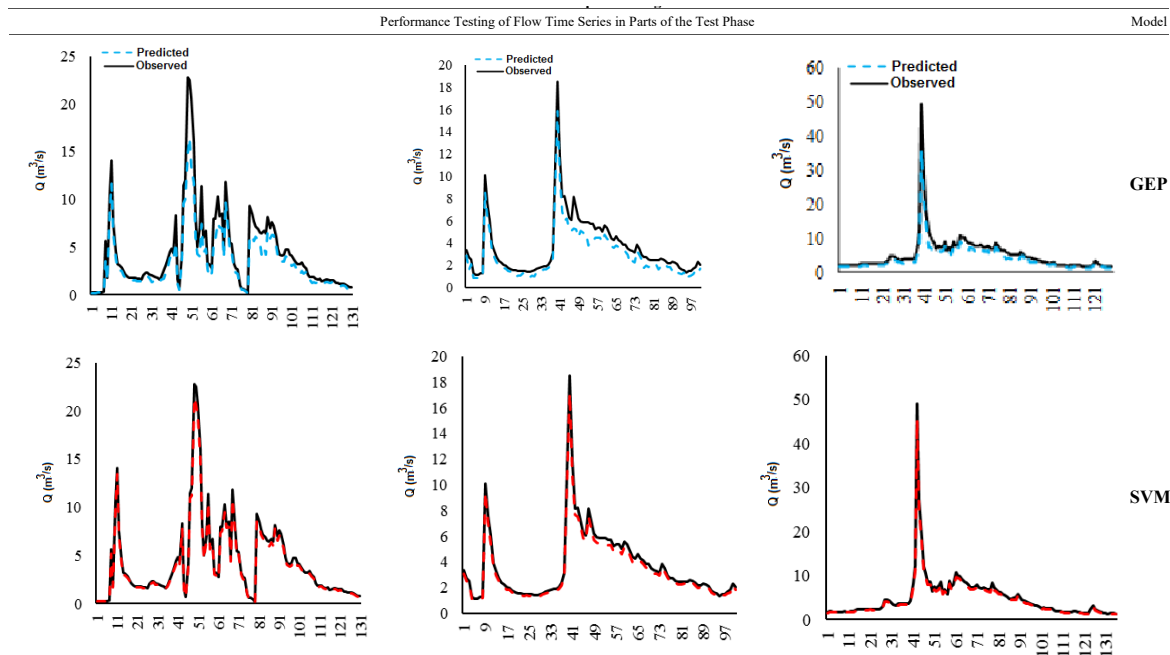
در این پژوهش برای مدل‌سازی آبدهی در آبخیز فریزی



شکل ۶- نمودار تیلور برای تعیین بهترین الگوی استفاده‌شده در مرحله‌ی آزمون در مدل GEP (الف) و SVM (ب).

Figure 6- Taylor Diagram to Determine the Best Model Used in the Test Phase: GEP (a) and SVM (b) Models

جدول ۴- شبیه‌سازی آبدهی در مرحله‌ی آزمایش با استفاده از مدل GEP و SVM.
Table 4- Flow simulation in the test phase using GEP and SVM methods.

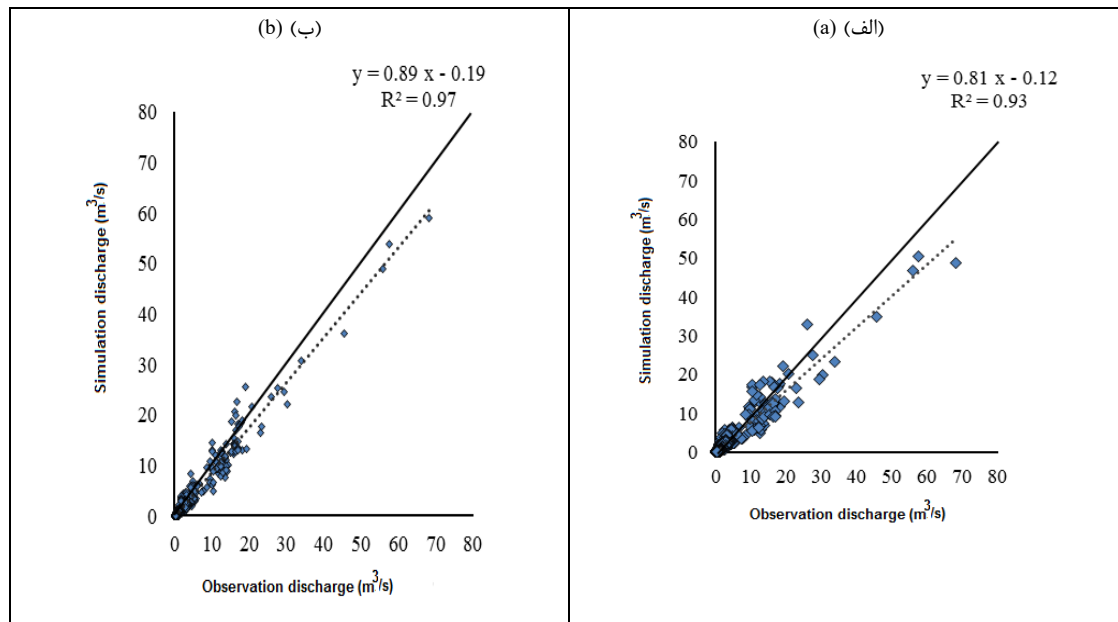


برآوردی و مشاهده‌ای به جز چند نقطه که اطراف خط ۱:۱ (خط نیمساز) نبودند، در دیگر نقاط روی خط برازش بودند. این یافته دلالت بر برابری اندازه‌های مشاهده‌ای و محاسبه‌ای داشت. همچنین نشان‌دهنده‌ی عملکرد مطلوب این مدل در برآورد اندازه‌های جریان رود فریزی بود. اما مدل GEP در برآورد اندازه‌های سیلابی عملکرد خوبی نداشت. یافته‌های این پژوهش با نتایج پژوهش‌های سالاری‌جزی و همکاران (۲۰۱۶)، دهقانی و همکاران (۲۰۱۷) و رحیمی و حافظ‌پرست‌مودت (۲۰۲۰) مطابقت داشت. براساس نتایج این پژوهش و با توجه به شکل ۸-ب)، مدل ماشین بردار پشتیبان در برآورد برخی اندازه‌های سیلابی عملکرد قابل قبول نداشت، به‌گونه‌ای که اندازه‌های مشاهده‌ای از خط نیمساز فاصله داشت و برآورد بیشتری را نشان داد. این نتایج با نتایج پژوهش‌های قربانی و همکاران (۲۰۱۶)، احمدی و همکاران (۲۰۱۶)، دهقانی و همکاران (۲۰۲۱) و خداخواه و همکاران (۲۰۲۲) همخوانی داشت. با توجه به شکل ۸، هر دو مدل روند تغییرات آبدهی را در به‌خوبی پیش‌بینی کردند و روش SVM آبدهی را در مقایسه با مدل GEP بهتر پیش‌بینی کرد. با توجه به نتایج این پژوهش مدل SVM در مقایسه با مدل GEP آبدهی سیلابی را بهتر پیش‌بینی کرد. یافته‌های این پژوهش با نتایج پژوهش‌های انجام شده در رود تلوار و رود خرخره‌چای مطابقت داشت (بایزیدی و همکاران ۲۰۱۸؛ بهمنش و همکاران ۲۰۱۷). ارزیابی نتایج شبیه‌سازی آبدهی در جدول ۵ آورده شده است. براساس نتایج به‌دست آمده، در دو مدل GEP و SVM ضریب نش-

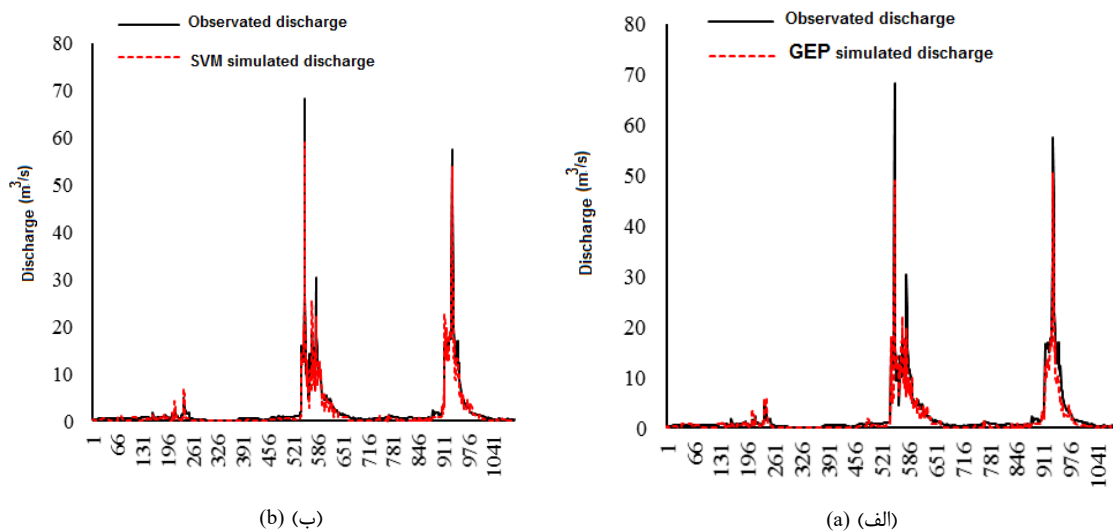
برای پیش‌بینی آبدهی جریان در آبخیز فریزی در مرحله‌ی آموزش در هر دو مدل مزبور از الگوی شماره‌ی ۵ که بهترین نتایج را در مقایسه با دیگر الگوها نشان داد، استفاده شد. پراکنش آبدهی شبیه‌سازی در مقابل آبدهی مشاهده‌ای با استفاده از دو مدل بیان ژن و ماشین بردار پشتیبان در شکل ۷ نشان داده شده است. براساس نتایج این پژوهش و با توجه به شکل ۷-الف، در مدل برنامه‌ریزی بیان ژن شیب خط برازش کمتر از یک (برابر با ۰/۸۱) بود، بنابراین این روش آبدهی جریان را کم‌تر برآورد کرد. براساس نتایج خط ۱:۱، در آبدهی‌های زیاد و سیلابی این روش کم‌برآورد بود. نتایج این پژوهش بیان‌گر آن است که دقت مدل GEP در برآورد آبدهی‌های کم در مقایسه با آبدهی‌های سیلابی بیشتر بود. براساس نتایج این پژوهش و با توجه به شکل ۷-ب، مدل ماشین بردار پشتیبان نیز آبدهی را کم برآورد کرد. مدل SVM نیز در آبدهی‌های کم برآورد بهتری را نشان داد. نتایج این پژوهش در پیش‌بینی آبدهی جریان با استفاده از مدل GEP و SVM با نتایج پژوهش‌های انجام شده در رود کارون، آبخیز دریاچه‌ی نمک، رود دز، رود کشکان و رود باراندوزچای مطابقت داشت (دهقانی و همکاران ۲۰۲۱؛ علی‌نژادی و همکاران ۲۰۲۱؛ شریفی‌گرم‌دره و همکاران ۲۰۱۹؛ نوذری و توکلی ۲۰۱۸؛ احمدی و همکاران ۲۰۱۶). نتایج مربوط به آبدهی مشاهده‌ای و شبیه‌سازی شده با مدل برنامه‌ریزی بیان ژن GEP و ماشین بردار پشتیبان SVM در شکل ۸ نشان داده شده است. براساس پراکنش نتایج مدل GEP و با توجه به شکل ۸-الف، اندازه‌های

شاخص کلینگ-گوپتا، به-ترتیب ۰/۶۹ و ۰/۷۹ بود، که بیان‌گر مطابقت بهتر نتایج آب‌دهی مشاهده‌ای با نتایج مدل SVM بود.

سات‌کلیم به-ترتیب ۰/۷۸ و ۰/۸۵ بود، که بیان‌گر عملکرد خیلی خوب مدل در پیش‌بینی بود. براساس نتایج آریبی، هر دو مدل آب‌دهی را کم برآورد کردند. براساس نتایج به‌دست آمده، در دو مدل SVM و GEP اندازه‌های



شکل ۷- پراکندگی اندازه‌های آب‌دهی مشاهده‌ای و شبیه‌سازی در مرحله‌ی آزمون در مدل برنامه‌ریزی بیان ژن (الف) و ماشین بردار پشتیبان (ب).
Figure 7-Dispersion of Observed and Simulated Discharge Values in the Test Phase: Gene Expression Programming Model (a) and Support Vector Machine (b).



شکل ۸- گروه زمانی آب‌دهی مشاهده‌ای و شبیه‌سازی در مرحله‌ی آزمون در مدل برنامه‌ریزی بیان ژن (الف) و ماشین بردار پشتیبان (ب).
Figure 8- Time Series of Observational Discharge and Simulation in the Test Phase: Gene Expression Programming Model (a) and Support Vector Machine (b).

جدول ۵- نتایج معیارهای ارزیابی مدل‌های GEP و SVM در پیش‌بینی آب‌دهی در مرحله‌ی آزمون.

Table 5-The Evaluation Criteria Results of GEP and SVM Models in Discharge Prediction During the Test Stage.

Parameter	GEP	SVM
R ²	0.93	0.97
RMSE	1.65	1.15
NS	0.78	0.85
KGE	0.69	0.79

نتیجه‌گیری و پیشنهادها

در این پژوهش، بررسی عملکرد مدل‌های برنامه‌ریزی بیان ژن (GEP) و ماشین بردار پشتیبان (SVM) در برآورد آب‌دهی جریان روزانه در رود فریزی در ایستگاه موشنگ استان خراسان رضوی انجام شد. برای پیش‌بینی آب‌دهی با مدل‌های مزبور از میانگین آب‌دهی روزانه ایستگاه موشنگ در دوره‌ی آماری ۱۳۹۰ تا ۱۳۹۹ استفاده شد. همچنین، برای مدل‌سازی از الگوهای مختلف آب‌دهی روزهای پیشین (گام تأخیر) استفاده شد. در این پژوهش، میان مدل‌های استفاده‌شده، بهترین الگوی ورودی، الگوی شماره ۵ بود که در آن متغیرهای آب‌دهی پیشین با پنج گام زمانی تأخیر استفاده شدند و در مرحله‌ی آموزش مدل‌های GEP و SVM در پیش‌بینی آب‌دهی روزانه ایستگاه موشنگ بهترین عملکرد را داشتند. در تبیین نتایج شکل ۸، می‌توان گفت که معادله‌ی به‌دست آمده از برنامه‌ریزی بیان ژن از ترکیب تصادفی مجموع تابع‌ها به‌دست می‌آید. همچنین اگر رابطه‌ی میان ورودی‌ها و

خروجی‌ها خطی باشد و از عملگرهای \sin ، \cos و... در استخراج رابطه استفاده شود، دقت مدل کاهش می‌یابد. در این پژوهش برای سهولت و افزایش دقت از عمل‌های چهارگانه‌ی ریاضی استفاده شد. در تبیین نتایج شکل ۸، می‌توان گفت که عملکرد ماشین بردار پشتیبان بر مبنای اصل استقرای کمینه‌سازی خطای ساختاری است. از این رو در شبیه‌سازی، با به‌کارگیری روش یادگیری با نظارت در تابع‌های RBF باعث شد این روش در برآورد سنج، سرعت بیشتر و خطای کمتری در مقایسه با دیگر تابع‌های کرنل داشته باشد. در این ایستگاه نتایج مدل SVM براساس معیار ارزیابی کلینگ-گوپتا و نش-ساتکلیف مطابقت بهتری داشت و کارایی آن در پیش‌بینی آب‌دهی روزانه خیلی خوب بود. از این رو، استفاده از مدل ماشین بردار پشتیبان و برنامه‌ریزی بیان ژن، به‌علت دقت خیلی خوب و نیز توانایی در ایجاد رابطه‌ی مطلوب میان سنج‌های ورودی و خروجی برای شبیه‌سازی میانگین و حداکثر آب‌دهی روزانه در ایستگاه موشنگ پیشنهاد می‌شود.

فهرست منابع

- Abdollahi S, Raeisi J, Khalilianpour M, Ahmadi F, Kisi O. 2017. Daily mean stream flow prediction in perennial and non-perennial rivers using four data driven techniques. *Water Resources Management*. 31(15): 4855-4874. <https://doi.org/10.1007/s11269-017-1782-7>.
- Ahmadi F, Radmanesh F, Mirabbasi R. 2016. Comparison between genetic programming and support vector machine methods for daily river flow forecasting (Case study: Barandoozchay River). *Water and Soil*. 28(6): 1162-1171. (In Persian). <https://doi.org/10.22067/jsw.v0i0.32406>.
- Ahmadyousefi S, Bahreman A, Sheikh V, Komaki CB. 2019. Determining the snow coefficient in order to simulate the snow melting in the Shemshak watershed using the WetSpa model. *Iranian Journal of Watershed Management Science*. 13(47): 1-8. (In Persian). <https://doi.org/20.1001.1.2008.9554.1398.13.47.3.1>
- Alinezhadi M, Mousavi SF, Hosseini K. 2021. Comparison of gene expression programming (GEP) and parametric and non-parametric regression methods in the prediction of the mean daily discharge of Karun river (A case study: Mollasani hydrometric station). *Journal of Water and Soil Science*. 25(1): 43-62. (In Persian). <https://doi.org/10.47176/jwss.25.1.1012>
- Azinmehr M, Bahreman A, Kabir A. 2016. Parameter sensitivity and uncertainty analysis of the model WetSpa in the flow hydrograph simulation using PEST in Dinvar Basin Karkheh. *Journal of Watershed Management Research*. 7(1): 82-72. (In Persian). <https://doi.org/10.18869/acadpub.jwmr.7.13.82>.
- Bahreman A, Corluy J, Liu Y, De Smedt F, Poórová J, Velcická L. 2005. Stream flow simulation by WetSpa model in Hornad river basin Slovakia. *Floods from defence to management*. London: Taylor-Francis Group. pp. 67-74.
- Bayazidi M, Asadzadeh F, Kaki M. 2018. The application of intelligent techniques for predicting daily flow at Telvar basin river. *Iranian journal of Ecohydrology*. 5(1): 203-213. (In Persian). <https://doi.org/10.22059/ije.2018.223015.386>.
- Behmanesh J, Mostafavi S, Zamanzad Ghavidel S. 2017. Use of soft calculations at estimation and prediction of environmental flow discharge (Case study: Khorkhoreh Chay river). *Journal of Civil and Environmental Engineering*. 47(3): 9-22. (In Persian).
- Dehghani R, Yonesi H, Torabi Poudeh H. 2017. Comparing the performance of support vector machines gene expression programming and bayesian networks in predicting river flow (Case study: Kashkan River). *Journal of Water and Soil Conservation*. 24(4): 161-177. (In Persian). <https://doi.org/10.22069/jwfs.2017.12398.2701>.
- Dehghani R, Torabi H, Younesi H, Shahinejad B. 2021. Application of wavelet support vector machine (WSVM) model in predicting river flow (Case study: Dez Basin). *Watershed Engineering and Management*. 13(1):98-110. (In Persian). <https://doi.org/10.22092/ijwmse.2020.128735.1748>.
- Delafrouz H, Ghaheeri A, Ghorbani MA. 2018. A novel hybrid neural network based on phase space reconstruction technique for daily river flow prediction. *Soft Computing*. 22(7):2205-2215. <https://doi.org/10.1007/s00500-016-2480-8>.
- Fatemi SE, Darabi Cheghabaleki S, Hafezparast M. 2022. The effect of preprocessing and reducing the input dimensions of the flow prediction model on optimized support vector regression by genetic algorithm. *Advanced Technologies in Water Efficiency*. 1(1): 24-47. (In Persian). <https://doi.org/10.22126/atwe.2021.6660.1002>.

- Ghorbani MA, Khatibi R, Goel A, FazeliFard MH, Azani A. 2016. Modeling river discharge time series using support vector machine and artificial neural networks. *Environmental Earth Sciences*. 75(8): 1-13. <https://doi.org/10.1007/s12665-016-5435-6>.
- Ghorbani MA, Khatibi R, Karimi V, Yaseen ZM, Zounemat-Kermani M. 2018. Learning from multiple models using artificial intelligence to improve model prediction accuracies: application to river flows. *Water Resources Management*. 32(13): 4201-4215. <https://doi.org/10.1007/s11269-018-2038-x>.
- Hussain D, Khan AA. 2020. Machine learning techniques for monthly river flow forecasting of Hunza River Pakistan. *Earth Science Informatics*. 13(3): 939-949. <https://doi.org/10.1007/s12145-020-00450-z>.
- Isazadeh M, Biazar S, Ashrafzadeh A, Khanjani R. 2019. Estimation of aquifer qualitative parameters in Guilans Plain using gamma test and support vector machine and artificial neural network models. *Journal of Environmental Science and Technology*. 21(2):1-21. (In Persian) <https://doi.org/10.22034/jest.2019.13946>.
- Khashei-Siuki A, Sarbazi M. 2015. Evaluation of ANFIS ANN and geostatistical models to spatial distribution of groundwater quality (Case study: Mashhad Plain in Iran). *Arabian Journal of Geosciences*. 8:903-912. <https://doi.org/10.1007/s12517-013-1179-8>.
- Khodakhah H, Aghelpour P, Hamed Z. 2022. Comparing linear and non-linear data-driven approaches in monthly river flow prediction based on the models SARIMA LSSVM ANFIS and GMDH. *Environmental Science and Pollution Research*. 29(15): 21935-21954. <https://doi.org/10.1007/s11356-021-17443-0>.
- Nivesh S, Negi D, Kashyap PS, Aggarwal S, Singh B, Saran B, Sihag P. 2022. Prediction of river discharge of Kesinga sub-catchment of Mahanadi basin using machine learning approaches. *Arabian Journal of Geosciences*. 15(16): 1-19. <https://doi.org/10.1007/s12517-022-10555-y>.
- Nozari H, Tavakoli F. 2018. Stream flow prediction using support vector machine based on discharge and precipitation time series on upstream stations (Case study: Taleh Zang hydrometric station). *Modeling in Engineering*. 16(54): 95-104. (In Persian) <https://doi.org/10.22075/jme.2017.11363.1112>.
- Pandhiani SM, Sihag P, Shabri AB, Singh B, Pham QB. 2020. Time-series prediction of stream flows of Malaysian rivers using data-driven techniques. *Journal of Irrigation and Drainage Engineering*. 146(7): 04020013. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001463](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001463).
- Rahimi B, Hafezparast Mavaddat M. 2020. Comparison of SVM GEP and IHACRES models in prediction of runoff changes due to climate change (Case study: Jamishan Dam). *Iranian Journal of Soil and Water Research*. 51(10): 2483-2499. (In Persian). <https://doi.org/10.22059/ijswr.2020.303779.668640>.
- Sahoo A, Samantaray S, Ghose DK. 2021. Prediction of flood in Barak River using hybrid machine learning approaches: a case study. *Journal of the Geological Society of India*. 97(2):186-198. <https://doi.org/10.1007/12594-021-1650-1>.
- Salarijazi M, Ghorbani K, Sohrabian E, Abdolhosseini M. 2016. Prediction of daily stream-flow using data driven models. *Iranian Journal of Irrigation and Drainage*. 10(4):479-488. (In Persian).
- Samadi M, Fathabadi A. 2019. Application of time series ANN and SVM models in forecasting the Gorgan dam inflow rate. *Environment and Water Engineering*. 4(4): 299-309. (In Persian). <https://doi.org/10.22034/jewe.2018.128256.1256>.
- Sattari M, Rezazadeh A, Safdari F, Ghahramanian F. 2016. Performance evaluation of M5 tree model and support vector regression methods in suspended sediment load modeling. *Water and Soil Resources Conservation*. 6(1): 109-124. (In Persian).

- Shaofu M, Al-Juboori AM, Alwan AH, Abdel-Salam ASG. 2021. On the investigation of monthly river flow generation complexity using the applicability of machine learning models. *Complexity*. pp. 1-14. <https://doi.org/10.1155/2021/3721661>
- Sharifi Garmdareh E, Vafakhah M, Eslamian S. 2019. Assessment the performance of support vector machine and artificial neural network systems for regional flood frequency analysis (a case study: Namak Lake watershed). *Journal of Water and Soil Science*. 23(1): 351-366. (In Persian). <https://doi.org/10.29252/jstnar.23.1.26>.
- Solgi A, Zareie H, Golabi M. 2017. Performance assessment of gene expression programming model using data preprocessing methods to modeling river flow. *Journal of Water and Soil Conservation*. 24(2):185-201. (In Persian) <https://doi.org/10.22069/jwfs.2017.11353.2573>.
- Taylor KE. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*. 106(D7): 7183-7192. <https://doi.org/10.1029/2000JD900719>.
- Zeinalie M, Golabi M, Niksokhan M, Sharifi M. 2020. Modeling daily river flow using simulator meta-models (Case study: Gamasiab River). *Journal of Environmental Science and Technology*. 22(4): 121-133. (In Persian). <https://doi.org/10.22034/jest.2020.33087.4089>.



Prediction of Ferizi River-Flow Using Data-Driven Models

Saber Jamali¹, Fereshte Rahimi Aghcheshme², Mohammad Javad Amiri^{3*}

1- PhD Candidate, Department of Water Science and Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

2- Ph.D. Candidate, Department of Range and Watershed management, Faculty of Natural Resource and Environment, Ferdowsi University of Mashhad, Mashhad, Iran

3- Associate Professor, Department of Water Science and Engineering, Faculty of Agriculture, Fasa University, Fasa, Iran

Extended Abstract

Introduction and Objective

Rivers serve as crucial sources of freshwater on Earth, and precise prediction of river flows plays a vital role in effective water resource management. To address this need, researchers have proposed various models. These models can be categorized into two types: (1) hydrological/hydraulic physically based models and (2) data-driven models. The cited papers provide evidence of the significance of employing data-driven models for hydrological issues, particularly in predicting river discharge.

Materials and Methods

In this study, the researchers utilized a gene expression programming model (GEP) and support vector machine (SVM) to predict the discharge of the Ferizi River. The dataset used in the study covered a period of ten years (2011-2020) and consisted of daily Ferizi river discharge readings obtained from the Moushang hydrometry station. The study area had an average altitude of 2171 meters above sea level, with a longitude range of 58° 49' 30" to 59° 4' 30" E and a latitude range of 36° 20' 1" to 36° 32' 1". For the prediction models, daily river discharge data from 1 to 5 days ahead were utilized as input variables for both the GEP and SVM models. To ensure data homogeneity and address any deficiencies, the researchers employed the run test and calculated the correlation coefficient between neighboring stations. Subsequently, the dataset was randomly divided into two groups: 80% for model training and 20% for model testing, as well as evaluating the modeling error. The performance of

Article Type: Research Article

*Corresponding Author E-mail: mj_amiri@fasau.ac.ir

Citation: Jamali, S., Rahimi Aghcheshme, F., Amiri M.j. 2024. Prediction of Ferizi river-flow Using Data-Driven Models. *Watershed Management Research*. 37(1): 81-97.

DOI: 10.22092/WMRJ.2023.360554.1504

Received: 29 November 2022, **Received in revised form:** 03 January 2023, **Accepted:** 19 March 2023,

Published online: 22 December 2023

Watershed Management Research, VOL. 36, No.4, Ser. No: 141, Winter 2024, pp. 81-97.

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



the models was assessed during the training and validation stages using various evaluation metrics, including root mean square error (RMSE), coefficient of correlation (R), bias, Kling Gupta (KGE), and Nash–Sutcliffe model efficiency (NSE). Additionally, in order to estimate the inflow to the Ferizi River using the SVM model, three common types of kernel functions in hydrology were examined, namely linear kernel, polynomial, and radial basis functions.

Results and Discussion

Among the different functions considered, the radial basis function was selected due to its lower error compared to other functions when applied to the variables. The best input model, Model No. 5, incorporated previous flow variables with a delay of five time steps. In the training phase, both the GEP and SVM models exhibited the highest performance in forecasting the daily flow of Moshang station. The performance of the applied models suggests that SVM (RSME = 1.15, R = 0.985, NSE = 0.85, and KGE = 0.79) demonstrates higher precision in the validation stage for river discharge prediction compared to GEP (RSME = 1.65, R = 0.964, NSE = 0.78, and KGE = 0.69).

Conclusion and Suggestions

This study has demonstrated that the utilization of soft computing techniques, such as SVM and GEP, is a powerful tool in predicting river flow. These techniques are capable of establishing a favorable relationship between input and output parameters, enabling accurate simulation of average and maximum daily flow.

Keywords: Artificial intelligence, Ferizi watershed, gene expression, runoff, soft computing