نشریه علمی

علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management
Science & Engineering

سال شانزدهم- شماره ۵۸- پاییز ۱۴۰۱

Vol. 16, No. 58, Fall 2022

کاربرد مدل‌های هوش مصنوعی در شبیه‌سازی کیفیّت آب بابلرود و سفیدرود

فاطمه آخونی پورحسینی[۱]، کیومرث ابراهیمی[۲]و محمدحسین امید[۳]

تاریخ دریافت : ۱۱ /۱۲/ ۱۴۰۰ تاریخ پذیرش:۲۸ /۰۳/ ۱۴۰۱

چکیده

میزان کل مواد جامد محلول[۴]عامل مهمّی در تعیین کیفیّت آب رودخانه‌ها می‌باشد. در این مطالعه، از چهار مدل مبتنی بر داده شامل سه مدل فرا ابتکاری علف هرز مهاجم[۵]، کلونی زنبورعسل[۶]، ازدحام ذرات[۷]و مدل ماشین‌بردارپشتیبان[۸] در پیش‌بینی کیفیّت آب بابلرود و سفیدرود استفاده شد. برای این منظور از داده‌های منتشرنشده‌ی ماهیانه‌ی کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pH و کل مواد جامد محلول از ایستگاه قرآن تالار بابلرود و ایستگاه پروریج‌آباد سفیدرود به ترتیب در بازه‌های زمانی ۱۳۵۷-۱۳۹٤ و ۱۳۹٤-۱۳۸۳ تجزیه‌وتحلیل شد. نتایج ارزیابی این چهار مدل بر مبنای معیارهای ضریب تبیین، جذر میانگین مربعات خطا و ضریب نش- ساتکلیف نشان داد که روش هیبریدی کلونی زنبورعسل در ایستگاه هیدرومتری قرآن تالار بابلرود و پروریج‌آباد سفیدرود در بخش صحّت‌سنجی با بیشترین ضریب تبیین به ترتیب برابر با ۰/۹۸۵و ۰/۹۸۹ و کمترین جذر میانگین مربعات خطای ۱۰/٤۹۳ و ۵/۲۸۹ میلی‌گرم بر لیتر و بیشترین ضریب نش-ساتکلیف مساوی ۰/۹۸۳ و ۰/۹۹۲ نسبت به سه مدل دیگر برتری بالا و سریعی در ارزیابی کیفیّت آب دارد.

کلیدواژه‌ها: *طبقه‌بندی، محیط‌زیست، مدل‌سازی، منابع آب، واسنجی.*

مقدمه

توسعه روزافزون فعالیت‌های کشاورزی و صنعتی و افزایش حجم فاضلاب‌های شهری موجب آلودگی رودخانه‌ها شده و کیفیّت آب این منابع حیاتی را مورد مخاطره جدّی قرار داده‌است [۷]. کلّ مواد جامد محلول در آب برابر مجموع غلظت همه یون‌های موجود در آب می‌باشد. تخمین و ارزیابی کیفیّت آب و مقدار مواد جامد محلول به‌منظور صیانت از کیفیّت منابع آب و همچنین بررسی و جلوگیری از آلودگی منابع آب‌های سطحی بسیار حائز اهمیت است [۷]. پایش مستمر پارامترهای شیمیایی، فیزیکی و بیولوژیکی بخش ضروری از برنامه‌های کنترل کیفیّت آب می‌باشد [۳]. مدل‌سازی و تخمین مقادیر پارامترهای کیفیّت آب دارای ارجحیتی همچون کاهش هزینه، صرفه‌جویی در زمان، کاهش نیاز به آزمایشگاه و تجهیزات اندازه‌گیری، کاهش تعداد نیروهای اجرایی، تولید داده‌های مصنوعی با دقّت بالا و در گستره مکانی بیشتر و بازسازی اطلاعات ازدست‌رفته می‌باشد [٤]. امروزه سامانه‌های هوشمند به‌طور گسترده برای پیش‌بینی پدیده‌های غیرخطی مورداستفاده قرار می‌گیرد [۷]. مدل ماشین‌بردار پشتیبان یک سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقّید است [۵]. در سال‌های اخیر استفاده از روش‌های هوشمند در مطالعات کیفیّت آب رودخانه‌ها مورد توجّه محققین قرارگرفته است. برخی از تحقیقات انجام‌شده در این زمینه عبارت‌اند از: شاهی نژاد و دهقانی [۱۸] برای تخمین اکسیژن محلول در آب رودخانه کامبرلند در سال‌های ۱۳۸۵ تا ۱۳۹۵ از روش‌های هوشمند شبکه عصبی موجک، ماشین بردارپشتیبان و برنامه‌ریزی بیان ژن استفاده کردند. درمجموع نتایج نشان داد که به لحاظ توانایی بالای شبکه عصبی موجک و حذف نویزهای سری‌های زمانی در تخمین پارامترهای کیفیّت آب رودخانه، این مدل می‌تواند راه‌کاری مناسب ارائه دهد. ادیب و همکاران [۱] جهت تخمین پارامترهای کیفیّت آب رودخانه سفیدرود از روش‌های ANFIS،GEPو-SVM LS در یک دوره آماری ٤۰ ساله استفاده کردند. نتایج حاصل از پژوهش ایشان نشان داد که هر سه روش هوش مصنوعی قابلیت بسیار زیادی در تخمین پارامترهای کیفی TDS، TH وECدارند. به صورتی که برای تخمین مواد جامد محلول روش ماشین بردارپشتیبان، برای تخمین EC روش GEP و برای تخمین THروش ANFIS بهترین شبیه‌سازی را انجام می‌دهد. این پژوهش نشان داد که با استفاده

۱- دانشجوی دکتری مهندسی منابع آب، گروه مهندسی آبیاری و آبادانی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، ایران
۲- نویسنده مسئول و استاد گروه مهندسی آبیاری و آبادانی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. پست الکترونیک: EbrahimiK@ut.ac.ir
۳- استاد گروه مهندسی آبیاری و آبادانی، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران

4.TDS (Total Dissolved Solid)
5.SVM-IWO
6.SVM-ABC
7.SVM-PSO
8. LS-SVM

از روش‌های هوشمند مصنوعی می‌توان غلظت پارامترهای کیفیّت آب را در صورت عدم اندازه‌گیری آن‌ها و بر اساس غلظت سایر پارامترهای کیفیّت محاسبه کرد. شریفی‌گرم‌درّه و همکاران [۱۹] برای تخمین دبی سیلاب در حوضه آبریز دریاچه نمک از اطلاعات ۵۵ ایستگاه آب‌سنجی استفاده کردند. ابتدا متغیرهای مؤثر با استفاده از روش WinGamma تعیین شد. درنهایت مدل‌سازی با استفاده از ماشین بردارپشتیبان، شبکه عصبی‌مصنوعی و رگرسیون چند متغیره غیرخطی انجام شد. نتایج پژوهش اخیر نشان داد که ماشین بردار پشتیبان نسبت به سایر روش‌های مطالعه شده دقّت بیشتری در پیش‌بینی دبی‌های لحظه‌ای داشته است.احترام و همکاران [٦] برای شبیه‌سازی سطح آب دریاچه ارومیه از روش ANFIS-SO و MLP-SO استفاده کردند. نتایج این تحقیق نشان داد که مدل هیبریدی ANFIS-SO نسبت به مدل‌های دیگر بهتر می‌تواند سطح آب دریاچه را شبیه‌سازی کند.کاووسی و خزیمه نژاد [۹] به بررسی عملکرد چهار روش، مدل‌سازی در شبیه‌سازی بارش– رواناب در هلیل رود سد جیرفت پرداختند و نتایج نشان داد که مدل ANFIS-PSO نسبت به سایر مدل‌ها در شبیه‌سازی دبی دقّت و کارایی بیشتری داشته است. راما و جایابراتی [۱۷] از الگوریتم علف هرز مهاجم برای تعیین اندازه و مکان بهینه واحدهای مولّد توزیع‌شده چندگانه در شبکه‌های توزیع استفاده نمودند. نتایج این شبیه‌سازی نشان‌دهنده کاربرد خوب و عملکرد مناسب الگوریتم علف‌های هرز در مقایسه با سایر روش‌ها بود. آخونی پورحسینی و اسدی [۲] برای پیش‌بینی تغییرات سطح ایستابی آبخوان دشت اردبیل از روش‌های ماشین بردار پشتیبان و شبکه بیزین استفاده کردند. نتایج نشان داد که ماشین بردار پشتیبان عملکرد بهتری نسبت به شبکه بیزین دارد. میرهاشمی و همکاران [١٤] به بررسی عملکرد ماشین بردار پشتیبان و الگوریتم جنگل تصادفی در پیش‌بینی جریان رودخانه خرم‌آباد پرداختند. نتایج از بهبود عملکرد مدل‌های مورداستفاده درنتیجه نویززدایی سیگنال جریان دارد. شنانی هویزه و زارعی [۱۵] به مقایسه الگوریتم‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان و حداکثر احتمال در استخراج نقشه کاربری اراضی در حوضه آبخیز ابوالعباس پرداختند. نتایج نشان داد که هر سه مدل قابلیّت تهیه کاربری اراضی با صحّت بالا دارا می‌باشند. همان‌طور که در تحقیقات بالا اشاره شد، کیفیّت آب بابلرود و سفیدرود و میزان آلودگی این رودخانه‌ها تحت تأثیر آلاینده‌های مختلفی ازجمله زه‌آب‌های کشاورزی و فاضلاب‌های شهری و صنعتی می‌باشد [۱۲]. لذا پایش وضعیت کیفیّت آب‌های سطحی رودخانه‌های موردمطالعه و بررسی روند تغییرات آن‌ها در طول زمان باهدف آگاهی از وضع حاکم و اتخاذ تمهیدات مختلف در راستای صیانت از کیفیّت منابع آب‌های سطحی ضروری است. در مقاله حاضر، برای شبیه‌سازی کیفیّت آب بابلرود و سفیدرود در دو ایستگاه هیدرومتری قرآن تالار و پروریج‌آباد، کارایی سه مدل هیبریدی ماشین‌بردارپشتیبان شامل علف‌هرز مهاجمSVM-IWO ، کلونی زنبورعسل SVM-ABC، ازدحام ذرات SVM-PSO و مدل ماشین‌بردارپشتیبان SVM LS- بررسی شد. لازم به ذکر است که استفاده از روش‌های هیبریدی علف‌هرز مهاجم، کلونی زنبورعسل برای این منظور تاکنون موردبررسی قرار نگرفته است.

مواد و روش‌ها

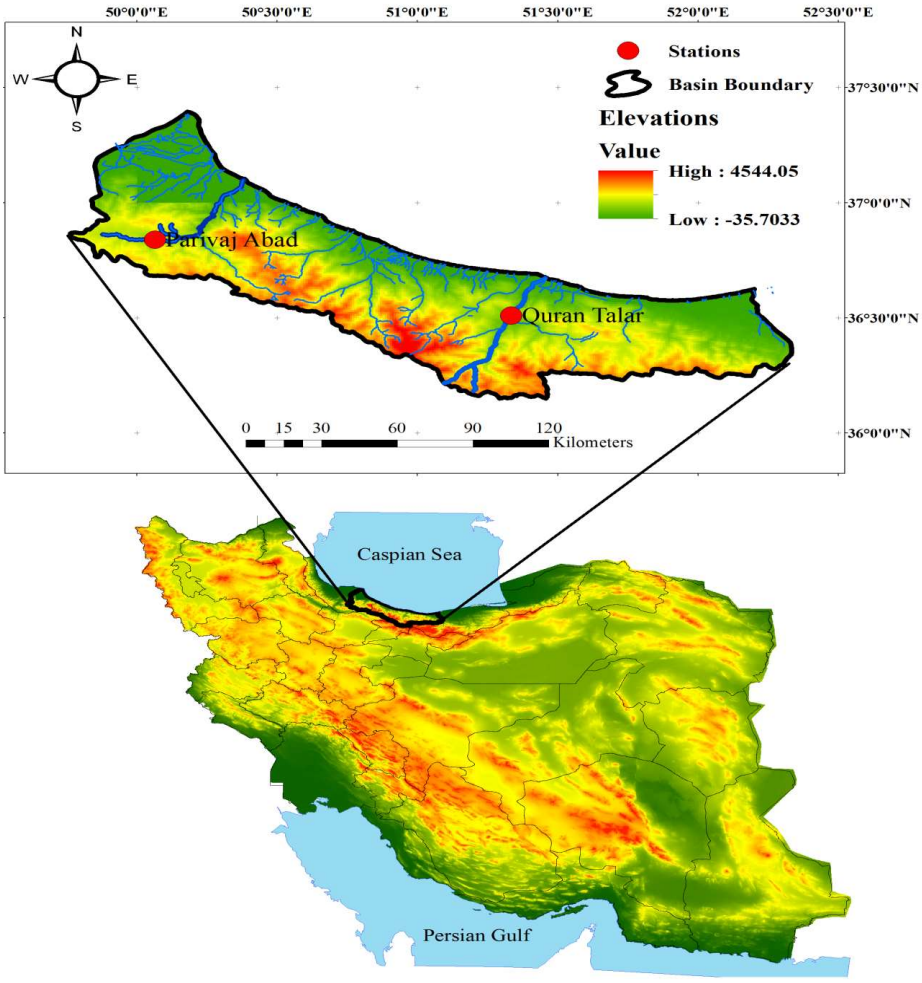
منطقه موردمطالعه و داده‌های مورداستفاده

بابلرود یکی از رودخانه‌های پر آب شمال کشور می‌باشد که در استان مازندران جاری است. طول بابلرود ۷۸ کیلومتر است. این رودخانه حوضه‌ای به مساحت ۱۷٤٦/٤۲ کیلومترمربع را زهکشی می‌کند. برای دست‌یابی به اطلاعات لازم در مورد ارزیابی کیفیّت آب بابلرود از داده‌های ماهیانه کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pHو مواد جامد محلول اندازه‌گیری شده توسط شرکت مدیریت منابع آب ایران از ۱۳۹٤–۱۳۵۷ در ایستگاه هیدرومتری قرآن تالار بابلرود با تعداد داده ٤٥٦ عدد استفاده شد. جانمایی ایستگاه هیدرومتری مذکور در شکل (۱) و نمودار داده‌های مشاهداتی مواد جامد محلول آن در شکل(۲) و مشخصّات آماری داده‌های کیفیّت آب آن نیز در جدول (۱) ارائه‌شده است. در این پژوهش همچنین از داده‌های ماهیانه کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pH و مواد جامد محلول ایستگاه هیدرومتری

جدول۱– مشخصّات پارامترهای آماری ایستگاه هیدرومتری قرآن تالار- بابلرود

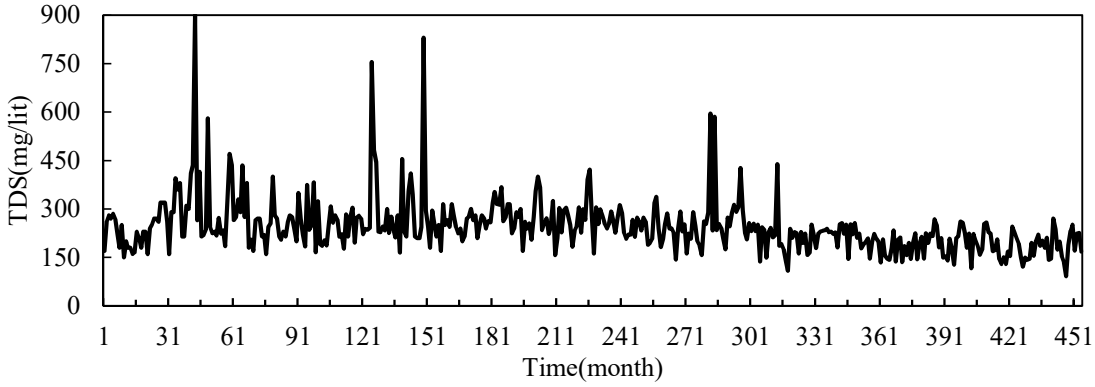
Table 1. Statistical parameters, Quran-Talar hydrometric station - Babolrood

ضریب تغییرات CV	انحراف معیار SD	حداقل Min	حداکثر Max	میانگین Mean	واحد Unit	پارامتر Parameter
0.34	84.09	92	900	244.82	mg/l	TDS
1.29	0.68	0.1	8.7	0.52	mg/l	Na
0.43	0.54	0.4	6.1	1.24	mg/l	Mg
0.32	0.66	0.55	8.4	2.04	mg/l	Ca
1.25	0.52	0.02	7.2	0.41	mg/l	SO_4
1.6	0.69	0.1	9.7	0.43	mg/l	Cl
0.28	0.83	0.8	9.9	2.9	mg/l	HCO_3
0.05	0.39	6.5	10.3	7.88	-	pH
0.34	127.64	146	1370	377.51	µmhos/cm	Ec



شکل ۱- موقعیت جغرافیایی منطقه موردمطالعه

Fig 1. The geographical location of the study area



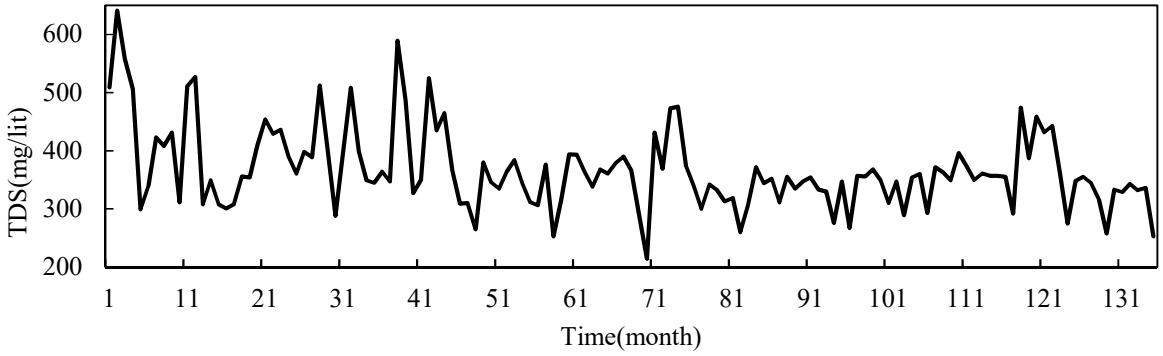
شکل ۲- نمودار مقادیر مشاهداتی مواد جامد محلول در ایستگاه هیدرومتری قرآن تالار- بابلرود

Fig 2. The Measured TDSdata, Quran-Talar hydrometric station- Babolrood

پروریج‌آباد واقع بر سفیدرود با تعداد داده ۱۳۶ عدد در طول دوره ۱۳۹٤-۱۳۸۳ اندازه‌گیری شده توسط شرکت مدیریت منابع آب ایران استفاده شد. جانمایی ایستگاه هیدرومتری مذکور در شکل (۱) و نمودار داده‌های مشاهداتی مواد جامد محلول آن در شکل(۳) و مشخصّات آماری داده‌های کیفیّت آب آن نیز در جدول (۲) ارائه‌شده است.

تحلیل حساسیت

به‌منظور بررسی دقیق‌تر پارامترهای اثرگذار در برآورد مقدار TDS با استفاده از مدل‌های علف هرز مهاجمSVM-IWO، کلونی زنبورعسلSVM-ABC، ازدحام ذراتSVM-PSO و مدل



شکل۳- نمودار مقادیر مشاهداتی مواد جامد محلول،ایستگاه هیدرومتری پروریج‌آباد - سفیدرود

Fig 3. The measured TDS data ,Parvarij-abad hydrometric station- Sefidrood

جدول۲- مشخصّات پارامترهای آماری ایستگاه هیدرومتری پروریج‌آباد - سفیدرود

Table 2. Statistical parameters, Parvarij-abad hydrometric station-Sefidrood

ضریب تغییرات CV	انحراف معیار SD	حداقل Min	حداکثر Max	میانگین Mean	واحد Unit	پارامتر Parameter
0.19	70.103	214	641	366.75	mg/l	TDS
1.12	0.91	0.1	8.8	0.81	mg/l	Na
0.32	0.54	0.5	3.1	1.71	mg/l	Mg
0.29	0.92	1.4	11.6	3.15	mg/l	Ca
0.69	0.7	0.1	3.7	1.01	mg/l	SO_4
1.06	0.89	0.2	9.6	0.83	mg/l	Cl
0.27	0.99	1.5	11	3.7	mg/l	HCO_3
0.33	0.26	7.1	9	7.94	-	pH
0.32	184.35	335	2340	583.96	µmhos/cm	Ec

ماشین‌بردارپشتیبانLS-SVMآنالیز حساسیت با استفاده از روش استات سافت (Statsoft) انجام شد. درروش استات‌سافت مقدار حساسیت برای هر ویژگی ورودی، از تقسیم نمودن خطای شبکه در غیاب متغیر ورودی موردنظر، برخطای شبکه در حضور همه متغیرهای ورودی بدست آمد. بدین معنی که ابتدا مدل با تمامی متغیرها ایجاد و بعد از رسیدن به بهترین عملکرد با کمترین خطا، مقدار شاخص خطا (RMSE) محاسبه‌شده و در نظر گرفته می‌شود. از نسبت شاخص خطا در حالت دوم به حالت اول، مقدار حساسیت خروجی نسبت به متغیر ورودی موردنظر محاسبه می‌شود [۱۱].

مدل حداقل مربعات ماشین بردار پشتیبان(LS-SVM)

اساس روش ماشین بردار پشتیبان برپایه تئوری یادگیری آماری ارائه‌شده است که از حداقل سازی ریسک ساختاری پیروی می‌کند. سانچز و همکاران [۲۰] روش حداقل مربعات ماشین بردار پشتیبان را ارائه دادند. این مدل برخلاف مدل SVM به‌جای مسئله برنامه‌نویسی درجه دوّم در حل از معادلات خطی استفاده می‌کند، بنابراین دقّت محاسباتی بیشتری نسبت به ماشین بردار پشتیبان کلاسیک دارد. مدل رگرسیون LS-SVM از رابطه زیر پیروی می‌کند:

$$\gamma_i = w^T.\varphi(x_i) + b \qquad (۱)$$

که در آن wوb به ترتیب مقادیر وزن‌ها و اریبی تابع رگرسیون هستند که از طریق حداقل سازی تابع هدف در رابطه (۲) تعیین می‌شود:

$$min w, e, b\ j(w,e) = \frac{1}{2}w^T w + \frac{Y}{2}\sum_{i=1}^{n} e_i^2 \qquad (۲)$$

با محدودیت رابطه(۳):

$$\gamma_i = w^T.\varphi(x_i) + b + x_i \quad (۳)$$

که در آن‌ها γ پارامترتنظیم‌کننده بخش خطاست و e_i خطای داده‌های آموزشی می‌باشد.

بهینه‌سازی ازدحام ذرات (PSO[1])

یک الگوریتم بهینه‌سازی مبتنی بر جمعیت و الهام گرفته‌شده از رفتار پرندگان و ماهی‌ها می‌باشد. در این الگوریتم هر راه‌حل فقط یک‌ذره در فضای جستجو می‌باشد و همه ذرات یک مقدار شایستگی دارند که توسط تابع هدف که باید بهینه شوند، ارزیابی می‌شود. در این الگوریتم، ابتدا برای هر متغیر در فضای جستجو به‌صورت تصادفی مقادیری تخصیص پیدا می‌کند. هر ذره یا پرنده با مقادیر اختصاص داده‌شده به‌تمامی متغیرها، بر اساس میزان شایستگی که همان مقدار تابع هدف هست، یک‌راه حل تولید می‌کند [۱۰].

(٤)

$$\vartheta_i^t + 1 = x\omega\vartheta_i^t + c1r1(p_i^t - x_i^t) + c2r2\,(p_i^t - x_i^t)$$

$$x_i^{t+1} = x_i^t + \vartheta_i^{t+1} \quad (۵)$$

در معادله x_i^tموقعیت یا مقدار هر ذره iدر فضای جستجو و t شماره تکرار هر ذره p_i^tموقعیت بهترین ذره در فضای جستجو که بسته به نوع جستجو متفاوت است. r1 و r2 عددهای تصادفی بین صفر و یک‌بر اساس تابع توزیع یکنواخت، c1و c2 ضریب‌های شناختی و اجتماعی، ω ضریب اینرسی می‌باشد و xضریب انقباض می‌باشد. مقادیر ضریب شناختی و اجتماعی با آزمون‌وخطا و نیز با نتایج به‌دست‌آمده از دیگر پژوهش‌ها تعیین می‌شود. این فرآیند بر اساس رسیدن به یک مقدار خاص یا تکرار معین ادامه می‌یابد. در PSO پارامترهای c1و c2 عامل یادگیری و قابل تنظیم هستند.

الگوریتم بهینه‌سازی علف‌های هرز مهاجم(IWO[2])

روش بهینه‌سازی علف‌های هرز مهاجم توسط محرابیان و لوکاس [۱٦]با الهام گرفتن از طبیعت معرفی و مورداستفاده قرار گرفت. طبق تعریف علف هرز گیاهی هست که در مکان‌های ناخواسته تولید و رشد می‌کند و برای گیاهان رزاعی افت جدی بوده و جلوی رشد آن‌ها را می‌گیرد. این الگوریتم دریافتن نقاط بهینه بسیار مؤثر و سریع هست و بر اساس ویژگی‌های اوّلیه و طبیعی علف‌های هرز در یک کلونی عمل می‌کند [۲۰]. مراحل انجام الگوریتم به ترتیب معرفی می‌شود.

تعیین مقدار جمعیت اوّلیه

ابتدا یک جمعیت محدود اوّلیه در فضای حل مسئله به‌صورت تصادفی تولید و پراکنده می‌شود [۱٦].

تولیدمثل

در این روش بهینه‌سازی هر عضو از جمعیت با توجّه به میزان توانایی‌هایشان دانه تولید می‌کنند. تعداد دانه‌هایی که هر گیاه می‌تواند تولید کند به‌طور خطی کمترین تعداد دانه ممکن تا بیشترین تعداد تغییر می‌کند و علف هرز با سازگاری بهتر دانه بیشتر تولید می‌کند. رابطه تولید تعداد دانه به‌صورت زیر می‌باشد [۱٦].

$$Seed_n = \frac{f - f_{min}}{f_{max} - f_{min}}(S_{max} - S_{min}) + S_{min} \quad (٦)$$

که در آن $Seed_n$ تعداد دانه‌های تولیدشده، f سازگاری علف‌های هرز فعلی، f_{min} ، f_{max} به ترتیب کمترین و بیشترین سازگاری از جمعیت فعلی ، s_{min}، s_{max} به ترتیب حداکثر و حداقل مقدار ممکن تولید دانه می‌باشد.

پراکندگی فضایی

در این مرحله دانه‌های تولیدشده به‌صورت تصادفی در فضای چندبعدی مسئله پراکنده می‌شود. تابع توزیع تصادفی، تابعی نرمال بوده و تضمین می‌کند دانه‌ها به‌صورت تصادفی تقسیم‌شده‌اند بسیار نزدیک به گیاه والدشان می‌باشد. مقدار انحراف معیار تابع توزیع نرمال در هر مرحله از مقدار اوّلیه تعریف‌شده $\sigma_{initial}$ تا مقدار نهایی σ_{final}کاهش پیدا می‌یابد. ارتباط بین پارامترهای بالا و انحراف معیار را می‌توان به‌صورت رابطه (۷) بیان کرد:

$$\sigma_{iter} = \frac{(iter_{max} - iter)^n}{(iter_{max})^n}\left(\sigma_{initial} - \sigma_{final}\right) + \sigma_{final} \quad (۷)$$

که در این رابطه $iter_{max}$بیشترین تعداد دفعات تکرار، σ_{iter} مقدار انحراف معیار در مرحله انجام عملیات و n شاخص غیرخطی بودن مدولاسیون یا شاخص نوسان غیرخطی می‌باشد.

حذف رقابتی

در الگوریتم‌های علف‌های هرز مهاجم بعد از چند مرحله تکرار، تعداد دانه‌های کلونی در اثر تولیدمثل به بیشترین حد خود (P_{max}) می‌رسد و سپس مکانیزمی برای حذف دانه‌های ضعیف به کار گرفته می‌شود.

الگوریتم کلونی زنبورعسل(ABC[3])

الگوریتم کلونی زنبورعسل مصنوعی، یک راهکار بهینه‌سازی است که رفتار یک کلونی زنبورعسل را شبیه‌سازی می‌کند و برای اوّلین بار در سال ۱۳۸٤ توسط کارابوگا[4] برای مسائل بهینه‌سازی ارائه شد. در این الگوریتم، یک منبع غذایی به‌عنوان هدفی در فضای جستجو تعریف می‌شود، به‌طوری‌که تعداد منابع غذایی در ابتدا برابر با تعداد زنبورهای موجود در کندوست.

مراحل اجرای الگوریتم به شرح زیر می‌باشد [۸]:

۱.تعیین اوّلیه منابع غذا

$$x_{ij} = x_j^{min} + rand(0.1)(x_j^{max} - x_j^{min}) \quad (۸)$$

1.Particle Swarm Optimization
2. Invasive Weed Optimization Algorithm
3. Artificial Bee Colony
4

به‌طوری‌که i=1,2,3,…SN، تعداد منابع غذایی=SN، j=1,2,3,…، D، تعداد پارامترها=D، x_j^{max} x_j^{min} به ترتیب حداکثر و حداقل مقدار پارامتر j می‌باشد.

۲. حرکت زنبورهای کارگر به سمت منابع غذا و تعیین منابع غذایی مجاور

$$V_{ij} = x_{ij} + \varphi_{ij}(x_{ij} - x_{kj}) \quad (۹)$$

که در این معادله j و k به ترتیب پارامترهای منتخب تصادفی و همسایگان تصادفی ، φ_{ij} یک عدد تصادفی بین ۱ و ۱- می‌باشد.

۳.تعیین یک منبع غذایی جدید به نام V_i و حذف x_i درصورتی‌که مقدار شهد منبع جدید V_i بهتر از منبع قبل x_i باشد.

٤. حرکت زنبورهای جستجو به سمت منابع غذا تعیین همسایه‌های جدید طبق مراحل ۲ و ۳:

$$p_i = \frac{fitness_i}{\sum_{j=1}^{SN} fitness_i} \quad fitness_i = \begin{cases} \frac{1}{1+fi_i} & fit \geq 0 \\ 1 + abs(fit_i) & fit < 0 \end{cases} \quad (۱۰)$$

به‌طوری‌که fit_i هزینه برای منبع غذا، x_i، $fitness_i$ مقدار شهد منبع غذا x_i می‌باشد.

۵. مراحل ۲ تا ٤ تا زمانی که مقدار بهینه یافت شود، تکرار می‌شود.

تجزیه‌وتحلیل داده‌ها

در تحقیق حاضر برای استفاده از داده‌های مذکور، ابتدا داده‌ها به‌صورت تصادفی به دو قسمت، ۷۰ درصد داده‌ها برای آموزش مدل و ۳۰ درصد داده‌ها برای بخش صحّت‌سنجی استفاده شد. به‌منظور افزایش کارایی مدل، همه داده‌ها ابتدا به شکل نرمال و با استفاده از رابطه (۱۱) بین دو عدد ۰ و ۱ استاندارد شدند.

$$X_n = \frac{X_{max} - X_i}{X_{max} - X_{min}} \quad (۱۱)$$

برای ارزیابی مدل‌های مورداستفاده از سه شاخص آماری معتبر و رایج شامل ضریب تبیین R^2، جذر میانگین مربعات خطا RMSE و ضریب نش-ساتکلیف NS_E استفاده شد. روابط شاخص‌های مذکور در زیر ارائه‌شده است. بر اساس این شاخص‌های کمّی، مدل برتر دارای مقدار حداکثر ضریب تبیین و ضریب نش- ساتکلیف (بیشترین مقدار NS_E برابر ۱ است) و حداقل RMSE خواهد بود.

$$R^2 = \left[\frac{\sum_N^{i=1}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_N^{i=1}(x_i - \bar{x})^2 . \sum_N^{i=1}(y_i - \bar{y})^2}}\right]^2 \quad (۱۲)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{N}} \quad (۱۳)$$

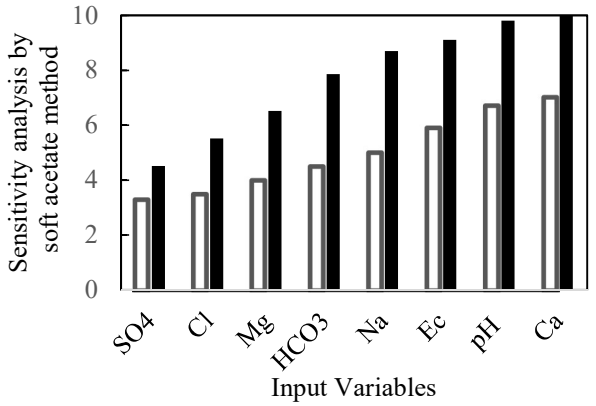
$$NS_E = 1 - \frac{\sum_N^{i=1}(x_i - y_i)^2}{\sum_N^{i=1}(x_i - \bar{x})^2} \quad (۱٤)$$

که در آن x_i داده‌های مشاهداتی، y_i داده‌های محاسباتی و $\bar{x}$ مقدار متوسط مشاهداتی می‌باشد.

نتایج

تحلیل حساسیت

نتایج تحلیل حساسیت با روش استات‌سافت شکل (٤) نشان داد در بین متغیرهای ورودی به ترتیب میزان تأثیر ورودی‌ها در بابلرود، ایستگاه هیدرومتری قرآن-تالار به ترتیب کلسیم، اسیدیته، هدایت الکتریکی، سدیم، بی‌کربنات، منیزیم، کلر و سولفات می‌باشد. در ایستگاه هیدرومتری قرآن-تالار به ترتیب کلسیم، اسیدیته، هدایت الکتریکی، سدیم، بی‌کربنات، منیزیم، کلر و سولفات می‌باشد. در سفیدرود در ایستگاه هیدرومتری پروریج‌آباد نیز به همین ترتیب میزان تأثیرپذیری پارامترهای ورودی به ترتیب کلسیم، اسیدیته، هدایت الکتریکی، سدیم، بی‌کربنات، منیزیم، کلر و سولفات می‌باشد. لذا در هر دو رودخانه پارامترهای فوق‌الذکر به‌عنوان ترکیب بهینه به مدل‌ها ارائه شد.



شکل٤-تحلیل حساسیت با روش استات سافت

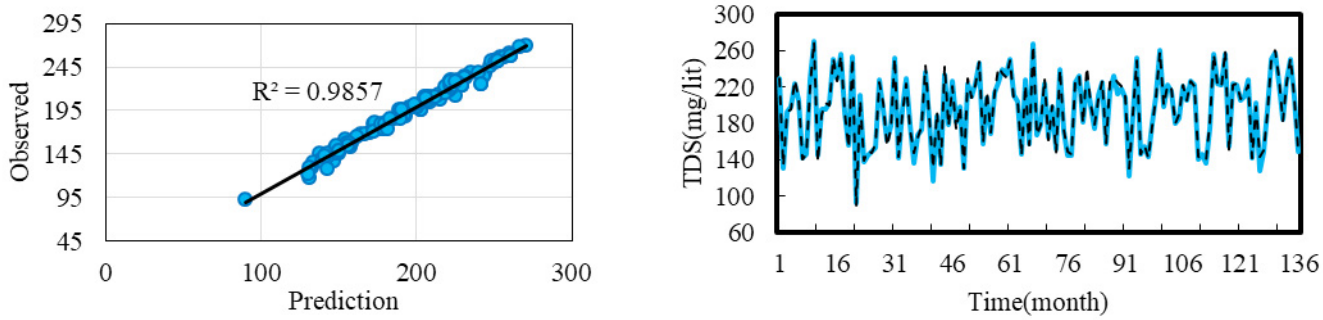
Fig 4. Analysis of acetate soft method

بعد از حذف داده‌های پرت، جهت بالا رفتن دقّت مدل‌سازی،داده‌ها به‌صورت نرمال تبدیل شد. سپس با استفاده از هیبریدهای مدل ماشین بردار پشتیبان(SVM-ABC،SVM-IWO و SVM-PSO) و LS-SVM مدل‌سازی برای ترکیب با ورودی‌های کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pH در ایستگاه‌های هیدرومتری قرآن تالار بابلرود و ایستگاه هیدرومتری پروریج‌آباد سفیدرود اجرا شد. نتایج مدل‌های مورداستفاده، در جدول (۳) ارائه‌شده است. با توجّه به جدول(۳) برای ایستگاه هیدرومتری قرآن تالار پارامترهای ورودی برای همه مدل‌ها شامل کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pH می‌باشد. با توجّه به معیارهای ارزیابی در هر مدل، بهترین مدل ایجادشده SVM-ABC می‌باشد. در شکل (۵) روند تغییرات داده‌های مشاهداتی و محاسبه‌شده توسط مدل بهینه SVM-ABC همچنین میزان پراکندگی خطاها در ایستگاه هیدرومتری قرآن تالار نشان داده‌شده است.

جدول۳- ارزیابی مدل‌های فرابتکاری و ماشین بردار پشتیبان در مراحل آموزش و صحّت سنجی بابلرود- ایستگاه هیدرومتری قرآن تالار

Table 3. Evaluation of hybrid models and Bayesian network models in training and verification stages, Goran-Talar hydrometric station-Babolrood

صحت‌سنجی Test			آموزش Train			پارامترهای ورودی Input parameters	مدل Model
$N.S_E$	RMSE (mg/l)	R^2	$N.S_E$	RMSE (mg/l)	R^2		
0.983	10.493	0.985	0.986	10.453	0.986	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-ABC
0.981	11.073	0.981	0.984	11.198	0.984	Na, Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-IWO
0.983	10.885	0.979	0.986	10.639	0.986	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-PSO
0.983	14.085	0.985	0.986	10.87	0.986	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	LS-SVM



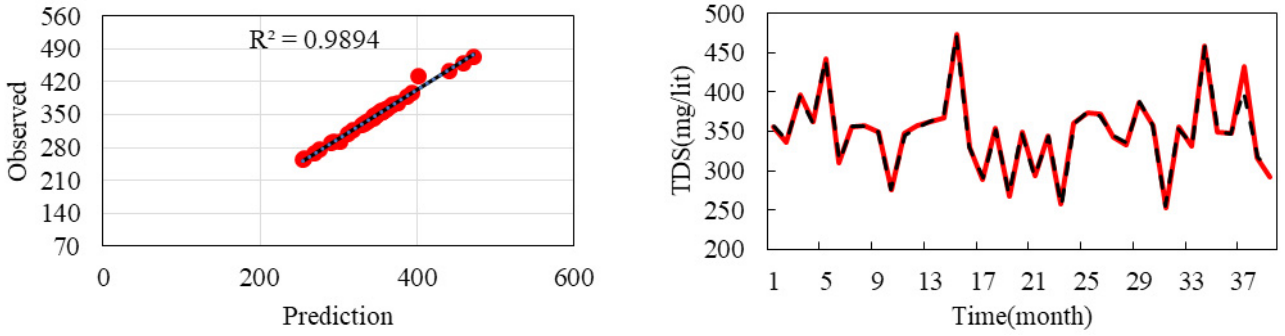
شکل۵- نتایج پیش‌بینی و دقّت مدل **SVM-ABC**در بخش صحّت سنجی ایستگاه قرآن_تالار

Fig 5. Predicted dataand accuracy of SVM-ABC model, Testing stage, Quran-Talar hydrometric Station

جدول٤- ارزیابی مدل‌های فرابتکاری و ماشین بردار پشتیبان در مراحل آموزش و صحّت سنجی ایستگاه هیدرومتری پروریج‌آباد- سفیدرود

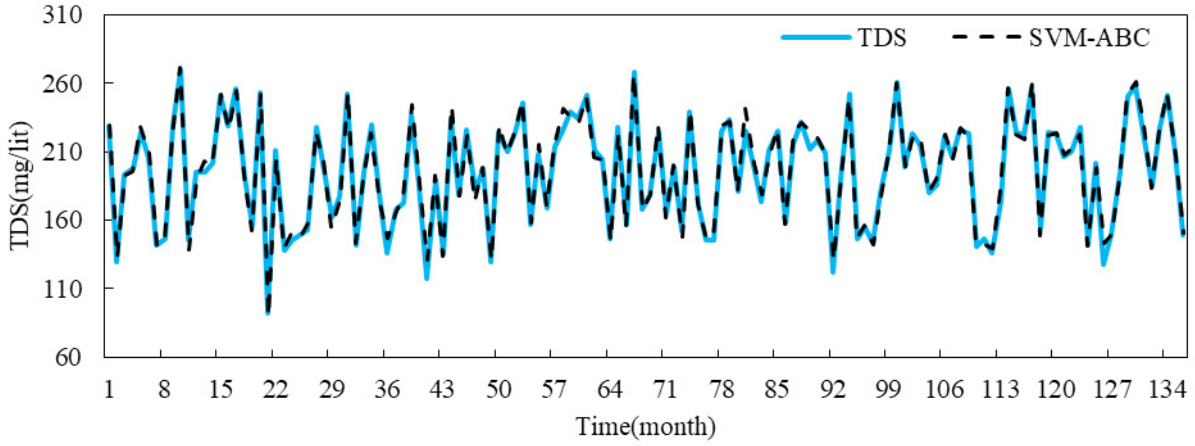
Table 4. Evaluation of hybrid and Bayesian network models, Training and verification stages, Parvarij-Abad hydrometric station-Sefidrood

صحّت‌سنجی Test			آموزش Train			پارامترهای ورودی Input parameters	مدل Model
$N.S_E$	RMSE (mg/l)	R^2	$N.S_E$	RMSE (mg/l)	R^2		
0.992	5.289	0.989	0.993	5.945	0.995	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-ABC
0.973	9.201	0.971	0.987	8.232	0.989	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-IWO
0.992	7.867	0.988	0.992	6.458	0.994	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	SVM-PSO
0.992	9.357	0.988	0.993	5.98	0.994	Na,Mg,So_4,Cl,Hco_3,Ca,pH,Ec	LS-SVM



شکل٦- پیش‌بینی و دقّت مدل SVM-ABCدر بخش صحّت سنجی ایستگاه پروریج‌آباد

Fig6. Predicted data and accuracسش, SVM-ABC model, Testing stage, Parvarij-Abad hydrometric Station



شکل٧- مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده مقادیر بهینه مدل SVM-ABC برای داده‌های صحّت‌سنجی در بابلرود ایستگاه قرآن تالار

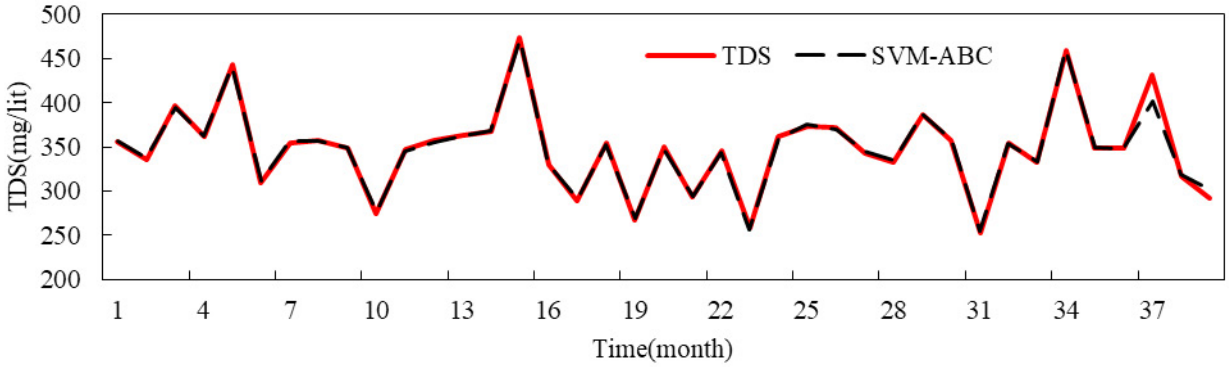
Fig7.Comparison between measured and predicted data, optimal values, SVM-ABC model, verification stage Babolrood Quran-Talar hydrometric Station

بر اساس جدول (٤) برای ایستگاه هیدرومتری پروریج‌آباد بهترین مدل ایجادشده، SVM-ABC می‌باشد. در شکل (٦) روند تغییرات داده‌های مشاهداتی و محاسبه‌شده توسط مدل بهینه در ایستگاه هیدرومتری پروریج‌آباد همچنین میزان پراکندگی خطاها ارائه‌شده است.

مقایسه عملکرد مدل‌ها

همان‌طور که از جداول(٣) و (٤) مشاهده می‌شود با انتخاب جواب بهینه هرکدام از مدل‌ها و مقایسه آن‌ها با یکدیگر در ایستگاه هیدرومتری قرآن تالار و پروریج‌آباد مشخص شد که هر چهار مدلSVM-ABC، SVM-IWO، SVM-PSO، LS-SVMبا دقّت خوبی و خطای کم‌تر می‌توانند مواد جامد محلول آب TDS را در رودخانه‌های بابلرود و سفیدرود شبیه‌سازی کنند. با توجّه به جدول (٣) مشاهده می‌شود در رودخانه بابلرود ایستگاه هیدرومتری قرآن تالار مدل SVM-ABC بیشترین دقّت و کمترین ریشه میانگین مربعات خطا در مرحله آموزش و صحّت‌سنجی را دارا می‌باشد. در شکل (٧) نتایج مقادیر مشاهداتی و محاسباتی مقادیر بهینه مدل SVM-ABC بخش صحّت‌سنجی در ایستگاه هیدرومتری قرآن تالار ارائه‌شده است.

با توجّه به جدول(٤) برای شبیه‌سازی TDS رودخانه سفیدرود ایستگاه هیدرومتری پروریج‌آباد مدل SVM-ABC با بیشترین دقّت و کمترین ریشه میانگین مربعات خطا در مرحله آموزش و صحّت سنجی دارا می‌باشد. در شکل (٨) مقادیر مشاهداتی و محاسباتی مقادیر بهینه مدل SVM-ABC بخش صحّت‌سنجی در ایستگاه هیدرومتری پروریج‌آباد را نشان می‌دهد.



شکل۸- مقایسه مقادیر مشاهداتی و پیش‌بینی‌شده مقادیر بهینه SVM-ABCمرحله صحّت‌سنجی سفیدرود ایستگاه پروریج‌آباد

Fig8. Comparison between measured and predicted data, optimal values, SVM-ABC model, verification stage, Parvarij-Abad hydrometric Station

بحث و نتیجه‌گیری

در این پژوهش کارایی پنج مدل داده محور LS-SVM، SVM-ABC، SVM-IWO، SVM-PSO در تخمین پارامتر مهم کیفیّت آب TDS،که اندازه‌گیری و کار آزمایشگاهی مربوط به آن پرهزینه و زمان‌براست، در ایستگاه‌های هیدرومتری قرآن تالار و پروریج‌آباد به ترتیب در بابلرود و سفیدرود ارزیابی شد. بدین منظور از داده‌های ماهانه کلسیم، منیزیم، بی‌کربنات، سدیم، سولفات، EC، pH به عنوان پارامترهای ورودی استفاده شد. در بخش مدل‌سازی ایستگاه قرآن تالار و در هردو بخش آموزش و صحّت‌سنجی SVM-ABCبهترین عملکرد را نسبت به روش‌های دیگر ارائه دادند. مدل‌های SVM-IWO، SVM-PSO، LS-SVM در مرتبه‌های بعدی قرار گرفتند. همچنین می‌توان گفت مدل‌های فرا ابتکاری با توجّه به کارایی مناسب در تخمین پارامترهای کیفیّت آب نسبت به مدلLS-SVM می‌توانند در اولویت گیرند.

تشکر و قدردانی

بدین‌وسیله از دانشگاه تهران و شرکت مدیریت منابع آب ایران به دلیل تأمین امکانات و داده‌های لازم جهت انجام این تحقیق و تهیه مقالات مربوطه تشکر و قدردانی می‌شود.

منابع

1. Adib, A. Farajpanah, H. Mahmoudian Shoushtari, M. Ahmadeanfar, I. 2020. Estimation of water quality parameters in the Sepidrood River by ANFIS, GEP and LS-SVM models. Journal of Water and Wastewater 31(5), 1-10. (In Persian)

2.AkhoniPourhosseini, F. Asadi, E.2017. Bayesian networks, Gamma Test, Groundwater level, model Least Squares Support Vector Machine, Plain Ardebil. Iran-Watershed Management Science & Engineering 11 (36) ,33- 42. (In Persian)

3.AkhoniPourhosseini, F. Ebrahimi, K. 2019. Shannon's Entropy Evaluation on Determination of Surface Water Quality) Case Studies: Karun and Babolrood Rivers). Journal of Water and Irrigation Management 2(9), 172-183. (In Persian)

4. Diamantopoulou, M.J. Antonopoulos, V.Z. Papamichail, D.M. 2005. The use of a neural network technique for the prediction of water quality parameters of Axios River in Northern Greece. Journal of Eur. Water. 11:55–62.

5. Dehghani, R. Ghorbani, M.A. Teshnehlab, M. RikhtehgarGheasi, A. Asadi, E. 2015. Comparison and Evalution of Bayesian Neural Network, Gene gramming, Support Vector Machine and Multiple expression pro Linear Regression in River Discharge Estimation (Case Study: Sufi Chay Basin).Journal of Irrigation Engineering and Water 20(5), 67-85. (In Persian).

6.Ehteram, M. Ferdowsi, A. Faramarzpour, M. Sami Al-Janabi, A.M. Al-Ansari, N. Dhanraj Bokde, N. Mundher Yaseen, Z. 2021. Hybridization of artificial intelligence models with nature inspired optimization algorithms for lake water level prediction and uncertainty analysis. Alexandria Engineering Journal. 60: 2193–2208.

7. Faryadi, S. Shahedi, K. Nabatpoor, M.2013. Investigation of Water Qualy Parameters in Tajan River using Multivariate Statistical Techniques. Journal of Watershed Management 6(3):75-92. (In Persian).

8. Karaboga, D. 2005. An idea based on honey bee swarm for numerical optimization. Technical Report-TR06, Department of Computer Engineering, Engineering Faculty. Erinyes University.

9. Kavosi, M. Khazinehneghad, H. 2021. Review and compare performance of 4 modeling methods LS-SVM, NN, GEP and ANFIS-PSO in Simulation of Rainfall - Runoff (Study Area: Halil River - Jiroft Dam). Iranian Journal of Irrigation and Water Engineering11 (43):96-110. (In Persian)

10. Kavosi, M. Khasheesioki, A. Pourrezabilandi M. Najafi,

M.H. 2018. Application of the new LSSVM-PSO simulation algorithm in the design of an optimal groundwater level monitoring network.Journal of Echo Hydrology 4(5). 13.9-1319. (In Persian)

11. Karimipour, A.R. Banitalebi, G.2020. Sensitivity analysis of meteorological data in estimating reference evapotranspiration with the minimum data using wavelet-neuro-fuzzy, ANN and ANFIS models. Journal of Water and soil resources conservation 9(3):48-70. (In Persian)

12. Mehrdadi, N. Aref, H. AsgharniAkbariya, A.2010. Investigation of the effects of environmental pollution on water qualy of Babolrood River. The first national conference on water resources management. Sari University, Iran. (In Persian).

13. MacKay, D.J.C. 1992. Bayesian Interpolation, Neural Computation, 4(2): 415-447.

14. Mirhashemi, H. Yarahmadi, D. Sharifi, S. Farzin, S. 2019. Improvement of Support Vector Machine and Random Forest Algorithm in Predicting Khorramabad River Flow Using Non-uniform De-Noising of data and Simplex Algorithm. Iran-Watershed Management Science & Engineering 13 (47):40-51. (In Persian).

15. ShananiHoveyzeh, M. Zarei, H.2016. Comparison of Three Classification Algorithms (ANN, SVM and Maximum Likelihood) for Preparing Land Use Map (Case Study: Abolabbas Basin). Iran-Watershed Management Science & Engineering 10 (33):73-84. (In Persian)

16. Mehrabian, A. R. Lucas, C. 2007. A Novel Numerical Optimization Algorithm Inspired from Weed Colonization. Journal of Ecological Informatics. 2(1): 355–366.

17. Rama, D. Jayabarathi, T. 2016. Optimal Placement and Sizing of Multiple Distributed Generating units in Distribution Networks by Invasive Weed Optimization Algorithm. Ain Shams Engineering Journal. 7(2): 683-694.

18. Shahinejad, B. Dehaghani, R.2018.Comparison of Wavelet Neural Network Models, Support Vector Machine, and Gene Expression Programming in Estimating the amount of Oxygen Dissolved in River.Iran-Water Resources Research 14(3):268-277. (In Persian)

19. Sharifi Garmdareh, E. Vafakhah, M. Eslamian, S. 2019. Assessment the Performance of Support Vector Machine and Artificial Neural Network Systems for Regional Flood Frequency Analysis (A Case Study: Namak Lake Watershed). Journal of Water and Soil Science 23 (1):351-366. (In Persian)

20. Sanchez, A.S. Nieto, P.J.G. Fernandez,P.R. Diaz, J.J.D. Iglesias-Rodr, F.J.2011. Application of an SVM-based regression modelto the air quality study at local scale in theAvilés urban area (Spain). Journal of Mathematical andComputer Modelling 54:1453–1466.

نشریه علمی
علوم و مهندسی آبخیزداری ایران
Iran-Watershed Management
Science & Engineering

Vol. 16, No. 58, Fall 2022
سال شانزدهم- شماره ۵۸- پاییز ۱۴۰۱

Evaluation of Intelligent Models in Water Quality Simulation of the Babolrood and Sefidrood Rivers, Iran

F. Akhoni Pourhosseini[1], K. Ebrahimi[2] and M. H. Omid[3]
Received: 02-03-2022 Accepted:18-06-2022

Abstract

The amount of Total Dissolved Solids (TDS) is an important factor in determining the water quality of rivers. The purposes of this paper were to investigate the efficiency of four intelligent models including, SVM-IWO, SVM-PSO, SVM-ABC and LS-SVM, in predicting water quality of Babolrood and Sefidrood rivers, north of Iran. To achieve this aim, unpublished measured monthly data including; Ca, Mg, HCO3, Na, SO4, EC, pH, and TDS from Quran-Talar station of Babolrood river and Parvarijabad station of Sefidrood river were analyzed in the periods of 1978-2015 and 2004-2015, respectively. Evaluation of the four above-mentioned models, based on the correlation coefficient (R2), Root Mean Square Error (RMSE)and the Nash-Sutcliff coefficient (NSE) showed that the SVM-ABC model in both Babolrood and Sefidrood rivers with the highest R2s equal to 0.985 and 0.989 and the least amounts of RMSE equal to 10.493mg/l and 5.289 mg/l and the highest Nash-Sutcliffe equal to 0.983 and 0.992, respectively, has better and faster performance compared to others models in assessing water quality.

Keywords: Calibration, Classification, Modelling, the environment, Water resources

1.Ph.D student of Water Resources Engineering, Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj, Iran
2.Corresponding Author and Professor, Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj, Iran. Email: EbrahimiK@ut.ac.ir
3. Professor, Department of of Irrigation and Reclamation Engineering, University of Tehran, Karaj, Iran