دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشنامه مدیریت حوزه آبخیز

ارزیابی تأثیر نرمال‌سازی توزیع احتمالاتی رواناب ماهانه بر عملکرد مدل‌های SVM و ANN در شبیه‌سازی جریان ماهانه رودخانه‌ها (مطالعه موردی: حوزه زرینه‌رود)

محمد عیسی‌زاده[۱]، حجت احمدزاده[۲] و محمدعلی قربانی[۳]

۱- دانشجوی کارشناسی ارشد مهندسی منابع آب، دانشگاه تبریز، (نویسنده مسوول: mohammadisazade@gmail.com)
۲ و ۳- دانشجوی دکتری و دانشیار، گروه علوم و مهندسی آب، دانشگاه تبریز
تاریخ دریافت: ۹۴/۲/۱ تاریخ پذیرش: ۹۵/۲/۱۸

چکیده

تخمین دقیق مقدار رواناب رودخانه‌ها از اقدامات اساسی در مدیریت منابع آب حوزه‌ها می‌باشد. شبکه عصبی مصنوعی(ANN) و ماشین بردار پشتیبان (SVM) از مهم‌ترین مدل‌های داده‌کاوی هستند که می‌توانند برای این منظور مدنظر قرار گیرند. با توجه به مبتنی بودن این مدل‌ها به داده‌ها ، نوع توزیع احتمالاتی داده‌ها می‌تواند تأثیر چشمگیری بر عملکرد آن‌ها در شبیه‌سازی جریان داشته باشد. برای بررسی این موضوع حوزه زرینه‌رود به‌عنوان منطقه مورد مطالعه انتخاب شده و ارزیابی‌های مورد نظر در سه ایستگاه آب‌سنجی واقع در آن صورت گرفت. در این راستا در ابتدا با استفاده از آزمون شاپیرو-ویلک تبعیت نمودن توزیع احتمالاتی داده‌های مشاهداتی اولیه از نرمال مورد بررسی قرار گرفته و سپس نرمال‌سازی توزیع داده‌ها انجام شد. در ادامه کارایی هر یک از مدل‌های ANN و SVM در شبیه‌سازی رواناب ماهانه سه ایستگاه آب سنجی برای داده‌های مشاهداتی و نرمال شده جریان مورد ارزیابی قرار گرفت. براساس نتایج این تحقیق مقادیر شاخص‌های ضریب همبستگی(CC) و جذر میانگین مربعات خطا (RMSE) در دوره صحت‌سنجی برای ایستگاه‌های صفاخانه، سنته و پل آنیان در مدل ANN به ترتیب ۰/۷۱، (m³/sec) ۵/۹۳، ۰/۸۰، (m³/sec) ۶/۵۸ و ۰/۸۲، (m³/sec) ۲۲/۹۰ بدست آمد. در مدل SVM برای این شاخص‌ها در ایستگاه های مذکور به ترتیب مقادیر ۰/۷۰، (m³/sec)٦/٣٤، ۰/۷۸، (m³/sec)۷/۰۲ و ۰/۷۹، (m³/sec)٢٤/٣١ حاصل شد. نتایج نشان داد که در مدل ANN استفاده از جریان نرمال شده در ایستگاه‌های پل آنیان، سنته و صفاخانه مقادیر CC را در دوره صحت‌سنجی به ترتیب ٦، ١٤ و ١١ درصد افزایش و مقادیر RMSE را به ترتیب ۹، ۱۹ و ٦ درصد کاهش می‌دهد. در مدل SVM پس از نرمال‌سازی داده‌ها مقادیر CC و RMSE فقط برای ایستگاه سنته به میزان ۱۰ و ١٦ درصد به ترتیب افزایش و کاهش می‌یابد. همچنین نتایج نشان داد که مدل ANN با ورودی‌های نرمال شده در هر سه ایستگاه آب‌سنجی عملکرد بهتری نسبت به مدل SVM دارد.

واژه‌های کلیدی: رواناب ماهانه، ANN، SVM، توزیع احتمالاتی نرمال، حوزه زرینه‌رود

مقدمه

انتخاب مدلی که بتواند جریان رودخانه را با دقت بالا و خطای کمتر شبیه‌سازی نماید کمک شایانی در مدیریت کارآمد طرح‌های منابع آبی دارد. استفاده از مدل‌های فیزیکی در حوزه‌هایی که در آن‌ها اطلاعات کافی از پارامترها و متغیرهای دخیل در جریان رودخانه در دسترس نباشد مقدور نمی‌باشد. مدل‌های داده‌کاوی از مهم‌ترین گزینه‌های مطرح جهت شبیه‌سازی رواناب در چنین حوزه‌هایی هستند (۱۸). این مدل‌های متکی به داده ساختار ریاضی مشخصی داشته و بدون اینکه فرآیند فیزیکی حاکم بر پدیده را شبیه‌سازی نمایند، توانایی تشخیص روابط پیچیده غیرخطی بین داده‌های ورودی و خروجی را دارا می‌باشند (۱۵). در مدل‌های داده‌کاوی بدلیل ماهیتی که دارند ساختار مدل، نوع و ترکیب داده‌های بکار گرفته شده در شبیه‌سازی تأثیر بسزایی در عملکرد آن‌ها دارند (۱۵). علاوه بر ترکیب داده‌های ورودی، توزیع احتمالاتی آنها از مهم‌ترین ویژگی‌هایی است که می‌تواند تأثیر چشمگیری در کارایی مدل‌های داده‌کاوی داشته باشد که در تحقیق حاضر به این موضوع مهم پرداخته شده است. شبکه عصبی مصنوعی (ANN)[۱] ماشین بردار پشتیبان (SVM)[۲] از مهم‌ترین مدل‌های داده‌کاوی هستند که در دهه‌های اخیر جهت تخمین جریان رودخانه‌ها مورد توجه محققین قرار گرفته‌اند. شبکه‌های عصبی مصنوعی به‌عنوان یکی از روش‌های جعبه سیاه این قابلیت را دارند که روابط منطقی بین پارامترهای ورودی و خروجی یک پدیده را بدست آورند. این مدل‌ها با تجزیه و تحلیل داده‌های ورودی و نتایج نظیر آن‌ها ارتباط منطقی بین داده‌ها برقرار می‌کنند که ممکن است غیرخطی و نامشخص باشد. سپس با استفاده از این ارتباط منطقی، شبیه‌سازی‌ها برای موارد مشابه انجام می‌گیرد (۱۱). ماشین‌های بردار پشتیبان، سیستم‌های یادگیری خاصی هستند که از یک فضای فرضیه به نام فضای مشخصه[۳]، که شامل توابع خطی با بعد زیاد است، استفاده می‌کند. ماشین‌های بردار پشتیبان می‌توانند برای پیش‌بینی (دسته‌بندی و رگرسیون) مورد استفاده قرار گیرند. این سیستم‌ها توسط الگوریتم‌های یادگیری که بر اساس تئوری بهینه‌سازی استوار است، واسنجی می‌شوند (۴).

تاکنون تحقیقات زیادی در زمینه کارایی هر یک از مدل‌های مذکور و مقایسه عملکرد آن‌ها با دیگر مدل‌ها در شبیه‌سازی جریان رودخانه‌ها صورت گرفته است. هانگ و همکاران (۸) دقت مدل‌های ANN و ARMA در پیش‌بینی جریان رودخانه آپالاچیکولا در ایالات متحده را مورد مقایسه قرار دادند. نتایج تحقیق آنان حاکی از عملکرد بهتر شبکه عصبی پرسپترون چند لایه (MLP)[۴] با الگوریتم آموزش پس انتشار خطا نسبت به مدل ARMA می‌باشد. میسرا و همکاران (۱۴) عملکرد مدل‌های SVM و ANN در پیش‌بینی رواناب و بار رسوب حوزه آبریز وامساهارا واقع در جنوب کشور هند را مورد مقایسه قرار دادند. نتایج نشان داد مدل SVM هم در پیش‌بینی رواناب و هم بار رسوب نسبت به ANN دارای عملکرد بهتری می‌باشد. در تحقیقی دیگر نوری و همکاران (۱۶) تأثیر انتخاب متغیرهای ورودی با استفاده از آنالیز مؤلفه اصلی بر عملکرد مدل SVM در پیش‌بینی دبی ماهانه رودخانه صوفی‌چای را مورد بررسی قرار دادند. همچنین لیو (۱۳) در مطالعه خود کارآیی مدل‌های

1- Artificial Neural Network 2- Support Vector Machine 3- Feature Space 4- Multi Layer Perceptron

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]
[DOI: 10.29252/jwmr.8.16.22]

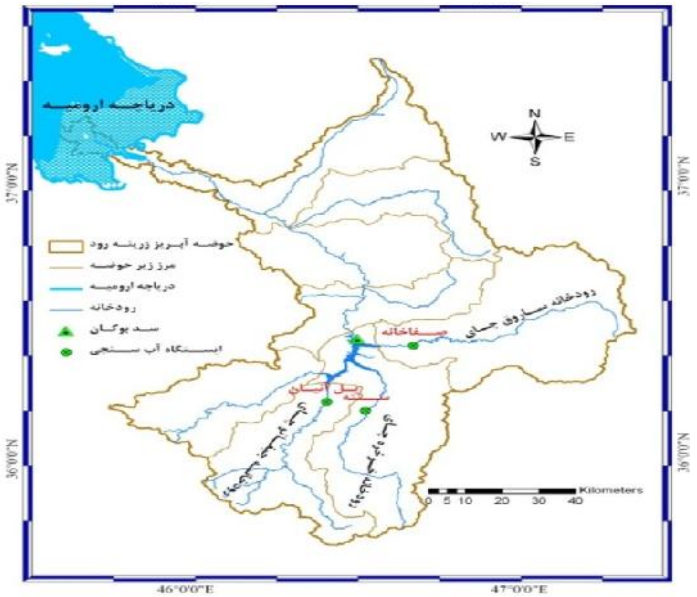
رگرسیونی، ARIMA و ANN را در پیش‌بینی جریان روزانه رودخانه وایت‌کلی‌کریک نزدیک ایالت پنسیلوانیا مورد ارزیابی قرار داده و دقت آن‌ها را با یکدیگر مقایسه کردند. نتایج تحقیق آنان نشان داد زمانی که داده‌ها از توزیع نرمال تبعیت کرده و دارای طول دوره آماری کافی باشند، مدل‌های خطی کمترین خطا را در پیش‌بینی جریان رودخانه دارا می‌باشند. آداموفسکی (۱) مدل‌های SVM و ANN را برای شبیه‌سازی بارش-رواناب در یک حوزه کوهستانی با داده‌های محدود در یوتارانچال هندوستان به کار گرفت. نتایج بررسی وی نشان از توانایی مدل SVM با تابع کرنل پایه شعاعی(RBF)[1] در پیش‌بینی رواناب مستقیم، جریان پایه و جریان کل این منطقه کوهستانی و پر از تپه داشت. در تحقیقی دیگر کاکائی لفادانی و همکاران (۹) سه مدل SVM، نروفازی و مدل هیدرولوژیکی مفهومی Mike11/NA را جهت شبیه‌سازی جریان روزانه حوزه اسکندری به‌کار گرفتند. نتایج نشان داد که هر سه مدل دارای عملکرد مناسب در منطقه مورد مطالعه می‌باشند. درنهایت مدل SVM با تابع کرنل RBF بالاترین ضریب همبستگی و کمترین میانگین مربعات خطا را در شبیه‌سازی جریان نسبت به دو مدل دیگر از خود نشان داد.

همانطور که قبلا نیز ذکر شد در مدل‌های داده‌کاوی علاوه بر ساختار بکار گرفته شده برای مدل‌ها، ترکیب و خصوصیات آماری داده‌های ورودی به این مدل‌ها تأثیر بسزایی در عملکرد آن‌ها دارد، توزیع احتمالاتی داده‌ها از مهم‌ترین آن‌ها می‌باشد. لذا ارزیابی تأثیر نرمال‌سازی توزیع احتمالاتی داده‌های ورودی بر عملکرد مدل‌های ANN و SVM هدف تحقیق حاضر را رقم می‌زند. برای این منظور حوزه‌های خرخره‌چای، ساروق‌چای و جیغاتوچای به‌عنوان مناطق مورد مطالعه انتخاب گردیدند.

مواد و روش‌ها

داده‌ها و منطقه مورد مطالعه

در این تحقیق حوزه‌های خرخره‌چای، ساروق‌چای و جیغاتوچای که از زیرحوزه‌های اصلی حوزه زرینه‌رود محسوب می‌شوند، به‌عنوان محدوده مورد مطالعه انتخاب گردیدند. این حوزه‌ها در در طول جغرافیایی '۵۳ º۴۵ تا '۲۰ º۴۷ شرقی و عرض جغرافیایی '۴۱ º۳۵ تا '۵۲ º۳۶ شمالی قرار دارد (شکل ۱). این سه حوزه سهم قابل توجهی در حجم آب ورودی به سد بوکان و در نتیجه تأمین نیازهای آبی پایین دست این سد دارد. ایستگاه‌های سنته، صفاخانه و پل آنیان که به ترتیب در حوزه‌های خرخره‌چای، ساروق‌چای و جیغاتوچای واقع شده‌اند از جمله ایستگاه‌های آب‌سنجی هستند که از طول دوره آماری قابل قبول برخوردار می‌باشند. مساحت حوزه بالادست هر یک از ایستگاه‌های آب‌سنجی مذکور به ترتیب بالغ بر ۱۲۳۲، ۲۲۰۹ و ۱۳۲۸ کیلومتر مربع می‌باشند. داده‌های مورد استفاده در این تحقیق شامل جریان ماهانه مشاهداتی در ایستگاه‌های آب‌سنجی سنته، صفاخانه و پل آنیان می‌باشند که ۷۵ درصد داده‌های دوره آماری (۱۳۸۴-۱۳۶۷) به‌عنوان داده‌های واسنجی و ۲۵ درصد دیگر (۱۳۹۰-۱۳۸۵) به‌عنوان داده‌های صحت‌سنجی هر یک از مدل‌ها در نظر گرفته شده است.



شکل ۱- موقعیت حوزه زرینه‌رود و ایستگاه‌های آب‌سنجی سنته، صفاخانه و پل آنیان
Figure 1. Location of Zarrinehrood basin and Santeh, Safakhaneh and Polanian hydrometric stations

1- Radial Basis Function

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

نرمال‌سازی توزیع احتمالاتی داده‌ها

پس از انجام آزمون شاپیرو–ویلک بر روی داده‌های مشاهداتی، چنانچه داده‌ها از توزیع نرمال تبعیت نکنند، می‌توان با بکارگیری روابط مختلف توزیع احتمالاتی آنها را به نرمال تبدیل نمود. از معادلات مختلفی جهت نرمال‌سازی داده‌ها استفاده می‌شود که لگاریتم‌گیری و جذر گرفتن از جمله مهم‌ترین آن‌ها می‌باشند (۱۷). لازم به ذکر است که معادلات بکار گرفته شده برای هر سری از داده می‌تواند متفاوت از سری‌های دیگر باشد (۱۷). در این تحقیق از روابط ۱ تا ۳، ۴ تا ۶ و ۷ تا ۹ به ترتیب جهت نرمال‌سازی توزیع احتمالاتی داده‌های رواناب ماهانه ایستگاه‌های سنته، صفاخانه و پل آنیان استفاده گردیده است.

$$P_{ij} = Ln(Q_{ij} + 0.052) \quad (۱)$$

$$W_{ij} = P_{ij} - \overline{Q}_j \quad (۲)$$

$$K_{ij} = \frac{W_{ij}}{SD_W} \quad (۳)$$

$$P_{ij} = Ln(Q_{ij} + 0.5) \quad (۴)$$

$$W_{ij} = P_{ij} - \overline{Q}_j \quad (۵)$$

$$K_{ij} = \frac{W_{ij}^{0/3}}{SD_W} \quad (۶)$$

$$P_{ij} = Ln(Q_{ij} + 0.115) \quad (۷)$$

$$W_{ij} = P_{ij} - \overline{Q}_j \quad (۸)$$

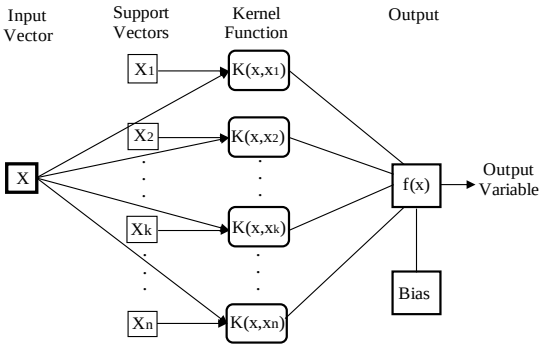
$$K_{ij} = \frac{W_{ij}}{SD_W} \quad (۹)$$

$$i = 1:n \quad , \quad j = 1:12$$

در این روابط n تعداد سال‌ها، Q_{ij} میزان جریان مشاهداتی در ماه jام از سال iام، $\overline{Q}_j$ میانگین دراز مدت جریان در ماه jام، SD_W انحراف معیار متغیر W_{ij} و K_{ij} جریان نرمال شده در ماه jام از سال iام می‌باشند. بعد از اعمال روابط فوق بر روی داده‌های اولیه غیرنرمال ایستگاه‌ها، شاخص‌های آماری میانگین، مقدار حداقل، مقدار حداکثر، انحراف معیار و چولگی برای هر یک از ایستگاه‌ها محاسبه شده و سپس آزمون شاپیرو–ویلک در سطح معنی‌داری[1] ۰/۰۵ جهت بررسی نرمال بودن داده‌ها انجام گرفت.

ماشین بردار پشتیبان (SVM)

اولین کاربرد این روش در مسائل آب توسط دیباک و همکاران (۶) با شبیه‌سازی بارش– رواناب ارائه شد. ماشین بردار پشتیبان یک سیستم یادگیری کار آمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد. شکل (۲) ساختار ماشین بردار پشتیبان را نشان می‌دهد.



شکل ۲– ساختار ماشین بردار پشتیبان (۲۰)

Figure 2. Support Vector Machine structure (20)

در مدل رگرسیون SVM تابعی مرتبط با متغیر وابسته y که خود تابعی از چند متغیر مستقل x است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی فرض می‌شود رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند f(x) به علاوه مقداری اغتشاش (خطای مجاز ε) مشخص شود.

$$f(x) = W^T.\phi(x) + b \quad (۱۰)$$

$$y = f(x) + noise \quad (۱۱)$$

چنانچه W بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و φ نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای f(x) است. این مهم با واسنجی مدل SVM توسط مجموعه‌ای از نمونه‌ها (مجموعه واسنجی) محقق می‌شود. برای محاسبه W و b لازم است تابع خطا (معادله ۱۲) در مدل SVM رگرسیونی با در نظر گرفتن شرایط مندرج (قیود) در معادله (۱۳) بهینه شود.

$$W^T.\phi(x_i) + b - y_i \le \varepsilon + \varepsilon_i^* \quad (۱۲)$$

$$\frac{1}{2}W^T.w + C\sum_{i=1}^{N}\varepsilon_i + C\sum_{i=1}^{N}\varepsilon_i^* y_i$$
$$-W^T.\varphi(x_i) - b \le \varepsilon + \varepsilon_i \quad (۱۳)$$
$$\varepsilon_i, \varepsilon_i^* \ge 0, i = 1,2,...,N$$

در معادلات بالا C عددی صحیح و مثبت است، که عامل تعیین جریمه در هنگام رخ دادن خطای واسنجی مدل می‌باشد، φ تابع کرنل، N تعداد نمونه‌ها و دو مشخصه ε_i^* و ε_i متغیرهای کمبود هستند.

در نهایت تابع SVM رگرسیونی را می‌توان به فرم زیر بازنویسی کرد:

$$f(x) = \sum_{i=1}^{N}\overline{\alpha_i}.\phi(x_i)^T.\phi(x) + b \quad (۱۴)$$

1- Significant Level

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

در معادله (۱۴)، $\overline{\alpha_i}$ میانگین ضرایب لاگرانژ می‌باشد. محاسبه $\varphi(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد. برای حل این مشکل روند معمول در مدل SVM رگرسیونی انتخاب یک تابع کرنل است.

ماشین بردار پشتیبان معمولا با سه تابع کرنل RBF، چند جمله‌ای درجه d و خطی مورد استفاده قرار می‌گیرد که روابط مورد استفاده در هر کدام از آن‌ها به ترتیب در زیر آورده شده است (۷،۱۰).

$$k(x,x_i)=\exp\left(-\|x-x_i\|^2/\sigma^2\right) \quad (۱۵)$$

$$k(x,x_i)=\left(t+\langle x,x_i\rangle\right)^d \quad (۱۶)$$

$$k(x,x_i)=x.x_i \quad (۱۷)$$

لازم بذکر است فرآیند محاسبات ماشین بردار پشتیبان، بر اساس کدنویسی در محیط متلب انجام شده و پارامترهای توابع کرنل از طریق سعی و خطا بهینه گردیدند.

شبکه عصبی مصنوعی با آموزش لونبرگ-مارکوارت

ANN یک سیستم پردازشگر توزیعی اطلاعات است که خصوصیات عملکرد آن شبیه به ساختار عصبی مغز انسان بوده و گسترش آن بر اساس قواعد زیر است.

– پردازش اطلاعات در عناصر منفردی به نام گره انجام و این گره‌ها در دسته‌هایی به نام لایه مرتب می‌شوند.

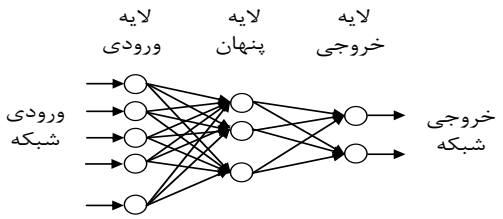
– سیگنال‌ها بین گره‌ها از طریق اتصالات[1] منتقل می‌شود.

– هر اتصال وزنی دارد که نشان‌دهنده نیروی اتصال آن است.

– هر گره مجموعه وزنی سیگنال‌ها را با به‌کار بردن یک تبدیل غیرخطی که تابع محرک[2] نامیده می‌شود به سیگنال خروجی تبدیل می‌کند (۲۱،۳).

می‌توان گفت که حدود ۹۰ ٪ از مدل‌های شبکه عصبی که در مسائل هیدرولوژی بکار گرفته شده‌اند از نوع شبکه‌های عصبی پیشرو[3] هستند (۵) که مدل MLP با الگوریتم آموزش لونبرگ–مارکوارت (LM)[4] از مهم‌ترین آنها بشمار می‌رود. معماری معمول این شبکه متشکل از سه بخش لایه ورودی، پنهانی و خروجی می‌باشد. تعداد نرون‌های موجود در لایه ورودی و خروجی بستگی به نوع مسئله دارد ولی تعداد گره‌های لایه پنهان با سعی خطا بدست می‌آید.

در شبکه‌های پیشرو، اتصالات میان گره‌هایی وجود دارد که در لایه‌های متفاوت قرار دارند. همچنین ورودی‌ها در لایه ورودی شبکه نمایش داده شده و عمل تحریک از ورودی به سمت خروجی شبکه شروع می‌شود (۳). شمای کلی آن در شکل (۳) آمده است.



شکل ۳– ساختار کلی شبکه عصبی مصنوعی (۲)

Figure 3. Artificial Neural Network general structure (2)

لازم به ذکر است که در تحقیق حاضر از تعداد نرون‌های ۱ تا ۲۰ برای لایه پنهان شبکه استفاده گردید تا از این طریق بتوان تاثیر ساختار مختلف شبکه بر عملکرد آن در شبیه‌سازی رواناب را مورد ارزیابی قرار دهد. از آنجا که متداول‌ترین نوع توابع محرک تابع تانژانت سیگموئید می‌باشد لذا در این مطالعه از آن برای نگاشت اطلاعات از لایه ورودی به لایه پنهان و از تابع محرک خطی برای نگاشت اطلاعات از لایه پنهان به لایه خروجی استفاده شد.

تعیین زمان تاخیر و معیارهای ارزیابی

توابع ACF[5] و PACF[6] از جمله معیارهای مطرح در تعیین تعداد ورودی‌های موثر در شبیه‌سازی متغیر مورد نظر می‌باشند (۱۹،۱۲). در تحقیق حاضر نیز از این توابع جهت تعیین تعداد تاخیرهای جریان ماهانه لازم که باید در ارزیابی کارایی مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان مورد بررسی قرار گیرند استفاده گردید.

با استفاده از روش‌های مختلفی می‌توان عملکرد مدل‌ها را مورد ارزیابی و مقایسه قرار داد، یکی از این روش‌ها استفاده از معیارهای ارزیابی می‌باشد، از جمله معیارهای ارزیابی پر کاربرد در هیدرولوژی، ضرایب همبستگی (CC)[7]، ریشه میانگین مربعات خطا (RMSE)[8] و معیار نش– ساتکلیف (NS)[9] می‌باشند که در این تحقیق نیز از این معیارها برای بررسی کارایی هر یک از مدل‌ها استفاده شده است.

نتایج و بحث

نرمال‌سازی داده‌ها

بعد از نرمال سازی داده‌ها، شاخص‌های آماری میانگین، حداقل و حداکثر جریان ماهانه، انحراف معیار و چولگی محاسبه گردیده و سپس آزمون شاپیرو–ویلک در سطح اطمینان ۰/۹۵ بر روی داده‌های مشاهداتی اولیه و نرمال شده انجام گرفت که نتایج آنها در جدول ۱ آورده شده است. لازم به ذکر است چنانچه طی آزمون شاپیرو– ویلک سطح معنی‌داری بیشتر از ۰/۰۵ بدست آید داده‌ها با سطح اطمینان ۰/۹۵ از توزیع نرمال تبعیت می‌کنند. بر این اساس از نتایج آمده در جدول ۱ می‌توان دریافت که توزیع احتمالاتی جریان ماهانه هر یک از سه ایستگاه آب‌سنجی در ابتدا نرمال نبوده ولی پس از اعمال معادلات ۱ تا ۹ توزیع آن‌ها به نرمال تبدیل گردیده است. به‌عنوان نمونه شکل ۴

1- Links 2- Activation Function 3- Feed forward neural networks 4- Levenberg-Marquardt
5- Auto Correlation Function 6- Partial Auto Correlation Function 7- Correlation Coefficient 8- Root Mean Square Error
9- Nash -Sutcliffe

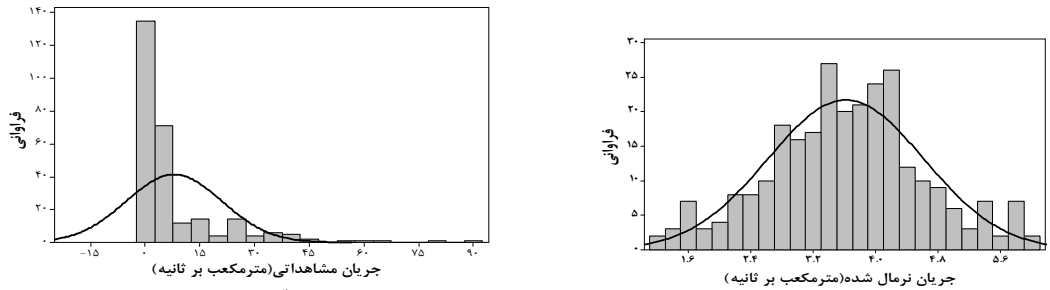
[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]
[DOI: 10.29252/jwmr.8.16.22]

هیستوگرام فراوانی داده‌های اولیه و نرمال شده را برای ایستگاه آب‌سنجی صفاخانه نشان می‌دهد.

جدول ۱- پارامترهای آماری داده‌های ماهانه جریان رودخانه‌ها

Table 1. Statistical parameters of monthly river flow data

ایستگاه	داده	داده‌های مشاهداتی				داده‌های مشاهداتی نرمال شده				شاپیرو-ویلک	
		جریان متوسط (m^3/s)	انحراف معیار	چولگی	جریان حداکثر (m^3/s)	جریان حداقل (m^3/s)	جریان متوسط (m^3/s)	انحراف معیار	چولگی	آماره	سطح معنی‌داری
صفاخانه	کل	۷/۸۵	۱۳/۱۳	۲/۹۲	۸۸/۱۳	۰/۰۰	۳/۶۲	۰۰/۱	-۰/۰۲	۰/۰۴	۰/۲۱
	واسنجی	۸/۶۴	۱۴/۳۰	۲/۷۶	۸۸/۱۳	۰/۰۰	۳/۶۵	۰۳/۱	۰/۰۴	۰/۰۳	۰/۴۳
	صحت سنجی	۵/۴۲	۸/۲۱	۲/۱۸	۳۵/۰۷	۰/۰۰	۳/۵۰	۰/۸۸	-۰/۴۳	۰/۱۰	۰/۰۶
سنته	کل	۸/۳۹	۱۳/۸۳	۲/۶۲	۸۹/۴۳	۰/۰۰	۰/۰۰	۰۰/۱	۰/۰۰	۰/۰۴	۰/۲۶
	واسنجی	۸/۸۹	۱۴/۶۵	۲/۵۵	۸۹/۴۳	۰/۰۰	۰/۰۶	۱/۰۱	-۰/۰۵	۰/۰۵	۰/۲۸
	صحت سنجی	۶/۸۵	۱۰/۹۲	۲/۶۰	۵۶/۹۰	۰/۰۰	-۰/۱۹	۰/۹۴	۰/۱۰	۰/۰۵	۰/۷۱
پل آنیان	کل	۱۷/۷۲	۳۰/۵۲	۳/۰۲	۲۱۳/۱۶	۰/۰۰	۰/۰۰	۸/۹۱	۰/۱۱	۰/۰۳	۰/۷۷
	واسنجی	۱۶/۸۱	۲۶/۸۲	۲/۵۱	۱۵۰/۶۷	۰/۰۰	۰/۷۰	۸/۸۴	-۰/۰۴	۰/۰۳	۰/۹۸
	صحت سنجی	۲۰/۴۹	۳۹/۸۹	۳/۱۷	۲۱۳/۱۶	۰/۰۰	-۲/۱۵	۸/۸۴	۰/۶۰	۰/۰۸	۰/۱۵



شکل ۴- هیستوگرام فراوانی جریان مشاهداتی اولیه و نرمال شده برای ایستگاه آب‌سنجی صفاخانه

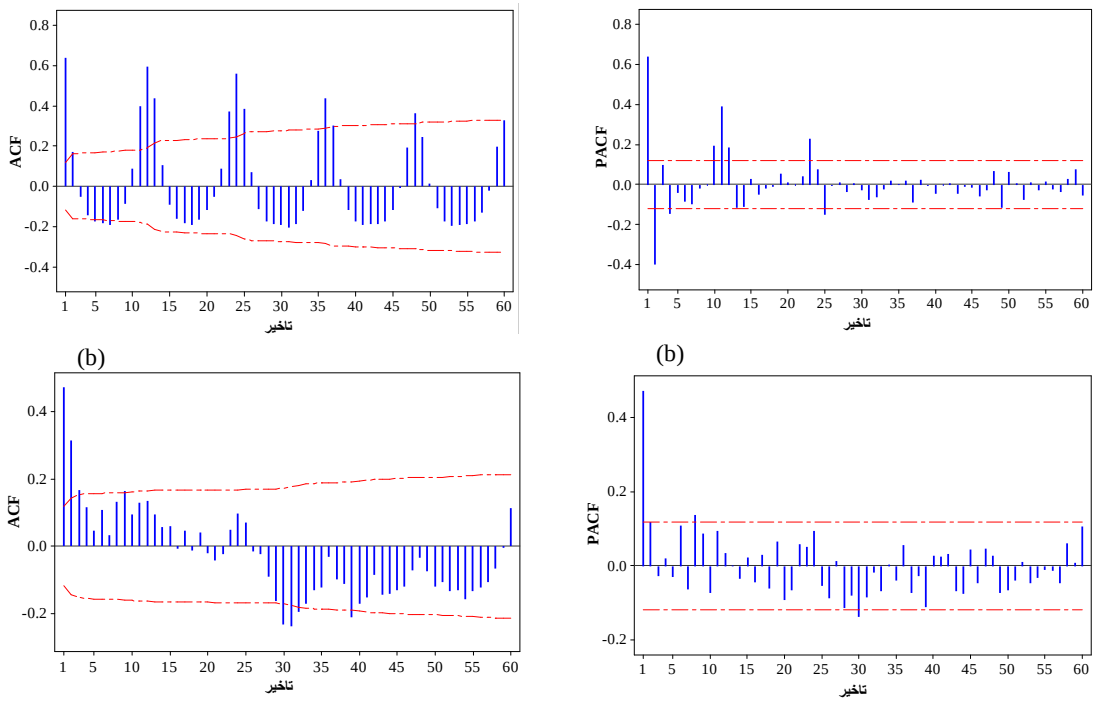
Figure 4. Frequency histogram of the initial and normalized observation flow for the Safakhaneh hydrometric station

تعیین تعداد ورودی مدل‌ها

توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) از جمله معیارهای مطرح در تعیین تعداد ورودی‌های موثر در شبیه‌سازی خودهمبسته متغیر مورد نظر می‌باشند. همچنین با استفاده از رفتار توابع مذکور می‌توان خواص ایستایی و فصلی بودن داده‌ها را مورد بررسی قرار داد (۱۲). ACF در واقع همان ضریب همبستگی بین سری زمانی داده‌های مشاهداتی یک متغیر با تاخیرهای مختلف آن می‌باشد. چنانچه تابع ACF بر حسب زمان تأخیر در یک دستگاه مختصات رسم گردد نمودار همبستگی‌نگار بدست می‌آید. تابع PACF که از حل پی‌درپی معادلات یول-واکر به‌ازای تأخیرهای مختلف بدست می‌آید، جهت بررسی همبستگی بین سری زمانی داده‌های مشاهداتی یک متغیر (xt) و سری تأخیر یافته آن (xt+k) با حذف وابستگی خطی مشترک بین متغیرهای (xt+1, xt+2, …, xt+k-1) بکار می‌رود (۱۷). به عنوان نمونه در شکل ۵ نمودار تابع‌های ACF و PACF مربوط به داده‌های جریان ماهانه مشاهداتی و نرمال شده ایستگاه آب‌سنجی صفاخانه نشان داده شده است.

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]



شکل ۵- نمودار توابع ACF و PACF جریان ماهانه مشاهداتی (a) و نرمال شده (b) در ایستگاه آب‌سنجی صفاخانه
Figure 5. diagrams of the ACF and PACF functions of the monthly observation flow (a) and normalized flow (b) at the Safakhaneh hydrometric station

مطابق با نمودار PACF، جریان مشاهداتی ۱، ۲، ۱۱، ۱۲ و ۱۳ ماه قبل، موثرترین جریان‌ها در برآورد جریان ماهانه مشاهداتی هر ماه می‌باشند و همچنین نمودار PACF جریان نرمال شده نشان می‌دهد که اولین جریان تاخیر یافته (درجه تاخیر برابر با یک) بیشترین تاثیر را در برآورد جریان ماهانه نرمال شده دارد. تغییرات نمودار ACF جریان مشاهداتی اولیه به‌علت دوره‌ای بودن جریان، در سال‌های مختلف مشابه به یکدیگر بوده و این نمودار حالت کسینوسی به‌خود گرفته است. ولی نمودار ACF مربوط به جریان نرمال شده به علت حذف روند میانگین و واریانس و همچنین حذف دوره‌ای بودن جریان‌ها بر اساس فرمول‌های ۴ تا ۶، حالت دوره‌ای نداشته و فقط مقادیر همبستگی تاخیرهای ۱ تا ۱۵ ماهه جریان خارج از باند استقلال قرار گرفته‌اند. بنابراین جهت بررسی دقت تابع‌های ACF و PACF و همچنین یکسان‌سازی تعداد و ترکیب ورودی مدل‌ها در حالت جریان مشاهداتی و نرمال شده؛ تعداد تاخیرهای لازم جهت شبیه‌سازی جریان برابر با ۱۵ ماه در نظر گرفته شد.

شبیه‌سازی رواناب با مدل SVM

در تحقیق حاضر از توابع کرنل چند جمله‌ای، پایه شعاعی و خطی جهت شبیه‌سازی جریان استفاده گردید و مقادیر بهینه پارامترهای مورد نیاز مدل SVM از طریق سعی و خطا برای هر کدام از توابع مذکور برای هر دو حالت جریان غیرنرمال و نرمال و به ازای ۱ تا ۱۵ ورودی بدست آمد. در نهایت عملکرد هر یک از سه تابع کرنل به ازای هر کدام از تعداد ورودی‌های مذکور توسط شاخص‌های آماری مورد ارزیابی قرار گرفت. در جدول ۲ بهترین مدل از بین ۴۵ مدل ارزیابی شده برای هر ایستگاه آب‌سنجی، به همراه نوع تابع کرنل مربوط به آن نشان داده شده است. مطابق با این جدول مدل SVM به ازای ۱۳ و ۱۱ جریان ورودی مشاهداتی و تابع کرنل چند جمله‌ای درجه ۴ به ترتیب در ایستگاه‌های صفاخانه و سنته و همچنین به‌ازای ۳ متغیر ورودی جریان مشاهداتی و تابع کرنل پایه شعاعی در ایستگاه پل آنیان. نتایج بهتری را نسبت به ۴۵ مدل دیگر ارزیابی شده در هر ایستگاه بدست آورده است.

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

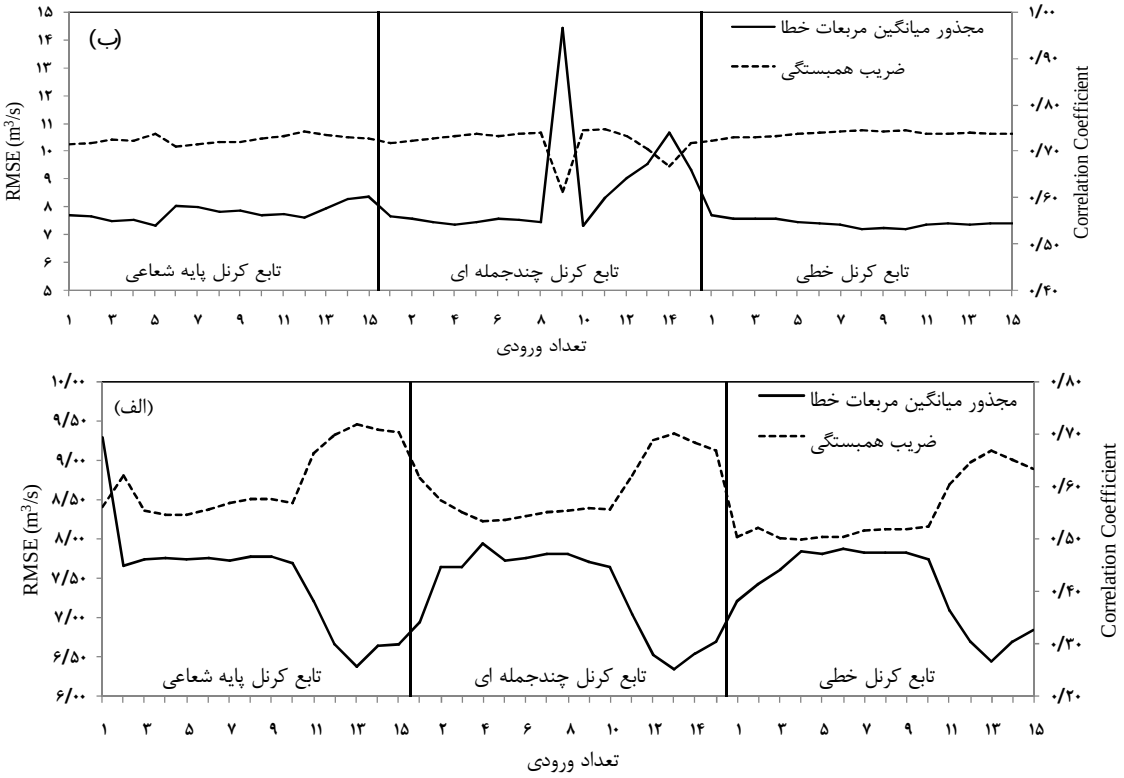
[DOI: 10.29252/jwmr.8.16.22]

جدول ۲- عملکرد مدل SVM در تخمین جریان ایستگاه‌های آب‌سنجی به ازای جریان مشاهداتی و نرمال شده

Table 2. Performance of the SVM model in estimation of hydrometrics stations flow for the observation and normalized flow

نوع داده‌ها	ایستگاه	بهترین تابع کرنل	بهترین تعداد ورودی	دوره واسنجی			دوره صحت سنجی		
				CC	RMSE (m^3/s)	NS	CC	RMSE (m^3/s)	NS
جریان مشاهداتی	صفاخانه	چند جمله ای	۱۳	۰/۸۳	۸/۱۰	۰/۶۹	۰/۷۰	۶/۳۴	۰/۳۹
	سنته	چند جمله ای	۱۱	۰/۸۰	۸/۹۶	۰/۶۴	۰/۷۸	۷/۰۲	۰/۵۸
	پل آنیان	پایه شعاعی	۳	۰/۷۷	۱۷/۳۴	۰/۵۹	۰/۷۹	۲۴/۳۱	۰/۶۲
جریان نرمال	صفاخانه	خطی	۱۰	۰/۷۲	۱۰/۱۰	۰/۰۶	۰/۷۵	۷/۲۱	۰/۴۴
	سنته	چند جمله ای	۱۴	۰/۸۱	۸/۸۲	۰/۵۱	۰/۸۶	۵/۸۸	۰/۷۳
	پل آنیان	پایه شعاعی	۷	۰/۸۱	۱۶/۳۴	۰/۵۹	۰/۷۸	۲۵/۵۳	۰/۱۳

با توجه به نتایج جدول ۲ مدل SVM به ازای ۱۰، ۱۴ و ۷ جریان ورودی نرمال شده به ترتیب در ایستگاه‌های آب‌سنجی صفاخانه، سنته و پل آنیان نتایج بهتری را نسبت ۴۵ مدل با ورودی نرمال بررسی شده در هر ایستگاه آب‌سنجی بدست آورده است. نتایج نشان داد که مدل SVM در شبیه‌سازی جریان ایستگاه صفاخانه و پل آنیان با داده‌های واقعی عملکرد بهتری را نسبت به شبیه‌سازی جریان با داده‌های دارای توزیع نرمال داشته است و استفاده از داده‌های نرمال شده فقط در ایستگاه سنته باعث بهبود نتایج شبیه‌سازی گردیده است. به‌عنوان نمونه در شکل ۶ برای ایستگاه صفاخانه تغییرات مقادیر CC و RMSE به ازای سه تابع کرنل در دوره صحت‌سنجی نشان داده شده است. است. RMSE در اکثر موارد به ترتیب روند صعودی و نزولی داشته نرمال شده و سه تابع کرنل برای ایستگاه صفاخانه نشان می‌دهد. با توجه به این شکل در توابع کرنل پایه شعاعی و خطی تغییرات CC و RMSE به ازای افزایش ورودی بسیار اندک بوده است که نشان دهنده حساسیت کمتر این مدل‌ها به تعداد ورودی‌های جریان نرمال شده می‌باشد.



شکل ۶- روند تغییرات CC و RMSE مربوط به شبیه‌سازی جریان با استفاده از مدل SVM در دوره صحت‌سنجی به ازای داده‌های ورودی جریان مشاهداتی(الف) و نرمال شده(ب) در ایستگاه صفاخانه

Figure 6. The trend of CC and RMSE variations in flow simulation using the SVM model during the verification period for the input data of observation flow (A) and normalized flow (B) at the Safahaneh station

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

شبیه‌سازی جریان با مدل ANN

در این مطالعه از شبکه عصبی پرسپترون چند لایه (MLP) جهت مدل سازی جریان ماهانه استفاده گردید. جهت تعیین مدل شبیه‌ساز مطلوب، جریان مشاهداتی و نرمال شده با تاخیرهای ۱ تا ۱۵ به عنوان ورودی شبکه عصبی مصنوعی در نظر گرفته شد. شبکه بکار گرفته شده دارای یک لایه پنهان بوده که در آن تعداد نرون‌ها از ۱ تا ۲۰ تغییر می‌یابد. در جدول ۳ بهترین مدل از بین ۳۰۰ مدل ارزیابی شده برای هر ایستگاه، به همراه بهترین تعداد نرون آنها نشان داده شده است.

جدول ۳- عملکرد مدل MLP در تخمین جریان ایستگاه‌های سنته، صفاخانه و پل آنیان به ازای جریان مشاهداتی و نرمال شده
Table 3. Performance of the MLP model in estimation of the Santeh, Safakhaneh and Polanian stations flow using the observation and normalized flow

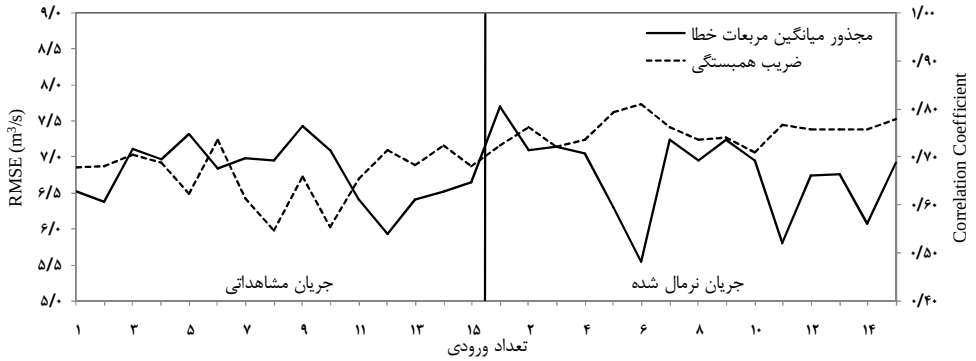
نوع داده ورودی	ایستگاه آب‌سنجی	بهترین تعداد ورودی	بهترین تعداد نرون	دوره واسنجی			دوره صحت سنجی		
				CC	RMSE (m^3/s)	NS	CC	RMSE (m^3/s)	NS
مشاهداتی	صفاخانه	۱۲	۴	۰/۸۰	۸/۹۰	۰/۶۳	۰/۷۱	۵/۹۳	۰/۴۷
	سنته	۱۱	۱۱	۰/۸۰	۸/۹۴	۰/۶۴	۰/۸۰	۶/۵۸	۰/۶۳
	پل آنیان	۳	۶	۰/۷۸	۱۷/۰۶	۰/۶۰	۰/۸۲	۲۲/۹۰	۰/۶۷
نرمال شده	صفاخانه	۶	۱۹	۰/۷۳	۱۰/۲۳	-۰/۲۷	۰/۸۱	۵/۵۵	۰/۶۰
	سنته	۶	۶	۰/۸۱	۸/۸۴	۰/۵۳	۰/۸۹	۵/۳۲	۰/۷۹
	پل آنیان	۱۴	۱۴	۰/۸۷	۱۳/۴۲	۰/۷۰	۰/۸۷	۲۰/۹۳	۰/۴۸

مطابق با جدول ۳ در ایستگاه صفاخانه برای جریان مشاهداتی مدل MLP با ۱۲ متغیر ورودی و ۴ نرون در لایه پنهان ((MLP(۴،۱۲) با مقادیر شاخص‌های CC و RMSE به ترتیب برابر با ۰/۷۱ و ۵/۹۳ مترمکعب بر ثانیه در دوره صحت‌سنجی عملکرد بهتری نسبت به سایر مدل‌های MLP دارد. برای جریان نرمال شده، مدل MLP با ۶ ورودی و ۱۹ نرون ((MLP(N۱۹،۶) با مقادیر شاخص‌های CC و RMSE به ترتیب برابر با ۰/۸۱ و ۵/۵۵ مترمکعب بر ثانیه در دوره صحت‌سنجی نسبت به سایر مدل‌های MLP بهترین کارایی را در تخمین جریان دارا می‌باشد. نتایج شبیه‌سازی جریان با مدل MLP و داده‌های نرمال شده نشان داد که استفاده از این نوع داده‌ها می‌تواند نقش بسزایی در افزایش CC و کاهش RMSE در کلیه ایستگاه‌های آب‌سنجی داشته باشد(جدول ۳ و ۴). در جدول ۴ مقدار افزایش CC و کاهش RMSE در شبیه‌سازی جریان با استفاده از داده‌های نرمال شده نسبت به مشاهداتی برای ایستگاه‌های آب‌سنجی مورد مطالعه نشان داده شده است.

جدول ۴- مقدار کاهش RMSE و افزایش CC به ازای جریان نرمال شده در شبیه‌سازی جریان دوره صحت‌سنجی
Table 4. The amount of RMSE reduction and CC increase per normalized flow in the simulation flow of verification period

شاخص آماری	ایستگاه آب‌سنجی		
	صفاخانه	سنته	پل آنیان
RMSE	۰/۰۶	۰/۱۹	۰/۰۹
CC	۰/۱۴	۰/۱۱	۰/۰۶

همچنین به عنوان نمونه در شکل ۷ عملکرد ۱۵ مدل MLP بکار گرفته شده جهت شبیه‌سازی جریان ماهانه ایستگاه صفاخانه به ازای بهترین تعداد نرون در لایه پنهان برای دو حالت داده‌های مشاهداتی و نرمال شده در دوره صحت‌سنجی آورده شده است.



شکل ۷- روند تغییرات CC و RMSE در دوره صحت سنجی به ازای ورودی‌های مختلف جریان مشاهداتی و نرمال شده
Figure 7. The trend of CC and RMSE variations in during the verification period for the various inputs of observation and normalized flow

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

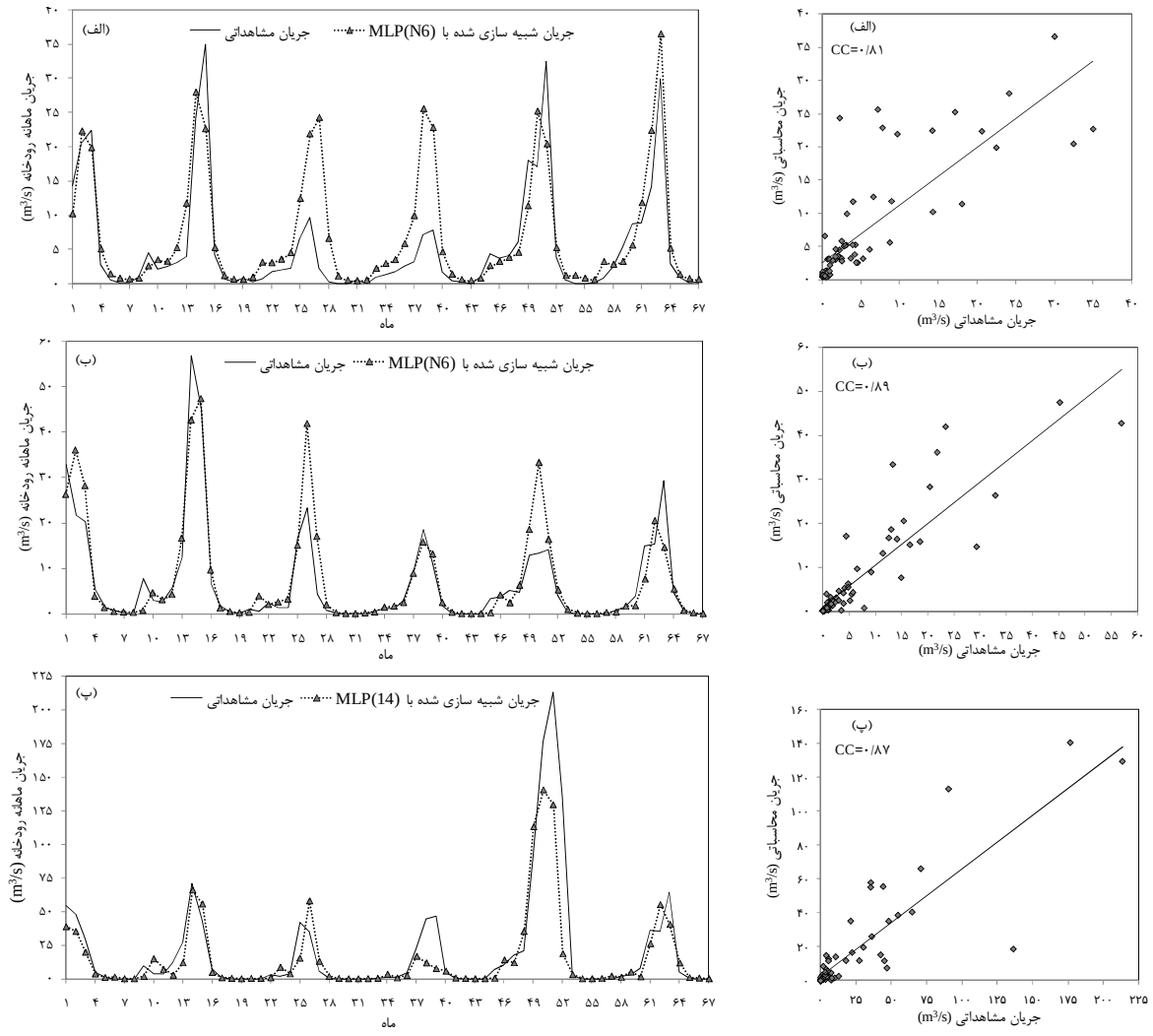
مقایسه عملکرد مدل‌های ANN و SVM

براساس نتایج بخش‌های قبل مدل‌های MLP(N۶)، MLP(N۶) و MLP(N۱۴) نسبت به سایر مدل‌های ارزیابی شده MLP و SVM، کم خطاترین شبیه‌سازی جریان را به ترتیب در ایستگاه‌های صفاخانه، سنته و پل آنیان دارا می‌باشند. در جدول ۵ خلاصه‌ای از عملکرد بهترین مدل‌های بکار رفته در هر ایستگاه آورده شده است. همچنین در شکل ۸ سری زمانی جریان ماهانه مشاهداتی و شبیه‌سازی شده توسط مدل‌های منتخب برای هر یک از ایستگاه‌های آب‌سنجی در دوره صحت‌سنجی نشان شده است.

جدول ۵- خلاصه عملکرد بهترین مدل شبیه‌سازی جریان در ایستگاه‌های آبسنجی مورد مطالعه

Table 5. Summary of performance of the best flow simulation model at the studied hydrometric stations

رودخانه	ایستگاه	نوع داده‌های ورودی	مدل	دوره واسنجی			دوره صحت سنجی		
				CC	RMSE (m^3/s)	NS	CC	RMSE (m^3/s)	NS
ساروق چای	صفاخانه	جریان نرمال شده	MLP(N6)	۰/۷۳	۱۰/۲۳	-۰/۲۷	۰/۸۱	۵/۵۵	۰/۶۰
خرخره چای	سنته	جریان نرمال شده	MLP(N6)	۰/۸۱	۸/۸۴	۰/۵۳	۰/۸۹	۵/۳۲	۰/۷۹
جیغاتو چای	پل آنیان	جریان نرمال شده	MLP(N14)	۰/۸۷	۱۳/۴۲	۰/۷۰	۰/۸۷	۲۰/۹۳	۰/۴۸



شکل ۸- جریان مشاهداتی و شبیه‌سازی شده مدل‌های منتخب در ایستگاه‌های صفاخانه (الف)، سنته (ب) و پل آنیان (پ)

Figure 8. Observation and simulated flow of selected models at the Safakhaneh (a), Santeh (b) and Polanian (c) stations

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

در تحقیق حاضر سعی گردید عملکرد مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان در شبیه‌سازی جریان ماهانه حوزه زرینه‌رود در سه ایستگاه صفاخانه، سنته و پل آنیان و همچنین تاثیر نرمال و غیرنرمال بودن جریان‌های ورودی بر عملکرد مدل‌ها مورد بررسی قرار گیرد. در این راستا ابتدا توزیع احتمالاتی داده‌های جریان مشاهداتی هر ایستگاه طبق روابط ۱ تا ۹ به جریان با توزیع نرمال تبدیل گردیده است. در ادامه جهت لحاظ نمودن تاثیر جریان با تاخیرهای ماهانه مختلف و همچنین ساختار بکار رفته در مدل‌ها بر عملکرد آنها در شبیه‌سازی جریان ماهانه، کارایی مدل‌های شبکه عصبی مصنوعی و SVM با تابع‌های کرنل مختلف به ازای تاخیرهای ۱ تا ۱۵ ماهه جریان رودخانه‌ها مورد ارزیابی قرار گرفت و نتایج زیر بدست آمد.

نتایج نشان داد مدل SVM به ازای ۱۳ و ۳ جریان ورودی مشاهداتی (غیر نرمال) نتایج بهتری را در شبیه‌سازی جریان دوره صحت‌سنجی ایستگاه‌های آب‌سنجی صفاخانه و پل‌آنیان با مقادیر شاخص‌های آماری CC و RMSE به ترتیب برابر با ۰/۷۰، ۶/۳۴(m^3/sec) و ۰/۷۹ ۲۴/۳۱(m^3/sec) داشته است. استفاده از داده‌های نرمال شده جریان نتایج بهتری در ایستگاه سنته با مقادیر شاخص‌های آماری CC و RMSE برابر با ۰/۸۶، ۵/۸۸(m^3/sec) داشته است، در واقع استفاده از داده‌های نرمال شده در مقایسه با حالت غیر نرمال باعث افزایش ۱۴ درصدی CC و کاهش ۶ درصدی RMSE در دوره صحت‌سنجی گردیده است.

در شبیه‌سازی جریان بوسیله شبکه عصبی مصنوعی، مدل‌های MLP(12)، MLP(11) و MLP(3) با داده ورودی مشاهداتی (غیرنرمال) به ترتیب جهت شبیه‌سازی جریان در ایستگاه‌های آب‌سنجی صفاخانه، سنته و پل آنیان انتخاب گردیدند که مقادیر شاخص‌های آماری CC و RMSE این ایستگاه‌ها در دوره صحت‌سنجی به ترتیب برابر با ۰/۷۱، ۵/۹۳ (m^3/sec)، ۰/۸۰، ۶/۵۸ (m^3/sec) و ۰/۸۲ ۲۲/۹۰ (m^3/sec) می‌باشند. مدل‌های MLP(N6)، MLP(N6) و MLP(N14) نیز جهت پیش‌بینی جریان ایستگاه‌های آب‌سنجی صفاخانه، سنته و پل‌آنیان به ازای داده‌های ورودی نرمال شده انتخاب گشتند؛ البته استفاده از داده‌های نرمال شده در کلیه ایستگاه‌های آب‌سنجی باعث افزایش CC و کاهش RMSE گردیده است که درصد این افزایش و کاهش در ایستگاه‌های آب‌سنجی صفاخانه، سنته و پل آنیان به ترتیب برابر با ۰/۱۴، ۰/۰۶، ۰/۱۱، ۰/۱۹، ۰/۰۶ و ۰/۰۹ می‌باشد.

با توجه به نتایج دوره صحت‌سنجی مدل‌های منتخب MLP (جدول ۳) می‌توان دریافت که عملکرد مدل MLP در شبیه‌سازی جریان نسبت به مدل‌های منتخب SVM (جدول ۲) به ازای داده ورودی مشاهداتی در هر یک از سه ایستگاه مذکور بهتر بوده است.

تابع PACF هر یک از سه ایستگاه آب‌سنجی، تاخیر اول جریان مشاهداتی و نرمال شده را موثرترین ورودی در شبیه‌سازی جریان معرفی می‌کند که در شبیه‌سازی‌های انجام شده و همچنین شکل‌های ۶ و ۷ نیز این مورد مشهود است. در حالت کلی مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان با یک ورودی غیرنرمال بین ۸۰ تا ۹۹ درصد و با یک ورودی نرمال نیز ۷۲ تا ۹۴ درصد مقدار خطای مدل منتخب مربوط به ایستگاه آب‌سنجی خود را پوشش می‌دهند.

با توجه به نتیجه‌گیری‌های فوق استفاده از مدل ANN به جای مدل SVM می‌تواند نتایج بهتری را در شبیه‌سازی جریان رودخانه‌ها به همراه داشته باشد. همچنین استفاده از ورودی‌های نرمال شده به جای ورودی‌های غیر نرمال در مدل‌های شبکه عصبی مصنوعی بسیار موثر بوده و باعث افزایش دقت و کاهش خطای شبیه‌سازی می‌گردد. استفاده از داده‌های نرمال شده همیشه باعث بهبود نتایج بدست آمده از مدل SVM نمی‌گردد، لذا برای دستیابی به بهترین نوع ورودی (جریان نرمال و غیر نرمال) در مدل SVM بایستی از سعی و خطا استفاده گردد.

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

منابع

1. Adamowski, J. 2013. Using support vector regression to predict direct runoff, base flow and total flow in a mountainous watershed whit limited data in Uttaranchal, India. Annals of Warsaw University of Life Sciences-SGGW. Land Reclamation, 45(1): 71-83.
2. Alvisi, S., G. Mascellani, M. Franchini and A. Bardossy. 2006. Water level forecasting through fuzzy logic and artificial neural network approaches. Hydrology and Earth System Sciences Discussions, 10(1): 1-17.
3. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. 2000. Artificial neural networks in hydrology, I: preliminary concepts. Journal of Hydrologic Engineering, 5(2): 115-123.
4. Cortes. C. and V. Vapnik. 1995. Support vector networks. Machine Learning, 20: 273-297.
5. Coulibaly, P., F. Anctil and B. Bobée. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. Journal of Hydrology, 230: 244-257.
6. Dibike, Y.B., S. Velikov, D. Solomatine and M.B. Abbort. 2001. Model induction with of support vector machines: introduction and applications. Journal of Computing in Civil Engineering, 15: 208-216.
7. Guo, B., S.R. Gunn, R.I. Damper and J.D. Nelson. 2008. Customizing kernel functions for svm-based hyper spectral image classification. IEEE Transactions on Image Processing, 17(4): 622-629.
8. Huang, W. and B.X. Chan-Hilton. 2004. Forecasting flows in Apalachicola River using neural networks. Hydrological Processes, 18(13): 2545-2564.
9. Kakaei Lafadani, E., A. Moghaddam Nia, A. Ahmadi, M. Jajarmizadeh and M. Ghafari. 2013. Stream flow simulation using SVM, ANFIS and NAM models (A case study). Caspian Journal of Applied Sciences Research, 2(4): 86-93.
10. Kavzoglu, T. and I. Colkesen. 2009. A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 11(5): 352-359.
11. Kouchakzadeh, M. and A. Bahmani. 2006. Assessment of artificial neural networks revenue in reducing required parameters for estimation of reference evapotranspiration. Journal of Agricultural Sciences, 11(4): 87-97.
12. Lin, J.Y., C.T. Cheng and K.W. Chau. 2006. Using support vector machines for long-term discharge prediction. Hydrological Sciences Journal, 51(4): 599-612.
13. Liu, G.Q. 2011. Comparison of regression and ARIMA models with neural network models to forecast the daily streamflow of White Clay Creek. PhD Thesis, University of Delaware, Newark, USA. 517 pp.
14. Misra, D., T. Oommen, A. Agarwal, S.K. Mishra and A.M. Thompson. 2009. Application and analysis of support vector machine based simulation for runoff and sediment yield. Biosystems Engineering, 103(9): 527-535.
15. Nayak, P.C., K.P. Sudheer, D.M. Rangan and K.S. Ramasastri. 2004. A neuro-fuzzy computing technique for modeling hydrological time series. Journal of Hydrology, 291(1): 52-66.
16. Nouri, R., A. Khakpour, M. Dehghani and A. Farrokhnia. 2010. Monthly river flow prediction using support vector machine based on principal components analysis. Journal of Water and Wastewater, 22(1): 118-129 (In Persian).
17. Safavi, H.R. 2009. Engineering hydrology. 4nd edn, Arkan Danesh publisher, Esfahan, Iran, 706 pp. (In Persian)
18. Samsudin, R., P. Saad and A. Shabri. 2011. River flow time series using least squares support vector machines. Hydrology and Earth System Sciences, 15(6): 1835-1852.
19. Yang, K.H., G.L. Shan and L.L. Zhao. 2006. Correlation coefficient method for support vector machine input samples. International Conference on Machine Learning and Cybernetics, 2857-2861 pp., Dalian, China.
20. Yoon, H., S.C. Jun, Y. Hyun, G.O. Bae and K.K. Lee. 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. Journal of Hydrology, 396(1–2): 128-138.
21. Zealand, C.M., D.H. Burn and S.p. Simonovic. 1999. Short term stream flow forecasting using artificial neural networks. Journal of Hydrology, 214(1–4): 32-48.

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]

Assessment of Normalization of Monthly Runoff Probabilistic Distribution impact on SVM and ANN Models Performance in Monthly River Flows Simulation (A Case Study: ZarrinehRud River Basin)

Mohammad Isazadeh[1], Hojat Ahmadzadeh[2] and Mohammad Ali Ghorbani[3]

1- M.Sc. Student of Water Resources Engineering, University of Tabriz,
(Corresponding author: mohammadisazade@gmail.com)
2 and 3- Ph.D. Student and of Associate Professor, Water Resources Engineering, University of Tabriz
Received: April 21, 2015 Accepted: May 7, 2016

Abstract

Accurate estimation of river flows is one of the fundamental activities in water resources management of river basins. Artificial neural network (ANN) and support vector machine (SVM) are the most important data mining models that can be considered for this purpose. Due to the data-based attribute of these models, probability distribution of data may have a considerable effects on their performance in river flow simulation. In order to, Zarrineh Rud River basin was selected as a study area and the investigations were done for three hydrometric stations located in this basin. In this regard, first monthly runoff probability distribution of stations were studies based on Shapiro- Wilk test and then normalization of data distribution were done. Then the performance of ANN and SVM models in monthly river flow simulation of three stations was evaluated for initial observed and normal data. Based on the results of this study, the values of 0.71, 5.93 (m^3/sec), 0.80, 6.58 (m^3/sec) and 0.82, 22.9 (m^3/sec) were obtained for correlation coefficient (CC) and root mean square errors (RMSE) indicators in the ANN model for Safakhaneh, Santeh and Polanian stations respectively in the testing period. In the SVM model, the values of 0.70, 6.34 (m^3/sec), 0.78, 7.02 (m^3/sec) and 0.79, 24.31 (m^3/sec) were obtained for these indicators in the mentioned stations respectively. The results showed that in river flow simulation by ANN model values of CC increase 6%, 14% and 11% and RMSE values decrease 9%, 19% and 6% for Polanian, Santeh and Safakhaneh stations respectively in the testing period due to normalization of data probability distribution. For SVM model, due to normalization of data probability, CC value increases 10% and RMSE value decrease 16% only for Santeh station. Also the results showed that the ANN model with normal input data has high performance in estimation of monthly river flow compared to the SVM model in each of the three hydrometric stations.

Keywords: ANN, Monthly River Flow, Normal Probabilistic Distribution, SVM, Zarrineh Rud Basin

[Downloaded from jwmr.sanru.ac.ir on 2026-11-08]

[DOI: 10.29252/jwmr.8.16.22]