

بررسی و مقایسه عملکرد ۴ روش مدل سازی LS-SVM، NN، GEP و ANFIS-PSO در شبیه سازی بارش - رواناب (منطقه مورد مطالعه: هلیل رود - سد جیرفت)

محدثه کاوسی^۱، حسین خزیمه نژاد^۲

تاریخ ارسال: ۱۳۹۸/۰۶/۲۲

تاریخ پذیرش: ۱۳۹۸/۱۱/۱۸

مقاله پژوهشی

چکیده

شبیه‌سازی و پیش بینی روند بارش-رواناب یکی از مهمترین زمینه‌های پژوهشی در هیدرولوژی است که در مدیریت و برنامه ریزی منابع آب رودخانه‌ها و مخازن سد و همچنین حفاظت از آن‌ها در زمان وقوع سیلاب نقش بسزایی دارد ولی به دلیل پیچیدگی‌های پارامترهای هیدرولوژیکی، انجام این مساله با چالش‌هایی رو به روست. در این مطالعه، از چهار مدل مبتنی بر داده LS-SVM، NN، GEP و ANFIS-PSO و همچنین از داده‌های ۷ ایستگاه هواشناسی بالادست سد جیرفت در استان کرمان در بازه زمانی ۱۳۷۸ تا ۱۳۹۴ برای شبیه سازی فرآیند تخمین دبی ورودی به سد جیرفت در استان کرمان استفاده شد. ابتدا با آزمون پیرسون ترکیب مناسبی از ورودی‌ها انتخاب و سپس شبیه سازی با هر روش صورت پذیرفت. نتایج پژوهش بر اساس دو معیار ضریب همبستگی (R)، ریشه متوسط مربعات خطا (RMSE) نشان داد که روش ANFIS-PSO با مقدار RMSE برابر ۰/۴۲ مترمکعب بر ثانیه و $R^2 = ۰/۹۵$ در بخش آموزش و RMSE و R^2 به میزان ۰/۹۷ مترمکعب بر ثانیه و ۰/۸۰ در بخش صحت سنجی بهترین روش از بین مدل‌های مورد بررسی بود و در هر دو بخش آموزش و صحت سنجی دقت بالاتری نسبت به سه مدل دیگر داشت. همچنین این مدل در شبیه سازی دبی‌های بیشینه نیز بهتر از سایر روش‌ها عمل نمود و خطای نسبی محاسبات آن بین ۳۳/۰۹ تا ۵۴/۲ برای دبی‌های مختلف به دست آمد.

واژه‌های کلیدی: بارش-رواناب، ANFIS-PSO، GEP، NN، LS-SVM

^۱ دانشجوی دکتری رشته منابع آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران.

Email: mohadesekavusi@gmail.com

^۲ دانشیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران.

Email: hkhozeymeh@birjand.ac.ir (نویسنده مسئول)

مقدمه

شبیه سازی جریان رودخانه جهت مدیریت و برنامه ریزی منابع آب در رودخانه‌ها و مخازن سدها و همچنین به منظور حفاظت از کناره‌های رودخانه در زمان وقوع سیلاب حائز اهمیت است. از سوی دیگر، یکی از نیازهای عمده در طراحی پروژه‌های مدیریت منابع آب و کنترل سیلاب، مقدار رواناب ناشی از بارندگی است که جهت برآورد آن از روش‌های مختلف استفاده می‌شود. با توجه به بروز سیلاب‌های گوناگون و خطرهای ناشی از رخداد آن که جامعه انسانی و سازه‌های سر راه خود را تهدید می‌کند، بررسی فرآیند بارش-رواناب^۱ و اطلاع از آن نمود^۲ جریان رودخانه‌ها و حوضه‌ها از اهمیت ویژه‌ای برخوردار است (علیزاده، ۱۳۸۵). هدف نهایی از مدل سازی بارش رواناب یافتن راه حلی با دقت مناسب برای برآورد رواناب بارش در یک حوضه آبریز است؛ ولی از آنجا که بارش-رواناب فرآیندی پیچیده و غیر خطی است، تغییرات مکانی و زمانی وسیعی دارد و به سادگی مدل سازی نمی‌شود. به طور کلی سه نگرش برای مدل سازی بارش رواناب ارائه شده است: تئوریک، مفهومی و سیستمی. دو مدل اول نیازمند اطلاعات پیچیده و وسیعی هستند، در حالی که روش سوم با توجه به عملکرد ساده تر، می‌تواند گزینه مناسبی در بحث مدیریت و تصمیم گیری به شمار آید (سعیدی، ۲۰۱۴).

روش‌های هوشمند یکی از روش‌های سیستمی بوده که تا کنون پژوهش‌های متعددی در همین راستا به انجام رسیده است نیک پور و همکاران (۱۳۹۵)، به شبیه‌سازی بارش-رواناب رودخانه خیاوچای با استفاده از مدل‌های برنامه‌ریزی بیان ژن، ماشین بردار پشتیبان و هیبرید موجک عصبی پرداختند. نتایج نشان از برتری مدل هیبرید موجک-عصبی نسبت به مدل‌های دیگر دارد. Aytok & Alp (2008)، با مدل‌های ANN و GEP بارش-رواناب روزانه رودخانه جونیتا در ایالت پنسیلوانیای آمریکا را شبیه‌سازی کردند. نتایج پژوهش، برتری مدل GEP را نسبت به ANN در تخمین رواناب روزانه رودخانه یادشده نشان داد. همچنین Misra et al. (2009)، در پژوهشی به

شبیه سازی رواناب و میزان رسوب روزانه، هفتگی و ماهانه در یک حوضه آبخیز در هند با استفاده از SVM اقدام نمودند و نتایج را با مقادیر به دست آمده از روش ANN مقایسه کردند. نتایج بیانگر قدرت SVM در آموزش، کالیبراسیون و اعتبارسنجی در مقایسه با ANN بود که در پیش بینی رواناب و رسوب دقت قابل قبولی داشت. Nayak et al. (2013)، برای مدل سازی بارش-رواناب و استخراج هیدروگراف‌های سیلاب حوضه آبریز مالاپرابها واقع در هندوستان از آمار روزانه دبی، بارش و تبخیر استفاده کردند. در پژوهش آن‌ها، مدل موجک-عصبی نسبت به شبکه‌های عصبی مصنوعی عملکرد بهتری نشان داد. Badrzadeh et al (2015)، عملکرد سه مدل ANN، ANFIS و WNN را در تخمین رواناب روزانه حوضه آبریزی واقع در استرالیا را مقایسه کردند. همچنین به منظور مدیریت سیلاب‌ها از داده‌های بارش-رواناب ساعتی دو ایستگاه بالادست و پایین دست در حوضه آبریز یادشده استفاده کردند. در نهایت، مدل WNN عملکرد بهتری را در مدل سازی بارش-رواناب و پیش‌بینی هیدروگراف‌های سیلاب به همراه داشت. نورانی و همکاران (۲۰۱۶)، برای اولین بار، کاربرد شبکه‌های عصبی عاطفی (هیجانی)^۳ را در مدل سازی بارش-رواناب مطرح کرد. منطقه‌ای که او مطالعه کرد، دو حوضه آبریز با شرایط اقلیمی متفاوت در کشور فرانسه بود. نتایج نشان داد استفاده از مدل یادشده سبب بهبود دقت نتایج در مراحل آموزش و صحت سنجی نسبت به شبکه‌های عصبی کلاسیک می‌شود. پوررضا بیلندی و همکاران (۱۳۹۳)، در پژوهشی به پیش‌بینی رواناب روزانه با مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM) پرداختند. نتایج نشان از برتری نسبی روش ماشین بردار پشتیبان به شبیه سازی مدل مفهومی Hymod دارد. در پژوهشی دیگر، باقرپور و همکاران (۱۳۹۷)، به پیش‌بینی رواناب با استفاده از مدل‌های جعبه سیاه و خاکستری پرداختند. مقایسه دقیق عملکرد کل مدل‌ها نشان داد که مدل‌های نروفازی و ماشین بردار پشتیبان رواناب را در

¹ Rainfall-Runoff

² Hydrograph

³ Emotional neural network

وقوع پدیده کاویتاسیون توانایی لازم را دارا می‌باشد. سلگی و همکاران (۱۳۹۶)، به بررسی عملکرد مدل برنامه‌ریزی بیان ژن با روش‌های پیش‌پردازش داده‌ها جهت مدل سازی جریان رودخانه گاماسیاب در مقیاس روزانه و ماهانه پرداختند. نتایج نشان داد که عملکرد مدل برنامه‌ریزی بیان ژن در دوره روزانه خوب ولی در دوره ماهانه با کاهش دقت همراه می‌باشد. همچنین مدل ترکیبی برنامه‌ریزی بیان ژن-موجک عملکرد بهتری نسبت به مدل ساده در هر دو دوره زمانی روزانه و ماهانه دارد. کیا و همکاران (۱۳۹۸)، کارایی سیستم فازی-عصبی را در برآورد رواناب ناحیه کوهستانی حوضه هراز مورد ارزیابی قرار دادند. نتایج نشان دهنده برتری روش ANFIS نسبت به MLR است. در مطالعه حاضر، پس از انتخاب ترکیب مناسبی از داده‌ها با استفاده از آزمون آماری که شامل چهار پارامتر دبی‌ها در زمان قبل، دما، بارش و تبخیر است، کارایی چهار روش هوشمند $LS-SVM^1$ ، NN^2 ، GEP^3 و $ANFIS-PSO^4$ که شامل دو روش ترکیبی و دو روش مجزاست، مقایسه و توانایی آن‌ها با یکدیگر در شبیه‌سازی میزان دبی ورودی به سد در یک منطقه خشک که تاکنون صورت نگرفته بود، بررسی گردید. در انتها نیز ارزیابی مدل‌ها در پیش بینی دبی‌های بیشینه نشان داده شد.

مواد و روش‌ها

منطقه مورد مطالعه

استان کرمان در جنوب شرقی فلات مرکزی و بین ۵۳ درجه و ۲۶ دقیقه تا ۵۹ درجه و ۲۹ دقیقه طول شرقی و ۲۵ درجه و ۵۵ دقیقه تا ۳۲ درجه عرض شمالی قرار داشته با مساحتی در حدود ۱۸۳۱۹۳ کیلومترمربع، وسیع‌ترین استان کشور به حساب می‌آید که بیش از ۱۱ درصد از خاک ایران را در بر گرفته است. حدود ۸۰٪ مساحت استان که در شرایط آب و هوایی خشک و فرا خشک قرار گرفته است. سد جیرفت در ۴۰ کیلومتری شمال غربی شهر

فصل‌های گرم با دقت کمتری نسبت به فصل‌های سرد پیش‌بینی کرده‌اند. Botsis et al. (2011)، بارش-رواناب روزانه حوضه آبخیزی واقع در کالیفرنیا را شبیه‌سازی کردند. مدل SVM با سه نوع تابع کرنل با مدل ANN مقایسه شد. در نهایت، مدل ماشین بردار پشتیبان شبیه سازی دقیق‌تری از بارش-رواناب داشت. صادقی و پوررضا بیلندی (۱۳۹۲)، در پژوهشی روش‌های بهینه‌سازی فراکوشی در تحلیل عدم قطعیت پارامترهای مدل مفهومی بارش-رواناب را با هم مقایسه کردند. نتایج نشان داد الگوریتم‌های PSO و SCE کارایی بهتری در تحلیل عدم قطعیت برآورد پارامترهای مدل بارش-رواناب نسبت به دو الگوریتم دیگر داشتند. اسحق تیموری و همکاران (۱۳۹۱)، به شبیه سازی فرآیند بارش-رواناب با استفاده از مدل wms پرداختند و این مدل را کارآمد معرفی نمودند. قربانی و همکاران (۱۳۹۵)، جریان ماهانه را در یک حوضه فاقد آمار در استان گلستان با استفاده از پارامترهای اقلیمی و فیزیوگرافی برآورد نمودند. با توجه به نتایج به‌دست آمده، برآورد جریان در ماه‌های پربارش نسبت به ماه‌های کم بارش در مدل رگرسیون چند متغیره و درخت تصمیم، از دقت بالاتری برخوردار بود. همچنین مدل درخت تصمیم نسبت به رگرسیون چند متغیره در تمام موارد نتایج بهتری داشت. در پژوهشی دیگر ناظری تهرودی و همکاران (۱۳۹۵)، رواناب ماهانه رودخانه نازلوچای واقع در غرب دریاچه ارومیه را با استفاده از داده‌های بارش و دبی با تاخیرهای یک تا پنج ماه بررسی کردند. در پژوهش آن‌ها مدل ANFIS نسبت به مدل‌های GP و SVM نتیجه بهتری به همراه داشت. قربانی و دهقانی (۱۳۹۵)، کارایی سه مدل شبکه عصبی بی‌زین (BNN)، SVM و GEP را در تخمین رواناب ماهانه رودخانه کاکارضا مورد مقایسه قرار دادند. از میان سه مدل مذکور، BNN بهترین نتیجه را به همراه داشت. امامقلی زاده و کریمی دمنه (۱۳۹۶)، در پژوهشی روش برنامه‌ریزی بیان ژن را در تعیین ضریب هوادهی دریچه تخلیه‌کننده تحتانی چهار سد البرز، ژاوه، گتوند علیا و جره مورد استفاده قرار داده و نشان دادند روش برنامه‌ریزی بیان ژن در تعیین ضریب هوادهی دریچه‌های تخلیه‌کننده تحتانی سدها به‌منظور جلوگیری از

¹ Least-squares support-vector machine

² Neural Networks

³ Gene Expression Programming

⁴ Adaptive neuro fuzzy inference system -Particle swarm optimization

انتخاب ترکیب مناسب

به منظور انتخاب ترکیبی مناسب از پارامترهای ورودی به مدل‌ها جهت افزایش دقت مدل سازی دبی که بازه هر یک در جدول (۲) آورده شده است، از آزمون‌های آماری استفاده شد و مقادیر ضرایب خودهمبستگی متقابل بین دبی و سه پارامتر دما، بارش و تبخیر و همچنین دبی در زمان‌های قبل و همراه با تاخیر، با آزمون پیرسون در نرم افزار SPSS محاسبه شد. این مقادیر در جدول (۳) مورد اشاره قرار گرفته اند. در این جدول، پارامترهای R_0 ، R ، E و Q به ترتیب بیانگر مقادیر اندازه گیری شده دما، بارش، تبخیر و دبی در ایستگاه‌های مورد بررسی می‌باشد و اندیس اعداد، شاخص گام زمانی قبل از زمان حاضر (t) است. با توجه به مقادیر به دست آمده، الگوی $R_0-T_2-E_2-Q_1$ یعنی استفاده از بارش یک گام قبل، دمای دوگام قبل، تبخیر دوگام قبل و دبی یک گام قبل، به عنوان بهترین الگو انتخاب و در مدل سازی‌های مدنظر مورد استفاده قرار گرفت.

جدول (۲): بازه داده های ورودی

متغیر	مقدار کمینه	مقدار کمینه
دما	۳/۶	۲۹/۷
بارش	۰	۶۳۳/۵
تبخیر	۶۲/۱	۵۵۵/۸
دبی	۰/۱	۶۱/۱

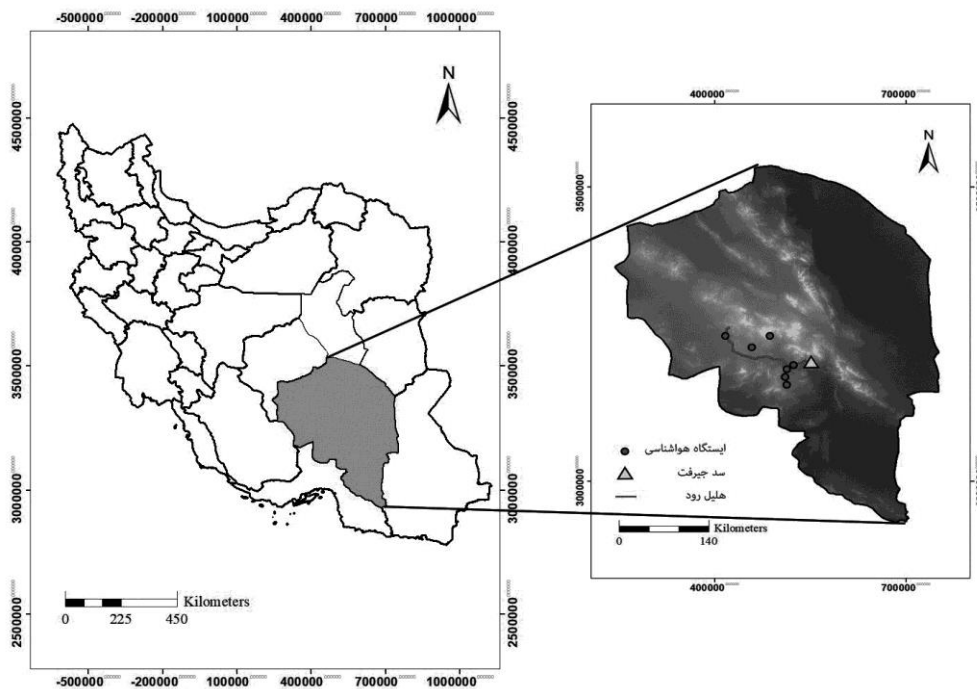
جیرفت، در محلی بنام تنگ نراب بر روی رودخانه هلیل رود احداث شده است و مساحت حوضه آبخیز بالا دست آن حدود هشتصد هزار هکتار می‌باشد.

در این پژوهش، از داده‌های هواشناسی ۷ ایستگاه بالادست سد جیرفت و نزدیک به رودخانه هلیل رود، شامل بارش، دما و تبخیر در بازه زمانی سالهای ۱۳۷۸ تا ۱۳۹۴ و همچنین داده‌های دبی از ایستگاه هیدرومتری نزدیک سد در همین بازه زمانی استفاده شد. در تمامی مدل سازی‌ها، ۷۰ درصد داده‌ها به عنوان داده‌های آموزش و ۳۰ درصد به عنوان داده صحت سنجی در نظر گرفته شد. مشخصات این ایستگاه‌ها آن‌ها در جدول (۱) آمده است. همچنین موقعیت سد جیرفت و ایستگاه‌های هواشناسی مورد نظر در شکل (۱) نشان داده شده است.

جدول (۱): مشخصات ایستگاه‌های هواشناسی مورد نظر در

بالادست سد جیرفت

نام ایستگاه	ارتفاع (متر)	طول جغرافیایی	عرض جغرافیایی
کناروئیه	۱۴۵۶	۵۷-۱۵-۰۹	۲۸-۵۱-۲۸
صالح آباد	۱۹۰۰	۵۷-۰۹-۰۰	۲۸-۴۷-۰۰
نمزد	۱۸۸۰	۵۸-۲۴-۰۰	۲۷-۵۹-۰۰
اسماعیل آباد	۱۸۰۰	۵۶-۵۹-۰۰	۲۹-۱۵-۰۰
دولت آباد	۱۰۷۰	۵۶-۳۹-۰۰	۲۸-۲۰-۰۰
فتح آباد	۲۱۳۰	۵۷-۱۳-۰۰	۲۸-۲۹-۰۰
بیدکردوئیه	۲۲۳۰	۵۶-۲۴-۲۲	۲۹-۱۸-۵۲



شکل (۱): موقعیت سد جیرفت و ایستگاههای هواشناسی مورد استفاده

جدول (۳): مقادیر همبستگی بین متغیرهای مورد بررسی و دبی

Q_2	Q_1	E_2	E_1	E_0	T_2	T_1	T_0	R_2	R_1	R_0	Q_t
۰/۴۸	۰/۸۴	-۰/۷۵	-۰/۵۰	-۰/۳۵	-۰/۷۳	-۰/۵۵	-۰/۳۳	۰/۳۲	۰/۴۸	۰/۸	

مجموعه داده‌های $\{x_i, y_i\}_{i=1}^N$ را خواهیم داشت. که $x_i \in R^n$ به عنوان بردار ورودی و $y_i \in R$ به عنوان بردار خروجی معرفی می‌شوند. تابع رگرسیون غیر خطی در مسائل تخمین به صورت رابطه (۱) می‌باشد:

$$y(x) = w^T \cdot \varphi(x) + b \quad (1)$$

که در آن: w ، b و T به ترتیب مقادیر وزن‌ها، بایاس تابع رگرسیون و نشان‌دهنده ترانهاده هستند. $\varphi(x)$ ترسیم غیر خطی ورودی‌ها در فضای ویژگی با ابعاد بالا است. حال مسئله رگرسیون غیرخطی بر اساس معادله بهینه‌سازی در رابطه (۲) حل می‌گردد:

$$\min_{w, e, b} j(w, e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 \quad (2)$$

در این پژوهش، از چهار مدل هوشمند جهت مدل سازی دبی ورودی به سد استفاده شد که مدل‌های LS-SVM، NN و ANFIS-PSO در محیط برنامه نویسی MATLAB و مدل GEP در نرم افزار GeneXproTools 5 اجرا گردیدند.

مدل LS-SVM

در سال ۱۹۹۹، سویکنس و وندوال مدل حداقل مربعات ماشین بردار پشتیبان را معرفی نمودند. این مدل محدودیت‌های یکسانی با روش SVM کلاسیک دارد اما نسبت به این روش از پیچیدگی محاسباتی کمتر و دقت و سرعت بالاتری برخوردار است. در واقع الگوریتمی پیشرفته تر از الگوریتم SVM می‌باشد. از این مدل هم در مسائل طبقه بندی و هم مسائل رگرسیونی استفاده شده است. چارچوب رگرسیون حداقل مربعات ماشین بردار پشتیبان به صورت زیر توضیح داده می‌شود:

در قسمت آموزش برای ثابت بودن شرایط در تعیین پارامترهای تابع تخمین، از روش بهینه سازی Simplex برای ۳ تابع کرنل انتخابی (Polynomial، Linear، RBF) استفاده گردید. تعداد الگوهای لازم برای آموزش مدل LS-SVM، ۷۵ درصد داده‌های موجود در نظر گرفته شد و از مابقی داده‌ها (۲۵ درصد داده‌ها) جهت صحت سنجی و ارزیابی مدل استفاده گردید. مقادیر متغیرهای این مدل در جدول (۴) آورده شده است.

جدول (۴): مقادیر بهینه پارامترهای ورودی مدل LS-SVM

پارامتر کرنل	پارامتر تنظیم کننده	تابع مورد استفاده
۱۰	۰/۶	RBF

مدل NN

این مدل، یک روش محاسباتی است که به کمک فرآیند یادگیری و با استفاده از پردازشگرهایی به نام نورون، نگاهی بین ورودی‌ها و خروجی‌ها ارائه می‌کند. پارامترهای موثر در مدل سازی شبکه عصبی شامل تعداد تکرار در مرحله آموزش، تعداد لایه‌های شبکه، تعداد نورون‌های هر لایه و نوع روش آموزش و تابع انتقال آن می‌باشد. به طور کلی، هرچه تعداد تکرار در آموزش بیشتر باشد، خطای شبیه سازی کمتر است. هرچند اگر تعداد تکرار از یک مقدار تجاوز کند، خطای دسته آزمایشی نیز به وجود می‌آید. پس یافتن تعداد بهینه تکرار و سایر پارامترهای مدل، می‌تواند از وقوع اینگونه خطاها جلوگیری نماید. در اینجا با اجرای ۵ تکرار با تعداد لایه‌ها، تابع انتقال و تعداد نورون‌های متفاوت بهترین مدل با توجه به نمایه‌های مورد نظر، انتخاب گردید. تعداد تکرار برای همه موارد ۱۰۰۰ در نظر گرفته شد و ۷۰ درصد داده‌ها جهت آموزش و ۳۰ درصد جهت صحت سنجی اعمال گردید.

مدل ANFIS-PSO

منطق فازی (FL) و سیستم استنتاج فازی (FIS) توسط لطفی زاده در سال ۱۹۶۵ مطرح شد که شیوه ای ساده برای تصمیم گیری براساس داده‌های مبهم یا ناقص است. سیستم استنتاج فازی از قوانین فازی، توابع عضویت فازی و عملیات فازی د-فازی تشکیل شده است که ساختار

بر اساس محدودیت رابطه (۳):

$$y_i = w^T \varphi(x_i) + b + e_i \quad (3)$$

که در آن: γ یا C پارامتر تنظیم کننده بخش خطا است و e میزان خطا را نشان می‌دهد. با حل آن با استفاده از شکل لاگرانژی از تابع هدف اصلی در رابطه (۴):

$$L(w, b, e; \alpha) = j(w, e) - \sum_{i=1}^N \alpha_i \{w^T \varphi(x_i) + b + e_i - y_i\} \quad (4)$$

α_i ضریب لاگرانژ است. بر اساس شرایط کان-تاکر (KKT) در رابطه (۵):

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^N \alpha_i \varphi_i(x_i)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i = 0$$

$$\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \quad (5)$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i = w^T \varphi(x_i) + b + e_i,$$

$$i = 1, \dots, N$$

بعد از حذف متغیرهای w و e راه حل از رابطه (۶)

تبعیت می‌نماید:

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \frac{I}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (6)$$

با استفاده از حقه کرنل در رابطه (۷):

$$\begin{aligned} \Omega_{il} &= \varphi(x_i)^T \varphi(x_l) \\ &= K(x_i, x_l) \quad i, l = 1, \dots, N \end{aligned} \quad (7)$$

پس مدل LS-SVM نتیجه شده، به منظور تابع تخمین به صورت رابطه (۸) نوشته می‌شود:

$$y(x) = \sum_{i=1}^N \alpha_i K(x, x_i) + b \quad (8)$$

در مرحله ورود داده‌ها به مدل LS-SVM، ترکیب بهینه پس از نرمال سازی داده‌ها مورد آموزش قرار گرفت.

(کروموزوم) رمزگذاری میشوند که بصورت غیرخطی و با شکل و اندازه متفاوت بصورت درخت بیان (expression trees) نمایان میگردند. ساختمان هر ژن به وسیله سر و دم آن کنترل میشود و این سازمان عملکردی و ساختاری ژن‌ها، GEP را برای ایجاد برنامه معتبر ضمانت میکند که حد اصلاح کروموزومها در آن اهمیت ندارد.

فرآیند گام به گام برنامه ریزی بیان ژن به صورت مراحل زیر است:

۱- یک جمعیت اولیه از توابع مرکب نشان دهنده مدل‌های پیش بینی، به صورت تصادفی در نظر گرفته می‌شود.

۲- هر یک از افراد جمعیت مذکور با استفاده از توابع برازش، مورد ارزیابی قرار می‌گیرد،

۳- هر تولید، مرحله‌ای را برای انتخاب یک جمعیت جدید دنبال می‌کند.

۴- گام سوم تا نیل به حداکثر تعداد تولید، تکرار خواهد شد.

در این روش در ابتدای فرآیند، هیچ گونه رابطه تابعی در نظر گرفته نشده و این روش قادر به بهینه سازی ساختار مدل و مولفه‌های آن می‌باشد.

در مدل سازی با روش بیان ژن نیز، ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد جهت صحت سنجی انتخاب شدند. همچنین با توجه به سعی و خطای صورت گرفته، توابع برگزیده شامل چهار عمل اصلی، لگاریتم، ln، سینوس و کسینوس، رادیکال و توان ۲ به عنوان توابع اولیه کار در این روش مورد استفاده قرار گرفتند. پارامترهای استفاده شده و نرخ آن‌ها در جدول (۶) آورده شده است.

نرمال سازی داده‌ها

به دلیل دامنه تغییرات زیاد ورودی‌ها و بالابردن دقت و سرعت مدل سازی، پیش از شروع مدل سازی، از رابطه (۹) جهت نرمال سازی داده‌ها استفاده شد:

$$X_n = 0.1 + 0.8 \times \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (9)$$

اصلی آن را تشکیل می‌دهند. ANFIS یک سیستم استنتاج فازی است که با استفاده از روش پس انتشار، خطا را به حداقل می‌رساند. روش کار ANFIS مشابه هر دو روش شبکه عصبی و منطق فازی است که به صورت ترکیبی به مدل سازی و کاهش خطا در آن می‌پردازد.

بهینه سازی ازدحام ذرات (PSO) یک الگوریتم بهینه سازی مبتنی بر جمعیت و الهام گرفته شده از رفتار پرندگان و ماهی‌ها است. در این الگوریتم، هر راه حل فقط یک ذره در فضای جستجو است و همه ذرات یک مقدار شایستگی دارند که توسط تابع هدف که باید بهینه شود، ارزیابی می‌شود. هر ذره i یک موقعیت X در فضای d بعدی مسئله دارد و سرعت آن V است که حرکت را هدایت می‌کند. ذره برای تثبیت بهترین موقعیت پیشین خود در هر تکرار، از حافظه‌ای با عنوان بردار P استفاده می‌کند.

در هر تکرار، ذره با داشتن دو مقدار با عنوان G -best و P -best که به معنی بهترین تجربه شخصی و گروهی می‌باشند، به روزسانی می‌شود. این فرآیند براساس رسیدن به یک مقدار خاص یا یک تکرار معین ادامه می‌یابد. در PSO پارامترهای $C1$ و $C2$ عامل‌های یادگیری و قابل تنظیم هستند. پارامترای مدل ترکیبی برای مساله مورد نظر طبق جدول (۵) و براساس تکرار و سعی و خطا انتخاب شد:

جدول (۵) - مقادیر بهینه پارامترهای ورودی مدل ANFIS-

PSO			
	جمعیت اولیه	C1	C2
تکرار	۲۰۰	۲/۱	۱/۹
	۵۰۰		

مدل GEP

برنامه ریزی بیان ژنتیک (GEP) که اولین بار توسط فریرا در سال ۱۹۹۹ معرفی شد، همانند GA و GP بصورت الگوریتم‌های ژنتیکی است. این روش نیز از افراد مجزا در جمعیت استفاده میکند و آنها را بر اساس برازش انتخاب کرده و بر روی آنها با استفاده از یک یا چند عملگر ژنتیکی تغییرات ژنتیکی اعمال مینماید. این الگوریتم، جزو (ژولش)های الگوریتم گردشی محسوب میشود که مبنای تمامی آنها براساس نظریه تکامل داروین استوار است. در GEP افراد بصورت ردیفهای علامتدار با طول ثابت

که در آن X_n نرمال شده اعداد، X_{max} مقدار بیشترین داده و X_{min} مقدار کمترین داده هستند.

جدول (۶): مقادیر بهینه پارامترهای ورودی مدل GEP

نرخ ترکیب تک نقطه ای	نرخ وارون سازی	نرخ جهش	تعداد ژن	اندازه راس	تعداد کروموزوم
۰/۳	۰/۱	۰/۰۳۵	۴	۹	۳۰

در این روابط، X_i و Y_i مقادیر مشاهداتی و محاسباتی می‌باشند.

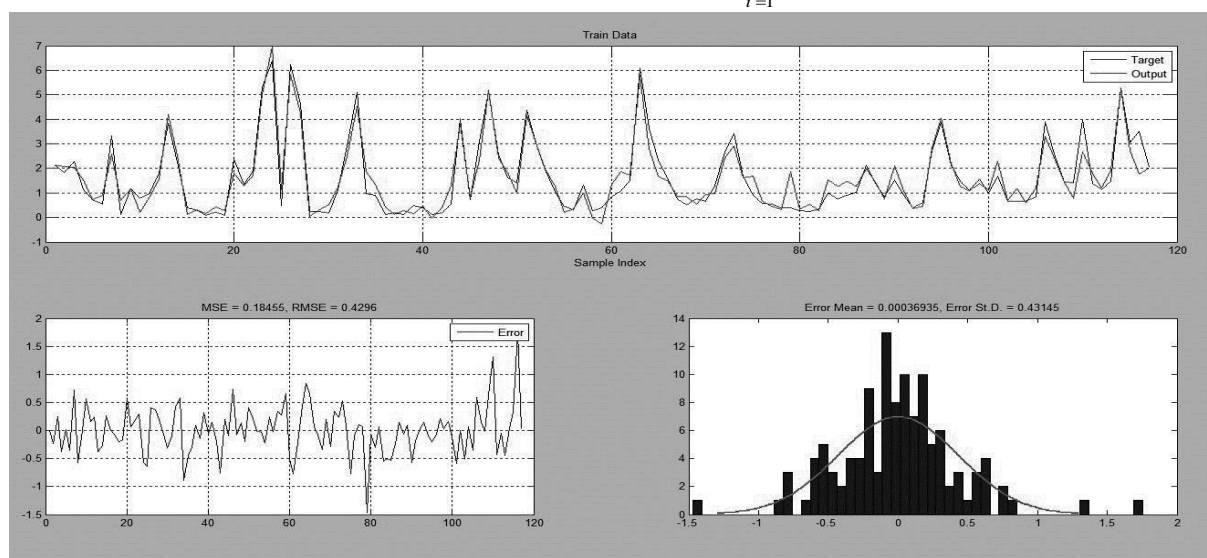
پس از اجرای مدل ANFIS-PSO، نتایج به صورت شکل‌های (۲) تا (۴) حاصل شد که در شکل‌های (۲) و (۳) روند تغییرات داده‌های مشاهداتی و محاسبه شده توسط مدل و همچنین میزان و پراکندگی خطاها نشان داده شده است. همینطور براساس جدول (۷) مدل بهینه ایجاد شده، در بخش آموزش دارای $RMSE = 0.42 m^3/s$ و $R^2 = 0.95$ و در بخش صحت سنجی به ترتیب دارای $RMSE$ و R^2 به میزان $0.97 m^3/s$ و 0.80 به دست آمد که مقادیر قابل قبولی می‌باشد.

نتایج و بحث

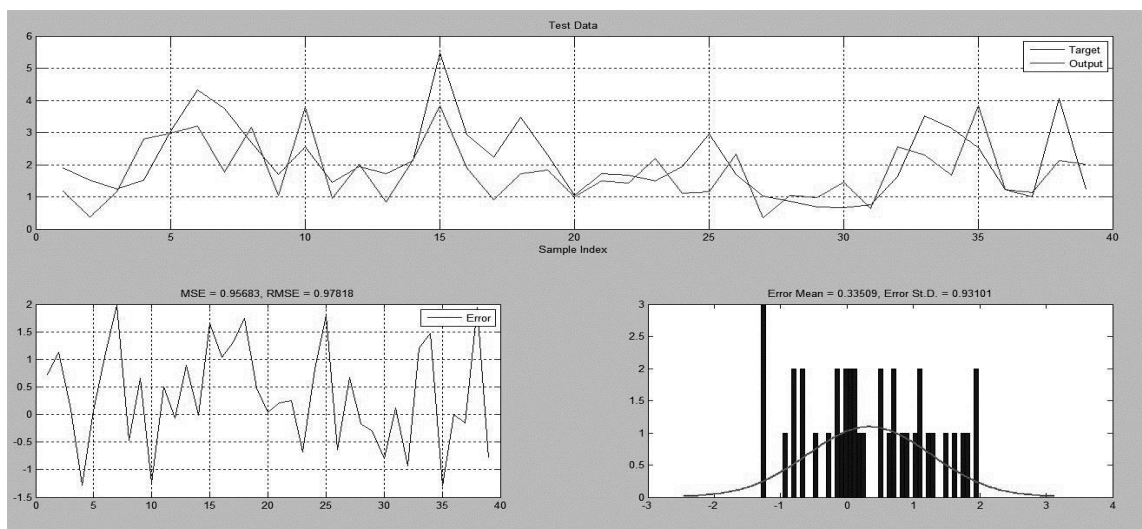
پس از حذف داده‌های پرت، جهت بالا رفتن دقت مدل سازی، داده‌ها به صورت نرمال تبدیل شد. در این پژوهش، به منظور ارزیابی دقت و کارایی مدل‌ها از ضریب همبستگی (R^2) و ریشه میانگین مربعات خطا (RMSE) استفاده گردید. روابط این نمایه‌ها در روابط ۱۰ و ۱۱ آورده شده است.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{N}} \quad (10)$$

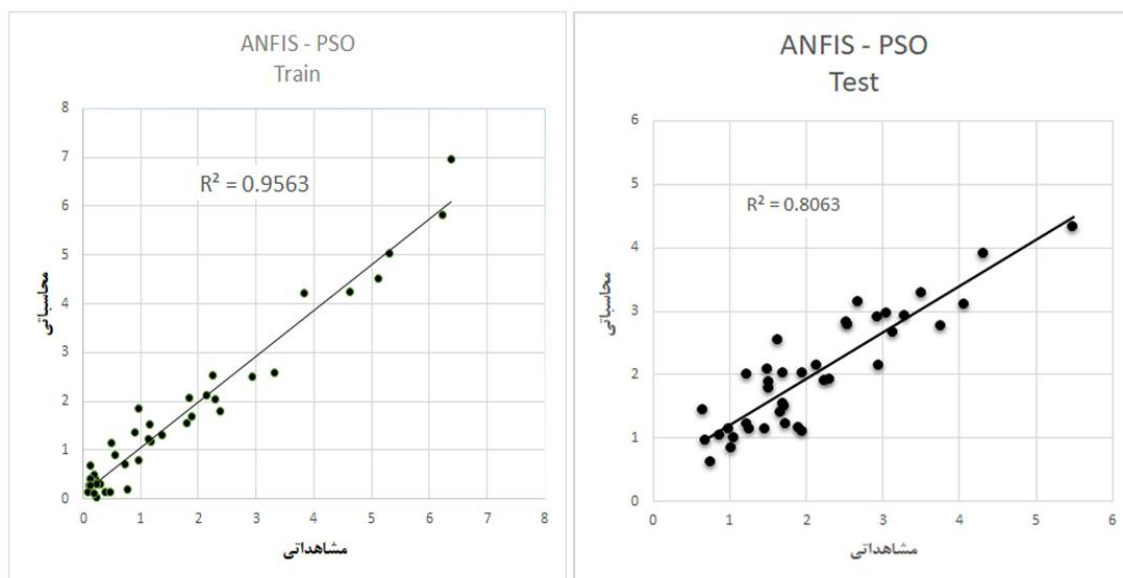
$$R^2 = \frac{(\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}))^2}{(\sum_{i=1}^n (X_i - \bar{X})^2)(\sum_{i=1}^n (Y_i - \bar{Y})^2)} \quad (11)$$



شکل (۲) - خروجی مدل ANFIS-PSO در بخش آموزش



شکل (۳): خروجی مدل ANFIS-PSO در بخش صحت سنجی



شکل (۴): دقت مدل سازی در روش ANFIS-PSO در بخش های آموزش و صحت سنجی

جدول (۷): مقادیر نمایه های آماری عملکرد مدل ANFIS-PSO در مراحل آموزش و صحت سنجی

تکرار	جمعیت اولیه	C1	C2	آموزش		صحت سنجی	
				R ²	RMSE(m ³ /s)	R ²	RMSE(m ³ /s)
۵۰۰	۲۰۰	۲/۱	۱/۹	۰/۹۵	۰/۴۲	۰/۸۰	۰/۹۷

همچنین جدول (۸) نشان می دهد با اجرای مدل LS-SVM مقدار R² و RMSE به ترتیب ۰/۷۸ و ۱/۰۲ m³/s و در بخش آموزش و ۰/۷۳ و ۱/۸ در بخش صحت سنجی به دست آمد.

جدول (۸): مقادیر نمایه های آماری عملکرد مدل LS-SVM در مراحل آموزش و صحت سنجی

تابع مورد استفاده	پارامتر تنظیم کننده	پارامتر کرنل	آموزش		صحت سنجی	
			R ²	RMSE(m ³ /s)	R ²	RMSE(m ³ /s)
RBF	۰/۶	۱۰	۰/۷۸	۱/۰۲	۰/۷۳	۱/۸

مخفی) به دست آمد به طوری که در هر اجرا تعداد تکرار ۱۰۰۰ در نظر گرفته شد. همینطور ساختار درختی مطلوب ایجاد شده توسط مدل GEP در شکل (۵) نشان داده شده است. براساس این نمودار، مدل از پارامترهای مانند بیشینه، کمینه، جذر و Ln برای رسیدن به مدل بهینه استفاده نموده است.

نتایج حاصل از روش‌های GEP و NN نیز در جدول (۹) و (۱۰) آورده شده است. براساس جدول (۱۰)، ترکیب بهینه ورودی به مدل شبکه عصبی، در ردیف ۵ آورده شده است که کمترین RMSE را در هر دو بخش و بیشترین R^2 را در بخش آموزش دارا می‌باشد. این ترکیب بهینه بعد از ۵ اجرای مختلف با پارامترهای ورودی متفاوت (تعداد لایه‌ها، تابع انتقال، روش آموزش و تعداد نرون‌های لایه

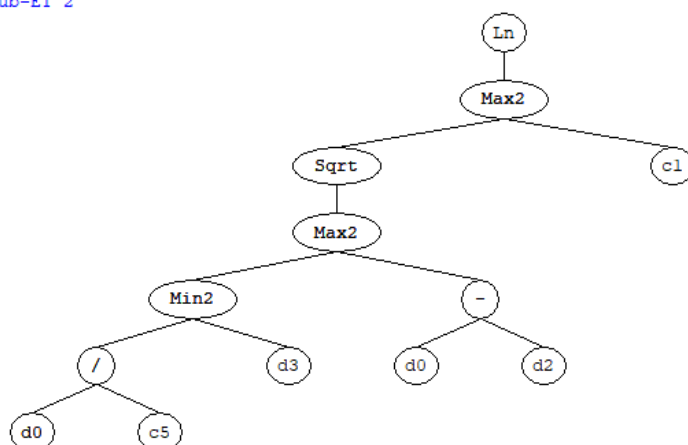
جدول (۹): مقادیر نمایه‌های آماری عملکرد مدل GEP در مراحل آموزش و صحت سنجی

تعداد	اندازه	تعداد ژن	نرخ	نرخ وارون	نرخ ترکیب	آموزش	صحت سنجی
کروموزوم	راس		جهش	سازی	تک نقطه‌ای	RMSE(m ³ /s)	RMSE(m ³ /s)
۳۰	۹	۴	۰/۰۳۵	۰/۱	۰/۳	۰/۷۵	۱/۹۵
						۱/۶	۰/۵۵
						R^2	R^2

جدول (۱۰): انتخاب پارامترهای مناسب در مدل NN با توجه به مقادیر نمایه‌های آماری در مراحل آموزش و صحت سنجی

تکرار	نرون لایه مخفی	روش آموزش	تابع انتقال	تعداد لایه	Train		Test	
					RMSE(m ³ /s)	R^2	RMSE(m ³ /s)	R^2
۱	۱۰	LM	TANSIG	۲	۰/۸	۰/۹۱	۱/۲۸	۰/۸۷
۲	۱۱	LM	LOGSIG	۲	۰/۷۵	۰/۸۹	۱/۲۳	۰/۸۲
۳	۱۱	LM	TANSIG	۳	۰/۵۳	۰/۸۳	۰/۸۶	۰/۸
۴	۱۲	LM	TANSIG	۳	۰/۴	۰/۸۴	۰/۶۴	۰/۹۱
۵	۱۳	LM	TANSIG	۴	۰/۳۵	۰/۹۲	۰/۵۷	۰/۸۲

Sub-ET 2



شکل (۵): ساختار درختی مدل خروجی

سنجی نیز مدل ANFIS-PSO بهترین پیش بینی را با توجه به نمایه‌های مورد بررسی داشته است و از بین سایر مدل‌ها، به ترتیب NN، LS-SVM و GEP عملکرد مناسبتری را ارائه دادند. به طور کلی در مجموع دو مقدار

بنابراین با توجه به نتایج به دست آمده، از مدل مورد بررسی، مدل ANFIS-PSO در بخش آموزش بهترین عملکرد را از خود نشان داده است و پس از آن مدل‌های NN، LS-SVM و GEP قرار گرفته اند. در بخش صحت

مرحله قبل به عنوان داده پرت حذف شده بودند) سنجیده شد. بنابراین از ۸ دبی بیشینه مشاهداتی استفاده و مقادیر محاسباتی و همچنین خطای نسبی هر یک محاسبه گردید. نتایج در این مرحله که در جدول (۱۱) آورده شده است، نیز نشان دهنده عملکرد بهتر مدل ANFIS-PSO نسبت به سایر موارد است. NN، LS-SVM و GEP هم در رتبه‌های بعدی قرار دارند.

R^2 و RMSE در دو بخش می‌توان بهترین عملکرد را به مدل سازی با ANFIS – PSO و ضعیف ترین عملکرد را به GEP نسبت داد که نتوانست مقادیر مناسبی را ارائه نماید.

ارزیابی مدل‌ها در پیش بینی دبی‌های بیشینه

به منظور ارزیابی مدل‌های ایجاد شده جهت برآورد موارد سیلابی، توانایی آن‌ها با بیشینه جریان‌ها (که در

جدول (۱۱): کارایی مدل‌های مورد مطالعه در برآورد بیشینه جریان‌های ورودی

خطای نسبی (درصد)				مقادیر پیش بینی شده (m^3/s)				دبی‌های بیشینه
GEP	LS-SVM	ANFIS-PSO	NN	GEP	LS-SVM	ANFIS-PSO	NN	مشاهداتی (m^3/s)
۷۲/۳۴	۶۵/۱۳	۳۴/۳۶	۴۱/۷۳	۲۱/۳	۱۶/۹	۴۰/۱	۳۵/۶	۶۱/۱
۷۰/۴۳	۶۵/۷۰	۵۴/۲۰	۶۵/۵	۱۶/۷	۱۴/۴	۲۲/۳	۱۶/۸	۴۸/۷
۶۰/۴۳	۴۸/۲۰	۳۳/۰۹	۴۸/۹۲	۱۴/۴	۱۱	۱۸/۶	۱۴/۲	۲۷/۸
۵۸/۱۱	۵۴/۷۱	۳۶/۲۲	۴۷/۹۲	۱۲	۱۱/۱	۱۶/۹	۱۳/۸	۲۶/۵
۵۹/۱۶	۵۷/۹۱	۴۲/۹۱	۴۵/۸۳	۱۰/۱	۹/۸	۱۳/۷	۱۳	۲۴
۵۶/۰۹	۵۳/۱۷	۴۱/۴۶	۳۶/۰۹	۹/۶	۹	۱۲	13.1	۲۰/۵
۶۵/۱۵	۵۱/۰۱	۳۷/۸۷	۴۴/۴۴	۹/۷	۶/۹	۱۲/۳	11	۱۹/۸
۷۱/۱۲	۵۸/۲۸	۴۷/۰۵	۴۵/۴۵	۷/۸	۵/۴	۹/۹	۱۰/۲	۱۸/۷

عملکرد را نسبت به سه روش دیگر ارائه نمود، به طوری که مقادیر RMSE و R^2 در بخش آموزش به ترتیب ۰/۴۲ و ۰/۹۵ و در بخش صحت سنجی به ترتیب ۰/۹۷ و ۰/۸ و به دست آمد. مدل‌های NN، LS-SVM و GEP در مرتبه‌های بعدی قرار گرفتند. همچنین در بخش پیش‌بینی دبی بیشینه نیز روش ANFIS-PSO با کمترین میزان خطاهای نسبی بین ۳۳/۰۹ تا ۵۴/۲ و بالاترین میزان همبستگی، عملکرد بهتری را داشت که در کل می‌توان این روش را مناسبترین روش این پژوهش گزارش کرد.

نتیجه‌گیری

در این پژوهش کارایی چهار روش داده محور ANFIS-PSO, LS-SVM, GEP, NN در تخمین دبی ورودی به سد جیرفت سنجیده شد. بدین منظور از داده‌های بارش، تبخیر، دما و مقادیر با تاخیر دبی، استفاده گردید. الگوی مناسب برای مدل سازی به وسیله آزمون همبستگی بین پارامترهای مختلف و دبی در نرم افزار SPSS حاصل شد که بهترین ترکیب $R_0-T_2-E_2-Q_1$ بود. با این ترکیب، در بخش مدل‌سازی و در هر دو قسمت آموزش و صحت سنجی، مدل ANFIS-PSO بهترین

منابع

- اسحاق تیموری، م. ع.، حبیب نژاد، م. کاویان، ع. و شاهدی، ک. ۱۳۹۱. شبیه سازی فرآیند بارش- رواناب در حوزه‌های با آمار کم با استفاده از مدل wms (مطالعه موردی؛ حوزه آبخیز چالوس). نشریه علمی پژوهشی مهندسی آبیاری و آب ایران، ۱۲-۲۵، (۱)۳.
- امامقلی‌زاده، ص. و کریمی‌دمنه، ر. ۱۳۹۶. کاربرد روش برنامه‌ریزی بیان ژن در تعیین ضریب هوادهی درپچه تخلیه‌کننده تحتانی سدها. مجله پژوهش‌های حفاظت آب و خاک. ۲۴(۱): ۲۸۶-۲۷۹.

- باقرپور، م.، سیدیان، م.، فتح آبادی، ا. و محمدی، ا. ۱۳۹۷. پیش‌بینی رواناب با استفاده از مدل‌های جعبه سیاه و خاکستری. تحقیقات منابع آب ایران، ۱۴(۵): ۲۱۹-۲۰۴.
- پوررضابیلندی، م.، خاشعی سیوکی، ع. و صادقی طبس، ص. ۱۳۹۳. پیش‌بینی رواناب روزانه با مدل حداقل مربعات ماشین بردار پشتیبان (LSSVM). نشریه پژوهش‌های حفاظت آب و خاک، ۲۱(۶): ۲۹۳-۳۰۴.
- سلگی، ا.، زارعی، ح. و گلابی، م. ر. ۱۳۹۶. بررسی عملکرد مدل برنامه‌ریزی بیان ژن با روش‌های پیش‌پردازش داده‌ها جهت مدل‌سازی جریان رودخانه. مجله پژوهش‌های حفاظت آب و خاک، ۲۴(۲): ۱۸۵-۲۰۱.
- صادقی طبس، ص. و پوررضابیلندی، م. ۱۳۹۲. مقایسه روش‌های بهینه‌سازی فراکاشی در تحلیل عدم قطعیت پارامترهای مدل مفهومی بارش-رواناب. مجله منابع طبیعی ایران، ۶۸(۲): ۵۵۲-۵۳۳.
- علیزاده، ا. ۱۳۸۵. اصول هیدرولوژی کاربردی، چاپ ششم، موسسه چاپ و انتشارات آستان قدس رضوی، دانشگاه امام رضا مشهد، ۷۳۴-۴۴۷.
- قربانی، م. ع. و دهقانی، ر. ۱۳۹۵. کاربرد شبکه‌های عصبی بیزین، ماشین بردار پشتیبان و برنامه‌ریزی بیان ژنی در تحلیل بارش - رواناب ماهانه (مطالعه موردی: رودخانه کاکارضا). علوم و مهندسی آبیاری، ۳۹(۲): ۱۳۸-۱۲۵.
- قربانی، خ. نیمه‌کلورزی، ز. سالاری جزی، م. و دهقانی، ا. ا. ۱۳۹۵. برآورد جریان ماهانه در حوضه‌های فاقد آمار با استفاده از پارامترهای اقلیمی و فیزیوگرافی حوضه! مجله پژوهش‌های حفاظت آب و خاک، ۲۳(۳): ۲۰۷-۲۲۴.
- کیاء، ع.، عمادی، ع. و غلامی، م. ع. ۱۳۹۸. مدل‌سازی بارش-رواناب با سیستم استنتاج فازی-عصبی تطبیقی (ANFIS) و رگرسیون خطی چندمتغیره (MLR). نشریه علمی پژوهشی مهندسی آبیاری و آب ایران، ۹(۴): ۳۹-۵۱.
- ناظری تهرودی، م.، هاشمی، س. ر.، احمدی، ف. و ناظری تهرودی، ز. ۱۳۹۵. بررسی دقت مدل‌های ANFIS، SVM و GP در مدل‌سازی مقادیر دبی جریان رودخانه. اکوهیدرولوژی، ۳(۳): ۳۴۱-۳۴۷.
- نیک‌پور، م.، ثانی خانی، ه.، محمودی بابلان، س. و محمدی، ع. ۱۳۹۵. کاربرد مدل‌های LSSVM, ANN, WNN و GEP در شبیه‌سازی بارش-رواناب رودخانه خیاوچای. مجله اکوهیدرولوژی، ۴(۲): ۶۲۵-۶۳۹.
- Aytek A, Alp M. An application of artificial intelligence for rainfall-runoff modeling. 2008. Journal of Earth System Science. 117 (2):145-155.
- Badrzadeh H, Sarukkalgige R, Jayawardena AW. Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. 2015. Journal of Hydrolog. 529:1633-1643.
- Botsis, D. Latinopoulos, P. and Diamantaras, K., 2011. Rainfall-Runoff Modeling Using Support Vector Regression and Artificial Neural Networks. In 12th International Conference on Environmental Science and Technology, Rhodes, Greece.
- Misra D, Oommen T, Agarwal A, Mishra S, Thompson A (2009) Application and analysis of support vector machine based simulation for runoff and sediment yield. Biosystems Engineering 103(4):527-535.
- Nayak PC, Venkatesh B, Krishna B, Sharad KJ. Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach. Journal of Hydrology. 2013; 493:57-67.
- Nourani V. An Emotional ANN (EANN) approach to modeling rainfall-runoff process. Journal of Hydrology. 2016; 544:267-277.
- Saeedi Farzad B. Intelligent simulation of rainfall-runoff using a semi-distributed model with time variables. Ph.D. thesis in civil engineering, Faculty of Engineering, University of Tabriz, Tabriz. 2014. [Persian]



Review and compare performance of 4 modeling methods LS-SVM, NN, GEP and ANFIS-PSO in Simulation of Rainfall - Runoff (Study Area: Halil River - Jiroft Dam)

Mohadeseh Kavusi¹, Hossein Khozayemeh Nejad²

Abstract

Simulation and prediction of rainfall-runoff process is one of the most important research areas in hydrology, that it is important in the management and planning of water resources of rivers and reservoirs and also their protection during the flood, but due to the complexity of the hydrological parameters, doing so have challenges. In this study, four data-driven method included LS-SVM, NN, GEP and ANFIS-PSO and data of 7 upstream meteorological stations of Jiroft Dam in Kerman province between 1999 and 2015 were used to simulate the input flow estimation process to Jiroft Dam. First combination of inputs was selected by Pearson test and then simulated by each method. The results based on two criteria of correlation coefficient (R), root mean square error (RMSE) showed that ANFIS-PSO method with $RMSE = 0.42 \text{ m}^3 / \text{s}$ and $R^2 = 0.95$ in education section and $RMSE=0.97 \text{ m}^3/\text{s}$ and $R^2 = 0.80$ in the test section was the best among the models studied and in both training and test sections, it was more accurate than the other three models. Also, this model performs better than other methods in simulation of maximum discharges and its relative calculation error was obtained between 33.09 and 54.2 for different discharges

Keywords: Rainfall-Runoff, LS-SVM, NN, GEP, ANFIS-PSO.

¹ PhD student of water resource, Department of Science and Water Engineering, School of Agriculture, University of Birjand, Birjand, Iran

² Associate Professor, Department of Science and Water Engineering, School of Agriculture, University of Birjand, Birjand, Iran (hkhozayemeh@birjand.ac.ir)



Review and compare performance of 4 modeling methods LS-SVM, NN, GEP and ANFIS-PSO in Simulation of Rainfall - Runoff (Study Area: Halil River - Jiroft Dam)

Mohadeseh Kavusi¹, Hossein Khozaymeh Nejad²

Introduction:

One of the basic needs in designing water resources management and flood control projects is the amount of runoff generated by rainfall. The ultimate goal of runoff precipitation modeling is to find an accurate solution for estimating runoff runoff in a catchment. But this process is complex and non-linear with large spatial and temporal variations and is not easily modeled. In general, three perspectives are presented for runoff precipitation modeling: theoretical, conceptual, and systemic. The first two models require complex and extensive information, but the third method has a simpler function and is a good option in the discussion of management and decision making.

In the present study, after selecting the appropriate combination of data using a statistical test that includes four parameters: pre-discharges, temperature, precipitation and evaporation, the efficiency of four intelligent methods LS-SVM, NN, GEP and ANFIS-PSO which include Two combined methods and two separate methods were compared and their ability to simulate the inflow rate to the dam in a dry area that had not been done before was investigated. Finally, the evaluation of the models in predicting the maximum discharges was shown.

Materials and methods area of study Jiroft Dam is located 40 km northwest of Jiroft city in Kerman province on the Halil River and the area of the watershed above it is about eight hundred thousand hectares. In this study, meteorological data from 7 stations upstream of Jiroft Dam and close to Halilrud River, including precipitation, temperature and evaporation in the period from 1999 to 2015, as well as discharge data from hydrometric stations near the dam in the same period were used. In all modelings, 70% of the data were considered as training data and 30% as validation data.

Method

After deleting the outdated data, the data was returned to normal to increase the modeling accuracy. Also, in order to evaluate the accuracy and efficiency of the models, correlation coefficient (R²) and root mean square error (RMSE) were used. In order to select the appropriate combination of input parameters to the models, statistical tests were used to increase the accuracy of flow modeling. Calculated in SPSS software. Parameters T, R, E and Q represent the measured values of temperature, precipitation, evaporation and flow, respectively, in the studied stations, and the index of numbers is the time step index before the present time (t). According to the obtained values, the pattern R0-T2-E2-Q1, ie the use of precipitation one step before, the temperature of the previous step, the evaporation of the previous step and the flow of one step before, is selected as the best model and used in modeling Took. Then, four intelligent models were used to model the inlet flow to the dam. LS-SVM, NN and ANFIS-PSO models were implemented in MATLAB programming environment and GEP model was implemented in GeneXproTools 5 software. In the data entry stage to the LS-SVM model, the optimal combination was normalized after data normalization. In the training section, in order to be stable in determining the parameters of the estimation function, the Simplex optimization method was used for 3 selected kernel functions (RBF, Linear, and Polynomial). Using the NN method, by performing 5 replications with the number of layers, the transfer function and the number of different neurons, the best model was selected according to the desired indices. The number of repetitions was considered 1000 for all cases. The ANFIS-PSO hybrid model was also coded in

¹ PhD student of water resource, Department of Science and Water Engineering, School of Agriculture, University of Birjand, Birjand, Iran

² Associate Professor, Department of Science and Water Engineering, School of Agriculture, University of Birjand, Birjand, Iran (hkhozaymeh@birjand.ac.ir)

MATLAB software and its main parameters were based on previous research and trial and error were considered. In modeling with gene expression method, due to trial and error, selected functions including four main operations, logarithm, ln, sine and cosine, radical and power 2 were used as the primary functions in this method.

Results and discussion

According to the results, among the 4 models reviewed, ANFIS-PSO model showed the best performance in training with $RMSE = 0.42 \text{ m}^3 / \text{s}^3$ and $R^2 = 0.95$, and then there were the NN, LS-SVM and GEP models, respectively. In terms of validation, ANFIS-PSO model with $RMSE = 0.97 \text{ m}^3 / \text{s}$ and $R^2 = 0.80$ was the best prediction according to the studied indicators and among other models, NN, LS-SVM and GEP, respectively provided better performance. Generally, based on the results of the two sections, the best performance to modeling with ANFIS - PSO and the least accurate performance for GEP. In order to evaluate the performance of the models in estimate flood cases, their ability was measured in forecasting the maximum flows (which were previously excluded as data outflows). Therefore, 8 observational maximum discharges were used and the computational values and the relative error of each were calculated. The results at this section also show better performance of the ANFIS-PSO model than other methods. NN, LS-SVM and GEP are in the next ranks.

Conclusion

In this study, the efficiency of four data-driven methods including: ANFIS-PSO, LS-SVM, GEP, NN in estimating the inflow to Jiroft Dam was measured with precipitation, evaporation, temperature and delay in flow data. The appropriate model for modeling by correlation test between different parameters and discharge in SPSS software was R0-T2-E2-Q1. With this combination, in both parts training and validation of the modeling department, the ANFIS-PSO model was the best performance compared to the other three methods, and the NN, LS-SVM and GEP models were in the next ranks. Also, in the maximum flow prediction, the ANFIS-PSO method was a better performance with the lowest relative error rate between 33.09 and 54.2 and the highest correlation rate, which generally can be introduced it the most appropriate method of this research.

Keywords: Rainfall-Runoff, LS-SVM, NN, GEP, ANFIS-PSO.

The most important references

- Aytek A, Alp M. An application of artificial intelligence for rainfall-runoff modeling. 2008. Journal of Earth System Science. 117(2):145-155.
- Badrzadeh H, Sarukkalige R, Jayawardena AW. Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. 2015. Journal of Hydrolog. 529:1633-1643.
- Botsis, D. Latinopoulos, P. and Diamantaras, K., 2011. Rainfall-Runoff Modeling Using Support Vector Regression and Artificial Neural Networks. In 12th International Conference on Environmental Science and Technology, Rhodes, Greece.
- Misra D, Oommen T, Agarwal A, Mishra S, Thompson A (2009) Application and analysis of support vector machine based simulation for runoff and sediment yield. Biosystems Engineering 103(4):527-535.
- Nayak PC, Venkatesh B, Krishna B, Sharad KJ. Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach. Journal of Hydrology. 2013;493:57-67.
- Nourani V. An Emotional ANN (EANN) approach to modeling rainfall-runoff process. Journal of Hydrology. 2016;544:267-277.
- Saedi Farzad B. Intelligent simulation of rainfall-runoff using a semi-distributed model with time variables. Ph.D. thesis in civil engineering, Faculty of Engineering, University of Tabriz, Tabriz. 2014