



Application of the Hybrid Model of Support Vector Machine-Algorithm Artificial Flora in Estimating the Daily Flow of Rivers (Case Study: Dez Basin)

R. Dehghani¹, H. Torabi Poudeh^{2*}, H. Younesi³,
and B. Shahinejad⁴

Abstract

In this study, the hybrid support vector machine-artificial flora algorithm method was developed and the results were compared with those of the support vector machine-wavelet model. The case study of Dez catchment area was used in order to estimate the flow rate of the rivers employing the daily discharge statistics from hydrometric stations located upstream of the dam in the statistical period from 2008 to 2018. The criteria of coefficient of determination, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Nash-Sutcliffe coefficient were used to evaluate and compare the models. The results showed that the combined structures provided acceptable results in river flow modeling. Also, comparison of the models based on the evaluation criteria and Taylor's diagram showed that the proposed hybrid support vector machine - artificial flora with the correlation coefficient ($R^2=0.933-0.985$), root-mean-square error (RMSE= 0.008-0.088 m³/s), mean absolute error (MAE= 0.004-0.040 m³/s), and Nash-Satcliffe coefficient (NS=0.951-0.995) performed better in estimating the daily flow rates of rivers.

Keywords: Artificial Flora Algorithm, Prediction, Dez Basin, Support Vector Machine.

Received: June 16, 2020

Accepted: June 27, 2020

بررسی عملکرد مدل هیبریدی ماشین بردار پشتیبان الگوریتم گیاهان مصنوعی در تخمین جریان دبی روزانه رودخانه‌ها (مطالعه موردی: حوضه دز)

رضا دهقانی^۱، حسن ترابی پوده^{۲*}، حجت‌اله یونسی^۳
و بابک شاه‌نژاد^۴

چکیده

در این پژوهش یک روش پیشنهادی هیبریدی تحت عنوان ماشین بردار پشتیبان - الگوریتم گیاهان مصنوعی مورد بررسی قرار گرفته و نتایج آن با مدل ماشین بردار پشتیبان - موجک مقایسه گردید. به منظور برآورد دبی رودخانه‌های حوضه آبریز دز، از آمار آبدی روزانه ایستگاه‌های هیدرومتری واقع در بالادست سد طی دوره آماری (۱۳۸۷-۱۳۹۷) استفاده شد. معیارهای ضریب تبیین، ریشه میانگین مربعات خطا، میانگین قدر مطلق خطا و ضریب نش ساتکلیف برای ارزیابی و مقایسه مدل‌ها مورد استفاده قرار گرفت. نتایج نشان داد ساختارهای ترکیبی نتایج قابل قبولی در مدل‌سازی دبی رودخانه ارائه می‌نمایند. مدل هیبریدی پیشنهادی ماشین بردار پشتیبان - گیاهان مصنوعی با ضریب همبستگی ($R^2=0.933-0.985$)، ریشه میانگین مربعات خطا (RMSE= 0.008-0.088 m³/s)، میانگین قدر مطلق خطا (MAE= 0.004-0.040 m³/s) و ضریب نش ساتکلیف (NS=0.951-0.995) عملکرد بهتری در تخمین جریان داشته و می‌تواند در زمینه پیش‌بینی دبی روزانه رودخانه‌ها مفید باشد.

کلمات کلیدی: الگوریتم گیاهان مصنوعی، تخمین، حوضه دز، ماشین بردار پشتیبان.

تاریخ دریافت مقاله: ۹۹/۳/۲۷

تاریخ پذیرش مقاله: ۹۹/۴/۷

1- Ph.D. Student of Water Structure, Faculty of Agriculture, Lorestan University, Iran.

2- Associate Professor, Department of Water Engineering, Lorestan University. Email: torabi.ha@lu.ac.ir

3- Assistant Professor, Department of Water Engineering, Lorestan University, Iran.

4- Assistant Professor, Department of Water Engineering, Lorestan University, Iran.

*- Corresponding Author

۱- دانشجوی دکتری سازه‌های آبی، دانشکده کشاورزی، دانشگاه لرستان.

۲- دانشیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان.

۳- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان.

۴- استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه لرستان.

*- نویسنده مسئول

بحث و مناظره (Discussion) در مورد این مقاله تا پایان پائیز ۱۳۹۹ امکانپذیر است.

(2016) Sedighi et al. جهت پیش‌بینی فرآیند بارش رواناب از مدل‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان در حوضه آبریز رودک واقع در شمال شرقی تهران با استفاده از اطلاعات تعداد ۹۲ سنجه مودیس طی سال آماری ۲۰۰۵-۲۰۰۳ نشان دادند مدل ماشین بردار پشتیبان توانایی قابل قبولی در تخمین رواناب دارد. در پژوهشی دیگر (2016) Ghorbani et al. جهت پیش‌بینی جریان روزانه رودخانه سایپرس (Cypress) تگزاس از مدل‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی استفاده نمودند و همچنین معیارهای ضریب همبستگی و ریشه میانگین مربعات خطا را جهت ارزیابی مدل‌ها به کار بردند و دریافتند مدل ماشین بردار پشتیبان عملکردی مناسبی جهت پیش‌بینی جریان رودخانه داشته و نسبت به مدل شبکه عصبی مصنوعی از دقت مطلوبی برخوردار است. (Shahinejad and Dehghani (2018) جهت تخمین میزان اکسیژن محلول در آب رودخانه کامبرلند واقع در ایالات متحده از مدل‌های شبکه عصبی موجک، ماشین بردار پشتیبان و برنامه‌ریزی بیان ژن استفاده کردند. نتایج نشان داد مدل شبکه عصبی موجک از عملکرد بهتری نسبت به سایر مدل‌ها برخوردار است. (2018) Bagherpour et al. عملکرد مدل‌های نروفازی و ماشین بردار پشتیبان جهت شبیه‌سازی فرآیند بارش- رواناب در رودخانه گرگان‌رود مورد بررسی قرار دادند. نتایج نشان داد مدل ماشین بردار پشتیبان دقت بیشتری نسبت به مدل نروفازی دارد. (2019) Samadianfard et al. عملکرد مدل هیبریدی پیشنهادی رگرسیون ماشین بردار پشتیبان- الگوریتم مگس نسبت به مدل درخت تصمیم در تخمین دبی رودخانه ونیار واقع در ایران مورد بررسی قرار دادند. نتایج نشان داد که مدل هیبریدی پیشنهادی رگرسیون ماشین بردار پشتیبان- الگوریتم دقت مناسب‌تری نسبت به مدل درخت تصمیم دارد. نتایج حاصل از پژوهش (2019) Adnan et al. که جهت پیش‌بینی دبی ماهانه رودخانه سوات (Swat) واقع در پاکستان از مدل‌های ماشین بردار پشتیبان و درخت تصمیم استفاده نمودند نشان داد مدل ماشین بردار پشتیبان کارایی مناسبی در پیش‌بینی جریان رودخانه‌ها دارد. (2020) Rajaei et al. از ترکیب تبدیل موجک با مدل‌های ماشین بردار پشتیبان، نروفازی، شبکه عصبی مصنوعی و برنامه‌ریزی ژنتیک جهت پیش‌بینی جریان روزانه رودخانه دان (Dunbe) واقع در صربستان استفاده نمودند نتایج نشان داد مدل ترکیبی ماشین بردار پشتیبان- موجک خطای کمتری نسبت به سایر مدل‌های ترکیبی مورد بررسی دارد. (2020) Alizadeh et al. پژوهشی به بررسی مدل هیبریدی ماشین بردار پشتیبان- موجک جهت پیش‌بینی جریان روزانه رودخانه سوریس (Souris) واقع در شمال ایالات متحده پرداختند و نتایج نشان داد مدل هیبریدی ماشین بردار پشتیبان- موجک کارایی و دقت مناسبی در پیش‌بینی جریان روزانه رودخانه از خود نشان داده

مهم‌ترین مسائل جهت مدیریت سیلاب و جلوگیری از صدمات اقتصادی و جانی ناشی از آن، برآورد دقیق جریان رودخانه‌ها است. بر این اساس دستیابی به روش‌های مطمئن پیش‌بینی جریان رودخانه‌ها به منظور برنامه‌ریزی در بهره‌برداری به موقع از منابع آب از اهمیت روزافزونی برخوردار است (Zhang et al., 1998). همچنین پیش‌بینی دقیق جریان رودخانه‌ها می‌تواند نقش به‌سزایی در برنامه‌ریزی و مدیریت منابع آب داشته باشد؛ اما فاکتورها و عوامل مختلفی بر روی این پدیده تأثیرگذارند که تحلیل آن‌ها را مشکل می‌سازد؛ لذا انتخاب مدلی که بتواند با استفاده از عوامل تأثیرگذار، جریان رودخانه را به طور قابل قبولی تخمین زند امری ضروری به نظر می‌رسد (Edossa and Edossa and (2005; Mohammadi et al., 2005). امروزه سامانه‌های هوشمند به طور گسترده جهت برآورد پدیده‌های غیرخطی مورد استفاده قرار می‌گیرند. یکی از روش‌هایی که در زمینه هیدرولوژی مورد توجه قرار گرفته است مدل ماشین بردار پشتیبان است این مدل دارای عملکرد مناسبی بوده و به منظور افزایش دقت و کاهش میزان خطای این مدل در سال‌های اخیر از الگوریتم‌های بهینه‌ساز جهت بهینه نمودن پارامترهای تنظیم این مدل که سبب افزایش عملکرد آن می‌شود، استفاده شده است. در الگوریتم‌های فراکاوشی به دلیل اضافه شدن سرعت‌هایی با مقادیر تصادفی به متغیرهای مسأله ممکن است مقادیر آن بطور ناخواسته به خارج از دامنه تعریفشان منتقل شود از سوی دیگر در سایر الگوریتم‌ها به دلیل استفاده از مقادیر گسسته برای متغیرهای مسأله جواب‌های به دست آمده در تمام تکرارها در دامنه مسأله قرار می‌گیرد و دستیابی به جواب بهینه سراسری بعضی مواقع طولانی می‌شود که سبب می‌گردد مسأله در جواب‌های بهینه محلی قرار گیرد (Chen and Zhu, 2008). از این رو الگوریتم گیاهان مصنوعی که ترکیبی از بهینه‌سازی پیوسته و گسسته است. در مسایل با دامنه بزرگ، زمان دستیابی به جواب بهینه جستجوی سراسری کاهش یافته و از جواب‌های بهینه محلی خارج می‌شود که این امر سبب توانایی قابل قبول این الگوریتم در حل مسائل غیر خطی با ابعاد بزرگ و سرعت مناسب در همگرایی به جواب بهینه می‌شود. از این رو در این پژوهش از ترکیب الگوریتم گیاهان مصنوعی با ماشین بردار پشتیبان به عنوان یک روش پیشنهادی استفاده شد. در سال‌های اخیر استفاده از مدل‌های هوشمند و هیبریدی در مطالعات پیش‌بینی جریان رودخانه‌ها مورد توجه محققین قرار گرفته است که از جمله می‌توان به موارد ذیل اشاره نمود: (2014) Hung et al. جریان ماهانه رودخانه ایستگاه هواکسی (Huaxian) چین را با استفاده از ماشین بردار پشتیبان پیش‌بینی نمودند و نتایج نشان داد مدل ماشین بردار پشتیبان دقت بالایی در پیش‌بینی جریان ماهانه رودخانه دارد.

تیره و سبزه تشکیل شده است. پیوستن آبراهه‌های متعدد و از جمله رودخانه ازنا در منطقه الیگودرز، رودخانه ماربره را پدید می‌آورند و این رودخانه در جهت غرب به طرف شهر دورود در استان لرستان جریان می‌یابد. حوضه آبریز سد دز در جنوب غربی ایران در منطقه نیمه خشک کوهستانی و بالادست حوضه کارون بین طول‌های " ۴۸° ۹' ۱۵ " تا " ۳۷° ۱۸' ۵۰ " شرقی و عرض‌های " ۳۱° ۳۵' ۵۱ " تا " ۳۴° ۷' ۴۶ " شمالی قرار گرفته است. که در شکل ۱ حوضه آبریز سد دز، در ایستگاههای منتخب که فاقد داده‌های مفقود شده و همگن می‌باشند از شرکت آب منطقه‌ای لرستان و سازمان آب و برق خوزستان در دسترس بود، نشان داده شده است.

۲-۲- ماشین بردار پشتیبان

ماشین بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده کرده و منجر به یک جواب بهینه کلی می‌گردد (Vapnik, 1995). 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 (Vapnik, 1998).

$$f(x) = W^T \cdot \phi(x) + b \quad (1)$$

$$y = f(x) + \text{noise} \quad (2)$$

است. Hussein and Ahmedkhan (2020) در پژوهشی با استفاده از مدل‌های ماشین بردار پشتیبان، شبکه عصبی مصنوعی و جنگل تصادفی به پیش‌بینی جریان رودخانه هانزا واقع در پاکستان پرداختند. نتایج نشان داد مدل جنگل تصادفی نسبت به مدل‌های مورد بررسی از عملکرد بهتری برخوردار است. در مجموع با توجه به پژوهش‌های انجام‌شده و ذکر این نکته که رودخانه‌های حوضه آبریز دز از مهمترین حوضه‌های آبریز کشور ایران و مهم‌ترین منبع تأمین‌کننده آب بخش‌های مختلف و نواحی مجاور خود از لحاظ کشاورزی و شرب می‌باشد و همچنین کاهش جریان رودخانه‌های این حوضه آبریز مشکلات زیادی در حوضه آبریز ایجاد کرده است، بنابراین، اهمیت شبیه‌سازی جریان رودخانه این حوضه و اقدامات مدیریتی جهت مدیریت آب آن بیش‌ازپیش ضروری است. لذا هدف این پژوهش پیش‌بینی جریان روزانه رودخانه‌های حوضه آبریز دز با استفاده از مدل پیشنهادی هیبریدی ماشین بردار پشتیبان- گیاهان مصنوعی و مقایسه نتایج آن با مدل هیبریدی ماشین بردار پشتیبان- موجک می‌باشد.

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

حوضه آبریز رودخانه دز به‌عنوان یک حوضه درجه ۳، زیرمجموعه‌ای از حوضه کارون بزرگ در کشور ایران محسوب می‌شود و در تقسیم‌بندی بزرگ‌تر در زیرمجموعه حوضه خلیج فارس و دریای عمان قرار می‌گیرد. سزار و بختیاری دوشاخه اصلی رودخانه دز هستند. رود سزار در شمالی‌ترین قسمت حوضه دز جریان دارد و از سه شاخه ماربره،

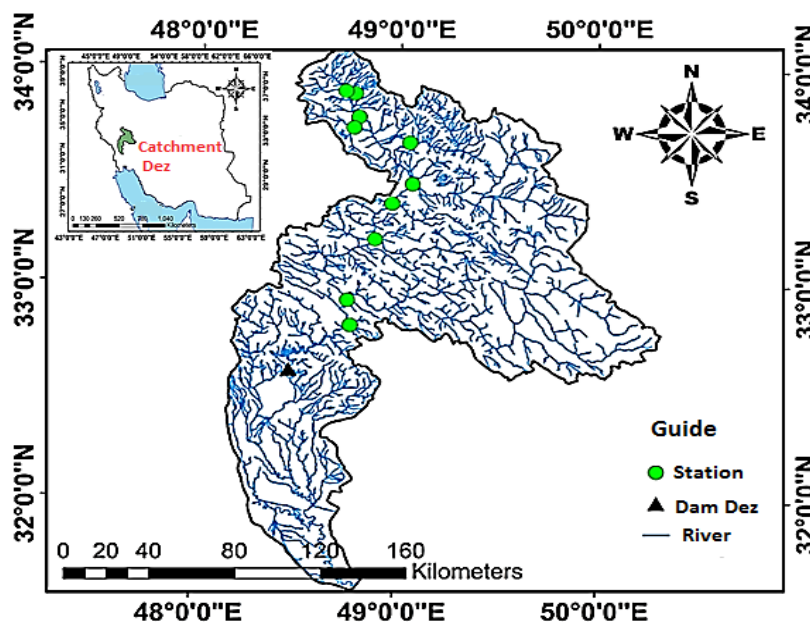


Fig. 1- The study area

شکل ۱- منطقه مورد مطالعه

Table 1- Specifications of the studied stations

جدول ۱- مشخصات ایستگاه‌های مورد بررسی

Number	Station	Longitude	latitude
1	Abgarmeh	48 ° 43' 19" "	33 ° 54' 7" "
2	Tirmarvak	49 ° 3' 46" "	33 ° 28' 37" "
3	Gelerod	48 ° 36' 23" "	33 ° 54' 8" "
4	Silakhor	48 ° 47' 52" "	33 ° 46' 43" "
5	Tire	49 ° 3' 46" "	33 ° 28' 37" "
6	Chamchit	47 ° 57' 58" "	33 ° 22' 43" "
7	Absardeh	48 ° 45' 17" "	33 ° 44' 13" "
8	Sezar	49 ° 33' 15" "	33 ° 54' 7" "
9	Talezang	48 ° 46' 17" "	32 ° 49'
10	Tangpang	48 ° 46'	32 ° 56'

(et al., 2007). در این پژوهش از این سه تابع کرنل استفاده شده است که در این بین کرنل توابع پایه شعاعی از عملکرد بهتری نسبت به سایر کرنل‌ها برخوردار بود، استفاده گردید. قابل ذکر است فرآیند محاسبات ماشین بردار پشتیبان براساس کدنویسی در محیط متلب انجام، و پارامترهای توابع کرنل از طریق سعی و خطا بهینه گردیدند.

$$k(x, x_j) = (t + x_i \cdot x_j)^d \quad (8)$$

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (9)$$

$$k(x, x_j) = x_i \cdot x_j \quad (10)$$

۳-۲- الگوریتم گیاهان مصنوعی (Artificial Flora)

۲-۳-۱- مابانی بیولوژیکی

گیاهان روش‌هایی جهت پخش دانه‌های خود داشته که می‌توان به دو روش خودبرافشانی و دیگربرافشانی تقسیم کرد. خودبرافشان به گیاهانی اطلاق می‌شود که دانه‌ها را خودشان پخش می‌کنند و دیگربرافشان به گیاهانی گفته می‌شود که دانه‌ها را با استفاده نیروهای خارجی پخش می‌کنند. خودبرافشانی شرایط مهاجرت مستقل گیاهان به محیط مناسب‌تر را فراهم می‌کند (Cheng et al., 2018). از طرف دیگر، دیگربرافشانی شرایط مهاجرت گیاهان به مناطق دورتر و نامعلوم را فراهم می‌کند. این روش‌های پخش دانه دامنه جستجوی فلور را گسترش داده و احتمال انقراض آن را کاهش می‌دهند. همچنین به دلیل تغییر آب و هوا، محیط طبیعی با شرایط سخت یا رقابت، منطقه توزیع فلور گسترش و کاهش می‌یابد و یا این که فلور مهاجرت می‌کند. با مهاجرت فلور به محیط جدید، هر یک از گونه‌های گیاهی موجود در فلور تکامل پیدا می‌کند. بنابراین، مهاجرت فلور می‌تواند منطقه توزیع را تغییر دهد و باعث تکامل، انقراض و تولد دوباره فلور شود. یک گیاه نمی‌تواند حرکت کند و هوشمند هم نیست، اما فلور با پخش دانه‌ها و تولید مثل می‌تواند بهترین مکان را برای زندگی پیدا کند. در مهاجرت و تولید مثل فلور، گیاه اصلی دانه‌ها را به طور تصادفی در فاصله معینی

چنانچه W^T ترانهاده بردار ضرایب و b ثابت مشخصه‌های تابع رگرسیونی و \emptyset نیز تابع کرنل باشد، آنگاه هدف پیدا کردن فرم تابعی برای $f(x)$ است. این مهم با آموزش مدل SVM توسط مجموعه‌ای از داده‌ها (مجموعه آموزش) محقق می‌شود (Misra et al., 2009). برای محاسبه W و b لازم است تابع خطا (معادله ۳) در مدل ε -SVM با در نظر گرفتن شرایط مندرج (قیود) در معادلات ۴ و ۵ کمینه شود (Humei, 2009):

$$\frac{1}{2} W^T \cdot W + C \sum_{i=1}^N \varepsilon_i + C \sum_{i=1}^N \varepsilon_i^* \quad (3)$$

$$W^T \cdot \emptyset(X_i) + b - y_i \leq \varepsilon + \varepsilon_i^* \quad (4)$$

$$y_i - W^T \cdot \emptyset(X_i) - b \leq \varepsilon + \varepsilon_i, \varepsilon_i, \varepsilon_i^* \geq 0, \quad i=1, 2, \dots, N \quad (5)$$

در معادلات بالا C عددی صحیح و مثبت است، که عامل تعیین جریمه در هنگام رخ دادن خطای آموزش مدل است. \emptyset 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:

$$f(x) = \sum_{i=1}^N \bar{\alpha}_i \emptyset(x_i)^T \cdot \emptyset(x) + b \quad (6)$$

در معادله ۶ $\bar{\alpha}_i$ میانگین ضرایب لاگراژ می‌باشد. محاسبه $\emptyset(x)$ در فضای مشخصه آن ممکن است بسیار پیچیده باشد (Yoon et al., 2011). برای حل این مشکل روند معمول در مدل SVM رگرسیون انتخاب یک تابع کرنل به صورت رابطه زیر است:

$$K(X_j, X) = \emptyset(X_j)^T \sqrt{b^2 - 4ac} \quad (7)$$

می‌توان از توابع مختلف کرنل برای ساخت انواع مختلف ε -SVM استفاده کرد. انواع توابع کرنل قابل استفاده در مدل SVM رگرسیونی عبارتند از: کرنل چندجمله‌ای^۱ و کرنل توابع پایه شعاعی^۲ (RBF) و کرنل خطی، به ترتیب طبق روابط ۸ تا ۱۰ محاسبه می‌گردند. با توجه به اینکه از پرکاربردترین توابع کرنل، کرنل پایه شعاعی، خطی و چند جمله‌ای می‌باشد (Vapnik and Chervonenkis, 1991; Basak)

$$d'_{1j} = d_{2j} \quad (12)$$

فاصله پخش والدین جدید، انحراف معیار بین موقعیت گیاه اصلی و موقعیت گیاه فرزند است.

$$d'_{2j} = \sqrt{\frac{\sum_{i=1}^N (P_{i,j} - P'_{i,j})^2}{N}} \quad (13)$$

رفتار پخش -

ابتدا الگوریتم فلور مصنوعی به طور تصادفی فلور اصلی را با N جواب تولید می‌کند، یعنی N گیاه در فلور وجود دارد. موقعیت گیاهان اصلی توسط ماتریس $P_{i,j}$ بیان می‌شود که در آن i بعد و j تعداد گیاهان فلور است:

$$P_{i,j} = \text{rand}(0,1) \times d \times 2 - d \quad (14)$$

که در آن d مقدار حداکثر حد منطقه و $\text{rand}(0,1)$ مجموعه‌ای از اعداد تصادفی است که به طور یکنواخت بین $(0,1)$ توزیع شده‌اند.

موقعیت گیاه فرزند با توجه به تابع پخش و به شرح زیر تولید می‌شود:

$$P'_{i,j \times m} = D_{i,j \times m} + P_{i,j} \quad (15)$$

که در آن m تعداد دانه‌هایی است که یک گیاه می‌تواند پخش کند، $P'_{i,j \times m}$ موقعیت گیاهان فرزند، $P_{i,j}$ موقعیت گیاه اصلی، و $D_{i,j \times m}$ یک عدد تصادفی با توزیع گاوسی با میانگین صفر و واریانس d_j هستند. اگر هیچ گیاه فرزندی زنده نماند، مطابق معادله (14) گیاه اصلی جدید تولید می‌شود.

رفتار انتخاب -

این که گیاهان فرزند زنده هستند یا نه با استفاده از احتمال بقاء و به شرح زیر تعیین می‌شود:

$$P = \left[\frac{F(P'_{i,j \times m})}{F_{\max}} \right] \times Q_x^{(i \times m - 1)} \quad (16)$$

که در آن $Q_x^{(i \times m - 1)}$ برابر Q_x به توان $(j \times m - 1)$ بوده و در آن Q_x احتمال انتخابی است. این مقدار باید بین صفر تا ۱ باشد. می‌توان دید که برازندگی گیاهان فرزند دورتر از گیاه اصلی، پایین‌تر است. Q_x توانایی جستجوی الگوریتم را تعیین می‌کند. Q_x باید برای مسأله‌ای که به راحتی می‌تواند به جواب بهینه محلی برسد، بزرگتر باشد. F_{\max} حداکثر برازندگی در فلور این نسل و $F(P'_{i,j \times m})$ برازندگی جواب j ام است.

۲-۴- تبدیل موجک

تبدیل موجک به عنوان روشی جایگزین برای تبدیل فوریه‌ی زمان کوتاه ارائه شده است و هدف از ارائه‌ی آن، غلبه بر مشکلات مربوط

پراکنده می‌کند. احتمال بقاء یک دانه با توجه به محیط خارجی متفاوت است. در یک محیط مناسب، یک گیاه زنده مانده و بعد از رسیدن دانه‌های خود را در اطراف پخش می‌کند. در محیط‌های دارای شرایط سخت، احتمال تکامل فلور برای سازگاری با محیط و یا انقراض آن در منطقه زیاد است. قبل از این که فلور در یک منطقه کاملاً منقرض شود، به دلیل احتمال تکثیر فلور در مناطق دیگر، دانه‌ها به روش دیگربرافشانی پراکنده می‌شوند. این دانه‌ها ممکن است به هر منطقه جدیدی که در آن احتمال از سرگیری تولید مثل فلور وجود دارد، برده شوند. از طریق تکثیر چند نسلی، فلور به مناسب‌ترین منطقه مهاجرت می‌کند. تحت این مکانیزم مهاجرت و تولید مثل، فلور وظیفه خود را برای یافتن محیط بهینه رشد از طریق تکامل، انقراض و تولد دوباره انجام می‌دهد (Cheng et al., 2018).

۲-۳-۲- نظریه الگوریتم گیاهان مصنوعی

الگوریتم گیاهان مصنوعی از چهار عنصر اساسی تشکیل شده است: گیاه اصلی، گیاه فرزند، موقعیت گیاه و فاصله پخش. گیاهان اصلی به گیاهانی گفته می‌شود که آماده پخش دانه هستند. گیاهان فرزند دانه گیاهان اصلی بوده و در آن لحظه نمی‌توانند دانه پخش کنند. موقعیت گیاه همان محل استقرار گیاه است. فاصله پخش به میزان فاصله پخش دانه اشاره دارد. در این الگوریتم سه الگوی اصلی رفتاری وجود دارد: رفتار تکامل، رفتار پخش و رفتار انتخاب (Pagie and Mitchell, 2002; Wiegand and Sarma, 2004; Rosin and Belew, 1995). رفتار تکامل به معنی احتمال تکامل گیاه برای سازگاری با رفتارهای محیطی است (Cartlidge and Balloc, 2004; Williams and Mitchell, 2005; Hillis, 1990). رفتار پخش به حرکت دانه‌ها اشاره دارد. این دانه‌ها می‌توانند از طریق خودبرافشانی یا دیگربرافشانی حرکت کنند. رفتار انتخاب نیز به معنای بقاء یا انقراض گیاه به دلایل زیست‌محیطی است. در شکل ۲ فلوچارت پیاده‌سازی الگوریتم گیاهان مصنوعی نشان داده شده است.

رفتار تکامل -

گیاه اصلی دانه‌ها را در یک دایره با شعاعی برابر فاصله پخش پراکنده می‌کند. این فاصله پخش از فاصله پخش گیاه مادر [والدین] و گیاه مادربزرگ [جد] به دست می‌آید.

$$d_j = d_{1j} \times \text{rand}(0,1) \times c_1 + d_{2j} \times \text{rand}(0,1) \times c_2 \quad (11)$$

که در آن d_{1j} فاصله پخش گیاه مادربزرگ، d_{2j} فاصله پخش گیاه مادر، c_1 و c_2 ضرایب یادگیری هستند. $\text{rand}(0,1)$ عددی با توزیع یکنواخت مستقل در بازه $(0,1)$ است. فاصله پخش مادربزرگ جدید برابر است با:

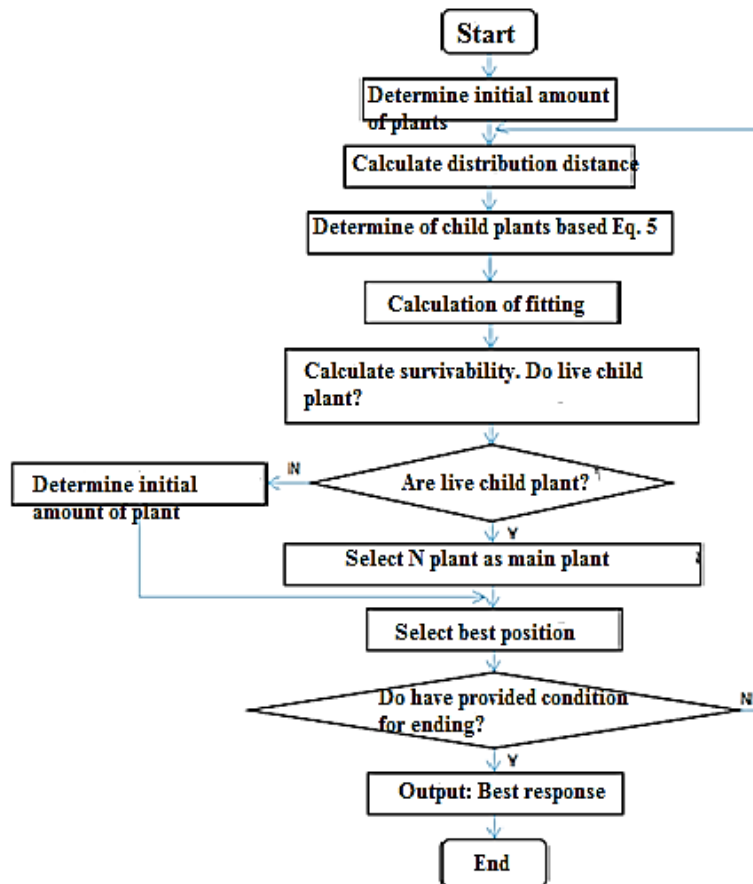


Fig. 2- Flowchart of artificial flora algorithm (Cheng et al., 2018)
 شکل ۲- فلوجارت الگوریتم گیاهان مصنوعی (Cheng et al., 2018)

نمونه‌های انتقال یافته^۳ و مقیاس شده‌ی^۴ سیگنال مادر هستند که نوساناتی در یک طول متناهی داشته و شدیداً میرا هستند. بر اساس این ویژگی مهم تبدیل موجک، می‌توان سری‌های زمانی نامانا و گذرا^۵ را به صورت موضعی مورد تجزیه و تحلیل قرار داد (Shin et al., 2005). تبدیل موجک به دو صورت پیوسته و گسسته تعریف می‌شود.

– تبدیل موجک پیوسته^۶ (CWT)

تبدیل موجک پیوسته‌ی تابع $f(t)$ به صورت روابط (۱۷) و (۱۸) تعریف می‌شود:

$$CWT_f^\psi(s, \tau) = \Psi_f^\psi(s, \tau) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-\tau}{s} \right) dt = \langle f(t), \psi_{s,\tau}(t) \rangle \quad (17)$$

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}} \psi \left(\frac{t-\tau}{s} \right) \quad (18)$$

رابطه‌ی (۱۸) یک رابطه با دو متغیر s و τ است که s پارامتر مقیاس (عکس فرکانس) و τ پارامتر انتقال می‌باشند. علامت * نیز نشان‌دهنده‌ی مزدوج مختلط است. ψ تابع پنجره یا موجک مادر و

به قدرت تفکیک‌پذیری فرکانس در تبدیل فوریه‌ی زمان کوتاه است. در تبدیل موجک همانند تبدیل فوریه‌ی زمان کوتاه، سیگنال مورد نظر به پنجره‌هایی تقسیم شده و تبدیل موجک بر روی هر کدام از این پنجره‌ها به صورت جداگانه انجام می‌گیرد (Wang et al., 2000). اما مهم‌ترین تفاوت آن‌ها در این است که در تبدیل موجک علاوه بر اینکه قدرت تفکیک فرکانس‌های یک سیگنال یا طول پنجره، متناسب با نوع فرکانس تغییر می‌کند، هم‌زمان عرض پنجره یا مقیاس فرکانس نیز متناسب با نوع فرکانس تغییر می‌کند. به عبارت دیگر، در تبدیل موجک به جای فرکانس، مقیاس وجود دارد. یعنی تبدیل موجک، نوعی تبدیل زمان-مقیاس است. بر همین اساس با استفاده از تبدیل موجک، در مقیاس‌های بالا سیگنال منبسط شده و جزئیات سیگنال قابل تجزیه و تحلیل است و در مقیاس‌های پایین سیگنال منقبض شده و کلیات سیگنال قابل بررسی می‌باشد (Wang et al., 2000). یک موجک به معنای موج کوچک، بخشی یا پنجره‌ای از سیگنال اصلی است که انرژی آن در زمان متمرکز شده است. با استفاده از تبدیل یا آنالیز موجک می‌توان یک سیگنال یا سری زمانی مادر را به موجک‌هایی با سطح تفکیک و مقیاس‌های مختلف تجزیه کرد. بنابراین موجک‌ها

بایستی سعی نمود مؤثرترین داده‌های مشاهداتی را به‌عنوان داده‌های آموزشی انتخاب کرد. با توجه به اینکه در تحقیق حاضر تأثیر توالی دبی روزهای قبل در پیش‌بینی دبی روز مدنظر بوده، صرفاً از داده‌های نرمال شده دبی جریان با توالی برگشتی تا ۴ روز به‌عنوان داده‌های آموزشی، به‌صورت ترکیب‌های مختلف مطابق جدول ۲ استفاده شده است که از این پس از آن‌ها به‌عنوان الگو نام برده خواهد شد و همچنین با توجه به اینکه مقدار همبستگی متقابل بین داده‌های ورودی و خروجی که مطابق شکل ۳ بیش از ۰/۷۵۰ حاصل شد، ترکیب‌های مختلفی از پارامترهای ورودی به‌منظور دستیابی به مدل بهینه جهت تخمین جریان روزانه حوضه آبریز دز استفاده گردید (Othman and Naseri, 2011). بدین منظور از داده‌های ایستگاه‌های هیدرومتری آبگرمه، گله رود، سیلاخور، آبسرد، تیره مروک، تیره دورود، چم چیت، سزار، تنگ پنج و تله زنگ واقع در حوضه آبریز دز که دارای ۳۶۵۰ رکورد ثبت‌شده طی بازه زمانی (۱۳۹۷-۱۳۸۷)، در مقیاس زمانی روزانه می‌باشد، استفاده شد. در نهایت تعداد ۲۹۲۰ رکورد برای آموزش و تعداد ۷۳۰ رکورد باقی‌مانده برای صحت‌سنجی مدل‌های مورد بررسی انتخاب شد. لازم به ذکر است جهت مدل‌سازی ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد باقی‌مانده جهت آزمون، به‌صورت تصادفی، که گستره وسیعی از انواع داده‌ها را پوشش دهد، انتخاب شد (Kisi et al., 2007; Nugy et al., 2002). همچنین در جدول ۳ خصوصیات آماری پارامترهای مورد استفاده نشان داده شده است.

۳-۱- نتایج مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی

در این پژوهش یک روش پیشنهادی هیبریدی حاصل از مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی مورد بررسی قرار می‌گیرد ابتدا به‌منظور کاهش دامنه تغییرات داده‌های جریان روزانه رودخانه‌های حوضه آبریز دز به همسان‌سازی اطلاعات ورودی و خروجی، بر روی داده‌ها استانداردسازی صورت گرفت. در مرحله بعد مقادیر بهینه مشخصه‌های مدل SVM شامل ε و C تعیین می‌گردد. همچنین در این پژوهش کرنل‌های متفاوتی مورد بررسی قرار گرفت و باتوجه به عملکرد کرنل‌ها، تابع کرنل مورد استفاده، تابع RBF انتخاب شد چراکه از دقت بهتری در برآورد جریان روزانه رودخانه‌ها برخوردار است (Lin et al., 2006; Liong and Sivapragasam, 2002). در این تابع نیز مشخصه γ بایستی تعیین گردد. بنابراین در حالت کلی برای پیش‌بینی جریان روزانه رودخانه‌ها به‌وسیله مدل SVM، لازم است که مقادیر بهینه سه مشخصه مذکور محاسبه گردد که بدین منظور سه مشخصه ε ، C و γ توسط الگوریتم گیاهان مصنوعی به بهترین مقادیر برازش داده می‌شوند.

$\frac{1}{\sqrt{|s|}} \Psi\left(\frac{t-\tau}{s}\right)$ موجک‌های حاصل از انتقال و تغییر مقیاس موجک مادر هستند. واژه‌ی مادر به این سبب به کار می‌رود که تمامی نسخه‌های انتقال یافته و مقیاس شده (موجک‌های دختر)، همگی از این تابع به دست می‌آیند. یعنی موجک مادر یک الگو برای سایر پنجره‌هاست. علامت (...) نیز نشان‌دهنده‌ی ضرب برداری دو تابع در فضای سیگنال است.

۲-۵- معیارهای ارزیابی

در این تحقیق به منظور ارزیابی دقت و کارایی مدل‌ها، از نمایه‌های ضریب تبیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، میانگین قدر مطلق خطا (MAE) و ضریب نش ساتکلیف (NS) طبق روابط زیر استفاده گردید (Zhu et al., 2007). بهترین مقدار برای این چهار معیار به ترتیب یک، صفر، صفر و یک می‌باشد.

$$R^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad -1 \leq R \leq 1 \quad (19)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad 0 \leq RMSE \leq 1 \quad (20)$$

$$MAE = \frac{1}{n} \sum_{i=1}^N |x_i - y_i| \quad 0 \leq MAE \leq 1 \quad (21)$$

$$NS = 1 - \frac{\sum_{i=1}^N (x_i - y_i)^2}{\sum_{i=1}^N (x_i - \bar{y})^2} \quad -\infty \leq NS \leq 1 \quad (22)$$

در روابط بالا، x_i و y_i به ترتیب مقادیر مشاهداتی و محاسباتی در گام زمانی i ام، N تعداد گام‌های زمانی، \bar{x} و \bar{y} نیز به ترتیب میانگین مقادیر مشاهداتی و محاسباتی می‌باشد.

۳- نتایج و بحث

یکی از مهم‌ترین مراحل در مدل‌سازی، انتخاب ترکیب مناسبی از متغیرهای ورودی است. از این‌رو، ابتدا همبستگی متقابل بین متغیرهای ورودی و خروجی محاسبه گردید و پارامترهای ورودی به‌منظور دستیابی به مدل بهینه جهت پیش‌بینی جریان رودخانه‌های حوضه آبریز دز انتخاب و در شکل ۳ ارائه شد. در این شکل، ستون‌های $Q(t-1)$ ، $Q(t-2)$ ، $Q(t-3)$ و $Q(t-4)$ (به ترتیب جریان رودخانه در زمان‌های $t-1$ ، $t-2$ ، $t-3$ ، $t-4$) به‌عنوان ورودی و $Q(t)$ (جریان رودخانه در زمان t) به‌عنوان خروجی مدل در نظر گرفته شد. به‌منظور آموزش ماهیت سازوکار حاکم بر پدیده نه‌تنها سبب پیچیدگی الگو و افزایش حافظه درگیر خواهد شد، بلکه سبب کاهش دقت مدل نیز می‌شود (Othman and Naseri, 2011). لذا در الگوسازی جریان رودخانه نیز

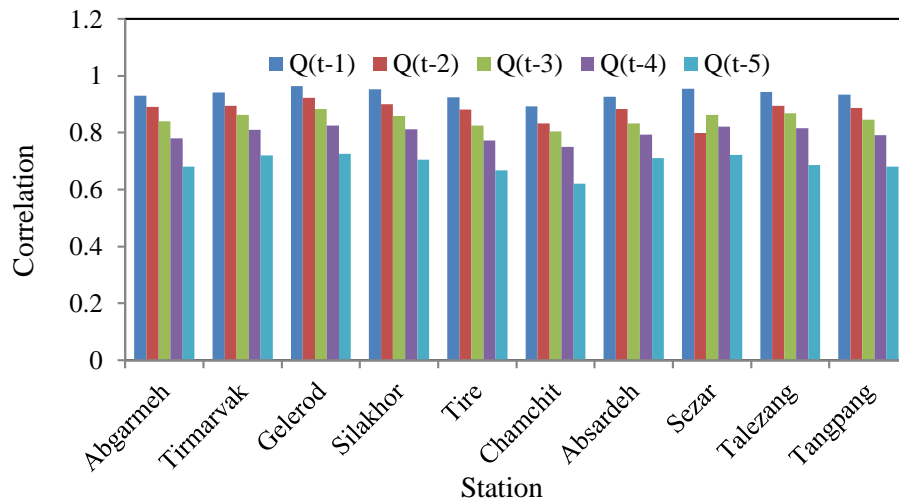


Fig. 3- The cross-correlation between input and output parameters

شکل ۳- همبستگی متقابل بین پارامترهای ورودی و خروجی

Table 2- Selected compositions of input parameter

جدول ۲- ترکیب‌های منتخب پارامترهای ورودی

Number	Input	Output
1	Q(t-1)	Q(t)
2	Q(t-1), Q(t-2)	Q(t)
3	Q(t-1), Q(t-2), Q(t-3)	Q(t)
4	Q(t-1), Q(t-2), Q(t-3), Q(t-4)	Q(t)

Table 3- Statistical characteristics of the parameters used

جدول ۳- خصوصیات‌های آماری پارامترهای مورد استفاده

Station	Minimum	Maximum	Mean	SD	Skew
Abgarmeh	0.14	16.10	1.61	1.26	1.83
Tire Marvak	0.00	8.87	0.71	1.23	2.99
Gelerod	0.22	7.74	1.74	1.12	1.53
Silakhor	0.00	149.00	4.12	6.58	7.55
Tireh	0.29	256.00	7.16	11.74	7.34
Chamchit	0.72	76.60	5.80	5.74	3.90
Absardeh	0.16	22.20	1.79	1.87	3.25
Sezar	0.00	1963.00	57.84	77.61	6.94
Talezang	0.00	5799.00	147.18	166.44	13.80
Tangpang	0.00	3191.00	86.50	119.26	11.42

می‌شود از سوی دیگر رفتار پخش شامل خودبرافشانی و دیگربرافشانی بوده که رفتار خودبرافشانی در اطراف خود و رفتار دیگربرافشانی در فضای بیشتری مکان بهینه را جستجو می‌نماید که این امر توانایی این الگوریتم در یافتن مکان بهینه بهبود می‌بخشد و همچنین سبب افزایش سرعت همگرایی به جواب بهینه می‌شود. نتایج حاصل بررسی مدل SVM در جدول ۵ آورده شد. در این جدول مدل هیبریدی پیشنهادی در ایستگاههای انتهایی حوضه آبریز به دلیل بهره‌گیری از دبی پایه در جریان رودخانه از دقت بیشتر و خطای کمتری برخوردار است. همچنین مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی

حال می‌توان از بین مدل‌های توسعه داده شده مدلی با کمترین خطا را تعیین کرده و مشخصه‌های آن را به‌عنوان مقادیر بهینه ϵ , C و γ انتخاب نمود که در جدول ۴ مقادیر بهینه متغیرهای مدل ماشین بردار پشتیبان که توسط الگوریتم گیاهان مصنوعی بهینه شده است نشان داده شد. الگوریتم گیاهان مصنوعی از رفتار مهاجرت و تولید مثل گیاهان الهام گرفته است که دارای سه رفتار اصلی شامل تکامل، پخش و انتخاب می‌باشد در رفتار تکامل فاصله انتشار گیاه فرزند براساس فاصله گیاه والد تکامل می‌یابد از آنجایی که فاصله پخش گیاه فرزند به والد وابسته نیست، این الگوریتم از راه حل بهینه محلی خارج

عملکرد بهتری نسبت به سایر توابع برخوردار بود. در جدول ۶ نتایج حاصل از مدل ترکیبی در ایستگاه‌های منتخب نشان داده شده است. مدل ماشین بردار پشتیبان-موجک ضریب همبستگی $(R^2=0/890-0/967)$ ، ریشه میانگین مربعات خطا $(0/045 \text{ m}^3/\text{s})$ و ضریب همبستگی $(R^2=0/911-0/984)$ در مرحله صحت‌سنجی از خود نشان داده است. در این جدول مدل هیبریدی در ایستگاه‌های گله‌رود، سیلاخور، سزار و تله زنگ از دقت بیشتر و خطای کمتری برخوردار است. در شکل ۵ نمودار پراکنش مدل هیبرید ماشین بردار پشتیبان-موجک مربوط به مرحله صحت‌سنجی نشان‌دهنده خط برازش مقادیر محاسباتی با بهترین خط $y=x$ است. در این شکل مقادیر تخمینی و مشاهداتی به‌غیر از چند نقطه که روی خط نیمساز قرار گرفته، که این امر دلالت بر برابری مقادیر مشاهداتی و محاسباتی براساس خط $(y=x)$ است.

همچنین همانطور که در این شکل مشهود است مدل هیبریدی در پیش‌بینی مقادیر میانی از عملکرد قابل قبولی برخوردار است به‌گونه‌ای که این مقادیر را نزدیک به مقدار واقعیشان تخمین زده است.

۳-۳- مقایسه عملکرد مدل‌ها

در ادامه با انتخاب جواب بهینه هر کدام از مدل‌های هیبریدی هوش مصنوعی و مقایسه آنها با یکدیگر مشخص گردید، هر دو مدل با دقت خوبی توانسته‌اند دبی حوضه آبریز دز را شبیه‌سازی کنند در شکل ۶ نمودار ضریب همبستگی حاصل از مدل‌های ماشین بردار پشتیبان-گیاهان مصنوعی و ماشین بردار پشتیبان-موجک در مرحله صحت‌سنجی نشان داده شده است در این شکل مدل ماشین بردار پشتیبان-گیاهان مصنوعی در کلیه ایستگاه‌های منتخب بدلیل بهینه

ضریب همبستگی $(R^2=0/933-0/985)$ ، ریشه میانگین مربعات خطا $(RMSE=0/008-0/088 \text{ m}^3/\text{s})$ ، میانگین قدرمطلق خطا $(MAE=0/004-0/040 \text{ m}^3/\text{s})$ و ضریب نش ساتکلیف $(NS=0/951-0/995)$ در مرحله صحت‌سنجی از خود نشان داده است. در شکل ۴ نمودار پراکنش مدل هیبریدی پیشنهادی مربوط به مرحله صحت‌سنجی نشان‌دهنده خط برازش مقادیر محاسباتی با بهترین خط $y=x$ است. در این شکل مقادیر تخمینی و مشاهداتی به غیر از چند نقطه که روی خط نیمساز قرار گرفته، که این امر دلالت بر برابری مقادیر مشاهداتی و محاسباتی براساس خط $(y=x)$ است. همچنین همانطور که در این شکل مشهود است مدل هیبریدی در پیش‌بینی مقادیر حداکثر و حداقل از عملکرد قابل قبولی برخوردار است بگونه‌ای که این مقادیر را نزدیک به مقدار واقعیشان تخمین زده است.

Table 4- Optimal variables of the support vector machine model

Algorithm	SVM Parameters		
	γ	ϵ	C
Artificial Flora (AF)	10	0.1	0.18

۳-۲- نتایج مدل ماشین بردار پشتیبان-موجک

به‌منظور بررسی نتایج حاصل از مدل ترکیبی، ابتدا با استفاده از تبدیل موجک، پارامترهای ورودی به زیرسیگنال‌ها تجزیه و سپس زیرسیگنال‌های مزبور، به عنوان ورودی به مدل ماشین بردار پشتیبان وارد می‌شوند و بدین ترتیب مدل ترکیبی ماشین بردار پشتیبان موجک حاصل می‌گردد. یکی از نکات مهم و اساسی در این پژوهش توابع موجک مختلف مورد بررسی قرار گرفت و موجک کلاه مکزیکی از

Table 5- The analysis of AF-SVM model results for selected stations
جدول ۵- تحلیل نتایج مدل ماشین بردار پشتیبان- گیاهان مصنوعی برای ایستگاه‌های منتخب

Station	Training				Testing			
	R^2	RMSE (m3/s)	MAE (m3/s)	NS	R^2	RMSE (m3/s)	MAE (m3/s)	NS
Abgarmeh	0.915	0.060	0.025	0.936	0.942	0.037	0.018	0.958
Tiremarvak	0.930	0.071	0.028	0.947	0.968	0.033	0.021	0.987
Gelerod	0.922	0.053	0.015	0.942	0.948	0.032	0.012	0.948
Silakhor	0.935	0.065	0.026	0.962	0.970	0.031	0.015	0.988
Tireh	0.912	0.065	0.030	0.928	0.933	0.042	0.020	0.951
Chamchit	0.850	0.118	0.076	0.874	0.898	0.088	0.040	0.917
Absardeh	0.918	0.052	0.017	0.942	0.938	0.035	0.015	0.936
Sezar	0.936	0.038	0.028	0.966	0.938	0.021	0.011	0.981
Talezang	0.946	0.029	0.020	0.971	0.972	0.015	0.008	0.989
Tangpanj	0.953	0.018	0.009	0.988	0.985	0.008	0.004	0.995

Table 6- The analysis of WSVM model results for selected stations
جدول ۶- تحلیل نتایج مدل ماشین بردار پشتیبان- موجک برای ایستگاه‌های منتخب

Station	Training			Testing				
	R ²	RMSE (m3/s)	MAE (m3/s)	NS	R ²	RMSE (m3/s)	MAE (m3/s)	NS
Abgarmeh	0.910	0.064	0.028	0.934	0.933	0.042	0.020	0.951
Tiremarvak	0.928	0.074	0.032	0.944	0.965	0.036	0.025	0.984
Gelerod	0.944	0.032	0.022	0.968	0.970	0.018	0.010	0.987
Silakhor	0.931	0.069	0.029	0.958	0.967	0.034	0.017	0.985
Tireh	0.908	0.067	0.032	0.925	0.930	0.045	0.022	0.948
Chamchit	0.846	0.126	0.084	0.868	0.890	0.094	0.045	0.911
Absardeh	0.915	0.058	0.021	0.937	0.935	0.038	0.018	0.932
Sezar	0.932	0.042	0.031	0.960	0.960	0.025	0.014	0.978
Talezang	0.924	0.048	0.033	0.941	0.957	0.028	0.017	0.973
Tangpanj	0.917	0.055	0.018	0.938	0.938	0.035	0.016	0.954

دو شاخص آماری ضریب همبستگی و انحراف استاندارد استفاده می‌کند (Taylor, 2001). هرچه مقدار پیش‌بینی شده از نظر ضریب همبستگی و انحراف استاندارد به ارزش مشاهده نزدیکتر باشد، پیش‌بینی بهتر صورت می‌گیرد. نمودار عملکرد تیلور که در شکل ۱۰ نشان می‌دهد که مدل ماشین بردار پشتیبان الگوریتم گیاهان مصنوعی دارای بالاترین راندمان و کارایی است، زیرا انحراف استاندارد پیش‌بینی شده نزدیکترین فاصله تا انحراف استاندارد از داده‌های مشاهداتی دارد و ضریب همبستگی بالاترین را نشان می‌دهد. مطابق کلیه معیارهای ارزیابی میتوان اظهار نمود مدل‌های مورد بررسی عملکرد مناسبی در برآورد دبی رودخانه‌ها دارند در این میان مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی دقت بیشتری دارد.

۴- نتیجه‌گیری

در این تحقیق سعی بر آن شد عملکرد مدل‌هایی جهت شبیه‌سازی جریان روزانه رودخانه‌های حوضه آبریز دز با استفاده از داده‌های روزانه مورد ارزیابی قرار گیرد. مدل‌های به کار گرفته شده شامل مدل هیبریدی ماشین بردار پشتیبان- گیاهان مصنوعی و ماشین بردار پشتیبان- موجک است. مقادیر جریان مشاهداتی با جریان پیش‌بینی شده با استفاده از معیارهای ارزیابی مورد مقایسه قرار گرفت. نتایج تحقیق را می‌توان به صورت زیر خلاصه نمود: نتایج نشان داد که در هر دو مدل مورد بررسی شامل مدل هیبریدی ماشین بردار پشتیبان- گیاهان مصنوعی و ماشین بردار پشتیبان- موجک، در ساختاری متشکل از ۱ تا ۴ تأخیر زمانی نتایج بهتری نسبت به سایر ساختارها ارائه می‌دهد. همچنین با توجه به معیارهای ارزیابی نتیجه شد که هر دو مدل مورد بررسی، می‌توانند با دقت نسبتاً بالایی میزان جریان روزانه رودخانه را پیش‌بینی نمایند. در این میان، مدل هیبریدی پیش‌بینی ماشین بردار پشتیبان- گیاهان مصنوعی دقت بیشتر و

نمودن مقادیر پارامترهای مدل توسط الگوریتم گیاهان مصنوعی از دقت و همپوشانی بهتری نسبت به سایر مدل‌های مورد بررسی برخوردار است. مدل ماشین بردار پشتیبان- موجک در ایستگاه‌های تیره مروک، تیره دورود، سیلاخور و آبسرده عملکرد مناسبی از خود نشان داده است. در تبیین این نتایج میتوان بیان نمود تبدیل موجکی با جدا نمودن سیگنال به فرکانس‌های بالا و پایین ویژگی‌های چند مقیاسی سیگنال را در اختیار داشته و دقت مدل را تا حد قابل توجهی، بالا می‌برد. از سوی دیگر طبق ریشه میانگین مربعات خطا که در شکل ۷ نشان داده شد مدل هیبریدی ماشین بردار پشتیبان- موجک در ایستگاه‌های گله‌رود و سزار خطای کمتری از خود نشان داده است. همچنین هر دو مدل مورد بررسی در ایستگاه چم چیت واقع در مرز مشترک استانهای لرستان و خوزستان خطای بیشتری نشان داده است. در شکل ۸ نمودار مقادیر مشاهداتی و محاسباتی مدل‌های مورد بررسی نسبت به زمان نشان در کلیه ایستگاه‌های مورد بررسی نشان داده شده است همان‌طور که مشاهده می‌گردد مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی در تخمین مقادیر حداقل و حداکثر قابلیت قابل قبولی از خود نشان داده است. همچنین مدل ماشین بردار پشتیبان- موجک در تخمین مقادیر میانی از عملکرد مناسبی برخوردار است به گونه‌ای که این مقادیر را نزدیک به مقادیر مشاهداتی تخمین زده است. در شکل ۹ نمودار خطای نسبی مدل‌های مورد بررسی نسبت به مقادیر مشاهداتی ترسیم شده است در این شکل مدل ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی نسبت به مدل ماشین بردار پشتیبان- موجک از خطای کمتری برخوردار است بگونه‌ای که در کلیه ایستگاه‌های مورد بررسی که در شکل ۸ مشاهده می‌گردد میزان خطای نسبی مدل ماشین بردار پشتیبان- موجک بیشتر می‌باشد. نمودارهای تیلور برای تجزیه و تحلیل و ارزیابی مدل‌های مورد استفاده در مطالعه مورد استفاده قرار گرفت، همان‌طور که در شکل ۱۰ نشان داده شده است. یک مزیت بارز نمودار تیلور این است که از

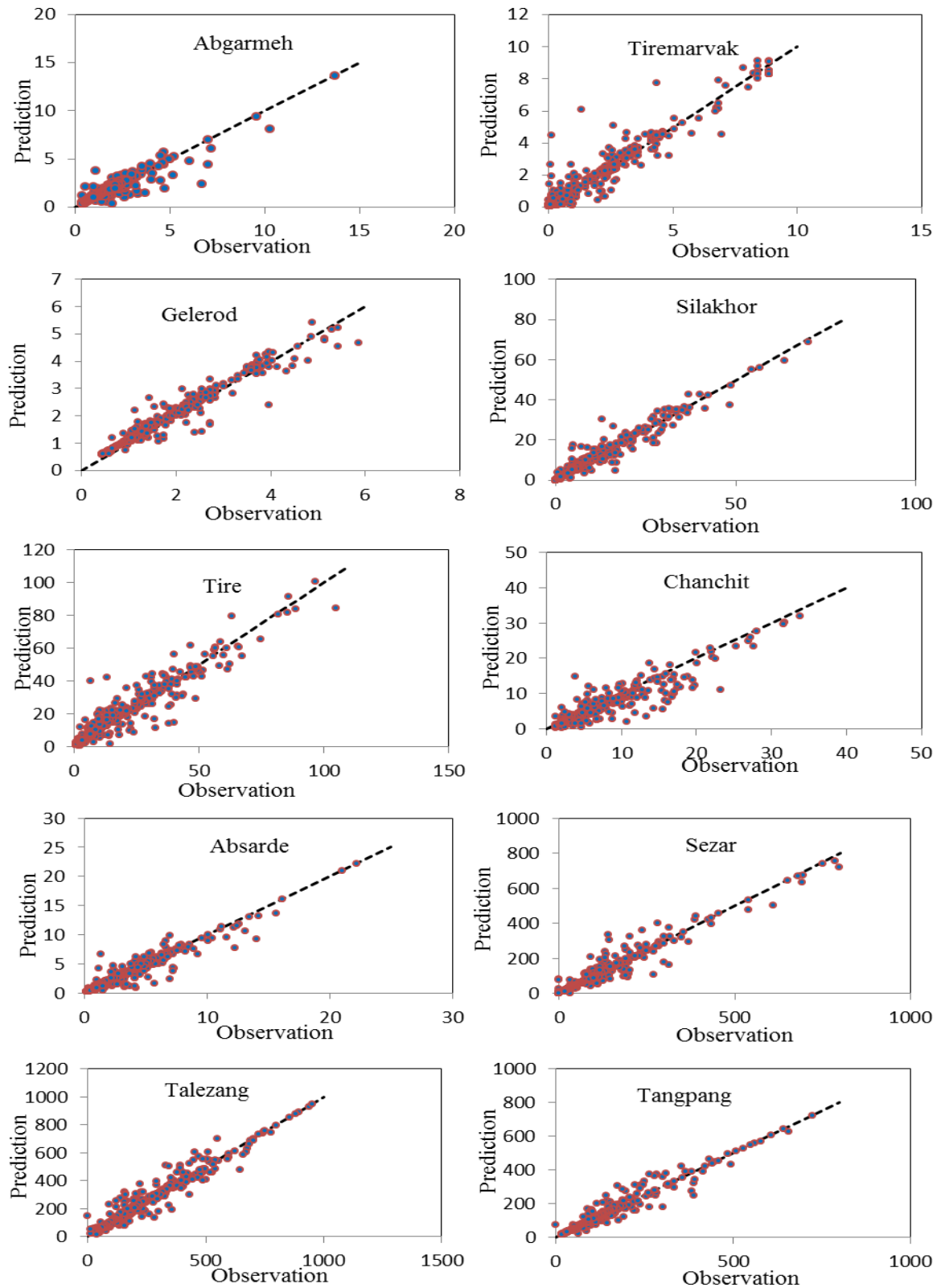


Fig. 4- Scatterplot of observational and computational values obtained from AF-SVM in the validation stage

شکل ۴- نمودار پراکنش مقادیر مشاهداتی و محاسباتی حاصل از مدل ماشین بردار پشتیبان - گیاهان مصنوعی در مرحله صحت‌سنجی

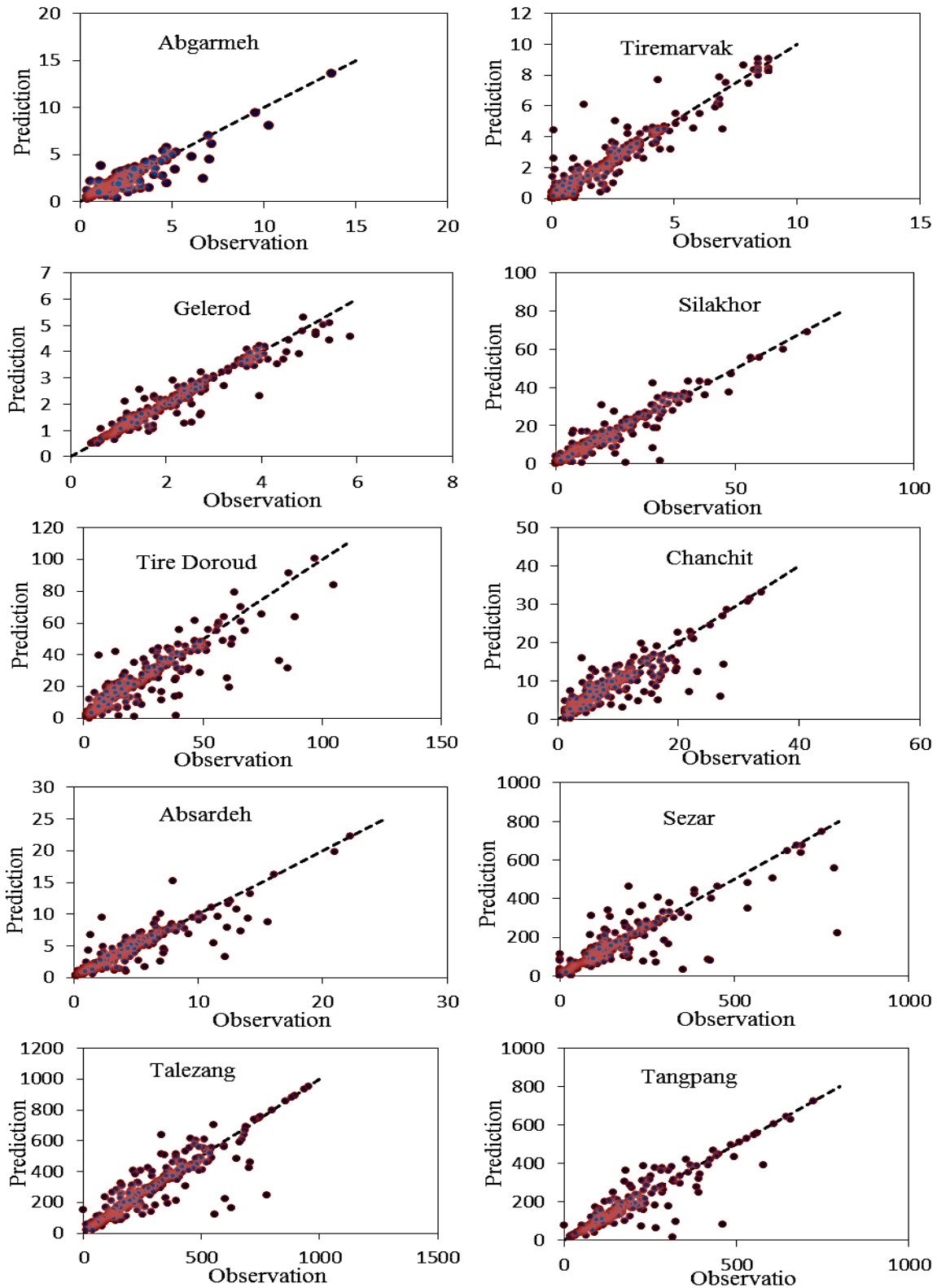


Fig. 5- Scatterplot of observational and computational values obtained from Wavelet-SVM in the validation stage

شکل ۵- نمودار پراکنش مقادیر مشاهداتی و محاسباتی حاصل از مدل ماتسین بردار پشتیبان- موجک در مرحله صحت‌سنجی

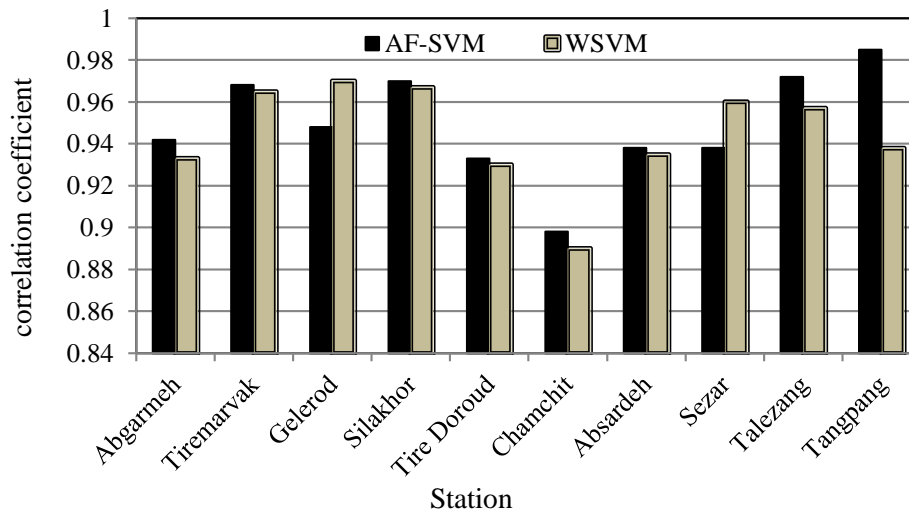


Fig. 6- The correlation coefficient of the studied models in the validation section

شکل ۶- میزان ضریب همبستگی مدل‌های مورد بررسی در بخش صحت‌سنجی

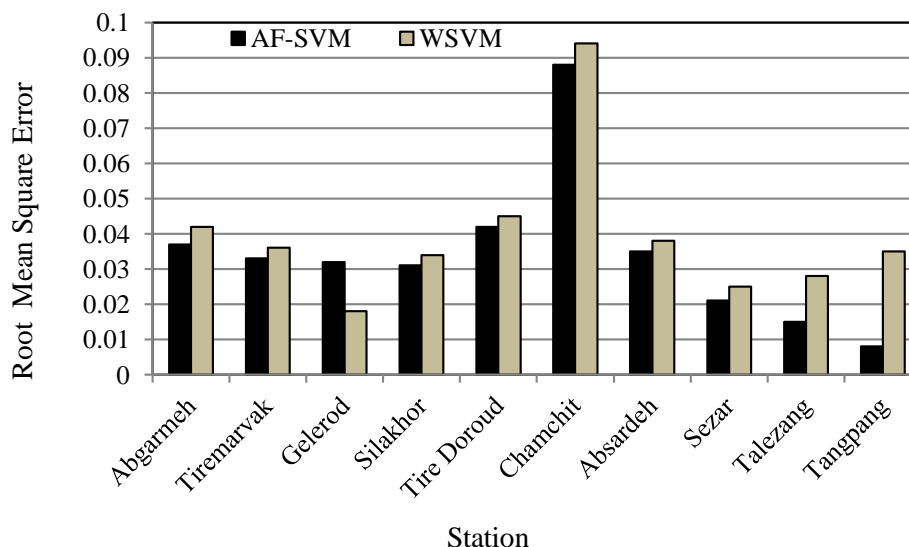


Fig. 7- Diagram of the root mean error squares of the models studied in the validation section

شکل ۷- نمودار ریشه میانگین مربعات خطا مدل‌های مورد بررسی در بخش صحت‌سنجی

با ابعاد بزرگ و سرعت مناسب در همگرایی به جواب بهینه می‌شود، دانست. در مجموع این تحقیق نشان می‌دهد استفاده از مدل ترکیبی ماشین بردار پشتیبان گیاهان مصنوعی می‌تواند در زمینه پیش‌بینی جریان روزانه رودخانه‌ها مؤثر باشد. و همچنین، با توجه به این‌که تصمیم‌گیری جهت بهره‌برداری از منابع آب و پیاده‌سازی استراتژی‌های مدیریتی در بسیاری مصارف (به‌ویژه کشاورزی و صنعت) برآورد دقیق جریان رودخانه‌ها وابسته است؛ لذا استفاده از مدل هیبریدی پیشنهادی جهت پیش‌بینی جریان رودخانه‌ها، می‌تواند به‌عنوان ابزاری مناسب در اتخاذ تصمیمات مدیریتی مورد استفاده قرار گیرد. همچنین پیشنهاد می‌گردد از مدل‌های هیبریدی ماشین بردار

خطای کمتری نسبت به مدل ماشین بردار پشتیبان- موجک از خود نشان داده است. همچنین نتایج مدل‌های مورد بررسی طبق نمودار تیلور نشان داد مدل هیبریدی پیشنهادی ماشین بردار پشتیبان- الگوریتم گیاهان مصنوعی دقت بیشتری دارد. در مجموع می‌توان بیان نمود دقت بالای مدل ترکیبی ناشی بهینه نمودن پارامترهای مدل ماشین بردار پشتیبان مطابق الگوریتم گیاهان مصنوعی به بهینه‌ترین مقادیر ممکن می‌باشد. این الگوریتم ترکیبی از بهینه‌سازی پیوسته و گسسته است در مسایل با دامنه بزرگ، زمان دستیابی به جواب بهینه جستجوی سراسری کاهش یافته و از جواب‌های بهینه محلی خارج می‌شود که این امر سبب توانایی قابل قبول این الگوریتم در حل مسائل غیر خطی

مقایسه شود و نیز کاربرد مدل پیشنهادی در سایر پدیده‌های هیدرولوژیکی مورد بررسی و ارزیابی قرار گیرد.

پشتیبان با الگوریتم‌های بهینه‌سازی جدید نظیر تفنگدار خلاق، اسکی، ازدحام مرغ، ازدحام گربه و غیره استفاده و نتایج با مدل پیشنهادی

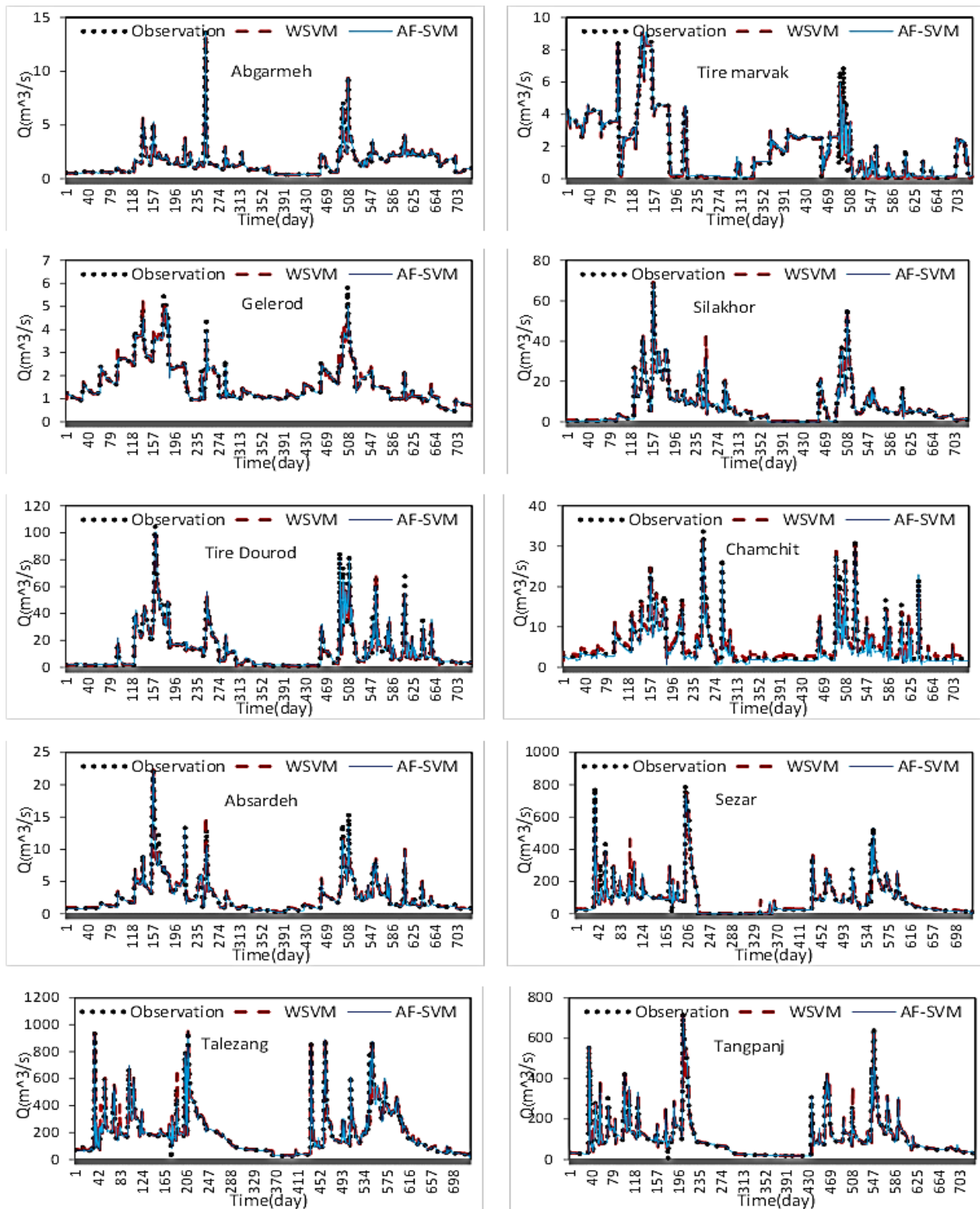


Fig. 8- Graph of computational and observational values of the studied models in the validation section

شکل ۸- نمودار مقادیر محاسباتی و مشاهداتی مدل‌های مورد بررسی در بخش صحت‌سنجی

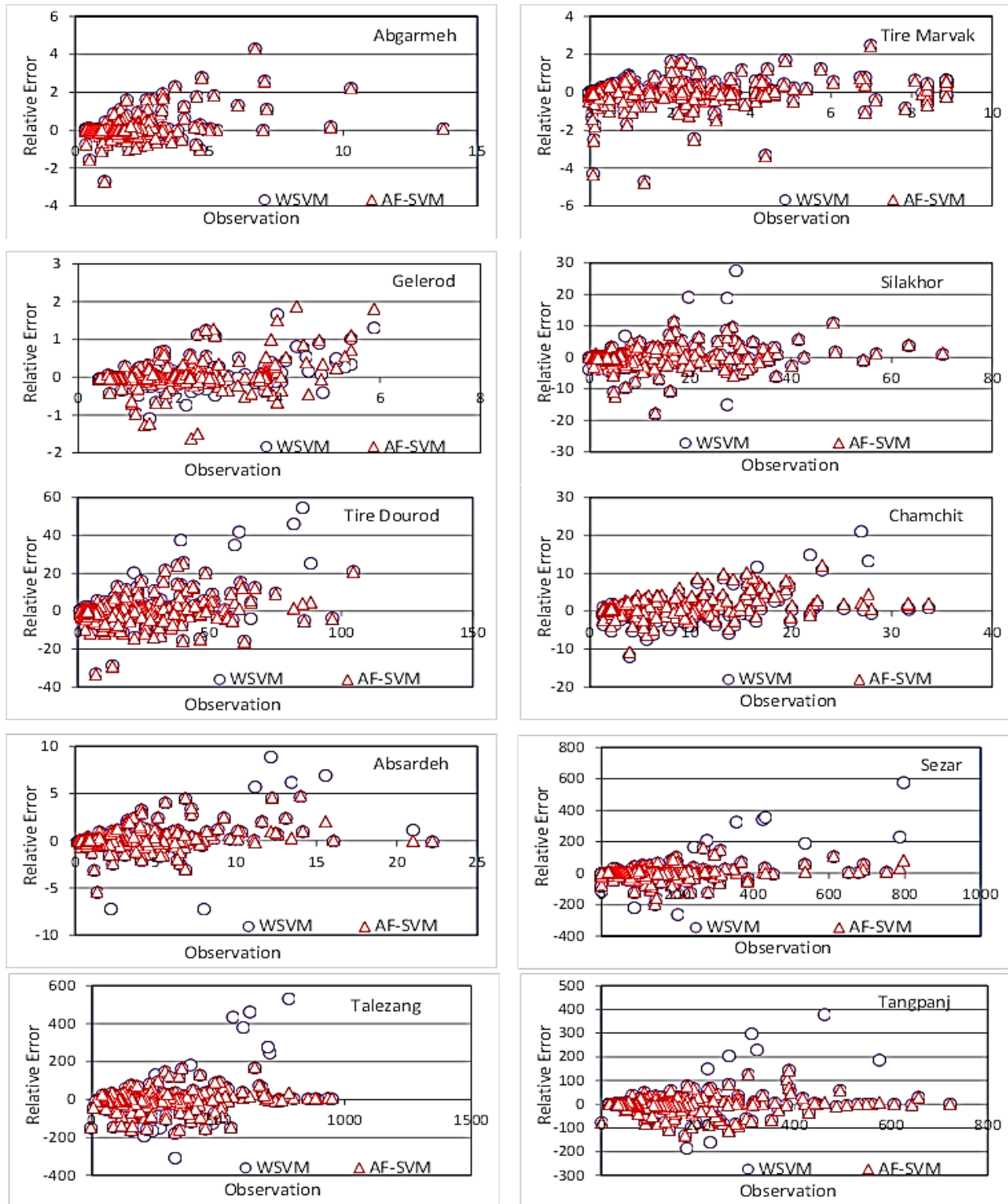
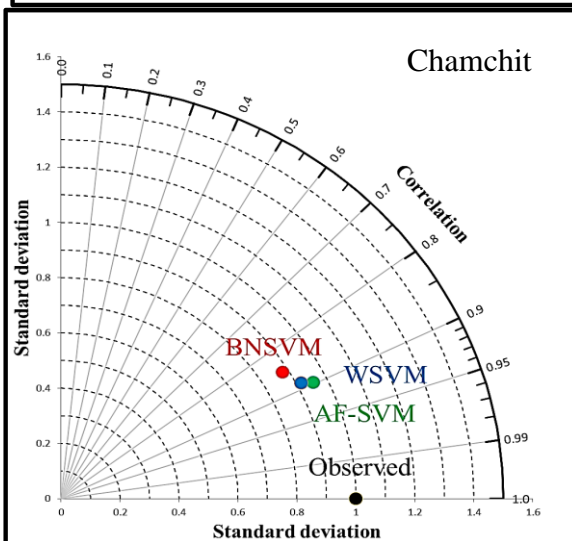
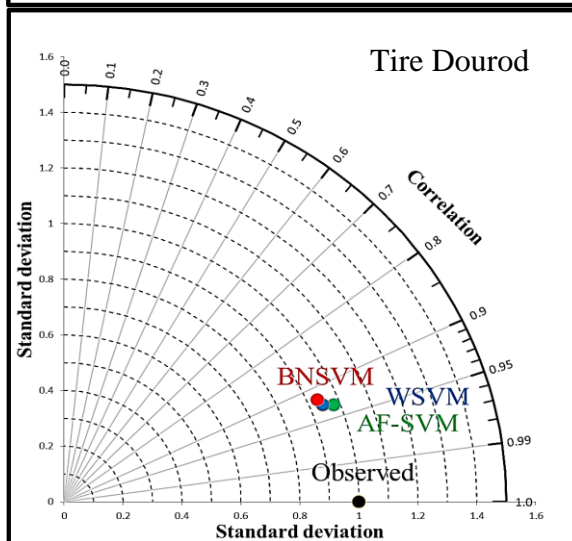
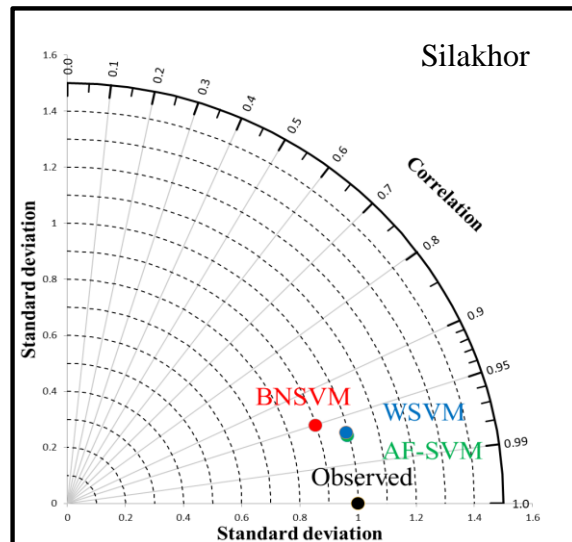
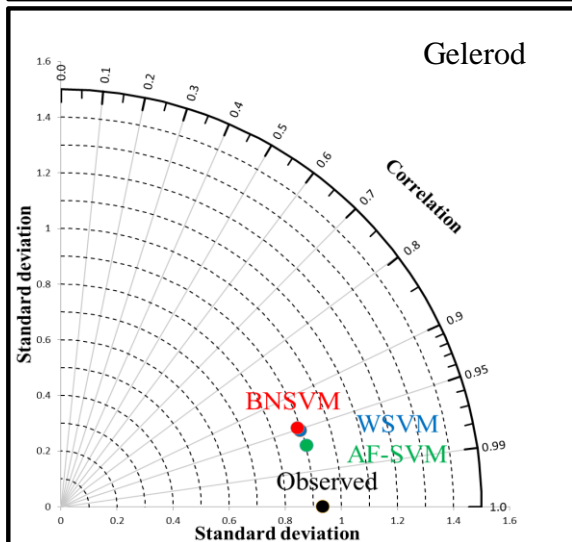
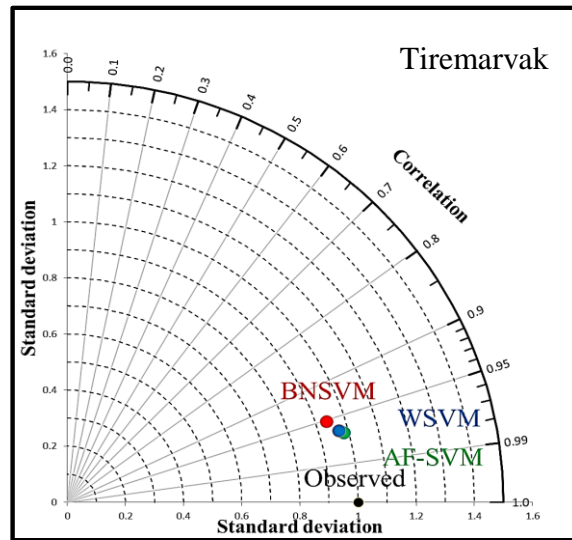
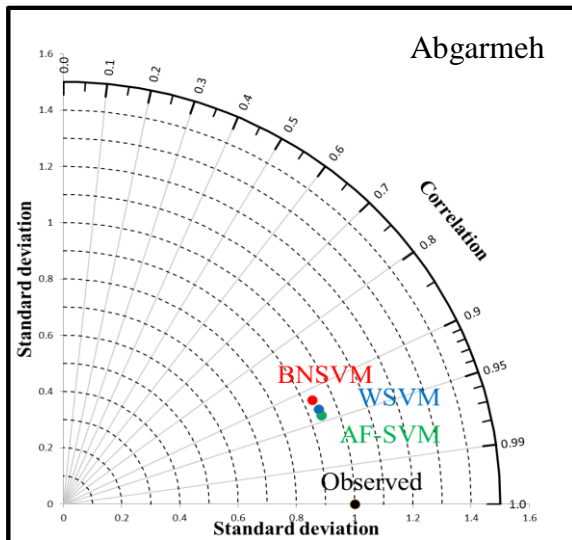


Fig. 9- The relative error diagram of the models studied in the validation section

شکل ۹- نمودار خطای نسبی مدل‌های مورد بررسی در بخش صحت‌سنجی



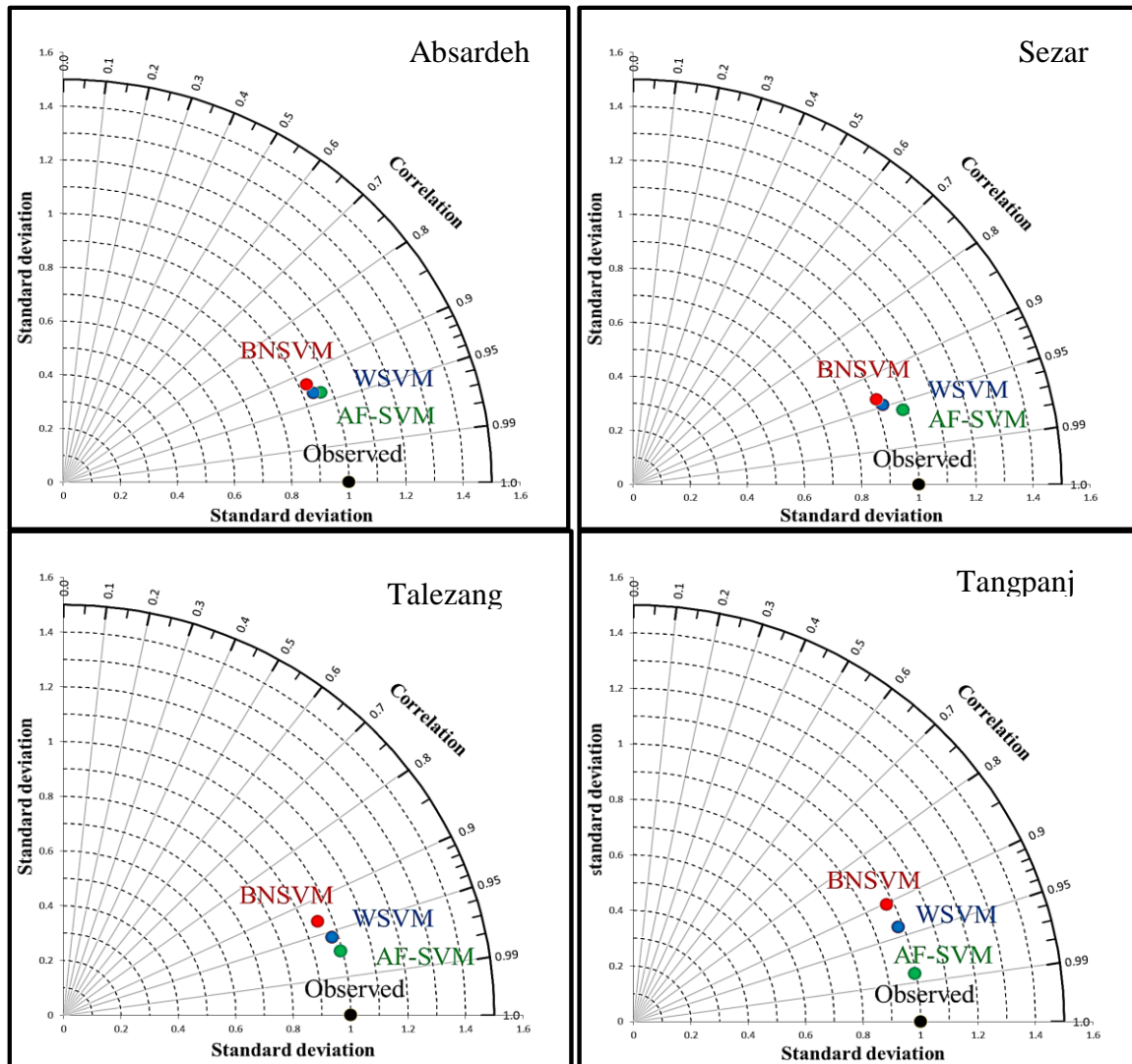


Fig. 10- Taylor diagram of the stations studied

شکل ۱۰- نمودار تیلور در ایستگاههای مورد بررسی

Alizadeh F, Gharamaleki A, Jalilzadeh M, Akhoundzadeh A (2020) Prediction of river stage-discharge process based on a conceptual model using EEMD-WT-LSSVM approach. *Water Resources* 47:41-53

Basak D, Pal S, Patranabis DC (2007) Support vector regression. *Neural Information Processing* 11:203-225

Cartlidge JP, Bulloc SG (2004) Combating coevolutionary disengagement by reducing parasite virulence. *Evolutionary Computation* 12(2):193-222

Chen H, Zhu Y (2008) Optimization based on symbiotic multi-species coevolution. *Journal on Applied Mathematics and Computation* 22(3):179-194

Cheng L, Wu X, Wang Y (2018) Artificial Flora (AF) optimization algorithm. *Applied Science* 329(8):2-22

پي نوشتها

- 1- Polynomial
- 2- Radial Basis Functions (RBF)
- 3- Translation
- 4- Dilation
- 5- Transient
- 6- Continues Wavelet Transform

۵- مراجع

Adnan R, Liang Z, Heddami S, Kermani M, Kisi O, Li B (2019) Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology* 19(4):432-448

- Othman F, Naseri M (2011) Reservoir inflow forecasting using artificial neural network. *International Journal of the Physical Sciences* 6(3):434-440
- Pagie L, Mitchell MA (2002) Comparison of evolutionary and coevolutionary search. *International Journal of Computational Intelligence and Application* 2:53-69
- Rosin CD, Belew RK (1995) Methods for competitive co-evolution. *Finding Opponents Worth Beating in Proceedings of the International Conference on Genetic Algorithms Pittsburgh*, 373-381
- Sedighi F, Vafakhah M, Javadi MR (2016) Rainfall-runoff modeling using support vector machine in snow-affected watershed. *Arabian Journal for Science and Engineering* 41(10):4065-4076
- Seyedian M, Bagherpour M, Fathabadi A, Mohammadi A (2018) Runoff prediction using black and gray box models. *Iranian Water Resources Research* 14(5):204-219 (In Persian)
- Shahinejad B, Dehghani R (2018) Comparison of wavelet neural network models, support vector machine and gene expression programming in estimating the amount of oxygen dissolved in rivers. *Iran-Water Resources Research* 14(3):226-238 (In Persian)
- Shin S, Kyung D, Lee S, Taik & Kim J, Hyun J (2005) An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications* 28(4):127-135
- Taylor E (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research* 106(7):7183-7192
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Vapnik VN (1998) *Statistical learning theory*. Wiley, New York
- Vapnik V, Chervonenkis A (1991) The necessary and sufficient conditions for consistency in the empirical risk minimization method. *Pattern Recognition and Image Analysis* 1(3):283-305
- Wang D, Safavi AA, Romagnoli JA (2000) Wavelet-based adaptive robust M-estimator for non-linear system identification. *AIChE Journal* 46(4):1607-1615
- Wiegand RP, Sarma J (2004) Spatial Embedding and loss of gradient in cooperative coevolutionary algorithms. In *Proceedings of the International Conference on Parallel Problem Solving from Nature, Berlin Germany* 43:912-921
- Williams N, Mitchell M (2005) Investigating the success of spatial coevolution. In *Proceedings of the 7th*
- Edossa DC, Babel MS (2012) Forecasting hydrological droughts using artificial neural network modeling technique South Africa. University of Pretoria, *Proceedings of 16th SANCIAHS National Hydrology Symposium*:1-10
- Ghorbani MA, Khatibi R, Geol A, Fazelifard MH, Azani A (2016) Modeling river discharge time series using support vector machine and artificial neural networks. *Environmental Earth Sciences* 75(4):675-685
- Ghorbani MA, Khatibi R, Karimi V, Yaseen ZM, Zounemat-Kermani M (2018) Learning from multiple models using artificial intelligence to improve model prediction accuracies: Application to river flows. *Water Resources Management* 32(13):4201-4215
- Hamel L (2009) *Knowledge discovery with support vector machines*. Hoboken, N.J. John Wiley
- Hillis WD (1990) Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D: Nonlinear Phenomena* 42:228-234, [CrossRef]
- Huang S, Chang J, Huang Q, Chen Y (2014) Monthly streamflow prediction using modified emd-based support vector machine. *Journal of Hydrology* 511(4):764-775
- Kisi O, Karahan M, Sen Z (2006) River suspended sediment modeling using fuzzy logic approach. *Hydrological Process* 20(2):4351-4362
- Lin JY, Cheng CT, Chau KW (2006) Using support vector machines for long-term discharge prediction. *Hydrolog Sciences Journal* 51(3):599-612
- Liong SY, Sivapragasam C (2002) Flood stage forecasting with support vector machines. *Journal of the American Water Resources Association* 38(4):173-186
- Misra D, Oommen T, Agarwa A, Mishra SK, Thompson AM (2009) Application and analysis of support vector machine based simulation for runoff and sediment yield. *Biosystems Engineering* 103(3):527-535
- Mohammadi K, Eslami HR, Dardashti SD (2005) Comparison of regression, ARIMA and ANN models for reservoir Inflow forecasting using snowmelt equivalent (A case study of Karaj). *Journal of Agricultural Science and Technology* 7:17-30
- Nagy H, Watanabe K, Hirano M (2002) Prediction of sediment load concentration in rivers using artificial neural network model. *Journal of Hydraulics Engineering* 128(3):558-559

- Zhao X, Chen X, Xu Y, Xi D, Zhang Y, Zheng X (2017) An EMD-based chaotic least squares support vector machine hybrid model for annual runoff forecasting. *Water* 9(3):140-153
- Zhu YM, Lu XX, Zhou Y (2007) Suspended sediment flux modeling with artificial neural network: An example of the longchuanjiang river in the upper yangtze catchment. *Geomorphology* 84(4):111-125
- Annual Conference on Genetic And Evolutionary Computation Washington 46:523-530
- Yoon H, Jun SC, Hyun Y, Bae GO, Lee KK (2011) A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology* 396(4):128-138
- Zhang G, Patuwo BE, Hu YM (1998) Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14(1):35-62