علوم و مهندسی آبیاری
جلد ۴۱، شماره‌ی٤، زمستان ۹۷، ص. ١٤٦-١٣٣

مدل‌سازی بارش – رواناب روزانه رودخانه دره‌رود واقع در استان اردبیل

محمد رضا نیک‌پور[1*]، هادی ثانی خانی[2]، سجاد محمودی بابلان[3] و سعید نسترنی عموقین[4]

۱*- نویسنده مسئول، استادیار گروه مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی. m_nikpour@uma.ac.ir
۲- استادیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه کردستان.
۳- دانش آموخته کارشناسی، گروه مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی.
۴- دانش آموخته کارشناسی، گروه مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی.

دریافت: ۹۵/۱۱/۱۰ بازنگری: ۹۶/۲/۳ پذیرش: ۹۶/۲/۹

چکیده

مدل‌سازی بارش- رواناب یکی از پراهمیت‌ترین موضوعات در مدیریت منابع آب‌های سطحی برای اتخاذ تدابیر مناسب در مواقع سیلاب و بروز خشک‌سالی‌ها است. در این تحقیق از مدل‌های شبکه‌های عصبی مصنوعی، برنامه‌ریزی بیان ژن، موجک- عصبی و حداقل مربعات ماشین بردار پشتیبان به‌منظور تخمین جریان روزانه رودخانه دره‌رود استفاده شد. داده‌های دبی و بارش روزانه ایستگاه مشیران واقع بر رودخانه مذکور و در بالادست سد عمارت به‌کار گرفته شد. نتایج نشان داد که الگوی تعریف شده بر اساس دبی روز قبل و بارش همان روز می‌تواند بهترین برآورد را از رواناب روزانه داشته باشد. هم‌چنین نتایج بیانگر عملکرد قابل قبول مدل‌ها و برتری مدل موجک- عصبی با بیشترین ضریب همبستگی (R=۰/۹۵۲)، کمترین ریشه میانگین مربعات خطا (RMSE=۱/۵۸۹) و ضریب نش ساتکلیف برابر ۰/۹۰۵ در مرحله صحت‌سنجی بود. در برآورد دبی بیشینه نیز مدل مذکور با میانگین خطای نسبی ۲۵/۹۷ درصد، از خطای کمتری نسبت به سایر مدل‌ها برخوردار بود.

کلید واژه‌ها: برنامه‌ریزی بیان ژن، تخمین جریان، حداقل مربعات ماشین بردار پشتیبان، شبکه عصبی مصنوعی، موجک- عصبی.

مقدمه

بارش- رواناب یکی از پیچیده‌ترین فرآیندهای هیدرولوژیکی است که از متغیرهای مختلف فیزیکی و هیدرولوژیکی تأثیر می‌پذیرد. لذا درک و پیش‌بینی فرآیندهای تولید رواناب و انتقال آن به نقطه خروجی حوضه آبخیز یکی از اساسی‌ترین مباحث علم هیدرولوژی محسوب می‌شود (Salajegheh et al., 2009). از سوی دیگر، با توجه به محدودیت منابع آب شیرین قابل استحصال، پیش‌بینی هرچه دقیق‌تر دبی جریان و تغییرات آن در طول سال از ارکان اساسی برنامه‌ریزی و مدیریت منابع آب‌های سطحی است (Nabizadeh et al., 2012). نظر به اهمیت موضوع، تاکنون تحقیقات متعددی در خصوص استفاده از مدل‌های مختلف هوشمند شامل شبکه‌های عصبی مصنوعی (ANNs)، برنامه‌ریزی بیان ژن (GEP)، فازی- عصبی (NF)، موجک- عصبی (WNN) و ماشین بردار پشتیبان (SVM) برای تخمین جریان (رواناب) روزانه و ماهانه رودخانه‌ها و همچنین روندیابی سیلاب صورت پذیرفته است.

Tokar و Markus (2000) جریان ماهانه سه حوضه متفاوت آبخیز را ازنظر شرایط آب و هوایی و خصوصیات فیزیوگرافی، با استفاده از ANNs و مدل مفهمومی تعادل آب (Watbal) مدل‌سازی کردند. در نهایت مدل ANNs نتیجه بهتری به‌همراه داشت. Jayawardena et al. (2005) با استفاده از برنامه‌ریزی ژنتیک (GP) و با به‌کارگیری داده‌های روزانه، بارش- رواناب دو حوضه نسبتاً بزرگ را در کشور چین شبیه‌سازی کردند که نتایج حاصل با داده‌های واقعی مطابقت خوبی داشت. Nourani et al. (2009) برای مدل‌سازی بارش- رواناب حوضه لیقوان چای از تلفیق تحلیل موجک با مدل ANN استفاده کردند. نتایج به‌دست‌آمده نشان داد که مدل پیشنهاد شده به دلیل استفاده از سری‌های زمانی با مقیاس چندگانه به‌عنوان لایه ورودی مدل ANN، قادر به پیش‌بینی رویدادهای بارندگی طولانی‌مدت و کوتاه‌مدت می‌باشد. Aytek و Alp (2008) با مدل‌های ANN و GEP بارش- رواناب روزانه رودخانه جونیاتا در ایالت پنسیلوانیای آمریکا را شبیه‌سازی کردند. نتایج تحقیق، برتری مدل GEP را نسبت به ANN در تخمین رواناب روزانه رودخانه مذکور نشان داد. Kisi (2008) عملکرد مدل ترکیبی WNN را با مدل NF در مدل کردن جریان ماهانه مقایسه و براساس نتایج تحقیق، مدل WNN را برای این منظور پیشنهاد نمود. Soltani et al. (2010) با کاربرد GEP در شبیه‌سازی فرآیند بارش- رواناب روزانه حوضه آبخیز لیقوان، با تعریف دو مجموعه عملگر ریاضی به این نتیجه دست یافتند که

مدل حاصل از مجموعه عملگرهای اصلی مناسب‌تر است. در سال‌های اخیر SVM به‌عنوان یک مدل جدید به‌منظور تخمین متغیرهای هیدرولوژیکی به‌کار گرفته می‌شود و موفقیت‌های قابل توجه در زمینه‌های مختلف، از جمله مدل‌سازی و پیش‌بینی بارش و رواناب ماهانه به همراه داشته است. Behzad et al. (2009) با استفاده از آمار یک ایستگاه هیدرومتری واقع بر روی رودخانه بختیاری، جریان روزانه را با در نظر گرفتن آمار روز قبل پیش‌بینی نمودند. نتایج نشان داد که SVM از عملکرد بهتری نسبت به ANNs برخوردار است. Yu et al. (2006) در مطالعه‌ای با استفاده از مدل SVM به پیش‌بینی سطح رقومی سیل در رودخانه لن- یانگ تایوان پرداختند. نتایج نشان داد که مدل به‌خوبی سطح آب رودخانه را برای ۱ تا ۶ ساعت بعد پیش‌بینی می‌کند. Taheri and Ghafouri (2012) رواناب روزانه یکی از ایستگاه‌های رودخانه کارون را با به‌کارگیری داده‌های بارش، رطوبت و دما مدل‌سازی نمودند. در تحقیق مذکور مدل SVM نتیجه بهتری را نسبت به روش یادگیری فعال به‌همراه داشت. 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 مورد مقایسه قرار گرفت. در نهایت SVM شبیه‌سازی دقیق‌تری از بارش- رواناب داشت. Ghorbani and Dehghani (2015) کارآیی 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. رودخانه دره‌رود به‌عنوان مهم‌ترین سرشاخه رود مرزی ارس در ایران، یکی از رودخانه‌های مهم استان اردبیل و یکی از منابع اصلی تأمین آب بخش‌های مختلف استان می‌باشد. از سوی دیگر روی رودخانه مذکور سد مخزنی عمارت در حال ساخت می‌باشد. بنابراین در تحقیق حاضر، نظر به موقعیت استراتژیک رودخانه مذکور و اهمیت مدیریت صحیح بهره‌برداری از آن، پیش‌بینی مقدار دبی جریان حائز اهمیت می‌باشد. با بررسی منابع انجام شده می‌توان گفت استفاده از روش‌های مدل‌سازی غیرخطی برای پیش‌بینی مقدار جریان رودخانه ضروری به نظر می‌رسد. لذا در این راستا مدل ANN به‌دلیل سادگی و کارآمدی بالا، مدل GEP به‌دلیل ارائه روابط صریح بین متغیرهای ورودی و خروجی، مدل LS-SVM به‌دلیل کارآیی و سرعت بالا و مدل WNN به‌دلیل استفاده از نویززدایی در ساختار داده‌ها، برای تخمین جریان ورودی به مخزن سد عمارت (واقع بر رودخانه دره‌رود، استان اردبیل) مورد ارزیابی قرار گرفت. در بین مدل‌های ذکر شده دامنه مطالعات انجام شده در زمینه مدل‌های WNN و LS-SVM به خصوص در مطالعات مدل‌سازی بارش- رواناب و هیدرولوژی، نسبتاً کم و محدود می‌باشد. در این تحقیق از ضرایب همبستگی موجود بین متغیرهای دبی جریان و بارش برای تعیین ورودی مدل‌های هوشمند استفاده شد.

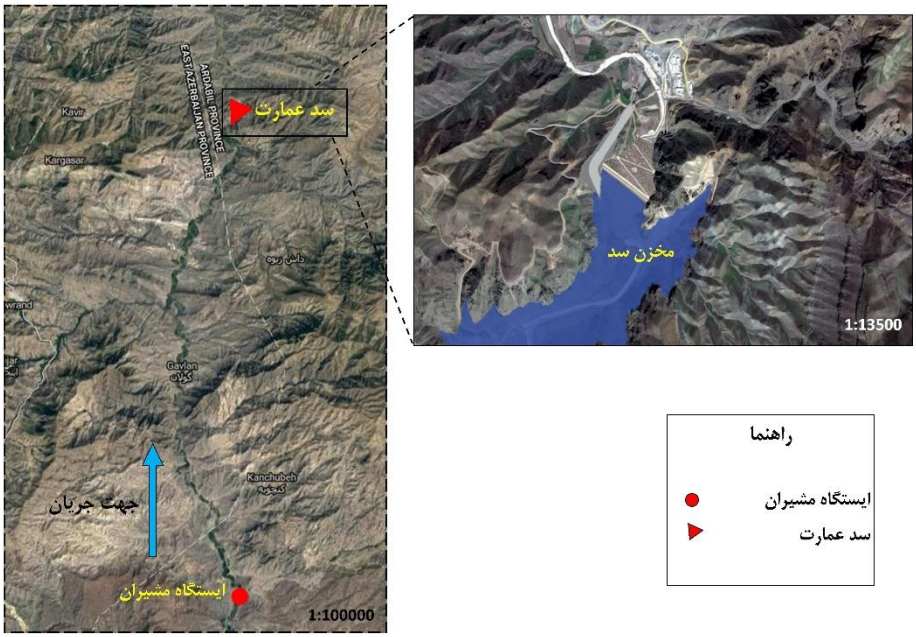
مواد و روش‌ها

منطقه مورد مطالعه

حوضه آبریز دره‌رود با وسعتی برابر ۱۴۳۱۰ کیلومتر مربع بخش قابل توجهی از استان اردبیل را تحت پوشش قرار می‌دهد. ۶۸ درصد وسعت حوضه در استان اردبیل و ۳۲ درصد آن در استان آذربایجان شرقی است. دره‌رود که یکی از رودخانه‌های مهم حوضه آبریز مذکور می‌باشد، به‌عنوان مهمترین سرشاخه رود مرزی ارس در ایران، از الحاق دو سرشاخه اصلی قره‌سو و اهرچای تشکیل می‌گردد (Shafiee-Jood et al., 2012). برای شبیه‌سازی رواناب روزانه دره‌رود از داده‌های آماری مربوط به ایستگاه هیدرومتری مشیران واقع در رودخانه مذکور و در بالادست سد عمارت طی دوره آماری ۲۶ ساله از سال ۱۳۶۷ تا سال ۱۳۹۳ استفاده گردید. لازم به ذکر است که بر اساس گزارش‌های موجود، سد عمارت در سال ۱۳۹۷ بهره‌برداری خواهد شد. ایستگاه هیدرومتری مشیران در ۴۷/۵۲ درجه طول جغرافیایی، ۳۸/۵۳ درجه عرض جغرافیایی، در ارتفاع ۷۸۰ متری از سطح دریا واقع شده است. شکل (۱) محدوده و موقعیت کلی منطقه مورد مطالعه را نشان می‌دهد. داده‌های بارش و رواناب روزانه از سازمان هواشناسی و شرکت آب منطقه‌ای استان اردبیل اخذ شد. تعداد کل نمونه‌ها در طول دوره آماربرداری، معادل ۹۳۴۰ داده می‌باشد. ۷۰ درصد داده‌های مذکور (۶۵۳۸ داده) به‌عنوان داده‌های آموزش و ۳۰ درصد داده‌ها (۲۸۰۲ داده) به‌عنوان داده‌های صحت‌سنجی در نظر گرفته شد. در جدول (۱) متغیرهای آماری مربوط به داده‌های اندازه‌گیری شده در بازه زمانی مذکور نشان داده‌شده است.

انتخاب الگوهای مناسب

به‌منظور انتخاب ترکیب مناسب از متغیرهای ورودی مدل‌ها، از آزمون‌های آماری استفاده شد و مقادیر ضرایب خود همبستگی جزئی متغیر دبی و همبستگی متقابل متغیرهای بارش و دبی محاسبه گردید. مقادیر همبستگی بین متغیرهای دبی و بارش و تأخیرهای آن‌ها در جدول (۲) ارائه شده است. در جدول مذکور، متغیرهای R و Q به ترتیب بیانگر مقادیر اندازه‌گیری شده بارش و دبی در ایستگاه هیدرومتری مشیران می‌باشد و اندیس t شاخص گام زمانی فعلی می‌باشد. با توجه به همبستگی متقابل و معنی‌دار بین متغیرهای ورودی و خروجی، الگوهای مختلفی برای تخمین رواناب روزانه رودخانه دره‌رود مطابق با جدول (۳) استفاده شد.



شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه

جدول ۱- خصوصیات آماری مقادیر دبی روزانه رودخانه مورداستفاده در بازه زمانی ۹۳-۱۳۶۷

متغیر	واحد	تعداد رکورد	میانگین	حداقل	حداکثر	انحراف معیار	ضریب تغییرات	چولگی
$R_{(t)}$	(mm)	۹۳۴۸	۰/۹۸۴	۰	۴۴/۵	۲/۹۹۲	۳/۰۴۰	۴/۸۶۱
$Q_{(t)}$	(m^3/day)	۹۳۴۸	۸/۶۳۲	۰	۱۸۵/۲	۱۰/۸۹۹	۱/۲۶۲	۳/۴۲۹

جدول ۲- مقادیر همبستگی بین متغیرهای بارش و دبی

	Q_{t-1}	R_t	R_{t-1}	R_{t-2}
Q_t	۰/۹۱۶	۰/۱۰۷	۰/۱۲۶	۰/۱۱۵

جدول ۳- ساختار الگوهای مورد استفاده در تحقیق حاضر

شماره الگو	ساختار ورودی	خروجی
۱	R_t	Q_t
۲	R_{t-1}, R_t	Q_t
۳	R_{t-2}, R_{t-1}, R_t	Q_t
۴	Q_{t-1}, R_t	Q_t
۵	Q_{t-1}, R_{t-1}, R_t	Q_t
۶	Q_{t-1},R_{t-2}, R_{t-1}, R_t	Q_t

مدل‌های هوشمند

در تحقیق حاضر برای اجرای مدل‌های ANNs، WNN و LS-SVM از کد نوشته شده توسط نویسندگان مقاله در محیط برنامه‌نویسی MATLAB و برای به‌کارگیری مدل GEP از نرم‌افزار GeneXproTools استفاده گردید. در ادامه به توضیح مختصر مدل‌های مذکور و ضرایب به‌کار رفته برای واسنجی مدل-ها پرداخته می‌شود.

شبکه‌های عصبی مصنوعی (ANNs)

ساختار کلی ANNs از شبکه بیولوژیکی مغز انسان الهام گرفته است. ANNs با پردازش روی‌داده‌ها قوانین کلی را فرامی‌گیرد و خروجی‌های شبکه با خروجی‌های مطلوب (اندازه‌گیری شده) مقایسه می‌شود. در ابتدا با استفاده از داده‌ها و مقادیر اندازه‌گیری شده موجود، شبکه آموزش داده می‌شود. متغیرهای مؤثر که در مدل‌سازی شبکه‌های عصبی نقش اساسی

دارند شامل تعداد تکرار و نوع تابع آموزش، تعداد لایه‌های شبکه و تعداد نرون‌های لایه میانی می‌باشد. گام مهم قبل از واردکردن داده‌ها به مدل، نرمال سازی داده‌ها می‌باشد. این عمل خصوصاً وقتی دامنه تغییرات ورودی‌ها زیاد باشد کمک شایانی به آموزش بهتر و سریع‌تر مدل می‌کند. اصولاً واردکردن داده‌ها به‌صورت خام باعث کاهش سرعت و دقت شبکه می‌شود (He et al., 2014). برای نرمال‌سازی داده‌های تحقیق حاضر از رابطه (۱) استفاده شد. در این رابطه، X_n مقدار نرمال شده ورودی X_i، X_{max} و X_{min} به ترتیب مقادیر حداکثر و حداقل داده‌ها می‌باشد. به‌منظور به حداقل رساندن خطای مدل‌سازی پس از انجام تحلیل حساسیت، توابع و مقادیر بهینه در نظر گرفته‌شده برای شبکه عصبی مورد استفاده در تحقیق حاضر به شرح جدول (۴) می‌باشد.

$$X_n = 0.1 + 0.8\frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \qquad (۱)$$

موجک- عصبی (WNN)

موجک، یک موج کوچک است که انرژی آن در ناحیه کوچکی متمرکزشده و ابزار مناسبی برای بررسی پدیده‌های گذرا است. موجک نوسان حداقلی دارد که به سمت صفر نزول می‌کند و این نزول باید در جهات مثبت و منفی واقع در دامنه‌اش محدود شود. این مشخصه باعث می‌شود که موجک انعطاف‌پذیر بوده و مانند یک تابع رفتار کند، بنابراین با توجه به رفتارهای موجی نامنظم پدیده‌های هیدرولوژیک، قابلیت مدل‌سازی این پدیده‌ها توسط امواج نامنظم (موجک‌ها) نسبت به امواج منظم (امواج سینوسی) بیشتر است (Nourani et al., 2009). موجک دارای سه مشخصه تعداد نوسان محدود، بازگشت سریع به صفر در هر دو جهت مثبت و منفی در دامنه خود و میانگین صفر است که شرط مقبولیت نامیده می‌شود و به شکل رابطه (۲) بیان می‌شود.

$$\int_{-\infty}^{+\infty}\phi(t)dt = 0 \qquad (۲)$$

در رابطه فوق، φ(t) تابع تبدیل موج (موجک مادر) است. عبارت مادر به این دلیل به کار رفته است که توابع (موجک‌های) متفاوت به وجود آمده بر اساس متغیرهای مقیاس و انتقال همگی از تابع پایه (موجک مادر) ناشی می‌شوند. تبدیل موجک عملگری است که یک تابع را به مجموعه‌ای از تبدیل موجک‌های مادر اصلاح‌شده تبدیل می‌کند که این تبدیل موجک‌های اصلاح‌شده همان تبدیل موجک دختر می‌باشند (Marofi et al., 2012). بعد از تعیین نوع موجک مادر، مهم‌ترین اقدام تعیین سطح تجزیه بود. استفاده از حداکثر سطح تجزیه موجب کند شدن فرآیند آموزش و در برخی موارد می‌تواند موجب کاهش دقت مدل نیز گردد. برای انتخاب سطح تجزیه تبدیل موجک گسسته، رابطه (۳) استفاده شد (Nourani et al., 2009). در این رابطه، L سطح تجزیه، N_S تعداد داده‌های سری زمانی و Int تابع جزء صحیح می‌باشد. در تحقیق حاضر از موجک‌های خانواده دابچیز (Daubechies) با سطح تجزیه ۳ استفاده گردید. در نهایت داده‌های حاصل از تجزیه توسط موجک پس از نرمال‌سازی (با استفاده از رابطه ۱) به‌عنوان ورودی شبکه عصبی تعریف شد.

$$L = Int(\log(N)) \qquad (۳)$$

برنامه‌ریزی بیان ژن (GEP)

روش برنامه‌ریزی بیان ژن، در سال ۱۹۹۹ توسط فریرا ارائه شد (Ferreira, 2001). این روش ترکیبی از روش‌های برنامه‌ریزی ژنتیک (GP) و الگوریتم ژنتیک (GA) بوده که در آن، کروموزوم‌های خطی و ساده با طول ثابت، مشابه با آنچه در الگوریتم ژنتیک استفاده می‌شود و ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک، ترکیب می‌شوند. ازآنجایی‌که در این روش تمام ساختارهای شاخه‌ای با اندازه اشکال متفاوت، در کروموزوم‌های خطی با طول ثابت کدگذاری می‌شوند، معادل این است که در این روش فنوتیپ و ژنوتیپ از هم جدا می‌شوند و سیستم می‌تواند از تمام مزایای تکاملی به سبب وجود آن‌ها بهره‌مند شود. اکنون باوجود اینکه فنوتیپ در GEP، همان نوع از ساختارهای شاخه‌ای مورد استفاده در GP را شامل می‌شود، اما ساختارهای شاخه‌ای که به‌وسیله GEP استنتاج می‌شوند (که بیان درختی نیز نامیده می‌شود) مبین تمامی ژنوم‌های مستقل هستند. به‌طور خلاصه می‌توان گفت در GEP بهسازی‌ها در یک ساختار خطی اتفاق افتاده و سپس به‌صورت ساختار درختی بیان می‌شود و این موجب می‌شود تنها ژنوم اصلاح‌شده به نسل بعد منتقل شده و نیازی به ساختارهای سنگین برای تکثیر و جهش وجود نداشته باشد (Ferreira, 2006). در این روش پدیده‌های مختلف با استفاده از مجموعه‌ای از توابع و مجموعه‌ای از ترمینال‌ها، مدل‌سازی می‌شوند. مجموعه توابع، معمولاً شامل توابع اصلی حسابی{+, -, ×, /}، توابع مثلثاتی و سایر توابع ریاضی و یا توابع تعریف شده توسط کاربر است که معتقد است، می‌تواند برای تفسیر مدل مناسب باشند. مجموعه ترمینال‌ها، از مقادیر ثابت و متغیرهای مستقل مسأله تشکیل شده‌اند (Ferreira, 2001). برای استخراج رابطه ریاضی برآورد مقدار دبی در این تحقیق علاوه بر چهار تابع اصلی حسابی، از توابع مثلثاتی شامل sin و cos و هم‌چنین عملگرهای √، power، Ln و Exp استفاده گردید. پس از تحلیل حساسیت مدل برنامه‌ریزی بیان ژن در تخمین دبی، متغیرهای مورداستفاده و نرخ آن‌ها در جدول (۵) ارائه‌شده است.

ماشین بردار پشتیان (SVM)

ایده اصلی SVM که به‌عنوان فرآیندهای رگرسیونی و طبقه‌بندی شناخته می‌شوند، توسط Vapinc (1998) مطرح گردید. رگرسیون بردار پشتیبان جهت توصیف رگرسیون با استفاده از ماشین بردار پشتیبان می‌باشد. در این روش، تلاش بر این است تا وابستگی بین مجموعه داده‌های ورودی $X=\left(\overrightarrow{x_1},\overrightarrow{x_2},...,\overrightarrow{x_n}\right)$ از فضای R^n و مقادیر خروجی $Y=\left(\overrightarrow{y_1},\overrightarrow{y_2},...,\overrightarrow{y_n}\right)$ که در آن $y_i \in R$ از طریق تابع $y(\overrightarrow{x_i})$ تخمین زده شود. در این روش فرض بر این است که نمونه‌ها به‌صورت مستقل و از یک توزیع احتمالاتی نامعین گرفته‌شده‌اند. در مدل SVM برای حل معادلات از برنامه‌ریزی دوگانه استفاده می‌شود. لذا حل مسائل در ابعاد بزرگ با استفاده از مدل مذکور، هزینه محاسباتی را به‌شدت افزایش می‌دهد. برای رفع این مشکل، Suykens et al. (2002) روش حداقل مربعات ماشین بردار پشتیبان (LS-SVM) را ارائه نمودند. در این مدل از معادلات خطی ساده برای حل مسائل استفاده می‌شود که به‌طور مؤثری پیچیدگی الگوریتم را کاهش می‌دهد. تابع رگرسیونی مورد استفاده برای تخمین مسائل به‌صورت رابطه (۴) می‌باشد.

$$y(x_i)=w^T.\phi(x_i)+b \qquad (۴)$$

که در آن $\phi(x_i)$ ترسیم غیرخطی ورودی‌ها در فضای ویژگی با ابعاد بالا می‌باشد. متغیرهای w و b به‌ترتیب مقادیر وزن‌ها و بایاس تابع رگرسیون هستند که از طریق حداقل سازی تابع هدف در رابطه (۵) تعیین می‌گردد.

$$\min_{w,e,b} j(w,e)=\frac{1}{2}w^T.w+\frac{\gamma}{2}\sum_{i=1}^{N}e_i^2 \qquad (۵)$$

با محدودیت:

$$y_i=w^T.\phi(x_i)+b+e_i \qquad i=1,2,...N \qquad (۶)$$

که در آن γ متغیر تنظیم‌کننده بخش خطا و e_i خطای داده‌های آموزش می‌باشد. در نهایت تابع تخمین مدل LS-SVM به‌صورت رابطه (۷) تعریف می‌شود.

$$y(x)=\sum_{i=1}^{N}a_iK(x_i,x_j)+b \qquad (۷)$$

در رابطه فوق، $K(x_i,x_j)$ تابع کرنل نامیده می‌شود که مطابق رابطه (۸)، به‌عنوان تابعی با ایجاد ضرب داخلی در فضای ویژگی معرفی می‌شود (Rezaei et al., 2014).

$$K(x_i,x_j)=\left\langle \varphi(x_i)\times\varphi(x_j)\right\rangle \quad i,j=1,2,...N \qquad (۸)$$

برای تحلیل حساسیت مدل LS-SVM توابع کرنل خطی، چندجمله‌ای و پایه شعاعی مورد آزمون قرار گرفت که از بین توابع مذکور، کرنل چندجمله‌ای از بالاترین سرعت و کمترین خطا برخوردار بود. ضمناً مقادیر متغیرهای تنظیم‌کننده و کرنل به‌ترتیب برابر ۲۰ و [۰/۲ و ۲] در نظر گرفته شد.

جدول ٤- مقادیر متغیرهای مورداستفاده در مدل شبکه‌های عصبی مصنوعی

متغیر	بهترین مقدار
نوع شبکه عصبی	Feed-forward back propagation
تعداد لایه پنهان	۱
الگوریتم آموزشی	Levenberg-Marquardt
تعداد نرون	۱-۱۰
تابع انتقالی برای لایه اول	TANSIG
تابع انتقالی برای لایه دوم	RURELIN
تعداد تکرار	۱۰۰۰

جدول ٥- مقادیر متغیرهای مورداستفاده در روش برنامه‌ریزی بیان ژن

تنظیمات کلی		عملکردهای ژنتیکی	
تعداد کروموزوم‌ها	۳۰	نرخ جهش	۰/۰۴۴
اندازه رأس	۷	نرخ وارون سازی	۰/۱
تعداد ژن در هر کروموزوم	۳	نرخ ترانهش درج متوالی	۰/۱
تعداد جمعیت تولیدی	۱۰۰۰	نرخ ترانهش ریشه درج متوالی	۰/۱
تابع پیوند	مجموع (+)	نرخ ترکیب تک نقطه‌ای	۰/۳

معیارهای ارزیابی مدل

در تحقیق حاضر به‌منظور ارزیابی دقت و کارآیی مدل‌ها از نمایه‌های ضریب همبستگی (R)، ریشه میانگین مربعات خطا(RMSE) و ضریب نش- ساتکلیف (NS) طبق روابط زیر استفاده گردید. بهترین مقدار برای معیارهای مذکور به ترتیب یک، صفر و یک می‌باشد.

$$R = \frac{\sum_{i=1}^{N}(Q_i^O - \overline{Q^O})(Q_i^P - \overline{Q^P})}{\sqrt{\sum_{i=1}^{N}(Q_i^O - \overline{Q^O})^2 \sum_{i=1}^{N}(Q_i^P - \overline{Q^P})^2}} \quad (۹)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Q_i^O - Q_i^P)^2}{N}} \quad (۱۰)$$

$$NS = 1 - \frac{\sum_{i=1}^{N}(Q_i^O - Q_i^P)^2}{\sum_{i=1}^{N}(Q_i^O - \overline{Q_i^O})^2} \quad (۱۱)$$

در روابط فوق، Q_i^O و Q_i^P به‌ترتیب مقادیر دبی مشاهداتی و محاسباتی توسط مدل در گام زمانی i ام، N تعداد گام‌های زمانی، $\overline{Q^O}$ و $\overline{Q^P}$ نیز به ترتیب میانگین مقادیر دبی مشاهداتی و محاسباتی می‌باشد. علاوه بر معیارهای فوق از نمودارهای پراکنش حول خط یک به یک و سری زمانی مقادیر مشاهداتی و محاسباتی برای مقایسه و تحلیل بیشتر نتایج استفاده شد.

نتایج و بحث

ارزیابی مدل‌ها بر اساس شاخص‌های آماری

در جدول‌های (۶) تا (۹) شاخص‌های آماری عملکرد مدل‌های مورد استفاده در تحقیق حاضر برای مراحل آموزش و صحت‌سنجی و به‌ازای الگوهای مختلف گزارش شده است. بر اساس نتایج گزارش شده در جدول‌های (۵) تا (۸)، برای مدل GEP، الگوی شماره ۵ و برای سایر مدل‌ها الگوی شماره ۴ به‌طور نسبی بهترین عملکرد را به‌همراه داشت. الگوی ۴ در مقایسه با سایر الگوهای ترکیبی تعداد داده‌های ورودی کمتری نیاز دارد و مزیت مهم آن کاهش زمان اجرای مدل‌ها و پیچیدگی آن‌ها می‌باشد. بنابراین جریان رودخانه دره‌رود در گام زمانی فعلی در حالت کلی متأثر از مقدار آن در روز قبل و بارش همان روز می‌باشد و استفاده از مقادیر بارش در گام‌های زمانی پیشین نیز تأثیری بر بهبود نتایج مدل‌ها در تخمین جریان رودخانه نداشت. از سوی دیگر ملاحظه می‌شود که در کلیه مدل‌ها، الگوهای ۱، ۲ و ۳ نتایج ضعیفی را به دنبال داشته است و این حاکی از آن است که در حوضه آبریز دره‌رود صرفاً با استفاده از مقادیر بارش به تنهایی نمی‌توان برآورد صحیحی از جریان رودخانه داشت. در تحقیقات مشابه سایر محققین در خصوص شبیه‌سازی جریان رودخانه نیز الگوی تعریف شده براساس ترکیب گام‌های زمانی فعلی و پیشین بارش- رواناب، بهترین نتیجه را به‌دنبال داشت که از آن جمله می‌توان به تحقیقات (Jayawardena et al., 2005); (Ghorbani and Dehghani, 2015); (Nourani et al., 2009) اشاره نمود. مقایسه نتایج مدل‌های مختلف به‌ازای بهترین الگو حاکی از آن است که مدل WNN با بیشترین ضریب همبستگی (R=۰/۹۵۲)، کمترین ریشه میانگین مربعات خطا (RMSE=۱/۵۸۹) و همچنین ضریب نش ساتکلیف برابر ۰/۹۰۵ در مرحله صحت‌سنجی عملکرد بهتری را نسبت به سایر مدل‌ها در تخمین جریان رودخانه از خود نشان داده است. علاوه بر آن، شاخص‌های آماری مربوط به مرحله آموزش مدل مذکور به مقادیر مطلوب نزدیک‌تر است. این درحالی است که عملکرد سایر مدل‌ها تقریباً مشابه یکدیگر می‌باشد. از سوی دیگر مقدار ضریب نش ساتکلیف مراحل آموزش و صحت‌سنجی کلیه مدل‌های مورد مطالعه در تحقیق حاضر به‌ازای الگوهای مطلوب، بزرگتر از ۰/۸ حاصل شده است که حاکی از عملکرد قابل قبول آن‌ها می‌باشد (He et al., 2014; Shu and Ouarda, 2008).

جدول ٦- مقادیر شاخص‌های آماری عملکرد مدل LS-SVM برای الگوهای مختلف در مراحل آموزش و صحت سنجی

الگو	آموزش			صحت‌سنجی		
	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف
۱	۰/۱۰۰	۱۲/۲۹۹	۰/۰۱۰	۰/۰۴۴	۷/۸۱۴	-۰/۰۱۷
۲	۰/۱۳۷	۱۲/۲۴۵	۰/۰۱۸	۰/۰۵۶	۷/۸۳۳	-۰/۰۳۲
۳	۰/۱۸۲	۱۲/۱۵۵	۰/۰۳۳	۰/۰۷۸	۷/۸۶۷	-۰/۰۵۳
۴	۰/۹۲۸	۴/۵۸۶	۰/۸۶۲	۰/۹۳۳	۱/۹۳۸	۰/۸۵۹
۵	۰/۹۳۰	۴/۵۱۹	۰/۸۶۶	۰/۹۲۹	۲/۰۰۹	۰/۸۴۹
۶	۰/۹۳۰	۴/۵۲۰	۰/۸۶۶	۰/۹۳۰	۱/۹۹۱	۰/۸۵۱

جدول ۷- مقادیر شاخص‌های آماری عملکرد مدل WNN برای الگوهای مختلف در مراحل آموزش و صحت سنجی

	آموزش			صحت‌سنجی		
الگو	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف
۱	۰/۲۳۸	۱۲/۰۲۷	۰/۰۵۶	۰/۱۳۱	۸/۰۶۰	-۰/۱۰۵
۲	۰/۲۵۱	۱۲/۰۱۷	۰/۰۶۲	۰/۱۳۲	۸/۰۰۸	-۰/۱۲۴
۳	۰/۳۳۶	۱۱/۶۳۹	۰/۱۱۳	۰/۱۳۴	۸/۱۱۱	-۰/۱۲۵
۴	۰/۹۷۴	۲/۸۰۹	۰/۹۴۸	۰/۹۵۲	۱/۵۸۹	۰/۹۰۵
۵	۰/۹۷۲	۲/۸۶۸	۰/۹۴۶	۰/۹۴۲	۱/۷۴۰	۰/۸۸۶
۶	۰/۹۷۴	۲/۸۴۱	۰/۹۴۷	۰/۹۴۷	۱/۷۲۸	۰/۸۸۹

جدول ۸- مقادیر شاخص‌های آماری عملکرد مدل GEP برای الگوهای مختلف در مراحل آموزش و صحت سنجی

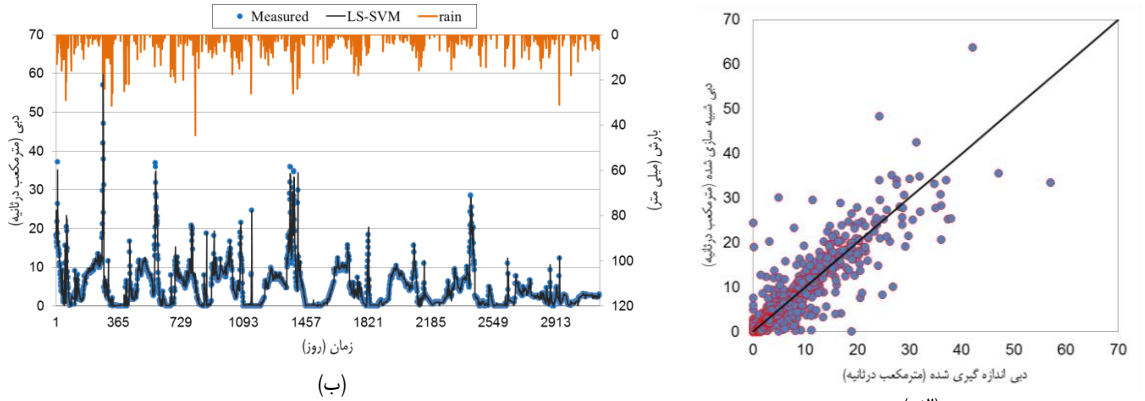
	آموزش			صحت‌سنجی		
الگو	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف
۱	۰/۱۷۵	۱۲/۱۳۹	۰/۰۳۰	۰/۰۷	۷/۴۶۳	-۰/۰۱۷
۲	۰/۲۰۹	۱۲/۰۵۲	۰/۰۴۳	۰/۰۹۵	۷/۷۹	-۰/۰۲۸
۳	۰/۱۲۷	۱۲/۲۲۹	۰/۰۱۶	۰/۰۱۴۵	۷/۷۰۶	-۰/۰۳۵
۴	۰/۹۱۶	۴/۹۳۹	۰/۸۳۹	۰/۹۲۸	۱/۹۸۴	۰/۸۵۳
۵	۰/۹۲۷	۵/۰۲۹	۰/۸۳۳	۰/۹۳۰	۱/۹۳۱	۰/۸۶۰
۶	۰/۷۸۰	۱/۴۱۰	۰/۸۳۴	۰/۹۲۸	۱/۹۳۶	۰/۸۵۹

جدول ۹- مقادیر شاخص‌های آماری عملکرد مدل ANNs برای الگوهای مختلف در مراحل آموزش و صحت سنجی

	آموزش			صحت‌سنجی		
الگو	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف	ضریب همبستگی (R)	ریشه میانگین مربعات خطا (m^3/s)	نش ساتکلیف
۱	۰/۱۳۰	۱۲/۲۵۴	۰/۰۱۷	۰/۰۷۷	۷/۷۷۲	-۰/۰۱۸
۲	۰/۲۴۰	۱۲/۰۰۰	۰/۰۵۷	۰/۰۸۷	۷/۷۷۰	-۰/۰۴۲
۳	۰/۲۲۳	۱۲/۰۵۰	۰/۰۴۹	۰/۱۲۷	۷/۷۶۳	-۰/۰۵۵
۴	۰/۹۳۰	۴/۳۵۰	۰/۸۷۵	۰/۹۳۰	۱/۹۱۱	۰/۸۶۰
۵	۰/۹۳۰	۴/۶۳۰	۰/۸۵۹	۰/۹۳۰	۱/۹۳۴	۰/۸۶۱
۶	۰/۹۲۸	۴/۶۱۰	۰/۸۶۱	۰/۹۳۰	۱/۹۳۲	۰/۸۶۱

ارزیابی مدل‌ها بر اساس نمودارهای پراکنش و سری زمانی

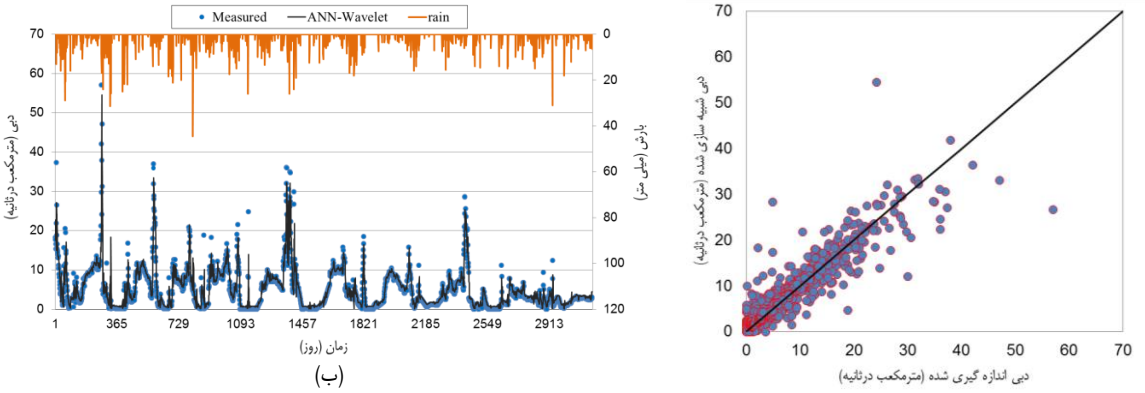
در شکل‌های (۲) تا (۵) نمودارهای پراکنش حول خط یک به یک (y=x) و سری زمانی مدل‌های مورد استفاده در تحقیق حاضر برای مرحله صحت‌سنجی و به‌ازای بهترین الگو نمایش داده شده است. همان‌طور که ملاحظه می‌شود در کلیه مدل‌ها پراکنش مناسبی حول محور یک به یک دیده می‌شود و در نمودارهای سری زمانی آن‌ها هم‌پوشانی قابل قبولی بین مقادیر مشاهداتی و شبیه‌سازی وجود دارد. از سوی دیگر با توجه به نمودارهای سری زمانی برای یک روز معین، تأثیر بارش بر مقدار جریان رودخانه کاملاً محسوس است، به‌طوری‌که فراز و نشیب مقدار دبی با نوسانات بارش همسو می‌باشد. برتری الگوی ۴ را نیز می‌توان به این موضوع مرتبط دانست.



(ب) (الف)

شکل ۲- نمودارهای حاصل از مدل LS-SVM به‌ازای بهترین الگو در مرحله صحت‌سنجی

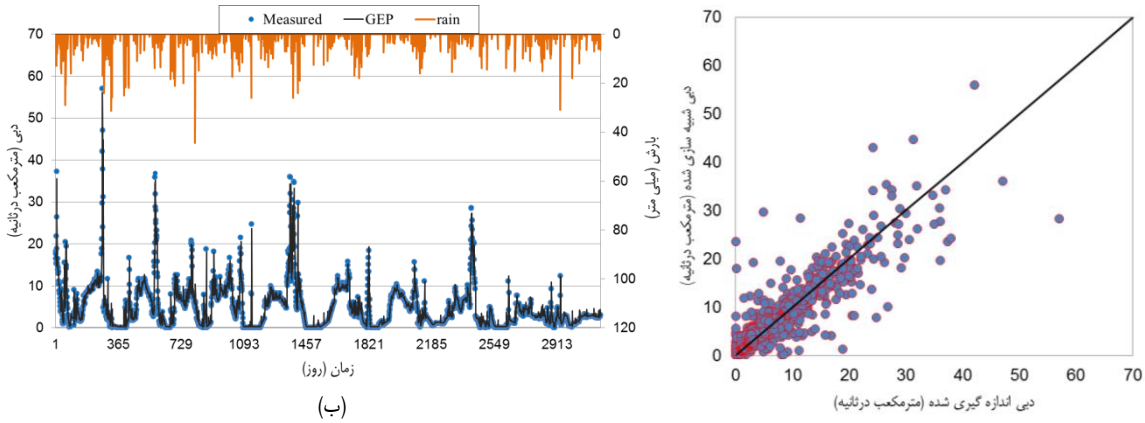
(الف)- پراکنش حول محور یک به یک (ب)- سری زمانی مقادیر مشاهداتی و شبیه‌سازی شده



(ب) (الف)

شکل ۳- نمودارهای حاصل از مدل WNN به‌ازای بهترین الگو در مرحله صحت‌سنجی

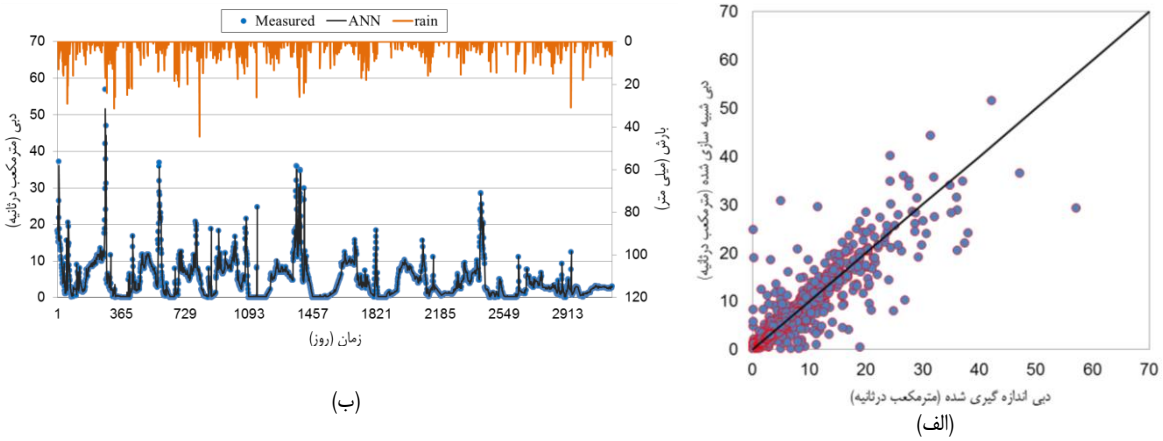
(الف)- پراکنش حول محور یک به یک (ب)- سری زمانی مقادیر مشاهداتی و شبیه‌سازی شده



(ب) (الف)

شکل ٤- نمودارهای حاصل از مدل GEP به‌ازای بهترین الگو در مرحله صحت‌سنجی

(الف)- پراکنش حول محور یک به یک (ب)- سری زمانی مقادیر مشاهداتی و شبیه‌سازی شده



(ب)

(الف)

شکل ۵- نمودارهای حاصل از مدل ANN به‌ازای بهترین الگو در مرحله صحت‌سنجی

(الف)- پراکنش حول محور یک به یک (ب)- سری زمانی مقادیر مشاهداتی و شبیه‌سازی شده

ارزیابی مدل‌ها در پیش‌بینی دبی بیشینه (سیلاب)

مقادیر بیشینه دبی به‌دست آمده از بهترین الگوهای مدل‌های مختلف و مقادیر مشاهداتی متناظر آنها به‌همراه قدر مطلق خطای نسبی هر مدل در جدول (۱۰) ارائه شده است. لازم به ذکر است که برای محاسبه خطای نسبی از رابطه (۱۳) استفاده شد. در رابطه مذکور، RE معرف خطای نسبی، Q_o و Q_c به‌ترتیب مقادیر دبی بیشینه مشاهداتی و شبیه‌سازی شده می‌باشد.

$$RE(\%) = \left| \frac{Q_o - Q_c}{Q_o} \right| \times 100 \qquad (۱۳)$$

تخمین دبی بیشینه جریان (دبی سیلاب) در بازه‌های پایین‌دست رودخانه، به‌منظور مدیریت و کنترل سیلاب بسیار حائز اهمیت می‌باشد. با توجه به نتایج جدول (۹) مشاهده می‌شود که خطای مدل‌ها در برآورد مقادیر دبی بیشینه تقریباً یکسان بوده و کمترین خطا به مقدار ۲۵/۹۷ درصد مربوط به مدل WNN می‌باشد.

جدول ۱۰- کارآیی مدل‌های مورد مطالعه در شبیه‌سازی دبی بیشینه جریان

مقادیر دبی بیشینه مشاهداتی	مقادیر شبیه‌سازی شده				خطای نسبی			
	ماشین بردار پشتیبان	برنامه‌ریزی بیان ژن	شبکه عصبی	موجک- عصبی	ماشین بردار پشتیبان	برنامه‌ریزی بیان ژن	شبکه عصبی	موجک- عصبی
۵۷/۰۰	۳۳/۵۰	۲۹/۱۸	۲۹/۳۷	۳۰/۴۶	۴۱/۲۳	۴۸/۸۱	۴۸/۴۷	۴۶/۵۶
۴۷/۰۶	۳۶/۵۱	۳۶/۰۷	۳۶/۶۳	۲۷/۴۰	۲۲/۴۲	۲۳/۳۵	۲۲/۱۶	۴۱/۷۸
۴۲/۰۶	۵۳/۷۵	۵۴/۵۶	۵۱/۶۸	۳۹/۰۲	۲۷/۷۹	۲۹/۷۲	۲۲/۸۷	۷/۲۳
۳۷/۸۸	۲۴/۲۹	۲۳/۱۳	۲۴/۲۷	۳۶/۸۸	۳۵/۸۸	۳۸/۹۴	۳۵/۹۳	۲/۶۴
۳۷/۲۴	۲۲/۲۰	۲۱/۰۵	۲۲/۱۲	۲۸/۷۷	۴۰/۳۹	۴۳/۴۷	۴۰/۶۰	۲۲/۷۴
۳۶/۹۰	۳۴/۸۸	۳۴/۲۶	۳۴/۹۹	۲۷/۵۰	۵/۴۷	۷/۱۵	۵/۱۸	۲۵/۴۷
۳۶/۰۰	۲۸/۸۹	۲۷/۸۴	۲۸/۹۶	۲۳/۸۱	۱۹/۷۵	۲۲/۶۷	۱۹/۵۶	۳۳/۸۶
۳۵/۹۵	۲۰/۷۳	۱۹/۶۱	۲۰/۶۰	۲۰/۸۵	۴۲/۳۴	۴۵/۴۵	۴۲/۷۰	۴۲/۰۰
۳۵/۹۰	۳۱/۴۷	۳۰/۵۶	۳۱/۵۵	۳۱/۷۹	۱۲/۳۴	۱۴/۸۷	۱۲/۱۲	۱۱/۴۵
				میانگین قدرمطلق خطا	۲۷/۵۱	۳۰/۴۹	۲۷/۷۳	۲۵/۹۷

نتیجه‌گیری

در تحقیق حاضر کارآیی مدل‌های مختلف هوشمند در تخمین جریان روزانه رودخانه دره‌رود مورد بررسی قرار گرفت. در این راستا حجم وسیعی از مقادیر دبی و بارش روزانه مشاهداتی طی دوره آماری ۲۶ ساله، در ایستگاه هیدرومتری مشیران واقع بر روی رودخانه مذکور مورد استفاده قرار گرفت. به‌منظور شبیه‌سازی جریان رودخانه شش الگوی متفاوت تعریف شد. در کلیه مدل‌ها با ترکیب بارش و دبی دقت تخمین جریان روزانه بهبود یافت. چنانچه بیشترین دقت مدل‌ها در ترکیب ورودی رواناب با یک تأخیر زمانی و بارش روزانه فراهم شد. این نشان دهنده تأثیرپذیری زیاد جریان حوضه آبریز دره‌رود از دبی روز قبل و بارندگی همان روز می‌باشد. در این راستا میزان تأثیرپذیری دبی جریان از دبی روز قبل به مراتب بیشتر از میزان بارندگی می‌باشد و لزوماً افزایش مقدار بارندگی موجب افزایش دبی نمی‌گردد. مقادیر ضرایب همبستگی (جدول ۲) نیز به لحاظ آماری حاکی از تأثیرپذیری کمتر دبی جریان از متغیر بارش می‌باشد. بررسی شاخص‌های آماری نشان داد که مدل‌های مورد استفاده در تحقیق حاضر از عملکرد قابل قبول در برآورد روزانه رودخانه دره‌رود برخوردار می‌باشند. همچنین نظر به دقت بالاتر مدل WNN می‌توان از مدل مذکور به‌عنوان ابزاری کارآمد و قابل اعتماد در تخمین جریان رودخانه مذکور استفاده نمود. کارآیی مدل‌های مختلف در پیش‌بینی دبی بیشینه نیز مورد بررسی قرار گرفت. نتایج حاصله دلالت بر این داشت که عملکرد مدل‌ها در این خصوص مشابه یکدیگر بوده و در عین حال مدل WNN در تخمین مقادیر پیک جریان در مقایسه با سایر مدل‌ها به‌صورت جزئی عملکرد بهتری را دارد. درمجموع با توجه به نتایج قابل قبول مدل‌های مورد استفاده در تحقیق حاضر، ضرایب و توابع به‌کار رفته برای واسنجی آن‌ها می‌تواند برای تخمین جریان حوضه آبریز دره‌رود بسیار مفید باشد.

منابع

1- Aytek, A. and Alp, M., 2008. An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science, 117*(2), pp.145-155.

2- Behzad, M. Asghari, K. Eazi, M. and Palhang, M., 2009. Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems with Applications, 36*(4), pp.7624-7629.

3- Botsis, D. Latinopoulos, P. and Diamantaras, K., 2011. Rainfall-Runoff Moeling Using Suport Vector Regression and Artificial Neural Networks. In *12th International Conference on Environmental Science and Technology, Rhodes, Greece.*

4- Ferreira, C., 2006. Automatically defined functions in gene expression programming. In *Genetic Systems Programming.* Springer Berlin Heidelberg.

5- Ferreira, C., 2001. Gene Expression Programming: a new adaptive algorithm for solving problems. *Complex Systems, 13*(2), pp.87–129.

6- Ghorbani, M.A. and Dehghani, R., 2015. Application of bayesian neural networks, support vector machines and gene expression programming analysis of rainfall - runoff monthly (case study: Kakarza river). *Irrigation Sciences and Engineering, 39*(2), pp.125–138. (In Persian).

7- He, Z. Wen, X. Liu, H. and Du, J., 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *Journal of Hydrology, 509*, pp.379-386.

8- Jayawardena, A.W. Muttil, N. and Fernando, T.M.K.G., 2005. Rainfall-runoff modeling using genetic programming. In *International Congress on Modeling and Simulation Society of Australia and New Zealand*.

9- Kisi, O., 2008. Stream flow forecasting using neuro-wavelet technique. *Journal of Hydrological Processes, 22*(20), pp.4142-4152.

10- Marofi, S. Amir-Moradi, K. and Parsafar, N., 2012. River flow prediction using Artificial Neural Network and Wavelet Neural Network models (Case study: Barandozchay River). *Water and Soil Science, 23*(3), pp.93–103. (In Persian).

11- Nabizadeh, M. Mosaedi, A. and Dehghani, A.A., 2012. Intelligent estimation of stream flow by Adaptive Neuro-Fuzzy Inference System). *Water and Irrigation Management, 2*(1), pp.69–80. (In Persian).

12- Nourani, V. Komasi, M. and Mano, A., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. *Journal of Water Resources Management, 23*(14), pp.2877-2894.

13- Rezaei, E. Khashei-Siuki, A. and Shahidi A., 2014. Design of groundwater level monitoring network, using the model of least squares support vector machine (LS-SVM). *Iranian Journal of Soil and Water Research, 45*(4), pp.389-396. (In Persian).

14- Salajegheh, A. Fathabadi, A. and Mahdavi, M., 2009. Investigation on the efficiency of neuro-fuzzy method and statistical models in simulation of rainfall-runoff process. *Journal of Range and Watershed Management, 62*(1), pp.65-79.

15- Shafiee-Jood, M. Abrishamchi, A. and Salavitabar, A., 2012. Evaluation of water resources development plans in multi-reservoir darehrood sub catchment system using performance indices. *Journal of Water and Wastewater, 3*, pp.22-32. (In Persian).

16- Shu, C. and Ouarda, T.B.M.J., 2008. Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *Journal of Hydrology, 349*(1), pp.31-43.

17- Soltani, A. Ghorbani, M.A. Fakheri-Fard, A. Darbandi, S. and Farsadizadeh, D., 2010. Genetic programming andi application in rainfall-runoff modeling. *Water and Soil Science, 20*(4), pp.62–71. (In Persian).

18- Suykens, J.A. De Brabanter, J. Lukas, L. and Vandewalle, J., 2002. Weighted least squares support vector machines, pp. robustness and sparse approximation. *Neurocomputing, 48*(1), pp.85-105.

19- Taheri, H. and Ghafouri M., 2012. Comparison between active learning method and support vector machine for runoff modeling. *Journal of Hydrology and Hydromechanics, 60*(1), pp.16-32.

20- Tokar, A.S. and Markus, M., 2000. Precipitation-runoff modeling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering, 5*(2), pp.156-161.

21- Vapnic, V.N., 1998. *Statistical Learning Theory*. Wiley, New York, USA.

22- Yu, P.S. Chen, S.T. and Chang, I.F., 2006. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology, 328*(3), pp.704-716.

Shahid Chamran University of Ahvaz
Irrigation Sciences and Engineering (JISE) **Vol. 41, No. 4, Winter 2019, p. 133-146**

EXTENDED ABSTRACT

Daily Rainfall – Runoff Modeling of Darreh-Rud River in Ardabil Province, Iran

M. R. Nikpour[1*], H. Sanikhani[2], S. Mahmodi Babelan[3] and S. Nastarani Amuqin[3]

1*- Corresponding Author, Assistant Professor, Department of Water Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. (*m_nikpour@uma.ac.ir*).
2- Assistant Professor, Department of Water sciences and Engineering, Faculty of Agriculture, University of Kurdistan, Sanandaj, Iran.
3- Graduate Student, Department of Water Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran.

Received:27 January 2017 Revised: 23 April 2017 Accepted: 29 April 2017

Keywords: Gene Expression programming, Estimation of Streamflow, Least-squared Support Vector Machines, Artificial Neural Networks, Hybrid of Wavelet-ANN.

Introduction

Rainfall- runoff is one of the most complicated hydrological processes that is affected by various physical and hydrological variables. Therefore, understanding and predicting runoff formation processes and their transfer to the outlet point of the watershed is one of the most important issues of hydrological sciences (Salajegheh et al., 2009). Botsis et al., (2011) simulated daily rainfall-runoff of a catchment in the California, USA. They compared the performance of SVM with three types of kernel function with ANN. Finally, the SVM had a more accurate simulation of rainfall-runoff. Nourani et al., (2009) applied the hybrid of wavelet-ANN to model rainfall-runoff of Lighvan-Chay basin in Iran. The results showed that the proposed model is capable to predict long-term and short-term rainfall events due to the use of the time series with multiple scales as input layer of ANN. Darreh-rud river as the most important branch of Aras border region in Iran, is one of the main rivers of Ardabil province and the main source of water supply in different parts of the province. On the other hand, the Emarat reservoir dam is under construction on the river. Therefore, ANN, WNN, GEP and LS-SVM models were evaluated for estimating the inflow of Emarat dam (located on Darreh-rud river, Ardabil province).

Methodology

In the present study, daily runoff and rainfall data from the Moshiran hydrometric station located in river and the upstream of Emarat dam for a period of 26 years (i.e. 1988-2014) was used. In order to select the appropriate combination of the input variables of the models, statistical tests were used and the partial auto-correlation coefficients of discharge and cross-correlation values between rainfall-discharge variables were calculated. The correlation values between the discharge, rainfall and related delays are presented in Table (1). In the table, the R and Q variables indicate the measured values of rainfall and discharge at Moshiran hydrometric station, respectively, and the t-index is the current time step. Considering the cross and significant correlation between input-output variables, different input patterns were used to estimate the daily runoff of the Darreh-rud river according to Table (2).

Table 1- The correlation values between rainfall-discharge variables

	Q_{t-1}	R_t	R_{t-1}	R_{t-2}
Q_t	0.916	0.107	0.126	0.115

Table 2- The structure of patterns used in this study

Pattern No.	*Input combination*	*output*
1	R_t	Q_t
2	R_{t-1}, R_t	Q_t
3	R_{t-2}, R_{t-1}, R_t	Q_t
4	Q_{t-1}, R_t	Q_t
5	Q_{t-1}, R_{t-1}, R_t	Q_t
6	Q_{t-1},R_{t-2}, R_{t-1}, R_t	Q_t

Results and Discussion

In the Tables (3-6), the statistical measures of the performance for the applied models in the present study are reported for the train and test stages based of different patterns. As can be seen from Tables (5-8), for GEP model, in the pattern No. 5 and for other models, in the pattern No. 4 provide the best performance. Pattern No. 4 requires a smaller number of inputs compared to other patterns, and its important advantage is to reduce the runtime of the models and their complexity. Therefore, the discharge of the river in the current time generally is influenced by its value on the previous day and the rainfall in the current time. It is noted that in all models, patterns No. 1-3 present poor results, which suggests that in the Darreh-rud watershed, using only rainfall data cannot provide proper estimation of run-off. Comparison of the results of different models using the best pattern indicates that the WNN model with the highest correlation coefficient (R = 0.982), the lowest root mean square error (RMSE = 1.589), and the NS equal to 0.905 in the validation step has shown better performance than other models in river flow estimation. In addition, the statistical criteria related to the training stage of the mentioned model are closer to the desired values. This is while the performance of other models is roughly the same. On the other hand, the NS values of training and validation phases of all the models applied in the present study are higher than 0.8 through optimal patterns, indicating their acceptable performance. The results showed that the error of the models in the estimation of the maximum discharge values was approximately the same and the least error (i.e. 25.97%) was related to the WNN model.

Table 3- The statistical measures of LS-SVM for diffrent patterns

Pattern No.	*Train*			*Test*		
	R	*RMSE (m^3/s)*	*NS*	*R*	*RMSE (m^3/s)*	*NS*
1	0.100	12.299	0.010	0.044	7.814	-0.017
2	0.137	12.245	0.018	0.56	7.833	-0.032
3	0.182	12.155	0.033	0.078	7.867	-0.053
4	**0.928**	**4.586**	**0.862**	**0.933**	**1.938**	**0.859**
5	0.930	4.159	0.866	0.929	2.009	0.849
6	0.930	4.520	0.866	0.930	1.991	0.851

Table 4- The statistical measures of WNN for diffrent patterns

Pattern No.	*Train*			*Test*		
	R	*RMSE (m^3/s)*	*NS*	*R*	*RMSE (m^3/s)*	*NS*
1	0.238	12.027	0.056	0.131	8.060	-0.105
2	0.251	12.017	0.062	0.132	8.008	-0.124
3	0.336	11.639	0.113	0.134	8.111	-0.125
4	**0.974**	**2.809**	**0.948**	**0.952**	**1.589**	**0.905**
5	0.972	2.868	0.946	0.942	1.740	0.886
6	0.974	2.841	0.947	0.947	1.728	0.889

Table 5- The statistical measures of GEP for diffrent patterns

Pattern No.	*Train*			*Test*		
	R	*RMSE (m^3/s)*	*NS*	*R*	*RMSE (m^3/s)*	*NS*
1	0.175	12.139	0.030	0.070	7.463	-.017
2	0.209	12.052	0.043	0.095	7.790	-0.028
3	0.127	12.229	0.016	0.015	7.706	-0.035
4	0.916	4.939	0.839	0.928	1.984	0.853
5	**0.927**	**5.029**	**0.833**	**0.930**	**1.931**	**0.860**
6	0.780	1.410	0.834	0.928	1.936	0.859

Table 6- The statistical measures of ANN for diffrent patterns

Pattern No.	*Train*			*Test*		
	R	*RMSE (m^3/s)*	*NS*	*R*	*RMSE (m^3/s)*	*NS*
1	0.130	12.254	0.017	0.077	7.772	-.018
2	0.240	12.000	0.057	0.087	7.770	-0.042
3	0.223	12.050	0.049	0.127	7.763	-0.055
4	**0.930**	**4.350**	**0.875**	**0.930**	**1.911**	**0.860**
5	0.930	4.630	0.859	0.930	1.934	0.861
6	0.928	4.610	0.861	0.930	1.932	0.861

Conclusions

In the present study, the efficiency of different intelligent models was investigated in daily river flow estimation of Darreh-rud river. In this regard, a long period of daily discharge and rainfall (i.e. 26 years) at the Moshiran hydrometric station located on the river were used. In order to simulate the streamflow, six different patterns were considered. For all models, the accuracy of daily discharge estimation was improved by combining rainfall and discharge variables. As the highest accuracy of the models in the prediction of run-off was provided with one lag of run-off and rainfall. This indicates the high impact of run-off in the Darreh-rud river from the previous day of run-off and current day of rainfall. In this regard, the effect of the discharge from the previous day of run-off was more than the rainfall, and necessarily increase the amount of rainfall does not increase the discharge. The values of correlation coefficients (Table 1) also statistically indicate the lower influence of discharge from rainfall variable. Statistical analysis showed that the models used in this study have acceptable performance in daily estimation of Darreh-rud river. Furthermore, due to the higher accuracy of the WNN model, we can use this model as an efficient and reliable tool to estimate the streamflow in the studied region. Also, the results indicate that the performance of the models for predicting maximum discharge is somewhat similar, while the WNN model has a better performance in this regard.

References

1- Botsis, D. Latinopoulos, P. and Diamantaras, K., 2011. Rainfall-Runoff Moeling Using Suport Vector Regression and Artificial Neural Networks. In 12th International Conference on Environmental Science and Technology, Rhodes, Greece.

2- Nourani, V. Komasi, M. and Mano, A., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. *Journal of Water resources management*, *23*(14), pp.2877-2894.

3- Salajegheh, A. Fathabadi, A. and Mahdavi, M., 2009. Investigation on the efficiency of neuro-fuzzy method and statistical models in simulation of rainfall-runoff process. *Journal of Range and Watershed Management*, *62*(1), pp.65-79.