

## پیش‌بینی جریان روزانه در حوضه آبریز رودخانه تلوار با مدل‌های هوشمند

مطلب بایزیدی<sup>۱\*</sup>، فرخ اسدزاده<sup>۲</sup>، مه‌ری کاکی<sup>۳</sup>

۱. استادیار گروه مهندسی آب دانشگاه آزاد اسلامی واحد سنندج

۲. استادیار گروه علوم خاک دانشگاه ارومیه

۳. کارشناسی ارشد مهندسی منابع آب، دانشگاه تبریز

(تاریخ دریافت ۱۳۹۶/۰۱/۲۸ تاریخ تصویب ۱۳۹۶/۰۹/۱۰)

## چکیده

دبی رودخانه‌ها به دلیل تغییرات درخور توجه زمانی و مکانی در مطالعات هیدرولوژی و برنامه‌ریزی مدیریت منابع آب اهمیت فراوانی دارد. در علوم مهندسی آب، سعی بر این است که با استفاده از سری زمانی داده‌های ثبت‌شده، اقدام به برآورد جریان رودخانه شود. افزایش دقت پیش‌بینی‌های یادشده می‌تواند در برنامه‌ریزی دقیق تخصیص منابع آب مؤثر باشد. در این پژوهش از سه روش برنامه‌ریزی بیان ژن، روش حداقل مربعات ماشین بردار پشتیبان و شبکه عصبی موجکی برای برآورد جریان رودخانه تلوار در حوضه آبخیز قروه- دهگلان استفاده شده و کارایی آنها مقایسه شد. به این منظور داده‌های ۱۰ ساله دبی رودخانه (۱۳۸۱-۱۳۹۱) به منظور مدل‌سازی استفاده شد. برای مدل‌سازی با روش‌های یادشده، ۷۵ درصد از داده‌ها به عنوان داده‌های آموزش و ۲۵ درصد نیز به منزله داده‌های اعتبارسنجی مدل استفاده شدند. نتایج به دست آمده نشان داد پیش‌بینی مدل جریان با تأخیر دو روز قبل جریان در رودخانه در هر سه روش دقت قابل قبولی دارد، به طوری که معیار ضریب تبیین در مرحله اعتبارسنجی در بهترین مدل به ترتیب برای برنامه‌ریزی بیان ژن، ماشین بردار پشتیبان و شبکه عصبی موجک برابر با ۰/۹۴، ۰/۹۲ و ۰/۹۰ بودند. همچنین، مقایسه سه مدل نشان داد روش ماشین بردار پشتیبان نسبت به دو روش دیگر عملکرد بهتری دارد و نیز دقت پیش‌بینی در این مدل با افزایش تعداد تأخیرها به چهار و پنج روز قبل افزایش می‌یابد.

**کلیدواژگان:** تخمین جریان، رودخانه تلوار، شبکه عصبی موجک، مدل برنامه‌ریزی بیان ژن، مدل LS-SVM.

## مقدمه

مه‌آب‌های سطحی، اهمیت فراوانی به‌ویژه در مناطق خشک و نیمه‌خشک دارد. در کشور ما نیز بیشتر رودخانه‌ها فصلی است و کمبود آب در پهنه وسیعی از کشور وجود دارد، بنابراین شناسایی و مدل‌سازی رفتار جریان‌های سطحی آب به‌منظور برنامه‌ریزی‌های بلندمدت و استفاده بیشتر و بهتر از پتانسیل آنها به‌عنوان یک نیاز اساسی محسوب می‌شود. از سوی دیگر، تخمین جریان رودخانه در مطالعات مربوط به مهندسی منابع آب، به‌عنوان مبنای اساسی برای پیاده‌سازی برنامه‌های کاربردی نظیر پیش‌بینی سیل، برنامه‌ریزی آبیاری، کنترل حجم رسوب و آب در مخازن و در نهایت تولید برقایی است [۱].

تغییرات سطوح جریان رودخانه‌ها متأثر از عوامل گوناگونی مانند حجم بارش و تخلیه مستقیم و یا غیرمستقیم آن از حوضه‌های مجاور، تبخیر، دما و برهم‌کنش بین جریان رودخانه‌ها با سفره‌های کم‌عمق زیرزمینی است. هر چند استفاده از متغیرهای یادشده می‌تواند برای ایجاد الگوهای پیچیده سطح جریان مفید باشد، مدل‌هایی که تغییرات جریان را بر اساس داده‌های تاریخی در دسترس، شبیه‌سازی می‌کنند نیز اهمیت فراوانی در هیدرولوژی و مدیریت منابع آب دارند [۲].

طی دو دهه اخیر استفاده از مدل‌های ریاضی و کامپیوتری در زمینه پدیده‌های غیرخطی از جمله متغیرهای مربوط به هیدرولوژی کاربرد گسترده‌ای یافته است. از جمله این روش‌ها می‌توان به روش‌های شبکه عصبی، هیبرید عصبی موجکی، ماشین بردار پشتیبان و برنامه‌ریزی بیان ژن در پیش‌بینی و برآورد جریان در حوضه‌های آبریز اشاره کرد. در این زمینه می‌توان به پژوهش‌هایی که در داخل و خارج از ایران نیز درباره پیش‌بینی با مدل‌های ریاضی و هوشمند انجام داده‌اند اشاره کرد؛ ویگام و کراپر [۳] الگوی بارش-رواناب روزانه را با استفاده از برنامه ژنتیک مدل‌سازی کردند. نتایج پژوهش آنها نشان داد روش‌های یادگیری توسط ماشین به‌ویژه زمانی که حرکت آب‌های سطحی و تلفات آب کم باشد، ابزاری مفید برای توسعه مدل‌های هیدرولوژیکی است. نیک‌بخت و شهبازی [۴] با استفاده از مدل بردار پشتیبان سری‌های زمانی و اطلاعات یک دوره ۴۵ ساله (۱۳۳۰-۱۳۷۴) از جریان رودخانه‌های کشف‌رود (ایستگاه آق‌دربند) و هریرود (ایستگاه پل خاتون) برای تولید جریان مصنوعی

استفاده کردند. نتایج پژوهش آنها نشان داد ماشین بردار پشتیبان در حفظ هم‌زمان ویژگی‌های ماهیانه و سالیانه عملکرد مطلوبی داشته است و با مقایسه نتایج شبیه‌سازی توسط HEC4 و همچنین ماشین بردار پشتیبان مشاهده شد که این روش در تولید جریان مصنوعی عملکرد به‌مراتب بهتری از خود نشان داده است. یو و همکارانش [۵] با استفاده از بردار پشتیبان به پیش‌بینی سیل در مقیاس زمانی واقعی برای رودخانه یانگ در تایوان اقدام کردند. نتایج اعتبارسنجی سیل‌های رخ داده نشان داد مدل‌های ارائه‌شده با شش ساعت تأخیر می‌توانند مؤثر باشند. اووکا [۶] نیز مدل شبکه عصبی موجکی را برای پیش‌بینی جریان ورودی به مخزن سد کمیر در ترکیه به کار برد و برای پیش‌بینی جریان ورودی به مخزن سد از مقیاس زمانی ماهانه با ترکیب موجک گسسته و بهینه‌سازی الگوریتم‌های مبتنی بر روش لوبنبرگ-مارکوارت استفاده کرد. نتایج پژوهش ایشان نشان داد شبکه عصبی موجکی یک مدل مناسب برای پیش‌بینی جریان ورودی به مخزن سد است و می‌تواند نسبت به مدل رگرسیون چندگانه عملکرد مطلوب‌تری داشته باشد. سایاگوی و همکارانش [۷] مدل‌های تخمین جریان براساس LS-SVM و تکنیک‌های مدل‌سازی MT, M5P براساس مقیاس روزانه و با استفاده از داده‌های هیدرولوژیکی و هواشناسی را توسعه دادند. نتایج پژوهش آنها نشان داد هر دو مدل طی فرایند آزمایش به‌خوبی به کار گرفته شدند و میزان استثنا در این روند بسیار کم بود. عملکرد مدل‌های LS-SVM در مقایسه با مدل M5P در حوضه‌های با مقیاس بزرگ بهتر حاصل شد به‌طوری که مدل LS-SVM ضریب همبستگی بیشتر و کمترین خطا به‌ویژه مجذور میانگین مربعات خطا را داشت. هه و همکارانش [۸] نیز به مقایسه عملکرد مدل‌های شبکه عصبی مصنوعی، نرو فازی و ماشین بردار پشتیبان در مناطق کوهستانی نیمه‌خشک پرداختند. نتایج پژوهش آنها نشان داد عملکرد هر سه روش بر مبنای مقیاس متنوعی از معیارهای کارایی در مراحل آموزش و اعتبارسنجی تفاوت درخور توجهی نکرده و نتایج ارزیابی عملکرد هر سه روش مطلوب بود هر چند در مجموع روش ماشین بردار پشتیبان با اندکی اختلاف عملکرد بهتری از خود نشان داده است. براساس نتایج یادشده، این پژوهشگران روش SVM را برای پیش‌بینی جریان رودخانه در مناطق با توپوگرافی پیچیده مانند مناطق کوهستانی، به‌عنوان روشی مطلوب ارزیابی کرده‌اند. شنفی و همکارانش [۹] با

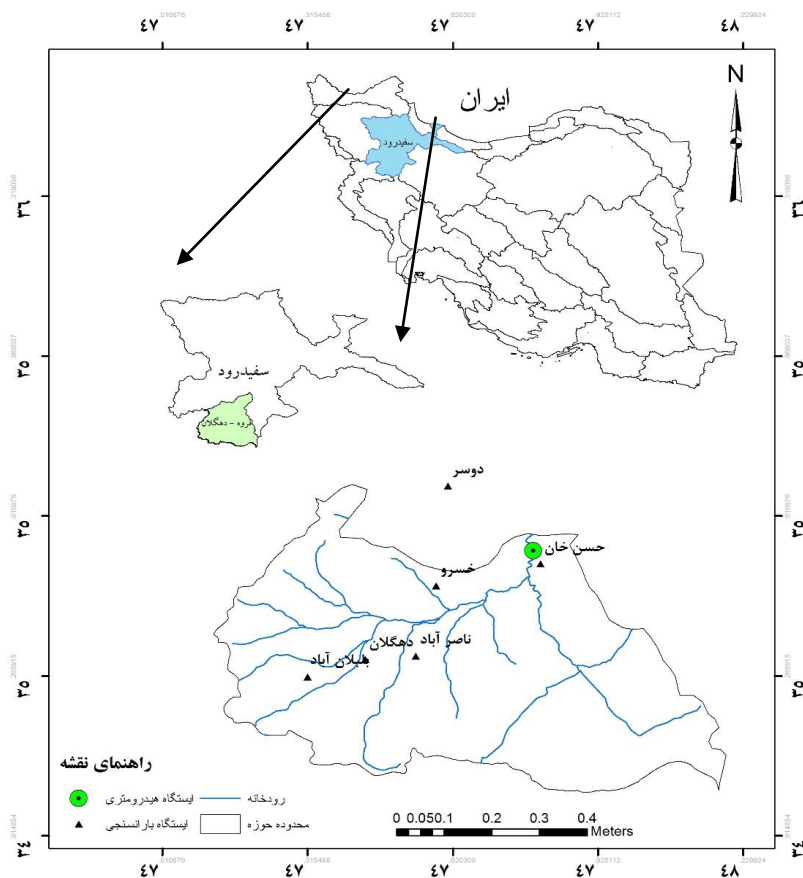
چهاردولی سرچشمه می‌گیرد و بعد از تلاقی با چم شور در بالادست قروه و رودهای دهگلان در حوضه بیجار با اوزن دره و رود شیرکش مخلوط می‌شود و در نهایت به رودخانه قزل‌اوزن می‌ریزد. مساحت این حوضه  $4136/5$  کیلومترمربع است و از نظر جغرافیایی بین طول‌های  $36^{\circ} 47'$  و  $36^{\circ} 47'$  شرقی و  $1^{\circ} 11'$  تا  $1^{\circ} 41'$  عرض شمالی واقع شده است. در شکل ۱ موقعیت حوضه آبخیز قروه-دهگلان و ایستگاه‌ها نشان داده شده است. متوسط بارندگی حوضه  $350$  میلی‌متر در سال و اقلیم آن براساس روش آمبرژه جزء نواحی نیمه‌خشک سرد است. پتانسیل آب‌های سطحی محدوده ضعیف است. دبی متوسط سالانه ایستگاه هیدرومتری تلوار-دهگلان در مرکز حوضه  $0/88$  مترمکعب بر ثانیه و خروجی حوضه در محل حسن‌خان  $5/9$  مترمکعب بر ثانیه است. در این پژوهش از داده‌های روزانه در بازه زمانی  $1381-1391$  ایستگاه هیدرومتری حسن‌خان استفاده شد. در جدول ۱ مشخصات آماری ایستگاه هیدرومتری حسن‌خان ارائه شده است.

استفاده از ماشین بردار پشتیبان با تبدیل موجک گسسته، بارش ماهانه در صربستان را برای دوره آماری  $1946-2012$  برآورد کردند. آنها از سه تکنیک شبکه عصبی، برنامه الگوریتم ژن (GP) و ماشین بردار پشتیبان با تبدیل موجک الگوریتم استفاده کردند. نتایج پژوهش آن‌ها نشان داد WT-SVM در مقایسه با مدل GP و شبکه عصبی عملکرد بهتری در برآورد بارش ماهانه دارد. در این پژوهش قابلیت سه مدل هوشمند برنامه‌ریزی بیان ژن، حداقل مربعات ماشین بردار پشتیبان (LS-SVM) و شبکه عصبی موجک برای تخمین جریان روزانه رودخانه تلوار بررسی شده است.

## مواد و روش‌ها

### منطقه مطالعه شده

حوضه آبخیز قروه-دهگلان یکی از زیرحوضه‌های حوضه آبخیز سفیدرود بزرگ است. شکل حوضه پهن است و زهکشی آن از طریق رودخانه چم تلوار انجام می‌شود که از کوه‌های کلیایی در جنوب دهگلان و پنجه‌علی در دشت



شکل ۱. موقعیت حوضه آبخیز قروه-دهگلان و ایستگاه‌ها در ایران

روش حداقل مربعات ماشین بردار پشتیبان (LS-SVM) ماشین بردار پشتیبان روشی است که در آن قوت‌های روش‌های آماری سنتی که بیشتر تئوری‌محور هستند و از نظر تجزیه و تحلیل ساده هستند، ترکیب می‌شود [۳]. در سال‌های اخیر رویکرد SVM<sup>۲</sup> به‌طور گسترده در حوضه‌های هیدرولوژی، مدیریت منابع آب و پیش‌بینی سری‌های زمانی استفاده شده است. در این پژوهش از جعبه ابزار LS-SVM1.8 در محیط MATLAB 2013 استفاده شد. الگوریتم LS-SVM در دو مرحله: ۱. آموزش و اعتبارسنجی؛ ۲. طراحی به‌کار گرفته شد. در این مطالعه از روش LS-SVM با قدرت تقریب دقیق دبی استفاده شد تا این روش بتواند خطای پیش‌بینی را به کمترین برساند. چارچوب رگرسیون حداقل مربعات ماشین بردار پشتیبان به این صورت است:

مجموعه داده‌های آموزشی را خواهیم داشت که  $X_i^p \in R^n$  به‌عنوان بردار ورودی و  $Y_i^p \in R^n$  به‌عنوان بردار خروجی معرفی می‌شوند. تابع رگرسیون غیرخطی در فضای وزن‌دهی اولیه به‌صورت رابطه ۱ قابل تعریف خواهد بود [۱۳]

$$y(x) = W^T \varphi(x_i) + b \quad (1)$$

در این رابطه  $W$  و  $b$  به‌ترتیب مقادیر وزن‌ها و بایاس تابع رگرسیون و  $\varphi(x_i)$  ترسیم غیرخطی ورودی‌ها در فضای ویژگی با ابعاد زیاد است [۱۴]. رگرسیون غیرخطی بر اساس رابطه بهینه‌سازی ۲ و با محدودیت رابطه ۳ حل می‌شود:

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

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

که در آنها،  $\gamma$ : پارامتر تنظیم‌کننده بخش خطاست و  $e_i$ : میزان خطای داده‌های آموزشی را نشان می‌دهد. در بخش اعتبارسنجی از Crossvalidate-10fold و تابع بهینه‌سازی simplex برای تعیین مقادیر و پارامترهای بهینه  $\gamma$  و  $\sigma$  در روش حداقل مربعات ماشین بردار پشتیبان استفاده شد.

#### شبکه عصبی موجک

شبکه‌های عصبی مبتنی بر ویولت که ویولت نامیده می‌شوند از ترکیب دو تئوری ویولت‌ها و شبکه‌های عصبی به‌وجود آمده‌اند [۱۵]. این شبکه‌ها هم از مزایا و

جدول ۱. مشخصات آماری میانگین دبی روزانه ( $m^3 s^{-1}$ ) رودخانه تلوار در ایستگاه حسن‌خان (۱۳۸۱-۱۳۹۱)

انحراف معیار	میانگین	واریانس	بیشینه	کمینه	متوسط
۲/۳۱	۱/۰۲	۵/۳۴	۶۹/۰۰	۰	دبی روزانه

#### مدل‌سازی دبی جریان

به‌منظور اجرای تحقیق، از داده‌های دبی روزانه حوضه آبخیز قروه-دهگلان استفاده شد. پس از بررسی داده‌ها در روش شبکه عصبی موجک برای کاهش حساسیت سامانه به بزرگی داده‌ها نرمال شده و در محدوده صفر تا یک قرار گرفتند [۱۰ و ۱۱]. سرانجام، از سال‌های آبی ۱۳۸۰-۱۳۸۱ تا ۱۳۸۷-۱۳۸۸ به‌عنوان داده‌های آموزشی (هفت سال) و از سال ۱۳۸۸-۱۳۸۹ تا ۱۳۹۰-۱۳۹۱ به‌عنوان داده‌های اعتبارسنجی (سه سال) استفاده شد. در فرایند مدل‌سازی با استفاده از برنامه‌ریزی بیان ژن از نرم‌افزار GeneXprotools 5، حداقل مربعات ماشین بردار پشتیبان و شبکه عصبی موجک با نرم‌افزار MATLAB2013 استفاده شد.

#### برنامه‌ریزی بیان ژن<sup>۱</sup>

این روش جزء روش‌های الگوریتم گردشی محسوب می‌شود که مبنای همه آنها براساس نظریه تکامل داروین استوار است [۱۲]. فرایند اجرای مدل‌سازی در برنامه‌ریزی بیان ژن به این ترتیب است که ابتدا یک جمعیت اولیه از توابع مرکب نشان‌دهنده مدل‌های پیش‌بینی، به‌صورت تصادفی در نظر گرفته می‌شود و هر یک از افراد جمعیت یادشده با استفاده از توابع برازش، ارزیابی می‌شوند. در هر مرحله از تولید توابع مراحل زیر برای انتخاب یک جمعیت جدید دنبال می‌شود: الف) یکی از عملگرهای عبور، جهش و کپی انتخاب می‌شود؛ ب) تعداد مناسبی از افراد جمعیت حاضر انتخاب می‌شوند؛ ج) از عملگر انتخابی برای تولید فرزند استفاده می‌شود؛ د) فرزند یادشده در یک جمعیت جدید وارد می‌شود؛ ه) مدل مد نظر با استفاده از برازش‌های مختلف ارزیابی می‌شود. در پژوهش حاضر از برنامه GeneXproTools5 برای توسعه و اجرای مدل‌های مبتنی بر برنامه‌ریزی ژنتیک استفاده شد.

$$W_{a,b}(f) = \tilde{f}_{(a,b)} = \Psi_{a,b}(t), f(t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (9)$$

در اینجا پارامتر  $a$  به‌عنوان پارامتر گسترش و متناسب با فرکانس است (به بیان دیگر، پارامتر تأخیر کوچک متناظر با فرکانس بالا و پارامتر تأخیر بزرگ متناظر با فرکانس پایین است) و پارامتر  $b$  به‌عنوان پارامتر انتقال و متناسب با زمان است.

در مجموع، گام‌های مهم در آموزش و صحت‌سنجی شبکه عصبی موجک به این ترتیب است: الف) داده‌های ورودی به دو دسته برای آموزش شبکه و صحت‌سنجی تقسیم می‌شوند؛ ب) موجک مادر با احراز شرایط یادشده پس از اعمال ضرایب انتقال و مقیاس مناسب به موجک فرزند تبدیل می‌شود؛ ج) توابع فعال‌سازی نورون‌های لایه پنهان شبکه عصبی با انواع موجک‌های فرزند جایگزین می‌شوند؛ د) شبکه عصبی ویولت ایجادشده با استفاده از مجموعه داده‌های مربوط به آموزش شبکه، آموزش داده می‌شود؛ ه) عملکرد کلی شبکه موجک با تجزیه و تحلیل چگونگی تخمین داده‌های صحت‌سنجی بررسی می‌شود و در نهایت با رضایت‌بخش بودن عملکرد شبکه مرحله آموزش به اتمام می‌رسد. در غیر این صورت، مراحل قبل تا دستیابی به بهترین حالت ارزیابی می‌شود [۱۷].

#### معیارهای ارزیابی مدل‌ها

برای روش‌های مدل‌سازی یادشده، ضرایب کارایی شامل  $R^2$ ، MAE، RMSE و NS (نش-ساتکلیف) محاسبه می‌شود و بر مبنای بهترین مقادیر ضرایب کارایی مدل مطلوب انتخاب شد. چگونگی محاسبه این معیارها به صورت روابط ۱۰-۱۳ است.

$$R^2 = \left[ \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right]^2 \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (11)$$

$$MAE = \left( \frac{\sum_{i=1}^n |O_i - P_i|}{n} \right) \quad (12)$$

$$NS = 1 - \left[ \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \right] \quad (13)$$

ویژگی‌های شبکه‌های عصبی برخوردارند و هم جذابیت‌ها و انعطاف‌پذیری‌ها و مبنای ریاضی مستحکم ویولت‌ها و آنالیز چندمقیاسی را دارند. در ویولت‌ها از دو دسته توابع ویولت و توابع مقیاس (scaling functions) استفاده می‌شود. یک خانواده از توابع مقیاس  $\phi(x)$  به صورت رابطه ۴ معرفی می‌شوند.

$$\phi_{m,k}(x) = \gamma^{-m/\gamma} \phi(\gamma^{-m}x - k) \quad m, k \in Z \quad (4)$$

اگر  $\phi(x)$  را یک تابع مقیاس مادر در نظر بگیریم، مقادیر  $2^m$  و  $k$  به ترتیب مربوط به گسترش و انتقال تابع مقیاس مادر هستند. توابع مقیاس رزولوشن  $m$ ،  $\phi_{m,k}(x)$  در اصل پایه‌های متعامد فضای برداری  $V_m$  در رزولوشن  $m$  هستند. به بیانی، فضای برداری  $V_m$  حاوی کلیه تقریب‌های تابع  $f(x)$  با رزولوشن  $m$  و توسط توابع  $\phi_{m,k}(x)$  است. بنابراین، فضاهای برداری  $V_m$  حاوی تقریب‌های مختلف تابع  $f(x)$  در رزولوشن‌های مختلف‌اند. حال اگر  $W_m$  را فضای برداری متعامد با  $V_m$  در رزولوشن  $m$  فرض کنیم آن‌گاه می‌توان دسته‌ای دیگر از پایه‌های متعامد فضای  $W_m$  که ویولت  $\Psi(x)$  نامیده می‌شوند را نیز به صورت رابطه ۵ بیان کرد [۱۴]:

$$\Psi_{m,k}(x) = \gamma^{-m/\gamma} \Psi(\gamma^{-m}x - k) \quad m, k \in Z \quad (5)$$

در حالت کلی همه توابع فیزیکی را می‌توان به کمک ویولت‌ها و توابع مقیاس به صورت روابط ۶ و ۷ بیان کرد [۱۵]:

$$f(x) = f_0(x) + \sum_{m=\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{m,k} \Psi_{m,k}(x) \quad (6)$$

$$f_0(x) = \sum_{k} a_{0,k} \phi_{0,k}(x) \quad (7)$$

روابط یادشده بیان می‌کنند که هر تابع فیزیکی را می‌توان ابتدا به کمک تعدادی توابع مقیاس در رزولوشن صفر تقریب زد و سپس به کمک توابع ویولت در رزولوشن‌های مختلف آن تقریب را تا میزان دقت مد نظر ادامه داد. شبکه عصبی ویولت بر مبنای روابط ۳ و ۴ شکل می‌گیرد و الگوریتم‌های آموزش مخصوص خود را دارد که در پژوهش وانگ و همکارانش [۱۶] به صورت کامل بررسی شده است. درخور یادآوری است در حالت کلی خانواده ویولت‌های پیوسته نیز به صورت رابطه ۸ بیان می‌شود:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad a, b \in R \quad (8)$$

و تبدیل ویولت برای توابع پیوسته به وسیله رابطه ۹ محاسبه می‌شود:

مقادیر در دامنه ۰/۳۶ — ۰/۷۵ نشان‌دهنده عملکرد رضایت‌بخش مدل است [۱۸ و ۱۹].

### یافته‌ها

در پژوهش حاضر برای مدل‌سازی جریان روزانه رودخانه تلوار با استفاده از برنامه‌ریزی ژن، حداقل مربعات ماشین بردار پشتیبان و شبکه عصبی مویک، داده‌های هفت سال (از یکم مهر ۱۳۸۱ تا سی‌ویکم شهریور ۱۳۸۸) برای آموزش و سه سال (از یک مهر ۱۳۸۹ تا سی‌ویکم شهریور ۱۳۹۱) به‌عنوان داده‌های اعتبارسنجی انتخاب شدند. با توجه به اینکه توالی دبی روزهای قبل در پیش‌بینی دبی امروز مد نظر بوده، صرفاً داده‌های دبی جریان با توالی برگشتی تا پنج روز قبل به عنوان داده‌های آموزشی، به‌صورت ترکیب‌های مختلف مطابق جدول ۲ استفاده شده است.

در روابط یادشده،  $P_i$  و  $O_i$  به ترتیب مقادیر مشاهده‌ای و پیش‌بینی شده توسط مدل،  $\bar{P}$  و  $\bar{O}$  به ترتیب میانگین مقادیر مشاهده‌ای و پیش‌بینی شده توسط مدل و  $n$  تعداد داده‌هاست. در روابط یادشده هرچه مقدار  $R^2$  (ضریب تبیین معادله خطی بین مقادیر پیش‌بینی شده و مشاهده‌ای) به یک نزدیک‌تر شود، نشان‌دهنده همبستگی بیشتر مقادیر پیش‌بینی شده و مشاهده‌شده خواهد بود. به این ترتیب حداقل بودن مقادیر RMSE و MAE نیز نشان‌دهنده عملکرد مطلوب‌تر مدل خواهد بود. برای ضریب NS نیز هر قدر مقدار این ضریب به یک نزدیک‌تر شود به این معنا خواهد بود که تناسب کاملی بین داده‌های مشاهداتی و شبیه‌سازی شده وجود دارد. مقادیر بزرگ‌تر از ۰/۷۵ برای ضریب NS به معنای عملکرد خوب مدل و

جدول ۲. الگوهای ورودی استفاده‌شده در مقیاس زمانی روزانه رودخانه تلوار

ردیف	الگوی ورودی مدل در مقیاس روزانه
۱	$Q(t) = f(Q(t-1))$
۲	$Q(t) = f(Q(t-1), Q(t-2))$
۳	$Q(t) = f(Q(t-1), Q(t-2), Q(t-3))$
۴	$Q(t) = f(Q(t-1), Q(t-2), Q(t-3), Q(t-4))$
۵	$Q(t) = f(Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5))$

درختی جریان رودخانه تلوار در مقیاس زمانی سه‌روزه نیز در شکل ۲ نشان داده شده است. نتایج آماری به‌دست‌آمده مطابق جدول ۳ نشان داد افزایش حافظه تأثیری در افزایش عملکرد نداشته است.

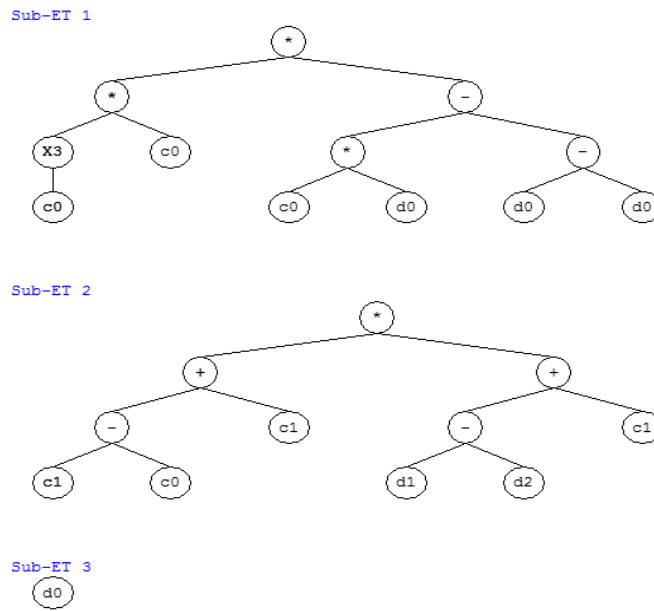
$$Q_t = -0.09116Q_{t-1} - 0.307007Q_{t-2} + 0.307007Q_{t-3} + 0.142203 \quad (14)$$

که در این معادله  $Q_t$  جریان روز  $t$  ام و  $Q_{t-1}$  نیز مقادیر جریان در یک روز قبل،  $Q_{t-2}$  جریان در دو روز قبل و  $Q_{t-3}$  جریان در سه روز قبل هستند.

نتایج معیارهای آماری روش برنامه‌ریزی بیان ژن در جدول ۳ آورده شده است. با توجه به این نتایج، می‌توان نتیجه گرفت که الگوی ۳ شامل دبی‌های یک، دو و سه روز قبل با کمترین خطا و مقدار ضریب تبیین  $R^2 = 0.916$  دقت زیادی دارد و معادله ریاضی و پارامترهای ژنی و مقادیر ثابت آن برای پیش‌بینی در دبی روز مد نظر به‌صورت پارامترهای ژنی  $d_1 = Q(i-1)$ ،  $d_2 = Q(i-2)$  و  $d_3 = Q(i-3)$  و مقادیر ثابت  $G_1C_1 = -0.463196$ ،  $G_2C_2 = -0.61938$  و  $G_3C_3 = -0.413696$  هستند (رابطه ۱۸). ساختار

جدول ۳. تحلیل‌های آماری نتایج برنامه‌ریزی بیان ژن برای الگوهای مختلف ورودی جریان روزانه ( $m^3s^{-1}$ ) رودخانه تلوار

الگو	آموزش				صحت‌سنجی			
	$R^2$	RMSE	MAE	NS	$R^2$	RMSE	MAE	NS
۱	۰/۹۰۵	۱/۱۹	۰/۲۳	۰/۴۲۶	۰/۹۲۱	۰/۲۶	۰/۱۰۶	۰/۸۱۳
۲	۰/۸۷۷	۱/۳۴	۰/۴۱	۰/۹۲۸	۰/۹۳۸	۰/۳۴	۰/۲۳	۰/۵۵۹
۳	۰/۸۸۳	۱/۳۱	۰/۲۸۳	۰/۸۱۴	۰/۹۱۶	۰/۲۸	۰/۱۵	۰/۸۴۱
۴	۰/۸۸۳	۱/۳۲	۰/۲۵	۰/۶۴۵	۰/۹۱۶	۰/۲۷	۰/۱۰۵	۰/۲۱۳
۵	۰/۸۹۴	۱/۲۴	۰/۳۳	۰/۵۹۶	۰/۹۱۱	۰/۳۴	۰/۱۶	۰/۱۷۹



شکل ۲. ساختار درختی برنامه‌ریزی بیان ژن در بهترین الگوی جریان روزانه رودخانه تلوار

جدول ۴. تحلیل‌های آماری نتایج مدل LS-SVM برای الگوهای مختلف ورودی جریان روزانه ( $m^3s^{-1}$ ) رودخانه تلوار

الگو	آموزش				صحت‌سنجی			
	R <sup>2</sup>	RMSE	MAE	NS	R <sup>2</sup>	RMSE	MAE	NS
۱	۰/۹۲۳	۰/۶۲۹	۰/۲۶۹	۰/۸۱۴	۰/۹۲۱	۰/۲۰۲	۰/۰۲۹	۰/۸۳۵
۲	۰/۸۹۲	۰/۴۴۸	۰/۱۵۱	۰/۹۰۵	۰/۹۲۳	۰/۲۰۲	۰/۱۱۶	۰/۸۳۶
۳	۰/۹۵۰	۰/۸۱۷	۰/۳۲۲	۰/۶۸۵	۰/۸۶۹	۰/۲۵۵	۰/۰۹۱	۰/۷۳۷
۴	۰/۹۰۳	۱/۲۴	۰/۲۶۵	۰/۷۹۹	۰/۹۲۰	۰/۲۰۱	۰/۰۶۷	۰/۸۳۵
۵	۰/۸۹۶	۱/۲۴۶	۰/۲۶۹	۰/۷۹۰	۰/۸۹۶	۰/۲۰۲	۰/۰۶۸	۰/۸۳۴

تنظیم پارامترهای شبکه استفاده می‌شود، به‌کار گرفته شد [۲۰]. در این روش با توجه به جدول ۵ مدل ۲ به‌صورت شبکه‌ای با ۳ نرون در لایه پنهان اول با بیشترین ضریب همبستگی  $R^2=0/90$  و کمترین مقدار ضرایب خطا  $RMSE=0/673$  در مرحله صحت‌سنجی، به‌عنوان بهترین مدل در فرایند شبیه‌سازی انتخاب شد.

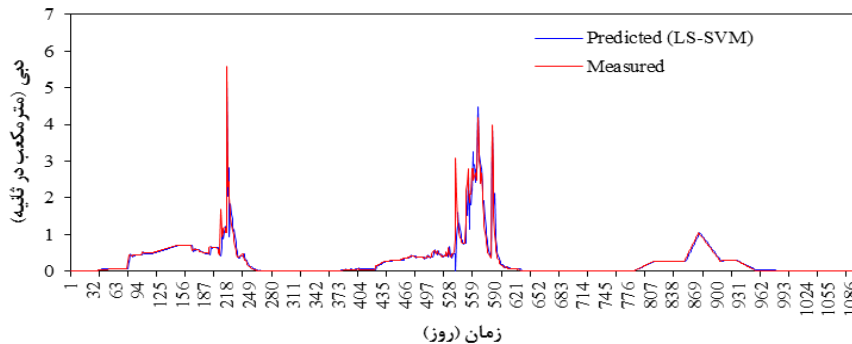
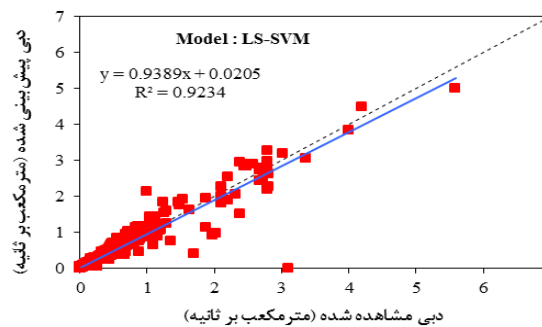
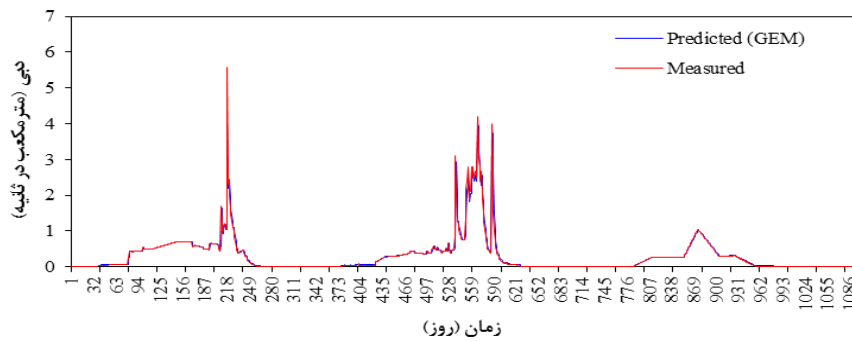
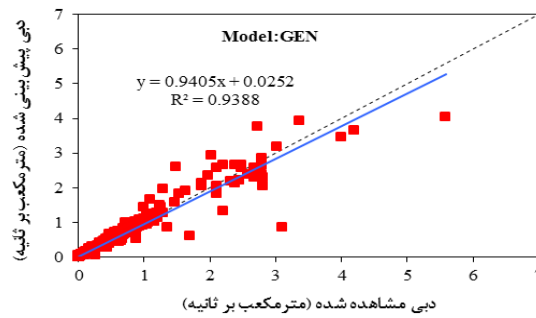
مقادیر مشاهده‌ای و پیش‌بینی‌شده دبی روزانه جریان در رودخانه تلوار به‌صورت نمودارهای صحت‌سنجی براساس مدل‌های سه‌گانه در شکل ۳ ارائه شده است. با توجه به نمودارها در بهترین مدل انتخابی مشاهده می‌شود، دقت در برآورد دبی‌های کمتر بیش از دبی‌های بزرگ‌تر است. در پژوهش احمدی و همکارانش [۲۱]، که به مقایسه روش‌های الگوریتم ژنتیک و ماشین بردار پشتیبان روی رودخانه باراندوزچای پرداختند، نتایج مشابهی گزارش شده است.

برای روش حداقل مربعات ماشین بردار پشتیبان، داده‌ها برای همسان‌سازی، استانداردسازی شده و سپس جریان رودخانه تلوار در مدل LS-SVM با تابع RBF همراه با مقدار بهینه مشخصه  $\gamma$  تخمین زده شد. طبق جدول ۴ نتایج نشان داد مدل با تأخیر دو روز با مقدار ضریب تبیین  $R^2=0/923$  بیشترین دقت برای شبیه‌سازی دارد.

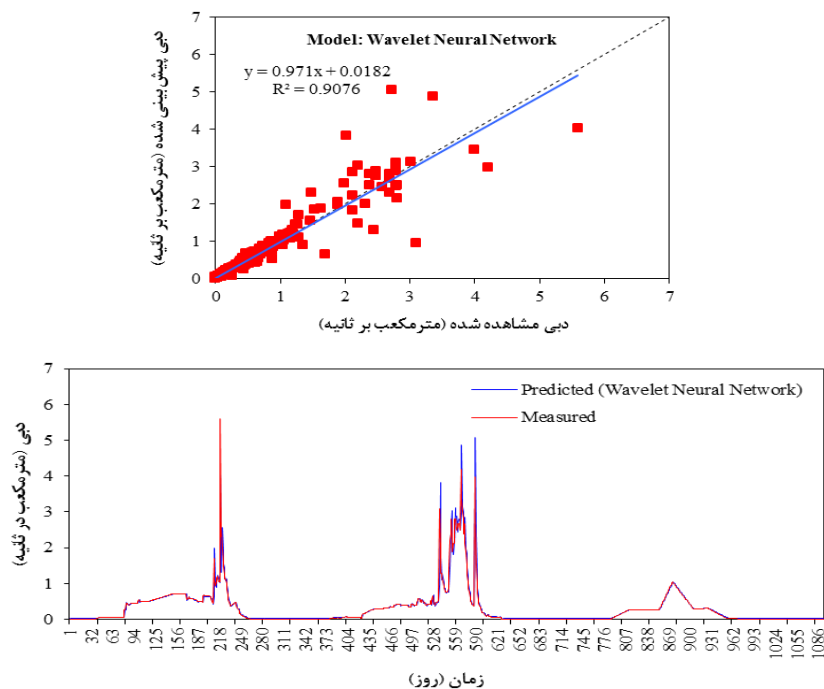
به‌منظور تخمین میزان جریان روزانه رودخانه تلوار از مدل شبکه عصبی موجک از نوع شبکه پرسپترون چندلایه با تعداد نورون‌های متفاوت نیز استفاده شد. در مدل عصبی موجکی ابتدا موجک مناسب (سیملت) انتخاب شد. سپس، با اعمال تبدیل روی داده‌ها، ضرایب تقریب و جزئیات آنها استخراج شد و داده‌ها به‌وسیله تابع موجک کلاه مکزیکی به‌عنوان توابع فعال‌سازی تبدیل شدند که مشتق دوم تابع گوسی است. برای آموزش شبکه، الگوریتم گرادیان نزولی که در یادگیری شبکه‌های عصبی و حداقل‌سازی مقدار خطا و

جدول ۵. تحلیل‌های آماری نتایج شبکه عصبی موجک برای الگوهای مختلف ورودی جریان روزانه ( $m^3s^{-1}$ ) رودخانه تلوار

الگو-ساختار	آموزش				صحت‌سنجی			
	R <sup>2</sup>	RMSE	MAE	NS	R <sup>2</sup>	RMSE	MAE	NS
۱-۶-۱	۰/۸۹۰	۰/۳۲۴	۰/۲۷۴	۰/۸۱۴	۰/۸۹۴	۰/۲۵۳	۰/۰۶۱۶	۰/۷۸۰
۲-۹-۱	۰/۸۸۶	۰/۵۸۷	۰/۲۷۷	۰/۸۱۸	۰/۹۰۷	۰/۶۳۷	۰/۰۵۵	۰/۸۰۰
۳-۸-۱	۰/۸۶۷	۰/۲۶۳	۰/۲۳۶	۰/۸۰۱	۰/۸۸۵	۰/۲۵۸	۰/۰۹۵	۰/۷۷۲
۴-۵-۱	۰/۸۷۳	۰/۴۱۷	۰/۲۴۸	۰/۸۰۵	۰/۸۹۵	۰/۵۲۴	۰/۳۴۱	۰/۳۴۷
۵-۴-۱	۰/۷۹۵	۰/۳۸۲	۰/۲۷۶	۰/۷۹۶	۰/۸۹۵	۰/۴۵۷	۰/۴۸۴	۰/۴۸۰







شکل ۳. نمودارهای ۱:۱ و ترسیم مقادیر مشاهداتی و مقادیر پیش‌بینی شده به‌دست آمده از سه مدل در مرحله صحت‌سنجی

قسمت با نتایج تحقیق احمدی و همکارانش [۲۱]، که روش‌های برنامه‌ریزی ژنتیک و ماشین بردار پشتیبان در پیش‌بینی جریان روزانه رودخانه باراندوزچای را مقایسه کردند، همخوانی دارد.

بهترین نتایج به‌دست آمده از الگوهای مختلف در روش برنامه‌ریزی بیان ژن در جدول ۳، مدل LS-SVM در جدول ۴ و روش شبکه عصبی موجک در جدول ۵ نشان داده شده است. چنانچه در جدول ۳ مشاهده می‌شود پیش‌بینی جریان روزانه در روش برنامه‌ریزی بیان ژن در مرحله صحت‌سنجی در الگوی دوم با تأخیر دو روز قبل دارای ضریب همبستگی  $0.93$  و میانگین مربعات خطا  $0.33$  (مترمکعب بر ثانیه) بهترین نتیجه ارائه داده است و در جدول‌های ۴ و ۵ نیز همین الگو در هر دو روش ماشین بردار پشتیبان و شبکه عصبی موجک به ترتیب با مقادیر ضریب همبستگی  $0.92$  و  $0.90$  بهترین نتایج را نشان داد. همچنین، با توجه به ضریب نش-ساتکلیف (NS) الگوی ۲ در مدل LS-SVM در مقایسه با دو مدل دیگر با مقدار  $0.84$  عملکرد بهتری را نشان داد. به‌طور کلی، از نتایج سه مدل پیداست که دقت پیش‌بینی جریان در دو مرحله آموزش و صحت‌سنجی در ایستگاه مطالعه شده در روش ماشین بردار پشتیبان نسبت به دو روش دیگر بیشتر است. سیدیان و همکارانش [۲۳] نیز، که با

### بحث و نتیجه‌گیری

دقت تخمین در مدل حداقل مربعات ماشین بردار پشتیبان نشان داد کارایی این مدل در جریان‌های با دبی بالاتر، بیشتر است. به علاوه، همان‌طور که در شکل ۳ نیز مشخص است تطابق مقادیر محاسباتی دبی جریان مربوط به مرحله صحت‌سنجی مدل شبکه عصبی موجک با مقادیر مشاهداتی وجود دارد. به‌گونه‌ای که با توجه به نمودار، قابلیت زیاد این مدل در تخمین بیشتر مقادیر مشهود است. در تبیین نتایج مدل شبکه عصبی موجک می‌توان بیان کرد که تبدیل موجک با جداکردن سیگنال به فرکانس‌های بالا و پایین ویژگی‌های چند مقیاس سیگنال را در اختیار داشته و دقت مدل را تا حد درخور توجهی، بالا می‌برد. سیگنال‌های بالاگذر و پایین‌گذر به‌وجود آمده از تجزیه موجک، برازش بسیار خوبی با معادلات مجموع سینوسی دارند که هر چه تعداد مراتب معادلات بیشتر می‌شود، دقت کار افزایش می‌یابد بسامدهای مراحل پایین‌تر تجزیه، نویز بیشتری دارند، ولی با افزایش سطح تجزیه از میزان نویز کاسته شده و سیگنال نرم‌تر می‌شود [۲۲].

از مقایسه نمودارهای شکل ۳ می‌توان نتیجه گرفت که برنامه‌ریزی ژنتیک دقت بیشتری نسبت به دو مدل دیگر و در تخمین دبی پیک نیز عملکرد بهتری داشته است. نتایج این

می‌شود، دقت کار افزایش می‌یابد. بسامدهای مراحل پایین‌تر تجزیه، خطای بیشتری دارند، ولی با افزایش سطح تجزیه افزایش نوروها از میزان نویز کاسته شده و سیگنال نرم‌تر می‌شود. با توجه به عملکرد مدل‌ها در پیش‌بینی منطقه مطالعه شده می‌توان به بررسی تأثیر پارامترهای هواشناسی نیز بر میزان جریان در منطقه پرداخت. بنابراین، لازم است تکنیک‌هایی مانند برنامه‌ریزی بیان ژن، شبکه عصبی، شبکه عصبی موجک، LS-SVM و... بیش از پیش در حوضه‌های منابع آب و توسعه پایدار پروژه‌های مربوط به آنها واکاوی شوند.

#### منابع

- [1]. Kisi O. Least squares support vector machine for modeling daily reference evapotranspiration. *Irrigation Science*; 2012.
- [2]. Ghorbani MA, Kisi O, Aalinezhad MA. Probe into the chaotic nature of daily stream flow time series by correlation dimension and largest Lyapunov methods *Applied Mathematical Modelling*. 2010; 34: 4050–4057.
- [3]. Whigham PA, Crpper PF. Modelling Rainfall-Runoff Using Genetic Programming. 2001. 33:707-721.
- [4]. Nikbakhtshabazi A. Application of support vector machine in river flow forecasting, *Iranian Hydraulic Conference*; 2009 [Persian]
- [5]. Yu PS, Chen ST, Chang IF. Support vector regression for real-time flood stage forecasting *Hydrology*. 2006; 328: 704-716.
- [6]. Okkan U. Wavelet neural network model for reservoir inflow prediction. *Journal of ScientiaIranica*; 2012, December 1445–1455
- [7]. Sayagavi V.G, Charhate Sh, Magar R. 2016. Estimation of Discharge Using LS-SVM and Model. *Journal of Water Resources and Ocean Science*. 2016; 5(6): 78-86.
- [8]. He Zh, Wen X, Liu H, Du J. 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. 2014; 509: 379-386.
- [9]. Shenify M, SeyedDanesh A, Gocić M, Surya R, Ainuddin T, Wahab WA, Gani A, Shamshirband Sh. Precipitation Estimation Using Support Vector Machine with Discrete Wavelet Transform, *Journal of water resources management*. 2016; 30(2):641-652.

استفاده از روش‌های سری زمانی و ماشین بردار پشتیبان جریان رودخانه حوضه آبخیز گرگان‌رود را پیش‌بینی کردند، مشخص کردند که روش ماشین بردار پشتیبان در سری‌های ماهانه و هفتگی نسبت به روش سری زمانی دقت بیشتری دارد. همچنین، تربیداتی و همکارانش [۲۴]، که به برآورد بارندگی ماهانه در هند با روش SVM پرداختند، پیشنهاد کردند که از SVM به‌عنوان گزینه‌ای برتر نسبت به سایر روش‌های معمول برای پیش‌بینی بارش می‌توان استفاده کرد. سایاگوی و همکارانش [۷] در بررسی مدل‌های LS-SVM و MT برای پیش‌بینی رواناب در حوضه کریشینا در هند دریافتند مدل LS-SVM مناسب‌تر است. از بررسی تحقیقات مختلف مشخص شد که روش ماشین بردار برای پیش‌بینی پارامترهای هیدرولوژی عملکرد مناسب‌تری نسبت به سایر مدل‌های پیش‌بینی دارد.

#### نتیجه‌گیری

تعیین جریان رودخانه‌ها یکی از مؤلفه‌های مهم و تأثیرگذار در مدیریت منابع آب حوضه آبریز است. هدف اصلی این مطالعه یافتن بهترین مدل هوشمند از مقایسه مدل‌های برنامه‌ریزی بیان ژن، حداقل مربعات ماشین بردار پشتیبان و شبکه عصبی موجک برای منطقه مطالعه شده بود. هر سه مدل با همه الگوهای طی فرایند شبیه‌سازی به‌خوبی عمل کردند. با توجه به نتایج سه مدل بهینه‌ترین الگو از بین پنج الگوی در نظر گرفته شده الگوی ۲ شبیه‌سازی با تأخیری دو روز قبل کمترین خطا و دقت زیاد داشته است و همچنین از نتایج سه مدل چنین مشاهده شد که دقت مدل‌سازی رفته‌رفته تا سه تأخیری زمانی دبی رو به بهبود است و بعد از آن کاهش می‌یابد. عملکرد مدل‌های LS-SVM در مقایسه با مدل‌های برنامه‌ریزی بیان ژن و عصبی موجک در دو مرحله آموزش و صحت‌سنجی در همه الگوها خوبی را داشتند، چراکه مدل LS-SVM ضریب همبستگی بیشتر و کمترین خطا به‌ویژه مجذور میانگین مربعات خطا را داشت. اگرچه مدل LS-SVM در برخی موارد میزان جریان روزانه را بیش از حد معمول پیش‌بینی می‌کند، نسبت به دو مدل دیگر با توجه به شاخص‌های عملکرد می‌توان آن را مناسب دانست زیرا بالاترین و پایین‌ترین پیک را به صورت منطقی لحاظ می‌کند. در روش مدل‌سازی شبکه عصبی موجک نشان داده شد که هر چه تعداد مراتب این معادلات مجموع سینوسی بیشتر

- [10]. Sarangi A, Bhattacharya A.K. Comparison of Artificial Neural Network and Regression Models for Sediment Loss Prediction from Banha Watershed in India. *Agricultural water management*. 2005; 28(4): 373-385.
- [11]. Solaimani K. Rainfall-Runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed). *American-Eurasian Journal of Agriculture and Environment, Science*. 2009; 5(6): 856-865.
- [12]. Borelli A, De Falco I, Della C A, Nicodemi M, Trautteur G. Performance of Genetic Programming to Extract the Trend in Noisy Data Series. *Physica A*. 2006; 370: 104-108.
- [13]. Suykens J A. K, Gestel T V, Brabanter J D, Moor B D, Vandewalle J. *Least Squares Support Vector Machines*. Copyright by World Scientific Publishing Co. Pie. Ltd. 2002; 58: 72-75.
- [14]. Shiloh Shah R. Support vector machines for classification and regression, Master of Science, computer science, McGill university Montreal, Quebec, 2007.
- [15]. Seif A. Developing an Expert System for Predicting Daily Reference Evapotranspiration Using Support Vector Machines (SVM) in Comparison with ANFIS, ANN and Empirical Methods. M. S. in Irrigation and drainage Engineering, TarbiatModares University, Tehran, 2010 [Persian]
- [16]. Wang D, Ding J. Wavelet network model and its application to the prediction of hydrology. *Nat Science*. 2003; 1(1): 67-71.
- [17]. Mellit A, MassiPavan A, Benghanem, M. Least squares support vector machine for short-term prediction of meteorological time series. *Theory Apple Climatol*. 2013; 111: 297-307.
- [18]. Safavi A, Romagnoli J.A. Comments on the non-linear black box modelling in system identification: a unified overview. *IFAC Journal of automatic*. 1997; 33(6): 1197-1198.
- [19]. Nash JE, Sutcliffe JV. River flow forecasting through conceptual models. Part 1 – a discussion of principles *J. Hydrology*. 1970; 10: 282-310.
- [20]. Menhaj MB. *Basics neural networks*. 1nd ed. Tehran : Amirkabir University of Technology; 2006 [ Persian]
- [21]. Ahmadi F, Radmanesh F, Mirabasi R. Comparison of genetic algorithm and support vector machine to predict daily river flow (Case Study: BarandozChay River). *Journal of Soil and Water (Food Science and Industry)*. 2014; 28(6): 1171-1116 [Persian]
- [22]. Toufani P, Mosaedi A, FakhriFard A. Prediction of Precipitation Applying Wavelet Network Model (Case study: Zarringol station, Golestan province, Iran), *Journal of Water and Soil* 2011; 25(5): 1217-1226 [ Persian]
- [23]. Seyedian SM, Solimani M, Kashani M. Predicting stream flow using data-driven model and time series. *Ecohydrology*. 2014; 1(3): 167-179 [Persian]
- [24]. Tripathi, Sh, Srinivas, VV, Nanjundiah RS. Downscaling of precipitation for climate change scenarios: A support vector machine approach. *Journal of Hydrology*. 2006; 330: 62-640.