

شیرگاه، تالار، کیاکلا و عرب خیل بدست آمده است. سپس شاخص NSFQI برآورد گردید. سپس به کمک روش‌های برنامه‌ریزی بیان ژنی و شبکه‌های عصبی مصنوعی، مدل‌هایی برای تعیین ارتباط بین پارامترهای کیفی آب و شاخص کیفیت آب با دقت بالا به دست آمد. برای ارزیابی عملکرد مدل‌ها از پارامترهای آماری نظیر، جذر میانگین مربعات خطا، میانگین مطلق خطا، شاخص پراکندگی و ضریب تعیین استفاده شد. نتایج به‌دست آمده نشان داد دقت شبکه‌های عصبی مصنوعی بر روی داده‌های مجموعه آزمون اندکی بالاتر از روش برنامه‌ریزی بیان ژنی در تخمین شاخص NSFQI است.

کلید واژه‌ها: مدلسازی، رودخانه تالار، برنامه‌ریزی بیان ژنی، شبکه‌های عصبی، شاخص NSFQI

مقدمه

رودخانه‌ها به عنوان مهم‌ترین منابع تامین و انتقال آب مصرفی بخش‌های صنعت، کشاورزی و شهری از اهمیت خاصی برخوردار هستند. توسعه روز افزون فعالیت‌های کشاورزی و صنعتی و افزایش حجم فاضلاب‌های شهری موجب آلودگی رودخانه‌ها گشته، به نحوی که کیفیت این منابع حیاتی آب را مورد مخاطره جدی قرار داده است. هم‌چنین به دلیل ورود پساب‌های صنعتی، کشاورزی و غیرکشاورزی و فاضلاب شهری به رودخانه‌ها و نیز کاهش نزولات جوی و افزایش مصرف آب در سال‌های اخیر، مشکل عمده آب رودخانه‌ها آلودگی بیش از حد آن بوده است.

بنا به اهمیت موضوع، تاکنون پژوهش‌های ارزشمندی در مورد ارزیابی، پهنه‌بندی کیفی رودخانه‌ها انجام شده است که از این میان می‌توان به سانجز و همکاران [۱۲] اشاره نمود، آنان شاخص WQI را در طول رودخانه گواداراما و مانزانارز مورد مطالعه قرار دادند. نمونه‌برداری در طول رودخانه برای ۲ سال (۲۰۰۳-۲۰۰۱) متوالی انجام شد و نمونه‌ها از ۶ ایستگاه برداشت گردید. آن‌ها در مطالعه خود از ۱۱ پارامتر شامل pH^۳، هدایت هیدرولیکی، TSS^۴

تخمین شاخص کیفیت آب رودخانه تالار با استفاده از شبکه‌های عصبی مصنوعی و برنامه‌ریزی بیان ژنی

برات مجردی^۱، فروغ علی‌زاده صنمی^۲ و مهرشاد صمدی^۳
تاریخ دریافت: ۱۳۹۵/۰۶/۰۴ تاریخ پذیرش: ۱۳۹۷/۰۱/۲۲

چکیده

رودخانه‌ها همواره در طول تاریخ به عنوان منبع مهم برای تامین آب آشامیدنی و کشاورزی مورد توجه جوامع انسانی قرار داشته و در شکل‌گیری تمدن‌های بشری بسیار موثر بوده‌اند. علاوه بر این به عنوان یک اکوسیستم دارای ارزش زیستگاهی بسیار بالایی می‌باشند. در حال حاضر اکثر رودخانه‌های شمال تحت تاثیر انواع دخالت‌های انسانی می‌باشد که باعث بروز انواع آلودگی‌ها، تغییر و تخریب اکوسیستم رودخانه‌ها شده است. از مهمترین عوامل آلودگی و دخالت‌های انسان می‌توان به آلودگی‌های ناشی از فاضلاب‌های صنعتی، شهری و روستایی؛ آلودگی‌های ناشی از تخلیه سموم مورد استفاده در کشاورزی، برداشت بی‌رویه آب رودخانه، تخریب پوشش گیاهی، احداث سدها و موانع زیر پل‌ها، مسدود شدن دهانه رودخانه، صید غیر مجاز و ... اشاره نمود. لذا با توجه به اهمیت رودخانه تالار در تامین آب کشاورزی و ...، هم‌چنین تخلیه آلاینده‌های متعدد به آن، شناسایی و ارزیابی کیفی آب رودخانه و ارائه رابطه‌ای برای تخمین آلودگی و کیفیت آب ضروری به نظر می‌رسد. حوضه آبریز تالار، دارای یک رودخانه اصلی و پنج رودخانه فرعی می‌باشد که هر کدام از زیر شاخه‌های متعددی تشکیل شده است. طول آبراهه اصلی ۱۵۱/۷۷ کیلومتر است و از جنوب و جنوب غربی حوضه سرمنشا می‌گیرد و از شمال غربی حوضه خارج می‌شود.

در این مطالعه داده‌ها از نوع مقطعی می‌باشد، ۷۲ نمونه در طی ماه‌های سال ۱۳۹۱-۱۳۹۲ از ۶ ایستگاه ورسک، پل سفید،

۱- استادیار گروه نقشه برداری، دانشگاه علم و صنعت ایران.

۲- نویسنده مسئول و دانشجوی دکتری مهندسی عمران گرایش آب و سازه هیدرولیکی، دانشگاه علم و صنعت ایران، پست الکترونیک:

Forough_alizadeh@civileng.iust.ac.ir.

۳- دانشجوی دکتری مهندسی عمران گرایش آب و سازه هیدرولیکی، دانشگاه علم و صنعت ایران.

۳- میزان اسیدی یا بازی بودن پ

۴- کل مواد جامد معلق در آب

آمونیاک، نیترات، نیتريت، فسفات، COD^۱، BOD5^۲، DO^۳ و درجه حرارت استفاده کردند و نتایج آن به صورت فصلی گزارش گردید. نتایج پژوهش آن‌ها نشان داد که مقدار شاخص WQI^۴ در ابتدای رودخانه گواداراما ۷۰ (کیفیت خوب) و در انتهای آن در حدود ۶۴ (کیفیت متوسط) می‌باشد. هم‌چنین مقدار شاخص WQI برای رودخانه مانزانارز در حدود ۶۵ گزارش گردید. فابیانو و همکاران [۵] در پژوهشی شاخص NSFQI^۵ را روی رودخانه‌های ماکوکو و کیوکسادا بررسی کردند. در این مطالعه نمونه برداری در طی ۲ سال انجام شد و پارامترهای کیفی شامل: pH، هدایت الکتریکی، TS، کلیفرم‌های مدفوعی، فسفات، کدورت، اکسیژن محلول، BOD5، قلیائیت، رنگ و مواد آلی مورد آنالیز قرار گرفت. نتایج نشان داد شاخص فوق، شاخص مناسبی برای پهنه‌بندی طول این ۲ رودخانه می‌باشد. روسلی و همکاران [۱۰] کیفیت آب رودخانه سبک واقع در مالزی را با استفاده از شاخص WQI مورد بررسی و پژوهش قرار دادند. نتایج حاصل از مطالعه آن‌ها نشان داد که رودخانه مورد مطالعه سطح پایینی از اکسیژن محلول و سطح بالایی از اکسیژن مورد نیاز شیمیایی و سرب را دارا می‌باشد و کیفیت آب رودخانه از نظر شاخص WQI در محدوده کلاس ۴ قرار گرفته است. واحداونیش و سندپی [۱۵] در پژوهشی کیفیت آب تالاب روب سیگار هند را با استفاده از شاخص NSFQI به صورت ماهیانه طی یک سال (۲۰۱۱-۲۰۱۲) بررسی نمودند. طبق نتایج حاصل از این مطالعه کیفیت آب تالاب روب سیگار در کل دوره مطالعه در محدوده کیفی متوسط قرار می‌گیرد. هم‌چنین بیشترین میزان شاخص (۶۶/۷۵) در ماه نوامبر و کمترین میزان شاخص (۵۵/۲۵) در ماه جولای مشاهده گردید. افندی و واردیاتنو [۴] وضعیت کیفی آب رودخانه چیمبیلایونگ واقع در استان بن تن را با استفاده از شاخص NSFQI و شاخص آلودگی (PI) بررسی نمودند. نمونه‌برداری از سه ایستگاه و در سه نوبت مطابق با فصل بارانی صورت گرفت. طبق نتایج حاصل از این مطالعه میزان شاخص NSFQI در محدوده ۸۸-۸۷ و میزان شاخص آلودگی در محدوده ۰/۷۸-۰/۵۶ قرار می‌گیرد. بر اساس این شاخص‌ها کیفیت آب رودخانه چیمبیلایونگ در محدوده خوب قرار می‌گیرد.

برنامه‌ریزی بیان ژن، برنامه‌ریزی ژنتیک از جدیدترین الگوریتم‌های فراکاوشی هستند که به دلیل دارا بودن دقت و عملکرد مناسب در الگوبندی، مورد توجه محققین قرار گرفته‌اند. در الگوی برنامه‌ریزی بیان ژن کروموزم‌های خطی و ساده با طول ثابت- مشابه با آنچه در الگوریتم ژنتیک استفاده می‌شود با ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک ترکیب می‌شوند. بنا به اهمیت موضوع، تاکنون پژوهشگران مختلفی در سراسر جهان اقدام به استفاده از برنامه‌ریزی بیان نموده‌اند

- ۱- اکسیژن مورد نیاز شیمیایی
- ۲- اکسیژن مورد نیاز بیولوژیکی در ۵ روز
- ۳- اکسیژن محلول در آب

4. Water Quality Index

5. Sanitation Foundation Water Quality Index

که از این میان به آیتک و کیشی [۲] اشاره نمود. آنان از روش برنامه‌ریزی ژنتیک برای شبیه‌سازی بار معلق در رودخانه تانگو واقع در ایالات متحده آمریکا بکار گرفتند. آن‌ها از مقادیر رسوب قبلی و جریان برای پیش‌بینی 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. امروزه علاقه‌ی فزاینده‌ای در توسعه نظری سامانه‌های پویای هوشمند که مبتنی بر داده‌های تجربی هستند ایجاد شده است. شبکه‌های عصبی مصنوعی جز این دسته‌اند که دانش یا قانون نهفته در ورای این اطلاعات را با پردازش روی داده‌های تجربی به ساختار شبکه منتقل می‌کنند. روش‌های هوشمند شبکه‌های عصبی مصنوعی نیز به منظور پیش‌بینی جریان روزانه در پژوهشی از دو روش شبکه عصبی مصنوعی و اتورگرسیون استفاده کرده‌اند.

با توجه به اهمیت بحث آلودگی و تعیین کیفیت آب رودخانه، نقش آن در مدیریت و برنامه‌ریزی منابع آب، مقایسه کاربرد روش‌های مختلف، در این پژوهش با استفاده از الگوهای برنامه‌ریزی بیان ژن و شبکه عصبی به شناسایی و ارزیابی کیفی آب رودخانه و ارائه رابطه‌ای برای تخمین آلودگی و کیفیت آب پرداخته شده است.

مواد و روش‌ها

منطقه مورد مطالعه: حوضه آبریز تالار، دارای یک رودخانه اصلی و پنج رودخانه فرعی می‌باشد که هر کدام از شاخه‌های متعددی تشکیل می‌شوند. طول آبراهه اصلی حدود ۷۷/۱۵۱ کیلومتر است که از جنوب و جنوب غربی حوضه سرمنشا می‌گیرند و از شمال غربی حوضه خارج می‌شود [۷]. در شکل ۱ محدوده رودخانه و در جدول ۱ سایر مشخصات رودخانه مورد مطالعه بیان شده است. برای تعیین ساختار شاخص کیفی آب‌های سطحی، اطلاعات حاصل از اندازه‌گیری پارامترهای کیفی آب، پایه و اساس کار می‌باشد.

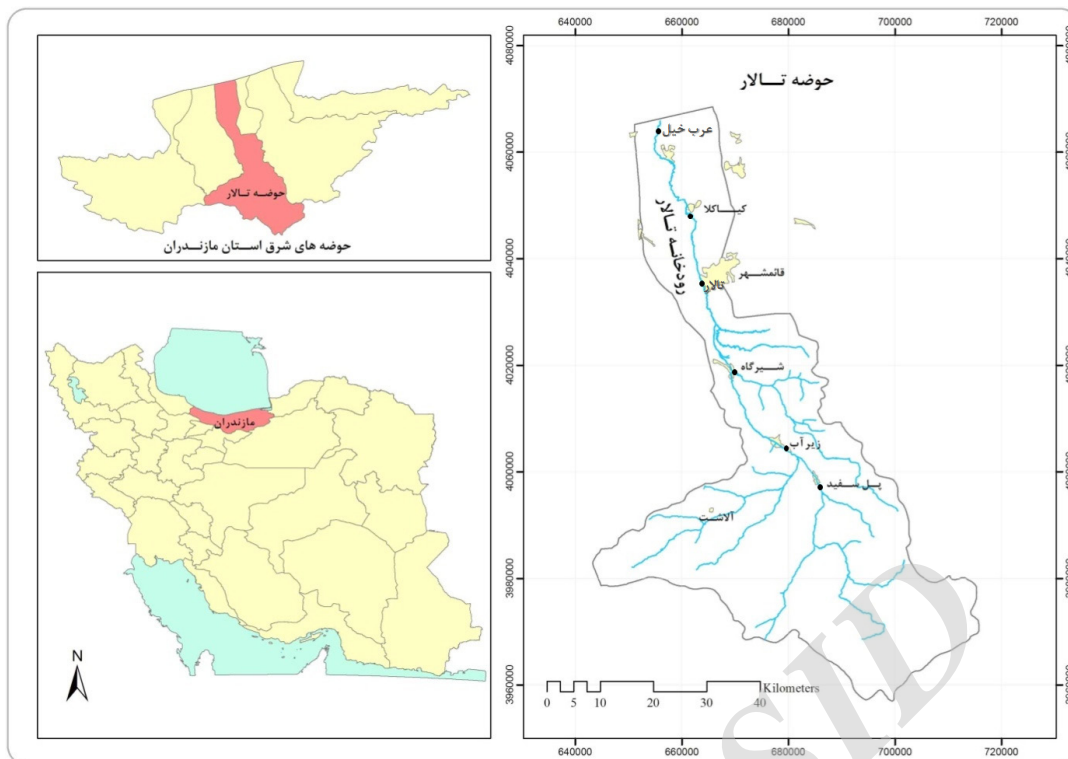
شاخص NSFQI: شاخص NSFQI که براساس ۹ پارامتر اصلی از جمله BOD5، اکسیژن محلول، کلیفرم مدفوعی، نیترات، pH، دما، TDS، فسفات کل و کدورت می‌باشد [۹].

$$NSFWQI = \sum W_i I_i \quad (1)$$

در این رابطه I_i معرف زیر شاخص W_i و W_i معرف ضریب وزنی شاخص I_i می‌باشد.

جدول ۲ طبقه‌بندی شدت آلودگی کیفیت آب را براساس شاخص‌های فوق از لحاظ مصارف عمومی و جدول ۳ پارامترهای وزن‌دهی به شاخص NSFQI را نشان می‌دهد.

با توجه به مطالب بیان شده در بخش مقدمه، هدف پژوهش کاربرد مدل‌های شبکه‌های عصبی مصنوعی و برنامه‌ریزی بیان ژنی به‌منظور تخمین کیفیت آب رودخانه تالار در ایستگاه‌های تعیین



شکل ۱- حوزه رودخانه تالار [۷]

Figure 1-Talar river watershed

جدول ۱- مشخصات رودخانه تالار

Table 1- Profile of the Talal River

ردیف	مشخصه حوضه	Charactestic watershed	No.
۱	وسعت حوضه آبریز	Basin area	۳۸/۲۹۰۵ 38.2905 Km ²
۲	میانگین آبدهی سالانه	Average annual discharge	۷۲/۳۳ متر مکعب بر ثانیه در ایستگاه شیرگاه 72.33 m ³ /s
۳	طول شاخه اصلی رودخانه	The main branch of the river length	۱۶۰ کیلومتر 160 Km
۴	ارتفاع متوسط	the average height	۱۶۹۹ متر 1699 m
۵	شیب متوسط حوضه	Basin mean slope	۵۴/۳۲ درصد 54.32 %
۶	جهت جریان	Flow direction	جنوب شرقی - شمال غربی Southeast - northwest
۷	تعداد ایستگاه‌های هیدرومتری	Number of hydrometric stations	۱۱ ایستگاه فعال 11 active stations

نمونه‌برداری از ایستگاه‌های مشخص شده در طی ۱۲ ماه از مهر سال ۱۳۹۱ الی شهریور ۱۳۹۲ انجام گرفت. تعداد نمونه‌های برداشتی در ۶ ایستگاه در طول رودخانه شامل ۷۲ نمونه بوده است. روش نمونه‌برداری از ۳۰ سانتی‌متر زیر سطح آب و در وسط رودخانه

شده براساس شاخص موسسه ملی بهداشت آمریکا و بررسی میزان آلودگی پارامترهای شاخص فوق در رودخانه می‌باشد. این ایستگاه‌ها به ترتیب در جدول ۴ معرفی گردیدند، همچنین موقعیت ایستگاه‌های مطالعاتی در شکل ۱ مشخص گردیده است.

جدول ۲- طبقه‌بندی شدت آلودگی رودخانه براساس شاخص NSFQI

Table 2: Classification of Pollution River by NSFQI index

مقدار شاخص NSFQI NSFWQI index	رنگ Color	وضعیت Condition
۹۱-۱۰۰ 91-100	آبی blue	عالی Excellent
۷۱-۹۰ 71-90	سبز green	خوب good
۵۱-۷۰ 51-70	زرد yellow	متوسط Middle
۲۶-۵۰ 26-50	نارنجی Orange	بد bad
۰-۲۵ 0-25	قرمز Red	خیلی بد Very bad

جدول ۳- پارامترهای وزن‌دهی در شاخص NSFQI

Table 3. Weighting parameters in NSFQI index

پارامتر parameter	اکسیژن محلول Do	کلیرم مدفوعی coliform	PH	فسفات Phosphate	دما Temperature	کدورت turbidity	BOD	نیترات Nitrate	TS
وزن Weight	0.17	0.16	0.11	0.1	0.1	0.08	0.11	0.1	0.07

جدول ۴- موقعیت ایستگاه‌های مورد مطالعه در طول مسیر رودخانه تالار [۷]

Table 4- Position of stations studied the Talar River

ایستگاه station	عرض جغرافیایی Latitude	طول جغرافیایی Longitude	ارتفاع متوسط حوزه Middle Elevation (m)	موقعیت ایستگاه‌های مورد مطالعه Location of stations studied
ورسک Versk	3793277	0678047	1608	ارتفاع متوسط حوزه
پل سفید polsefid	3998373	0685267	637	طول جغرافیایی
شیرگاه shrqah	4018780	0669383	253	عرض جغرافیایی
تالار Talar	4039124	0662855	36	ارتفاع متوسط حوزه
کیاکولا Kiakola	4047543	0662205	6	طول جغرافیایی
عرب خیل Arabkhill	4062293	0654998	1	عرض جغرافیایی

نرون‌های بیولوژیکی مغز انسان الهام گرفته و با ایجاد نگاشت‌های غیر خطی میان پارامترهای مستقل و پارامترهای هدف قادر به کشف ارتباط بین پارامترهای دخیل در مساله می‌باشند [۱۴].

از متداولترین انواع شبکه‌های عصبی، می‌توان به پرسپترون چند لایه اشاره کرد، که در حوزه‌های متنوعی از علوم مهندسی

می‌باشد. برداشت، نگهداری و سنجش با استفاده از روش‌های استاندارد متد در آزمایشگاه آب منطقه‌ای و حفاظت محیط‌زیست استان مازندران انجام گردیده است.

شبکه‌های عصبی مصنوعی: شبکه‌های عصبی مصنوعی^۱ از فعالیت

1. Artificial Neural Networks (ANNs)

$$F(x) = \frac{1}{1+e^{-x}} \quad (6)$$

در روابط بالا W_{ij} وزن اتصال بین نرون j در لایه و نرون i در لایه قبلی و X_i مقدار ورودی یا مقدار خروجی از نرون i در لایه قبلی، θ_j مقدار آستانه در گره j و O_j خروجی از نرون j است. در شکل ۲ عملکرد یک نرون نشان داده شده است.

برای به دست آوردن خروجی شبکه، لازم است که در ابتدا وزن‌های اتصالات بین نرون‌ها محاسبه شوند. این امر با الگوریتم‌های آموزشی انجام می‌شود.

الگوریتم پس انتشار خطا روشی برای پیدا کردن وزن‌های بهینه در یک شبکه پیش‌خور چند لایه است؛ و در دو مرحله اجرا می‌شود. اولین مرحله فاز رو به جلو است. در این مرحله، شبکه نمونه‌های آموزشی را دریافت کرده و با استفاده از وزن‌های اولیه، خروجی شبکه را پیش‌بینی می‌کند، سپس نتایج به دست آمده با مقادیر هدف مقایسه شده و خطا از رابطه (۷) محاسبه می‌شود.

$$E = \frac{1}{2} \sum_{n=1}^n (O_n - t_n)^2 \quad (7)$$

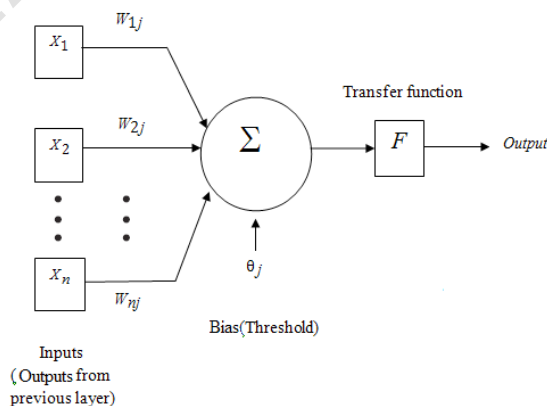
در رابطه بالا O_n و t_n ، به ترتیب خروجی شبکه و متغیر هدف از n امین گره خروجی است.

سپس فاز دوم یعنی فاز برگشت به عقب اجرا می‌شود. در این مرحله خطا به طور معکوس برای تنظیم و اصلاح وزن‌ها به شبکه تغذیه شده و به عقب منتشر می‌شود؛ و این فرآیند تاجایی ادامه پیدا می‌کند تا خطا حداقل شود.

فرآیند آموزش پس از رسیدن خطا، به سطح قابل قبول، یا پس از تکرار به تعداد تعیین شده، یا عدم کاهش خطا در تکرارهای متوالی متوقف می‌شود؛ در این حالت شبکه آموزش دیده است و با استفاده از رابطه (۸) وزن‌های نهایی ثابت می‌شوند.

$$\Delta W(k+1) = \alpha \Delta w(k) - \eta \frac{\sum E}{\sum W} \quad (8)$$

در رابطه بالا $\Delta w(k)$ تغییر وزن در تکرار k ام، $\Delta w(k+1)$ تغییر وزن در تکرار $k+1$ ام، α ضریب مومنتم و η نرخ یادگیری است.



شکل ۲ - عملکرد یک نرون
Figure 2 - Function of a Neuron

با موفقیت به کار برده شده‌اند. شبکه چند لایه با تغذیه رو به جلو (پیش‌خور) شامل یک لایه ورودی، یک یا چند لایه میانی (پنهان) و یک لایه خروجی است. تعداد لایه‌های پنهان، بستگی به پیچیدگی مساله داشته و با سعی و خطا به دست می‌آیند؛ از این رو همواره مساله‌ای مورد بحث بوده است [۶]. همچنین تعداد نرون‌ها در لایه ورودی و خروجی به تعداد پارامترهای ورودی (متغیرهای پیش‌بینی کننده) و پارامترهای خروجی (پارامترهای هدف) مساله بستگی دارد. علاوه بر این تعداد نرون‌ها در لایه‌های میانی نیز به پیچیدگی مساله مربوط بوده و با روش سعی و خطا به دست می‌آیند. در سال‌های اخیر تلاش‌هایی جهت انتخاب تعداد بهینه لایه‌های پنهان و تعداد نرون‌های این لایه انجام شده است. مطالعات و کارهای تئوریک بر روی پرسپترون چند لایه نشان داده است که کاربرد یک لایه پنهان برای مدل‌سازی و تقریب هر تابع پیچیده غیرخطی مناسب است. از این رو در این مطالعه از یک لایه پنهان استفاده شده است. هم‌چنین تعداد نرون‌های لایه پنهان از معیار ارائه شده توسط راجرز و دولا [۱۱] و هج - نلسون [۸] به دست می‌آیند.

$$N^H \leq 2N^I + 1 \quad (2)$$

$$(3)$$

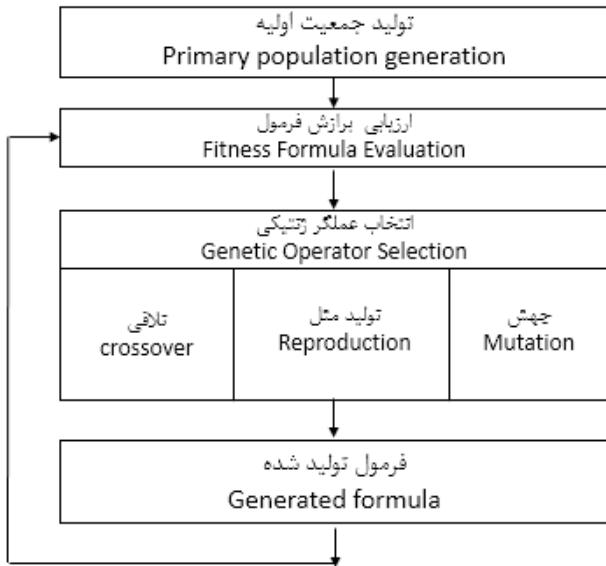
$$N^H \leq \frac{N^R}{(N^I + 1)}$$

در روابط بالا N^H تعداد نرون‌ها در لایه پنهان، N^I تعداد متغیرهای ورودی و N^R تعداد نمونه‌های آموزشی است. در فرآیند آموزش شبکه، هر ورودی نرون، در یک وزن اولیه ضرب شده و یک مقدار آستانه θ_j به مجموع آن‌ها اضافه می‌شود، رابطه (۴)، سپس نتیجه حاصل با عبور از یک تابع انتقال، خروجی O_j را نتیجه می‌دهد رابطه (۵). متداووترین تابع انتقال سیگموئید (F) است، که با رابطه (۶) تعریف می‌شود:

$$I_j = \sum_{i=1}^n W_{ij} X_i + \theta_j \quad (4)$$

$$O_j = F(I_j) \quad (5)$$

ه) مدل مورد نظر با استفاده از تابع برازش مورد ارزیابی واقع می‌شود.
 ۴- گام سوم تا نیل به حداکثر تعداد تولید، تکرار خواهد شد. فرم کلی گام‌های اجرایی برنامه‌ریزی ژنتیک در شکل ۳ نشان داده شده است.



شکل ۳- برنامه‌ریزی بیان ژنی
 Figure 3- gene expression planning

برای ارزیابی عملکرد و صحت‌سنجی مدل‌ها و مقایسه نتایج حاصل از روش‌های مختلف از شاخص‌های آماری نظیر ضریب تعیین R^2 ، جذر میانگین مربعات خطا $RMSE$ ، میانگین مطلق خطا MAE و شاخص پراکندگی SI استفاده شده است. ضریب تعیین شاخصی است که به منظور تعیین نوع همبستگی و میزان ارتباط بین دو متغیر به کار می‌رود و مقدار آن بین ۰ و ۱+ نوسان می‌کند، $(0 \leq R^2 \leq 1)$. اگر $R^2 = 1$ باشد، نشانگر این است که بین متغیرهای مورد بررسی همبستگی کامل و مستقیم وجود دارد و اگر $R^2 = 0$ این است که بین دو متغیر همبستگی وجود ندارد.

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \quad (9)$$

جذر میانگین مربعات خطا و خطای مطلق میانگین از شاخص‌های آماری هستند که بیانگر میزان خطا، بین داده‌های مشاهداتی و پیش‌بینی

برنامه‌ریزی بیان ژنی: برنامه‌ریزی بیان ژنی تعمیم یافته روش الگوریتم ژنتیک می‌باشد که برای اولین بار توسط جان کزا [۱۳] براساس تئوری داروین ارائه شد. برنامه‌ریزی ریزی بیان ژنی یک تکنیک برنامه‌ریزی خودکار می‌باشد که راه حل مسئله را با استفاده از برنامه کامپیوتری ارائه می‌کند. در این روش در ابتدای فرآیند، هیچگونه رابطه تابعی در نظر گرفته نشده و این روش قادر به بهینه‌سازی ساختار مدل و مولفه‌های آن می‌باشد. برخلاف الگوریتم ژنتیک روی ساختار درختی، فرمول‌ها به جای سلسله ارقام دودویی عمل می‌کند. ساختارهای درختی از مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) تشکیل می‌شوند.

قبل از مراحل اجرایی برنامه‌ریزی ژنتیک گام‌های مقدماتی زیر باید توسط کاربر تعیین شوند. ۱- مجموعه ترمینال‌ها (متغیرهای مساله، اعداد ثابت تصادفی)، ۲- مجموعه عملگرهای ریاضی مورد استفاده در فرمول‌ها، ۳- انتخاب معیار برازش مناسب (R , R^2 , $RMSE$, MSE و ...) برای سنجش برازش فرمول‌ها، ۴- تعیین پارامترهای کنترل کننده اجرایی برنامه (اندازه جمعیت، احتمال مربوط به بکارگیری اعمال ژنتیکی و جزئیات دیگر مربوط به اجرای برنامه) و ۵- معیار پایان اجرای برنامه و ارائه نتایج (مثل تعداد تولید جمعیت جدید، تعیین یک مقدار مشخص برای برازش فرمول‌ها که اگر میزان برازش برابر یا بیشتر از آن مقدار شد، اجرا متوقف شود). فرآیند اجرایی گام به گام برنامه‌ریزی ژنتیک به صورت مراحل زیر است:

- ۱- تولید یک جمعیت اولیه از فرمول‌ها که از ترکیب تصادفی مجموعه توابع (عملگرهای ریاضی مورد استفاده در فرمول‌ها) و ترمینال‌ها (متغیرهای مسئله و اعداد ثابت) ایجاد می‌شوند.
- ۲- هر یک از افراد جمعیت مذکور با استفاده از ملاک‌های برازش مورد ارزیابی قرار می‌گیرند.
- ۳- تولید یک جمعیت جدید از فرمول‌ها، بدین منظور مراحل زیر برای تولید یک جمعیت جدید دنبال می‌شوند: الف) یکی از عمل‌های ژنتیکی تلاقی، جهش و تولید مثل انتخاب می‌شود. این سه عمل ژنتیکی مهمترین عمل‌های ژنتیکی مورد استفاده در برنامه‌ریزی ژنتیک می‌باشند. عمل‌های دیگری مثل اصلاح ساختار و ... نیز با احتمال کمتر مورد استفاده قرار می‌گیرند.
- ب) تعداد مناسبی از افراد جمعیت حاضر انتخاب می‌شوند. انتخاب فرد یا افرادی از جمعیت مذکور به صورت احتمالاتی می‌باشد که در این انتخاب احتمالاتی منفردهای با برازش بهتر به منفردهای نامرغوب ترجیح داده می‌شوند و این بدان معنی نیست که حتما منفردهای نامرغوب حذف می‌شوند.
- ج) از عمل ژنتیکی انتخاب شده برای تولید فرزند (فرمول جدید) استفاده می‌شود.
- د) فرزند (فرمول جدید) تولید شده در یک جمعیت جدید وارد می‌شود.

1. Coefficient of determination (R^2)
2. Root Mean square error (RMSE)
3. Mean absolute error (MAE)
4. Scatter Index (SI)

Table 5- Final parameters of the gene expression programming

رابطه استخراج شده	پارامتر	parameter
relationship		
6	اندازه سر	Head size
30	تعداد کروموزوم‌ها	Number of Crunches
3	تعداد ژن‌ها	Number of genes
e, +, -, /, ×, توان, e	عملگرهای ریاضی به کار برده شده	Math Operators Used
+	عملگر ریاضی بین ژن‌ها	Math Operator between genes
0.044	سرعت جهش	mutation
0.1	سرعت وارونگی	Speed of inversion
0.3	سرعت تلاقی یک نقطه‌ای	The speed of a point crossing
0.3	سرعت تلاقی دو نقطه‌ای	The speed of a point crossing
0.1	سرعت تلاقی ژن‌ها	The speed of crossing the genes
0.1	سرعت جابجایی ژن‌ها	The speed of the genes

$$NSFWQI = F(D, E, H, S, BOD, T, D, P, Tur)$$

که در رابطه فوق پارامترها به ترتیب نشان دهنده شاخص تعیین کیفیت آب، میزان اکسیژن محلول، کلروفورم مدفوعی، پی‌اچ، جامدات محلول، اکسیژن مورد نیاز بیولوژیکی در ۵ روز، دما، نیترات، فسفات و کدورت می‌باشند.

جهت توسعه مدل برنامه ریزی بیان ژنی، داده‌ها را به صورت تصادفی به دو گروه آموزش (۷۰ درصد داده‌ها، ۴۸ داده) و آزمون (۳۰ درصد داده‌ها، ۲۴ داده) تقسیم می‌نماییم. این کار در بیشتر مطالعات مبتنی به روش‌های نوین محاسباتی مانند شبکه‌های عصبی مصنوعی، برنامه ریزی بیان ژنی، مدل‌های درختان تصمیم انجام می‌شود. در جدول ۵ پارامترهای نهایی مدل برنامه‌ریزی بیان ژنی برای بدست آوردن رابطه‌ای از این مدل ارائه شده است.

برای مدل‌سازی رابطه کیفیت آب رودخانه تالار با استفاده از برنامه‌ریزی بیان ژنی دو مجموعه از عملگرهای ریاضی، مجموعه یک و دو به ترتیب، مجموعه یک: [+, -, /, ×, Exp, Sin, Cos] در نظر گرفته شده است که مجموعه یک، عملگرهای پیش فرض نرم‌افزار (معمولا بهترین عملگرها) و مجموعه دو، [+, -, /, ×] چهار عمل اصلی (عملگرهایی که مدل حاصل ساده‌ترین است) می‌باشد. با مقایسه پارامترهای آماری نتایج حاصل از دو عملگر ریاضی، نتیجه می‌شود که دقت آن‌ها نزدیک به هم بوده، بنابراین مدل حاصل از عملگرهای ریاضی مجموعه دو به علت سادگی مدل، به عنوان مدل انتخاب گردید (رابطه ۱۳).

$$NSFWQI = (D + 9.6) + \left[\text{Arc tan} \left(\frac{D \times 9.9}{Tur \times P} \right) \times D \right] + \text{Cos} \left(\frac{(-7.5 \times BOD)^3}{(P + H)} \times T \right) + 8.0 + \ln(D)$$

(۱۳)

شده هستند، و در کاربردهای مهندسی فراوان استفاده می‌شوند.

$$RMSE = \sqrt{\frac{\sum_i (x_i - y_i)^2}{n}} \quad (10)$$

$$MAE = \frac{1}{n} \sum_i |x_i - y_i| \quad (11)$$

شاخص پراکندگی نشان دهنده چگونگی پراکندگی بین مقادیر مشاهداتی و پیش‌بینی شده است، و به صورت زیر تعریف می‌شود.

$$S = \frac{\sqrt{\frac{\sum_i (x_i - y_i)^2}{n}}}{x} \quad (12)$$

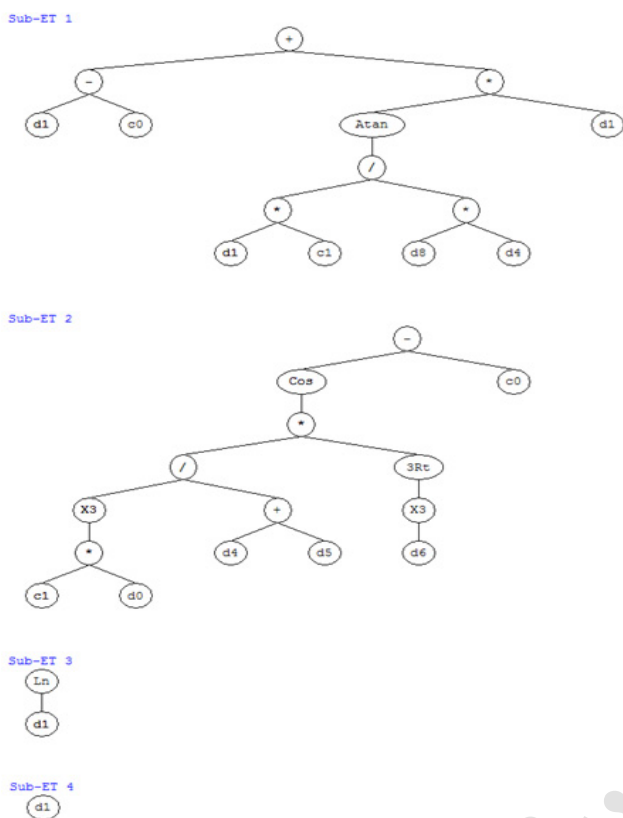
در روابط فوق x_i مبین پارامتر مشاهداتی و y_i مبین پارامتر پیش‌بینی شده و x و y به ترتیب نشان‌دهنده میانگین مقادیر مشاهداتی و پیش‌بینی شده است. n نیز تعداد داده‌ها می‌باشد. مدل بهینه، مدلی است که دارای ضریب همبستگی بالاتر و کمترین مقدار شاخص پراکندگی، جذر میانگین مربعات خطا و میانگین مطلق خطا باشد.

نتایج و بحث

نتایج آزمایش‌ها پس از بررسی صحت و پایش داده‌ها تحلیل گردید و با توجه به روابط فوق‌الذکر شاخص کیفیت آب محاسبه شد. سپس با استفاده از مدل شبکه‌های عصبی مصنوعی و روش برنامه‌ریزی بیان ژنی به منظور تخمین شاخص کیفیت آب رودخانه تالار استفاده گردید.

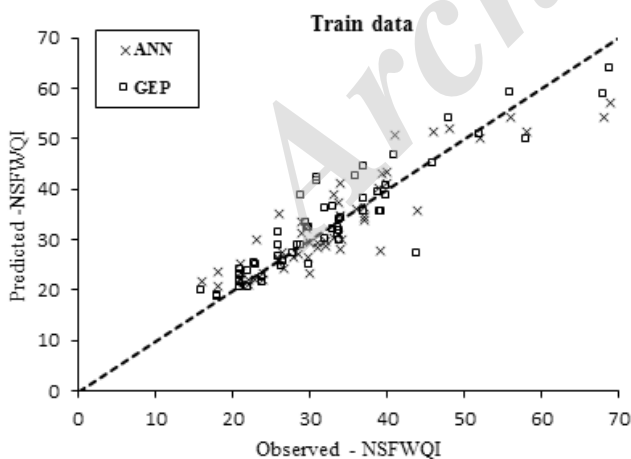
برای انجام این پژوهش شاخص NSFWQI با استفاده از پارامترهای زیر و با استفاده از روش شبکه‌های عصبی مصنوعی و برنامه ریزی بیان ژنی تخمین زده می‌شود.

شکل ۷ محور افقی شماره داده‌ها و محور قائم شاخص NSFWQI را برای داده‌های مشاهداتی و تخمین زده شده توسط روش‌های برنامه‌ریزی بیان ژن و شبکه عصبی نشان می‌دهد. همانطور که مشاهده می‌شود شکل بیان کننده همبستگی بالای بین داده‌های مشاهداتی و محاسباتی با



شکل ۴- رابطه درختی برنامه ریزی بیان ژنی

Figure 4. Tree Relationship of gene expression planning



شکل ۵- مقایسه نتایج مشاهداتی و محاسباتی برای داده‌های مجموعه آموزش

Figure 5- Comparison of the observational and computational results for the training set data

جدول ۶ نتایج پارامترهای آماری را برای رابطه تعیین شده شاخص کیفیت آب رودخانه تالار در دو مرحله آموزش و آزمون نشان می‌دهد. با توجه به بالا بودن ضریب همبستگی و قابل پذیرش بودن میزان خطا، فرمول‌های استخراج شده توسط برنامه‌ریزی ژنتیک مورد قبول می‌باشد.

به منظور مدلسازی با روش شبکه‌های عصبی مصنوعی از همان داده‌هایی که برای آموزش و آزمون روش برنامه ریزی بیان ژنی استفاده شد، به کار رفت. با استفاده از روش آزمون و خطا بهترین آرایش شبکه بدست آمد، که نتایج آن در جدول ۷ نشان داده شده است. ضریب ممتوم ۰/۹، نرخ یادگیری ۰/۲۵ و آرایش بهینه شبکه‌های عصبی مصنوعی $9 \times 10 \times 1$ (۹ لایه ورودی، ۱۰ لایه پنهان و ۱ لایه خروجی) به دست آمد. با توجه به جدول ۷، بالا بودن مقدار ضریب همبستگی (۰/۹۲۹) برای داده‌های آموزش و ۰/۹۲۶ برای داده‌های آزمون) و قابل پذیرش بودن میزان خطا، می‌باشد، مدل به دست آمده از روش شبکه‌های عصبی مصنوعی قابل پذیرش می‌باشد.

شکل‌های ۵ و ۶ نتایج بدست آمده با استفاده از برنامه‌ریزی بیان ژنی و شبکه‌های عصبی مصنوعی را برای داده‌های مشاهداتی و محاسباتی نشان می‌دهد. همانطور که مشاهده می‌کنید شکل‌ها بیان کننده همبستگی بالای بین داده‌های مشاهداتی و محاسباتی با استفاده از روش‌های شبکه عصبی و برنامه‌ریزی بیان ژنی می‌باشند.

جدول ۶- شاخص‌های آماری مجموعه آموزش و آزمون برای مدل برنامه‌ریزی بیان ژنی

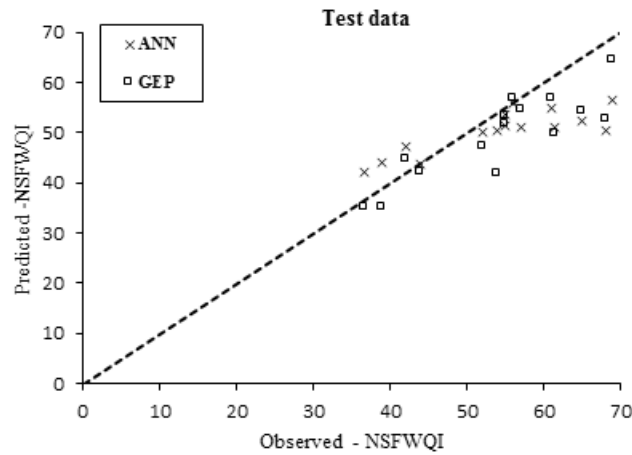
Table 6 - Statistical Index of Training and Testing for Gene Expression Program Model

	R^2	RMSE	MAE	δ
آموزش	0.873	4.92	3.4	0.131
Training				
آزمون	0.840	5.77	4.34	0.15
Testing				

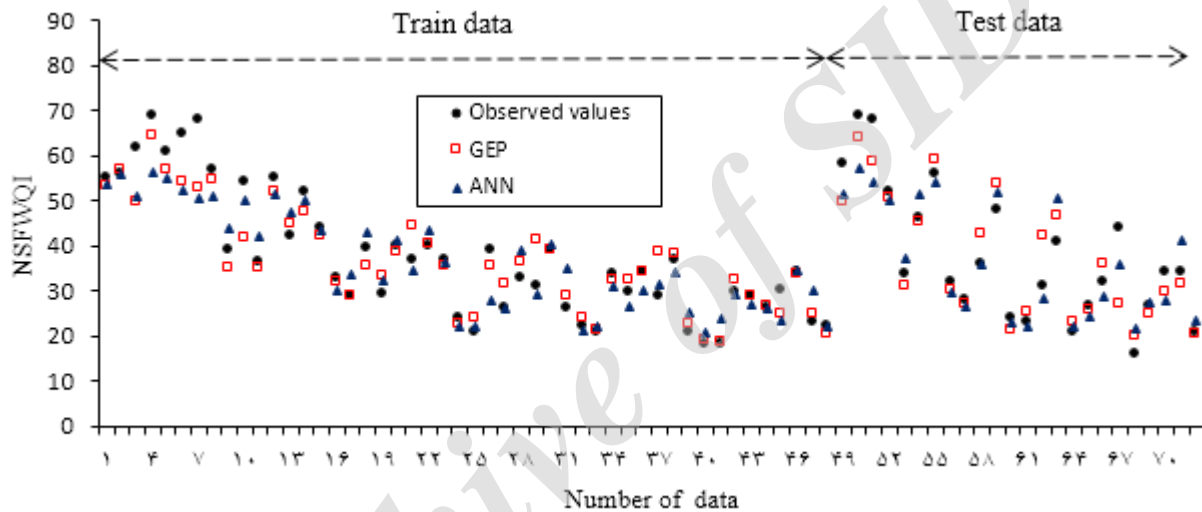
جدول ۷- شاخص‌های آماری مجموعه آموزش و آزمون برای مدل شبکه‌های عصبی مصنوعی

Table 7 - Statistical Index of Training and Testing for Artificial Neural Network Models

مجموعه	R^2	RMSE	MAE	δ
آموزش	0.842	5.492	4.01	0.146
Training				
آزمون	0.849	5.60	4.30	0.149
Testing				



شکل ۶- مقایسه نتایج مشاهداتی و محاسباتی برای داده‌های مجموعه آزمون
Figure 6. Comparison of observational and computational results for test set data



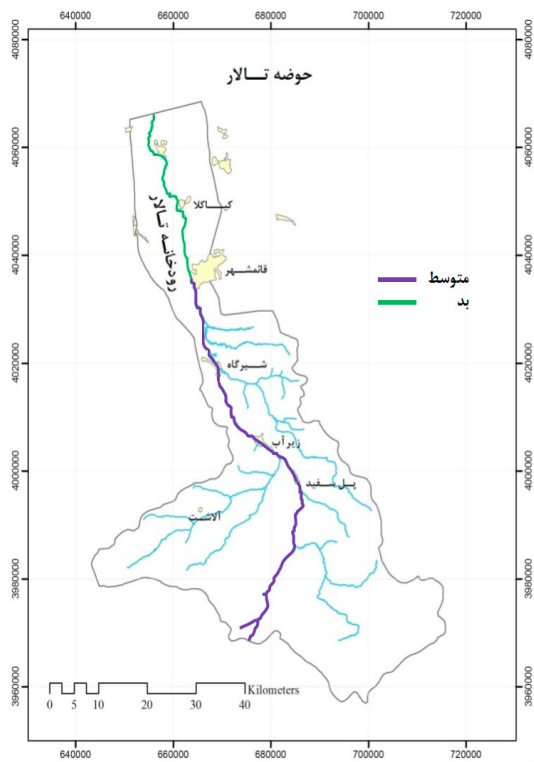
شکل ۷- روند تغییرات داده‌های مشاهداتی در مقایسه با تخمین‌های شبکه عصبی و برنامه‌ریزی بیان ژن
Figure 7- Changes in observational data compared to neural network estimates and gene expression planning

علاوه بر این مدل برنامه ریزی بیان ژنی قادر به ارائه روابط ریاضی می‌باشد که می‌تواند توسط دیگر کاربران در محاسبات استفاده شود. در حالی که مدل شبکه‌های عصبی مصنوعی مانند یک جعبه سیاه می‌باشد که برخلاف برنامه ریزی بیان ژنی رابطه صریحی بین پارامترهای موثر در مساله ارائه نمی‌دهد.

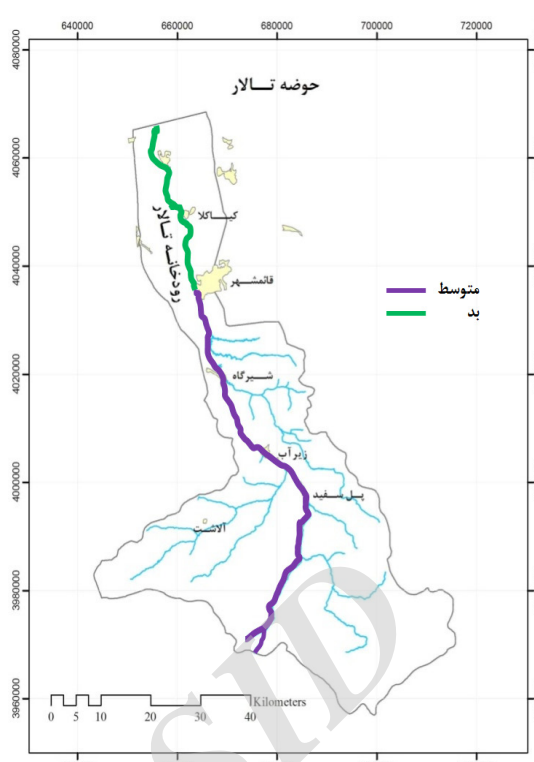
به طور کلی آلاینده‌های موجود در منطقه به دو دسته نقطه‌ای و غیرنقطه‌ای تقسیم می‌شوند. آلودگی نقطه‌ای معمولاً در نقاط معین و مشخص وارد آب‌های پذیرنده می‌شوند، چه به صورت تک نقطه‌ای و یا چند نقطه‌ای که حامل مواد آلاینده هستند. این آلاینده‌ها نسبتاً حجیم و دارای کیفیت پایداری هستند ولی آلودگی‌های غیرنقطه‌ای را به سختی می‌توان تشخیص داد و تحت کنترل قرار در آورد. آلودگی‌های نقطه‌ای شامل خروجی فاضلاب شهری و صنعتی،

استفاده از روش‌های شبکه عصبی و برنامه‌ریزی بیان ژنی می‌باشند. در ادامه کیفیت آب رودخانه تالار در طول مسیر براساس شاخص NSFWQI مورد پهنه‌بندی قرار گرفت. نقشه‌های به دست آمده در محیط ArcGIS 10 برای پهنه‌بندی کیفی آب در طول رودخانه براساس شاخص فوق در دوره‌های و ماه‌های مختلف سال در ایستگاه‌های مورد نظر در شکل ۲ داده شده است. بر این مبنا، سه گروه کیفیت خوب، متوسط و بد بر طبق طبقه‌بندی شاخص کیفیت آب به ترتیب با رنگ‌های آبی پررنگ، بنفش و سبز نشان داده شده است. مقایسه نتایج پژوهش حاضر با نتایج پژوهش رضانی و همکاران تطابق خوبی را نشان داد.

بررسی نتایج فوق نشان داد که دقت مدل‌های شبکه‌های عصبی مصنوعی و برنامه‌ریزی بیان ژنی بسیار به هم نزدیک می‌باشد.



ب



الف

شکل ۸- نقشه پهنه‌بندی کیفی رودخانه تالار براساس شاخص NSFQI. الف) معادله ۱؛ ب) معادله ۱۳
Figure 8 - quality zoning map based on Talar River in NSFQI index, a) Equation 1; b) Equation 13

سیلاب‌ها و هم‌چنین تخلیه و تلبار مواد زائد جامد در کنار رودخانه باشد. این نتایج با نتیجه رضانی و همکاران که بر روی رودخانه تالار پژوهشی انجام دادند منطبق می‌باشد. به طور کلی می‌توان نتایج را اینگونه تحلیل نمود که رودخانه تالار در طول مسیر خود از ایستگاه ورسک تا خروجی دریای خزر شاهد افزایش بار آلودگی و کاهش کیفیت آب بوده است.

آنالیز حساسیت

به‌منظور تعیین اهمیت نسبی پارامترهای مستقل بر شاخص NSFQI، مدل‌های جدیدی با استفاده از حذف تک تک پارامترهای موثر بر شاخص NSFQI توسعه می‌یابند. به این صورت که ابتدا یک پارامتر موثر بر شاخص NSFQI از رابطه کلی حذف شده و سپس با بقیه پارامترهای موثر مدل‌هایی با استفاده از شبکه‌های عصبی مصنوعی توسعه می‌یابند و در انتها عملکرد و دقت مدل‌های توسعه یافته به‌وسیله مجموعه آزمون ارزیابی می‌شوند. در جدول ۵ مدل‌ها و نتایج آنالیز حساسیت آورده شده است. همان‌طور که از نتایج این جدول مشخص است پارامتر Do بیشترین و پارامتر TS کمترین اهمیت را در تخمین شاخص NSFQI دارند.

شرابه و پساب‌های محل دفن زائدات، پساب ناشی از فعالیت‌های دامداری و ... می‌باشند و آلودگی‌های غیرنقطه‌ای آب برگشتی از مزارع، پساب کشاورزی و ... می‌باشند. با بررسی‌های انجام شده در طول رودخانه تالار، ۲۳ مورد آلاینده نقطه‌ای شناسایی گردید.

با توجه به نتایج حاصل از پژوهش رودخانه تالار در معرض آلودگی شدید قرار گرفته است که می‌توان این آلودگی را اینگونه تحلیل نمود. میزان اکسیژن محلول طی مدت بررسی از ایستگاه ورسک به سمت شیرگاه با روند افزایشی روبه‌رو بوده است، ولی در ایستگاه‌های کیاکلا و عرب‌خیل به دلیل ورود فاضلاب شهرهای قائمشهر، کیاکلا، بهنمیر و مرکز دفن زباله کیاکلا و کشتارگاه‌های مرغ‌شهرداری و پرستو و هم‌چنین شهرک صنعتی سنگتاب به طور چشمگیری کاهش پیدا کرده است. بیشترین میانگین BOD مربوط به ایستگاه‌های کیاکلا و عرب‌خیل بوده است که دلیل آن افزایش اکسیژن‌خواهی بیوشیمیایی مواد آلی موجود در فاضلاب‌های تصفیه‌نشده، کودها و زائدات جانوری می‌باشد. بررسی شاخص NSFQI نشان می‌دهد آب رودخانه در ایستگاه ورسک دارای کیفیت متوسط و سایر ایستگاه‌ها دارای کیفیت بد می‌باشد که این امر می‌تواند ناشی از تخلیه مستقیم فاضلاب‌های روستایی، شهری و صنعتی به داخل رودخانه و ورود آن‌ها به داخل رودخانه از طریق

Table 5 - Sensitivity analysis of effective parameters on NSFQI index using artificial neural network model

model مدل	R^2	RMSE	MAE
$NSFWQI = F(E, H, S, BOD, T, D, P, Tur)$	0.618	8.923	7.560
$NSFWQI = F(D, H, S, BOD, T, D, P, Tur)$	0.639	8.688	6.558
$NSFWQI = F(D, E, S, BOD, T, D, P, Tur)$	0.665	8.363	6.806
$NSFWQI = F(D, E, H, S, BOD, T, D, Tur)$	0.668	8.325	6.301
$NSFWQI = F(D, E, H, S, T, D, P, Tur)$	0.721	8.208	6.124
$NSFWQI = F(D, E, H, S, BOD, T, P, Tur)$	0.734	7.945	5.590
$NSFWQI = F(D, E, H, S, BOD, D, P, Tur)$	0.747	7.320	5.543
$NSFWQI = F(D, E, H, S, BOD, T, D, P)$	0.763	6.832	5.232
$NSFWQI = F(D, E, H, BOD, T, D, P, Tur)$	0.789	6.632	4.447

Ankara, Turkey.

4- Effendi, H. and Wardiatno, R. Y., 2015. Water Quality Status of Ciambulung River, Banten Province, Based on Pollution Index and NSF-WQI. *Procedia Environmental Sciences*, 24: 228-237.

5- Fabiano, D., Santos, S., Altair, B., Moreira. and Sonia, M., 2008. Water quality index as a simple indicator of aquaculture effects on aquatic bodies. *ecological indicators* 8, 476 – 484.

6- Hornik K, Stinchcombe M, and White H, (1989). "Multilayer feedforward networks are universal approximators". *Neural Networks*. p. 359-366.

7-Hoseyni, M., Jafari, H., Salari, F., 2011. Measuring the Remnants of Organochlorine Organic Toxins in the Hall and Tajan Rivers, Research Project of Regional Water Organization of Mazandaran Province, pages 122 and 126

8- Hecht-Nielson R. (1987). "Kolmogorov's mapping neural network existence theorem". In *Proceedings of the first IEEE international joint conference on neural networks*, New York, pp.3-11.

9- Oram, B., 2011. Calculating NSF Water Quality Index. Wilkes University Center for Environmental Quality GeoEnvironmental Sciences and Engineering Department.

16- Puri, P. J., Yenkie, M. K. N., Sangal, S. P, Gandhare, N. V., Sarote, G. B. and Dhanorkar, D. B., 2011. Surface water (lakes) quality assessment in Nagpur City (India) based on water quality index (WQI). *Rasayan Journal of Chemistry*, 4(1): 43-48.

10- Rosli, N. A., Zawawi, M. H. and Bustami, R. A., 2012.

نتیجه گیری

با توجه به اهمیت رودخانه تالار در تامین آب کشاورزی و پرورش ماهی اطراف این رودخانه و همچنین تخلیه آلاینده‌های متعدد به آن، ارزیابی کیفی آب این رودخانه ضروری به نظر می‌رسید. از طرفی ارائه رابطه‌ای تجربی میان عوامل تعیین کننده آلودگی آب باعث می‌گردد تا هرگونه تصمیم‌گیری مدیریتی که اثرات زیست‌محیطی آن به صورت مستقیم و یا غیر مستقیم متوجه آب‌های سطحی کشور باشد، با آگاهی بیشتری اتخاذ گردد. در این مطالعه با استفاده از روش برنامه‌ریزی بیان ژنی رابطه‌ای برای برآورد میزان آلودگی آب رودخانه تالار تعیین گردید سپس با استفاده از روش شبکه‌های عصبی مصنوعی نیز مدل‌سازی انجام گردید. نتایج به دست آمده از این مطالعه نشان داد که دو روش مورد استفاده دارای شاخص‌های آماری نزدیک به هم می‌باشند. شاخص‌های آماری نشان داد که دقت روش شبکه‌های عصبی مصنوعی و برنامه‌ریزی بیان ژن بسیار به هم نزدیک می‌باشند و مدل شبکه‌های عصبی مصنوعی قادر است اندکی بهتر شاخص NSFQI را برآورد نماید. علاوه بر این برنامه ری بی بیان ژنی با ارائه رابطه ریاضی می‌تواند جهت تخمین شاخص NSFQI توسط دیگر محققین مورد ارزیابی قرار گیرد.

منابع

1- Aminpur shiani, S., mohammadi, M., khaledian, M., mir roshndel, A., 2016. Water quality evaluation of Gazroudbar River using NSFQI and Liou indices. 3.; 8 (1) :65-78.

2- AYTEK A., and KISI O. 2008. A genetic programming approach to suspended sediment modeling. *Journal of Hydrology*, 351:288-298

3- Dogan E., Isik S., Toluk T., and Sandalci M. 2007. Daily streamflow forecasting using artificial neural networks. Pp. 448-459. *International Congress River Flood Management*.

14- Samadi, M., Jabbari, E., Azamathulla, H.M. and Mojallal, M., 2015. Estimation of scour depth below free overfall spillways using multivariate adaptive regression splines and artificial neural networks. *Engineering Applications of Computational Fluid Mechanics*, 9(1), pp.291-300.

15- Vaheedunnisha, S. and Sandeep, K., 2013. Water Quality Assessment of RoopSagar Pond of Satna Using NSF-WQI. *International Journal of Innovative Research in Science, Engineering and Technology*, 2(5): 1386-1388.

Salak River Water Quality Identification and Classification According to Physico-Chemical Characteristics. *Procedia Engineering*, 50: 69-77.

11- Rogers LL, and Dowla FU. (1994). "Optimization of ground water remediation using artificial neural networks with parallel solute transport modeling ". *Water resource research* 30(2), pp. 457-481.

12-Sanchez, E., Manuel, F., Colmenarejo, V., Angel, M., Garcı, T. and Rafael, B., 2007. Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution.

13-Salamati S. Dehghani A A. Meftah Halaghi M. and Zahiri A R. 2015. Estimation of Discharge Coefficient of Combined Weir-Gate Structure by Genetic Programming. *J. of Water and Soil Conservation*, Vol. 22(2).

Archive of SID

Abstract

Estimation of Water Quality Index in Talar River Using Gene Expression Programming and Artificial Neural Networks

B. Mojaradi¹, F. Alizadeh Sanami² and M. Samadi³

Received:2016/08/25 Accepted : 2018/04/11

Historically, rivers as a major source to supply drinking water and agriculture are considered in human societies as a very effective in the formation of civilizations. At present, most of the northern rivers that are under the influence of human, cause a variety of changes and destruction of ecosystem. Pollution and human interfaces could be the most important factors to pollution caused by waste water of industrial, urban and rural, pollution resulting from disposal of pesticides used in agriculture, destruction of vegetation, construction of dam, and barriers under bridge, blocking the mouth of the river, illegal fishing. Hence, considering the importance of Talar River for supply agriculture water, and also pour pollution to in, to identify and assess the river water quality and provide the necessary relationship to estimate pollution and water quality. In current study, 72 samples were used during the years of 1391 to 1392 from 6 stations namely, weresk, pole sefid, Shirgah, Talar, Kiakola and Arabkhil. Then, NSFQI index was estimated. Finally, with applying gene-expression programming and artificial neural network, the obtained models for determination of relation between water quality parameters and water quality index are used with high accuracy. To evaluate the performance of models, statistical parameter such as, root mean square error, mean absolute error, scatter index and correlation coefficient were used. Results showed that the proper functioning of artificial neural networks and planning methods in estimating gene expression index is NSFQI. The obtained results revealed appropriate performance of that artificial neural and gene-expression programming to estimate NSFQI index.

Keywords: Modeling, Talar River, Gene expression programming, Artificial neural networks, NSFQI index.

1. Assistant professor of water engineering and hydraulic structures, Faculty of civil engineering, Iran University of science and technology. Extension Organization, and PhD Candidate, University of Tehran, Corresponding Author, Email: kamali_kourosh@yahoo.com

2. Ph.D student of water resources engineering, Faculty of civil engineering, Iran university of science and technology. Corresponding author Email: Forough_alizadeh@civileng.iust.ac.ir. Extension Organization, and PhD Candidate, University of Tehran, Corresponding Author, Email: kamali_kourosh@yahoo.com

3. Ph.D student of water resources engineering, Faculty of civil engineering, Iran university of science and technology.