



## مدل سازی رواناب رودخانه صوفی چای با استفاده از ماشین بردار پشتیبان و شبکه عصبی مصنوعی

فاطمه آخونی پورحسینی<sup>۱</sup> و صابره دربندی<sup>۲</sup>

۱- دانشجوی کارشناسی ارشد مهندسی منابع آب، دانشگاه تبریز، (نویسنده مسوول: fateme.pourhosseini@yahoo.com)

۲- استادیار گروه مهندسی آب، دانشگاه تبریز

تاریخ پذیرش: ۹۵/۶/۱۴

تاریخ دریافت: ۹۵/۲/۲

### چکیده

شبیه سازی دقیق فرآیند رواناب می تواند نقش بسزایی در مدیریت منابع آب و مسائل مربوطه داشته باشد. پیچیدگی ذاتی این فرآیند استفاده از مدل های فیزیکی و عددی را مشکل می نماید. در سال های اخیر کاربرد مدل های هوشمند به عنوان ابزاری توانمند در علم هیدرولوژی افزایش یافته است. هدف این مطالعه کاربرد آزمون گاما برای انتخاب ترکیب بهینه متغیرهای ورودی در مدل سازی رودخانه صوفی چای می باشد. مدل سازی جریان آب رودخانه با استفاده از تعداد نقاط بهینه متغیرهای منتخب با روش های شبکه عصبی مصنوعی و ماشین بردار پشتیبان می باشد. نتایج آزمون گاما نشان داد که رواناب رودخانه با شش تأخیر زمانی، نتایج بهتری به منظور پیش بینی ارائه می دهد. شبیه سازی رواناب با استفاده از دو مدل ماشین بردار پشتیبان و شبکه عصبی مصنوعی نشان داد که بهترین ساختار ورودی برای پیش بینی رواناب ماه بعد، تا شش تأخیر خواهد بود. از میان دو مدل با ساختار ورودی یکسان، مدل ماشین بردار پشتیبان کارایی نسبتاً بالایی نسبت به شبکه عصبی مصنوعی داشته است.

واژه های کلیدی: آزمون گاما، رواناب، رودخانه صوفی چای، شبکه عصبی مصنوعی، ماشین بردار پشتیبان

### مقدمه

آب یکی از عوامل مهم در توسعه پایدار هر منطقه می باشد. به نظر می رسد با توجه به شرایط اقلیمی، کمیت و کیفیت آب قابل دسترس برای توسعه اقتصادی و زیست محیطی در اکثر نقاط ایران کافی نمی باشد (۱۷). در سال های گذشته وقوع سیلاب ها و خشکسالی ها همواره امنیت زندگی و اموال مردم ساکن در پایین دست هر رودخانه ای را تهدید و خسارات فراوانی را به بار آورده هست. بنابراین با مدل سازی و پیش بینی دقیق دبی رودخانه در فصول مختلف، می توان اولین گام را در مدیریت علمی منابع آب های سطحی منطقه برداشت (۱۷). در فرآیند تشکیل رواناب حوزه متغیرهای متعددی دخالت دارند. اثر متقابل این متغیرها باعث می شود که این فرآیند به یک پدیده غیرخطی و پیچیده تبدیل شود. از طرفی برای اینکه بتوان یک مدل مناسب برای پیش بینی میزان دبی جریان آب رودخانه ارائه نمود لازم هست تا حد امکان تمامی عوامل مؤثر اندازه گیری در مدل دخالت داده شود. در بیشتر مواقع اندازه گیری غلط، ثبت اشتباه داده، کمبود دقت و منابع مالی جهت احداث ایستگاه های هیدرومتری و نهایتاً کافی نبودن طول سری داده ها باعث می شود که محقق قادر به تهیه مدل با دقت مناسب نباشد (۱۷). تاکنون مطالعات زیادی برای مدل سازی و پیش بینی جریان رودخانه با روش های مختلف صورت گرفته است. با این حال هنوز مسائل حل نشده فراوانی در این زمینه وجود دارد. انتخاب بهترین ترکیب برای دبی جریان یکی از این مسائل می باشد (۱۷). یکی از مراحل مهم و پیچیده برای مدل سازی غیرخطی، پیش پردازش داده های ورودی به منظور انتخاب ترکیبی مناسب از آن در مدل ها می باشد که سبب کاهش مراحل سعی و خطا و شناخت مهم ترین متغیرهای بر پدیده مورد نظر در مدل سازی می باشد. رمضان و همکاران

(۱۳) برای مدل سازی تابش روزانه ی خورشید در حوزه آبریز برو در جنوب غربی انگلستان از چهار پارامتر هواشناسی شامل تابش های برون زمینی خورشیدی، میانگین دمای هوا، میانگین سرعت روزانه باد و عمق بارش روزانه استفاده کردند. آن ها برای انتخاب بهترین ورودی به مدل از آزمون گاما استفاده کردند. نتایج تحقیق آن ها نشان داد که بهترین ترکیب پارامترهای ورودی به مدل، ترکیبی است که در آن هر چهار پارامتر ذکر شده در ورودی مدل وجود داشته باشد. لعدانی و همکاران (۱۱) جهت پیش بینی حجم رسوبات معلق رودخانه دویرج، واقع در استان ایلام از روش آزمون گاما جهت تعیین متغیرهای ورودی به مدل ماشین بردار پشتیبان استفاده نمودند و نتیجه گرفتند که عملکرد مدل GT-SVM در مرحله آزمون بهتر از عملکرد مدل SVM می باشد. رمضان و همکاران (۱۴) با استفاده از داده های روزانه بارش و رواناب و روش هیبرید موجک رواناب حوزه آبخیز برو را در انگلستان مدل سازی نمودند. آن ها از داده های روزانه بارش و رواناب بازمان تأخیرهای مختلف، ترکیب های مختلفی را به عنوان متغیرهای ورودی مورد بررسی قرار دادند و با استفاده از آزمون گاما نتیجه گرفتند سه تأخیر در رواناب و یک تأخیر در بارش بهترین ترکیب ورودی برای مدل سازی رواناب رودخانه هست. دیباک و همکاران (۶) قابلیت های مدل SVM را در خصوص پیش بینی های هیدرولوژیکی نشان دادند. آن ها با استفاده از این روش در زمینه ی دسته بندی داده های سنجش اذدور و مدل سازی بارش- رواناب و مقایسه آن با روش های شبکه عصبی مصنوعی به نتایج خوبی برای پیش بینی دست یافتند. باتسیس و همکاران (۱) بارش- رواناب روزانه در کالیفرنیا را شبیه سازی کردند مدل ماشین بردار پشتیبان با سه نوع تابع کرنل با مدل شبکه عصبی مصنوعی مورد مقایسه قرار گرفت، در نتیجه این تحقیق ماشین بردار پشتیبان شبیه سازی

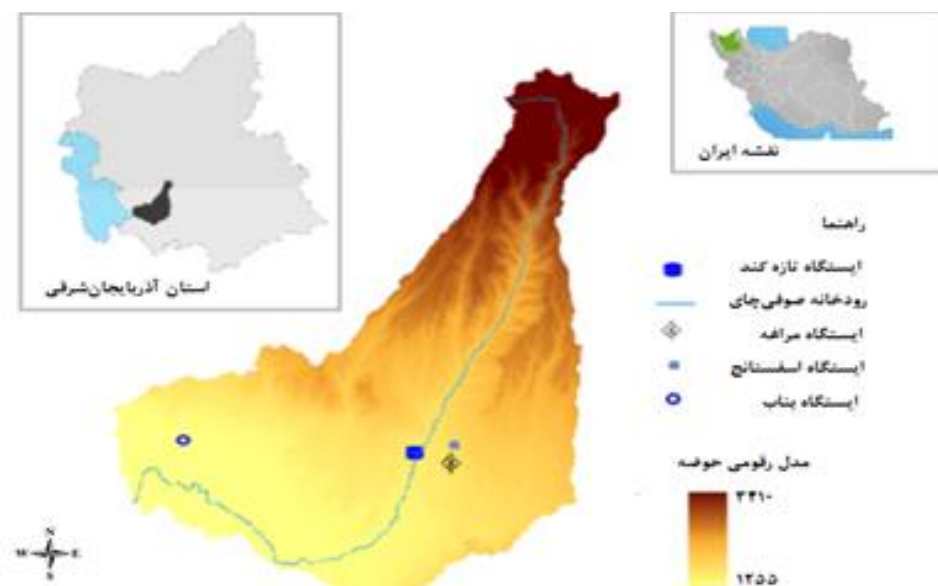
اصلی، بهبود عملکرد مدل SVM را به همراه خواهد داشت. در مدل PCA SVM به جای ۱۸ ورودی تنها از ۵ مؤلفه به عنوان ورودی به مدل استفاده شد که این واقعیت مزیت اصلی این مدل به شمار می رود. در مطالعه حاضر برای یافتن بهترین ترکیب از متغیرهای ورودی روش آزمون گاما انتخاب شد و با کمک ماشین بردار پشتیبان و شبکه عصبی مصنوعی به پیش بینی رواناب رودخانه صوفی چای پرداخته شد.

### مواد و روش ها

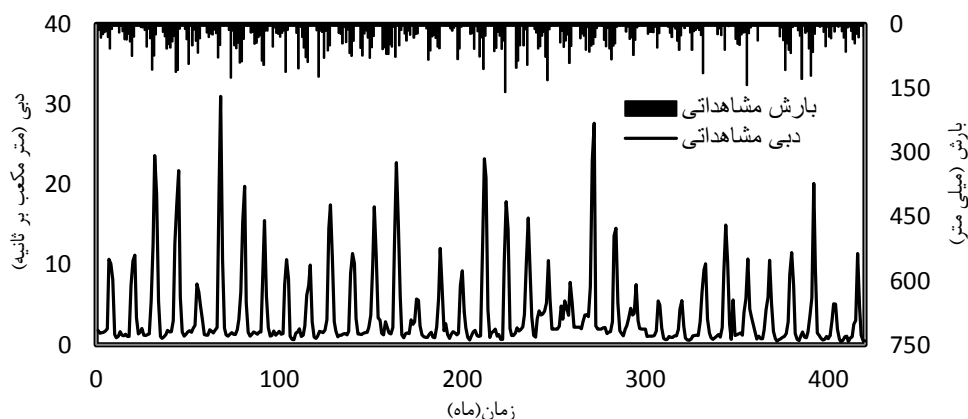
#### منطقه مورد مطالعه و داده های مورد استفاده

حوزه صوفی چای از دامنه جنوبی کوهستان سهند سرچشمه گرفته و از شمال به حوزه آبریز ليقوان و کند چای، از شرق به حوزه آبریز مردق چای، از غرب به حوزه آبریز قلعه چای و از جنوب بعد از عبور از محدوده سد علویان به دریاچه ارومیه منتهی می شود. رودخانه صوفی چای با طول ۲۰ کیلومتر در داخل حوزه آبریز صوفی چای قرار دارد. یکی از نزدیک ترین ایستگاه هیدرومتری به این حوزه، ایستگاه هیدرومتری تازه کند می باشد. ایستگاه تازه کند در مختصات جغرافیایی در طول  $36^{\circ} 16'$  و در عرض  $37^{\circ} 29'$  قرار دارد. مساحت حوزه آبریز تا محل ایستگاه هیدرومتری تازه کند سد علویان در حدود ۲۵۰ کیلومتر مربع و شیب متوسط وزنی آن  $7/3$  درصد و زمان تمرکز حوزه به روش کریپچ  $8/75$  ساعت می باشد. بر اساس آمار و اطلاعات ایستگاه ها، متوسط بارندگی سالانه ۳۳۰ میلی متر می باشد که حدود ۳۰ درصد بارندگی در زمستان، بیش از ۴۰ درصد آن در بهار، ۲ درصد در تابستان و حدود ۲۵ درصد در پاییز می باشد. داده های مورد استفاده در این تحقیق شامل مقادیر ماهانه رواناب در طول ۴۰ سال (مهر ۱۳۵۲ - شهریور ۱۳۹۱) برای ایستگاه هیدرومتری تازه کند می باشد. شکل ۱ موقعیت حوزه آبریز صوفی چای و ایستگاه مذکور را نشان می دهد. نمودار تغییرات داده های رواناب و بارش ماهانه رودخانه صوفی چای در ایستگاه تازه کند در شکل ۲ و مشخصات آماری ایستگاه باران سنجی نیز در جدول ۱ ارائه شده است.

دقیق تری از بارش - رواناب داشت. یون و جون (۲۱) دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان را به منظور پیش بینی تغییرات سطح آب زیرزمینی بکار گرفتند، به این نتیجه رسیدند که مدل ماشین بردار پشتیبان دارای عملکردی بهتر از شبکه عصبی مصنوعی است. کیسی و مسات (۹) مطالعه ای به منظور پیش بینی جریان ماهانه دو ایستگاه گردلی و ایساکوی<sup>۲</sup> در منطقه شرقی دریای سیاه ترکیه با استفاده از مدل موجک - ماشین بردار پشتیبان انجام دادند. نتایج این تحقیق افزایش دقت مدل هیبریدی موجک - ماشین بردار پشتیبان را نسبت به مدل ماشین بردار پشتیبان به همراه داشت. طاهری و همکاران (۲۰) دو روش یادگیری فعال و ماشین بردار پشتیبان را جهت شبیه سازی جریانات روزانه رودخانه کارون بکار برده و به نتایج مشابهی برای این دو روش رسیدند و قابلیت بالای ماشین بردار پشتیبان در شبیه سازی رواناب را مطرح نمودند. لعدانی و همکاران (۱۰) سه مدل ماشین بردار پشتیبان و سیستم استنتاج فازی - عصبی تطبیقی (ANFIS) و مدل هیدرولوژیکی مفهومی Mike11/NA را در شبیه سازی جریان روزانه حوزه اسکندری به کار گرفتند. نتایج نشان داد که هر سه مدل دارای عملکرد مناسب در منطقه مورد مطالعه می باشند. در نهایت مدل ماشین بردار پشتیبان با بالاترین ضریب همبستگی و کمترین میانگین مربعات خطا عملکرد بهتری نسبت به دو مدل دیگر داشت. فرج زاده و همکاران (۷) با استفاده از مدل های سری زمانی و شبکه عصبی مصنوعی به پیش بینی جریان ماهانه و بارش حوزه دریاچه ارومیه پرداختند. نتایج حاصل از پیش بینی جریان حاکی از دقت خوب هر دو مدل در تخمین جریان ماهانه دارد. البته طبق این تحقیق تفاوت چندانی بین دو مدل مقایسه در تخمین جریان وجود ندارد. نوری و همکاران (۱۲) جریان ماهانه رودخانه صوفی چای را از طریق آنالیز مؤلفه اصلی (PCA) بر مبنای ماشین بردار پشتیبان مدل سازی نموده و عملکرد مدل های SVM و PCA-SVM را مورد ارزیابی قرار دادند. نتایج این تحقیق نشان داد که پیش پردازش متغیر ورودی به ماشین بردار پشتیبان با استفاده از آنالیز مؤلفه



شکل ۱- مشخصات جغرافیایی ایستگاه‌های هیدرومتری در حوزه صوفی چای (نوری، ۱۳۸۹)  
Figure 1. The geographical location and hydrometric stations in sofichy basin



شکل ۲- مقادیر مشاهداتی بارش و رواناب ماهانه ایستگاه تازه کند در طول آماری  
Figure 2. The observed values of monthly precipitation and runoff Tazekand station during statistical period

جدول ۱- مشخصات آماری رواناب ماهانه ایستگاه هیدرومتری تازه کند

| پارامتر | کل    | آموزش<br>(مترمکعب بر ثانیه) | صحت سنجی<br>(مترمکعب بر ثانیه) |
|---------|-------|-----------------------------|--------------------------------|
| میانگین | ۳/۷۸  | ۴/۰۸                        | ۳/۰۷                           |
| حداقل   | ۰/۳۵  | ۰/۵۵                        | ۰/۳۵                           |
| حداکثر  | ۳۰/۹۸ | ۳۰/۹۸                       | ۲۰/۱۳                          |

سری مربوط با تأخیر یک‌روزه  $Q_{t-1}$ ، دوازده‌روزه  $Q_{t-2}$ ، سه‌روزه  $Q_{t-3}$ ، چهار‌روزه  $Q_{t-4}$ ، پنج‌روزه  $Q_{t-5}$  و شش‌روزه  $Q_{t-6}$  به‌عنوان متغیر ورودی و  $Q_t$  به‌عنوان متغیر خروجی به مدل معرفی گردید.

به‌منظور تخمین رواناب ماه‌های آینده، از اطلاعات رواناب ماه‌های قبل استفاده گردید. با استفاده از ماتریس همبستگی سری زمانی با تأخیر زمانی شش ماه قبل، به‌عنوان ورودی به مدل‌ها معرفی شد (جدول ۲). مقدار رواناب ماه بعد به‌عنوان خروجی مدل‌ها تعیین گردید. همچنین از شش متغیر شامل

جدول ۲- همبستگی متقابل بین پارامترهای ورودی و خروجی

Table 2. Cross-correlation coefficient between input and output parameters

| $(Q \text{ m}^3/\text{s})_{t-6}$ | $(\text{m}^3/\text{s})Q_{t-7}$ | $(\text{m}^3/\text{s})Q_{t-4}$ | $(\text{m}^3/\text{s})Q_{t-3}$ | $(\text{m}^3/\text{s})Q_{t-2}$ | $(\text{m}^3/\text{s})Q_{t-1}$ | $(\text{m}^3/\text{s})Q_t$ |
|----------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|----------------------------|
| -۰/۲۵                            | -۰/۲۶                          | -۰/۲۷                          | -۰/۱۹                          | ۰/۱۱                           | ۰/۶۴                           |                            |

### آزمون گاما

آزمون گاما یک روش غیرمتغیری است که نتایج آن صرف نظر از روش های خاص برای مدل سازی به کار می رود. این روش را به عنوان ابزاری مناسب برای یافتن بهترین ترکیب از متغیرهای ورودی به مدل غیرخطی می توان در نظر گرفت. در واقع آزمون گاما تخمینی از حداقل میانگین مربعات خطای آن بخش داده ها است که با مدل خطی نمی توان آن را تعیین کرد. فرض می شود مجموعه داده های مشاهداتی به صورت زیر است (۱۵).

$$\{(x_i - y_i), 1 \leq i \leq M\}. \quad (۱)$$

که در آن  $x_i$  بردار مشاهداتی ورودی،  $y_i$  خروجی مشاهداتی و  $M$  تعداد کل مشاهدات می باشد. از بردار  $x_i$  (در فضای  $m$  بعدی) برای مدل سازی  $y_i$  (در فضای  $R$ ) استفاده می شود. بنابراین رابطه ی بین ورودی و خروجی به شرح زیر در نظر گرفته می شود:

$$y = f(x) + r \quad (۲)$$

که در آن  $f$  یک تابع هموار  $r$  متغیر تصادفی نشان دهنده خطا می باشد. فرض شد برای بردار ورودی  $x_i$ ، مجموعه  $k$ ،  $N[i]$  (به نام مجموعه نزدیک ترین همسایگی  $i$ ،  $1 \leq k \leq p$ ) به نام مجموعه نزدیک ترین همسایگی هست (۱۸).

جدول ۳- توابع کرنل

Table 3. Kernel functions

| رابطه ریاضی   | توابع کرنل              |
|---|-------------------------|
| $K(x_i, x_j) = x_i^T \cdot x_j$                                 | تابع کرنل خطی           |
| $K(x_i, x_j) = (x_i^T \cdot x_j + \tau)^d$                      | تابع کرنل چندجمله ای    |
| $K(x, x_i) = \exp\left(\frac{- x_i - x_j ^2}{2\sigma^2}\right)$ | تابع کرنل با پایه شعاعی |

### روش کار

به منظور تعیین ترکیب بهینه متغیرهای ورودی در مدل های ماشین بردار پشتیبان و شبکه عصبی مصنوعی ابتدا با استفاده از آزمون گاما، مقدار گاما برای حالت استفاده از تمام متغیرهای ورودی به دست آمد. سپس یکی از متغیرهای ورودی حذف و مقدار گاما برای ترکیب مفروض محاسبه شد. آنگاه متغیر مذکور دوباره به مدل وارد و متغیر دیگری حذف شد و مقدار گاما برای ترکیب جدید به دست آمد. این کار برای همه متغیرها تکرار و در جدولی ثبت شد.

جدول ۴ نتایج آزمون گاما را در ایستگاه تازه کند نشان می دهد. در این جدول ترکیب شماره یک، ترکیبی است که در آن همه متغیرهای ورودی حضور دارد. در ترکیبات دیگر همه متغیرها به جز یکی در مدل حاضر هستند. مثلاً در ترکیب

### مدل شبکه های عصبی مصنوعی (ANNs)

شبکه عصبی مصنوعی در علوم مختلف از جمله فیزیک، علوم مهندسی، پزشکی و ... گسترش و انواع مختلفی از شبکه ها و الگوریتم های آموزش توسعه یافتند. حدود ۹۰ درصد شبکه هایی که در مسائل مختلف هیدرولوژی کاربرد دارد شبکه های پیشرو هستند (۲). معماری معمول این شبکه متشکل از سه لایه ورودی، پنهان و خروجی است. تعداد نرون های موجود در لایه ورودی و خروجی، بستگی به نوع مسئله دارد ولی تعداد گره های لایه پنهان با سعی و خطا دست می آید. در شبکه های پیشرو، اتصالات میان گره هایی وجود دارد که در لایه های متفاوت قرار دارند. همچنین ورودی ها در لایه ورودی شبکه نمایش داده شده و عمل تحریک از ورودی به سمت خروجی شبکه شروع می شود (۵).

گامای آن‌ها از ترکیب شماره ۱ بیشتر می‌باشد. بنابراین متغیرهای غایب این ترکیبات شامل شش متغیر  $Q_{t-2}$ ،  $Q_{t-3}$ ،  $Q_{t-4}$ ،  $Q_{t-5}$ ،  $Q_{t-6}$  به‌عنوان متغیر مؤثر در مدل‌سازی رواناب تشخیص داده شدند. همچنین متغیر  $Q_{t-1}$  به‌عنوان مهم‌ترین متغیر در این ترکیب می‌باشد. وقتی این متغیر از ترکیب حذف می‌شود مقدار  $(\Gamma)$  نسبت به ترکیب شماره یک (یعنی زمانی که همه متغیرها حضور دارند) بیشترین مقدار را به خود اختصاص می‌داد.

شماره دو، همه متغیرها به‌جز متغیر  $Q_{t-6}$  حاضر می‌باشد. بعد از اینکه مقدار گاما برای تمام ترکیب‌های جدول ۴ به دست آمد مقدار آماره  $(\Gamma)$  هریک از ترکیبات شماره‌ی ۲ تا ۷ با مقدار آماره گامای مربوطه به ترکیب یک مقایسه شد. ترکیباتی که در آن‌ها مقدار گامای مدل با ترکیبات مفروض بیشتر از مقدار گامای نظیر ترکیب شماره ۱ بود، به‌عنوان مهم‌ترین متغیر در پیش‌بینی رواناب تشخیص داده شدند و از آن‌ها برای مدل‌سازی استفاده گردید (۱۲). به‌طوری‌که از جدول ۴ مشاهده می‌شود شش ترکیب وجود دارد که مقدار

جدول ۴- نتایج آزمون گاما در مدل‌سازی رواناب ایستگاه هیدرومتری مورد مطالعه

Table 4. Gamma test results in modeling runoff in Tazekand hydrometric station

| ترکیب | متغیر غایب در مدل | گاما  | شیب A | خطای استاندارد | نسبت V | تعداد داده‌ها |
|-------|-------------------|-------|-------|----------------|--------|---------------|
| ۱     | -                 | ۰/۰۶۵ | ۰/۰۶۳ | ۰/۰۰۷          | ۰/۲۶۲  | ۴۲۶           |
| ۲     | $Q_{t-6}$         | ۰/۰۰۷ | ۰/۰۶۳ | ۰/۰۰۹          | ۰/۲۸۲  | ۴۲۶           |
| ۳     | $Q_{t-5}$         | ۰/۰۶۹ | ۰/۰۵۲ | ۰/۰۰۵          | ۰/۲۷۷  | ۴۲۶           |
| ۴     | $Q_{t-4}$         | ۰/۰۶۸ | ۰/۰۵۹ | ۰/۰۰۶          | ۰/۲۷۲  | ۴۲۶           |
| ۵     | $Q_{t-3}$         | ۰/۰۶۷ | ۰/۱۳۵ | ۰/۰۰۸          | ۰/۲۵۴  | ۴۲۶           |
| ۶     | $Q_{t-2}$         | ۰/۰۷۰ | ۰/۱۰۳ | ۰/۰۰۸          | ۰/۲۸۰  | ۴۲۶           |
| ۷     | $Q_{t-1}$         | ۰/۱۳۴ | ۰/۰۴۲ | ۰/۰۱۴          | ۰/۵۳۶  | ۴۲۶           |

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad (7)$$

$$E = 1 - \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (8)$$

که در روابط ۶ الی ۸،  $x_i$  مقدار دبی مشاهداتی،  $y_i$  مقدار دبی پیش‌بینی شده و  $\bar{x}$  و  $\bar{y}$  به ترتیب نشانگر مقادیر میانگین داده‌های متناظر می‌باشند. مدلی بهترین نتیجه را خواهد داشت که برای این سه معیار به ترتیب مقادیری نزدیک به یک، صفر و یک را به همراه داشته باشد.

### نتایج و بحث

به‌منظور مدل‌سازی داده‌های رواناب با استفاده از دو مدل SVM و ANN بخشی از داده‌ها جهت آموزش و بخشی دیگر جهت صحت سنجی مورد استفاده قرار گرفت. در جداول ۵ و ۶ مقادیر شاخصه‌ای آماری محاسبه‌شده مربوط به نتایج بخش آموزش و صحت سنجی در هر دو مدل SVM و ANN ارائه شده است.

جهت استفاده از داده‌های رواناب ماهانه ایستگاه تازه کند رودخانه صوفی چای در مدل‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی ابتدا داده‌ها به‌صورت تصادفی به دو قسمت تقسیم‌بندی شدند. ۷۰ درصد برای آموزش مدل و ۳۰ درصد بقیه برای بخش صحت سنجی استفاده گردید (۱۶). به‌منظور افزایش کارایی مدل، همه داده‌ها ابتدا به شکل نرمال بین دو عدد صفر و ۱ استاندارد شدند سپس آن‌ها به مقادیر اولیه بعد از استفاده در شبیه‌سازی برگشتند.

$$X_n = \frac{X_i - X_{max}}{X_{min} - X_{max}} \quad (5)$$

در این رابطه  $X_i$  مقادیر داده‌های مشاهداتی،  $X_n$  داده‌های استاندارد شده،  $X_{min}$  و  $X_{max}$  به ترتیب معرف داده‌های حداقل و حداکثر می‌باشد.

برای مقایسه داده‌های مشاهداتی و پیش‌بینی‌شده از شاخص‌های آماری ضریب تبیین ( $R^2$ )، ریشه متوسط مربعات خطا (RMSE) و معیار نش ساتکلیف (E)<sup>۱</sup> استفاده گردید.

$$R^2 = \left[ \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \cdot \sum_{i=1}^N (y_i - \bar{y})^2}} \right]^2 \quad (6)$$

جدول ۵- ارزیابی مدل ماشین بردار پشتیبان در شبیه‌سازی رواناب ایستگاه هیدرومتری تازه کند

Table 5. Evaluation of Support Vector Machine to predict runoff in Tazekand hydrometric station

| ترکیب | مدل        | پارامترهای ورودی  | آموزش          |       |                         | صحت سنجی       |       |                         |
|-------|------------|---|----------------|-------|-------------------------|----------------|-------|-------------------------|
|       |            |   | R <sup>2</sup> | E     | RMSE(m <sup>3</sup> /s) | R <sup>2</sup> | E     | RMSE(m <sup>3</sup> /s) |
| ۱     | پایه شعاعی | Q <sub>t-1</sub>  | ۰/۴۴۹          | ۰/۴۴۹ | ۳/۸۰۲                   | ۰/۳۹۴          | ۰/۳۵۸ | ۲/۹۹۳                   |
|       | چندجمله‌ای | Q <sub>t-1</sub>  | ۰/۴۴۹          | ۰/۴۴۹ | ۳/۸۰۳                   | ۰/۳۸۶          | ۰/۳۵۷ | ۲/۹۹۵                   |
|       | خطی        | Q <sub>t-1</sub>  | ۰/۴۲۲          | ۰/۴۲۱ | ۳/۸۹۷                   | ۰/۳۵۷          | ۰/۳۴۷ | ۳/۰۲۰                   |
| ۲     | پایه شعاعی | Q <sub>t-1</sub> , Q <sub>t-2</sub>   | ۰/۷۳۹          | ۰/۷۳۹ | ۲/۱۸                    | ۰/۵۳۸          | ۰/۴۲۸ | ۲/۸۲۵                   |
|       | چندجمله‌ای | Q <sub>t-1</sub> , Q <sub>t-2</sub>   | ۰/۷۳۹          | ۰/۷۲۶ | ۲/۶۸۰                   | ۰/۴۸۲          | ۰/۳۸۱ | ۲/۹۴۰                   |
|       | خطی        | Q <sub>t-1</sub> , Q <sub>t-2</sub>   | ۰/۵۹۲          | ۰/۵۹۲ | ۳/۲۷۰                   | ۰/۴۶۷          | ۰/۳۱۷ | ۲/۸۰۶                   |
| ۳     | پایه شعاعی | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub>  | ۰/۷۵۶          | ۰/۷۵۶ | ۲/۵۳۰                   | ۰/۵۷۹          | ۰/۴۹۸ | ۲/۶۴۷                   |
|       | چندجمله‌ای | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub>  | ۰/۷۴۰          | ۰/۷۳۹ | ۲/۶۱۷                   | ۰/۵۳۵          | ۰/۴۶۲ | ۲/۷۴۱                   |
|       | خطی        | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub>  | ۰/۵۹۷          | ۰/۵۹۷ | ۳/۲۵۱                   | ۰/۴۵۷          | ۰/۴۲۴ | ۲/۸۳۷                   |
| ۴     | پایه شعاعی | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub>                                       | ۰/۷۵۹          | ۰/۷۵۹ | ۲/۵۱۳                   | ۰/۵۷۵          | ۰/۴۷۷ | ۲/۷۰۲                   |
|       | چندجمله‌ای | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub>                                       | ۰/۷۳۱          | ۰/۷۳۱ | ۲/۶۵۷                   | ۰/۵۳۷          | ۰/۴۳۹ | ۲/۸۰۷                   |
|       | خطی        | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub>                                       | ۰/۶۰۶          | ۰/۶۰۶ | ۳/۲۱۴                   | ۰/۴۶۱          | ۰/۴۱۷ | ۲/۸۰۲                   |
| ۵     | پایه شعاعی | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub>                    | ۰/۷۶۲          | ۰/۷۶۲ | ۲/۴۹۷                   | ۰/۵۷۵          | ۰/۴۹۱ | ۲/۶۶۶                   |
|       | چندجمله‌ای | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub>                    | ۰/۷۳۶          | ۰/۷۳۴ | ۲/۶۴۱                   | ۰/۵۴۷          | ۰/۴۳۷ | ۲/۸۰۳                   |
|       | خطی        | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub>                    | ۰/۶۱۳          | ۰/۶۱۳ | ۳/۸۶                    | ۰/۴۷۷          | ۰/۴۲۹ | ۲/۸۲۵                   |
| ۶     | پایه شعاعی | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub> , Q <sub>t-6</sub> | ۰/۷۷۸          | ۰/۷۷۷ | ۲/۴۱۸                   | ۰/۵۸۸          | ۰/۵۰۳ | ۲/۶۳۴                   |
|       | چندجمله‌ای | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub> , Q <sub>t-6</sub> | ۰/۷۴۳          | ۰/۷۴۳ | ۲/۵۹۵                   | ۰/۵۵۸          | ۰/۴۴۱ | ۲/۷۹۴                   |
|       | خطی        | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub> , Q <sub>t-6</sub> | ۰/۶۱۹          | ۰/۶۱۹ | ۳/۱۶۱                   | ۰/۴۸۸          | ۰/۴۳۳ | ۲/۸۱۴                   |

جدول ۶- نتایج شبکه عصبی مصنوعی استفاده‌شده در مراحل آموزش و صحت‌سنجی

Table 6. Results artificial neural network used in the training and validation

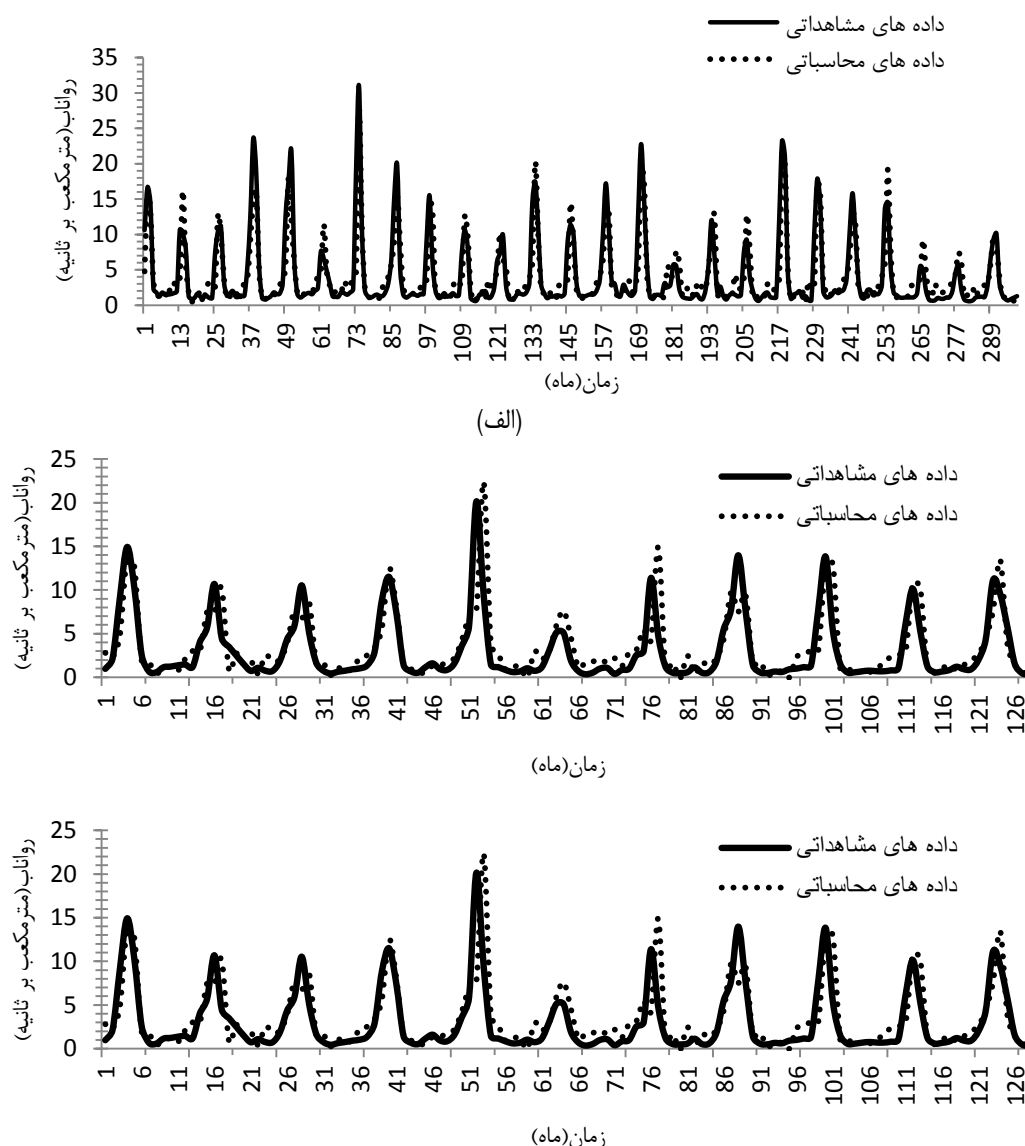
| ترکیب | تعداد نرون پنهان | تعداد تکرار (EPOCH) | پارامترهای ورودی  | آموزش          |       |                         | صحت سنجی       |       |                         |
|-------|------------------|---------------------|---|----------------|-------|-------------------------|----------------|-------|-------------------------|
|       |                  |                     |   | R <sup>2</sup> | E     | RMSE(m <sup>3</sup> /s) | R <sup>2</sup> | E     | RMSE(m <sup>3</sup> /s) |
| ۱     | ۱۲               | ۱۳                  | Q <sub>t-1</sub>  | ۰/۵۴۵          | ۰/۵۳۱ | ۳/۵۰۷                   | ۰/۳۶۴          | ۰/۰۹۶ | ۲/۵۵۲                   |
| ۲     | ۱۰               | ۱۲                  | Q <sub>t-1</sub> , Q <sub>t-2</sub>   | ۰/۷۵۵          | ۰/۷۵۴ | ۳/۵۳۹                   | ۰/۵۱۶          | ۰/۱۲۷ | ۳/۹۷۰                   |
| ۳     | ۱۳               | ۱۴                  | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub>  | ۰/۷۱۲          | ۰/۶۳۴ | ۳/۰۹۹                   | ۰/۵۷۴          | ۰/۲۸۶ | ۴/۲۳۹                   |
| ۴     | ۲۰               | ۲۲                  | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub>                                       | ۰/۷۱۰          | ۰/۷۰۹ | ۲/۸۰۳                   | ۰/۵۳۳          | ۰/۱۱۶ | ۳/۵۱۴                   |
| ۵     | ۷                | ۱۰                  | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub>                    | ۰/۷۲۸          | ۰/۷۲۸ | ۲/۷۱۲                   | ۰/۵۴۹          | ۰/۰۱۶ | ۳/۷۰۷                   |
| ۶     | ۱۷               | ۲۰                  | Q <sub>t-1</sub> , Q <sub>t-2</sub> , Q <sub>t-3</sub> , Q <sub>t-4</sub> , Q <sub>t-5</sub> , Q <sub>t-6</sub> | ۰/۷۷۴          | ۰/۷۳۳ | ۲/۴۷۴                   | ۰/۵۷۹          | ۰/۲۲۴ | ۳/۲۹۱                   |

نشان می‌دهند که ترکیب مدنظر نسبت به سایر ترکیبات مدل‌سازی شده با روش‌های دیگر و دیگر توابع کرنل از دقت بیشتری برخوردار است. در شکل ۳، مقادیر مشاهداتی و پیش‌بینی‌شده برای ترکیب ششم نشان داده شده است. اصولاً نوع تابع کرنل مناسب باید به روش سعی و خطا تعیین گردد (۱۹)؛ با این حال در اکثر تحقیقات به این موضوع اهمیت داده نشده و تابع کرنل RBF به‌طور پیش‌فرض به‌عنوان تابع کاربردی مدل، مورد استفاده قرار گرفته‌است. با توجه به این

همان‌طور که در جدول ۵ مشاهده می‌شود، میزان R<sup>2</sup> در بخش آموزش ششمین ترکیب مدل ماشین بردار پشتیبان برابر با ۰/۷۷۸ به دست آمده که نسبت به ترکیبات قبلی بزرگ‌تر است. همچنین میزان E و RMSE برابر با ۰/۷۷۷ و ۲/۴۱۸ است. در بخش صحت‌سنجی نیز وضعیت به همین ترتیب است. مقدار R<sup>2</sup> در ششمین ترکیب این بخش، بزرگ‌تر از ترکیبات قبلی و برابر با ۰/۵۵۸ است. میزان E و RMSE برابر با ۰/۵۰۳ و ۲/۶۳۴ هستند. شاخص‌های آماری یادشده

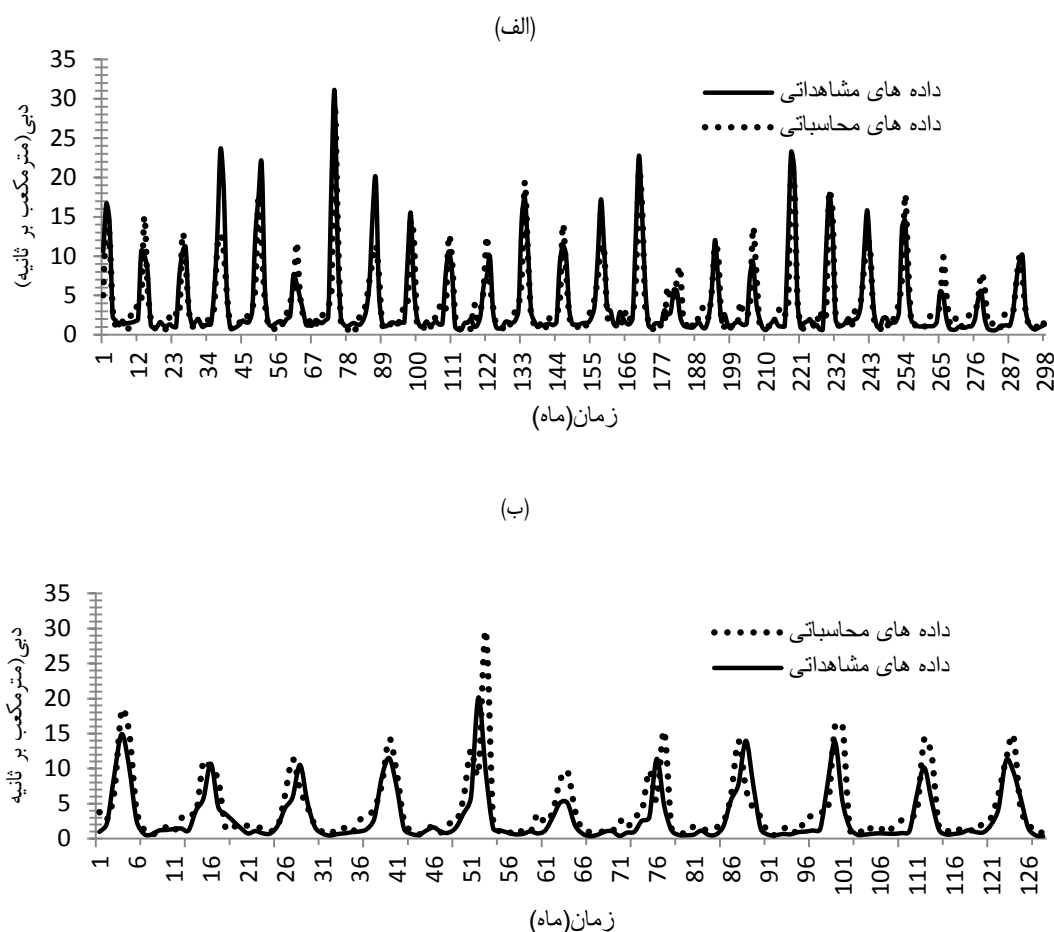
جدول ۶ در بخش آموزش میزان  $R^2$  برابر با ۰/۷۷۴ است که نسبت به ترکیبات قبلی بزرگتر می باشد. همچنین میزان E و RMSE برابر با ۰/۷۷۳ و ۲/۴۷۴ است. در بخش صحت سنجی نیز وضعیت به همین ترتیب است. مقدار  $R^2$  در ششمین ترکیب این بخش، بزرگ تر از ترکیبات قبلی و برابر با ۰/۵۷۹ است. میزان E و RMSE برابر با ۰/۲۲۴ و ۳/۲۹۱ هستند. شاخص های آماری یادشده نشان می دهند که ترکیب مدنظر نسبت به سایر ترکیبات مدل سازی شده با این روش از دقت بیشتری برخوردار است. در شکل ۴، مقادیر مشاهداتی و پیش بینی شده برای ترکیب ششم نشان داده شده است.

که این پیش فرض ممکن است در تمام موارد صحیح نباشد، در این تحقیق از روش سعی و خطا استفاده شد. در جدول ۵ ملاحظه می گردد که در همه ترکیب ها، مدل ماشین بردار پشتیبان با تابع کرنل پایه شعاعی نسبت به مدل با توابع کرنل خطی و چندجمله ای، مقادیر شبیه سازی شده ی بهتری ارائه می دهد، بنابراین می توان اطمینان داشت که تخمین ارائه شده توسط تابع کرنل RBF نسبت به توابع دیگر دقیق تر است. نتایج حاصل از شبیه سازی با شبکه عصبی مصنوعی با مینیمم میزان یادگیری<sup>۱</sup> ۰/۰۰۱ و ماکزیمم میزان یادگیری<sup>۲</sup> ۰/۳ و ضریب مومنتوم<sup>۳</sup> ۰/۸ نشان می دهد ترکیب ششم در



شکل ۳- مقادیر مشاهداتی و پیش بینی شده میزان رواناب ایستگاه تازه کند رودخانه صوفی چای با بهترین ترکیب ماشین بردار پشتیبان (ترکیب ۶) الف- آموزش ب- صحت سنجی

Figure 3. The observed and predicted values of Tazekand station with the best combination of SVM (compound6) the training and validation



شکل ۴- مقادیر مشاهداتی و محاسباتی میزان رواناب ایستگاه تازه کند با بهترین ترکیب شبکه عصبی مصنوعی (ترکیب ششم): (الف) آموزش (ب) صحت سنجی +++

Figure 4. The observed and predicted values of runoff Tazekand station with the best combination of ANN . (compound6) the training and validation

رودخانه‌ها پیشنهاد می‌کند. این پیشنهاد با نتایج حاصل از مطالعه چوبی و همکاران (۴) و دو تحقیق انجام شده توسط یو و همکاران (۲۳، ۲۲) مطابقت دارد. چوبی و همکاران (۴) با استفاده از دو مدل ماشین بردار پشتیبان و شبکه عصبی مصنوعی به پیش‌بینی جریان ماهانه در رودخانه نارمادا<sup>۱</sup> در هند پرداختند. نتایج این تحقیق نشان از برتری ماشین بردار پشتیبان نسبت به شبکه عصبی مصنوعی در پیش‌بینی جریان ماهانه رودخانه مورد نظر داشت. یو و همکاران (۲۳) برای مدل‌سازی سیلاب رودخانه لن‌یانگ<sup>۲</sup> در تایوان از مدل ماشین بردار پشتیبان استفاده کردند. نتایج حاصل، حاکی از دقت خوب ماشین بردار پشتیبان در تخمین سیلاب بود. با دقت بیشتر به شکل‌های ارائه شده برای هر دو روش می‌توان نتیجه گرفت که عملکرد دو روش ماشین بردار پشتیبان و شبکه عصبی مصنوعی از نظر تحلیلی آماری به هم نزدیک است. از نظر گرافیکی و تخمین دبی‌های کمینه و بیشینه، هر دو مدل در پیش‌بینی دبی بیشینه خطای بیشتری دارند اما در پیش‌بینی دبی متوسط و کمینه عملکرد، مدل ماشین بردار پشتیبان به مراتب بهتر از شبکه عصبی مصنوعی است.

در این تحقیق، بهترین ترکیب متغیرهای ورودی با روش آزمون گاما برای مدل‌سازی جریان ماهانه حوزه صوفی چای انجام گرفت. برای اینکار سری‌های مورد استفاده در ورودی مدل، که داده‌های جریان ماهانه با تاخیرهای مختلف بودند، نرمال شدند. سپس از بین شش متغیر ورودی ترکیب بهینه متغیرها با استفاده از آزمون گاما بدست آمد. ترکیب منتخب شامل شش متغیر  $Q_{t-1}$ ,  $Q_{t-2}$ ,  $Q_{t-3}$ ,  $Q_{t-4}$ ,  $Q_{t-5}$  و  $Q_{t-6}$  می‌باشد. نتایج این تحقیق با نتایج رمضان و همکاران (۱۳) مطابقت دارد. ایشان اذعان داشته‌اند با توجه به قابلیت بالای آزمون گاما در کاهش مراحل سعی و خطا در تعیین ورودی‌ها به مدل کمک بسزایی خواهد شد. هدف بعدی، مدل‌سازی جریان رودخانه با استفاده از دو روش ماشین بردار پشتیبان و شبکه عصبی مصنوعی بود. نتایج مدل‌سازی با استفاده از سه معیار  $R^2$  و  $E$  و  $RMSE$  مورد ارزیابی قرار گرفت. نتایج نشان داد که مدل ماشین بردار پشتیبان با تابع کرنل پایه شعاعی از دقت بیشتری نسبت به مدل شبکه عصبی مصنوعی برخوردار است. نتایج حاصله در مجموع روش ماشین بردار پشتیبان را به عنوان یک روش صحیح و دقیق برای پیش‌بینی جریان



## منابع

1. Botsis, D., P. Latinopoulos and k. Diamantaras. 2011. Rainfall-Runoff Moeling Using Suport Vector Regression and Artificial Neural Networks. CEST 2011-Rhodes, Greece.
2. Coulibaly, P., F. Anetil and B. BoBee. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, 230: 244-257.
3. Cristianini, N. and J. Shawe-Taylor. 2000. *An Introduction to Support Vector Machines*, Cambridge University Press, New York.
4. Choubey, V., S. Mishra and S.K. Pandey. 2014. Time Series Data Mining in Real Time Surface Runoff Forecasting through Support Vector Machine. *International Journal of Computer Applications*, 98(3): 23-30.
5. Dawson, C.W., R.J. Abrahart, A.Y. Shamseldin and R.L. Wibly. 2006. Flood estimation at ungauged sites using artifitial neural networks. *Journal of Hydrology*, 319: 391-409.
6. Dibike, Y.B., S. Velickov, D.P. Solomatine and M.B. Abbott. 2001. Model Induction with Support Vector Machines: Introduction and Application, ASCE. *Journal of Computing in Civil Engineering*, 15(3): 208-216.
7. Farajzadeh, J., A. Fakheri Fard and S. Lotfi. 2014. Modeling of monthly rainfall and runoff of Urmia lake basin using feed-forward neural network and time series analysis model. *Water Resources and Industry*, 7(8): 38-48 (In Persian).
8. Jones, A.J. 2004. New tools in non-linear modeling and prediction. *Computational Management Science* 1: 109-149.
9. Kisi, O. and C. Mesut. 2011. A wavelet-support vector machine conjunction model for monthly stream flow forecasting. *Journal of Hydrology*, 399: 132-140.
10. Lafadani, E., A. Moghadamnia, A. Ahmadi, M. Jajarmizadeh and M. Ghafari Gosheh. 2013. Stream flow simulation using SVM, ANFIS and NAM models (A Case study). *Caspian Journal of Applied Sciences Reaserch*, 2(4): 86-93 (In Persian).
11. Lafadani, E., A. Moghadamnia, A. Ahmadi and H. Ebrahimi. 2014. Evaluate the effect of processing the input variables to the model of support vector machine to predict the volume of sediment gamma test. *Iranian Journal of Natural Resources*, 67(2): 289-303 (In Persian).
12. Noori, R., A. Karbassi and M. Sabahi. 2009. Evaluation of PCA and gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management*, 91: 767-771 (In Persian).
13. Remesan, R., M.A. Shamim and D. Han. 2008. Model data selection using gamma test for daily solar radiation estimation *Hydrological Processes*, 22: 4301-4309.
14. Remesan, R., M.A. Shami, D. Han and J. Mathew. 2009. Runoff prediction using an integrated hybrid modeling scheme. *Journal of Hydrology*, 372: 48-60.
15. Rezaee, E., A. Shahidi, A. Khasheh and H. Riyahi Madvar. 2014. Application of Least Squares Support Vector Machine Model for Water Table Simulation (Case Study: Ramhormoz plain). *Iranian Journal of Irrigation and Drainage*, 4: 510-520 (In Persian).
16. Roozbahani, A., B. Zahraie and M. Tabesh. 2012. Risk Analysis of Drinking water system by aggregation of Fuzzy Fault Tree Analysis with Bayesian networks and Dempster-shafer theory ,10<sup>th</sup> International Conference on Hydro informatics ,Humborg ,Germany.
17. Sharifi, A., Y. Dinpashoh, A. Fakheri-Fard and A. Moghaddamniya. 2013. Optimum Combination of Variables for Runoff Simulation in Amameh Watershed using Gamma test. *Journal of Water and Soil*, 23(4): 59-72 (In Persian).
18. Seifi, A. and H. Riahi-Madvar. 2012. Input Variable Selection in expert systems based on hybrid Gamma Test-Least Square Support Vector Machine, ANFIS and ANN models. Provisional chapter. INTECH.
19. Seifi, A. 2010. 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, Tarbiat Modares University, Tehran.
20. Taheri, H. and M. Ghafouri. 2012. Comparison between active learning method and support vector machine for runoff modeling. *Journal hydraulic Hydromech*, 1: 16-32 (In Persian).
21. Yoon, H. and S.C. Jun. 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396: 128-138.
22. Yu, P.S., S.T. Chen and I.F. Chang. 2005. Flood stage forecasting using support vector machines. *Geophysical Research Abstracts*, 7: 41-76.
23. Yu, P.S., S.T. Chen and I.F. Chang. 2006. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328: 704-716.

## **Sofichay River Runoff Modeling using Support Vector Machine and Artificial Neural Network**

**Fateme Akhoni Pourhosseini<sup>1</sup> and Sabereh Darbandi<sup>2</sup>**

---

1- M.Sc. Student, Department of Engineering, Faculty of Agriculture, University of Tabriz  
(Corresponding Author: fateme.pourhosseini@yahoo.com)

2- Assistant Professor, Department of Engineering, Faculty of Agriculture, University of Tabriz  
Received: April 21, 2016                      Accepted: September 4, 2016

---

### **Abstract**

Accurate simulation runoff process can have a significant role in water resources management and related issues. The inherent complexity of this process makes difficult the use of physical and numerical models. In recent years, application of intelligent models is increased a powerful tool in hydrological modeling. The aim of this study was the application of the Gamma test to select the optimal combination of input variables for runoff modeling in Sofi Chay. Streamflow modeling was performed based on the optimum number of the selected variables using the artificial neural network (ANN) and Support vector machine (SVM) methods. Gamma test results showed that monthly runoff with six antecedent runoff values provide better results to predict. Runoff simulation using support vector machines and artificial neural network models also showed that the best input structure will be delayed until six to predict of next month runoff. Among to models with the same input structure, support vector machine have relatively high efficiency compared to artificial neural network.

**Keywords:** Artificial Neural Network, Gamma Test, Runoff, SofiChay River, Support Vector Machine