

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده: کشاورزی

گروه آب و خاک

پایان نامه کارشناسی ارشد

برآورد پارامترهای کیفی رودخانه های تجن و بابلرود مازندران با استفاده از مدل برنامه

ریزی بیان ژن GEP

مهدی یوسفی

اساتید راهنما :

دکتر هادی قربانی

دکتر صمد امامقلی زاده

بهمن ۱۳۹۴

تقدیم بہ

مقدس ترین واژه مادر لغت نامہ دلم، مادر مہربانم کہ زندگیم را دیون مہر و عطوفت آن می دانم

پدر، مہربانی مشفق، برادر و حامی

ہمسرم کہ نشانہ لطف الہی در زندگی من است

برادر و خواہرم بہر اہل ہمشکی و پشتونہ ہای زندگیم

و پسر مہدیار کہ آیندہ ام از آن اوست

تشکر و قدردانی

شنای بی پایان ذات اقدس یکرانی را که آفرینش را دستمایه زیستن قرار داد و زیستن را موقوف آموختن.

حال که به لطف قلم صنع آورش این پژوهش به پایان رسید بر خود لازم می دانم از تمامی بزرگان و بزرگ اندیشانی که به نوعی از الطاف وجودیشان بهره برده ام نهایت تشکر و قدردانی را داشته باشم.

این پژوهش اگر سلامتی یافته، در سایه بهره گیری از دانش و راهنمایی خردمندانه استاد ارجمند، جناب آقای دکتر مادی قربانی بوده است. همواره پاسکزار ایشان و قدرشناس راهبری پیش برنده و همراهی عالمانه شان خواهیم بود.

استاد ارجمند، جناب آقای دکتر صد اما مقلی زاده را پاس می گویم که علاوه بر آموختن نکته های ماندگار از محضر ایشان در مدت شاکردی، افتخار بهر یابی از نظرت فاضلانه ایشان را به عنوان استاد راهنما داشته ام.

هم چنین کمال امتنان خویش را تقدیم استاید فریخته داور می دارم که مژده نجاتی ارزشمند در جهت تقویت اثر خوانند بود. هر چند نهایت پاس من در این گفتار، حکم چشیدنی را دارد به قدر نمئی برای کسی که او را توان کشیدن آب دریا نیست.

در نهایت از خانواده محترم و سر اوار شنایم به خاطر همه صبور و بزرگواریشان کمال پاس و تشکر را دارم.

و از خداوند بزرگ که در راه آموختنم پیش برد و مجال اندیشیدن و نوشتنم داد به آرزوی خواهم، عنایت خویش را برای کارهای دیگر و بهتر شامل عالم فرماید.

چکیده

اکوسیستم‌های آبی از جمله رودخانه‌ها به عنوان یکی از اجزای سازنده این کره خاکی به شمار آمده که نه تنها زیستگاه‌های مهمی را برای موجودات آبی تشکیل می‌دهند، بلکه منبع مهمی از غذا را برای مصارف انسان شامل می‌شوند هم اکنون بر همگان ثابت شده است که کیفیت آب علاوه بر کمیت آن دارای اهمیت فوق‌العاده می‌باشد. واژه کیفیت آب نیز در توضیح خصوصیات فیزیکی، شیمیایی و بیولوژیکی برای استفاده برای هدف خاص مورد استفاده قرار می‌گیرد. مطالعات کیفیت آب به علت کاهش چشمگیر منابع آبی در سال‌های اخیر بسیار با اهمیت تر گردیده است. از طرفی اندازه‌گیری پارامترهای کیفیت آب بسیار پرهزینه بوده و اگر بتوان بعضی پارامترها را از روی پارامترهای دیگر اندازه‌گیری شده برآورد کرد، یک مسئله مهم می‌باشد در پژوهش حاضر به منظور برآورد پارامترهای کیفی آب رودخانه‌های تجن و بابلرود مازندران، روش برنامه ریزی بیان ژن (GEP) توسعه داده شده است. لذا اطلاعات هیدرومتری این دو رودخانه مربوط به سالهای ۱۳۵۰-۱۳۹۰ جمع‌آوری شد و مورد بررسی قرار گرفت. برای اجرای مدل برنامه ریزی بیان ژن ابتدا با استفاده از نرم افزار SPSS و با روش Stepwise Regression مدل‌های پیشنهادی تهیه شد، سپس با استفاده از نرم افزار Gene xprotools مدل‌ها با شاخص‌های آماری (ضریب همبستگی، مجذور میانگین مربعات خطا و میانگین قدر مطلق خطاها) مورد سنجش قرار گرفت. نتایج اجرای مدل برنامه ریزی بیان ژن در دو بخش آزمون و آموزش نشان داد که برنامه ریزی بیان ژن قابلیت بالایی در تخمین TDS و SAR رودخانه‌های بابلرود و تجن با ارائه رابطه ریاضی دقیق دارا می‌باشد.

کلمات کلیدی: برنامه ریزی بیان ژن، رودخانه تجن، رودخانه بابلرود، برآورد TDS، برآورد SAR

مقالات مستخرج از پایان نامه:

۱- بررسی کیفی آب رودخانه تجن از جنبه مصرف کشاورزی، اولین همایش ملی کیفیت

منابع آب و توسعه پایدار، آبان ۱۳۹۴

۲- پیش‌بینی شاخص کیفی کل مواد محلول (TDS) رودخانه تجن با استفاده از برنامه ریزی

بیان ژن (GEP)، اولین همایش ملی کیفیت منابع آب و توسعه پایدار، آبان ۱۳۹۴

فهرست مطالب

۱-۱-تعریف مسئله	۲
۲-۱-روش انجام تحقیق	۳
۳-۱-اهداف و نوآوری	۴
۱-۲-مقدمه	۶
۲-۲-پیشینه مطالعات برآورد پارامترهای کیفی آب	۷
۳-۲-پیشینه مطالعات رودخانه های بابلرود و تجن	۹
۴-۲-پیشینه مطالعات روش های هوشمند	۱۰
۲-۳-مناطق مورد مطالعه	۱۶
۱-۲-۳-رودخانه تجن	۱۶
۲-۲-۳-رودخانه بابلرود	۱۸
۳-۳-داده های مورد استفاده	۱۹
۴-۳-آشنایی با مفاهیم زیست شناختی طبیعی	۲۱
۵-۳-الگوریتم های تکاملی	۲۱
۱-۵-۳-الگوریتم ژنتیک (GA)	۲۳
۱-۱-۵-۳-عملگرهای الگوریتم ژنتیک	۲۳
۱-۱-۵-۳-عملگر انتخاب و تولید مثل	۲۳
۲-۱-۵-۳-عملگر تقاطع (ترکیب)	۲۴
۱-۲-۱-۵-۳-ترکیب تک نقطه ای	۲۴
۲-۲-۱-۵-۳-ترکیب دو نقطه ای	۲۴
۳-۲-۱-۵-۳-ترکیب یکنواخت	۲۵
۳-۱-۵-۳-عملگر جهش	۲۶
۲-۵-۳-برنامه ریزی ژنتیک (GP)	۲۷
۱-۲-۵-۳-عملگر ترکیب	۲۸
۲-۲-۵-۳-عملگر جهش	۲۹
۳-۲-۵-۳-عملگر جایگشت	۳۰
۳-۵-۳-برنامه ریزی بیان ژن (GEP)	۳۲
۱-۳-۵-۳-الگوریتم های برنامه ریزی بیان ژن	۳۲
۲-۳-۵-۳-ساختار برنامه ریزی بیان ژن	۳۵

۳۶.....	۳-۳-۵-۳ ساختار افراد در GEP
۳۶.....	۳-۳-۵-۳-۱ ORF و بیان درختی ژن ها
۳۸.....	۳-۳-۵-۲ ساختارهای ژن ها در برنامه ریزی بیان ژن
۴۱.....	۳-۳-۵-۴ تابع برازش
۴۲.....	۳-۶-۶ مراحل اجرای مدل
۴۲.....	۳-۶-۱ تعیین همبستگی بین داده ها
۴۳.....	۳-۶-۲ برنامه ریزی بیان ژن
۴۳.....	۳-۶-۱-۲ تعیین تابع برازش
۴۴.....	۳-۶-۲-۲ انتخاب مجموعه ترمینال ها و مجموعه توابع
۴۴.....	۳-۶-۲-۳ انتخاب ساختار کروموزوم
۴۴.....	۳-۶-۲-۴ انتخاب تابع پیوند
۴۵.....	۳-۶-۲-۵ انتخاب عملگرهای ژنتیکی
۴۷.....	۳-۶-۲-۶ الگوریتم برنامه ریزی بیان ژن
۴۹.....	۳-۷ شاخص های آماری سنجش مدل ها
۵۲.....	۴-۱ مقدمه
۵۲.....	۴-۲-۱ ارزیابی کیفی آب از جنبه مصرف کشاورزی
۵۲.....	۴-۲-۱-۱ ارزیابی کیفی رودخانه تجن
۵۳.....	۴-۲-۲-۱ ارزیابی کیفی رودخانه بابلرود
۵۴.....	۴-۳-۱ رودخانه تجن
۵۵.....	۴-۳-۱-۱ برآورد شاخص کیفی کل مواد محلول
۵۵.....	۴-۳-۱-۱-۱ همبستگی و رگرسیون
۵۷.....	۴-۳-۱-۲ برنامه ریزی بیان ژن
۶۵.....	۴-۲-۲-۲ برآورد نسبت جذب سدیم SAR
۶۵.....	۴-۲-۲-۱-۱ همبستگی و رگرسیون
۶۶.....	۴-۲-۲-۲-۲ برنامه ریزی بیان ژن
۷۰.....	۴-۳-۱ رودخانه بابلرود
۷۰.....	۴-۳-۱-۱ برآورد شاخص کیفی کل مواد محلول TDS
۷۲.....	۴-۳-۱-۲ برنامه ریزی بیان ژن
۸۰.....	۴-۲-۳-۲ برآورد نسبت جذب سدیم SAR
۸۰.....	۴-۲-۳-۱ همبستگی و رگرسیون

۸۱	 ۲-۲-۳-۴- برنامه ریزی بیان ژن
۹۰	 ۱-۵- نتیجه گیری
۹۲	 ۲-۵- پیشنهادات
۹۳	 منابع

فهرست اشکال

شکل (۱-۳)	نقشه محدوده مورد مطالعه رودخانه تجن	۱۷
شکل (۲-۳)	محدوده مورد مطالعه رودخانه بابلرود	۱۹
شکل (۳-۳)	ترکیب تک نقطه ای	۲۴
شکل (۴-۳)	ترکیب دو نقطه ای	۲۵
شکل (۵-۳)	ترکیب یکنواخت	۲۶
شکل (۶-۳)	جهش	۲۶
شکل (۷-۳)	بیان درختی رابطه (۱-۳)	۲۷
شکل (۸-۳)	ترکیب خطی در برنامه ریزی ژنتیک	۲۹
شکل (۹-۳)	جهش درختی در برنامه ریزی ژنتیک	۳۰
شکل (۱۰-۳)	عملگر جایگشت در برنامه ریزی ژنتیک	۳۰
شکل (۱۱-۳)	یک واقعه فرضی از جهش نقطه ای در برنامه ریزی ژنتیک و ساختار نامعتبر درخت ایجاد شده	۳۱
شکل ۱۲-۳	چرخ گردان	۳۴
شکل (۱۳-۳)	کدگذاری به شکل رشته خطی و کد برداری به صورت ET یک کروموزوم با دو ژن در GEP	۳۶
شکل (۱۴-۳)	بیان درختی رابطه $F(x)=\sin(x)+\sqrt{x^2+y}$	۳۷
شکل (۱۵-۳)	بیان درختی رابطه (۳-۳)	۳۸
شکل (۱۶-۳)	بیان درختی رابطه (۵-۳)	۳۹
شکل (۱۷-۳)	بیان درختی رابطه (۶-۳)	۴۰
شکل (۱۸-۳)	بیان درختی رابطه (۷-۳)	۴۱
شکل (۱۹-۳)	کدگذاری سه ژن، b پیوند سه ژن	۴۵
شکل (۲۰-۳)	الگوریتم برنامه ریزی بیان ژن (فریرا، ۲۰۰۱)	۴۸
شکل (۱-۴)	نمودار ویلکاکس رودخانه تجن	۵۳
شکل (۲-۴)	نمودار ویلکاکس رودخانه بابلرود	۵۴
شکل (۳-۴)	نمودار اجرای الگو ۵، برآورد TDS رودخانه تجن در مرحله آموزش	۵۹
شکل (۴-۴)	نمودار اجرای الگو ۵، برآورد TDS رودخانه تجن در مرحله آزمون	۶۰
شکل (۵-۴)	نمودار درختی الگو ۵ برآورد TDS رودخانه تجن با استفاده از برنامه ریزی بیان ژن	۶۱
شکل (۶-۴)	نمودار درختی الگو ۱ برآورد TDS رودخانه تجن با استفاده از برنامه ریزی بیان ژن	۶۲
شکل (۷-۴)	نمودار مقایسه برآورد TDS به سه روش	۶۴
شکل (۸-۴)	نمودار اجرای الگو ۲ برآورد SAR رودخانه تجن در مرحله آموزش	۶۷

- شکل (۹-۴) نمودار اجرای الگو ۲ برآورد SAR رودخانه تجن در مرحله آزمون ۶۷
- شکل (۱۰-۴) نمودار درختی برآورد SAR رودخانه تجن با استفاده از برنامه ریزی بیان ژن ۶۸
- شکل (۱۱-۴) نمودار اجرای الگو ۸ برآورد TDS رودخانه بابلرود در مرحله آموزش ۷۴
- شکل (۱۲-۴) نمودار اجرای الگو ۸ برآورد TDS رودخانه بابلرود در مرحله آزمون ۷۵
- شکل (۱۳-۴) نمودار درختی برآورد TDS رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن ۷۶
- شکل (۱۴-۴) نمودار درختی الگوی ۱ برآورد TDS رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن ۷۸
- شکل (۱۵-۴) نمودار مقایسه برآورد TDS به سه روش ۸۰
- شکل (۱۶-۴) نمودار اجرای الگو ۵ برآورد SAR رودخانه بابلرود در مرحله آموزش ۸۳
- شکل (۱۷-۴) نمودار اجرای الگو ۵ برآورد SAR رودخانه بابلرود در مرحله آزمون ۸۳
- شکل (۱۸-۴) نمودار درختی برآورد SAR رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن ۸۴
- شکل (۱۹-۴) نمودار درختی الگوی ۱ برآورد SAR رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن ۸۶

فهرست جداول

جدول (۱-۳)	پارامتر های آماری اندازه‌گیری های کیفی بر روی رودخانه تجن	۲۰
جدول (۲-۳)	پارامتر های آماری اندازه‌گیری های کیفی بر روی رودخانه بابلرود	۲۰
جدول (۱-۴)	ضریب همبستگی متغیرهای مورد مطالعه رودخانه تجن	۵۶
جدول (۲-۴)	الگو های ورودی برآورد TDS اسخراج شده از SPSS23	۵۷
جدول (۳-۴)	مقادیر پارامترهای GEP در برآورد TDS و SAR رودخانه های تجن و بابلرود	۵۸
جدول (۴-۴)	نتایج مدل برنامه‌ریزی بیان ژن برای الگوهای ورودی برآورد TDS رودخانه تجن	۵۸
جدول (۵-۴)	مقادیر ثابت در بیان درختی الگوی ۵، برآورد TDS رودخانه تجن	۶۱
جدول (۶-۴)	مقادیر ثابت در بیان درختی الگوی ۱، برآورد TDS رودخانه تجن	۶۳
جدول (۷-۴)	مقایسه شاخص های آماری سه روش برآورد TDS	۶۴
جدول (۸-۴)	الگو های ورودی برآورد SAR اسخراج شده از SPSS23	۶۵
جدول (۹-۴)	نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی، برآورد SAR رودخانه تجن	۶۶
جدول (۱۰-۴)	مقادیر ثابت در بیان درختی الگوی ۲ برآورد SAR رودخانه تجن	۶۹
جدول (۱۱-۴)	ضریب همبستگی متغیرهای مورد مطالعه رودخانه بابلرود	۷۱
جدول (۱۲-۴)	الگو های ورودی برآورد TDS بابلرود اسخراج شده از SPSS	۷۲
جدول (۱۳-۴)	نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی برآورد TDS رودخانه بابلرود	۷۳
جدول (۱۴-۴)	مقادیر ثابت در بیان درختی الگوی ۸ برآورد TDS رودخانه بابلرود	۷۶
جدول (۱۵-۴)	مقادیر ثابت در بیان درختی الگوی ۸ برآورد TDS رودخانه بابلرود	۷۸
جدول (۱۶-۴)	مقایسه شاخص های آماری سه روش برآورد TDS	۷۹
جدول (۱۷-۴)	الگو های ورودی برآورد SAR رودخانه بابلرود اسخراج شده از SPSS	۸۱
جدول (۱۸-۴)	نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی برآورد SAR رودخانه بابلرود	۸۲
جدول (۱۹-۴)	مقادیر ثابت در بیان درختی الگوی ۵ برآورد SAR رودخانه بابلرود	۸۵
جدول (۲۰-۴)	مقادیر ثابت در بیان درختی الگوی ۱ برآورد SAR رودخانه بابلرود	۸۶

فصل اول

کلیات

۱-۱- تعریف مسئله

توسعه و حیات تمدن های بشری به آب وابسته است. حجم آب در کره زمین بسیار زیاد می باشد، اما حجم آبی که قابلیت استفاده در بخش های شرب، صنعت، کشاورزی را دارا می باشد؛ در مقایسه با آن بسیار ناچیز می باشد. از همین حجم اندک آبی که در خشکی ها و آبهای سطحی و زیرزمینی قابل دستیابی می باشند، نیز به دلیل شرایط کیفی دائما کاسته می شود. به عبارت دیگر اگرچه به لحاظ کمی، حجم آب در چرخه هیدرولوژی زمین تقریبا ثابت است ولی به دلیل تنزل کیفیت شیمیایی، حجم آب قابل استفاده در حال کم شدن می باشد به همین دلیل پیش بینی و پایش کیفی منابع آب و ارزیابی و تحلیل روند تغییرات آن از اهمیت بسزایی برخوردار می باشد. تاکنون روابط گوناگون و پیچیده ای به منظور پیش بینی مولفه های کیفی منابع آب ارائه شده است، لیکن تعداد زیادی از پارامترهای موثر بر کیفیت آب دارای رابطه ی پیچیده ی غیرخطی با یکدیگر می باشند که روشهای سنتی عملکرد مناسبی برای حل این چنین مسائلی نداشته اند (حسامی افشار و همکاران، ۱۳۹۲).

آب یک منبع حیاتی است که سلامتی و زنده ماندن اکوسیستم و بشر وابسته به آن می باشد (نجاه و همکاران، ۲۰۱۳). در نتیجه مطالعه کیفیت منابع آب از مهم ترین اولویت های حفاظت محیط زیست و از عامل های مهم در دستیابی به توسعه پایدار در بیش تر جوامع است (آذرمدل و همکاران، ۱۳۹۳). در حال حاضر کمیت و کیفیت منابع آبی در جهان یک موضوعی نگران کننده برای جوامع مختلف می باشد. امروزه به خاطر کاهش مقدار آب موجود در دنیا، رو به آب های با کیفیت پایین آوردند. بنابراین ارزیابی تغییرات درازمدت کیفیت آب یک چالش اساسی می باشد (دیامونتوپولا و همکاران، ۲۰۰۵). با این حال مطالعات کیفیت آب به علت کاهش چشمگیر منابع آبی در سال های اخیر بسیار با اهمیت تر گردیده است. از طرفی اندازه گیری پارامترهای کیفیت آب بسیار پر هزینه بوده و اگر بتوان بعضی پارامترها را از روی پارامترهای دیگر اندازه گیری شده برآورد کرد، یک مسئله مهم می باشد (کاظم زاده و همکاران، ۱۳۹۳).

اکوسیستم های آبی از جمله رودخانه ها به عنوان یکی از اجزای سازنده این کره خاکی به شمار آمده که نه تنها زیستگاه های مهمی را برای موجودات آبی تشکیل می دهند، بلکه منبع مهمی از غذا را برای مصارف انسان شامل می شوند. هم اکنون بر همگان ثابت شده است که کیفیت آب علاوه بر کمیت آن دارای اهمیت فوق العاده می باشد. واژه کیفیت آب در توضیح خصوصیات فیزیکی، شیمیایی و بیولوژیکی برای استفاده برای هدف خاص مورد استفاده قرار می گیرد (Khalil et al., 2011).

۱-۲- روش انجام تحقیق

در زمینه مدیریت بهتر برای حفظ کیفیت آب، مدل های بسیار زیادی گسترش یافته است. بیشتر این مدل ها نیازمند پارامتر های ورودی بسیاری هستند که یا دسترسی به آنها مشکل است و یا اینکه اندازه گیری آنها محتاج صرف هزینه و زمان زیادی می باشد (میثاقی و محمدی، ۱۳۸۳). در سال های اخیر استفاده از تکنیک های هوش مصنوعی شامل شبکه عصبی مصنوعی، نرو فازی، برنامه ریزی بیان ژن و ماشین پشتیبان بردار برای پیش بینی و برآورد پارامترهای مختلف زیاد شده است (ستاری و همکاران، ۲۰۱۳).

مطالعات زیادی در مورد پیش بینی کیفیت آب انجام شده است که برای مثال می توان به مطالعات دیامونتوپولا و همکاران جهت پیش بینی پارامتر های کیفیت آب در رودخانه اکسیوس (شمال یونان) با استفاده از تکنیک شبکه عصبی مصنوعی اشاره نمود. آنها بیان نمودند که شبکه عصبی مصنوعی برای پیش بینی پارامترهای کیفیت آب بسیار مناسب بوده و همچنین یک ابزار کارا برای برآورد مقادیر اندازه گیری نشده یا مفقوده می باشد.

زمان زاد قویدل و منتصری (۲۰۱۴) کاربرد روش های مختلف داده مبنا را در پیش بینی کل جامدات محلول در زاینده رود ارزیابی کردند. آنها به این نتیجه رسیدند که مدل های ANFIS-SC, GEP ANN, ANFIS-GP می توانند به طور موثر در پیش بینی تغییرات کل جامدات محلول استفاده شوند. هم چنین در بین روش های مختلف هوش مصنوعی روش GEP نسبت به مدل های دیگر در پیش

بینی این پارامتر برتری داشته است. در تحقیق حاضر از روش برنامه ریزی بیان ژن به این منظور استفاده شده است. برنامه ریزی بیان ژن (GEP) روشی جدید در فناوری هوش مصنوعی است که در سال ۱۹۹۹ توسط فریرا ابداع گردید. این روش توسعه یافته الگوریتم ژنتیک (GA) و برنامه ریزی ژنتیک (GP) است (فریرا، ۲۰۰۱). برای اجرای مدل از داده های هیدرومتری رودخانه های بابلرود و تجن که شامل هدایت الکتریکی آب، اسیدیته، سدیم، کلسیم، پتاسیم، نسبت جذب سدیم، منیزیم، کلر، سولفات، بیکربنات و دی و کل جامدات محلول بود استفاده شد.

۱-۳-اهداف و نوآوری

در پژوهش حاضر از روش برنامه ریزی بیان ژن (GEP) به عنوان یک الگوریتم تطبیقی جدید، برای برآورد پارامترهای کیفی آب رودخانه های تجن و بابلرود استفاده شده است. با توجه به آنچه که قبلاً اشاره شد، حجم آب قابل استفاده در حال کم شدن می باشد به همین دلیل پیش بینی و پایش کیفی منابع آب و ارزیابی و تحلیل روند تغییرات آن از اهمیت بسزایی برخوردار می باشد. از طرفی اندازه گیری پارامترهای کیفیت آب بسیار پرهزینه بوده و اگر بتوان بعضی پارامترها را از روی پارامترهای دیگر اندازه گیری شده برآورد کرد، یک مسئله مهم می باشد. از این رو، در صورتیکه نتایج حاصل از برآورد پارامترهای کیفی آب به کمک برنامه ریزی بیان ژن موفقیت آمیز باشد، می توان از این روش به عنوان روشی پیشنهادی در برآورد پارامترهای کیفی آب رودخانه ها نام برد.

فصل دوم

سابقه و پیشینه تحقیق

۲-۱- مقدمه

امروزه با توجه به رشد جمعیت، افزایش روز افزون فعالیت های کشاورزی و صنعتی و افزایش قابل توجه حجم فاضلاب های شهری آلودگی منابع آب به ویژه آبهای سطحی مانند رودخانه ها بیش از پیش مورد توجه قرار گرفته به طوریکه در بسیاری از مناطق با خطر جدی روبرو است (بانژاد و همکاران، ۱۳۹۲).

مواد محلول در آب ممکن است از نظر ماهیت آلی یا معدنی باشند. مواد غیر آلی (معدنی) حل شده در آب شامل مواد معدنی، فلزات و گازها می باشند. بعضی از مواد آلی به صورت ذرات کلوییدی هستند اما بیشتر مواد آلی به صورت محلول هستند. آلاینده های آلی ممکن است باعث بو، رنگ و طعم نامطبوع آب شوند. مواد حاصل از تجزیه گیاهان، مواد شیمیایی آلی و گازهای آلی، اجزای آلی محلول در آب را تشکیل می دهند. بسیاری از مواد حل شده در آب نامطلوب هستند. برخی از ترکیبات شیمیایی ممکن است سمی باشند و برخی از اجزای آلی محلول به اثبات رسیده است که سرطانزا هستند. البته باید توجه داشت که تمامی مواد محلول در آب نامطلوب نیستند. آب تقطیر شده که کاملاً خالص باشد، بی مزه است.

یکی از معیارهای بررسی کیفی آب مقدار کل جامدات محلول در آب است. منظور از TDS کل جامدات محلول است که می توان به سادگی با تبخیر حجم معینی از آب و اندازه گیری وزن املاح باقیمانده تعیین کرد. واحد سنجش TDS بر حسب میلی گرم در لیتر (mg/lit) و یا قسمت در میلیون (ppm) بیان می شود و عامل مهمی در کیفیت آب بوده و اثر زیادی در جابجایی و تبدیل شیمیایی و یونیزه شدن مواد دارد (اگدرنژاد و همکاران، ۱۳۹۳).

از جمله مهمترین پارامتر های کیفی آب شاخص نسبت جذب سدیم است که برای تعیین مطلوب بودن آب آبیاری استفاده می گردد. برای تعیین میزان سدیم قابل تبادل از فاکتور SAR استفاده می شود. املاح باعث افزایش فشار اسمزی در فضاهای بین سلولی و درون سلولی و محیط ریشه

می‌شود که در این میان یون سدیم و کلر نقش بیشتری دارد. اندازه‌گیری سدیم در آب بر حسب میلی‌گرم و گرم در لیتر، میلی‌اکی‌والان در لیتر و یا ppm انجام می‌شود. اگر چه اندازه‌گیری سدیم بر حسب گرم، خود معرف میزان سدیم است. از آنجا که میزان یون سدیم در رقابت با املاح دیگر خصوصاً کلسیم و منیزیم قرار می‌گیرد به لحاظ اختلاف وزن مولکولی بین عناصر، واحد اکی‌والان در لیتر معرف مناسب‌تری است که امکان مقایسه املاح مختلف را از نظر ترکیبات شیمیایی میسر می‌نماید. (یوسفی و همکاران، ۱۳۹۴). به طور کلی مقادیر بالای SAR سبب نامناسب شدن آب آبیاری می‌گردد، آبیاری بوسیله آب با SAR بالا مستلزم اصلاح خاک برای جلوگیری از آسیب‌های دراز مدت به خاک است (Michael et al., 2008).

۲-۲- پیشینه مطالعات برآورد پارامترهای کیفی آب

محور توسعه پایدار، انسان سالم است و سلامت انسان در گرو بهره‌مندی از آب آشامیدنی مطلوب می‌باشد بدون تامین آب سالم جایی برای سلامت مثبت و رفاه جامعه، وجود ندارد. لذا آگاهی از روند تغییرات، شبیه‌سازی و پیش‌بینی کیفیت آب با توجه به اهمیت آن حائز اهمیت می‌باشد (فریدونی و امیراحمدی، ۱۳۹۳).

معمولاً یک رابطه‌ی خطی بین EC و TDS وجود دارد و از این جهت اندازه‌گیری قابلیت هدایت الکتریکی، شاخص مناسبی برای تعیین کل مواد حل شده در آب است (جلیلیان و امینی، ۱۳۸۶). بطور کلی طبق پژوهشهای انجام شده رابطه تجربی (۱-۲) بین EC و TDS وجود دارد (تیرامالین و همکاران ۲۰۰۹).

$$TDS = (0.55 \text{ to } 0.7)EC \quad (1-2)$$

تیرامالین و همکاران (۲۰۰۹) در مطالعه خود در منطقه ساحلی مالزی به بررسی ویژگیهای فیزیکی شیمیایی آب- های زیرزمینی این منطقه برای مصرف آبیاری پرداختند. در این پژوهش همبستگی ۰/۶۳ بین هدایت الکتریکی و کل مواد محلول آب پیدا شد. بتوشا و همکاران در مطالعه خود بر روی کیفیت آبهای زیرزمینی منطقه غربی کویماتور هند به وسعت ۱۰۰ کیلومتر مربع بر اساس رابطه‌ی (۲-۲) با ضریب تبیین ۰,۹۵ را برای منطقه بدست آوردند.

$$TDS(mg/l) = 353 + 0.506 * EC(\mu s/cm) \quad (2-2)$$

مفتاح وصادقی (۱۳۸۸) به منظور بررسی تاثیر سد گلستان بر کیفیت آب گرگانرود، چهار پارامتر TDS,EC,SAR و سختی را در نظر گرفتند و به این نتیجه رسیدند که به دلیل برداشت آب از دریاچه های تحتانی سد، میزان املاح در ایستگاه خروجی سد افزایش یافته و مجددا کاهش می‌یابد.

امید کرمی و عبدالرحیم هوشمند (۱۳۹۲)، در تحقیقی به منظور صحت سنجی و برآورد ارتباط بین EC و TDS آب رودخانه کارون در فصول پرآب و فصول کم آب با استفاده از نرم افزار SPSS 18 نشان دادند که رابطه $TDS = 0.619(EC)$ با ضریب تعیین ۰,۸۵۸ و رابطه $TDS = 0.606(EC)$ با ضریب تعیین ۰,۸۶ برای فصول کم آب و نیز این دو رابطه توسط آزمون آنالیز واریانس رگرسیونی تأیید شده و باقیمانده های حاصل از هریک، توسط آزمون t-جفتی با مقدار صفر مقایسه شدند.

پارسافر و زارع ایبانه (۱۳۸۸)، در دشت همدان-بهار مطالعاتی انجام دادند که نشان می‌دهد در چاه های آب این دشت همبستگی خوبی بین هدایت الکتریکی و مجموع کاتیون ها و آنیون ها وجود دارد. از بین این کاتیون ها، منیزیم و از بین آنیون ها کلر بیشترین اثر را در میزان هدایت الکتریکی چاه های آب دشت همدان-بهار دارد.

دودکانلوی میلان و همکاران (۱۳۹۰)، در تحقیقی نشان دادند که رابطه TDS با EC در دشت جمال در شهرپور بصورت $TDS(mg/l) = 0.581 \times EC(\mu S/cm)$ و در فاصله زمانی پاییز تا مرداد رابطه به صورت $TDS(mg/l) = 0.604 \times EC(\mu S/cm)$ است. درآبخوان شمالی رابطه EC با TDS در مرداد ماه

بصورت مدل خطی همراه با عرض از مبدأ به صورت $(TDS(mg/L) = EC0.396(\mu S/cm) + 140)$ درحالی که در اوقات دیگر این رابطه با مدل خطی بدون عرض از مبدأ قابل توجیه است.

۲-۳- پیشینه مطالعات رودخانه های بابلرود و تجن

رودخانه تجن یکی از پرآبترین رودخانه های استان مازندران است که از ارتفاعات جنوب شهرستان ساری سرچشمه می گیرد. رضویان در سال ۱۳۸۳ طی تحقیقی با بررسی اثرات کارخانه چوب و کاغذ مازندران بر رودخانه تجن افزایش بار آلودگی رودخانه تجن تأکید داشته است. وی با نتایج خود نشان داد که در سال های ۸۲ و ۸۳، تمام قسمت های رودخانه برای مصارف شرب، صنایع غذایی و سایر استفاده های بهداشتی نامناسب بوده است. بدون تردید، رشد سریع جمعیت و شهرنشینی به همراه توسعه بی رویه روستایی مبتنی بر انجام فعالیت های کشاورزی با بهره وری پائین، اکوسیستم های آسیب پذیر نوار باریک سواحل جنوبی دریاچه خزر را طی دهه های اخیر، مورد تهدید جدی قرار داده است. این درحالی است که فعالیت های معدنی در سرشاخه رودخانه ها، توسعه شبکه خطوط مواصلاتی، احداث سازه های آبخیزداری و نیز توسعه صنعتی و فعالیت های اقتصادی مزید بر علت شده و در نهایت، تمرکز بسیاری از انواع آلاینده های محیط زیستی را در پهنه های آبی و خاکی حوزه آبریز رودخانه تجن به دنبال داشته است (لوپزه، ۱۳۸۴).

ورسه و همکاران طی تحقیقی در سال ۱۳۹۳ نتیجه گرفتند که به طور کلی، کیفیت فیزیکی و شیمیایی آب رودخانه تجن در تمامی مسیر رودخانه و در اکثر ایستگاه های هیدرومتری در شاخه های اصلی و فرعی آن دارای شرایط نسبتاً عادی (غیر از زمان وقوع سیل و یا در زمان هائی که رژیم آبدهی رودخانه در وضعیت کم آبی فصلی قرار دارد) از جهت مصارف کشاورزی و صنعت (به جز صنایع غذایی) بوده است.

آزموده و همکاران (۱۳۹۲) در تحقیقی با عنوان ارزیابی کارایی شبکه عصبی مصنوعی در پیش بینی شاخص کیفی کل مواد محلول آب رودخانه تجن ساری نشان دادند که کارایی مدل های شبکه عصبی

به مراتب دقیقتر از رگرسیون خطی می‌باشد. دلیل این امر نیز سیستم پردازش گسترده و موازی و در نظر گرفتن تبدیل های غیرخطی در شبکه عصبی است.

رودخانه بابلرود در شمال ایران و در جنوب دریای مازندران قرار دارد و بخشی از حوزه آبریز دریای مازندران را تشکیل می‌دهد. خسروی و همکاران (۱۳۹۱) در تحقیقی با عنوان بررسی روند تغییرات زمانی پارامترهای کیفیت آب رودخانه بابلرود نشان دادند که کل پارامترهای کیفیت آب با دبی رابطه عکس داشته به جز اسیدیته که با دبی فاقد رابطه معنی داری می‌باشد. بنابراین مقدار این پارامترها در تابستان که دبی افت می‌کند به مقدار زیادی در آب رودخانه زیاد می‌شود.

بختیاری و همکاران در تحقیقی تحت عنوان بررسی دو الگوریتم یادگیری شبکه عصبی مصنوعی در تخمین بار معلق رسوب رودخانه بابلرود از دو الگوریتم on lin back propagation و MemoryQuasi-Newton limited برای آموزش داده ها استفاده کردند که نتایج بدست آمده از الگوریتم on line back propagation دقیق تر از نتایج بدست آمده از الگوریتم دیگری بود.

جروقی و همکاران (۱۳۸۸) در تحقیقی تحت عنوان شبیه سازی کیفی رودخانه بابلرود با استفاده از نرم افزار MIKE11 به این نتیجه دست یافتند که این رودخانه با این کیفیت و دبی حاضر برای آبیاری و کشاورزی مناسب می‌باشد به شرطی که آلاینده های اضافی در رودخانه تخلیه نشود و یا در صورت تخلیه فاضلابها تصفیه شوند.

۲-۴- پیشینه مطالعات روش های هوشمند

با توجه به زمان بر بودن و هزینه بر بودن اندازه گیری پارامترهای کیفی، امروزه از مدل های هوشمند مانند شبکه عصبی مصنوعی ، الگوریتم ژنتیک و ... به طور وسیعی در مطالعات مربوط به تخمین پارامترهای مختلف منابع آب استفاده می شود و در پژوهش های مختلف دقت بالای این روش نسبت به روش های تجربی و رگرسیونی مشاهده شده است (Kuo et al., 2004., Kuo et al., 2007).

موسوی جهرمی و گلایی (۲۰۰۸) در پژوهشی کارایی شبکه عصبی مصنوعی برای شبیه سازی پارامترهای کیفی SAR ، TDS و EC برای برخی ایستگاه های رودخانه کارون با دقت بیش از ۹۰ درصد بررسی کردند. در پژوهشی دیگر کارایی مدل های شبکه عصبی و Recurrent Neural Network را برای تخمین شاخص کیفی کل جامدات محلول در رودخانه تلخه در استان آذربایجان شرقی مورد بررسی قرار گرفت (Asadollahfardi et al., 2012).

در زمینه مدیریت حفظ کیفیت آب مدل های بسیاری گسترش یافته است. بیشتر این مدل ها نیازمند پارامترهای ورودی متعددی هستند که دسترسی به آن ها مشکل و نیازمند صرف وقت و هزینه بسیار بوده و ارزیابی دقیق و گسترش آن کار دشواری است. بسیاری از مدل های پایه آماری کیفیت آب بر روابط خطی بین متغیرها استوار بوده که به خوبی رفتار پیچیده و واقعی را پیش بینی نمی کنند؛ بنابراین با توجه به مشکلات موجود، امروزه سیستم های هوشمند به طور گسترده برای پیش بینی پدیده های غیرخطی مورد استفاده قرار می گیرد، که برنامه ریزی بیان ژن به عنوان یک تکنیک برنامه ریزی خودکار است که راه حل مسئله را با استفاده از برنامه ریزی کامپیوتر ارائه کرده و عضوی از خانواده الگوریتم تکاملی می باشد (صالحی، ۱۳۹۰). طی سالهای اخیر استفاده از روش برنامه ریزی بیان ژن در مطالعات کیفیت و کمیت آب رودخانه مورد توجه محققین قرار گرفته است، که از جمله می توان به موارد ذیل اشاره نمود:

قربانی و صالحی در سال ۱۳۹۰ تغییرات سطح آب زیرزمینی و تاثیر آن بر برخی ویژگیهای کیفی آب زیرزمینی نظیر کل املاح محلول، غلظت یون منیزیم، غلظت یون کلسیم و شوری را با استفاده از برنامه ریزی بیان ژن در یک دوره ۳۱ ساله بررسی کردند نتایج حاکی از همبستگی قابل قبول نوسانات سطح آب زیرزمینی با کل املاح محلول و غلظت یون کلسیم بوده در حالیکه برای شوری و غلظت یون منیزیم معنی دار نمی باشد.

کرمی و گلابی (۱۳۹۱) در پژوهشی به منظور شبیه سازی و پیش بینی پارامترهای کیفی آب با استفاده از شبکه عصبی مصنوعی نشان دادند که در شبیه سازی TDS بیشترین تاثیر به ترتیب مربوط به SO_4 ، Ca، K، Na و Cl و کمترین ناشی از Q و CO_3 بوده است. در شبیه سازی EC بیشترین تاثیر به ترتیب مربوط به Ca، Cl و SO_4 کمترین ناشی از HCO_3 ، K، Q، Mg و CO_3 بوده است. در شبیه سازی SAR بیشترین تاثیر به ترتیب مربوط به Ca، Na و Mg و کمترین ناشی از K و CO_3 ، Q بوده که با توجه به رابطه ریاضی محاسبه SAR نتیجه به دست آمده قابل قبول می باشد. یکسان بودن برخی از یون های تاثیر گذار بر EC و TDS نیز می تواند تایید کننده رابطه مستقیم و خطی بین این دو پارامتر باشد.

حسامی افشار و همکاران (۱۳۹۲) در پژوهشی با عنوان پیش بینی غلظت TDS رودخانه ی زرينه رود بر مبنای کمترین تعداد پارامتر ورودی به کمک برنامه ریزی ژنتیک دریافتند که تاکنون روابط گوناگون و پیچیده ای به منظور پیش بینی مولفه های کیفی منابع آب ارائه شده است، لیکن تعداد زیادی از پارامترهای موثر بر کیفیت آب دارای رابطه پیچیده ی غیرخطی با یکدیگر می باشند که روشهای سنتی عملکرد مناسبی برای حل این چنین مسائلی نداشته اند. هدف از این پژوهش پیش بینی مقدار غلظت املاح محلول رودخانه ی زرينه رود (منتهی به دریاچه ی ارومیه) به کمک برنامه ریزی ژنتیک و با در نظر گرفتن کمترین مقدار ورودی می باشد. برای تحقق این امر از روش رگرسیون گام به گام استفاده گردید. نتایج مربوط به رگرسیون گام به گام نشان داد که در بین ورودی های در نظر گرفته شده، پارامترهای Na^+ ، Ca^{2+} و Mg^{2+} دارای بیشترین سطح معنی داری نسبت به دیگر پارامترها در پیش بینی مقدار TDS می باشند. همچنین نتایج مربوط به برنامه ریزی ژنتیک نیز حاکی از آن می باشد که می توان در بررسی ها و تخمین پارامتر غلظت املاح محلول از آن به عنوان یک ابزار مناسب استفاده نمود.

فرودفام و همکاران (۱۳۸۸) به منظور پیش بینی جریان روزانه رودخانه ليقوان از مدل برنامه ریزی ژنتیک استفاده کردند. نتایج این تحقیق نشان داد روش برنامه ریزی بیان ژن (GEP) از دقت بسیار بالایی نسبت به روش شبکه های عصبی مصنوعی و مدل های سری زمانی برخوردار است.

سلطانی و همکاران (۱۳۸۹) با کاربرد برنامه ریزی ژنتیک در مدلسازی فرآیند بارش رواناب روزانه حوضه آبریز ليقوان، با تعریف دو مجموعه عملگر ریاضی و با محاسبه ضریب همبستگی و میانگین مربعات خطا به این نتیجه دست یافتند، مدل حاصل از مجموعه عملگر های اصلی، مناسب تر است. دربندی و همکاران (۱۳۹۰) جریان ورودی به مخزن سد علویان را با استفاده از برنامه ریزی بیان ژن مورد بررسی قرار داده و نتایج حاصله بیانگر کارایی مناسب و دقت بالای برنامه ریزی ژنتیک در پیش بینی جریان ورودی به سد می باشد.

اگدرنژاد و همکاران (۱۳۹۳) در تحقیقی تحت عنوان مقایسه کارایی مدل های شبکه عصبی مصنوعی و رگرسیون برای پیش بینی مقدار کل جامدات محلول (TDS) رودخانه کارون شبکه عصبی مصنوعی را ابزاری مناسب برای تعیین پارامترهای کیفی آب رودخانه کارون دانستند و این مدل را با دقت قابل قبول و سرعت بالا قادر به مشخص کردن پارامتر TDS دانسته و می تواند جایگزین مناسبی برای روش های آماری و سبب صرفه جویی اقتصادی گردد. و همچنین با توجه به ارتباط بین مقادیر مشاهده شده و مقادیر پیش بینی شده، از بین روابط رگرسیونی خطی، نمایی، توانی و لگاریتمی بیشترین ضریب همبستگی و کمترین مقدار RMSE مربوط به تابع خطی بود.

کاظم زاده و همکاران (۱۳۹۳) در تحقیقی مدل های m5 و MLP را در برآورد میزان کل جامدات محلول حوزه آبخیز ليقوان مورد ارزیابی قرار دادند. به این منظور داده های انتخاب شده از دو ایستگاه آبسنجی به صورت ماهانه طی دوره ۱۳۵۰-۱۳۸۹ بوده است. برای ارزیابی مدل ها، داده ها به دو قسمت شامل آموزش ۷۰ درصد داده ها و آزمون ۳۰ درصد داده ها تقسیم شدند. پارامترهای ورودی در شش سناریو مختلف با توجه به درجه همبستگی با خروجی ایجاد شدند. نتایج همبستگی نشان

داد که پارامتر HCO_3 بیشترین همبستگی با مقادیر TDS در دو ایستگاه لیقوان و هروی داشته است در صورتیکه کمترین همبستگی را در بین ورودی‌های مدل‌ها SO_4^{2-} داشته است. با توجه به معیارهای ارزیابی مدل‌ها MLP, M5 توانایی بالایی در پیش‌بینی کل جامدات محلول داشته و همچنین در بیشتر موارد نتایجی یکسان را در این دو روش نشان دادند. به عبارتی نتایج این پژوهش نشان داد که تقریباً کارایی مدل‌های MLP, M5 در پیش‌بینی مقادیر کل جامدات محلول یکسان می‌باشد.

سارانی و همکاران (۱۳۹۱) در پژوهشی به ارزیابی و مقایسه عملکرد شبکه‌های عصبی مصنوعی MLP و GFF در شبیه‌سازی شاخص کیفی SAR رودخانه سیستان پرداختند به منظور شبیه‌سازی آب رودخانه سیستان در طی سالهای ۱۳۷۵ تا ۱۳۸۹ در ایستگاه پایاب سد کهک با استفاده از پارامترهای کلر، شوری، قلیائیت و کل مواد محلول با استفاده از شبکه عصبی مصنوعی با الگوریتم‌های MLP و GFF صورت پذیرفت. با ارزیابی تمامی مدل‌ها نشان دادند که، مقادیر پیش‌بینی شده SAR توسط شبکه MLP با یک لایه پنهان با مقادیر واقعی SAR همبستگی بیشتری دارد و با بررسی و آنالیز حساسیت بر روی بهترین شبکه، فاکتور کیفی شوری (EC) بیشترین تاثیر و فاکتور pH کمترین تاثیر را در نتایج پیش‌بینی شبکه داشته است.

فصل سوم

مواد و روش انجام پژوهش

۳-۱- مقدمه

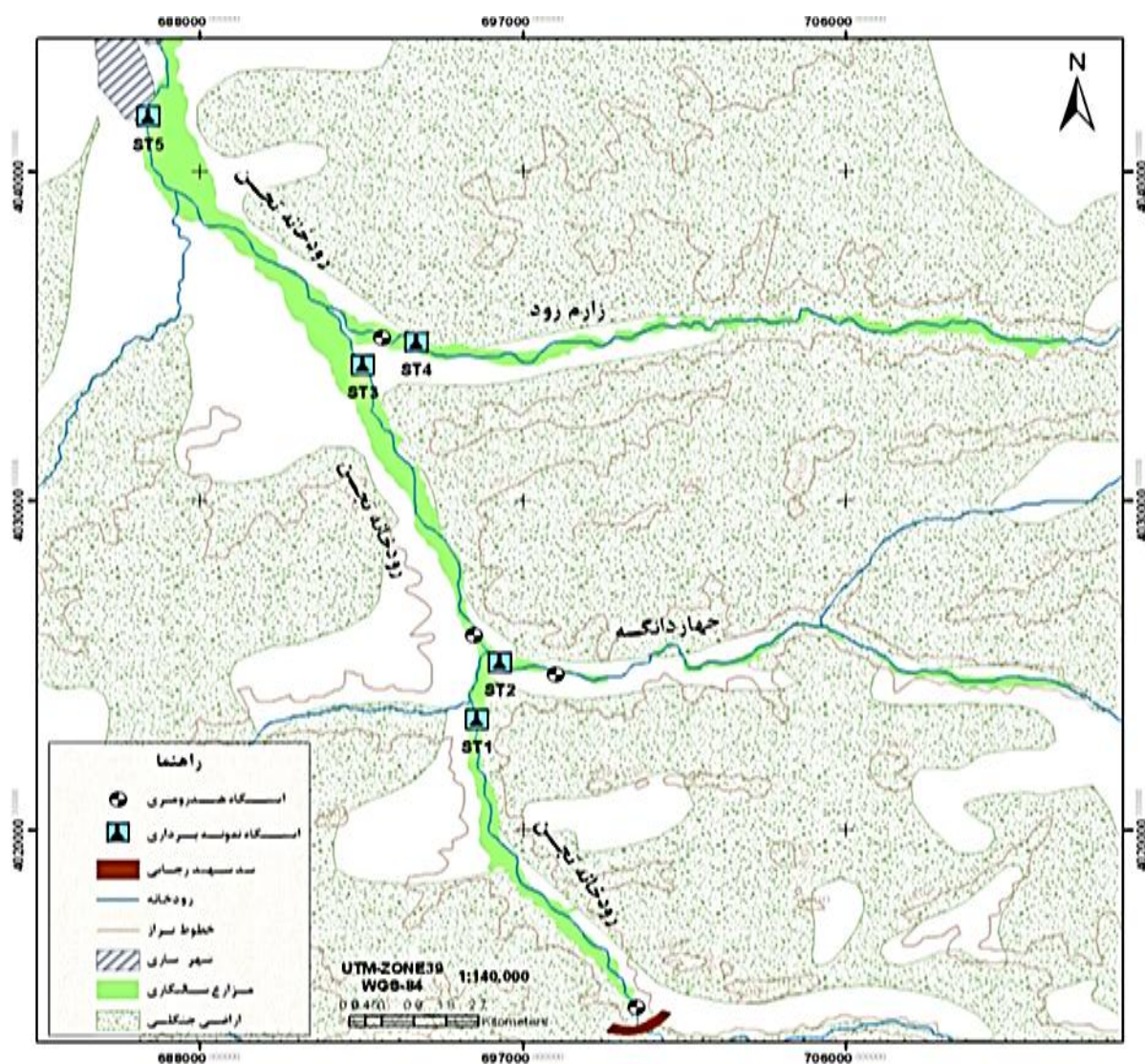
امروزه مدل های بسیاری برای بررسی و پیش بینی کیفیت آب مورد استفاده قرار می گیرند که بیشتر آن ها نیازمند اطلاعات ورودی فراوان و غیرقابل دسترس هستند و یا اندازه گیری این اطلاعات صرف هزینه های زمانی و مکانی فراوانی را به دنبال خواهد داشت. طی سالهای اخیر، محققان از روش های هوشمند در برآورد پارامتر های کیفی آب استفاده کرده اند که در مقایسه با روش های مرسوم رگرسیونی از دقت بالاتری برخوردار است. در این فصل ابتدا شرح مختصری از رودخانه های مورد مطالعه ارائه می گردد، سپس مفاهیم زیست شناسی و پارامترهای مورد استفاده در استخراج مدل ارائه شده و در ادامه روش برنامه ریزی بیان ژن معرفی شده است.

۳-۲- مناطق مورد مطالعه

۳-۲-۱- رودخانه تجن

رودخانه تجن یکی از پرآبترین رودخانه های استان مازندران است که از ارتفاعات جنوب شهرستان ساری سرچشمه می گیرد. رقوم سرشاخه های تجن ۲۵۰۰ متر از سطح دریا ارتفاع دارند. حوضه آبریز تجن در حدود ۴۱۴۷/۲۲ کیلومتر مربع وسعت دارد و شیب متوسط آن ۰/۸۵ درصد است. این حوضه آبریز در موقعیت جغرافیایی ۵۳ درجه و ۷ دقیقه الی ۵۳ درجه و ۴۲ دقیقه طول شرقی و ۵۳ درجه و ۵۶ دقیقه و ۳۶ درجه و ۱۷ دقیقه عرض شمالی قرار دارد. جریان آب رودخانه تجن در محدوده کوهستانی به صورت دره_ رودخانه است و شیب بستر در بخش کوهستانی ۱۰-۴ درجه می باشد. رودخانه تجن از میان شهر ساری گذر مینماید. این رودخانه در محل شهرستان ساری وارد پهنه ساحلی خزر شده و در ناحیه فرح آباد به دریای خزر میریزد. طول رودخانه اصلی ۱۵۱/۵۷ کیلومتر بوده و شیب بستر آن از ۵ تا ۱۷ در هزار تغییر میکند (مرادی و زمانی نوری ۱۳۹۲). زیست بوم رودخانه تجن در نیم قرن اخیر به واسطه قرار گرفتن مرکز استان مازندران در میانه مسیر آن همواره مورد تعرض جدی قرار داشته است. پاک تراشی جنگل در بخش دودانگه (محمدآباد و فریم صحرا)

ومتعاقب آن فرسایش شدید خاک و حمل آن به بخش پائین دست اراضی جنگلی، احداث سد سلیمان تنگه (شهید رجائی)، احداث و رونق نسبی جاده کیاسر - شه میرزاد، توسعه اراضی کشاورزی و چرای بی رویه دام در مناطق نامبرده، توسعه سریع و غیراصولی شهرستان ساری در سنوات اخیر، دفع غیر اصولی زباله های شهری در حوضه آبریز رودخانه تجن و بالاخره احداث کارخانه چوب و کاغذ مازندران که سالانه چند صد هزارتن پساب تولید کاغذ را مستقیماً وبدون هرگونه فیلتراسیون به داخل رودخانه واریز می کند، گوشه ای از بحران آلودگی زیست محیطی رودخانه تجن در سال های اخیر می باشد (لویزه، ۱۳۸۴).

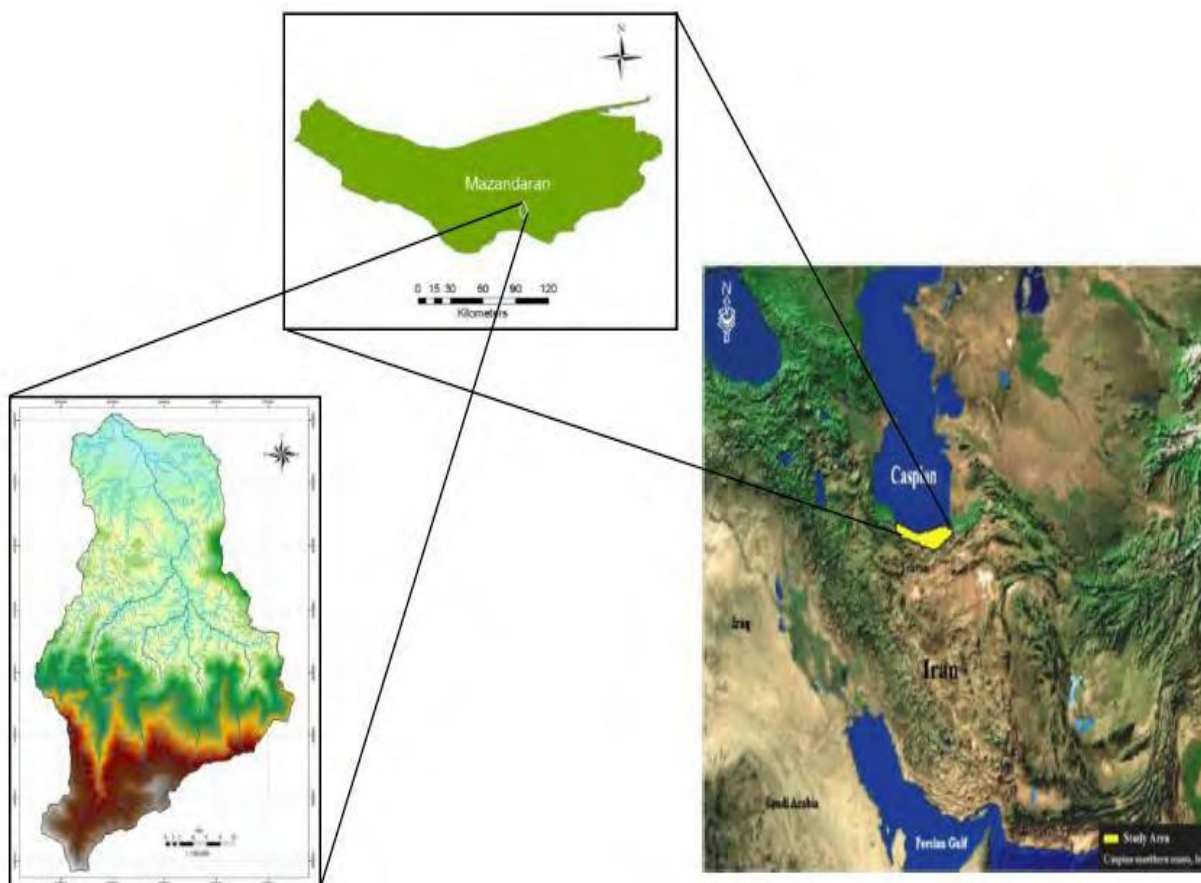


شکل (۳-۱) نقشه محدوده مورد مطالعه رودخانه تجن

۳-۲-۲- رودخانه بابلرود

رودخانه بابلرود در شمال ایران و در جنوب دریای مازندران قرار دارد و بخشی از حوزه آبریز دریای مازندران را تشکیل می‌دهد. این رودخانه از دامنه شمالی البرز سرچشمه گرفته و از دو شاخه اصلی خود به اسامی بابلک و کرسنگ تشکیل می‌گردد. رژیم این رودخانه عمدتاً باران بوده و کمتر تحت تأثیر ذوب برف قرار دارد. در بالاتر از محلی به نام پل محمد حسن خان در اطراف بابل، از سمت غرب رودخانه های سجادرود و کلارود به آن متصل می‌شوند. قبل از پیوستن کلارود به رودخانه بابلرود، آب مسیل‌های متالون و بزود به آن ریخته می‌شوند. آب رودخانه گرم‌رود که به جهت احداث سد دارای اهمیت خاصی می‌باشد در حوالی شهر بابل به رودخانه بابلرود افزوده می‌شود. از آب بابلرود در دشتهای اطراف جهت مصارف کشاورزی برداشت می‌شود. این رودخانه در بابل‌سر وارد دریای مازندران می‌شود. منشأ آبی این حوزه در ارتفاعات برخلاف حوزه های دیگر که برفی است، فقط باران است و دارای جریانهای برف مقطعی و ناپایدار است و در حوزه های کم ارتفاع و دشت نیز طبیعتاً بارانی است. بیش از ۶۰ درصد آب این رودخانه از شاخه های فرعی تأمین می‌شود و دارای رژیم تقریباً دائمی محسوب می‌شود. هر چند در سالهای اخیر به دلایل کاهش میزان بارندگی و به خصوص افزایش فواصل باران و همچنین افزایش سطح زیر کشت حوزه، کل جریان آبی در مقاطعی از سال قطع می‌گردد که همین مسئله موجب قطع حیات موجودات زنده می‌شود، اما در بعضی از مقاطع سال نیز حالت طغیانی داشته و بسیاری از بته ها و حواشی رودخانه را شسته و در مناطق پایینتر رسوب می‌گذارد. بستر رودخانه بابلرود در قسمتهای بالادست متشکل از سیستمهای خاکستری و مارنی است که با ماسه و سنگهای آهکی و خاکستری تیره همراه است و در پایین دست از سنگهای دولومیتی و آهکی پوشیده شده است. ضریب فرسایش این رودخانه نسبت به رودخانه های دیگر کمتر است (یعنی حدود ۱/۰) و بار رسوبی آن به ۳۹۲۰۰۰ تن میرسد که در مسیر، خصوصاً در پایین دستها به صورت شن و ماسه برداشت می‌شود اما هنگام طغیان و یا جریانهای شدید، مصالح ریزدانه تر به سمت مصب هدایت میشوند و به همین خاطر غالباً در مصب رسوبگذاری وجود دارد که البته برخی از این رسوبات

مربوط به خاکریزها و سدهای خاکی موقت بستر رودخانه برای جمع آوری و هدایت آب برای کارهای کشاورزی است که به هنگام بارندگی به همراه جریان آب به سمت مصب هدایت می‌شوند (میرزایی و همکاران، ۱۳۹۰).



شکل (۳-۲) محدوده مورد مطالعه رودخانه بابلرود

۳-۳- داده های مورد استفاده

به منظور برآورد پارامترهای کیفی آب رودخانه های تجن و بابلرود به کمک روش برنامه ریزی بیان ژن، اطلاعات ۱۲ متغیر کیفی ماهانه آب برای یک دوره ۴۰ ساله (۹۰-۱۳۵۰) از شرکت آب منطقه ای استان مازندران جمع آوری گردید. نتایج آماری مربوط به میانگین متغیرهای مورد مطالعه در طول

دوره آماری موردنظر در جداول (۱-۳) و (۲-۳) ارائه شده است. به منظور برآورد پارامترها با استفاده از برنامه ریزی بیان ژن ۲۰٪ داده ها را به مرحله آموزش و ۸۰٪ داده ها را نیز به مرحله آزمون اختصاص دادیم.

جدول (۱-۳) پارامترهای آماری اندازه‌گیری‌های کیفی بر روی رودخانه تجن

TDS	SAR	K	Na	Mg	Ca	SO4	Cl	HCO3	pH	EC	Q	پارامتر
mg/lit		meq/lit								Micromohs/cm	m³/s	
443.37	0.870	3.05	1.44	2.15	3.22	1.49	1.28	3.96	7.79	686.12	12.34	میانگین
121.236	0.647	0.0219	1.19	0.673	0.789	0.730	1.061	0.974	0.23	185.44	29.025	انحراف معیار
14698.21	0.419	0	1.41	0.454	0.620	0.530	1.127	0.950	0.09	34390.62	842.47	واریانس
30	0.040	0.010	0.080	0.50	1	0.100	0.150	1.40	6.60	240	0.01	حداقل
1560	7.81	0.360	16.20	4.70	7.40	6.00	15.50	9.80	8.90	2400	489.41	حداکثر

جدول (۲-۳) پارامترهای آماری اندازه‌گیری‌های کیفی بر روی رودخانه بابلرود

TDS	SAR	K	Na	Mg	Ca	SO4	Cl	HCO3	pH	EC	Q	پارامتر
mg/lit		meq/lit								Micromohs/cm	m³/s	
340.01	0.802	0.041	1.252	1.546	2.402	0.576	1.249	3.335	7.628	525.344	10.535	میانگین
181.21	0.762	0.043	1.357	0.744	1.044	0.643	1.434	1.210	0.323	275.637	22.291	انحراف معیار
32839.35	0.580	0.0018	1.842	0.554	1.089	0.414	2.058	1.465	0.105	75976.133	496.909	واریانس
81	0.082	0.01	0.1	0.3	0.1	0.02	0.1	0.1	6.8	128	0.01	حداقل
1394	4.409	0.9	8.7	6.3	8.4	7.2	9.7	7.4	10.3	2080	295	حداکثر

۳-۴- آشنایی با مفاهیم زیست شناختی طبیعی

موجودات زنده از مجموعه ای سلول تشکیل می‌شوند، هر سلول نیز حاوی مجموعه ای مشابه شامل یک یا چند کروموزوم است. کروموزوم ها از رشته هایی به نام DNA تشکیل شده اند که حامل نقشه کد تشکیل موجود زنده است. هر کروموزوم به طور مفهومی از ژن هایی که هر یک کد تشکیل پروتئین به خصوصی را حمل می‌کنند تشکیل شده است. هر ژن نماینده ویژگی به خصوصی از موجود زنده مانند رنگ چشم است. صورت های مختلف این ویژگی مانند آبی، سبز و یا قهوه ای اصطلاحاً آلل نامیده می‌شود. موجودات زنده کروموزوم های متعددی در هر سلول دارند و جایگاه هر ژن در کروموزوم معین است. مجموعه کروموزوم ها یا مواد ژنتیکی را ژنوم می‌نامند. به مجموعه معینی از ژن های یک ژنوم، عبارت ژنوتیپ اطلاق می‌گردد. هر دو موجودی که ژنوم های مشابه داشته باشند، ژنوتیپ مشابه دارند. ژنوتیپ ها پس از رشد موجود زنده، مشخصات فیزیکی و ذهنی مانند رنگ پوست، رنگ چشم، قد و میزان هوش را ایجاد می‌نمایند که اصطلاحاً فنوتیپ نامیده می‌شوند (البرزی، ۱۳۸۸).

۳-۵- الگوریتم های تکاملی

الگوریتم های تکاملی مجموعه ای از روش های ابتکاری هستند که با اشکال متنوع و پیچیدگی های مختلف، تحت نام کلی محاسبات تکاملی مورد استفاده قرار می‌گیرند. برتری این روش ها، انعطاف پذیری در بهینه سازی توابع مختلف است (خلج، ۱۳۹۰). الگوریتم های ژنتیک یکی از انواع الگوریتم های تکاملی هستند که از علم زیست شناسی مانند وراثت، جهش، انتخاب ناگهانی، انتخاب طبیعی و ترکیب الهام گرفته اند (مطیع قادر و همکاران، ۱۳۸۹). این الگوریتم ها از لحاظ ساختاری به سه دسته الگوریتم ژنتیک (GA)، برنامه ریزی ژنتیک (GP) و برنامه ریزی بیان ژن (GEP) تقسیم بندی می‌شوند که تفاوت اساسی سه الگوریتم GA, GP و GEP به ساختار افراد باز می‌گردد: در الگوریتم ژنتیک اعضا رشته های خطی با طول ثابتند، در برنامه ریزی ژنتیک اعضا رشته های غیر خطی با

شکل و اندازه متفاوت هستند و به شکل درخت تجزیه بیان می‌شوند، درخت تجزیه (کد گذاری درختی) که به برنامه نویسی ژنتیکی شهرت دارد، شیوه ای از کدگذاری است که در برنامه های کامپیوتری و برنامه ریزی تکاملی به کار می‌رود، در این شیوه کدگذاری هر کروموزوم مانند درختی است که هر شاخه آن نشان دهنده یک شی یا بخشی از برنامه کامپیوتری است (مطیع قادر و همکاران، ۱۳۸۹). در برنامه ریزی بیان ژن، اعضا رشته های خطی با طول ثابت هستند که سپس به شکل غیر خطی با اندازه و شکل و بیان درختی تعریف می‌گردند. از جمله مهمترین مزایای GEP آن است که کروموزوم های خطی، فشرده، نسبتا کوچک و آسان برای عملیات ژنتیکی تعریف شده اند، هم چنین بیان درختی منحصر کروموزوم های مربوطه که بر اساس سازگاری انتخاب شده اند را بیان می‌نماید، ایجاد این اعضای منحصر به فرد این قابلیت را برای الگوریتم فراهم میکند که با کارایی بالا، از تکنیک های تطبیقی موجود پیشی بگیرد (Ferreira., 2001).

تفاوت اساسی آنها به صورت زیر برشمرده می‌شوند:

- ۱- الگوریتم ژنتیک (GA) با افرادی شامل کروموزوم های با طول ثابت و بدون بیان پیچیده است. در این سیستم، کروموزوم ها طبق برتری فردی و بر اساس شایستگی حفظ می‌شوند.
- ۱- برنامه ریزی ژنتیک (GP) با افرادی با ساختار شاخه ای با اندازه ها و اشکال متفاوت است که در مقایسه با الگوریتم ژنتیک، عوامل بیشتری را در نظر می‌گیرند.
- ۲- برنامه ریزی بیان ژن (GEP) با افرادی کدگذاری شده به شکل کروموزوم های خطی با طول ثابت است که به شکل ساختار شاخه ای و با اندازه و اشکال متفاوت بیان می‌گردند. در این سیستم ها، کروموزوم ها به وسیله برتری عوامل سببی روی فنوتیپ (ساختار شاخه ای) حفظ می‌شوند.

۳-۵-۱- الگوریتم ژنتیک (GA)

الگوریتم ژنتیک در سال ۱۹۶۰ توسط جان هلند ابداع شد. الگوریتم های ژنتیک بر نظریه تکامل داروین استوار می باشند و از این نظریه در سیستم های کامپیوتری استفاده می کنند. در این روش راه حل های مسئله در رشته هایی با طول ثابت و به صورت ۰ و ۱ کد گذاری می شوند که کروموزوم نام دارند، فرآیند حل مسئله با مجموعه ای از این کروموزوم ها آغاز می گردد که جمعیت اولیه نامیده می شوند و جواب های حاصل برای تولید جمعیت بعدی مورد استفاده قرار می گیرند. انتخاب جواب ها از بین کل جواب ها (والدین) به منظور ایجاد جواب های جدید (فرزندان) براساس شایستگی یا میزان برازش در دامنه مسئله صورت می پذیرد. طبیعی است جواب های سازگارتر، شانس بیشتری برای ابقا و انتخاب مجدد دارند. این فرآیند در نهایت به تکامل جمعیتی از اعضا ختم می گردد که نسبت به اعضای اولیه جمعیت (والدین)، با محیط سازگاری بیشتری دارند (البرزی، ۱۳۸۸؛ مطیع قادر و همکاران، ۱۳۸۹؛ خلج، ۱۳۹۰).

۳-۵-۱-۱- عملگرهای الگوریتم ژنتیک

۳-۵-۱-۱-۱- عملگر انتخاب و تولید مثل

در مرحله انتخاب، ابتدا به کمک تابع برازش، به هر عضو جمعیت، ارزشی نماینده میزان سازگاری آن تعلق می گیرد. تابع برازش میزان برازندگی و عملکرد هر عضو جمعیت را در حل مسئله ارزیابی می کند. در جهان طبیعی، این تابع نشان دهنده قدرت و توانایی موجودات برای زنده ماندن در محیط اطراف خودشان است.

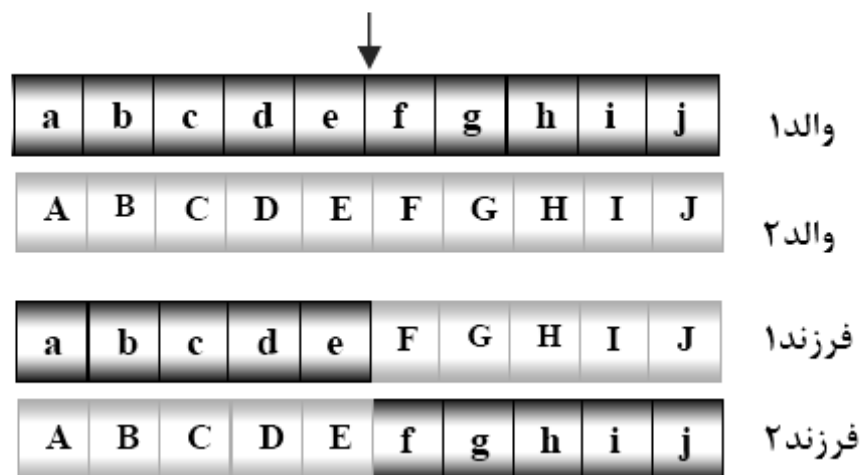
پس از ارزش گذاری، اعضای سازگارتر، شانس بیشتری برای انتخاب خواهند داشت. هنگامی که میزان سازگاری هر عضو تعیین شود، اعضا را می توان با احتمالی متناسب با میزان سازگاری نسبی آن ها برای تولید نسل بعدی انتخاب کرد.

۳-۵-۱-۱-۲-عملگر تقاطع (ترکیب)

این عملگر امکان ترکیب ویژگی های دو کروموزوم را برای تولید نسل جدید فراهم می کند. در این مرحله به اندازه نرخ ترکیب در کروموزوم انتخاب شده و با یکدیگر ترکیب می شوند. ترکیب انواع مختلف دارد که در ادامه به آنها اشاره می شود:

۳-۵-۱-۱-۲-ترکیب تک نقطه ای

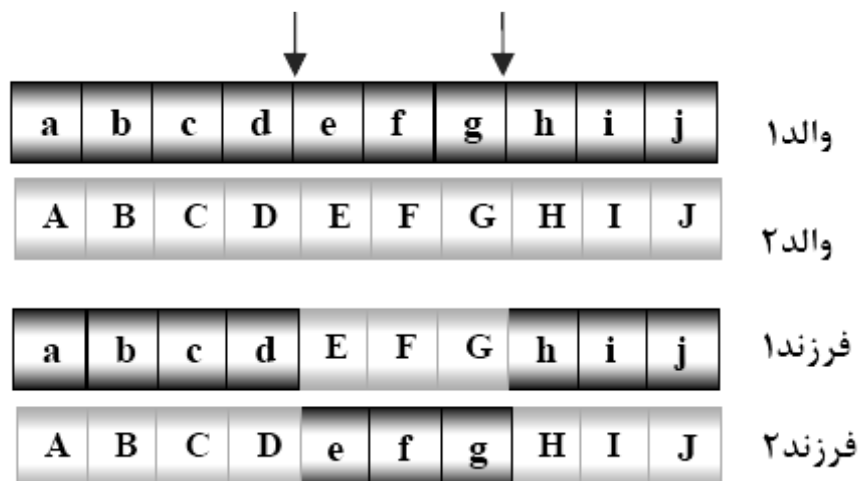
رایج ترین نوع ترکیب، ترکیب تک نقطه ای است. در این روش همانطور که در شکل (۳-۱) نشان داده شده است، کروموزوم های والدین در یک نقطه تصادفی به دو قسمت تقسیم شده و هر فرزند با انتخاب قسمت اول از کروموزوم یک والد و قسمت دوم کروموزوم والد بعدی ساخته می شود.



شکل (۳-۳) - ترکیب تک نقطه ای

۳-۵-۱-۱-۲-ترکیب دو نقطه ای

به کمک این عملگر می توان ژن هایی که خصوصیات مناسب دارند را حفظ نمود. در این روش، پس از انتخاب والدین و انتخاب دو نقطه تصادفی بر روی رشته های کروموزوم، محتویات بین دو نقطه تبادل می گردد (شکل ۳-۲).



شکل (۳-۴) - ترکیب دو نقطه ای

۳-۵-۱-۱-۲-۳- ترکیب یکنواخت

در ترکیب یک و دو نقطه ای، انتقال خصوصیات به صورت جابه جایی قسمتی از رشته کروموزوم مورد نظر انجام می گیرد. اما در شرایطی که ژن ها با خصوصیات مناسب در طول کروموزوم پراکنده باشند، این ترکیب مفید تر از دو روش قبل عمل می کند و خصوصیات مناسب بعضی از ژن ها را که در طول رشته کروموزوم پراکنده شده اند را منتقل می نماید. در ترکیب یکنواخت، به منظور تولید فرزندان، بعضی از بیت ها در رشته والدین به طور تصادفی جابه جا می گردند (شکل ۳-۳).

a	b	c	d	e	f	g	h	i	j	والد ۱
0	0	1	0	1	1	0	1	0	0	ماسک
A	B	C	D	E	F	G	H	I	J	والد ۲
A	B	c	D	e	f	G	h	I	J	فرزند ۱
a	b	C	d	E	F	g	H	i	j	فرزند ۲

شکل (۳-۵) - ترکیب یکنواخت

۳-۵-۱-۱-۳-عملگر جهش

پس از انجام عملیات ترکیب بر روی کروموزوم ها، نوبت به عملگر جهش کی رسد این عملگر، به جمعیت ثانویه ویژگی هایی می دهد که ممکن است در جمعیت اولیه وجود نداشته باشد. همچنین مانع از همگرایی سریع الگوریتم ژنتیک خواهد شد. عملگر جهش در کروموزوم هایی که به صورت صفر و یک کد گذاری شده اند، بیت های تشکیل دهنده کروموزوم را به صورت تصادفی و با احتمال ناچیزی مثلا ۰,۰۰۱ تا ۰,۰۱ تغییر میدهد (شکل ۳-۴). افزایش نرخ جهش، به الگوریتم آزادی بیشتری برای جستجو در فضای مسئله می دهد. عملگر جهش بر روی کروموزوم های نسل آخر اعمال نمی شود.

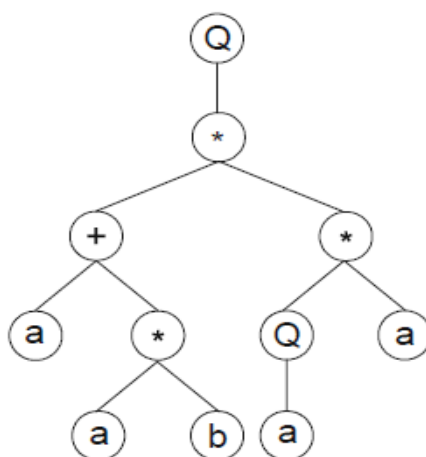
A	B	C	D	E	F	G	H	کروموزوم اولیه
A	B	C	d	E	F	G	H	کروموزوم جهش یافته

شکل (۳-۶) - جهش

۳-۵-۲- برنامه ریزی ژنتیک (GP)

برنامه ریزی ژنتیک ابتدا در سال ۱۹۸۵ توسط کرامر ابداع شد و پس از آن در سال ۱۹۹۲ توسط کوزا توسعه پیدا کرد. این روش با ایجاد ساختارهای غیر خطی (درخت تجزیه)، با اندازه های و اشکال متفاوت، قادر به ایجاد راه حل های تناوبی برای مسائل با طول ثابت می باشد. به منظور آشنایی با درخت تجزیه در شکل (۳-۵) ساختار درختی رابطه (۳-۱) نشان داده شده است (Q ریشه دوم است).

$$\sqrt{(a + (a \times b)) \times (\sqrt{a} \times a)} \quad (۳-۱)$$



شکل (۳-۷) - بیان درختی رابطه (۳-۱)

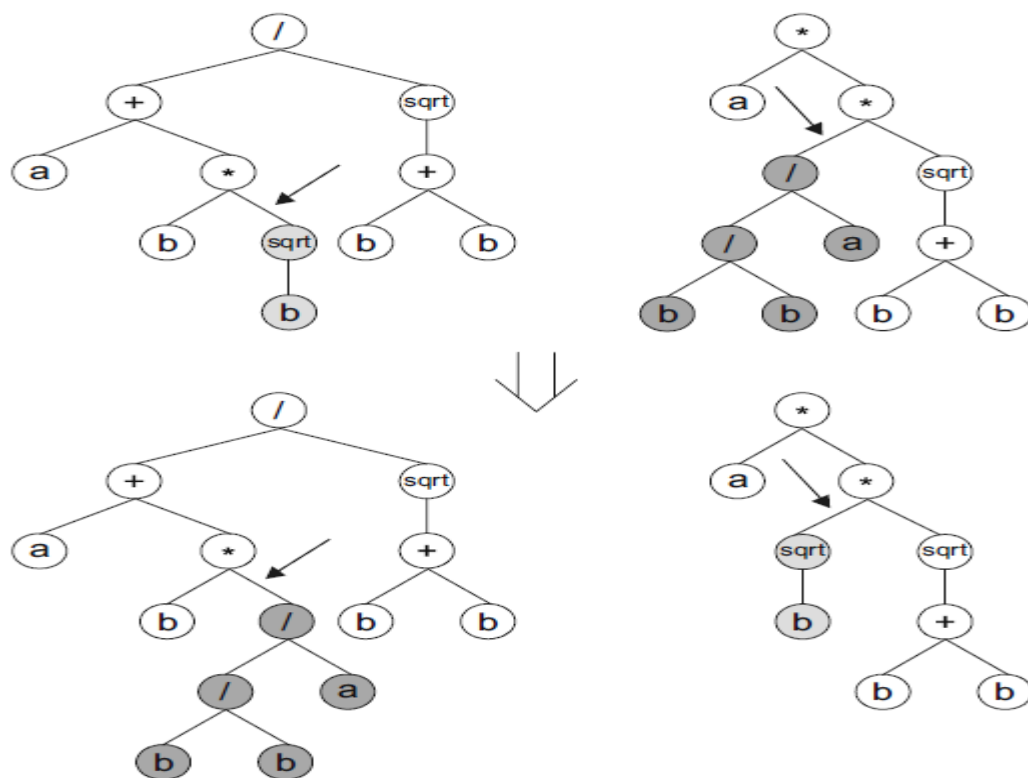
الفبای مورد استفاده در GP، متفاوت تر از ۰ و ۱ در الگوریتم ژنتیک است و سیستم انطباق پذیر تری را ایجاد می نماید. ساختارهای غیر خطی در GP مانند کروموزوم های GA دارای نقش دوگانه ژنوتیپ- فنوتیپ هستند. در GP عملگرهای ژنتیکی مستقیماً بر روی درختان تجزیه عمل می کنند، این موضوع در نگاه اول مفید قلمداد می گردد اما در حقیقت این امر بسیار محدود کننده است و تنها در دامنه محدودی از بهسازی ها امکان پذیر خواهد بود (کریمی، ۱۳۹۳)

عملگرهای ژنتیکی در GP شامل سه عملگر ترکیب، جهش و جایگشت می‌باشد که در ادامه به طور مختصر معرفی خواهد شد:

۳-۵-۲-۱- عملگر ترکیب

این عملگر رسیدن به جواب بهینه عمومی را با ترکیب کردن مقادیر متغیر فرزندان تامین می‌کند (شکل ۳-۶). در بیشتر موارد تنها از این عملگر در GP استفاده می‌شود، در این مورد برای ایجاد فرزندان، شاخه‌هایی به طور تصادفی بر روی دو درخت والد انتخاب شده و سپس به طور تصادفی معاوضه می‌شوند. هدف از اعمال این عملگر، استنتاج راه حل های پیچیده تری است که از نظر قواعد ساختار صحیحی داشته باشند (فریرا، ۲۰۰۱).

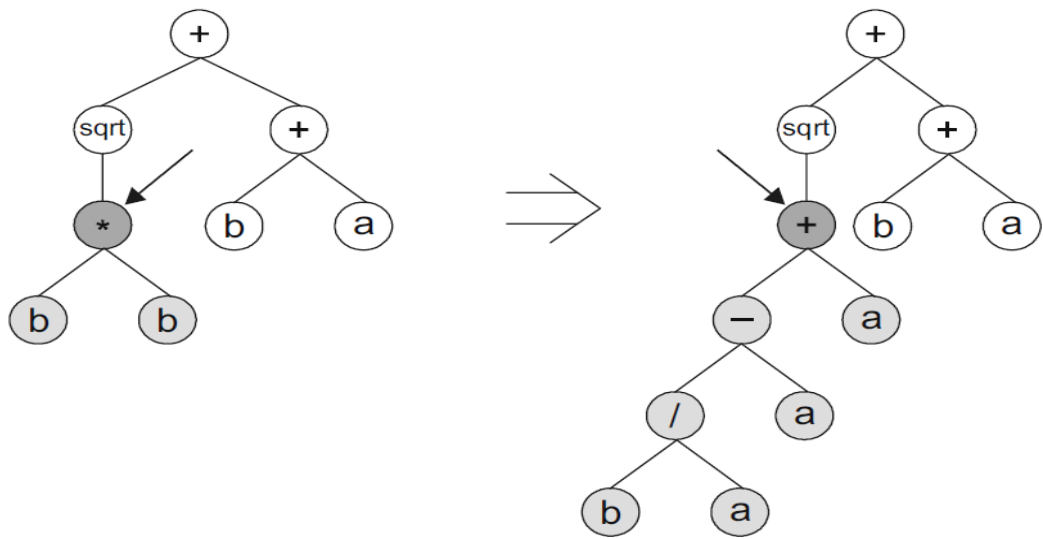
در حقیقت، ترکیب به معنای باز ترکیب اطلاعات مابین کروموزوم ها می‌باشد. این عملگر، خصایص والدین را برای ساختن فرزندان ترکیب می‌کند تا کروموزوم های بهتری ایجاد شوند. به طور معمول این عملگر روی یک جفت از کروموزوم ها عمل می‌کند و دو فرزند برای هر جفت تولید می‌شود. عملگر ترکیب می‌تواند روی چندین والد نیز عمل کند که در این صورت خصیصه های بیش از دو والد را برای تولید فرزندان، ترکیب می‌نماید. علاوه بر این، عملگر ترکیب می‌تواند بیش از دو فرزند را برای هر گروه از والدین تولید کند. وظیفه اصلی این عملگر بهبود برازندگی جمعیت می‌باشد و به منظور حفظ تنوع و گوناگونی جمعیت، مانع از به ارث بردن تنها ژن های خوب خواهد شد (مطیع قادر و همکاران، ۱۳۸۹).



شکل (۳-۸)- ترکیب خطی در برنامه ریزی ژنتیک

۳-۵-۲- عملگر جهش

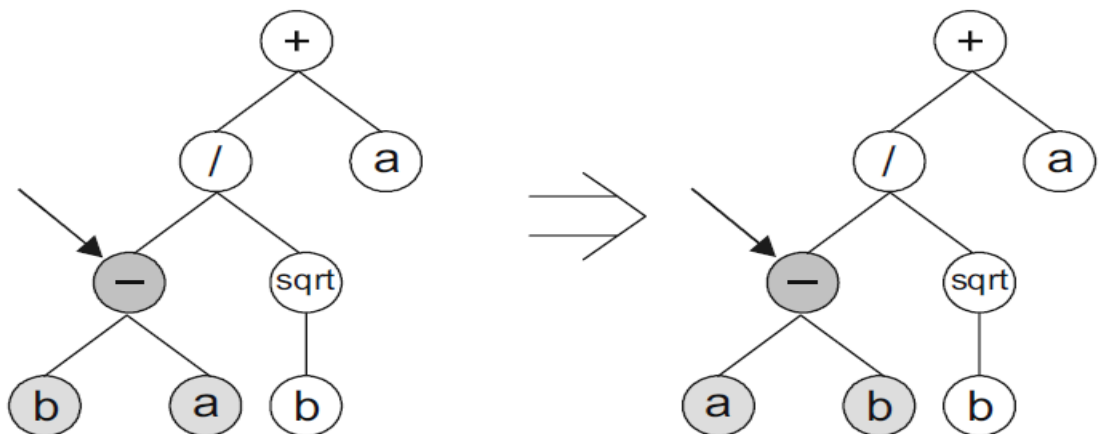
عملگر جهش یک گره در درخت تجزیه انتخاب کرده و شاخه زیرین آن را با یک شاخه جدید تولید شده به صورت تصادفی جابه جا می‌نماید (۳-۷). در این حالت شکل درخت تجزیه دستخوش تغییرات زیادی نخواهد شد. این عملگر دستیابی برنامه ریزی ژنتیک را به جواب های بهینه تضمین می‌نماید (فریرا، ۲۰۰۱).



شکل (۹-۳) - جهش درختی در برنامه ریزی ژنتیک

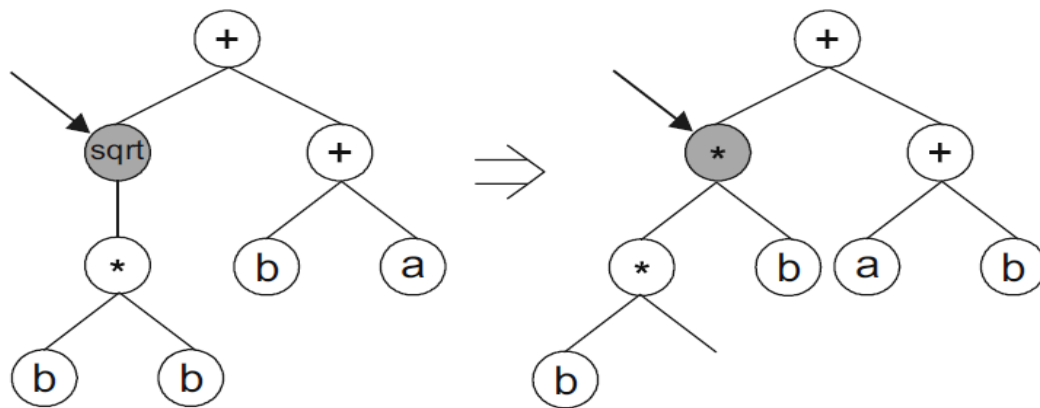
۳-۲-۵-۳- عملگر جایگشت

این عملگر محافظه کارترین عملگر در برنامه ریزی ژنتیک است، روش اعمال آن به این صورت است که در ابتدا متغیرهایی به طور تصادفی روی درخت تجزیه انتخاب شده و به شکل اتفاقی با هم معاوضه می‌شوند (شکل ۸-۳). در این حالت شکل کلی درخت تجزیه بدون تغییر باقی می‌ماند (فریرا، ۲۰۰۱).



شکل (۱۰-۳) - عملگر جایگشت در برنامه ریزی ژنتیک

در روش برنامه ریزی ژنتیک استفاده از عملگرهای جهش، ترانهش و یا وارون سازی با تکرار زیاد، موجب ایجاد ساختارهای نامعتبر و نادرست خواهد گردید و فضای را به طور کشف نشده باقی می‌گذارد (فریرا، ۲۰۰۱). در شکل (۳-۹) نمونه ای از جهش نقطه ای در GP نشان داده شده است که درخت ایجاد شده ساختار صحیحی ندارد.



شکل (۳-۱۱) - یک واقعه فرضی از جهش نقطه ای در برنامه ریزی ژنتیک و ساختار نامعتبر درخت ایجاد شده

اگرچه کوزا عملگرهای ترکیب، جهش و جایگشت را به عنوان عملگرهای اصلی GP معرفی کرده است، اما در عمل، عملگر ترکیب بیشترین نقش را داراست و استفاده از این عملگر، مانع از ورود مواد ژنتیکی جدید به منبع ژنتیکی GP خواهد شد.

از جمله مزیت های برنامه ریزی ژنتیک در مقایسه با سیستم های عصبی آن است که در برنامه ریزی ژنتیک ابتدا ساختار بلوک ها (متغیرهای ورودی، هدف و مجموعه توابع) تعریف شده و سپس ساختار بهینه مدل و ضرایب طی فرآیند آموزش تعیین می‌شوند، همچنین GP به طور خودکار قادر به گزینش متغیر های ورودی با بیشترین تاثیر در مدل می‌باشد. از آنجاییکه در برنامه ریزی ژنتیک، درختان تجزیه هم نقش ژنوتیپی دارند و هم به عنوان فنوتیپ عمل می‌کند، از مقایسه GP و GA

می‌توان نتیجه گرفت که هر دو روش توانای بیان ساده ندارند و در همه موارد از کل درخت تجزیه به عنوان راه حل استفاده می‌شود (فریرا، ۲۰۰۱).

۳-۵-۳- برنامه ریزی بیان ژن (GEP)

روش برنامه ریزی بیان ژن، ترکیب و توسعه یافته روش‌های GA و GP است که در سال ۱۹۹۹ توسط فریرا ابداع شد. در این روش، کروموزوم‌های خطی و ساده با طول ثابت، مشابه با الگوریتم ژنتیک و ساختارهای شاخه‌ای با اندازه و اشکال متفاوت مشابه با درختان تجزیه در برنامه ریزی ژنتیک ترکیب می‌شوند. به طور خلاصه می‌توان بیان داشت که در این روش فنوتیپ و ژنوتیپ از هم جدا شده و سیستم قادر خواهد بود از تمام مزایای تکاملی بهره گیرد. با وجود اینکه فنوتیپ در GEP مشابه ساختار شاخه‌ای GP می‌باشد، اما ساختار شاخه‌ای در GEP که بیان درختی نیز نامیده می‌شود، بیانگر تمام ژنوم‌های مستقل است.

در مجموع، به طور خلاصه می‌توان بیان داشت که در GEP بهسازی‌ها در یک ساختار خطی اتفاق می‌افتد و سپس به صورت ساختار درختی بیان می‌شود، این امر سبب خواهد شد که تنها ژنوم اصلاح شده به نسل بعد انتقال یابد و نیازی به ساختارهای سنگین تکثیر و جهش نباشد (فریرا، ۲۰۰۴).

۳-۵-۳-۱- الگوریتم‌های برنامه ریزی بیان ژن

اولین مرحله در GEP، تشکیل جمعیت اولیه از راه حل‌هاست. سپس کروموزوم‌ها به بیان درختی (نشان داده می‌شوند) در مرحله بعد باید کارایی یا میزان سازگاری هر عضو از جمعیت کروموزوم‌ها توسط تابع برازش ارزیابی شود. اگر چه در حل مسایل به روش GEP نیازی به دانستن دقیق ساختار ریاضی مسایل نیست، اما باید به روشی میزان مطلوبیت جواب‌های بالقوه مسئله (کروموزوم‌ها) و میزان سازگاری آن‌ها را تعیین کنیم و با حذف جواب ضعیف‌تر و حفظ جواب‌های سازگارتر، پاسخ بهینه مسئله نزدیک شویم.

در پدیده تکامل و در نظام انتخاب طبیعی این نقش بر عهده طبیعت است که پاسخ های سازگارتر را بر دیگر جواب ها ترجیح می دهد. در سیستم GEP معیارهای توقف یک و یا ترکیب دلایل زیر می باشد:

- معیار دستیابی به هدف: عدم مشاهده تفاوت محسوس بین پاسخ های بهینه طی چند نسل، می تواند معیاری برای دستیابی به پاسخ بهینه باشد.
- معیار تکرار: اگر پس از تعداد نسل معین به پاسخ بهینه دست یابیم، فرآیند تولید نسل متوقف خواهد شد.
- معیار زمان: اگر پس از طی زمانی معین، در بهترین جواب حاصل از تولید نسل بهبودی حاصل نگردد، محاسبات متوقف خواهد شد.

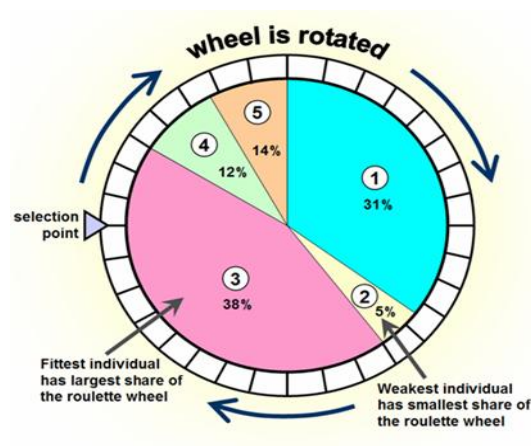
به طور کلی مراحل اصلی الگوریتم های تکاملی ذکر شده را می توان به صورت زیر خلاصه کرد:

فرایند با تولید تصادفی کروموزوم ها از جمعیت اولیه آغاز می شود. سپس کروموزوم های ایجاد شده در GP و GEP به صورت درختی بیان می گردند. در ادامه به کمک تابع هدف میزان برازندگی و کارایی هر فرد ارزیابی می شود، پس از آن اعضای سازگارتر انتخاب و اصلاح شده تا فرزندان جدیدتر ایجاد شوند، فرزندان ایجاد شده نیز تحت فرآیند توسعه ای قرار می گیرند تا پاسخ خوب و مناسب پیدا شود (فریرا، ۲۰۰۴).

لازم به ذکر است که منظور از تولید مثل، تکثیر و اصلاحات ژنتیکی است. منظور از تکثیر، حفظ چندین فرد مناسب از نسل حاضر برای نسل بعد است. واضح است که تکثیر به تنهایی نمی تواند نشانگر تغییرات باشد و تغییرات ژنتیکی به واسطه عملگرهایی از قبیل جهش، واریاسیون، ترانزیشن IS، ترانزیشن ریشه RIS، ترانزیشن های ژنی، ترکیب تک نقره ای، ترکیب دو نقطه ای و ترکیب ژنی انجام می شود (فریرا، ۲۰۰۱).

در GEP از روش چرخ رولت برای انتخاب اعضا استفاده می شود. در این روش، اعضا براساس میزان سازگاری نسبی گزینش می گردند، برای این منظور ابتدا میزان سازگاری خام تمام اعضا را جمع کرده،

سپس عدد حقیقی بدست آمده (SUM) رابه نسبت احتمال مورد نظر انتخاب (میزان سازگاری های نسبی) اعضا تقسیم بندی می کنیم. با این شرایط هر عضو در یکی از این تقسیمات منظور می شود. برای این منظور از دایره ای که در شکل (۳-۱۰) آمده است استفاده می کنیم. محیط دایره به اندازه SUM و اندازه هر تکه به اندازه سازگاری نسبی یا خام عضو مربوطه است. اگر این دایره را مانند فرفره به چرخش در آوریم، احتمال قرارگیری نشانه ثابت در کنار هر یک از مقاطع به اندازه آن مقطع یا سازگاری نسبی آن ارتباط خواهد داشت (البرزی، ۱۳۸۸). مثلاً در شکل (۳-۱۰) عضو شماره ۳ بیشترین احتمال انتخاب و عضو شماره ۲ کمترین احتمال گزینش را خواهد داشت.



شکل ۳-۱۲ چرخ گردان

در GEP بر خلاف GA و GP چندین عملگر ژنتیکی به منظور اصلاح نسل مورد استفاده قرار می گیرند که در ادامه معرفی خواهند شد (لوپز و وینرت، ۲۰۰۴).

۳-۵-۲- ساختار برنامه ریزی بیان ژن

برنامه ریزی بیان ژن، با خلق کروموزوم ها و بیان درختی، دارای بینشی گسترده در ایجاد راه حل های جدید برای محاسبات تکاملی است. لذا به منظور کد گذاری اطلاعات و بیان کروموزوم ها، یک زبان جدید موسوم به گاروا توسط فریرا خلق شد. همچنین، ساختار کروموزوم ها به گونه ای طراحی شده است که شرایط ایجاد چندین زن که هر یک برای بیان درختی یا برنامه کوچکتر کد گذاری می شود را فراهم کند (فریرا، ۲۰۰۱).

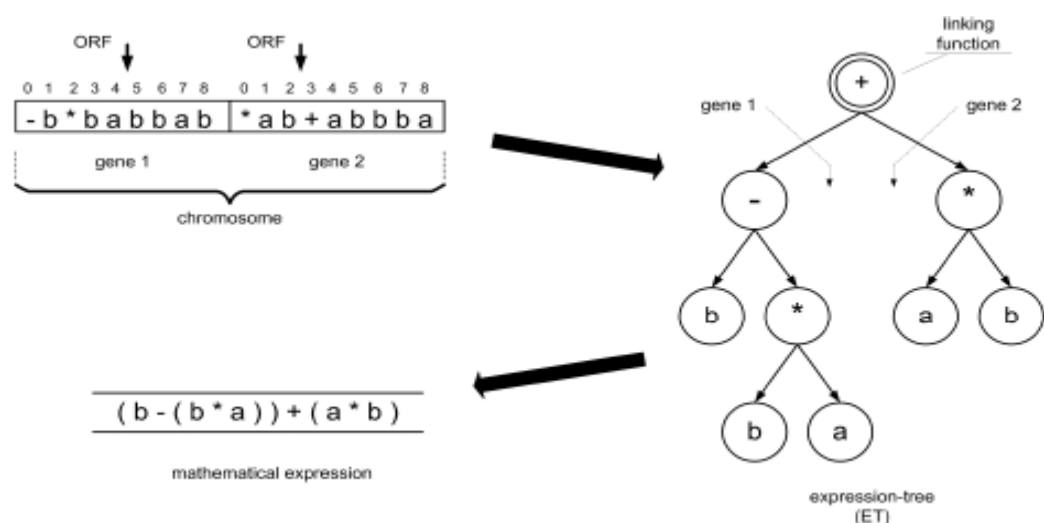
سیستم سلولی GEP به شکل خودکار و بسیار ظریف، مجموعه توابع تعریف شده را به کار می گیرد (فریرا، ۲۰۰۶). ساختار ساده و در عین حال قابل تغییر و تحول ژنهای GEP، علاوه بر متصور ساختن کد گذاری برنامه ها، شرایط تکامل را نیز فراهم می سازند. همچنین با ساختاری تطبیق پذیر و به کار گیری مجموعه ای قدرتمند از عملگرهای ژنتیکی، به گونه ای مناسب فضای راه حل را جستجو می نمایند. درست مشابه با آنچه در طبیعت دیده می شود، عملگرهای GEP نیز ساختارهای صحیح و مناسب را جستجو می کنند و به طور قابل ملاحظه ای منطبق بر تنوع ژنتیکی هستند (فریرا، ۲۰۰۴).

در ساختار فنوتیپ GEP نیز مانند GP، درختان توابع و ترمینال وجود دارد، با این تفاوت که در GP فرایند ترجمه (کدگذاری ژنها) وجود ندارد و فنوتیپ و ژنوتیپ دارای یک ساختار هستند، اما در GEP، فنوتیپ از ترجمه یک کروموزوم خطی بدست می آید. به عبارت دیگر، در GEP ژنومها کروموزوم های خطی هستند و فنوتیپ ها، بیان درختی (ET) می باشند که به صورت مجزا وجود دارند. بیان درختی، تحت فرایند ترجمه از کروموزوم ها استنتاج می شود. در واقع کروموزوم ها تحت انتخاب و عملگرهای ژنتیکی قرار می گیرند و میزان سازگاری آنها با بیان درختی نمایش داده می شود (ویلسون، ۲۰۰۸).

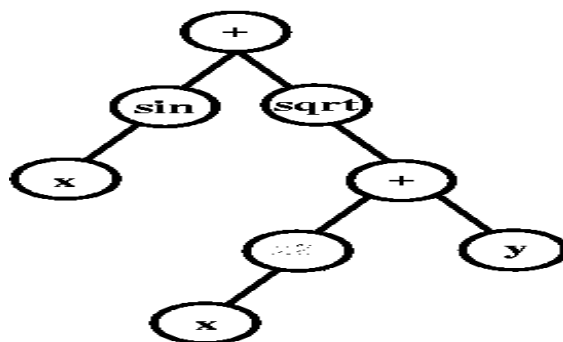
۳-۵-۳-۳- ساختار افراد در GEP

۳-۵-۳-۱- ORF و بیان درختی ژن ها

در GEP، ژنها به شکل درختی بیان و کد گذاری می‌شوند. در مورد کروموزوم های چند ژنی تمامی بیان درختی ها به وسیله تابع پیوند، از محل گره ریشه خود به یکدیگر متصل می‌شوند. هر ژن، یک ناحیه کدگذاری به نام ORF (چهارچوب تفسیر باز) یا بیان K دارد که پس از کد برداری و رمز گشایی به صورت ساختار درختی (ET) بیان می‌شود و نمایش دهنده یک راه حل برای مسئله است. شکل (۳-۱۱) نشان می‌دهد که چگونه یک کروموزوم با دو ژن به صورت یک رشته خزی کد گذاری شده و سپس به صورت ET بیان شده است. دو ژن نشان داده شده در شکل، مشابه با توالی های ژن های بیولوژی، نواحی کدگذاری شده و نشده دارند (لوپز و وینرت، ۲۰۰۴).



شکل (۳-۱۳) - کدگذاری به شکل رشته خطی و کد برداری به صورت ET یک کروموزوم با دو ژن در GEP مثال زیر برای درک بهتر از نحوه کدگذاری ORF و کد برداری به شکل ET ارائه شده است. به عنوان مثال رابطه ریاضی $F(x) = \sin(x) + \sqrt{x^2 + y}$ را در نظر بگیرید، دیاگرام و بیان درختی آن به صورت شکل (۳-۱۲) نشان داده شده است.



شکل (۳-۱۴) - بیان درختی رابطه $F(x) = \sin(x) + \sqrt{x^2 + y}$

این دیاگرام فنوتیپ افراد را در GEP نمایش می‌دهد. ژنوتیپ افراد نیز با تفسیر فنوتیپ و خواندن ET از چپ به راست و بالا به پایین به صورت عبارت (۳-۲) بدست خواهد آمد.

0 1 2 3 4 5 6 7 8

(۳-۲)

+ sin sqrt x + ^ y x 2

بیان (۳-۲) یک ORF است که در + (موقعیت ۰) شروع شده و در ۲ (موقعیت ۸) خاتمه می‌یابد. فرایند معکوس، یعنی ترجمه بیان k به صورت بیان درختی (ET) هم به سادگی صورت می‌پذیرد.

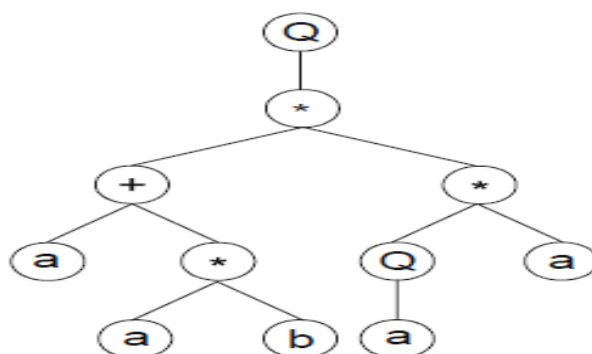
بیان ORF زیر را در نظر بگیرید.

01 234 5678 90

(۳-۳)

Q*+*a*Qaaba

نقطه شروع ORF متناظر با گره ریشه از ET است که در مثال فوق، + می‌باشد. در ادامه، توابع به صورت شاخه‌های افزایشی به آن متصل می‌گردند، تا جایی که شناسه توابع وجود داشته باشد. عمل اضافه شدن زمانی خاتمه می‌یابد که خط انتهایی که فقط شامل ترمینال‌ها (متغیرها یا مقادیر مورد استفاده در یک مسئله) است، تشکیل گردد. بیان ET رابطه فوق به شکل زیر می‌باشد (فریرا، ۲۰۰۱).



شکل (۳-۱۵)- بیان درختی رابطه (۳-۳)

در GEP کروموزوم ها طول ثابت دارند و از یک یا چند ژن با طول مساوی تشکیل شده اند، لذا تنها طول ORFs در GEP تغییر می کند و طول ژنها ثابت باقی می ماند. با اینکه در GEP محل شروع همیشه اولین نقطه از یک ژن است، اما نقطه انتهایی همیشه منطبق بر آخرین محل یک ژن نیست. معمولاً برای ژن های GEP، نواحی کد گذاری نشده در قسمت بعد از نقطه انتهایی وجود دارد. نواحی کد گذاری نشده ف ماهیت GEP و قدرت استنتاج آن هستند، چون اجازه اصلاحات و بهسازی های ژنوم را بدون محدودیت و با استفاده از هر عملگر ژنتیکی می دهند و همیشه از نظر قواعد برنامه های صحیحی را بدون نیاز به قواعد ویرایشی پیچیده یا روش های تحمیلی در جهت اجرای عملگر های ژنتیکی تولید می نمایند. این بزرگترین تفاوت بین GEP و GP بدون ژنوم های خطی است. (فریرا، ۲۰۰۱).

۳-۵-۳-۲- ساختارهای ژن ها در برنامه ریزی بیان ژن

یک کروموزوم از یک یا چند ژن تشکیل شده است و هر ژن به یک سر و یک دنباله تقسیم می شود. اندازه سر (h)، به وسیله کاربر تعیین می شود اما اندازه دنباله (T)، به صورت تابعی از h و پارامتر n بدست می آید. پارامتر n ، بیشترین تعداد پارامتر مستقل یافت شده در مجموعه توابع مورد استفاده است. رابطه (۳-۴) طول دنباله را با توجه به پارامتر های دیگر تعیین می کند (لوپز و وینرت، ۲۰۰۴).

$$t = h(n-1) + 1 \quad (4-3)$$

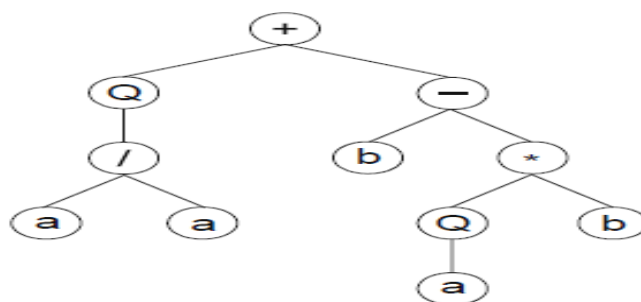
مسایل رگرسیونی با استفاده از مجموعه ای از توابع و ترمینال ها مدلسازی می‌شوند. مجموعه توابع معمولاً شامل توابع اصلی حسابی $\{+, -, *, /\}$ ، توابع مثلثاتی یا هر نوع تابع ریاضی دیگر و یا توابع تعریف شده توسط کاربر است. توابع و ترمینال ها در بخش سر ژن وجود دارند و دنباله تنها شامل ترمینال هاست (لوپز و وینرت، ۲۰۰۴).

به عنوان مثال ژنی را به صورت $\{Q, \times, \div, +, -, a, b\}$ در نظر بگیرید. در این مورد تعداد متغیر های مستقل (n) با ۲ برابر است. برای مقادیر طول سر برابر با ۱۰ و طول دنباله برابر ۱۱، طول ژن برابر ۲۱ خواهد بود این ژن در عبارت (۵-۳) نشان داده شده است که در آن، قسمت ابتدایی سر و قسمت انتهایی دنباله است و رنگ ها تفکیک شده اند (فریرا، ۲۰۰۱).

$$012345678901 \ 2345678 \ 90 \quad (5-3)$$

+Q-/ b*aaQba abaabbbaa ab

و پس از کدبرداری، بیان درختی آن مطابق شکل (۱۴-۳) خواهد بود:



شکل (۱۶-۳)-بیان درختی رابطه (۵-۳)

در این مورد، ORF در موقعیت ۱۰ خاتمه می‌یابد، اما ژن در موقعیت ۲۰ خاتمه می‌یابد که کوقعیت کد گذاری نشده را نشان می‌دهد.

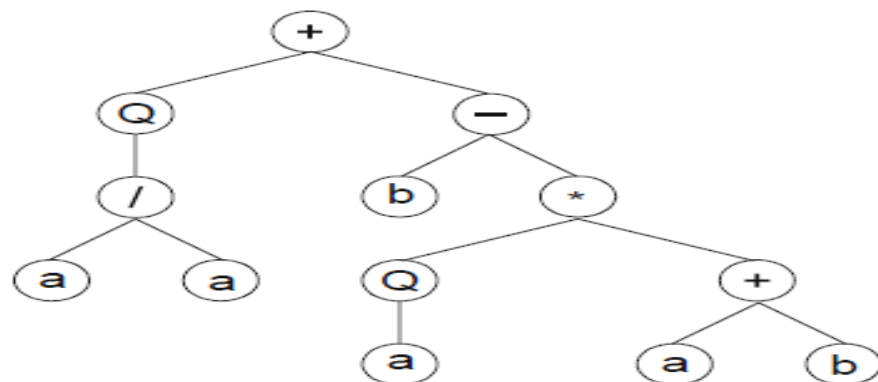
حال در موقعیت ۹ یک جهش را در نطفه بگیرید که سبب تغییر (b) به (+) شده است. پس از ژن به صورت عبارت (۶-۳) نوشته خواهد شد (فریرا، ۲۰۰۱).

012 3456789 0123456 7890

(۶-۳)

+Q-/b*aaQ+aabaabbbaaab

شکل (۱۵-۳) بیان درختی رابطه حاصل را نشان می‌دهد.



شکل (۱۷-۳) - بیان درختی رابطه (۶-۳)

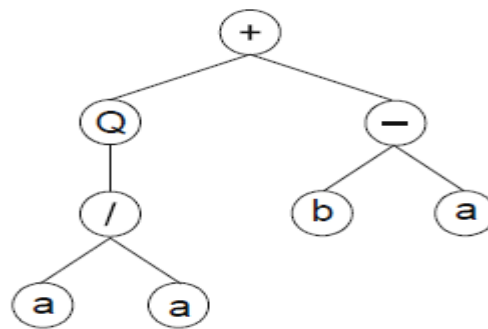
بر اساس جهش حاصل، موقعیت انتهایی ORF، دو موقعیت به سمت راست جابه جا شده است (موقعیت ۱۲). حال اگر در رابطه (۳-۳)، در اثر جهش در موقعیت ۵، (*) به (a) تغییر کند، مطابق با رابطه (۷-۳)، طول ORF، بجای افزایش، کاهش پیدا می‌کند.

01 234567890 1234 567890

(۷-۳)

+Q-/baaaQbaabaabbbaaab

و بیان درختی آن مطابق شکل زیر خواهد بود:



شکل (۳-۱۸) - بیان درختی رابطه (۳-۷)

در این مثال، ORF، در موقعیت ۸ خاتمه می‌یابد و ET اولیه، کوتاه شده و به سه گره می‌رسد.

هر ژن، با وجود طول ثابت توانایی کدگذاری برای ETs با اندازه و اشکال متفاوت دارد. ساده ترین نوع آن فقط از یک گره (زمانیکه اولین المان از یک ژن ترمینال باشد) و بزرگترین نوع آن از تعداد زیادی گره، به اندازه طول ژن (زمانیکه تمام المان های بخش سر توابع با تعداد حداکثر متغیر های مستقل باشند) تشکیل می‌شود (فریرا، ۲۰۰۱).

۳-۵-۴- تابع برازش

یکی از موارد مهم در برنامه ریزی بیان ژن تعیین تابع برازش به منظود یافتن راه حلی است که در تمامی مراحل برازش برای یک خطای معین به خوبی عمل کند. برای برخی کاربردهای ریاضی، استفاده از خطای نسبس یا مطلق کوچک برای یافتن راه حل مناسب، موثر است. اما اگر محدوده انتخابی بینهایت کوچک باشد، جمعیت به کندی استنتاج شده و در نتیجه برای یافتن راه حل مناسب ناتوان خواهد بود. به علاوه اگر محدوده انتخابی وسیع باشد، راه حل های متعددی با برازش حداکثر ظاهر می‌گردد که از راه حل های مناسب دور خواهند بود. به منظور رفع این مشکل یک راه حل تکاملی اختراع شده است که اجازه یافتن راه حل مناسب را بدون توقف تکامل می‌دهد. بنابراین، سیستم قادر به یافتن راه حل مناسب با کمترین خطا خواهد بود (فریرا، ۲۰۰۱).

۳-۶- مراحل اجرای مدل

در پژوهش حاضر از روش برنامه ریزی بیان ژن برای برآورد پارامترهای کیفی آب رودخانه های بابلرود و تجن استفاده شده است. به منظور اجرای روش برنامه ریزی بیان ژن نرم افزار GeneXpro Tools4 (فریرا، ۲۰۰۱) مورد استفاده قرار گرفته است که در ادامه مراحل پیش بینی به کمک این نرم افزار ارائه می گردد.

۳-۶-۱- تعیین همبستگی بین داده ها

در این مطالعه پس از بررسی های لازم به منظور مدل سازی شاخص کیفی (TDS) از اطلاعات و داده های ۱۲ متغیر کیفی ماهانه آب استفاده شد. این متغیرها شامل هدایت الکتریکی آب، اسیدیته، سدیم، کلسیم، پتاسیم، نسبت جذب سدیم، منیزیم، کلر، سولفات، بیکربنات و دی به عنوان متغیر مستقل و TDS به عنوان متغیر وابسته در نظر گرفته شد. با استفاده از نرم افزار SPSS23، همبستگی بین داده ها را مورد سنجش قرار دادیم و سپس به تحلیل رگرسیون بین متغیر های مستقل و وابسته پرداختیم که به این منظور از روش گام به گام (Stepwise Regression) استفاده کردیم. و نیز برای مدل سازی نسبت جذب سدیم همانند شاخص کیفی کل مواد محلول از ۱۲ متغیر کیفی استفاده شد با این تفاوت که نسبت جذب سدیم به عنوان متغیر وابسته و بقیه پارامتر ها به عنوان پارامتر مستقل در نظر گرفته شدند.

۳-۶-۲- برنامه ریزی بیان ژن

۳-۶-۲-۱- تعیین تابع برازش

مقدار کمی که نمایش دهنده کیفیت یک کروموزوم به عنوان یک جواب از مسئله است و در مقابل هزینه قرار دارد و میزان شایستگی یک کروموزوم را در بین جمعیت کروموزومی نشان می‌دهد، به عنوان تابع برازش شناخته می‌شود. در واقع تابع برازش شاخصی برای انتخاب جفت های مناسب و جفت گیری آنها در جمعیت کروموزوم ها فراهم می‌آورد. فرایند GEP مانند سایر الگوریتم ها زمانی خاتمه می‌یابد که به ضابطه تعیین شده ای مثل میزان معین خطای میانگین کارایی اعضای جمعیت رسیده باشد و یا چرخه فرآیند به تعداد دلخواه ادامه یافته باشد در اینصورت بهترین راه حل یافت شده، گزارش داده می‌شود و در غیر اینصورت، بهترین راه حل از نسل حاضر نگه داشته می‌شود و بقیه راه حل ها به فرایند گزینشی واگذار می‌شوند که براساس آن بهترین افراد شانس بیشتری برای تولید فرزندان دارند. با تکرار روند ذکر شده و با پیشرفتن نسل به جلو، کیفیت جمعیت بهبود یافته و به پاسخ بهینه نزدیک خواهد شد (لوپز و وینرت، ۲۰۰۴).

در واقع تابع برازش برای ارزیابی عملکرد هر برنامه خروجی بر مبنای خطای نسبی انتخاب می‌شود و آن را به وسیله ی یک محدوده انتخابی و میزان دقت کاوش می‌کند. از لحاظ ریاضی، تابع برازش F_i به صورت رابطه (۳-۸) بیان می‌شود (فریرا، ۲۰۰۶).

$$f_i = \sum_{j=1}^n \left(R - \left| \frac{P_{(ij)} - T_j}{T_j} \cdot 100 \right| \right) \quad (۳-۸)$$

که در آن R محدوده انتخابی، P_{ij} مقدار پیش بینی شده به وسیله برنامه انفرادی i برای مورد برازش j (از میان n برازش) و T_j مقدار هدف برای مورد برازش J است. عبارت داخل قدر مطلق متناظر با مقدار خطای نسبی است و دقت نامیده می‌شود. چنانچه $P_{ij} = T_j$ باشد، خطا صفر خواهد بود. بنابر

این برای یک برازش کامل که در آن خطا صفر باشد، مقدار f_i از یک برنامه انفرادی i برابر با $f_i = f_{\max}$ و $Nr =$ خواهد بود (فریرا، ۲۰۰۶).

۳-۲-۶-۳- انتخاب مجموعه ترمینال ها و مجموعه توابع

مرحله دوم شامل انتخاب مجموعه ترمینال ها و مجموعه توابع برای ایجاد کروموزوم ها است. برای برقراری رابطه بین داده های ورودی و خروجی لازم است توابعی تعریف شود، در پژوهش حاضر علاوه بر چهار عمل اصلی $\{+, -, \times, \div\}$ ، توابع $\{x^2, x^3, x^{1/3}, x^{1/2}, \exp, \log, \cos, \sin, \tan^{-1} x\}$ نیز مورد استفاده قرار گرفته اند. لازم به یادآوری است که در ساختار کروموزوم ها، توابع و ترمینال ها در بخش سر ژن قرار می گیرند و ترمینال ها در قسمت دنباله قرار دارند.

۳-۲-۶-۳- انتخاب ساختار کروموزوم

مرحله سوم انتخاب ساختار کروموزوم ها می باشد. این مرحله شامل تعیین طول سر و تعداد ژن ها می باشد. در پژوهش حاضر با سعی و خطا بر روی تعداد کروموزوم ها، ژن ها و اندازه سر، ساختار کروموزوم ها و اندازه دنباله از رابطه (۳-۴) تعیین گردید.

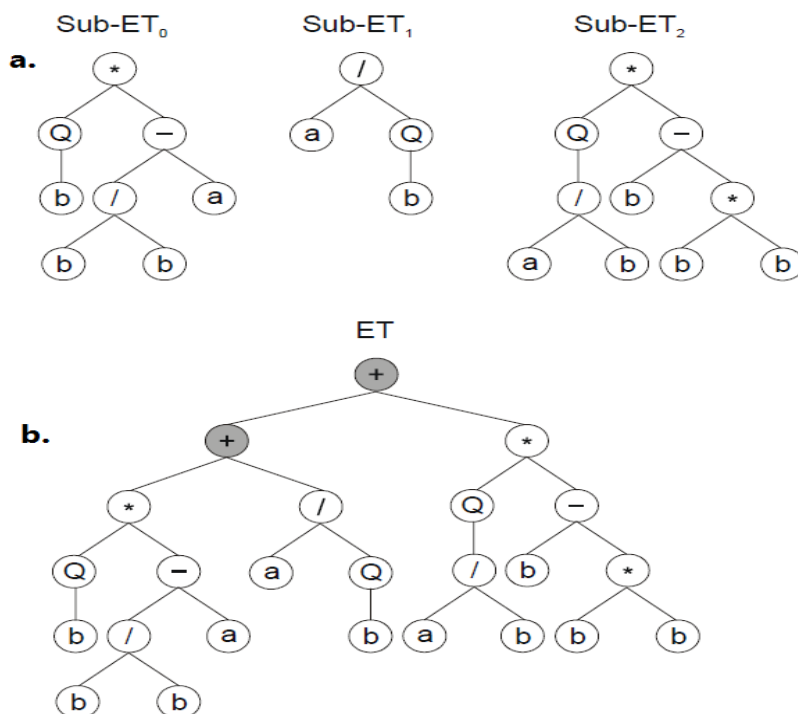
در سعی و خطا به منظور تعیین تعداد کروموزوم ها توجه به این نکته ضروری است که در تعداد کم کروموزوم ها، امکان شکل گیری فضای جستجو وجود ندارد، هم چنین تعداد بسیار زیاد کروموزوم ها روند جستجو را کند خواهد کرد. بر اساس تحقیقات، جمعیت های مناسب بین ۲۰ تا ۳۰ کروموزوم دارند. البته گاه جمعیت هایی با ۵۰ تا ۱۰۰ کروموزوم نیز بهترین جواب را داده اند. تحقیقات دیگر نیز نشان داده اند که اندازه جمعیت باید بر اساس نوع مسئله و کدگذاری آن تعریف گردد و افزایش بیشتر آن بی فایده خواهد بود و کمکی به حل مسئله نخواهد کرد (مطیع قادر و همکاران، ۱۳۸۹).

۳-۲-۶-۴- انتخاب تابع پیوند

کروموزوم ها در برنامه ریزی بیان ژن، معمولا از چندین ژن با طول مساوی تشکیل شده اند. برای هر مسئله یا هر اجرا، تعداد ژن ها نیز مانند طول سر قابل انتخاب است. هر ژن، به صورت یک زیر بیان

درختی کدگذاری می‌شود. استفاده از کروموزوم های چندژنی، جهت استنتاج راه حل در مسایل پیچیده مناسب تر است، نتایج استفاده از هر دو نوع سیستم های تک ژنی و چند ژنی، نشان می‌دهد که در اکثر موارد سیستم‌های چندژنی موثرتر از سیستم های تک ژنی عمل می‌کند (خلج، ۱۳۹۰). در پژوهش حاضر نیز از سیستم چندژنی استفاده شده است که در آخر ژن های انتخابی از طریق تابع مجموع به یکدیگر پیوند داده می‌شوند.

به عنوان مثال شکل (۳-۱۹) تابع پیوند یک سیستم سه ژنی را نشان داده است (فریرا، ۲۰۰۶).



شکل (۳-۱۹) - a) کدگذاری سه ژن، b) پیوند سه ژن

۳-۶-۲-۵- انتخاب عملگرهای ژنتیکی

هدف از عملیات تکثیر، نگهداری اعضای یک نسل برای تولید نسل بعدی است. در حقیقت، این یک عملیات احتمالی است که در طول انتخاب با استفاده از چرخ رولت رخ می‌دهد و افراد تکثیر شده در معرض فعالیت عملگر های زنتیکی از قبیل ترکیب، جهش و غیره قرار خواهند گرفت. یکی از عکسگرهایی که در برنامه ریزی بیان ژن مورد استفاده قرار میگیرد عملگر جهش است. هدف عملگر

جهش، بهسازی تصادفی داخل مجموعه کروموزوم ها است. با این عمل کروموزوم های جدیدی که به احتمال زیاد در کل جمعیت وجود نداشته اند به وجود می آیند (مطیع قادر و همکاران، ۱۳۸۹).

به منظور بهبود عملکرد عملگر جهش، می توان از تابیری در آن استفاده کرد. به عنوان مثال می توان گرایش عملگر را نسبت به اعضای ضعیف تر افزایش داد تا ضمن حفظ خصوصیات اعضای برآورنده تر، شانس جستجو در اعضای ضعیف افزایش یابد. در این الگوریتم ها، به دلیل نخبه گرایی بر روی بهترین کروموزوم هر نسل جهش انجام نمی شود و بهترین جواب بدون هیچ تغییری به نسل بعد انتقال می یابد، این امر سبب می گردد بهترین جواب نسل حاضر حداقل به خوبی جواب نسل پیشین باشد.

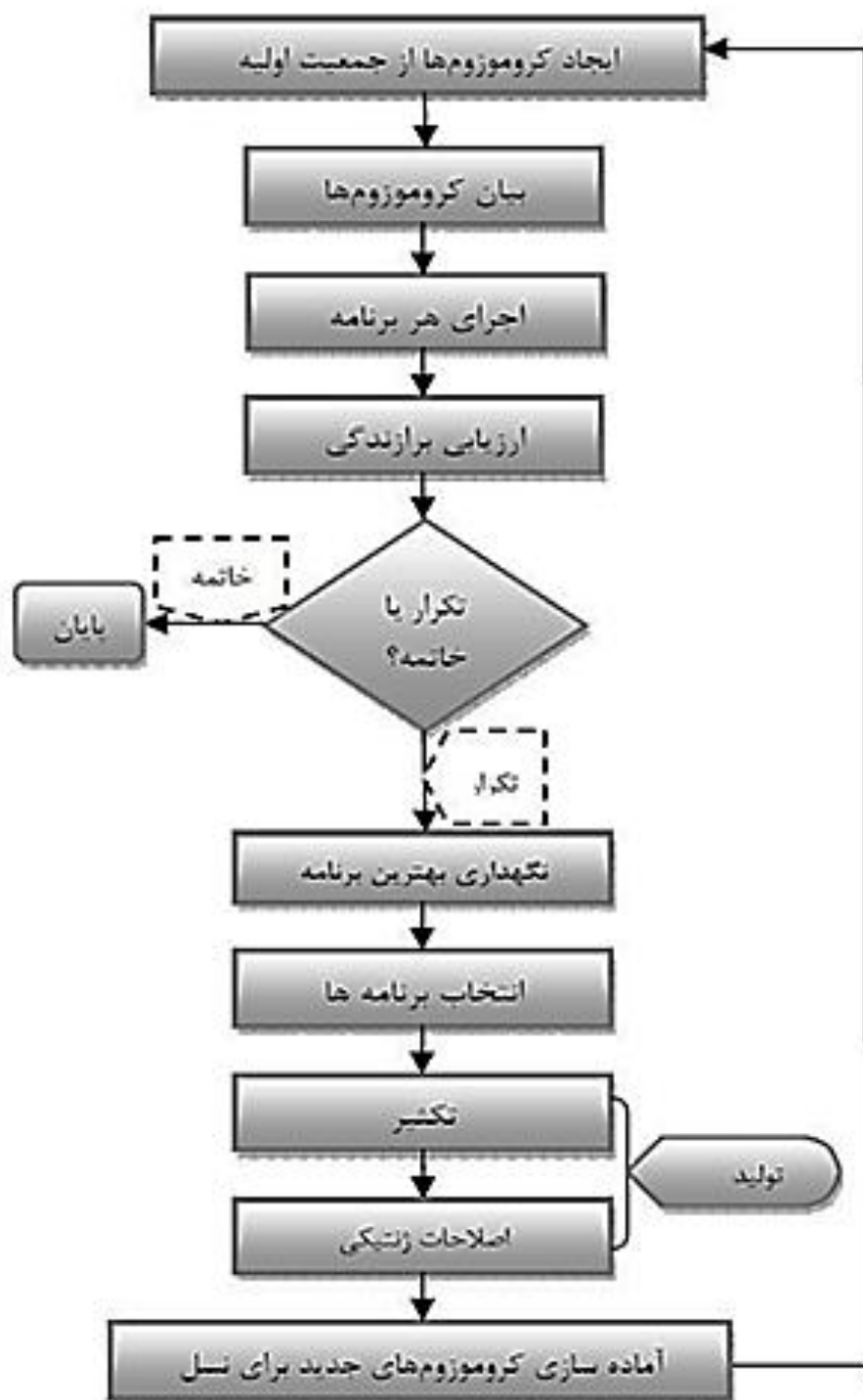
عملگر دیگری که در برنامه ریزی بیان ژن مورد استفاده قرار میگیرد عملگر ترکیب است. در برنامه ریزی بیان ژن علاوه بر ترکیب تک نقطه ای و دونقطه ای، ترکیب ژنی نیز کاربرد دارد. در این نوع از ترکیب، ژن ها به طور کامل ترکیب می شوند. اگرچه در ترکیب تک نقطه ای، فرآیند انتخاب و برش به صورت تصادفی شکل می گیرد، اما بیشترین تمایل به تغییر بخش های ابتدایی کروموزوم ها دیده می شود، لذا کیفیت کروموزوم از ابتدا تا انتها به تدریج کاهش می یابد. از این رو علاوه بر ترکیب تک نقطه ای، ترکیب دونقطه ای و ژنی نیز مورد استفاده قرار می گیرد (خلج، ۱۳۹۰).

در برنامه ریزی بیان ژن دو عملگر ترانهش یا جابه جایی مورد استفاده قرار می گیرند که عبارتند از ترانهش IS و RIS (ریشه IS). عامل IS، ردیفی از المان ها با اندازه های متغیر است که از یک نقطه شروع و از داخل ژنوم انتخاب شده و استخراج می گردد. در ژنوم مکان دیگری با عنوان نقطه درج انتخاب می شود. این نقطه باید داخل قسمت سر ژن باشد، المان های IS از قسمت سر حذف می شوند و ساختار تشکیلات حفظ می شود. در واقع عملگر IS، ترانهش یافت شده در ساختار ژنوم های بیولوژیکی را شبیه سازی می کند. عملگر RIS نیز مشابه با ترانهش IS عمل می کند، با این تفاوت که اولین المان ردیف درج شده، بایستی تابع باشد و نقطه درج نیز می بایست اولین المان از ژن (ریشه) باشد (لوپز و وینرت، ۲۰۰۴).

۳-۶-۲-۶- الگوریتم برنامه ریزی بیان ژن

حل مسئله به کمک GEP از الگوریتم زیر پیروی میکند:

- ۱- فرایند با تولید تصادفی کروموزوم ها از تعداد معینی افراد (جمعیت اولیه) آغاز می گردد.
 - ۳- کروموزوم ها به شکل درختی بیان می گردند.
 - ۴- میزان برازندگی هر فرد تعیین می شود.
 - ۵- در صورت حصول شرایط مطلوب، فرایند تولید نسل متوقف شده و جمعیت موجود به عنوان بهترین نسل انتخاب می شود، در غیر این صورت:
 - ۶- بهترین افراد نسل حاضر بدون تغییر باقی می مانند.
 - ۷- بقیه افراد جمعیت حاضر براساس عملکرد انتخاب می شوند.
 - ۸- بر روی جمعیت انتخاب شده، عملگرهای ژنتیکی و تکثیر صورت می پذیرد تا فرزندان با ویژگی های جدید رشد کنند.
 - ۹- فرزندان جدید نیز در یک چرخه، تحت فرایند های توسعه ای، تکثیر، انتخاب براساس برازندگی یا اصلاحات ژنتیکی قرار می گیرند.
- این فرایند ها بر روی تعداد معینی از نسل ها تکرار می گردد تا به پاسخ بهینه دست پیدا کنیم.
- در شکل (۳-۲۰) الگوریتم برنامه ریزی بیان ژن نشان داده شده است:



شکل (۳-۲۰) - الگوریتم برنامه ریزی بیان ژن (فریرا، ۲۰۰۱)

۳-۷- شاخص های آماری سنجش مدل ها

به منظور ارزیابی و بررسی کارایی روش استفاده شده برای پیش بینی TDS، در این تحقیق از معیارهای، ضریب همبستگی (R^2)، مجذور میانگین مربعات خطا (RMSE) و میانگین قدر مطلق خطاها (MAE) استفاده شده است.

$$R = \left[\frac{\sum_{i=1}^N (TDS_{io} - \overline{TDS_o})(TDS_{ie} - \overline{TDS_e})}{\sqrt{\sum_{i=1}^N (TDS_{io} - \overline{TDS_o})^2 \sum_{i=1}^N (TDS_{ie} - \overline{TDS_e})^2}} \right] \quad (۹-۳)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (TDS_{io} - TDS_{ie})^2} \quad (۱۰-۳)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |TDS_{io} - TDS_{ie}| \quad (۱۱-۳)$$

در روابط بالا $\overline{TDS_o}$ و $\overline{TDS_e}$ میانگین شاخص کیفی کل مواد محلول برآوردی و مشاهده ای و TDS_{io} و TDS_{ie} مقادیر مشاهده ای و برآوردی شاخص کل مواد محلول و N تعداد کل نمونه میباشند. و برای نسبت جذب سدیم SAR نیز از این رابطه ریاضی ها استفاده شد و بجای TDS از SAR استفاده شد.

فصل چهارم

بحث و نتایج

۴-۱- مقدمه

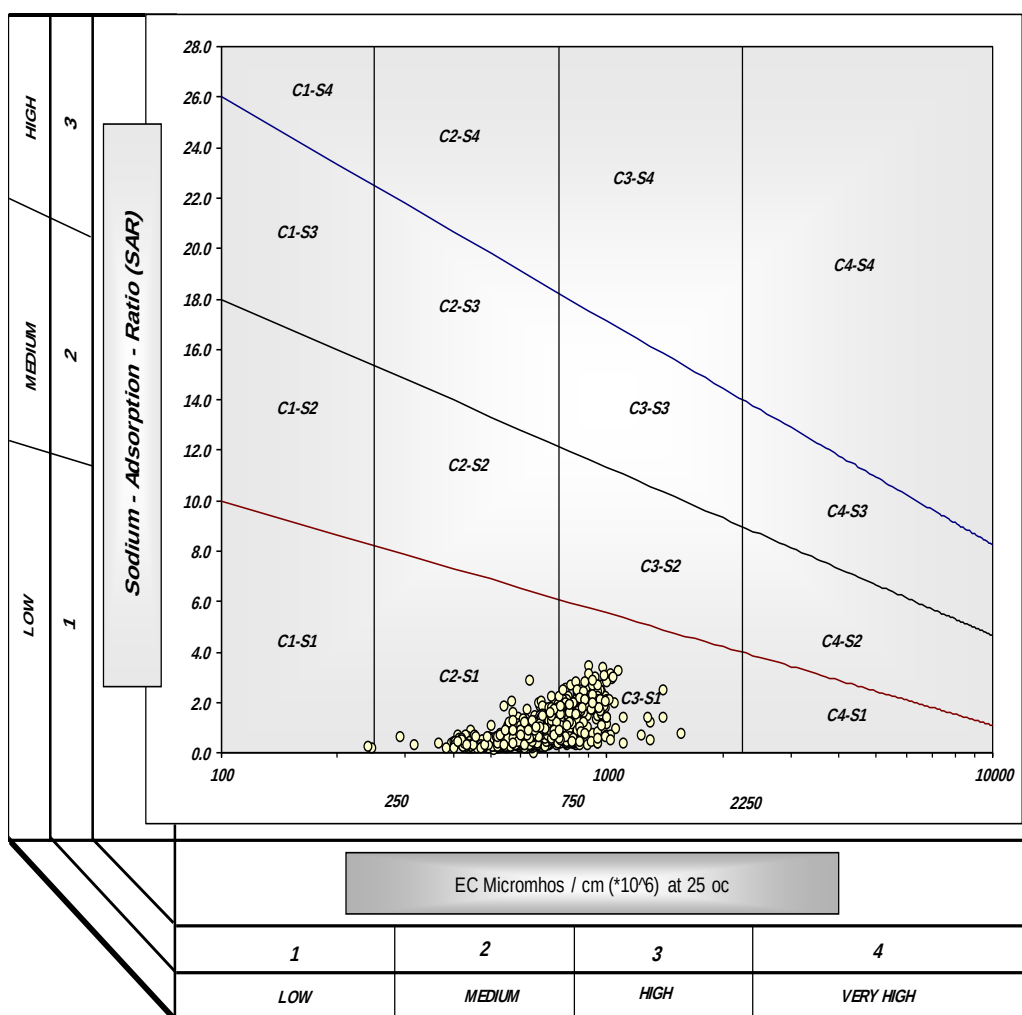
در این فصل ابتدا کیفیت آب دو رودخانه تجن و بابلرود از جنبه مصرف کشاورزی با استفاده از نمودار ویلکاکس مورد بررسی قرار گرفته است و سپس پارامترهای شاخص کیفی کل مواد محلول و نسبت جذب سدیم رودخانه تجن مورد بررسی قرار گرفت و بعد از آن به بررسی دو پارامتر بالا در رودخانه بابلرود پرداخته شد. که با توجه به نمودارهای درختی بدست آمده از نرم افزار معادلات ریاضی برآورد هر دو پارامتر با ضریب همبستگی بالا محاسبه شد.

۴-۲- ارزیابی کیفی آب از جنبه مصرف کشاورزی

طبقه بندی کیفیت آب برای استفاده در آبیاری توسط نمودار ویلکاکس بر مبنای نسبت جذب سدیم و هدایت الکتریکی صورت می گیرد.

۴-۲-۱- ارزیابی کیفی رودخانه تجن

با توجه به EC و SAR رودخانه تجن نمودار ویلکاکس رودخانه به صورت شکل (۴-۱) می باشد.

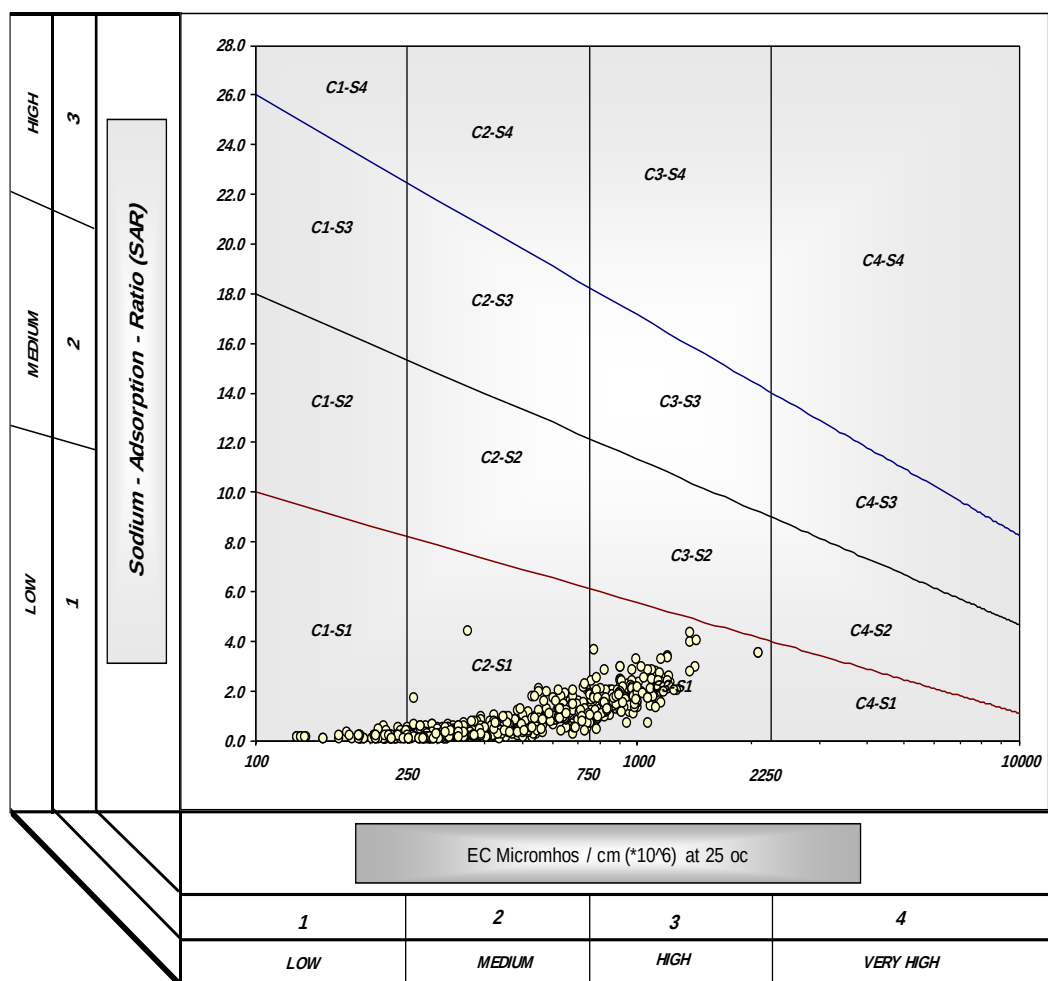


شکل (۴-۱)- نمودار ویلکاکس رودخانه تجن

با توجه به نمودار ویلکاکس رودخانه تجن بطور متوسط آب این رودخانه از نظر کشاورزی در وضعیت خوب (کلاس C2S1، C3S1) قرار دارد. لذا استفاده از آب این رودخانه برای مصارف کشاورزی هیچ مانعی ندارد.

۴-۲-۲- ارزیابی کیفی رودخانه بابلرود

با توجه به هدایت الکتریکی EC و نسبت جذب سدیم SAR رودخانه بابلرود نمودار ویلکاکس رودخانه به صورت شکل (۴-۲) می باشد.



شکل (۴-۲) - نمودار ویلکاکس رودخانه بابلرود

با توجه به نمودار ویلکاکس رودخانه بابلرود بطور متوسط آب این رودخانه از نظر کشاورزی در وضعیت خوب (کلاس C1S1, C2S1, C3S1) قرار دارد. لذا استفاده از آب این رودخانه برای مصارف کشاورزی هیچ مانعی ندارد.

۴-۳- رودخانه تجن

به منظور برآورد TDS رودخانه تجن داده های ۱۲ متغیر کیفی ماهانه آب استفاده شد. این متغیرها شامل هدایت الکتریکی آب، اسیدیت، سدیم، کلسیم، پتاسیم، نسبت جذب سدیم، منیزیم، کلر،

سولفات، بیکربنات و دبی به عنوان متغیر مستقل و TDS به عنوان متغیر وابسته در نظر گرفته شد. و برای برآورد SAR رودخانه تجن نیز از تمامی متغیرهای بالا استفاده شد با این تفاوت که پارامتر

Q	EC	pH	HCO ₃	Cl	SO ₄	Ca	Mg	Na	K	SAR	TDS
---	----	----	------------------	----	-----------------	----	----	----	---	-----	-----

نسبت جذب سدیم به عنوان متغیر وابسته و شاخص کیفی کل مواد محلول به متغیرهای مستقل اضافه شده است.

۴-۳-۱- برآورد شاخص کیفی کل مواد محلول

۴-۳-۱-۱- همبستگی و رگرسیون

با استفاده از نرم افزار SPSS23، همبستگی بین داده ها مورد سنجش قرار گرفت که نتایج آن در جدول (۴-۱) مشاهده می شود. و سپس به تحلیل رگرسیون بین متغیرهای مستقل و وابسته پرداخته شد که به این منظور از روش گام به گام (Stepwise Regression) استفاده شد. در روش گام به گام ۵ الگوی ورودی پیشنهاد شد که در جدول (۴-۲) ارائه شده است.

جدول (۴-۲) - الگو های ورودی برآورد TDS استخراج شده از SPSS23

الگو	پارامتر ورودی
۱	EC
۲	EC,SO ₄
۳	EC,SO ₄ ,SAR
۴	EC,SO ₄ ,SAR,HCO ₃
۵	EC,SO ₄ ,SAR,HCO ₃ ,Cl

۴-۳-۱-۲- برنامه ریزی بیان ژن

در پژوهش حاضر از روش برنامه ریزی بیان ژن برای برآورد پارامتر های کیفیت آب بهره گرفته شد. به این منظور ۵ الگوی ورودی متفاوت تعریف و اجرا شد. با هدف یکسان سازی فضای مقایسه و صحت اجرای مدل، پارامترهای GEP به طور همسان تعیین شدند که به طور خلاصه در جدول (۴-۳) ارائه گردیده است. در جدول (۴-۴) ضرایب همبستگی، جذر میانگین مربعات خطا و متوسط خطای میانگین حاصل از اجرای ترکیب های مذکور در هر مرحله آموزش و آزمون، ارائه گردیده است. مدل مطلوب و بهینه مدلی است که با دارا بودن بیشترین ضریب همبستگی، کمترین میزان خطا را به خود اختصاص دهد.

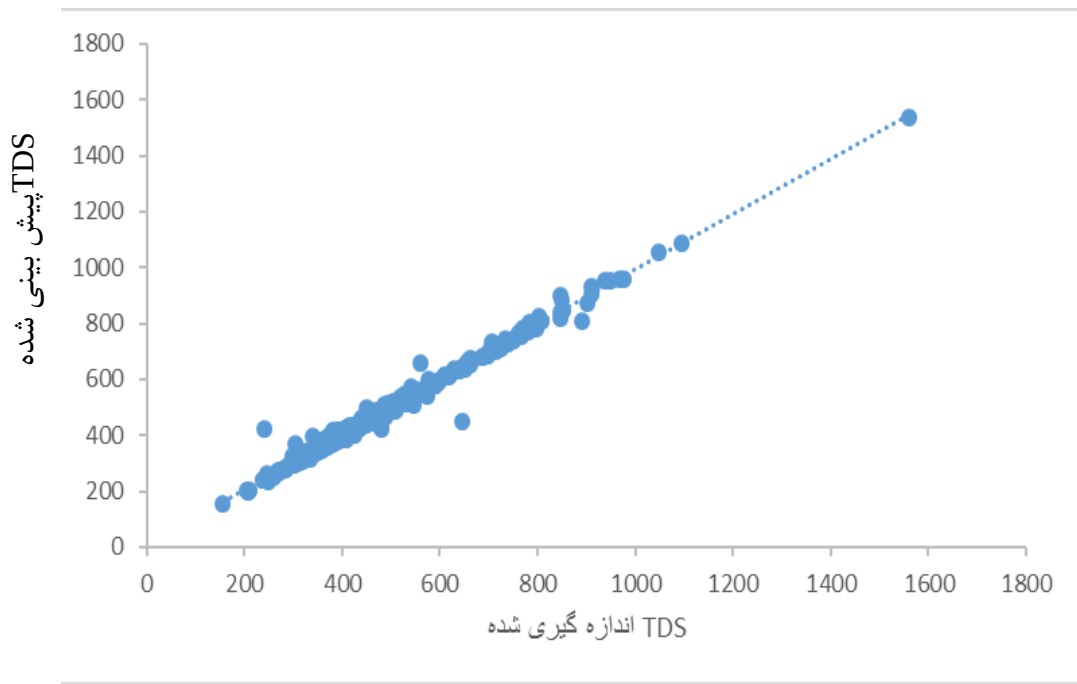
جدول (۳-۴) - مقادیر پارامترهای GEP در برآورد TDS و SAR رودخانه های تجن و بابلرود

تنظیمات کلی			عملگرهای ژنتیکی
تعداد کروموزوم ها	۳۰	نرخ جهش	۰/۴۴
اندازه راس	۸	نرخ وارون سازی	۰/۱
تعداد ژنها در هر کروموزوم	۳	نرخ ترانهش درج متوالی	۰/۱
تعداد جمعیت تولیدی	۱۰۰۰	نرخ ترانهش ژن	۰/۱
تابع پیوند	جمع(+)	نرخ ترکیب تک نقطه ای	۰/۳
معیار خطا	RMSE	نرخ ترکیب ژن	۰/۱

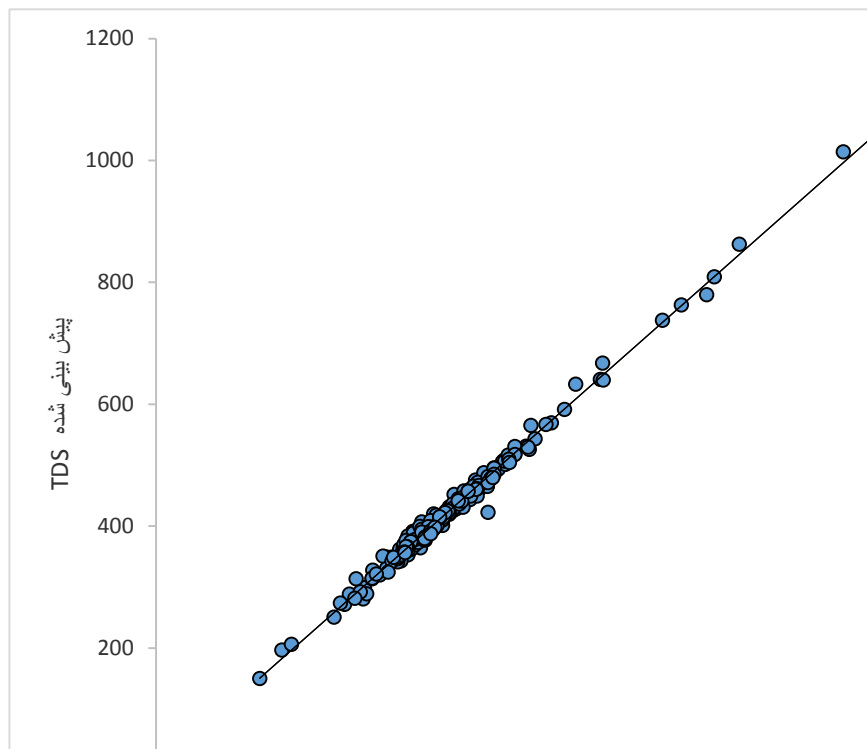
جدول (۴-۴) - نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی برآورد TDS رودخانه تجن

الگو	مرحله اجرا	R^2	RMSE	MAE
1	آموزش	0.9917	10.06	7.32
	آزمون	0.9808	17.17	6.78
2	آموزش	0.9924	9.7	7.06
	آزمون	0.9806	17.3	6.99
3	آموزش	0.9921	9.82	7.14
	آزمون	0.9807	17.16	6.79
4	آموزش	0.9926	9.5	6.67
	آزمون	0.9807	17.2	6.54
5	آموزش	0.9927	9.5	6.83
	آزمون	0.9808	17.06	6.16

مقایسه نتایج جدول فوق نشان داد که الگو ۵ با پارامترهای ورودی Cl و HCO_3 ، SAR، SO_4 ، EC با دارا بودن بیشترین ضریب همبستگی برابر ۰/۹۹۲۷ در آموزش و ۰/۹۸۰۸ در آزمون و کمترین خطا مدل بهینه می‌باشد. نتایج حاکی از آن است که الگوی مورد نظر با دارا بودن ضریب همبستگی ۰/۹۸۰۸ در بخش آزمون، از توانایی بالایی در برآورد TDS برخوردار است و به عبارتی نتایج قابل اطمینانی را در پیش بینی شاخص کیفی کل مواد محلول ارائه می‌کند در شکل های (۳-۴) و (۴-۴) نمودارهای همبستگی پارامترهای اندازه گیری شده و پیش بینی شده در هر دو مرحله آموزش و آزمون و همچنین ساختار ORF ژن های حاصل از اجرای الگوی پنجم نشان داده شده است.



شکل (۳-۴) نمودار اجرای الگو ۵، برآورد TDS رودخانه تجن در مرحله آموزش



شکل (۴-۴) نمودار اجرای الگو ۵، برآورد TDS رودخانه تجن در مرحله آزمون

ساختار ORF زن های مدل برتر به زبان کاروا در زیر نشان داده شده است:

$$*Cos.d\backslash.Sin.X^3.Exp.Atan.Cos.c\backslash.d^2.c\backslash.d^2.d^3.d^3.d\backslash.d\backslash.d^2$$

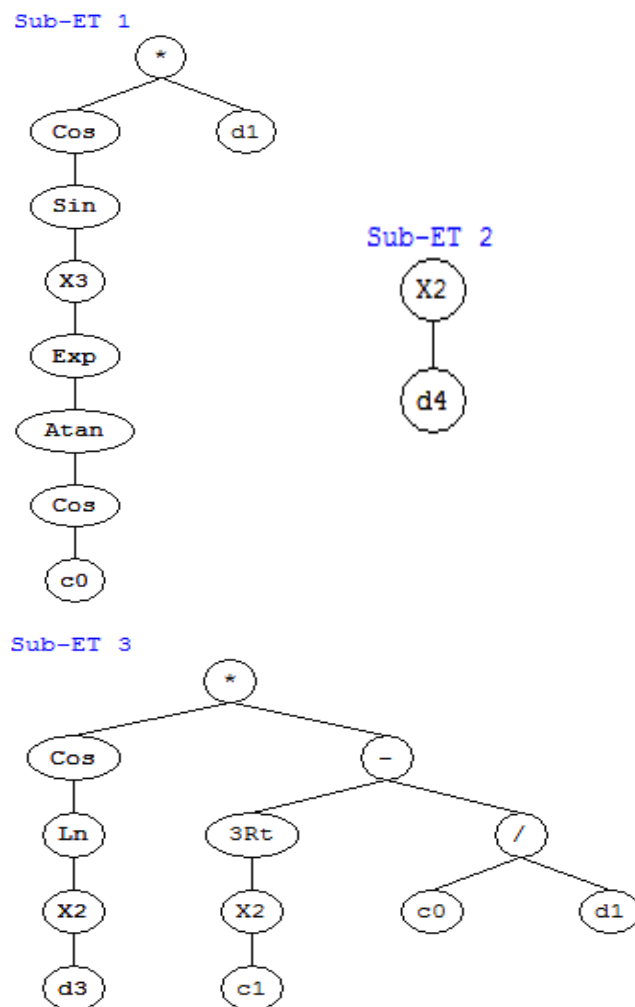
+

$$X^2.d^4.-.d\backslash.d^4./.^3Rt.*.c\backslash.d\backslash.d^2.c\backslash.c\backslash.c\backslash.d\backslash.c\backslash$$

+

$$*Cos.-.Ln.^3Rt./X^2.X^2.c\backslash.d\backslash.d^3.c\backslash.c\backslash.c\backslash.d^4.d\backslash.d\backslash$$

در ادامه بیان درختی الگوی فوق نشان داده شده است. که با استفاده از جدول (۵-۴) رابطه ریاضی برآورد TDS رودخانه تجن را طبق رابطه ریاضی (۱-۴) بدست آوردیم.



شکل (۵-۴) نمودار درختی الگوی ۵ برآورد TDS رودخانه تجن با استفاده از برنامه ریزی بیان ژن

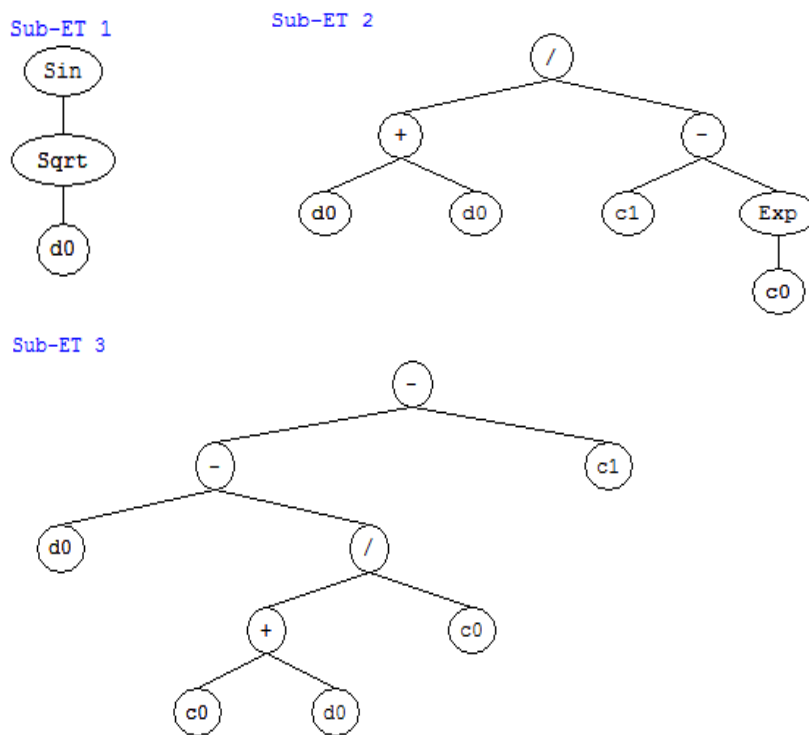
جدول (۵-۴) - مقادیر ثابت در بیان درختی الگوی ۵، برآورد TDS رودخانه تجن

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
-۷/۳۶۹	-۹/۵۵۴	۴/۶۶۵	-۹/۵۵۴	۴/۱۳۴	۰/۹۱۰

$$TDS = (0/573117 - EC) + (SO_4)^2 + \left[\cos(\ln(SAR^2)) \left(3/787 + \frac{9/5539}{EC} \right) \right]$$

رابطه ریاضی (۱-۴) برآورد شاخص کیفی کل مواد محلول رودخانه تجن

در رابطه ریاضی فوق $SAR_{SO_4,EC}$ شرکت دارند تا بالاترین ضریب همبستگی و کمترین میزان خطا بدست آید. ولی با توجه به اینکه هدف دست یافتن به رابطه ریاضی‌ای است که بتواند با کمترین پارامتر ورودی و بالاترین مقدار ضریب همبستگی و نیز کمترین میزان خطا TDS را برآورد کند بنابراین از الگو ورودی ۱ استفاده شد که با توجه به جدول (۴-۴) تفاوت چندانی در R^2 و میزان خطا با الگوی ۵ نداشت و پس از اجرای مدل نتایج زیر حاصل شد:



شکل (۴-۶) نمودار درختی الگو ۱ برآورد TDS رودخانه تجن با استفاده از برنامه ریزی بیان ژن

با توجه به شکل (۴-۶) نمودار درختی رودخانه تجن و با استفاده از جدول (۴-۶) مقادیر ثابت برنامه ریزی بیان ژن رابطه ریاضی نهایی برآورد TDS رودخانه تجن طبق رابطه ریاضی (۴-۲) بدست آمده است.

جدول (۴-۶) - مقادیر ثابت در بیان درختی الگوی ۱، برآورد TDS رودخانه تجن

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
-۳/۴۷۶	۳/۱۲۸	۸/۵۳۰	۴/۰۸۳۷	۰/۴۶۱۱	-۸/۷۷۹

$$TDS = \sin \sqrt{EC} + \frac{2EC}{-50.834} + EC - \left(\frac{3.128 + EC}{3.128} \right) + 3.476$$

رابطه ریاضی (۴-۲) برآورد شاخص کیفی کل مواد محلول الگوی ۱ رودخانه تجن

رابطه ریاضی پیشنهادی نشان می‌دهد که TDS با EC رابطه خوبی دارد و این در تطابق با نتایج سایر محققین مانند جلیلیان و امینی (۱۳۸۶)، بتوشا و همکاران (۲۰۰۹)، تیرامالین و کوریان (۲۰۰۹) و همکاران (۲۰۰۹)، کرمی و هوشمند (۱۳۹۲)، دودکانلوی میلان و همکاران (۱۳۹۰) و کرمی و گلابی (۱۳۹۰) می‌باشد.

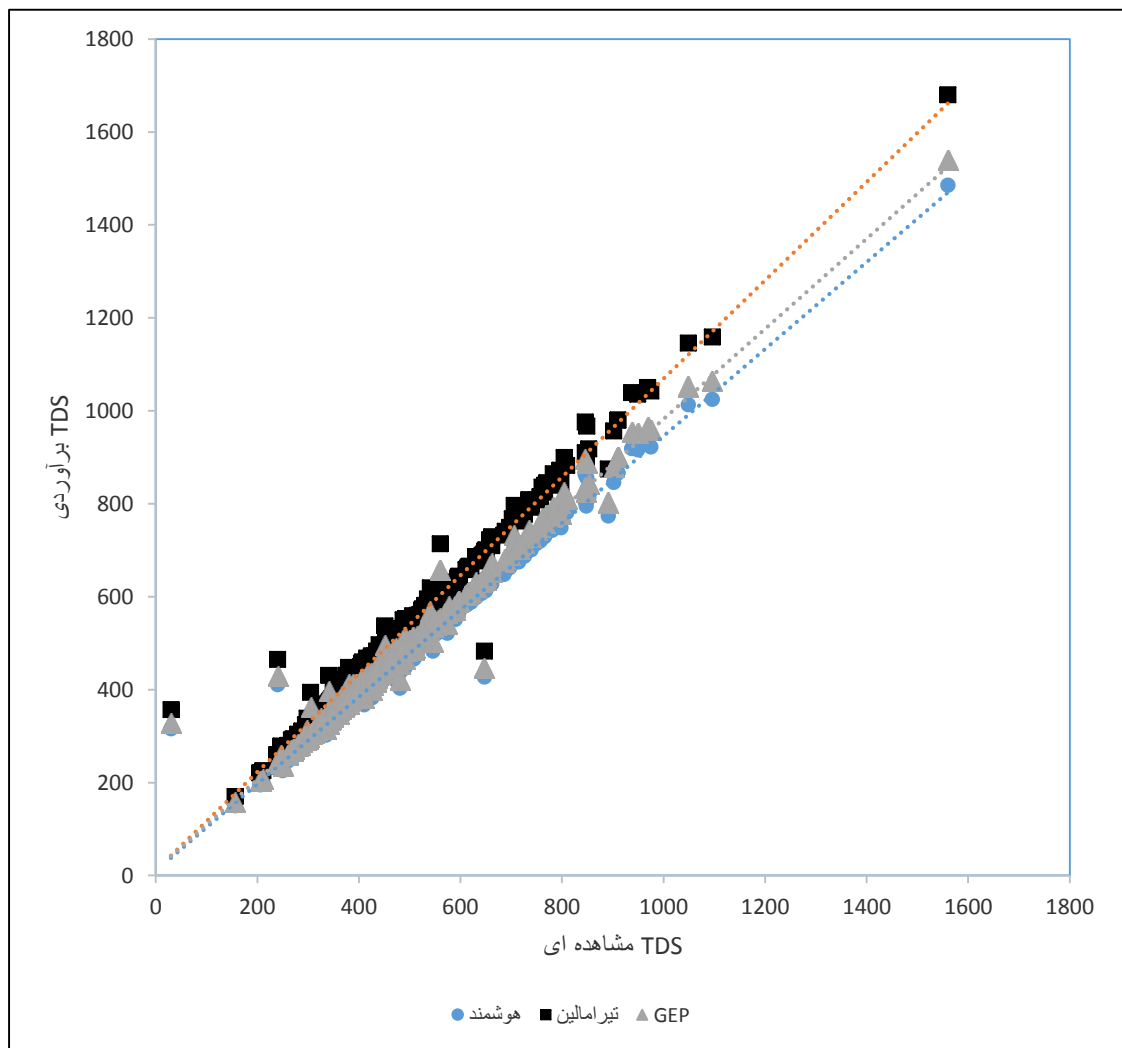
با توجه به روابط بدست آمده از پژوهش‌های کرمی و هوشمند و نیز تیرامالین و کوریان (۲۰۰۹) و همکاران در ذیل به مقایسه این دو رابطه با رابطه (۴-۲) بدست آمده از الگوی ۱ مدل برنامه ریزی بیان ژن پرداخته شده است.

$$TDS = (0.55 \text{ to } 0.7)EC$$

تیرامالین و کوریان ۲۰۰۹

$$TDS = 0.619(EC)$$

کرمی و هوشمند ۱۳۹۲



شکل (۷-۴) نمودار مقایسه برآورد TDS به سه روش

جدول (۷-۴) مقایسه شاخص های آماری سه روش برآورد TDS

RMSE	R^2	
41.92829	0.980778	تیرمالین و کوریان ۲۰۰۹
26.17054	0.980778	کرمی و هوشمند ۱۳۹۲
17.17337	0.980814	برنامه ریزی بیان ژن

با توجه به نمودار (۷-۴) و جدول (۷-۴) می‌توان نتیجه گرفت که روش برنامه ریزی بیان ژن با ارائه رابطه ریاضی پیشنهادی (۲-۴) به خوبی توانسته با کمترین میزان RMSE و بیشترین میزان همبستگی نسبت به دو رابطه دیگر شاخص کیفی کل مواد محلول را برآورد کند.

۲-۲-۴- برآورد نسبت جذب سدیم SAR

۱-۲-۲-۴- همبستگی و رگرسیون

همبستگی بین داده های رودخانه تجن را قبلا در جدول (۱-۴) نشان دادیم و تحلیل رگرسیون بین متغیر های مستقل و وابسته را نیز با نرم افزار SPSS23 انجام دادیم. که به این منظور از روش گام به گام (Stepwise Regression) استفاده کردیم. در روش گام به گام ۴ الگوی ورودی پیشنهاد شد که در جدول (۸-۴) ارائه شده است.

جدول (۸-۴)- الگو های ورودی برآورد SAR استخراج شده از SPSS23

الگو	پارامتر ورودی
۱	Na
۲	Na,EC
۳	Na, EC,Cl
۴	Na,EC,Cl,Q

۴-۲-۲-برنامه ریزی بیان ژن

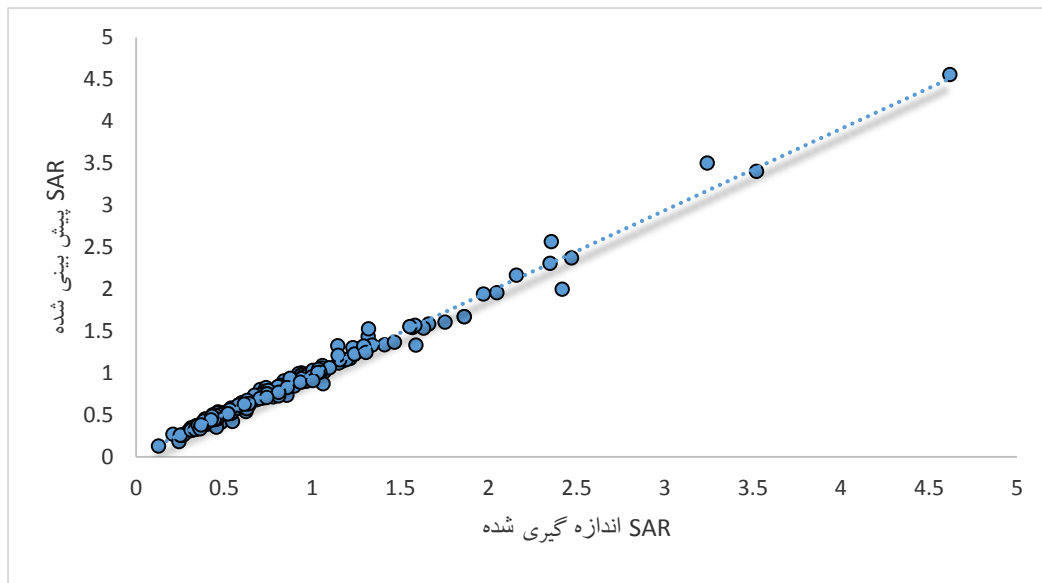
الگوهای ورودی جدول فوق را در نرم افزار GeneXproTools وارد کردیم و در جدول (۹-۴) ضرایب همبستگی، جذر میانگین مربعات خطا و متوسط خطای میانگین حاصل از اجرای ترکیب های مذکور در هر مرحله آموزش و آزمون، ارائه گردیده است. مدل مطلوب و بهینه مدلی است که با دارا بودن بیشترین ضریب همبستگی، کمترین میزان خطا را به خود اختصاص دهد.

جدول (۹-۴)- نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی، برآورد SAR رودخانه تجن

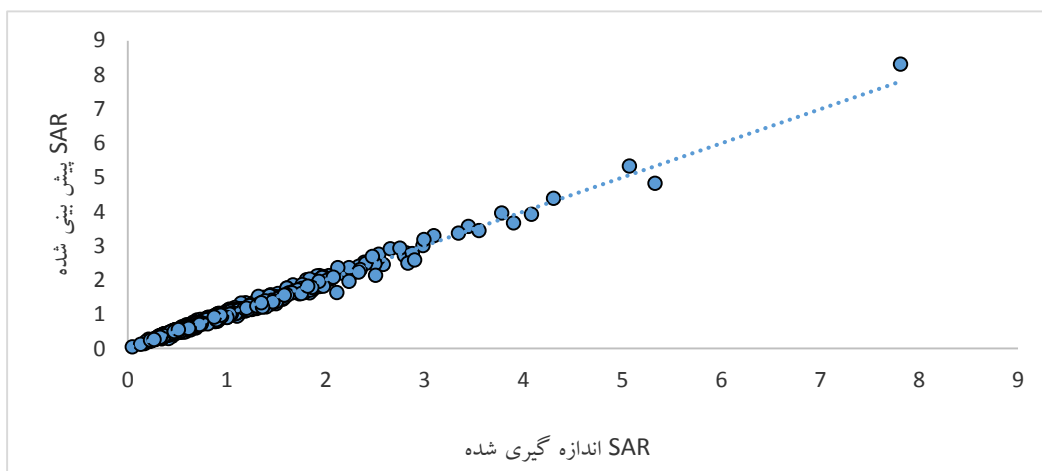
الگو	مرحله اجرا	R^2	RMSE	MAE
1	آموزش	۰/۹۷۲	۰/۱۰۱	۰/۰۶۲
	آزمون	۰/۹۶۴	۰/۱۳۱	۰/۰۷۷
2	آموزش	۰/۹۸۷	۰/۰۶۵	۰/۰۴۰
	آزمون	۰/۹۸۹	۰/۰۶۹	۰/۰۴۳
3	آموزش	۰/۹۷۷	۰/۰۸۷	۰/۰۵۵
	آزمون	۰/۹۸۴	۰/۰۸۵	۰/۰۵۸
4	آموزش	۰/۹۷۶	۰/۰۹۰	۰/۰۵۶
	آزمون	۰/۹۸۳	۰/۰۸۶	۰/۰۵۹

مقایسه نتایج جدول فوق نشان داد که الگو ۲ با پارامترهای ورودی EC_{Na} با دارا بودن بیشترین ضریب همبستگی برابر ۰/۹۸۷ در آموزش و ۰/۹۸۹ در آزمون و کمترین خطا و همچنین با داشتن پارامتر وردی کم مدل بهینه می باشد. نتایج حاکی از آن است که الگوی مورد نظر با دارا بودن ضریب

همبستگی بیشتر، توانایی بالایی در برآورد SAR دارد و به عبارتی نتایج قابل اطمینانی را در پیش بینی نسبت جذب سدیم ارائه می‌کند در شکل های (۴-۸) و (۴-۹) نمودارهای همبستگی پارامتر های اندازه گیری شده و پیش بینی شده در هر دو مرحله آموزش و آزمون و همچنین ساختار ORF ژن های حاصل از اجرای الگوی ۲ نشان داده شده است.



شکل (۴-۸) نمودار اجرای الگو ۲ برآورد SAR رودخانه تجن در مرحله آموزش



شکل (۴-۹) نمودار اجرای الگو ۲ برآورد SAR رودخانه تجن در مرحله آزمون

ساختار ORF ژن های مدل برتر به زبان کاروا در زیر نشان داده شده است:

$d \cdot X_3 \cdot + \cdot - \cdot \text{Exp} \cdot X_3 \cdot + \cdot d \cdot c \cdot d \cdot d \cdot d \cdot d \cdot c \cdot d \cdot$

+

$\cdot \text{Exp} \cdot X_3 \cdot + \cdot d \cdot c \cdot c \cdot c \cdot d \cdot d \cdot d \cdot d \cdot d \cdot c \cdot d \cdot$

+

$\cdot / \cdot \cdot \cdot c \cdot d \cdot , 3Rt \cdot 3Rt \cdot d \cdot c \cdot c \cdot d \cdot d \cdot d \cdot c \cdot d \cdot$

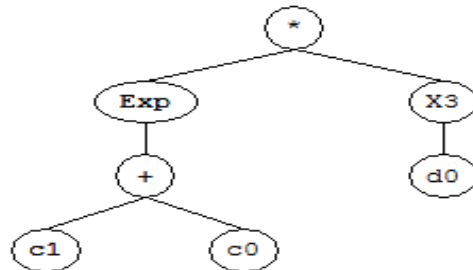
در ادامه بیان درختی الگوی فوق نشان داده شده است. که با استفاده از جدول (۴-۱۰) رابطه ریاضی

برآورد SAR رودخانه تجن را طبق رابطه ریاضی (۴-۳) بدست آوردیم.

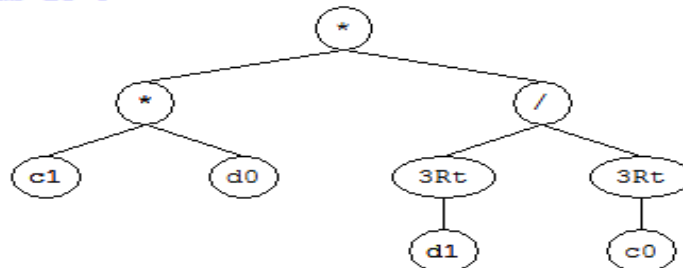
Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل (۴-۱۰) نمودار درختی برآورد SAR رودخانه تجن با استفاده از برنامه ریزی بیان ژن

جدول (۴-۱۰) - مقادیر ثابت در بیان درختی الگوی ۲ برآورد SAR رودخانه تجن

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
-۰/۰۸۲	۶/۹۰۶	-۴/۹۴۴	-۳/۰۶۴	-۴/۷۴۲	-۳/۰۶۴

$$SAR = \left[NA + [0.000333(NA)^3] + \left[(-0.081665NA) \left(\frac{EC^{\frac{1}{3}}}{1.9043} \right) \right] \right]$$

رابطه ریاضی برآورد نسبت جذب سدیم رودخانه تجن (۴-۳)

رابطه فوق طبق الگو ۲ که کاملاً مطابق هدف ما مبنی بر کمترین پارامتر ورودی ، بالاترین ضریب

همبستگی و کمترین میزان خطا بود بدست آمد و مطابق با رابطه ریاضی محاسبه SAR

($SAR = \frac{NA}{\sqrt{\frac{Ca+Mg}{2}}}$) که بیشترین تاثیر را در این رابطه سدیم دارا می باشد و استفاده از EC در رابطه

که یک پارامتر زود یافت می باشد توانست صرفه جویی وقت و هزینه را در برآورد SAR محقق سازد.

۴-۳- رودخانه بابلرود

به منظور برآورد TDS رودخانه بابلرود داده های ۱۲ متغیر کیفی ماهانه آب استفاده شد. این متغیرها شامل هدایت الکتریکی آب، اسیدیته، سدیم، کلسیم، پتاسیم، نسبت جذب سدیم، منیزیم، کلر، سولفات، بیکربنات و دی به عنوان متغیر مستقل و TDS به عنوان متغیر وابسته در نظر گرفته شد. و برای برآورد SAR رودخانه بابلرود نیز از تمامی متغیرهای بالا استفاده شد با این تفاوت که پارامتر نسبت جذب سدیم به عنوان متغیر وابسته و شاخص کیفی کل مواد محلول به متغیرهای مستقل اضافه شده است.

۴-۳-۱- برآورد شاخص کیفی کل مواد محلول TDS

۴-۳-۱-۱- همبستگی و رگرسیون

با استفاده از نرم افزار SPSS23، همبستگی بین داده ها را مورد سنجش قرار دادیم که نتایج آن در جدول (۴-۱۱) مشاهده می شود. و سپس به تحلیل رگرسیون بین متغیرهای مستقل و وابسته پرداختیم که به این منظور از روش گام به گام (Stepwise Regression) استفاده کردیم. در روش گام به گام ۸ الگوی ورودی پیشنهاد شد که در جدول (۴-۱۲) ارائه شده است.

جدول (۴-۱۱) ضریب همبستگی متغیرهای مورد مطالعه رودخانه بابلرود

	Q	TDS	EC	pH	HCO ₃	Cl	SO ₄	Ca	Mg	Na	K	SAR
Q	1											
TDS	-0.037	1										
EC	-0.038	0.998**	1									
pH	0.043	-0.29**	-0.28**	1								
HCO ₃	-0.037	0.853**	0.857**	-0.34**	1							
Cl	-0.034	0.919**	0.916**	-0.24**	0.661**	1						
SO ₄	-0.024	0.698**	0.696**	-0.18**	0.426**	0.556**	1					
Ca	-0.041	0.853**	0.856**	-0.34**	0.825**	0.729**	0.575**	1				
Mg	-0.008	0.798**	0.801**	-0.24**	0.754**	0.661**	0.610**	0.531**	1			
Na	-0.044	0.932**	0.929**	-0.27**	0.706**	0.939**	0.643**	0.691**	0.679**	1		
K	-0.003	0.393**	0.394**	-0.058	0.317**	0.350**	0.284**	0.352**	0.310**	0.341**	1	
SAR	-0.027	0.882**	0.879**	-0.21**	0.664**	0.900**	0.592**	0.619**	0.633**	0.984**	0.325**	1

جدول (۴-۱۲) - الگو های ورودی برآورد TDS بابلرود استخراج شده از SPSS

الگو	پارامتر ورودی
۱	EC
۲	EC, Na
۳	EC, Na, SAR
۴	EC, Na, SAR, pH
۵	EC, Na, SAR, pH, SO ₄
۶	EC, Na, SAR, pH, SO ₄ , Cl
۷	EC, Na, SAR, pH, SO ₄ , Cl, HCO ₃
۸	EC, Na, SAR, SO ₄ , Cl, HCO ₃

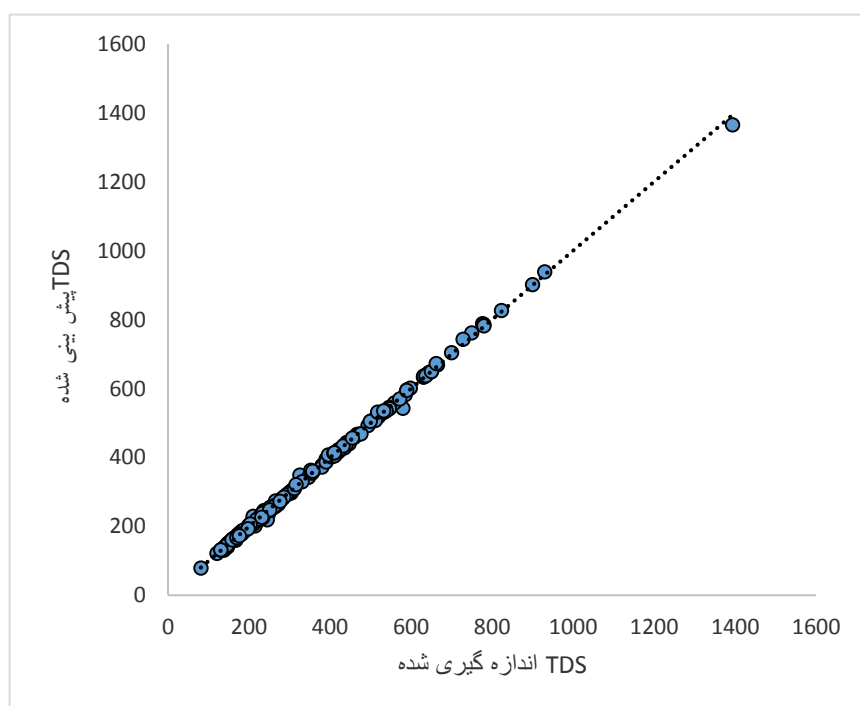
۴-۳-۱-۲- برنامه ریزی بیان ژن

الگو های ورودی جدول (۴-۱۲) را در نرم افزار GeneXproTools وارد کردیم و در جدول (۴-۱۳) ضرایب همبستگی، جذر میانگین مربعات خطا و متوسط خطای میانگین حاصل از اجرای ترکیب های مذکور در هر مرحله آموزش و آزمون، ارائه گردیده است. مدل مطلوب و بهینه مدلی است که با دارا بودن بیشترین ضریب همبستگی، کمترین میزان خطا را به خود اختصاص دهد.

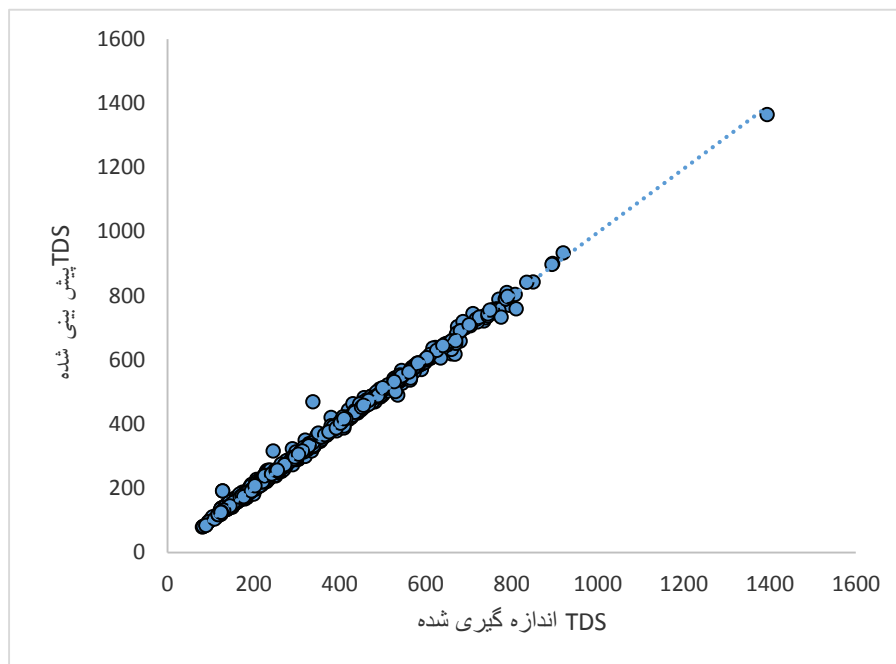
جدول (۴-۱۳) - نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی برآورد TDS رودخانه بابلرود

الگو	مرحله اجرا	R^2	RMSE	MAE
۱	آموزش	۰/۹۹۸	۶/۴۶۲	۴/۱۸۷
	آزمون	۰/۹۹۶	۱۰/۶۸۷	۵/۸۲۴
۲	آموزش	۰/۹۹۹	۵/۸۸۷	۳/۵۴۲
	آزمون	۰/۹۹۶	۱۰/۵۴۳	۵/۴۴۵
۳	آموزش	۰/۹۹۸	۶/۴۰۷	۴/۲۴۹
	آزمون	۰/۹۹۶	۱۰/۶۹۸	۵/۹۹۰
۴	آموزش	۰/۹۹۹	۵/۸۳۰	۳/۵۶۶
	آزمون	۰/۹۹۶	۱۰/۴۸۹	۵/۴۶۳
۵	آموزش	۰/۹۹۸	۶/۶۶۰	۴/۱۶۰
	آزمون	۰/۹۹۶	۱۱/۱۰۳	۶/۲۱
۶	آموزش	۰/۹۹۹	۵/۸۰۲	۳/۵۸۳
	آزمون	۰/۹۹۶	۱۰/۵۸۱	۵/۴۲۹
۷	آموزش	۰/۹۹۹	۵/۵۸۴	۳/۴۹۶
	آزمون	۰/۹۹۶	۱۰/۵۶۲	۵/۶۴۷
۸	آموزش	۰/۹۹۹	۶/۰۷۵	۳/۸۶۴
	آزمون	۰/۹۹۷	۱۰/۲۸۷	۵/۴۵۰

مقایسه نتایج جدول فوق نشان داد که الگو ۸ با پارامترهای ورودی EC , Na , SAR , SO_4 , Cl , HCO_3 با دارا بودن بیشترین ضریب همبستگی برابر ۰/۹۹۹ در آموزش و ۰/۹۹۷ در آزمون و کمترین خطا مدل بهینه می‌باشد. نتایج حاکی از آن است که الگوی مورد نظر با دارا بودن ضریب همبستگی بالا از توانایی خوبی نسبت به برآورد TDS برخوردار است و به عبارتی نتایج قابل اطمینانی را در پیش بینی شاخص کیفی کل مواد محلول ارائه می‌کند در شکل های (۴-۱۱) و (۴-۱۲) نمودارهای همبستگی پارامترهای اندازه گیری شده و پیش بینی شده در هر دو مرحله آموزش و آزمون و همچنین ساختار ORF ژن های حاصل از اجرای الگوی ۸ نشان داده شده است.



شکل (۴-۱۱) نمودار اجرای الگو ۸ برآورد TDS رودخانه بابلرود در مرحله آموزش



شکل (۴-۱۲) نمودار اجرای الگو ۸ برآورد TDS رودخانه بابلرود در مرحله آزمون

ساختار ORF ژن های مدل برتر به زبان کاروا در زیر نشان داده شده است:

$$-.-.-Cos.d2.d4./+.d1.c1.c1.d2.d0.d3.c1.d0.d3$$

+

$$-.-.+d2.+Cos.3Rt.Atan.d4.d3.c1.d4.c0.d0.d5.d0.d5$$

+

$$.+d2.-.-.*.X2.Cos.d1.c0.d3.d4.d2.d5.d3.d2.c0$$

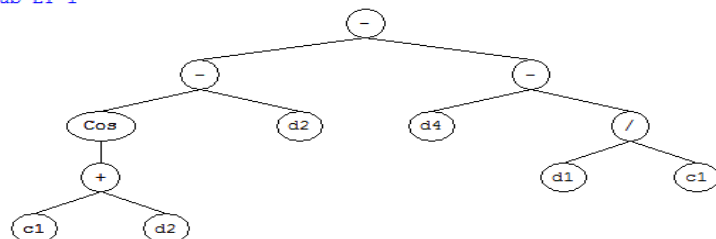
در ادامه بیان درختی الگوی فوق نشان داده شده است. که با استفاده از جدول (۴-۱۴) مقادیر ثابت

بیان درختی الگوی ۸، رابطه ریاضی برآورد TDS رودخانه بابلرود طبق رابطه ریاضی (۴-۴) بدست آمد.

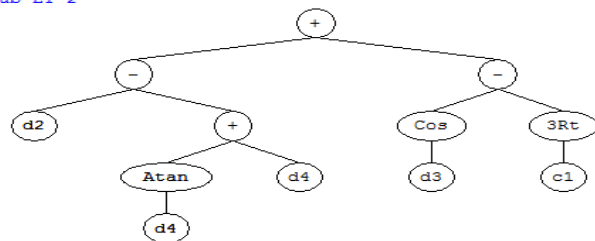
جدول (۴-۱۴) - مقادیر ثابت در بیان درختی الگوی ۸ برآورد TDS رودخانه بابلرود

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
۸/۴۱۲	۸/۳۸۳	۹/۶۳۱	۷/۱۱۵	-۲/۷۰۵	۹/۶۴۸

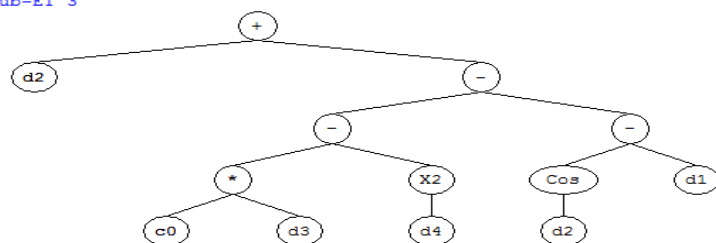
Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل (۴-۱۳) نمودار درختی برآورد TDS رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن

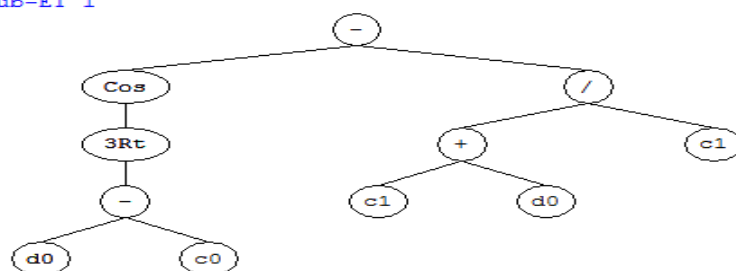
$$TDS = \left[\left[\left(\cos(HCo3 - 2.704865) - (HCo3) \right) - \left(SAR - \left(\frac{EC}{-2.704865} \right) \right) \right] + [HCo3 - (\tan^{-1} SAR + SAR) + (\cos NA - 2.1276)] + [HCo3 + ((8.382721NA) - SAR^2) - (\cos HCo3 - EC)] \right]$$

رابطه ریاضی (۴-۴) برآورد شاخص کیفی کل مواد محلول رودخانه بابلرود

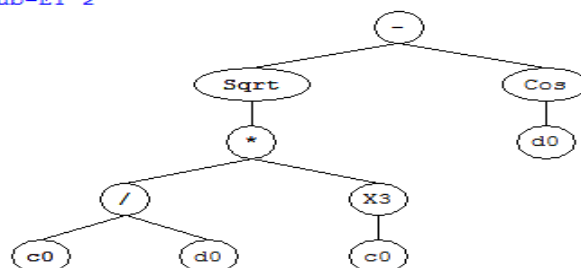
با توجه به رابطه ریاضی بدست آمده از الگوی ۸ برآورد TDS رودخانه بابلرود، پارامترهای EC, Na, SAR, HCO₃ با شرکت در رابطه ریاضی فوق بالاترین میزان ضریب همبستگی و کمترین خطا را بدست آوردند ولی یکی از اهداف ما که استفاده از کمترین پارامتر می باشد محقق نشده است به همین دلیل از الگوی ۱ استفاده می کنیم زیرا تفاوت زیادی در میزان ضریب همبستگی و خطا ندارد و این تفاوت قابل اغماض می باشد.

با اجرای مدل الگوی ۱ نمودار درختی زیر بدست آمد :

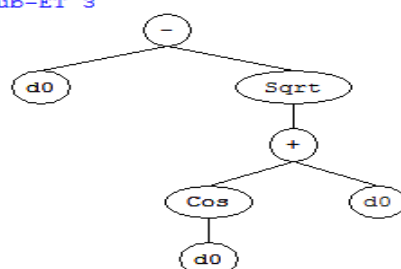
Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل (۴-۱۴) نمودار درختی الگوی ۱ برآورد TDS رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن

با توجه به نمودار درختی و جدول (۴-۱۵) مقادیر ثابت بیان درختی، رابطه ریاضی منطقی با کمترین پارامتر ورودی طبق رابطه ریاضی (۴-۵) بدست آمد.

جدول (۴-۱۵) - مقادیر ثابت در بیان درختی الگوی ۸ برآورد TDS رودخانه بابلرود

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
-۲/۲۵۹	۲/۴۰۲	-۵/۶۷۵	۹/۲۳۳	۳/۱۸۱	۷/۴۸۵

$$TDS = \cos(\sqrt[3]{EC - 7.485}) - \frac{3.181 + EC}{3.181} + \sqrt{\frac{7267.27}{EC}} - \cos EC + EC - \sqrt{\cos EC + EC}$$

رابطه ریاضی (۵-۴) برآورد شاخص کیفی کل مواد محلول الگوی ۱ رودخانه بابلرود

رابطه ریاضی پیشنهادی نشان می‌دهد که TDS با EC رابطه خوبی دارد و این در تطابق با نتایج سایر محققین مانند جلیلیان و امینی (۱۳۸۶)، بتوشا و همکاران (۲۰۰۹)، تیرامالین و کوریان ۲۰۰۹ و همکاران (2009)، کرمی و هوشمند (۱۳۹۲)، دودکانلوی میلان و همکاران (۱۳۹۰) و کرمی و گلابی (۱۳۹۰) می‌باشد.

با توجه به روابط بدست آمده از پژوهش های کرمی و هوشمند و نیز تیرامالین و کوریان ۲۰۰۹ و همکاران در ذیل به مقایسه این دو رابطه با رابطه (۲-۴) بدست آمده از الگوی ۱ مدل برنامه ریزی بیان ژن رودخانه بابلرود پرداخته شده است.

$$TDS = (0.55 \text{ to } 0.7)EC$$

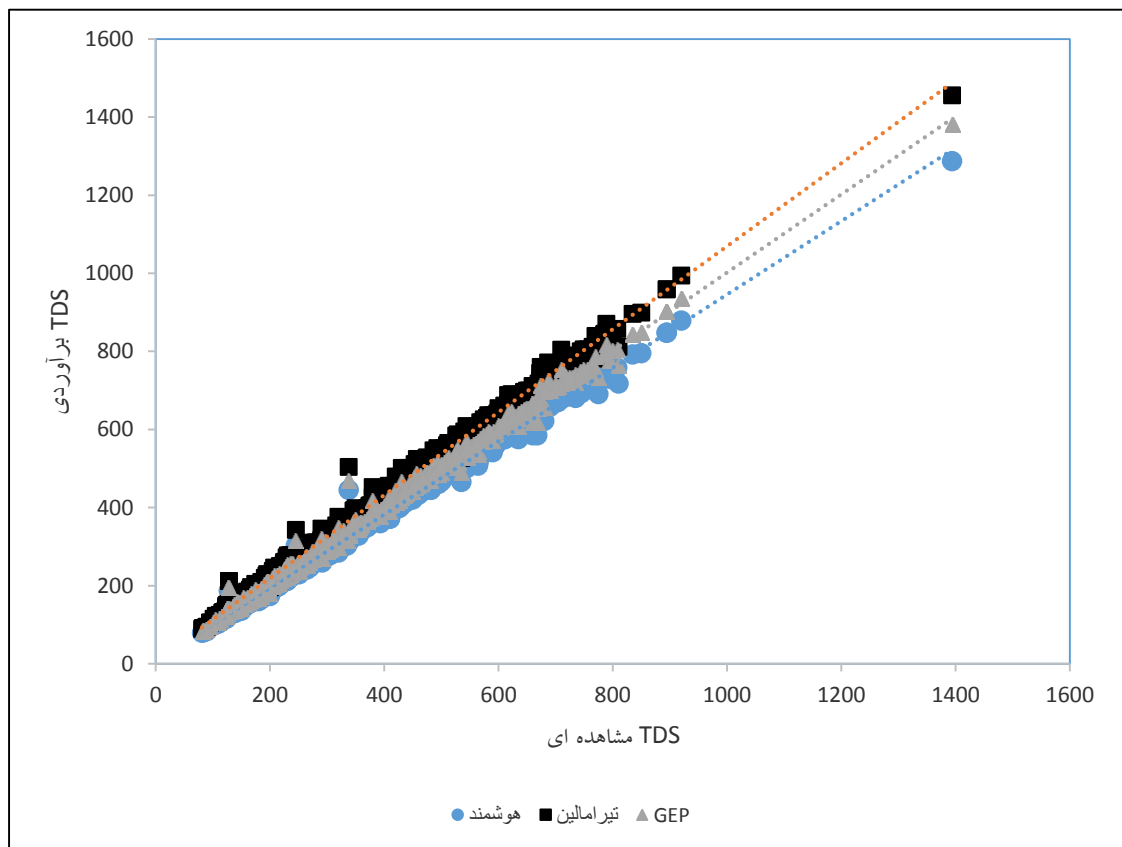
تیرامالین و کوریان ۲۰۰۹

$$TDS = 0.619(EC)$$

کرمی و هوشمند ۱۳۹۲

جدول (۱۶-۴) مقایسه شاخص های آماری سه روش برآورد TDS

RMSE	R ²	
32.14229	0.996452	تیرامالین و کوریان ۲۰۰۹
20.85255	0.996452	کرمی و هوشمند ۱۳۹۲
10.68704	0.996485	برنامه ریزی بیان ژن



شکل (۴-۱۵) نمودار مقایسه برآورد TDS به سه روش

با توجه به نمودار (۴-۱۵) و جدول (۴-۱۶) می توان نتیجه گرفت که روش برنامه ریزی بیان ژن با ارائه رابطه ریاضی پیشنهادی (۴-۵) به خوبی توانسته با کمترین میزان RMSE و بیشترین میزان همبستگی نسبت به دو رابطه دیگر شاخص کیفی کل مواد محلول را برآورد کند.

۴-۳-۲- برآورد نسبت جذب سدیم SAR

۴-۳-۲-۱- همبستگی و رگرسیون

همبستگی بین داده ها قبلا در جدول (۴-۹) نشان داده شد و تحلیل رگرسیون بین متغیرهای مستقل و وابسته نیز با نرم افزار SPSS23 انجام شده است. که به این منظور از روش گام به گام (Stepwise Regression) استفاده شده است. در روش گام به گام ۵ الگوی ورودی پیشنهاد شد که در جدول (۴-۱۳) ارائه شده است.

جدول (۴-۱۷) - الگو های ورودی برآورد SAR رودخانه بابلرود اسخراج شده از SPSS

الگو	پارامتر ورودی
۱	Na
۲	Na , TDS
۳	Na , TDS, HCO ₃
۴	Na , TDS, HCO ₃ ,Q
۵	Na , TDS, HCO ₃ ,Q ,Cl

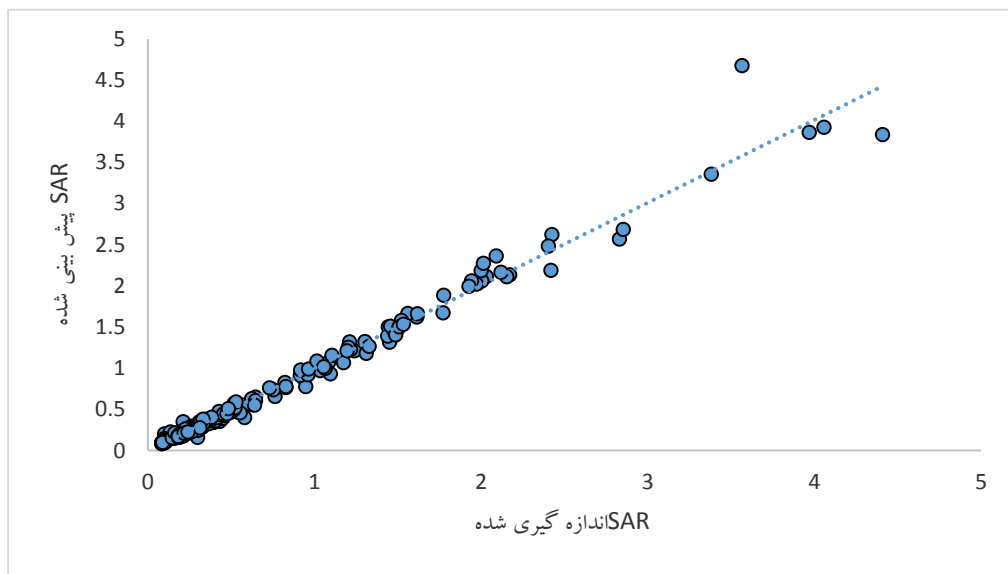
۴-۳-۲- برنامه ریزی بیان ژن

الگو های ورودی جدول فوق را در نرم افزار GeneXproTools وارد کردیم و در جدول (۴-۱۴) ضرایب همبستگی، جذر میانگین مربعات خطا و متوسط خطای میانگین حاصل از اجرای ترکیب های مذکور در هر مرحله آموزش و آزمون، ارائه گردیده است. مدل مطلوب و بهینه مدلی است که با دارا بودن بیشترین ضریب همبستگی، کمترین میزان خطا را به خود اختصاص دهد.

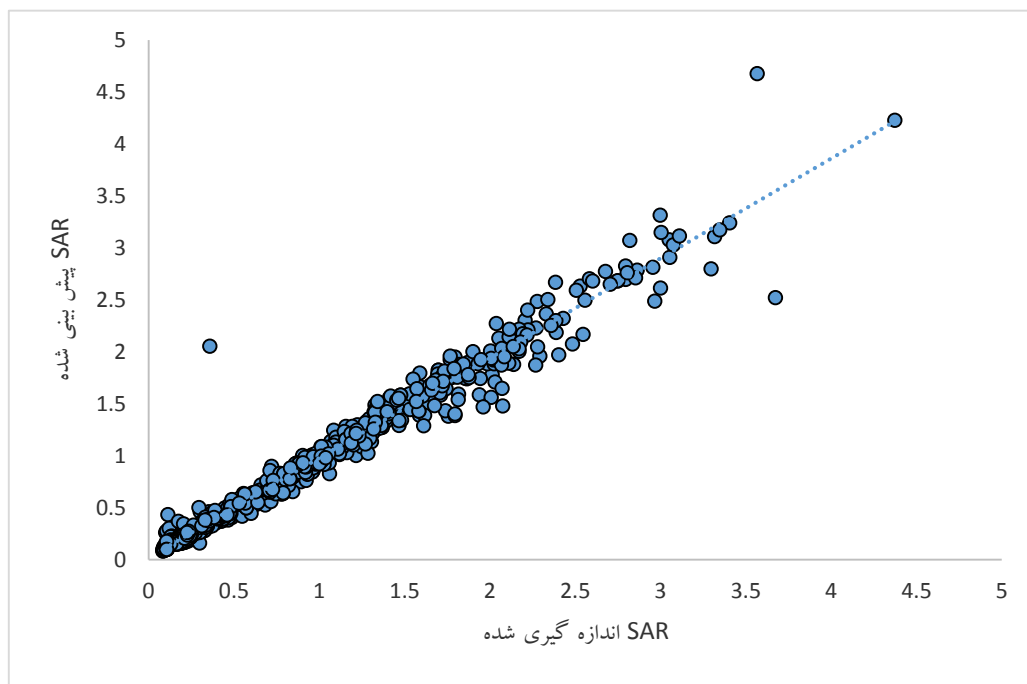
جدول (۴-۱۸) - نتایج مدل برنامه ریزی بیان ژن برای الگوهای ورودی برآورد SAR رودخانه بابلرود

الگو	مرحله اجرا	R^2	RMSE	MAE
1	آموزش	0.973	0.133	0.079
	آزمون	0.969	0.136	0.089
2	آموزش	0.974	0.131	0.064
	آزمون	0.964	0.143	0.083
3	آموزش	0.952	0.175	0.108
	آزمون	0.955	0.158	0.109
4	آموزش	0.943	0.205	0.097
	آزمون	0.958	0.163	0.097
5	آموزش	0.982	0.109	0.050
	آزمون	0.972	0.125	0.063

مقایسه نتایج جدول فوق نشان داد که الگو ۵ با پارامترهای ورودی Na , TDS , HCO_3 , Q , Cl با دارا بودن بیشترین ضریب همبستگی برابر ۰/۹۸۲ در آموزش و ۰/۹۷۲ در آزمون و کمترین خطا مدل بهینه می‌باشد. نتایج حاکی از آن است که الگوی مورد نظر با دارا بودن ضریب همبستگی بیشتر، توانایی بالایی در برآورد SAR دارد و به عبارتی نتایج قابل اطمینانی را در پیش بینی نسبت جذب سدیم ارائه می‌کند در شکل های (۴-۱۶) و (۴-۱۷) نمودارهای همبستگی پارامترهای اندازه گیری شده و پیش بینی شده در هر دو مرحله آموزش و آزمون و همچنین ساختار ORF ژن های حاصل از اجرای الگوی ۵ نشان داده شده است.



شکل (۴-۱۶) نمودار اجرای الگو ۵ برآورد SAR رودخانه بابلرود در مرحله آموزش



شکل (۴-۱۷) نمودار اجرای الگو ۵ برآورد SAR رودخانه بابلرود در مرحله آزمون

ساختار ORF ژن های مدل برتر به زبان کاروا در زیر نشان داده شده است:

$/d0 \cdot \text{Atan} \cdot - \cdot \text{Ln} \cdot X2 \cdot * \cdot \text{Atan} \cdot d4 \cdot d1 \cdot d3 \cdot d1 \cdot d2 \cdot c1 \cdot d0 \cdot c0 \cdot c0$

+

$\text{Sqrt} \cdot / \cdot d0 \cdot - \cdot d1 \cdot \text{Sin} \cdot \text{Cos} \cdot + \cdot c0 \cdot d0 \cdot d0 \cdot d1 \cdot d3 \cdot d1 \cdot d1 \cdot c1$

+

$/d0 \cdot c0 \cdot d1 \cdot / \cdot \text{Sin} \cdot X3 \cdot d4 \cdot d4 \cdot d1 \cdot d4 \cdot c1 \cdot d4 \cdot d1 \cdot d4 \cdot d4 \cdot c0$

در ادامه بیان درختی الگوی فوق نشان داده شده است. که با استفاده از جدول (۴-۱۹) رابطه ریاضی برآورد SAR رودخانه تجن را طبق رابطه ریاضی (۴-۶) بدست آوردیم.

Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل (۴-۱۸) نمودار درختی برآورد SAR رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن

جدول (۴-۱۹) - مقادیر ثابت در بیان درختی الگوی ۵ برآورد SAR رودخانه بابلرود

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
۵/۷۶۳	-۶/۴۰۳	-۲/۴۹۱	-۷/۱۸۹	-۲/۴۹۱	-۷/۱۸۹

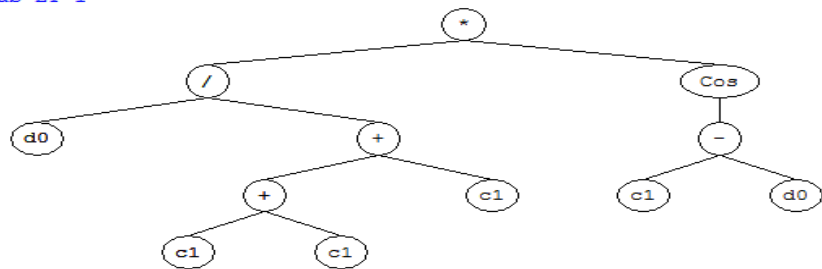
$$SAR = \left[\frac{NA}{\tan^{-1}[(\ln CL * TDS) - (\tan^{-1} Q)^2]} + \left[\frac{NA}{TDS - (\sin \cos(-7.189178 + NA))} \right] + \frac{NA}{-6.403137} \right]$$

رابطه ریاضی (۴-۶) برآورد نسبت جذب سدیم SAR رودخانه بابلرود

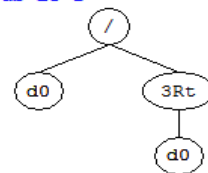
در رابطه ریاضی فوق Na,Cl,Q,TDS شرکت دارند تا بالاترین ضریب همبستگی و کمترین میزان خطا بدست آید. ولی با توجه به اینکه هدف دست یافتن به رابطه ریاضی‌ای است که بتواند با کمترین پارامتر ورودی و بالاترین مقدار ضریب همبستگی و نیز کمترین میزان خطا SAR را برآورد کند بنابراین از الگو ورودی ۱ استفاده شد که تفاوت چندانی در R^2 و میزان خطا با الگوی ۵ نداشت و این تفاوت قابل چشم پوشی می‌باشد.

پس از اجرای مدل الگوی ۱ نمودار درختی زیر حاصل شد:

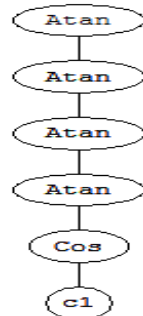
Sub-ET 1



Sub-ET 2



Sub-ET 3



شکل (۱۹-۴) نمودار درختی الگوی ۱ برآورد SAR رودخانه بابلرود با استفاده از برنامه ریزی بیان ژن

با توجه به نمودار درختی و جدول (۴-۲۰) مقادیر ثابت بیان درختی، رابطه ریاضی منطقی با کمترین پارامتر ورودی طبق رابطه ریاضی (۴-۷) بدست آمد.

جدول (۴-۲۰) مقادیر ثابت در بیان درختی الگوی ۱ برآورد SAR رودخانه بابلرود

G3C1	G3C0	G2C1	G2C0	G1C1	G1C0
۸/۱۲۱	-۹/۷۱۵	۳/۲۸۲	۵/۸	۵/۹۱۱	۸/۵۸۶

$$sar = \left[\frac{NA}{17.733} \times \cos(5.911 - NA) + \frac{Na}{\sqrt[3]{NA}} + 0.5113 \right]$$

رابطه ریاضی (۷-۴) برآورد نسبت جذب سدیم SAR الگوی ۱ رودخانه بابلرود

با توجه به رابطه ریاضی پیشنهادی با استفاده از پارامتر سدیم که مطابق رابطه ریاضی SAR رابطه مستقیم با نسبت جذب سدیم دارد هر دو هدف کم بودن پارامترهای ورودی و بالا بودن ضریب همبستگی بدست آمد که نشان دهنده کارایی بالا مدل برنامه ریزی بیان ژن در برآورد نسبت جذب سدیم می باشد.

فصل پنجم

نتیجه گیری و پیشنهادات

۵-۱- نتیجه گیری

در پژوهش حاضر از روش برنامه ریزی بیان ژن (GEP) در برآورد پارامترهای TDS و SAR رودخانه‌های تجن و بابلرود مازندران استفاده شده است که به این منظور، اطلاعات ۱۲ متغیر کیفی ماهانه آب برای یک دوره ۴۰ ساله (۱۳۹۰-۱۳۵۰) از شرکت آب منطقه‌ای استان مازندران جمع‌آوری گردید و برای اجرای مدل ۲۰ درصد داده‌ها به مرحله آموزش و ۸۰ درصد داده‌ها مورد آزمون قرار گرفتند.

با توجه به نمودارهای درختی مدل‌ها و نیز روابط ریاضی بدست آمده از آنها می‌توان متغیرهای با تاثیر بالا در برآورد پارامترهای مورد بحث این پژوهش را اینطور بیان کرد که:

در برآورد TDS رودخانه تجن از الگو ۵ با متغیرهای ورودی EC, SO_4, SAR, HCO_3, Cl استفاده شده است که نرم افزار تنها متغیرهای SAR و EC, SO_4 را مورد استفاده قرار داد و به بالاترین ضریب همبستگی، کمترین میزان خطا دست یافت. با توجه به اینکه با این متغیرهای ورودی بالاترین ضریب همبستگی را بدست آمده است و بهترین معادله ریاضی برآورد بدست آمد ولی یک هدف مهم دیگر که کم بودن تعداد پارامترهای مورد استفاده در معادله می‌باشد را می‌توان با معادله‌ی دوم برآورد شاخص کیفی کل مواد محلول بدست آورد که تفاوت چندانی در مقادیر R^2 و خطا با معادله اول نداشتند و تنها از پارامتر EC استفاده کرده است که یک پارامتر زودیافت می‌باشد.

در برآورد TDS رودخانه بابلرود از الگو ۸ با متغیرهای ورودی $EC, Na, SAR, SO_4, Cl, HCO_3$ استفاده شده است که مدل برنامه ریزی بیان ژن تنها متغیرهای EC, Na, SAR, HCO_3 را مورد استفاده قرار داد و به بالاترین ضریب همبستگی، کمترین میزان خطا دست یافت. ولی در بعضی مواقع بدلیل نداشتن تمام پارامترها و دیر یافت بودن آنها باید از تفاوت‌های ناچیز آنها در مقادیر شاخص‌های آماری چشم‌پوشی کرد به همین دلیل از الگوی ورودی ۱ استفاده شده است که تفاوت چندانی در

میزان ضریب همبستگی و مقدار خطا با الگوی ۸ نداشت و تنها از پارامتر زودیافت EC استفاده کرده است.

در برآورد SAR نسبت جذب سدیم رودخانه بابلرود از الگو ۵ با متغیرهای ورودی Na , TDS, HCO_3 , Q, Cl استفاده شده است که مدل برنامه ریزی بیان ژن متغیرهای Na , TDS, Q, Cl را مورد استفاده قرار داد و از HCO_3 بدلیل کاهش ضریب همبستگی در مدل و افزایش خطا استفاده نکرد. ولی همانطور که قبلا اشاره شد کم بودن پارامترهای ورودی در زمانیکه بدست آوردن پارامترهای بالا وقت گیر و نیاز به هزینه‌ی زیاد باشد امر مهمی تلقی می‌شود بنابراین از الگوی ۱ که تفاوت چندانی در مقدار ضریب همبستگی و خطا با الگو ۵ نداشت استفاده شد و رابطه ریاضی خوبی با استفاده از پارامتر سدیم بدست آمد که با رابطه ریاضی SAR در تطابق می‌باشد.

در برآورد SAR نسبت جذب سدیم رودخانه تجن از الگو ۲ با متغیرهای ورودی Na و EC استفاده شده است که استفاده از سدیم با رابطه ریاضی SAR تطابق دارد و نیز استفاده از EC بدلیل زودیافت بودن این پارامتر کمک زیادی در بالا بردن کارایی مدل برنامه ریزی بیان ژن و بالا بردن ضریب همبستگی و کاهش خطا کرده است و هدف کم بودن پارامترهای مورد استفاده در معادله و نیز هدف دیگر که بالاترین ضریب همبستگی و کمترین میزان خطا بوده را بدست آورده است.

این نتایج نشان دهنده عملکرد بالای روش برنامه ریزی بیان ژن در برآورد پارامترهای کیفی رودخانه های بابلرود و تجن می‌باشد.

و نیز با توجه به نمودارهای ویلکاکس رودخانه‌های تجن و بابلرود این دو رودخانه از جنبه مصرف کشاورزی خوب ارزیابی می‌شوند.

۵-۲- پیشنهادات

۱- پژوهشگران استفاده از روش های هوشمند مخصوصا روش برنامه ریزی بیان ژن را در برآورد پارامترهای کیفی رودخانه ها مد نظر قرار دهند.

۲- با توجه به اینکه در پژوهش حاضر تنها به برآورد نسبت جذب سدیم و شاخص کیفی کل مواد محلول پرداخته شده است لذا کارایی سیستم های هوشمند در برآورد دیگر پارامترها نیز بررسی شود.

۳- روابط ریاضی بدست آمده مختص این دو رودخانه است و برای رودخانه های دیگر باید دوباره مدلسازی انجام گیرد.

۴- با توجه به اینکه در این پژوهش تنها از پارامترهای کیفی آب به منظور برآورد TDS و SAR استفاده شده است لذا میتوان پارامترهای دیگر مانند دما، باران و غیره را نیز برای برآورد مورد بررسی قرار داد.

منابع

--آذرمدل ح.، مصطفی‌زاده ر. و قاسمی.ا.، ۱۳۸۹، ارزیابی شبکه ی ایستگاه های سنجش کیفیت آب سطحی رودخانه ی گرگانرود استان گلستان، مجله علمی-پژوهشی علوم و مهندسی آبخیز داری ایران، سال چهارم، شماره ی ۱۰، ص ۵۷

-آزموده ع.، سلیمانی ک. و مهدوی ف.، ۱۳۹۲، ارزیابی کارایی شبکه عصب مصنوعی در پیشبینی شاخص کیفی کل مواد محلول TDS آب رودخانه تجن ساری، فصلنامه مدیریت و برنامه ریزی محیط زیست، سال سوم شماره ۲، ص ۲۷.

- آگدرنژاد ا.، اسدی کپورچال ص.و پرهیزکار م.، ۱۳۹۳، مقایسه کارایی مدل های شبکه عصبی مصنوعی و رگرسیون برای پیش بینی مقدار کل جامدات محلول (TDS) رودخانه کارون (ایستگاه پل فولاد)، کنفرانس ملی علوم و مهندسی محیط زیست، دانشگاه شهید چمران اهواز

-البرزی م.، (۱۳۸۸)، الگوریتم ژنتیک، چاپ اول، انتشارات علمی دانشگاه صنعتی شریف، تهران، ص ۱۴، ۱۳، ۲۷.

- بختیاری م؛ محمدی مجد ع.ر؛ حسن زاده ز. و بهرامی یار احمدی م، ۱۳۸۸، مقایسه دو الگوریتم یادگیری شبکه عصبی مصنوعی در تخمین بار معلق رسوب رودخانه بابلرود، دهمین سمینار سراسری آبیاری و کاهش تبخیر، کرمان، دانشگاه شهید باهنر

پارسافر، ن. زارع ابیانه، ح. ۱۳۸۷. "بررسی روابط رگرسیونی بین هدایت الکتریکی و کاتیون ها و آنیون های چاه های آب دشت همدان-بهار". همایش ملی بحران آب در کشاورزی و منابع طبیعی، دانشگاه آزاد شهر ری، شهر ری، ایران.

-جروقی س.، جباری ا. و اعلمی م.ت.، ۱۳۸۸، شبیه سازی کیفی رودخانه بابلرود با استفاده از نرم افزار MIKE 11، هشتمین کنگره بین المللی مهندسی عمران، دانشگاه شیراز

- جلیلیان م. و امینی ژ.، ۱۳۸۶، بررسی رابطه بین فاکتور EC و TDS در آب رودخانه زاینده رود و آب آشامیدنی شهر اصفهان،/اولین کنفرانس مهندسی برنامه ریزی و مدیریت سیستم های محیط زیست، تهران، دانشگاه تهران، دانشکده محیط زیست

-حسامی افشار م.، خلیلی ک.، علیقلی نیا ت.و زمانزاد قویدل س.، ۱۳۹۲، پیشبینی غلظت TDS رودخانه ی زرينه رود بر مبنای کمترین تعداد پارامتر ورودی به کمک برنامه ریزی ژنتیک، سومین کنفرانس بین المللی برنامه ریزی و مدیریت محیط زیست، دانشگاه تهران

-- خسروی خ.، وهاب زاده ق.، صفری ع.، کریمی ف. و جلیلیان ف.، ۱۳۹۱، بررسی روند تغییرات زمانی پارامترهای کیفیت آب رودخانه بابلرود، سومین همایش ملی مقابله با بیابانزایی و توسعه پایدار تالاب های کویری ایران، اراک

-خلج م،(۱۳۹۰)، پایان نامه ارشد، روندیابی سیلاب رودخانه با استفاده از برنامه ریزی ژنتیک و موج دینامیکی، دانشکده کشاورزی، دانشگاه رازی کرمانشاه

-دربندی، ص؛ عباسپور، ا و شعبی نوبریان، م، ۱۳۹۰، پیش بینی جریان ورودی به مخزن سد علویان با استفاده از برنامه ریزی ژنتیک، پنجمین کنفرانس سراسری آبخیزداری و مدیریت منابع آب و خاک کشور، کرمان، انجمن مهندسی آبیاری و آب ایران

-دودکانلوی خ، هنربخش ا، داودیان دهکردی ع.ر، پژوهش م، بیگی هرچگانی ح و استواری ی، ۱۳۹۰، تخمین TDS آب زیرزمینی دو آبخوان به کمک EC و ارتباط دو پارامتر با یکدیگر ، اولین کنگره ملی علوم و فناوریهای نوین کشاورزی دانشگاه زنجان

-رضویان ف.، پایان نامه ارشد، ۱۳۸۳، بررسی اثرات کارخانه چوب و کاغذ مازندران بر رودخانه تجن، دانشگاه آزاد اسلامی، واحد علوم تحقیقات تهران.

-سارانی ن.، سلطانی ج.، معاشری س ع.، ۱۳۹۱، ارزیابی و مقایسه عملکرد شبکه های عصبی مصنوعی MLP و GFF در شبیه سازی شاخص کیفی SAR رودخانه سیستان. همایش ملی جریان و آلودگی آب، دانشگاه تهران

-سلطانی،ع. قربانی،م ع. فاخری فرد،ا. دربندی،ص و فرسادی زاده،د، ۱۳۸۹، برنامه ریزی ژنتیک و کاربرد آن در مدلسازی فرآیند بارش رواناب. مجله دانش آب و خاک. جلد ۲۰ شماره ۴، ص ۶۱

- صالحی ا.، ۱۳۹۰، پایان نامه ارشد، تخمین مکانی-زمانی تغییرات سطح آب زیرزمینی با استفاده از زمین آمار و مدل های هوشمند، دانشکده کشاورزی، دانشگاه تبریز

- فربودفام. ن، قربانی. م ع، اعلمی م، پیش بینی جریان رودخانه با استفاده از برنامه ریزی ژنتیک (مطالعه موردی: حوضه آبریز رودخانه ليقوان) مجله دانش کشاورزی. جلد ۱۹ شماره ۴، ص ۱۰۷

- فریدونی م. و امیراحمدی ع.ر.، ۱۳۹۳، شبیه سازی کل جامدات محلول در آب (TDS) با استفاده از سیستم استنتاج فازی- عصبی- موجک (Conjoined Wavelet ANFIS Networks) و مقایسه آن با سیستم استنتاج فازی- عصبی تطبیقی (ANFIS) (مطالعه موردی: مخزن سد درودزن)/ولین کنفرانس ملی مهندسی عمران و توسعه پایدار/ایران، بصورت الکترونیکی، مرکز راهکارهای دستیابی به توسعه پایدار

- قربانی، م ع و صالحی، ا.، ۱۳۹۰، استفاده از برنامه ریزی بیان ژن در بررسی تغییرات داده های کیفی آب زیرزمینی با نوسانات سطح آب در دشت برخوار اصفهان، ششمین کنگره ملی مهندسی عمران، سمنان، دانشگاه سمنان

- کاظم زاده م.، عظیمی و. و خلیقی سیگارودی ش.، ۱۳۹۳، ارزیابی کارایی مدل m5 و mlp در برآورد میزان کل جامدات محلول (مطالعه موردی : حوزه آبخیز ليقوان)، نشریه علمی- پژوهشی علوم و مهندسی آبخیزداری ایران، سال هشتم، شماره ۲۷، ص ۱-۱۰

-کرمی ا. و هوشمند ع. ۱۳۹۲. صحت سنجی و برآورد ارتباط بین EC و TDS آب رودخانه کارون در فصول پرآب و فصول کم آب، چهارمین همایش ملی مدیریت شبکه های آبیاری و زهکشی، دانشگاه شهید چمران اهواز، دانشکده مهندسی علوم آب

-کرمی ب و گلابی م ، ۱۳۹۱، شبیه سازی و پیش بینی پارامترهای کیفی آب با استفاده از شبکه عصبی مصنوعی، روش فازی-عصبی و رگرسیون آماری (مطالعه موردی رودخانه کارون، استان خوزستان)، نهمین سمینار بین المللی مهندسی رودخانه، اهواز، دانشگاه شهید چمران

-کریمی دمنه ر.، ۱۳۹۳، پایان نامه ارشد، پیش بینی بار رسوب معلق سد کرخه به کمک برنامه ریزی بیان ژن، دانشکده کشاورزی، دانشگاه صنعتی شاهرود

- کریمی دمنه، ر ؛ امامقلی زاده، ص ؛ اژدری، خ و زلّقی، ا ، ۱۳۹۲، برآورد بار رسوب معلق ورودی به سد کرخه به کمک برنامه ریزی بیان ژن و شبکه عصبی مصنوعی، اولین همایش ملی الکترونیکی کشاورزی و منابع طبیعی پایدار، تهران، موسسه آموزش عالی مهر اروند، گروه ترویجی دوستداران محیط زیست

- لویزه، ف. (۱۳۸۴) "بررسی منشا و میزان تمرکز آلاینده های فلزی سنگین در حوضه آبریز رودخانه های تالار تجن نکارود"، بیست و یکمین گردهمایی علوم زمین.

- مرادی ش. و زمانی نوری ع. ر، ۱۳۹۲، بررسی روند تغییرات کیفی آب رودخانه تجن در بازه زمانی ۱۳۸۸-۱۳۵۲، پنجمین کنفرانس مدیریت منابع آب ایران، پردیس فنی و مهندسی شهید عباسپور دانشگاه شهید بهشتی

-مطیع قادر ح، لطفی ش. و سید اسفهلان م.، ۱۳۸۹، مروری بر برخی از روش های بهینه سازی هوشمند، چاپ اول، انتشارات دانشگاه آزاد اسلامی واحد شبستر، تبریز

- مغربی، م.، تجربی، م.، ۱۳۸۸، مدل سازی سیر و حرکت باکتری کلیفرم در رودخانه جاجرود، هشتمین کنگره بین المللی مهندسی عمران دانشگاه شیراز

- مفتاح هلقی م. و صادقی ز.، ۱۳۸۸، تاثیر سد گلستان بر کیفیت آب گرگانرود، چهارمین کنفرانس سراسری آبخیز داری و مدیریت منابع آب و خاک، کرمان

- میثاقی ف. و محمدی ک.، ۱۳۸۳، پیش بینی تغییرات کیفیت آب رودخانه زاینده رود با استفاده از شبکه های عصبی مصنوعی، دومین کنفرانس ملی دانشجویی منابع آب و خاک، شیراز، دانشکده کشاورزی دانشگاه شیراز

- میرزایی م.، امینی راد ح.، گلبابایی کوتنایی ف. و نصراللهی هریکنده ع.، ۱۳۹۰، مدلسازی کیفی پارامترهای DO و BOD ، pH رودخانه بابلرود توسط نرم افزار QUAL2K پنجمین همایش ملی و نمایشگاه تخصصی مهندسی محیط زیست، دانشگاه تهران

- ورسه س.، پناهی م.، خضری م. و محمدی فاضل ا.، ۱۳۹۳، بررسی اثر فعالیت های صنعتی بر نوسان BOD، COD و TSS در رودخانه تجن، فصلنامه علوم و مهندسی محیط زیست، دوره ۱، شماره ۲، ص

۴۵

-یوسفی م.، قربانی ه. و امامقلی زاده ص. (۱۳۹۴) بررسی کیفی آب رودخانه تجن مازندران از جنبه مصرف کشاورزی اولین همایش ملی کیفیت منابع آب و توسعه پایدار، آبان ۹۴، دانشگاه اراک

-Asadollahfardi, A., Taklifi, Gh. and Ghanbari, A. (2012). Application of artificial neural network to predict TDS in Talkheh Rud River. Journal of Irrigation and Drainage Engineering. 138(4):363–70.

- Bathusha, M.I. and Saseetharan, M.K, (2009) “Correlation and Regression study on Coimbatore City West Zone Groundwater Quality Characteristics”, IE(I) Journal- EN, Vol.89, pp.8-13
- Diamantopoulou, M.J. Antonopoulos, V.Z. and Papamichail, D.M. 2005. The use of a neural network technique for the prediction of water quality parameters of Axios River in Northern Greece. European Water Resources Association. 11(12) 55-62.
- Ferreira C.,(2001), Gene Expression Programing: A New Adaptive Algorithm for Solving Problems, Complex systems, 13(2), 87-129
- Ferreira C.,(2004), Gene Expression Programing and the evolution of computer programs,. recent development in biologically inspired computing,. 82-103
- Ghorbani. M.A., Singh, V.P., Kashani, M., and Kashani, A. 2012. Modelling pan evaporation using genetic programming. Journal of Statistics: Advances in Theory and Application, pp.1-23.
- Guven, A. 2009. Linear genetic programming for time-series modelling of daily flow rate. Journal Earth System Science, vol 118, 157-173.
- Khalil, B., Ouarda, T.B.M.J., St- Hilaire, A., 2011, Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis, Journal of Hydrology, 405, PP. 277–287.
- Kuo, Y-M., Liu, C-W. And Lin, K-H. Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of Blackfoot disease in Taiwan. (2004). Water Research. 38(1):148-58.

- Kuo, J-T., Hsieh, M-H., Lung, W-S. And She,N. Using artificial neural network for reservoir eutrophication prediction. 2007. Ecological Modeling. 200(1-2):171-77.
- lopes H.S. and Weinret W.R., (2004) ,EGYPSIS: An enhanced gene expression programing approach for symbolic regression problems, Journal of Applied Mathematics and Computer Science, 14(3), 375-384
- Musavi-Jahromi, SH. and Golabi, M. (2008). Application of artificial neural networks in the river water quality modeling: Karoon River, Iran. Journal of Applied Sciences. 8:2324-28.
- Michael SAM, Khepar SD, Sondhi SD (2008). Water wells and pumps. McGraw-Hill, New Delhi.
- Najah, A. El-Shafie, A. Karim, O.A. and El-Shafie, A.H. 2013. Application of artificial neural networks for water quality prediction. Neural Comput & Applic. 22: 187–S201.
- Sattari, M.T. Pal, M. Apaylin, H. Ozturk, F. 2013. M5 Model Tree Application in Daily River Flow Forecasting in Sohu Stream, Turkey. Water Resources. 40: 223-242.
- Thirumalini S and Kurian J. 2009. Correlation between electrical conductivity and total dissolved solids in Natural waters. Malaysian Journal of Science 28 (1): 55-61.
- Wilson W.S., (2008), Classifeir conditions using gene expression programing, IlliGAL report, 4998, 206-217
- Zaman Zad Ghavidel, S. and Montaseri, M. 2014. Application of different data-driven methods for the prediction of total dissolved solids in the Zarinehroud basin. Stoch Environ Res Risk Assess. 28: 2101–2118.

Abstract

Water ecosystem including rivers is one of the components of the earth. It is considered not only as an important settlement for aquatic organisms but also as an important source of food for human beings. It is now taken for granted that the quality of water is as much important as its quantity. The term water quality is also used in physical, chemical and biological explanation for special purposes. Water quality studies have become increasingly important due to considerable decrease in water resources. However, measuring water quality parameters is highly expensive. Thus, estimating some parameters from other already measured parameters is of crucial importance. In this study, GEP model was developed to estimate water quality parameters of Tajan and Babolrood Rivers in Mazandaran. Therefore, hydrometric data on these two rivers during 1971-2011 was collected and examined. To implement GEP model, the proposed models were first provided using SPSS and stepwise regression, and then using GeneXproTools, they were examined based on R^2 , RMSE and MAE. The results of implementing GEP model in testing and training phases showed that GEP model is of high capability in estimating TDS and SAR of Tajan and Babolrood Rivers, offering precise mathematical relationships.

Keywords: GEP, Tajan River, Babolrood River, Estimate TDS, Estimate SAR



Shahrood University of Technology
Faculty of Agriculture
Department of Water & Soil

**Prediction of Water Quality Parameters in the Tajan and Babolrood
Rivers Using GEP Model**

Mahdi Yusefi

Supervisors:
Dr.Hadi Ghorbani
Dr.Samad Emamghlizadeh

February 2016