

برآورد مقدار جامدات محلول رودخانه با استفاده از محاسبات نرم (مطالعه موردی: بالادست سد بوکان)

سروین زمان زاد قويدل^۱ - کامران زینال زاده^{۲*}

تاریخ دریافت: ۱۳۹۳/۰۹/۰۲

تاریخ پذیرش: ۱۳۹۴/۰۳/۰۴

چکیده

زندگی بشر در کره زمین در گرو تأمین آب مناسب از نظر کمی و کیفی قرار دارد. در سال‌های اخیر استفاده از مدل‌های ریاضی برای شبیه‌سازی کیفیت آب رودخانه‌ها توسعه زیادی یافته که به دلیل پیچیدگی و تعدد فرآیندهای کیفی منابع آب سطحی، استفاده از محاسبات نرم روشی نو جهت پیش‌بینی کیفیت آب رودخانه‌ها می‌باشد. در این مطالعه، قابلیت روش شبکه‌های عصبی مصنوعی با الگوریتم آموزشی لونیگ-ماکواریت، سامانه استنتاجی عصبی-فازی تطبیقی مبتنی بر مدل دسته‌بندی تفریقی و برنامه‌ریزی بیان ژن برای مدل‌بندی خصوصیات کیفیت آب رودخانه در بالادست سد مخزنی بوکان واقع در حوضه آبریز زرینه‌رود و برآورد میزان جامدات محلول (TDS) ارزیابی و مقایسه شدند. بدین منظور از داده‌های کیفیت آب دو ایستگاه هیدرومتری آنیان و صفاخانه در بالادست سد مخزنی بوکان با طول دوره آماری ۱۸ ساله (۱۳۷۲-۱۳۸۹) استفاده گردید. ابتدا در یک آنالیز اولیه آماری بر اساس رگرسیون گام‌به‌گام، پارامترهای مؤثر و اصلی کیفیت آب در پیش‌بینی TDS تعیین گردید. نتایج بیانگر عملکرد قابل قبول هر سه روش شبکه عصبی مصنوعی، سامانه استنتاجی عصبی-فازی تطبیقی و برنامه‌ریزی بیان ژن برای مدل‌بندی میزان جامدات محلول در بالادست سد می‌باشد. برنامه‌ریزی بیان ژن، سامانه استنتاجی عصبی-فازی تطبیقی و شبکه عصبی مصنوعی با پنج پارامتر ورودی بیشترین دقت را داشتند. برنامه‌ریزی بیان ژن علاوه بر عملکرد بالا، در ایستگاه‌های آنیان و صفاخانه به ترتیب با ضرایب همستگی ۰/۹۶۲ و ۰/۹۷۱ و ریشه میانگین مربعات خطا ۱۲/۸۲ mg/L و ۲۹/۰۸ mg/L، روابط ریاضی جهت برآورد مقدار جامدات محلول در رودخانه‌های بالادست سد را نیز ارائه داد.

واژه‌های کلیدی: بیان ژن، جامدات محلول، زرینه‌رود، شبکه عصبی مصنوعی، عصبی-فازی

مقدمه

مدل‌های شبیه‌سازی در قالب مدل‌های قطعی و غیرقطعی ارائه می‌شوند. در بخش مدل‌های شبیه‌سازی قطعی، مدل‌های هوشمند و در بخش مدل‌های شبیه‌سازی غیرقطعی، مدل شبیه‌سازی مونت کارلو را می‌توان اشاره کرد. همچنین، در ارتباط با مدل‌های بهینه‌سازی می‌توان به مدل‌های تک هدفه، مدل‌های چند هدفه، مدل‌های حل اختلاف و مدل‌های نوین بهینه‌سازی اشاره نمود. بیشتر این مدل‌ها نیازمند پارامترهای ورودی بسیار زیادی هستند که یا دسترسی به آن‌ها مشکل است و یا نیازمند صرف هزینه و زمان زیادی می‌باشد. در سال‌های اخیر استفاده از مدل‌های هوش مصنوعی^۳ در زمینه‌های مختلف مهندسی آب مورد توجه بسیاری از محققین قرار گرفته است (۷، ۸، ۱۰ و ۲۰). شبکه‌های عصبی مصنوعی^۴، سامانه استنتاج عصبی-فازی تطبیقی^۵ و برنامه‌ریزی بیان ژن^۶ نمونه‌های بارزی از

امروزه در دنیا آب و منابع آب، یکی از پایه‌های اصلی توسعه پایدار به شمار می‌روند. در مواردی که رودخانه‌ها به‌عنوان یکی از منابع تأمین کننده نیاز جوامع بشری مطرح می‌گردند، علاوه بر کمیت و میزان آورد رودخانه، کیفیت آب نیز جزء پارامترهای مهم مورد توجه قرار می‌گیرد. همچنین پارامترهای کیفی آب یکی از مولفه‌هایی می‌باشند که در برنامه‌ریزی‌ها بایستی به‌دقت شبیه‌سازی شده و تخمین زده شود (۱۷). با توجه به اینکه بحران آب و کاهش کیفیت آن از مهم‌ترین دغدغه‌ها و مشکلات کشورهای مختلف جهان خصوصا در دو دهه اخیر گردیده است، لذا مطالعه و تحقیق در این زمینه اهمیت فراوانی دارد. در زمینه مدیریت بهتر برای حفظ کیفیت آب مدل‌های بسیار زیادی گسترش یافته است. مدل‌های شبیه‌سازی و بهینه‌سازی کیفیت آب سطحی، طیف وسیع و گسترده‌ای دارند.

3- Artificial Intelligence Methods

4- Artificial Neural Networks (ANNs)

5- Adaptive neuro fuzzy inference system (ANFIS)

6- Gene expression programming

۲- دانشجوی دکتری و استادیار گروه مهندسی آب، دانشگاه ارومیه

(Email: k.zeinalzadeh@urmia.ac.ir)

مدل‌های هوش مصنوعی می‌باشند. برعکس بسیاری از مدل‌های پایه آماری کیفیت آب، که با فرض رابطه خطی بین مقادیر واقعی و مقادیر تخمینی عمل می‌نمایند، روش‌های مذکور که جز روش‌های محاسبات نرم^۱ (برگرفته از طبیعت) می‌باشند، قادر به ارائه رابطه غیرخطی حاکم بر مسئله می‌باشند. تاکنون تحقیقات متعددی در زمینه روش‌های هوشمند برای تخمین دقیق مقدار پدیده‌های مختلف هیدرولوژیکی از جمله برآورد پارامترهای کیفیت آب، میزان بار رسوب رودخانه، تبخیر از تشت، تبخیر و تعرق پتانسیل گیاه مرجع، بارندگی، رواناب، تغییرات سطح تراز دریاچه و ... انجام شده است.

منتصری و زمان‌زاد قویدل (۱۸) عملکرد دو مدل هوشمند شامل مدل فازی-عصبی تطبیقی مبتنی بر روش افراز شبکه و مدل برنامه‌ریزی بیان ژن را در پیش‌بینی جریان رودخانه در حوضه آبریز زربینه رود مورد بررسی قرار دادند. نتایج به دست آمده حاکی از آن بود که هر دو روش فازی-عصبی و برنامه‌ریزی بیان ژن می‌توانند به طور موفقیت‌آمیزی جهت پیش‌بینی جریان رودخانه مورد استفاده قرار گیرند. در این راستا مدل برنامه‌ریزی بیان ژن به دلیل ارائه رابطه ریاضی حاکم بر مساله مورد بررسی نسبت به مدل فازی-عصبی تطبیقی برتری داشت. آیتک و آلپ (۲) برای مدل‌سازی فرآیند بارش-رواناب از برنامه‌ریزی بیان ژن و دو مدل مختلف از شبکه‌های عصبی مصنوعی استفاده کردند. آن‌ها برای انجام مدل‌سازی از داده‌های سه ایستگاه باران‌سنجی و یک ایستگاه هیدرومتری در حوضه‌ای از ایالت پنسیلوانیا کشور آمریکا استفاده کردند. نتایج به دست آمده دقت و کارایی مدل برنامه‌ریزی ژنتیک را در مقایسه با مدل‌های عصبی مصنوعی نشان داد. آیتک و کیشی (۳) روش برنامه‌ریزی ژنتیک را جهت شبیه‌سازی بار معلق در رودخانه تانگو^۲ واقع در ایالات متحده آمریکا بکار گرفتند. آن‌ها از مقادیر رسوب قبلی و جریان جهت پیش-بینی بار معلق استفاده کردند. نتایج حاصل شده با روش برنامه‌ریزی ژنتیک حاکی از عملکرد بهتر آن در مقایسه با روش‌های منحنی سنج و رگرسیون خطی چندگانه بود.

گاون و گونال (۹) به منظور تخمین میزان آبشستگی در پایین-دست سازه‌های هیدرولیکی بدون بعد به‌عنوان ورودی مدل، استفاده کردند. نتایج تحقیق آن‌ها نشان از عملکرد بهتر روش برنامه‌ریزی ژنتیک در مقایسه با روابط استخراج شده با روش رگرسیون خطی و غیرخطی بود. ساین و همکاران (۲۴) از شبکه‌های عصبی مصنوعی با استفاده از کلیه پارامترهای کیفیت آب در دسترس برای مدل‌سازی میزان اکسیژن محلول و اکسیژن مورد نیاز بیولوژیکی در رودخانه گمتی هند استفاده کردند. نتایج تحقیق نشان داد که روش شبکه

عصبی مصنوعی به‌عنوان یکی از روش‌های قابل استفاده با عملکرد مناسب برای برآورد پارامترهای کیفیت آب رودخانه‌ها به‌شمار می‌رود. نجاح و همکاران (۱۹) توانایی شبکه عصبی را در پیش‌بینی شاخص‌های کیفی آب رودخانه جوهر مالزی و توانایی برآورد مقدار EC و TDS و کدورت در این حوضه را مورد تأکید قرار دادند. هشمی و همکاران (۱۱) به‌منظور ریزمقیاس نمودن مقادیر بارش روزانه در حوضه رودخانه کلوسا در کشور کانادا از مدل برنامه‌ریزی بیان ژن استفاده نمودند. بدین منظور آن‌ها از ۲۶ متغیر اقلیمی مختلف در ورودی مدل برنامه‌ریزی بیان ژن استفاده نمودند و بهترین ترکیب ورودی را استخراج کردند. مقایسه نتایج مدل برنامه‌ریزی بیان ژن با مدل‌های آماری ریزمقیاس نمایی نشان از عملکرد و دقت بالای روش برنامه‌ریزی بیان ژن داشت. تراوره و گاون (۲۵) از مدل برنامه‌ریزی بیان ژن جهت تخمین تبخیر و تعرق در کشور ساحلی بورکینافاسو استفاده نمودند. آن‌ها از داده‌های هواشناسی شش ایستگاه واقع در نواحی مختلف کشور استفاده کردند و جهت مدل‌سازی، ترکیبات ورودی مختلف داده‌های هواشناسی را بکار بردند. نتایج به‌دست‌آمده حاکی از دقت بالای روش برنامه‌ریزی در فرآیند مدل‌سازی تبخیر و تعرق بود. شیری و همکاران (۲۳) با مقایسه روش برنامه‌ریزی ژنتیک و سامانه استنتاجی عصبی-فازی تطبیقی جهت مدل‌سازی تبخیر تعرق مرجع روزانه به این نتیجه رسیدند که عملکرد برنامه‌ریزی ژنتیک در این زمینه بهتر از سامانه استنتاجی عصبی-فازی می‌باشد. کیشی و همکاران (۱۴) جهت پیش‌بینی تغییرات سطح تراز دریاچه از برنامه‌ریزی ژنتیک، سامانه استنتاجی عصبی-فازی تطبیقی، شبکه عصبی مصنوعی و روش میانگین متحرک استفاده کردند. نتایج بیانگر بهتر بودن برنامه‌ریزی ژنتیک نسبت به سایر روش‌ها بود. کیشی و همکاران (۱۵) برای مدل‌سازی بارش-رواناب حوضه آبریز در کشور ترکیه از روش‌های محاسبات نرم (شبکه‌های عصبی مصنوعی، سامانه استنتاجی عصبی-فازی تطبیقی و برنامه-ریزی ژنتیک) استفاده کردند. نتایج، برتری برنامه‌ریزی ژنتیک را نسبت به سایر روش‌ها نشان داد. هم‌چنین این روش معادله ریاضی حاکم بین ورودی‌ها و خروجی‌های مدل را در اختیار قرار داد.

سد مخزنی بوکان تأمین کننده آب شرب شهر بوکان می‌باشد و کیفیت آب مخزن این سد از اهمیت فوق‌العاده‌ای برخوردار است. در این مطالعه مقدار جامدات محلول (TDS) بالادست سد مخزنی بوکان در دو ایستگاه آنیان و صفاخانه با استفاده از روش شبکه‌های عصبی مصنوعی با الگوریتم آموزشی لوبنبرگ-مارکواریت^۳، سامانه استنتاجی عصبی-فازی تطبیقی مدل دسته‌بندی تفریقی^۴ و برنامه‌ریزی بیان ژن برآورد می‌گردد. جهت ارزیابی نتایج مدل‌های مورد مطالعه،

مقدار جامدات محلول (TDS) آب رودخانه روش‌های شبکه عصبی مصنوعی و عصبی-فازی تطبیقی در محیط نرم‌افزاری MATLAB کدنویسی شده و برنامه‌ریزی بیان ژن با استفاده از برنامه GeneXpro به کار گرفته شده است.

روش شبکه‌های عصبی مصنوعی (ANNs)

شبکه‌های عصبی مصنوعی بر اساس ارتباطات داخلی بسیار گسترده، همانند سیستم عصبی و مغز انسان بنا نهاده شده‌اند. شبکه‌های عصبی مصنوعی جزء سیستم‌های دینامیکی می‌باشند که با پردازش داده‌های تجربی، قانون نهفته در ورای اطلاعات را به ساختار شبکه منتقل می‌کنند. به همین خاطر آن‌ها را سیستم‌های هوشمند می‌نامند (۱۲). ساختار کلی مدل‌های شبکه عصبی مصنوعی مطابق شکل ۲ از سه لایه تشکیل شده است: (۱) لایه ورودی: در این لایه اطلاعات ورودی به مدل معرفی می‌شوند. (۲) لایه و یا لایه‌های مخفی: در این لایه اطلاعات پردازش می‌شوند. (۳) لایه خروجی: نتایج حاصل از عملکرد مدل تولید می‌شود. ساختار یک شبکه عصبی مصنوعی با تعیین تعداد لایه‌ها، تعداد نرون‌ها در هر لایه، تابع محرک (کنترل-کننده خروجی هر نرون)، روش آموزش، الگوریتم تصحیح وزن‌ها و نوع مدل، تعیین می‌شود (۴).

در این مطالعه از یک ساختار سه لایه (لایه ورودی، پنهان و خروجی) و الگوریتم آموزشی لونیبرگ-مارکوارت، تابع محرک متغیر لوگ-سیگموئید^۲ و تانژانت-سیگموئید^۳ در لایه پنهان، تابع محرک خطی^۴ در لایه خروجی، تعداد متغیر نرون‌ها و تعداد تکرار ۱۰۰۰ به منظور مدل‌بندی مقدار جامدات محلول در آب حوضه آبریز زرينه‌رود به ازای پارامترهای مختلف کیفیت آب و جریان رودخانه‌ای مطابق نمودار شماتیک شکل ۲ استفاده گردید.

سامانه استنتاجی عصبی-فازی تطبیقی (ANFIS)

سامانه استنتاجی عصبی-فازی تطبیقی برای اولین بار توسط جانگ (۱۹۹۳) معرفی شد (۱۳). این سامانه ابزاری کارآمد جهت تخمین توابع پیوسته حقیقی در دامنه‌ای محدود می‌باشد که به ازای هر درجه‌ای از صحت قابل استفاده است. این سیستم پیاده‌سازی یک سیستم فازی سوگنو^۵ به صورت ساختار شبکه‌ای پیش‌رونده است (۲۲).

اگر فرض کنیم یک سامانه استنتاج فازی دارای دو ورودی Ca و HCO_3 و یک خروجی TDS باشد (مطابق شکل ۳)، مدل فازی مرتبه اول سوگنو با دو قانون فازی اگر-آنگاه به صورت زیر بیان می-

پارامترهای آماری ضریب همبستگی، ریشه میانگین مربعات خطا و میانگین قدرمطلق خطا استفاده شدند. در این مطالعه، روش‌های محاسبات نرم به‌ویژه سامانه استنتاجی فازی-عصبی تطبیقی و برنامه-ریزی بیان ژن که به‌ندرت در برآورد پارامترهای کیفیت آب به کار گرفته شدند، به‌عنوان جایگزین مدل‌های معمول در شبیه‌سازی پارامترهای کیفی آب، استفاده و مورد توجه قرار گرفته‌اند.

مواد و روش‌ها

مشخصات حوضه آبریز و داده‌ها

حوضه آبریز زرينه‌رود در موقعیت جغرافیایی ۴۷° ۲۴' تا ۴۵° ۴۷' طول شرقی و ۳۵° ۴۱' تا ۳۷° ۴۳' عرض شمالی، با وسعتی معادل ۱۳۶۸۵ کیلومترمربع در محل تخلیه به دریاچه ارومیه، در جنوب دریاچه مذکور واقع شده است. رودخانه زرينه‌رود به‌عنوان آبراهه اصلی این حوضه آبریز، یکی از بزرگ‌ترین و پرآب‌ترین رودخانه‌های حوضه آبریز دریاچه ارومیه و یکی از رودخانه‌های اصلی کشور است. در این مطالعه از داده‌های کیفیت آب ایستگاه‌های هیدرومتری آبیان و صفاخانه، واقع در بالادست سد مخزنی بوکان در طول دوره نمونه-برداری ۱۸ ساله (۱۳۸۹-۱۳۷۲) استفاده شده است. موقعیت جغرافیایی و مشخصات عمومی ایستگاه‌های هیدرومتری در شکل ۱ و جدول ۱ ارائه شده است.

جدول ۲ مشخصات آماری داده‌های کیفیت آب و جریان رودخانه‌ای در دو ایستگاه آبیان و صفاخانه را نشان می‌دهد. مقادیر این جدول نشان می‌دهد که حداکثر جامدات محلول و سایر پارامترهای کیفیت آب در ایستگاه صفاخانه بیش‌تر از ایستگاه آبیان می‌باشد. برای تعیین پارامترهای ورودی مدل‌ها جهت تخمین مقدار (TDS) ابتدا ضرایب همبستگی پیرسون در سطح معنی‌داری ۹۵ درصد با استفاده از نرم‌افزار SPSS، بین پارامترهای کمی و کیفی آب در دو ایستگاه مورد مطالعه محاسبه گردید. سپس پارامترهایی که دارای همبستگی معنی‌دار بودند بر اساس رگرسیون گام‌به‌گام^۱ به ترتیب اولویت تأثیر بر مقدار جامدات محلول رودخانه (ضریب همبستگی مابین پارامترهای ورودی و مقدار جامدات محلول) به عنوان ورودی مدل‌ها تعیین شدند. جهت آماده‌سازی داده‌ها برای ورود به شبکه عصبی مصنوعی، ابتدا داده‌های منتخب با استفاده از رابطه ۱ نرمال‌سازی شدند.

$$X_N = \frac{X_R - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

که در آن، X_N : داده نرمال شده، X_R : مقدار داده مشاهده‌ای، X_{\min} : کم‌ترین مقدار داده‌های مشاهده‌ای و X_{\max} : بیشترین مقدار داده‌های مشاهده‌ای می‌باشند. بعد از آماده‌سازی داده‌های اولیه برای مدل‌بندی

2- Log-Sigmoid

3- Tan-Sigmoid

4- Linear Transfer Function

5- Sugeno

1- Stepwise Regression

شود:

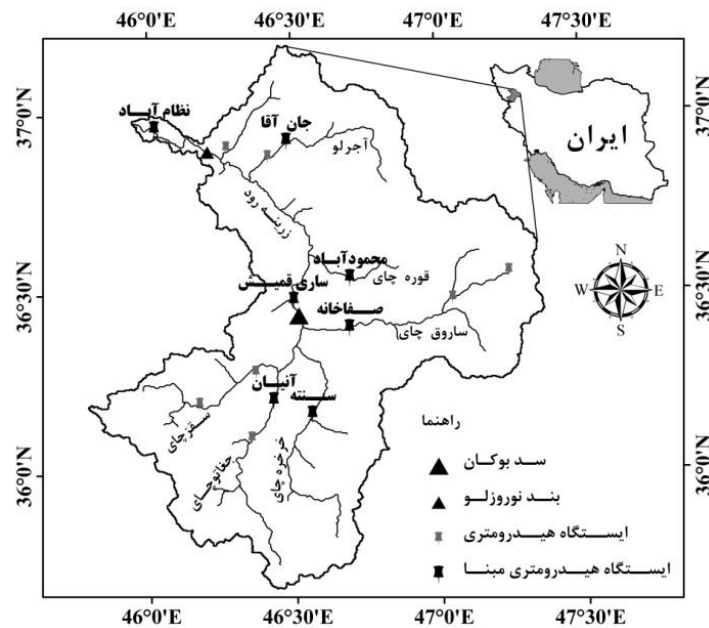
HIGH (زیاد) باشد آنگاه:

$$TDS = p_1HCO_3 + q_1Ca + r_1 \quad (2) \quad \text{قانون ۱: اگر } HCO_3 \text{ در حالت LOW (کم) و Ca در حالت}$$

جدول ۱ - مشخصات عمومی حوضه آبریز زرينه رود در ايستگاه‌های هيدرومتری مورد مطالعه

Table 1- The Characteristics of studied hydrometric stations in the Zarinehroud basin

ايستگاه Station	رودخانه River	مساحت Area(km ²)	ارتفاع از سطح دريا Elevation(m)	طول جغرافيايي Longitude (°E)	عرض جغرافيايي latitude(°N)	تعداد نمونه برداری ها Number of sampling
آنيان Anyan	جیغاتوچای Jighatuchay	۱۲۲۱	۱۴۶۰	۴۶ ۲۶'	۳۶ ۱۲'	۱۱۸
صفاخانه Safakhaneh	ساروق چای Saroughchay	۲۲۱۹	۱۴۷۵	۴۶ ۴۲'	۳۶ ۲۴'	۱۱۴



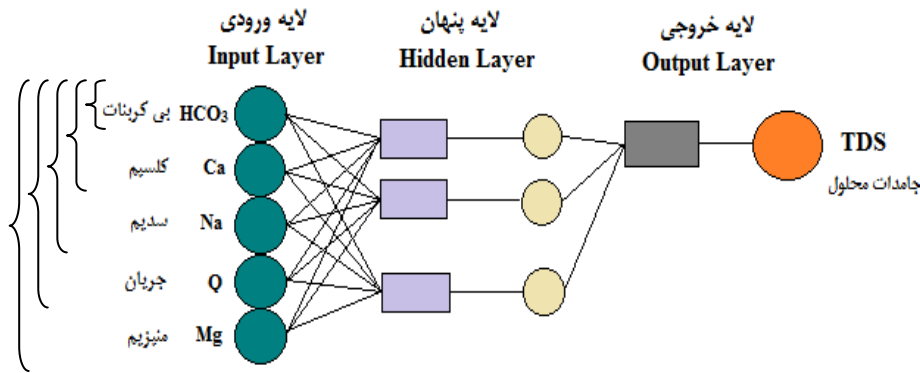
شکل ۱- موقعیت جغرافیایی منطقه مورد مطالعه

Figure 1- The location of the study area

جدول ۲- پارامترهای آماری مربوط به داده‌های ایستگاه‌های هيدرومتری آنيان و صفاخانه

Table 2- Statistical parameters of Anyan and Safakhaneh stations data

Parameters	پارامترها	Na	Mg	Ca	HCO ₃	Q	TDS
Unit	واحد	(meq.L ⁻¹)	(meq.L ⁻¹)	(meq.L ⁻¹)	(meq.L ⁻¹)	(m ³ .s ⁻¹)	(mg.L ⁻¹)
ايستگاه آنيان Anyan Station	Min	۰/۱۰	۰/۱۰	۱/۰	۰/۹۰	۰/۰۱	۱۰۴/۰
	Max	۰/۷۵	۲/۵۰	۴/۰۰	۴/۹۰	۳۶۳/۹۶	۳۴۴/۵۰
	Average	۰/۲۷	۰/۶۵	۲/۱۹	۲/۴۳	۳۲/۱۲	۱۷۷/۷۵
	Standard deviation	۰/۱۳	۰/۴۳	۰/۴۸	۰/۶۹	۵۳/۳۳	۴۷/۱۱
ايستگاه صفاخانه Safakhaneh Station	Min	۰/۱	۰/۶	۱/۵	۱/۷	۰/۰۳	۱۶۹
	Max	۳/۱	۵/۶	۵/۱	۵/۸	۱۷۵/۱۷	۶۵۰
	Average	۰/۸۹	۲/۲۵	۳/۳۷	۳/۷۶	۱۴/۶۱	۳۸۵/۵۱
	Standard deviation	۰/۵۵	۰/۸۸	۰/۷۸	۰/۸۴	۳/۶۶	۱۰۸/۱۴



شکل ۲- نمودار شماییک یک شبکه عصبی چند لایه با ورودی‌های مدل

Figure 2- ANN structure with inputs of model

داروین است. این الگوریتم‌ها اقدام به تعریف یک تابع هدف در قالب معیارهای کیفی نموده و سپس تابع یاد شده را برای مقایسه جواب‌های مختلف حل مسئله در یک فرآیند گام‌به‌گام تصحیح ساختار داده‌ها به کار می‌گیرند و در نهایت جواب مناسب را ارائه می‌نمایند. روش برنامه‌ریزی ژنتیک جدیدترین شیوه از بین روش‌های الگوریتم تکاملی می‌باشد که به دلیل دارا بودن دقت کافی از کاربرد بیشتری برخوردار است (۱ و ۶).

فرآیند گام‌به‌گام حل یک مسأله با استفاده از برنامه‌ریزی ژنتیک متشکل از پنج مرحله به شرح زیر می‌باشد: (۱) انتخاب مجموعه ترمینال، که همان متغیرهای مستقل مسأله و متغیرهای حالت سامانه می‌باشد، (۲) انتخاب مجموعه توابع؛ که شامل عملگرهای حسابی، توابع آزمون و توابع بولی می‌باشد، (۳) شاخص اندازه‌گیری دقت مدل که بر مبنای آن می‌توان مشخص نمود که توانایی یک مدل در حل یک مسأله خاص تا چه اندازه می‌باشد، (۴) مؤلفه‌های کنترل؛ مقادیر مؤلفه‌های عددی و متغیرهای کیفی که برای کنترل اجرای برنامه‌های به کار بسته می‌شوند.

قانون ۲: اگر HCO_3 در حالت LOW (کم) و Ca در حالت MEDIUM (متوسط) باشد آنگاه:

$$TDS = p_2 HCO_3 + q_2 Ca + r_2 \quad (3)$$

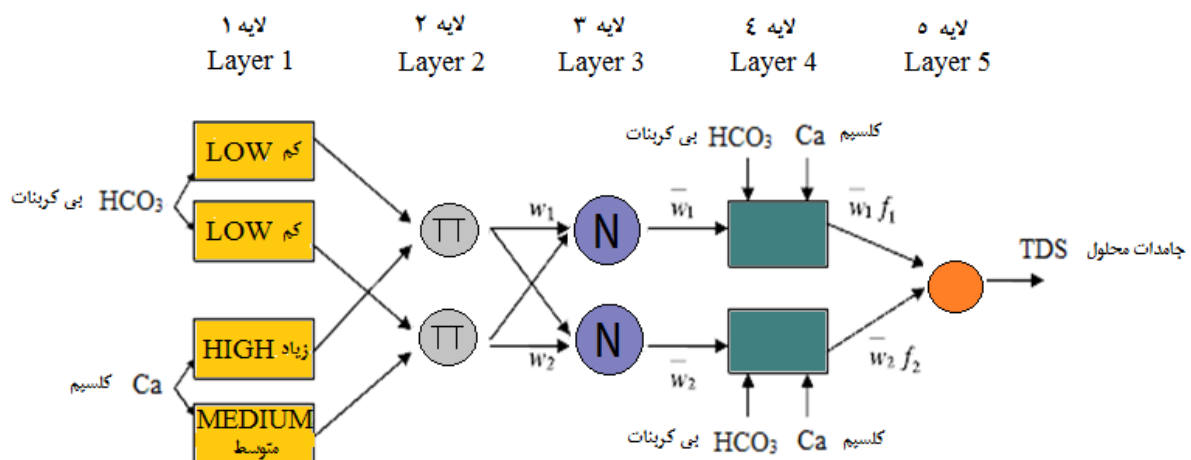
ساختار ANFIS دارای پنج لایه می‌باشد که در شکل ۳ نشان داده شده است.

روش آموزش اصلی در این سیستم، روش پس انتشار خطا می‌باشد (۲۲). مدل دسته‌بندی تفریقی ترکیبی از دسته‌بندی تفریقی و سامانه استنتاج عصبی-فازی تطبیقی می‌باشد. روش دسته‌بندی تفریقی شکل توسعه یافته روش دسته‌بندی نقاط مرتفع می‌باشد. چپو (۵) روش اصلاح شده دسته‌بندی نقاط مرتفع را جهت کاهش پیچیدگی‌های محاسباتی روش قبلی پیشنهاد نمود.

در این مطالعه، در روش سامانه استنتاجی عصبی-فازی تطبیقی با توجه به شعاع دسته متغیر با تعداد تکرار ۱۰۰ و ورودی‌های مشابه با شبکه عصبی مصنوعی، بهترین مدل‌ها تعیین گردید.

برنامه‌ریزی بیان ژن

روش برنامه‌ریزی ژنتیک جزو روش‌های الگوریتم گردشی محسوب می‌شود که مبنای این روش‌ها بر اساس نظریه تکمیل



شکل ۳- نمای شماییک مدل سامانه استنتاجی عصبی-فازی تطبیقی

Figure 3- Adaptive neuro-fuzzy inference systems structure

جدول ۳- پارامترهای به کار گرفته شده در برنامه ریزی بیان ژن
Table 3- The parameters of GEP model that used in

تعداد کروموزومها Number of chromosomes	۳۰	نرخ برگشت Inversion rate	۰/۱	نرخ ترکیب تک نقطه‌ای One point recombination rate	۰/۳
طول هر راس Head size	۸	نرخ ترانژیشن ژنی Gene transposition rate	۰/۱	نرخ ترکیب دو نقطه‌ای Two point recombination rate	۰/۳
تعداد ژن‌ها Number of Genes	۳	نرخ جهش Mutation rate	۰/۰۴	نرخ ترکیب ژنی Gene recombination rate	۰/۱

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (TDS_{io} - TDS_{ie})^2} \quad (5)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |TDS_{io} - TDS_{ie}| \quad (6)$$

که در آنها، TDS_{io} : مقادیر مشاهداتی، TDS_{ie} : مقادیر برآوردی، $\sqrt{\quad}$ ، $\sqrt[3]{\quad}$ ، \ln ، e^x ، x^2 ، x^3 ، $\sin x$ ، $\cos x$ ، $\arctag x$ ، $+$ ، $-$ ، \times ، \div در این مطالعه از عملگرهای اصلی جهت ساخت درخت تجزیه می‌باشد. در این مطالعه از عملگرهای $\sqrt{\quad}$ ، $\sqrt[3]{\quad}$ ، \ln ، e^x ، x^2 ، x^3 ، $\sin x$ ، $\cos x$ ، $\arctag x$ عملگرهای ریاضی پیش فرض مدل GeneXpro می‌باشند، استفاده به عمل آمد. به طور خلاصه پارامترهای استفاده شده در هر بار اجرای مدل جهت طی فرآیند مدل سازی مقدار جامدات محلول در بالادست سد مخزنی بوکان، که اغلب مقادیر پیش فرض نرم افزار می‌باشند، در جدول ۳ ارائه شده است.

نتایج و بحث

تعیین الگوهای ورودی مدلها

با محاسبه ضریب همبستگی بین جامدات محلول و سایر پارامترهای موجود، پنج الگوی ورودی مدل‌های شبیه سازی شده به ترتیب: ۱. بی کربنات (HCO_3)، ۲. بی کربنات و کلسیم (HCO_3 , Ca)، ۳. بی کربنات، کلسیم و سدیم (HCO_3 , Ca, Na)، ۴. بی کربنات، کلسیم، سدیم و میزان جریان در رودخانه (HCO_3 , Ca, Na, Q)، ۵. بی کربنات، کلسیم، سدیم، میزان جریان و منیزیم (HCO_3 , Ca, Na, Q, Mg) حاصل از روش رگرسیون گام به گام، در نظر گرفته شدند. در مرحله بعد، ۷۵ و ۲۵ درصد داده‌های ایستگاه‌های مورد نظر به طور تصادفی^۱ به ترتیب برای آموزش و آزمون مدل‌ها، در دو مجموعه تفکیک شده‌اند.

مقایسه نتایج مدلها

به منظور یافتن ساختار بهینه شبکه‌های عصبی مصنوعی، پارامترهایی همانند تعداد نرون‌ها (۱ تا ۱۰ نرون) در لایه پنهان، توابع محرک در لایه‌های پنهان و خروجی به روش سعی و خطا مورد آزمون قرار گرفت. بهترین مدل بر اساس حداقل میزان مربعات خطا، تعداد تکرار ۱۰۰۰ و سرعت آموزش ۰/۰۱ برای ایستگاه‌های مورد مطالعه تعیین گردید.

۵) شروط توقف اجرای برنامه؛ که معیاری برای حصول به نتایج و توقف اجرای برنامه می‌باشد. اولین مرحله در مدل سازی، انتخاب تابع برازش مناسب می‌باشد که در این تحقیق با بررسی انواع توابع برازش، تابع RMSE به عنوان تابع برازش برتر در جهت مدل بندی انتخاب شد. مرحله بعدی انتخاب عملگرهای اصلی جهت ساخت درخت تجزیه می‌باشد. در این مطالعه از عملگرهای $+$ ، $-$ ، \times ، \div ، $\sqrt{\quad}$ ، $\sqrt[3]{\quad}$ ، \ln ، e^x ، x^2 ، x^3 ، $\sin x$ ، $\cos x$ ، $\arctag x$ عملگرهای ریاضی پیش فرض مدل GeneXpro می‌باشند، استفاده به عمل آمد. به طور خلاصه پارامترهای استفاده شده در هر بار اجرای مدل جهت طی فرآیند مدل سازی مقدار جامدات محلول در بالادست سد مخزنی بوکان، که اغلب مقادیر پیش فرض نرم افزار می‌باشند، در جدول ۳ ارائه شده است.

معیارهای ارزیابی مدلها

جهت ارزیابی مدل‌های مورد استفاده از سه شاخص آماری معتبر و رایج شامل ضریب همبستگی (R)، ریشه میانگین مربعات خطا ($RMSE$) و قدر مطلق میانگین خطا (MAE) و یک روش گرافیکی (نمودار پراکندگی داده‌های برآوردی و مشاهداتی) استفاده گردید. هر چند ضریب همبستگی به عنوان یکی از شاخص‌های بسیار مهم در ارزیابی عملکرد مدل‌ها محسوب می‌شوند ولی این شاخص به تنهایی به عنوان معیاری جهت مقایسه عملکرد مدل‌ها کافی نبوده و لازم است علاوه بر آن، شاخص‌های دیگری نظیر ریشه میانگین مربعات خطا، قدر مطلق میانگین خطا و برازش گرافیکی پراکندگی داده‌های مشاهداتی و برآوردی برای ارزیابی مدل‌ها استفاده شود (۱۶ و ۲۰).

روابط شاخص‌های کمی سه گانه مذکور در فرمول‌های ۴ تا ۶ نشان داده شده است. بر اساس این شاخص‌های کمی، مدل برتر دارای مقدار حداکثر $|R|$ و حداقل $RMSE$ و MAE خواهد بود.

$$R = \frac{\sum_{i=1}^N (TDS_{io} - \overline{TDS_o})(TDS_{ie} - \overline{TDS_e})}{\sqrt{\sum_{i=1}^N (TDS_{io} - \overline{TDS_o})^2 \sum_{i=1}^N (TDS_{ie} - \overline{TDS_e})^2}} \quad (4)$$

جدول ۴- مشخصات مدل‌های بهینه شبکه‌های عصبی مصنوعی و پارامترهای آماری محاسبه شده جهت تخمین مقدار جامدات محلول در دوره آزمون

Table 4- The Characteristics and statistical parameters of ANNs optimal models for TDS predictions during the test period

ایستگاه Station	نام مدل Model	ترکیبات ورودی مدل Model input Combination	ساختار مدل Structure of Model	خروجی Transfer function in hidden and output layers	R	RMSE (mg/L)	MAE (mg/L)
آنیان Anyan	ANN1	HCO ₃	۱-۵-۱	Tangsig-purelin	0/885	21/95	16/12
	ANN2	HCO ₃ , Ca	۲-۵-۱	Logsig-purelin	0/897	20/83	14/88
	ANN3	HCO ₃ , Ca, Na	۳-۳-۱	Logsig-purelin	0/923	17/17	13/80
	ANN4	HCO ₃ , Ca, Na, Q	۴-۴-۱	Logsig-purelin	0/931	17/54	15/14
	ANN5	HCO ₃ , Ca, Na, Q, Mg	۵-۳-۱	Tangsig-purelin	0/942	15/11	13/46
صفاخانه Safakhaneh	ANN1	HCO ₃	۱-۴-۱	Tangsig-purelin	0/577	86/01	71/91
	ANN2	HCO ₃ , Ca	۲-۳-۱	Tangsig-purelin	0/608	84/48	70/99
	ANN3	HCO ₃ , Ca, Na	۳-۲-۱	Logsig-purelin	0/878	49/16	39/74
	ANN4	HCO ₃ , Ca, Na, Q	۴-۲-۱	Logsig-purelin	0/904	46/30	36/83
	ANN5	HCO ₃ , Ca, Na, Q, Mg	۵-۲-۱	Tangsig-purelin	0/945	35/26	26/46

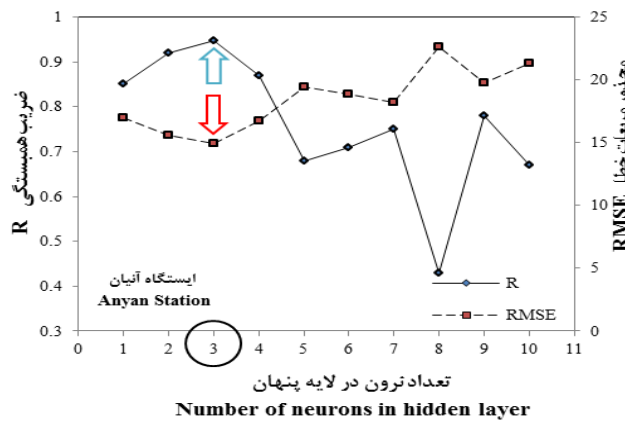
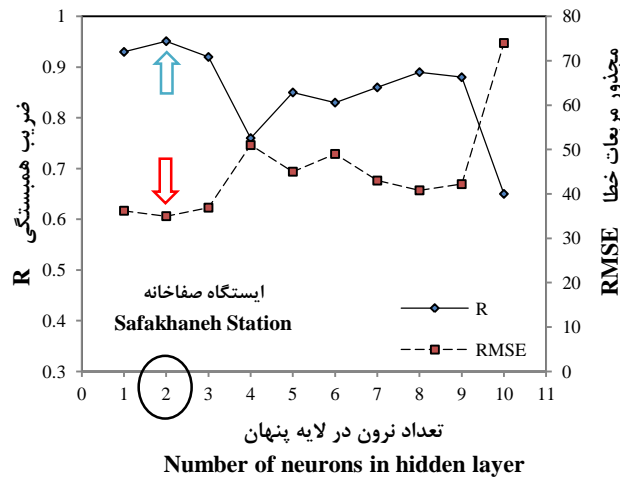
جدول ۵- مشخصات مدل‌های بهینه فازی-عصبی تطبیقی و پارامترهای آماری محاسبه شده جهت تخمین مقدار جامدات محلول در دوره آزمون

Table ۵-The Characteristics and statistical parameters of ANFIS optimal models for TDS predictions during the test period

ایستگاه Station	نام مدل Model	ترکیبات ورودی مدل Model input Combination	تعداد قوانین فازی Number of fuzzy rules	شعاع دسته- بندی Radii value	R	RMSE (mg/L)	MAE (mg/L)
آنیان Anyan	ANFIS-SC1	HCO ₃	۵	۰/۱۸	0/896	18/83	15/08
	ANFIS-SC2	HCO ₃ , Ca	۳	۰/۳	0/910	18/41	14/10
	ANFIS-SC3	HCO ₃ , Ca, Na	۳	۰/۴۵	0/924	17/27	13/74
	ANFIS-SC4	HCO ₃ , Ca, Na, Q	۲	۰/۳۹	0/926	17/09	13/81
	ANFIS-SC5	HCO ₃ , Ca, Na, Q, Mg	۲	۰/۴	0/948	14/94	13/60
صفاخانه Safakhaneh	ANFIS-SC1	HCO ₃	۳	۰/۵۲	0/586	85/79	71/860
	ANFIS-SC2	HCO ₃ , Ca	۳	۰/۶۲	0/619	83/38	67/13
	ANFIS-SC3	HCO ₃ , Ca, Na	۲	۰/۷۵	0/896	48/60	39/03
	ANFIS-SC4	HCO ₃ , Ca, Na, Q	۲	۰/۷۱	0/923	43/79	34/18
	ANFIS-SC5	HCO ₃ , Ca, Na, Q, Mg	۲	۰/۷	0/951	34/97	26/49

در بازه صفر تا ۱ می‌باشد، انجام گرفت. بهترین مدل با استفاده از سامانه استنتاجی نوع سوگنو، اندازه گام ۰/۰۲، اندازه گام افزایشی ۱/۱، اندازه گام کاهش ۰/۹ و الگوریتم آموزشی هیبرید^۲ که ترکیب روش گرادیان کاهش و حداقل مربعات خطا می‌باشد، تعیین گردید. مشخصات مدل‌های بهینه شبکه عصبی مصنوعی و سامانه استنتاجی عصبی-فازی تطبیقی در جداول ۴ و ۵ آورده شده است. شکل ۴ تغییرات R و RMSE مدل ANN5 را بر اساس تعداد نرون در لایه

الگوریتم لونبرگ-مارکوارت در دسته الگوریتم‌های سریع پس-انتشار که از تکنیک‌های استاندارد بهینه‌سازی عددی استفاده می‌کند، قرار گرفته و به دلیل استفاده از ماتریس ژاکوبین به جای ماتریس هشین در ساختار محاسباتی، پیچیدگی محاسباتی در آن کمتر بوده و دارای کارایی بسیار بالایی نسبت به سایر الگوریتم‌ها می‌باشد. هم-چنین در روش عصبی-فازی تطبیقی مدل دسته‌بندی تفریقی، تحلیل حساسیت مدل بر مبنای پارامتر شعاع دسته^۱ که محدوده تغییرات آن



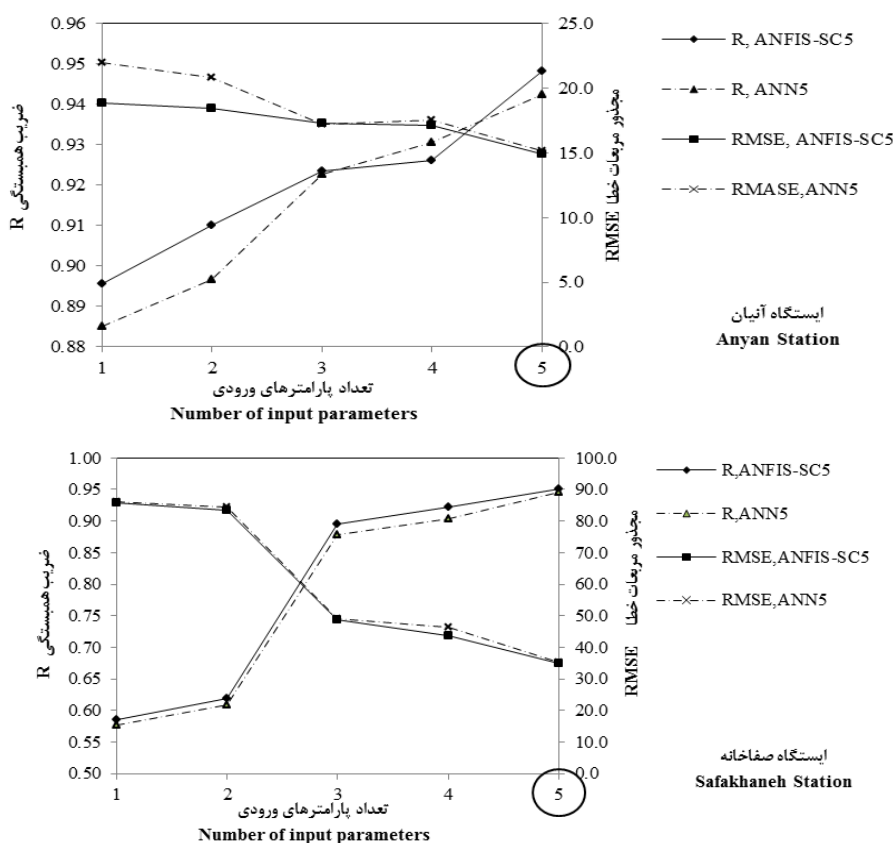
شکل ۴- نمودار تغییرات R و RMSE مدل ANN5 بر اساس تعداد نرون در لایه پنهان برای ایستگاه‌های آنیان و صفاخانه

Figure4- Variation of R and RMSE versus the number of neurons in hidden layer of ANN5 model for Anyan and Safakhaneh stations

بر اساس جداول ۴ و ۵ و شکل ۴ مدل بهینه شبکه عصبی مصنوعی و سامانه استنتاجی عصبی-فازی تطبیقی در دو ایستگاه آنیان و صفاخانه به ترتیب با تعداد نرون ۳ و ۲، تابع محرک تانژانت سیگموئید در لایه پنهان، تابع محرک خطی در لایه خروجی، شعاع دسته‌ای ۰/۴ و ۰/۷ همراه با ارائه دو قانون فازی به‌دست آمد. بر اساس نتایج مشاهده می‌گردد که در ایستگاه‌های آنیان و صفاخانه، هر دو روش شبکه عصبی مصنوعی و سامانه استنتاجی فازی-عصبی تطبیقی دارای عملکرد قابل قبول برای تخمین مقدار جامدات محلول داشتند. ولی به ترتیب اولویت، سامانه استنتاجی عصبی-فازی تطبیقی نسبت به شبکه عصبی مصنوعی عملکرد مناسب‌تری داشت. شکل ۵ تغییرات R و RMSE بر اساس تعداد پارامترهای ورودی مدل‌های ANN و ANFIS-SC در ایستگاه‌های آنیان و صفاخانه را نشان می‌دهد. بررسی شکل مذکور نشان می‌دهد که مدل‌های ANN5 و

ANFIS-SC5 با پنج پارامتر ورودی، دارای حداکثر میزان R و حداقل میزان RMSE و بهترین عملکرد در بین سایر الگوهای ورودی می‌باشند. در شکل ۶ نمودارهای مقادیر مشاهداتی و برآوردی توسط شبکه‌های عصبی مصنوعی با الگوریتم آموزشی لونیگ-مارکواریت (ANN5) و سامانه استنتاجی عصبی-فازی تطبیقی مدل دسته‌بندی تفریقی (ANFIS-SC5)، به ترتیب برای ایستگاه‌های آنیان و صفاخانه در دوره آزمون نشان داده شده است. پس از تعیین بهترین الگوی ورودی توسط دو روش مذکور، برنامه‌ریزی بیان ژن برای ارائه روابط ریاضی حاکم بر مسئله با هدف تخمین مقدار جامدات محلول در رودخانه به کار گرفته شد.

بر اساس جداول ۴ و ۵ و شکل ۴ مدل بهینه شبکه عصبی مصنوعی و سامانه استنتاجی عصبی-فازی تطبیقی در دو ایستگاه آنیان و صفاخانه به ترتیب با تعداد نرون ۳ و ۲، تابع محرک تانژانت سیگموئید در لایه پنهان، تابع محرک خطی در لایه خروجی، شعاع دسته‌ای ۰/۴ و ۰/۷ همراه با ارائه دو قانون فازی به‌دست آمد. بر اساس نتایج مشاهده می‌گردد که در ایستگاه‌های آنیان و صفاخانه، هر دو روش شبکه عصبی مصنوعی و سامانه استنتاجی فازی-عصبی تطبیقی دارای عملکرد قابل قبول برای تخمین مقدار جامدات محلول داشتند. ولی به ترتیب اولویت، سامانه استنتاجی عصبی-فازی تطبیقی نسبت به شبکه عصبی مصنوعی عملکرد مناسب‌تری داشت. شکل ۵ تغییرات R و RMSE بر اساس تعداد پارامترهای ورودی مدل‌های ANN و ANFIS-SC در ایستگاه‌های آنیان و صفاخانه را نشان می‌دهد. بررسی شکل مذکور نشان می‌دهد که مدل‌های ANN5 و



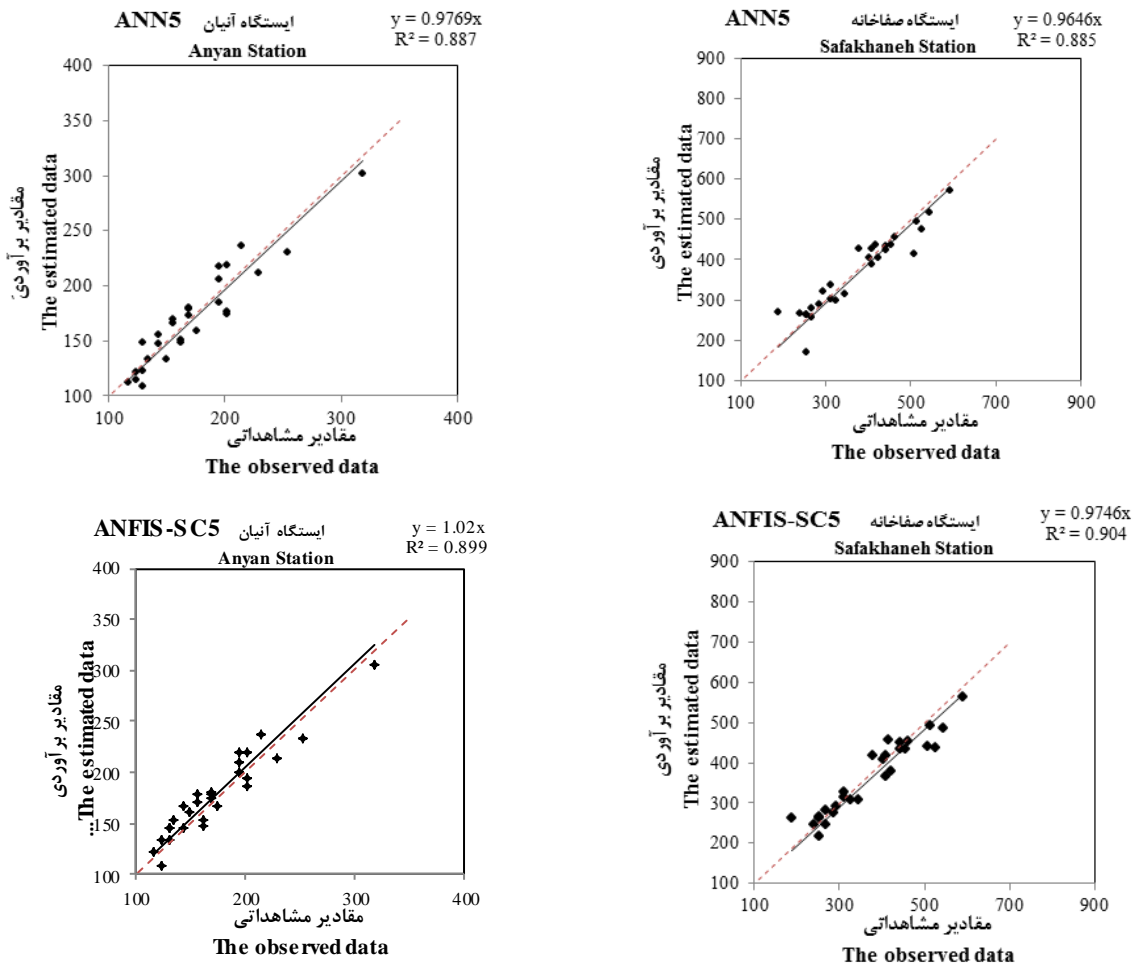
شکل ۵- تغییرات R و RMSE بر اساس تعداد پارامترهای ورودی مدل های ANN و ANFIS-SC در ایستگاه های آنیان و صفاخانه
 Figure 5- Variation of R and RMSE values versus input numbers of ANN and ANFIS-SC models for Anyan and Safakhaneh stations

خروجی و متغیرهای ورودی، جهت تخمین مقدار جامدات محلول در بالادست سد، آن را از سایر مدل های هوشمند متمایز نموده که این مسأله در فرآیند پیش بینی اهمیت بسزایی دارد. معادلات حاصل از بهترین مدل ها برای دو ایستگاه مورد مطالعه به زبان برنامه نویسی MATLAB در جدول ۷ و ساختار درخت تجزیه مربوط به معادلات در شکل ۸ ارائه شده است.

جهت ارزیابی داده های به دست آمده از بهترین مدل (برنامه ریزی بیان ژن)، نمودار بین باقیمانده اختلاف حاصل از داده های مشاهداتی و برآوردی به ازای مقادیر برآوردی برای داده های مرحله آزمون ترسیم گردید (شکل ۹). بر اساس این نمودارها مشاهده می شود که برای هر دو ایستگاه مورد مطالعه، داده های تولیدی حول محور تقریباً افقی بوده و دارای شیب کمی می باشد، که این امر حاکی از تولید داده های تصادفی، مستقل و بدون روند مدل ها می باشد. لذا می توان گفت که مدل های هوش مصنوعی توانایی آن را دارند که به عنوان یک مدل شبیه ساز تغییرات کیفیت آب در یک رودخانه، مورد استفاده قرار گیرند، و با سایر مدل های شبیه سازی کیفی آب رقابت نمایند.

جدول ۶ مشخصات مدل های بهینه برنامه ریزی بیان ژن و پارامترهای آماری محاسبه شده را نشان می دهد. شکل ۷ نمودار و منحنی های مقادیر مشاهداتی و برآوردی مدل GEP5 در ایستگاه های مطالعاتی برای داده های مرحله آزمون را نشان می دهد. معادله رگرسیونی برازش داده شده بین مقادیر برآوردی و مشاهداتی معادله $y=ax$ می باشد. با توجه به معادلات رگرسیونی مستخرج، ضریب a در روش برنامه ریزی بیان ژن برای دو ایستگاه مورد مطالعه به عدد واحد (نیمساز ربع اول) نزدیک تر بوده، که حاکی از کاهش میزان انحرافات داده های برآوردی می باشد. با وجود اینکه ضریب همبستگی بین داده های مشاهداتی و برآوردی الگوی شماره پنج در هر سه روش مدل شبکه عصبی مصنوعی، عصبی-فازی تطبیقی و برنامه ریزی بیان ژن در سطح معنی داری ۹۵ درصد معنی دار می باشد، ولی روش برنامه ریزی بیان ژن دارای ضریب همبستگی بالاتری نسبت به دو روش دیگر می باشد. هم چنین این روش در تخمین مقادیر حداکثر جامدات محلول در هر دو ایستگاه مورد مطالعه نیز عملکرد مناسبی داشته است.

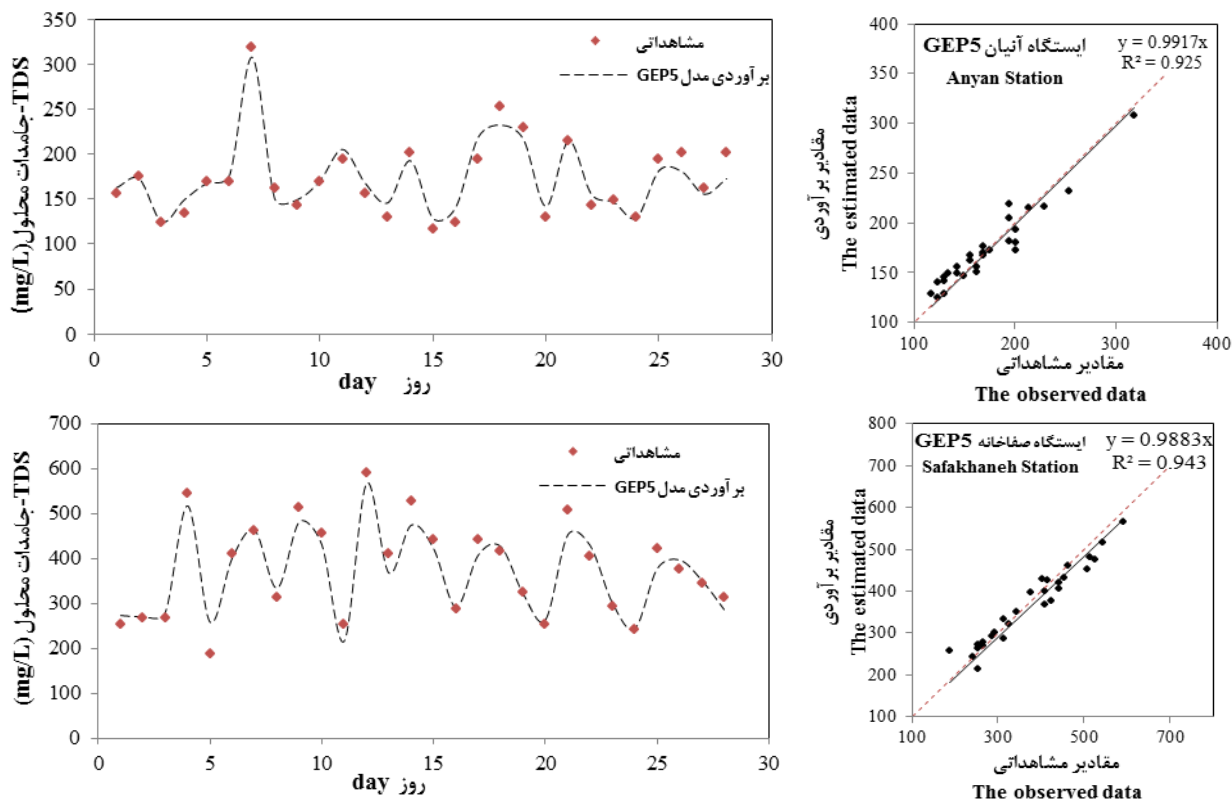
قابلیت مدل برنامه ریزی بیان ژن در ارائه معادله ریاضی بین متغیر



شکل ۶- نمودار مقادیر مشاهداتی و برآوردی مدل ANN5 و ANFIS-SC5 در ایستگاه‌های مطالعاتی برای داده‌های مرحله آزمون
Figure 6- Observed and estimated of ANN5 and ANFIS-SC5 models for study stations during the test period

جدول ۶- مشخصات مدل‌های بهینه برنامه‌ریزی بیان ژن و پارامترهای آماری محاسبه شده جهت تخمین مقدار جامدات محلول در دوره آزمون
Table 6- The Characteristics and statistical parameters of GEP optimal models for TDS predictions during the test period

نام ایستگاه Station	نام مدل Model	ترکیبات ورودی مدل Model input Combination	R	RMSE (mg/L)	MAE (mg/L)
آنیان Anyan	GEP1	HCO ₃	0/982	18/06	14/77
	GEP2	HCO ₃ , Ca	0/910	17/55	13/12
	GEP3	HCO ₃ , Ca, Na	0/937	16/01	12/92
	GEP4	HCO ₃ , Ca, Na, Q	0/941	15/88	12/43
	GEP5	HCO ₃ , Ca, Na, Q, Mg	0/962	12/82	10/60
صفاخانه Safakhaneh	GEP1	HCO ₃	0/623	75/71	67/18
	GEP2	HCO ₃ , Ca	0/738	67/77	57/99
	GEP3	HCO ₃ , Ca, Na	0/918	40/22	39/03
	GEP4	HCO ₃ , Ca, Na, Q	0/952	35/45	29/76
	GEP5	HCO ₃ , Ca, Na, Q, Mg	0/971	29/08	23/19



شکل ۷- نمودار و منحنی‌های مقادیر مشاهداتی و برآوردی مدل GEP5 در ایستگاه‌های مطالعاتی برای داده‌های مرحله آزمون
Figure 7- Observed and estimated plots of GEP5 models for study stations during the test period

جدول ۷- معادلات حاصل از بهترین مدل‌ها برای دو ایستگاه مورد مطالعه

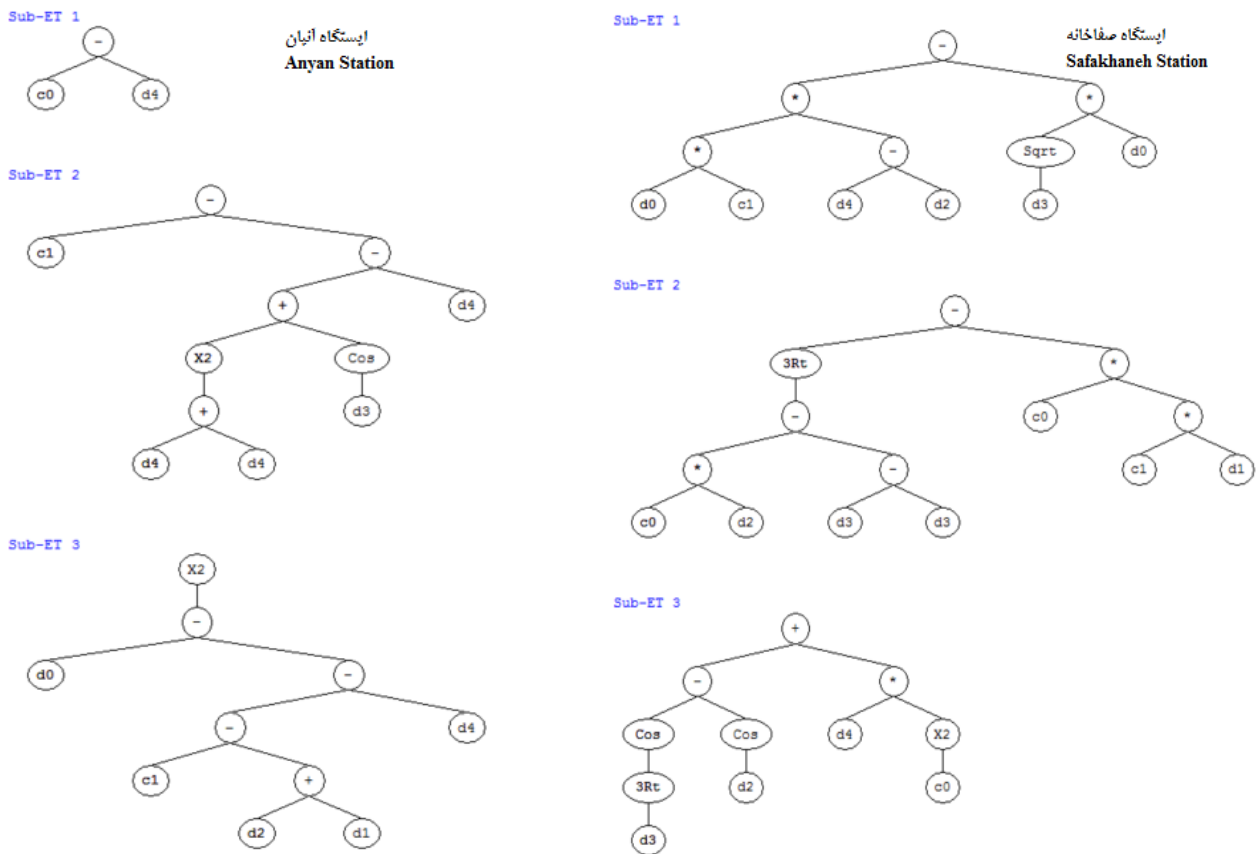
Table 7- Equations of best models for two study station

ایستگاه Station	معادلات صریح و ضرایب ثابت Explicit equations and constants
آنیان Anyan	$a_0 = -8.0755, a_1 = -0.548554, a_2 = -0.144653, a_3 = -6.059173, a_4 = -6.425171, a_5 = -8.312713$ $Y_1 = (a_0 - (Mg));$ $Y_2 = Y_1 + (a_3 - (((2Mg)^2) + \cos(Q)) - Mg);$ $TDS = Y_2 + ((HCO_3 - ((a_5 - (Ca + Na)) - Mg))^2)$
صفاخانه Safakhaneh	$a_0 = -6.264007, a_1 = -8.100647, a_2 = 4.72403, a_3 = 9.290619, a_4 = -1.629944, a_5 = -8.100647$ $Y_1 = ((Mg + Ca) * (a_1^2));$ $Y_2 = Y_1 + (((Q^{1/3}) - a_3) * (HCO_3 + Na)) + ((a_3 * a_2) * (2Na));$ $TDS = Y_2 + (1 - (((Ca^2) * Mg) - a_4))$

نتیجه‌گیری کلی

در این مطالعه جهت تخمین مقدار جامدات محلول در بالادست سد مخزنی بوکان، از روش‌های شبکه عصبی مصنوعی با الگوریتم آموزشی پس انتشار خطای لونیگ-مارگارت، سامانه استنتاجی عصبی-فازی تطبیقی مدل دسته‌بندی تفریقی و برنامه ریزی بیان ژن استفاده شد.

با توجه به ارزیابی عملکرد مدل‌های مورد مطالعه می‌توان نتیجه گرفت که با وجود عملکرد بهتر روش برنامه‌ریزی بیان ژن نسبت به سایر روش‌های هوش مصنوعی و ارائه رابطه ریاضی حاکم بر مسئله، هر سه روش را می‌توان به عنوان ابزار مناسب برای تخمین پارامترهای کیفیت آب رودخانه در بالادست سد مخزنی بوکان معرفی نمود.

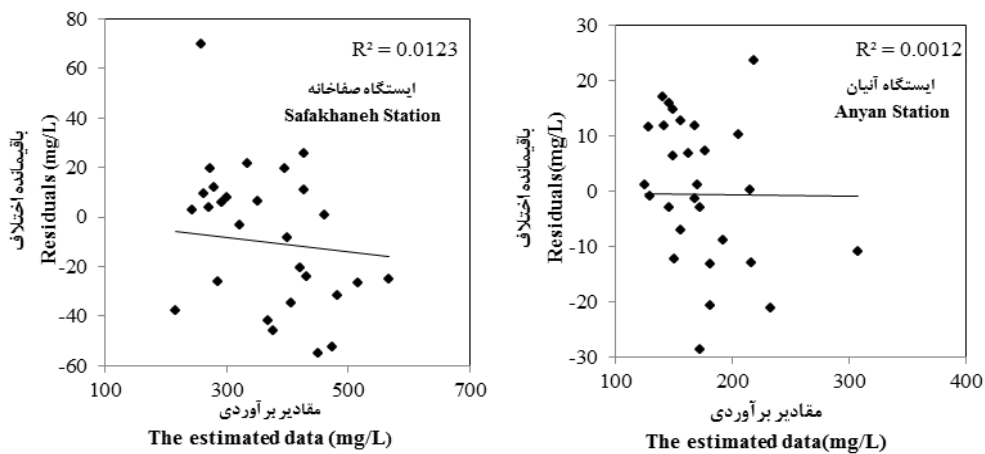


شکل ۸- درخت تجزیه حاصل از GEP5 برای ایستگاه‌های مطالعاتی

* پارامترهای ورودی d_0, d_1, d_2, d_3, d_4 به ترتیب معرف $\text{HCO}_3, \text{Ca}, \text{Na}, \text{Q}$ و Mg می‌باشند.

Figure 8- Expression tree of GEP5 for study stations

*Terminal sets d_0, d_1, d_2, d_3, d_4 denote $\text{HCO}_3, \text{Ca}, \text{Na}, \text{Q}$ and Mg , respectively



شکل ۹- نمودارهای باقیمانده اختلاف و برآوردی مدل GEP5 برای داده‌های آزمون در دو ایستگاه مطالعاتی

Figure 9- Plots of the residuals versus GEP5 models computed during the test period for study stations

بندی تفریقی و برنامه‌ریزی بیان ژن می‌تواند به عنوان ابزاری توانمند برای تخمین پارامترهای کیفی آب رودخانه‌ها مورد استفاده قرار گیرند.

بر اساس نتایج حاصل، می‌توان بیان کرد که روش‌های شبکه‌های عصبی مصنوعی، سامانه استنتاجی عصبی-فازی تطبیقی، مدل دسته-

آزمایشگاهی کاری پرهزینه و زمان‌بر است، باشند. علاوه بر این ارائه معادله ریاضی بین متغیر وابسته و سایر متغیرهای مستقل توسط مدل برنامه‌ریزی بیان ژن از مهمترین مزیت‌های این مدل بوده و جهت پیش‌بینی‌های آتی اهمیت بسزایی دارد. لذا مزیت ذکر شده، این مدل را نسبت به سایر مدل‌های هوشمند متمایز می‌کند. همچنین به ترتیب اولویت، برنامه‌ریزی بیان ژن، روش سامانه استنتاج عصبی-فازی تطبیقی و شبکه عصبی مصنوعی با افزایش ضرایب همبستگی و کاهش مجذور مربع خطاها و میانگین قدرمطلق خطاها بهترین عملکرد را در طول دوره مطالعاتی ۱۸ ساله نشان دادند. در نهایت پیشنهاد می‌گردد، عملکرد این روش‌ها بر آورد مقدار جامدات محلول سایر رودخانه‌های کشور نیز بررسی و با نتایج مطالعات قبلی مقایسه و تحلیل شود.

نتایج به دست آمده از این تحقیق در زمینه تخمین پارامترهای کیفی آب رودخانه‌ها با نتایج تحقیقات سنگرور و همکاران (۲۱) در تخمین میزان COD و نیز ساین و همکاران (۲۴) در تخمین میزان BOD و COD همخوانی دارد. همچنین، نتایج این تحقیق، بر اساس تحقیقات انجام شده در مورد عملکرد بهتر روش برنامه‌ریزی بیان ژن نسبت به سایر روش‌های هوش مصنوعی، در زمینه پدیده‌های هیدرولوژیکی با نتایج تحقیقات بسیاری از محققین از جمله آیتک و آلپ (۲)، شیری و همکاران (۲۳)، کیشی و همکاران (۱۴ و ۱۵) مطابقت دارد. روش‌های مذکور قادر به رقابت با سایر تکنیک‌های معمول در زمینه مدل‌سازی بوده و به سبب سادگی از لحاظ عدم نیاز به دانش بالایی میانی ریاضی الگوریتم، می‌توانند جایگزینی مناسب برای روش‌های مرسوم اندازه‌گیری پارامترهای کیفیت آب، که در شرایط معمول اندازه‌گیری

منابع

- 1- Alvisi S., Mascellani G., Franchini M., and Bardossy A. 2005. Water level forecasting through fuzzy logic and artificial neural network approaches, *Journal of Hydrology and Earth System Sciences*, 2: 1107-1145.
- 2- Aytok A., and Alp M. 2008. An application of artificial intelligence for rainfall runoff modelling, *Journal of Earth Systems Science*, 117 (2):145-155.
- 3- Aytok A., and Kisi O. 2008. A genetic programming approach to suspended sediment modelling, *Journal of Hydrology*, 351:288-298.
- 4- Caudill M. 1987. Neural networks primer: Part I, *AI Expert*, 2(12), 46-52.
- 5- Chiu S.L. 1995. Extracting fuzzy rules for pattern classification by cluster estimation, p. 1-4. In: *The 6th International Fuzzy Systems Association World Congress*.
- 6- Farboudnam N., Ghorbani M.A., and Alami M.T. 2009. River Flow Prediction Using Genetic Programming (Case Study: Lighvan River Watershed), *Journal of Soil and Water*, 19(1): 107-123. (in Persian with English abstract)
- 7- Goyal M.K., and Ojha C.S.P. 2011. Estimation of scour downstream of a ski-jump bucket using support vector and M5 model tree, *Water Resources Management*, 25(9): 2177-2195.
- 8- Guven A. 2009. Linear genetic programming for time-series modeling of daily flow rate, *Journal of Earth System Science*, 118 (2): 137-146.
- 9- Guven A., and Gunal M. 2008. Genetic programming approach for prediction of local scour downstream hydraulic structures, *Journal of Irrigation and Drainage Engineering*, 134(2):7 241-249.
- 10- Guven A., and Talu N.E. 2010. Gene-expression programming for estimating suspended sediment in Middle Euphrates Basin, Turkey, *CLEAN-Soil Air Water*, 38(12): 1159-1168.
- 11- Hashmi M.Z., Shamseldin A.Y., and Melville B.W. 2011. Statistical downscaling of watershed precipitation using Gene Expression Programming (GEP), *Environmental Modelling & Software*, 26:1639-1646.
- 12- Jain S.K., Das A., and Srivastava D.K. 1999. Application of ANN for reservoir inflow prediction and operation, *Journal of Water Resources Planning and Management*, ASCE, 125(5): 263-271.
- 13- Jang J.S.R. 1993. ANFIS: adaptive-network-based fuzzy inference system, *IEEE Transactions on Systems. Man and Cybernetics*, 23 (3): 665-685.
- 14- Kisi O., Shiri J., Nikoofar B. 2012. Forecasting daily lake levels using artificial intelligence approaches, *Computers & Geosciences*, 41: 169-180.
- 15- Kisi O., Shiri J., and Tombul M. 2013. Modeling rainfall-runoff process using soft computing techniques, *Computers & Geosciences*, 51: 108-117.
- 16- Legates D.R. and McCabe G.J. 1999. Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, 35 (1): 233-241.
- 17- Misaghi F., and Mohammady k. 2004. Forecasting quality variouse of ZayandehRood river water by using artificial neural networks. *The 2th National Student Conference in water and soil resources*, College of Agriculture, Shiraz University. (in Persian with English abstract)
- 18- Montaseri M., and Zaman Zad Ghavidel S. 2014. River Flow Forecasting by Using Soft computing. *Journal of Water and Soil*, 28 (2): 394-405. (in Persian with English abstract)
- 19- Najah A., Elshafie A., Karim O., and Jaffar O. 2009. Prediction of Johor river water quality parameters using artificial neural networks, *European Journal of Scientific Research*, 28: 422-35.

- 20- Sanikhani H., and Kisi O. 2012. River Flow Estimation and Forecasting by Using Two Different Adaptive Neuro-Fuzzy Approaches, *Water Resources Management*, 26: 1715–1729.
- 21- Sengorur B., Dogan E., Koklu R., and Samandar A. 2006. Dissolved oxygen estimation using artificial neural network for water quality control, *Fresenius Environmental Bulletin*, 15: 1064–1067.
- 22- Shafaei Y., Farzaneh M., and Teshnehlab M. ۲۰۰۲. Modeling of producing trip by using Adaptive Neuro-Fuzzy. *Issue of Engineering Faculty*, 36(3): 361-170,(in Persian with English abstract)
- 24- Shiri J., Kisi O., Landaras G., Lopez J.J., Nazemi A.H., and Stuyt L.C.P.M. 2012. Daily referenc evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain), *Journal of Hydrology*, 414- 415, 302–316.
- 25- Sighn K.P., Basant A., Malik A., and Jain G. 2009. Artificial neural network modeling of the river water quality-A case study, *Ecological Modelling*, 220: 888–895.
- 26- Traore S., and Guven A. 2012. Regional-specific numerical models of evapotranspiration using gene-expression programming interface in Sahel, *Water Resources Management*, 26(15):4367-4380.

Estimation of Rivers Dissolved Solids TDS by Soft Computing (Case Study: Upstream of Boukan Dam)

S. Zaman Zad Ghavidel¹ - K. Zeinalzadeh^{2*}

Received: 11-23-2014

Accepted: 05-25-2015

Introduction: A total dissolved solid (TDS) is an important indicator for water quality assessment. Since the composition of mineral salts and discharge affects the TDS of water, it is important to understand the relationship of mineral salts composition with TDS.

Materials and Methods: In this study, methods of artificial neural networks with Levenberg-Marquardt training algorithm, adaptive neuro fuzzy inference system based on Subtractive Clustering and Gene expression programming were used to model water quality properties of Zarrineh River Basin at upstream of Boukan dam, to be developed in total dissolved solids prediction. ANN and ANFIS programs code were written using MATLAB programming language. Here, the ANN with one hidden layer was used and the hidden nodes' number was determined using trial and error. Different activation functions (logarithm sigmoid, tangent sigmoid and linear) were tried for the hidden and output nodes and the GeneXpro Tools 4.0 were used to obtain the equation of the best models. Therefore, water quality data from two hydrometer stations, namely Anyan and Safakhaneh hydrometer stations were used during the statistical period of 18 years (1389-1372). In this research, for selecting input variables to the data driven models the stepwise regression method was used. In the application, 75% of data set were used for training and the remaining, 25% of data set were used for testing, randomly. In this paper, three statistical evaluation criteria, correlation coefficient (R), the root mean square error (RMSE) and mean absolute error (MAE), were used to assess model's performances.

Results and Discussion: By applying stepwise method, the first significant (at 95% level) variable entered to the model was the HCO_3 . The second variable that entered to the model was Ca. The third and fourth ones were Na and Q respectively. Mg was finally entered to the model. The optimal ANN architecture used in this study consists of an input layer with five inputs, one hidden and output layer with three and two neurons for Anyan and Safakhaneh hydrometer stations, respectively. Similar ANN, ANFIS-SC5 model had the best performance. It is clear that the ANFIS with 0/4 and 0/7 radii value has the highest R and the lowest RMSE for Anyan and Safakhaneh hydrometer stations, respectively. Various GEP models have been developed using the input combinations similar ANN and ANFIS models. Comparing the GEP5 estimations with the measured data for the test stage demonstrates a high generalization capacity of the model, with relatively low error and high correlation. From the scatter plots it is obviously seen that the GEP5 predictions are closer to the corresponding measured TDS than other models. As seen from the best straight line equations (assume the equation as $y=ax$) in the scatter plots that the a coefficient for GEP5 is closer to 1 than other models. In addition to previous operation, Gene expression programming offered mathematical relationships in the stations of Anyan and Safakhaneh with the correlation coefficients, respectively 0.962, 0.971 and with Root-mean-square errors, respectively 12.82, 29.08 in order to predict dissolved solids (TDS) in the rivers located at upstream of the dam. The obtained results showed the efficiency of the applied models in simulating the nonlinear behavior of TDS variations in terms of performance indices. Overall, the GEP model outperformed the other models. For all of applied models, the best result was obtained by application of input combination (5) including HCO_3 , Ca, Na, Q and Mg. The results are also tested by using t test for verifying the robustness of the models at 95% significance level. Comparison results indicated that the poorest model in TDS simulation was ANN especially in test period. The observed relationship between residuals and model computed TDS values shows complete independence and random distribution. It is further supported by the respective correlations for GEP5 models ($R^2 = 0.0011$ for Anyan station and $R^2 = 0.0123$ for safakhaneh station) which are negligible small. Plots of the residuals versus model computed values can be more informative regarding model fitting to a data set. If the residuals appear to behave randomly it suggests that the model fits the data well. On the other hand, if non-random distribution is evident in the residuals, the model does not fit the data adequately. On the base of these results, we propose

1, 2- PhD Candidate and Assistant Professor, Department of Water Engineering, Urmia University
(* - Corresponding Author Email: k.zeinalzadeh@urmia.ac.ir)

GEP, ANFIS-SC and ANN methods as effective tools for the computation of total dissolved solids in river water, respectively.

Conclusion: It can be concluded that the ANN, ANFIS-GP, ANFIS-SC and GEP models can be considered as promising tools for forecasting TDS values, based on water quality parameters. It is notable from the results that the prediction accuracy of all applied models increases by increasing the number of input combinations. With attention to the aim of current research that is presenting the feasibility of artificial intelligence techniques for modeling TDS values, it is notable that the results presented in this paper are for research purpose and applying the abstained results for real-world needs some complicated steps and building artificial intelligences methods, based on complete data and parameters maybe affected the TDS values.

Keywords: Gene Expression, Dissolved Solids, Zarrineh River, Artificial Neural Networks