

پیش‌بینی جریان ماهانه با استفاده از رویکرد برنامه ریزی بیان ژن (مطالعه موردی: حوضه آبخیز چهل‌چای)

عبدالرحیم لطفی^۱، رمضان بهترک^۲، رضا نقوی^۳، حامد تقی‌زاده (نویسنده مسئول)^۴

^۱مدیرکل منابع طبیعی و آبخیزداری استان گلستان، گرگان Gol.Modir@frw.ir

^۲سرپرست معاونت فنی و آبخیزداری استان گلستان، گرگان Gol.Mohandesi@frw.ir

^۳معاون اداره مهندسی و مطالعات منابع طبیعی و آبخیزداری گلستان، گرگان reza.naghavy@gmail.com

^۴رئیس هیئت مدیره شرکت مهندسی مشاور ترازآب جامع، گرگان Hamed.t.nyu@gmail.com

چکیده

نیاز روزافزون به آب سبب گردیده است که برنامه‌ریزی‌های مدیریتی به منظور کنترل مصرف آب در آینده از اهمیت بیشتری برخوردار باشد. با پیش‌بینی نمودن جریان رودخانه‌ها علاوه بر مدیریت بهره‌برداری از منابع آب، می‌توان حوادث طبیعی نظیر سیل و خشکسالی را نیز پیش‌بینی و مهار نمود. همچنین می‌توان از نتایج یک مدل شبیه‌سازی به منظور بررسی صحت داده‌ها و یا اصلاح و تکمیل آن‌ها استفاده کرد. روش‌های متعددی همچون مدل‌های سری‌زمانی، شبکه‌های عصبی مصنوعی، منطق فازی و برنامه‌ریزی بیان ژن (GEP) برای پیش‌بینی جریان رودخانه به کار می‌رود. در مطالعه حاضر به منظور پیش‌بینی جریان رودخانه چهل‌چای از روش GEP استفاده شد. داده‌های جریان ماهانه این رودخانه در محل ایستگاه آبسنجی لزوره در طی سال‌های ۱۳۴۵ تا ۱۳۹۶ برای ایجاد مدل استفاده شد که ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد برای صحت‌سنجی مدل به کار رفت. جهت بررسی و ارزیابی نقش حافظه در افزایش یا کاهش دقت، الگوهای ورودی متعددی به مدل‌ها معرفی شد و در نهایت مشخص گردید که با افزایش حافظه، دقت مدل‌سازی نیز افزایش می‌یابد به طوری که بیشترین دقت در الگوی سوم به دست می‌آید و سپس با افزایش حافظه از دقت مدل‌سازی کاسته می‌شود.

واژه‌های کلیدی

الگو، تاخیر بهینه، ساختار درختی، نمودار PACF.

۱. مقدمه

در حالت کلی روش‌های پیش‌بینی جریان رودخانه را می‌توان به دو دسته تقسیم نمود، دسته اول روش‌هایی است که بر اساس داده‌های حاصل از متغیرهای مختلف مثل روش‌های رگرسیونی بدست می‌آیند و دسته دوم روش‌هایی که بر اساس فرآیند مورد نظر حاصل می‌گردند. مطابق تقسیم‌بندی گویندراجو [10] مدل‌هایی که امروزه در هیدرولوژی مورد استفاده قرار می‌گیرند، شامل مدل‌های ریاضی-فیزیکی، مدل‌های ژئومورفولوژیکی و مدل‌های تجربی می‌باشند. دسته اول مدل‌ها، براساس خصوصیات فیزیکی سیستم هستند که به صورت معادلات دیفرانسیل ارائه می‌شوند. اما دسته دوم بر مبنای خصوصیات ژئومورفولوژیکی سیستم هیدرولوژیکی مورد نظر بیان می‌گردد. مدل‌های تجربی بدون توجه به پارامترها، سعی در ایجاد رابطه‌ای بین داده‌های ورودی و خروجی دارند که این مدل‌ها به مدل‌های جعبه سیاه معروف هستند. برنامه‌ریزی ژنتیک^۱ برای اولین بار توسط کوزا [12] ارائه گردید. این روش جزو روش‌های الگوریتم گردشی محسوب می‌شود که مبنای آنها براساس نظریه تکامل داروین استوار است. الگوریتم‌های یاد شده اقدام به تعریف یک تابع هدف در قالب معیارهای کیفی نموده و سپس تابع یاد شده را برای مقایسه جواب‌های مختلف حل مسئله در یک فرآیند گام به گام تصحیح ساختار داده‌ها به کار می‌گیرند و در نهایت، جواب مناسب را ارائه می‌نمایند [2]. روش برنامه‌ریزی بیان ژن^۲ جدیدترین شیوه از بین روش‌های الگوریتم تکاملی می‌باشد که به دلیل دارا بودن دقت کافی، از کاربرد بیشتری برخوردار است [9]. از این روش به طور موفقیت آمیزی می‌توان در شرایطی که تشخیص ارتباط داخلی میان متغیرهای وابسته بسیار پیچیده باشد، یافتن اندازه و یا شکل متغیر نهایی بسیار سخت بوده و بخش عمده مسئله را تشکیل دهد، قوانین ریاضی متداول قادر به ارائه روش‌های حل تحلیل نباشند ولی حصول جواب‌های تقریبی کیفیت کند و اصلاحات جزئی و متداول اجرایی بر روی نتایج به راحتی قابل لحاظ و اندازه‌گیری باشد، می‌توان بهره جست [3]. خو و همکاران [11] رواناب ساعتی حوضه آرگوال فرانسه را با استفاده از روش برنامه‌ریزی ژنتیک، پیش‌بینی نمودند. بررسی نتایج نشان داد که روش مذکور، توانایی پیش‌بینی کاملاً دقیق رواناب را در تمام فواصل زمانی (به ویژه فواصلی که کوتاه‌تر از زمان تمرکز حوضه هستند) را دارد. همچنین نتایج با مطالعه انجام شده که در آن روش‌های خودهمبستگی و فیلترکالمن استفاده شده، مقایسه گردید و نتایج تحقیق نشان داد که برنامه ریزی ژنتیک ابزار بهتری برای پیش‌بینی رواناب است. دوگان و همکاران [5] از دو روش شبکه‌های عصبی مصنوعی و روش اتورگرسیون برای پیش‌بینی جریان طغیانی روزانه در حوضه ساکارا استفاده کردند. نتایج حاکی از دقت بالای مدل شبکه‌های عصبی مصنوعی بازگشت کننده در پیش‌بینی جریان طغیانی روزانه می‌باشد. آیتک و کیشی [4] روش جدید GEP را برای فرمول‌بندی رابطه رسوبات معلق و دبی روزانه در دو ایستگاه رودخانه تانگو در مونتانا به کار گرفتند. مدل‌های بدست آمده با استفاده از GEP با دو روش رگرسیون خطی و نمودارهای شدت رسوب در تخمین بار رسوبات معلق مقایسه شد. نتایج نشان داد که فرمول به دست آمده از GEP، در مقایسه با دو روش دیگر از کارایی بیشتری برخوردار است و توانایی این روش را در ایجاد ابزار مفید در حل مسائل ویژه در هیدرولوژی از قبیل تخمین رسوبات معلق تأیید می‌کند. قربانی و همکاران [9] پیش‌بینی سطح آب دریا را با استفاده از سری-زمانی پیشین در غرب استرالیا در مقیاس‌های زمانی مختلف با استفاده از روش برنامه ریزی ژنتیک انجام دادند. نتایج نشان داد که هر دو روش عملکرد خوبی دارند. نوحه‌گر و همکاران [1] به منظور پیش‌بینی جریان روزانه رودخانه حوضه آبخیز معرف امامه از مدل‌های برنامه‌ریزی ژنتیک و شبکه عصبی مصنوعی استفاده کردند. این محققان گزارش نمودند که روش GP، از میان مدل‌های مختلف به کار رفته، خطای کمتری داشته و به عنوان روش مناسب جهت پیش‌بینی جریان روزانه معرفی گردید. مهر [13] با استفاده از مدل تلفیقی الگوریتم ژنتیک- برنامه‌ریزی بیان ژن جریان ماهانه رودخانه شاور واقع در حوضه آبخیز سفید رود را پیش‌بینی نمودند. نتایج نشان داد که هیبرید GEP و الگوریتم ژنتیک نسبت به مدل ساده GEP از عملکرد بسیار بهتری برخوردار بود. مهر و نورانی [14] با استفاده از روش برنامه‌ریزی بیان ژن چندگانه، فرآیند بارش و رواناب حوضه آبخیز هالدیزن واقع در کشور ترکیه را بررسی نموده و گزارش نمودند که مدل مذکور از دقت بسیار بالایی در برآورد جریان برخوردار است. با توجه به پیشینه پژوهش مشاهده می‌شود که پیش‌بینی جریان رودخانه توجه محققان بسیاری را به خود جلب نموده است. بنابراین در مطالعه حاضر سعی گردید تا با استفاده از روش برنامه‌ریزی بیان ژن جریان رودخانه چهل‌چای در مقیاس ماهانه پیش‌بینی گردد.

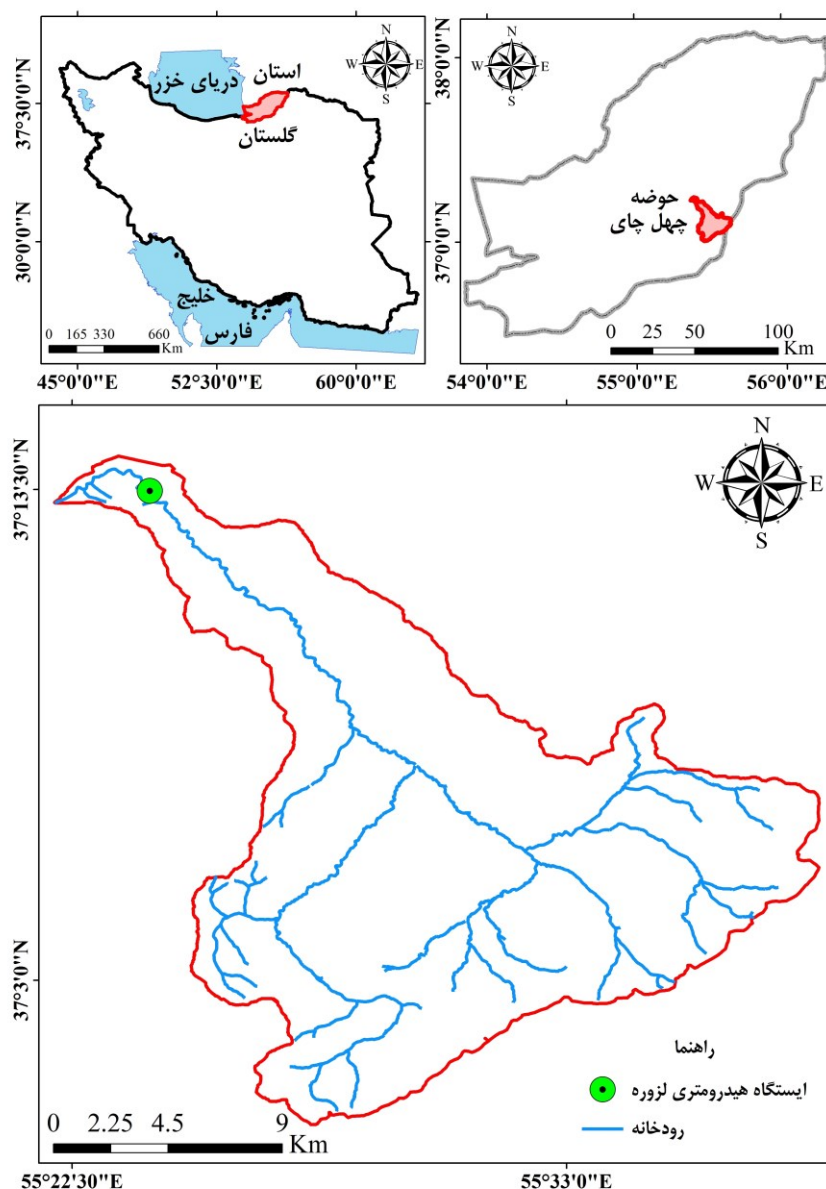
¹ - Genetic Programming

² - Genetic Expression Programming

۲. مواد و روش‌ها

۱.۲. داده‌ها و منطقه مورد مطالعه

حوضه آبخیز چهل‌چای از نظر موقعیت سیاسی در استان گلستان و در محدوده شهرستان مینودشت قرار گرفته است. این حوضه آبخیز جزئی از آبخیز دریای خزر و حوضه رود ارس است و بر اساس تقسیم‌بندی تماب جزو محدوده مطالعاتی گلستان با کد ۱۶۰۱ می‌باشد. در این مطالعه برای پیش‌بینی جریان ماهانه رودخانه چهل‌چای از داده‌های ایستگاه هیدرومتری لزوره واقع در خروجی حوضه استفاده شده است. این ایستگاه فعالیت خود را از سال ۱۳۴۵ آغاز نموده و تا ۱۳۹۶ اطلاعات آن در دسترس می‌باشد. به عبارت دیگر ۵۱ سال آمار ثبت شده در این ایستگاه موجود بوده و داده‌های مفقوده در آن مشاهده نمی‌شود. متوسط آورد این رودخان در دوره آماری مورد مطالعه ۲/۰۳ مترمکعب در ثانیه محاسبه شد. موقعیت مطالعاتی حوضه آبخیز در کشور و استان گلستان در شکل ۱ نشان داده شده است. همچنین مکان ایستگاه هیدرومتری لزوره نیز در این شکل مشخص شده است.



شکل ۱. موقعیت حوضه آبخیز چهل‌چای و ایستگاه هیدرومتری منتخب

۲.۲. برنامه ریزی بیان ژن

روش برنامه ریزی بیان ژن، در سال ۱۹۹۹ توسط فریرا [6] ابداع شد. این روش ترکیبی از روش‌های برنامه‌ریزی ژنتیک و الگوریتم ژنتیک بوده که در آن، کروموزوم‌های خطی و ساده با طول ثابت، مشابه با آنچه که در الگوریتم ژنتیک استفاده می‌شود و ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت، مشابه با درختان تجزیه در برنامه‌ریزی ژنتیک، ترکیب می‌شوند. از آنجایی که در این روش تمام ساختارهای شاخه‌ای با اندازه اشکال متفاوت، در کروموزوم‌های خطی با طول ثابت کد گذاری می‌شوند، معادل این است که در این روش فنوتیپ و ژنوتیپ از هم جدا می‌گردند و سیستم می‌تواند از تمام مزایای تکاملی به سبب وجود آنها بهره‌مند شود. اکنون با وجود اینکه فنوتیپ در GEP، همان نوع از ساختارهای شاخه‌ای مورد استفاده در GP را شامل می‌شود، اما ساختارهای شاخه‌ای (بیان درختی نیز نامیده می‌شود) که به وسیله GEP استنتاج می‌شوند مبین تمامی ژنوم‌های مستقل هستند. به طور خلاصه می‌توان نتیجه گرفت که در GEP بهسازی‌ها در یک ساختار خطی اتفاق افتاده و سپس به صورت ساختار درختی بیان می‌شود و این موجب می‌گردد تنها ژنوم اصلاح شده به نسل بعد منتقل شده و نیازی به ساختارهای سنگین برای تکثیر و جهش وجود نداشته باشد [8].

فرآیند گام به گام حل یک مسئله با استفاده از برنامه‌ریزی بیان ژن متشکل از ۵ مرحله به شرح زیر می‌باشد [9]:

- (۱) انتخاب مجموعه ترمینال: که همان متغیرهای مستقل مسئله و متغیرهای حالت سامانه می‌باشند.
 - (۲) انتخاب مجموعه توابع: که شامل عملگرهای حسابی، توابع آزمون و توابع بولی می‌باشد.
 - (۳) شاخص اندازه‌گیری دقت مدل که بر مبنای آن می‌توان مشخص نمود که توانایی مدل در حل یک مسئله خاص تا چه اندازه می‌باشد.
 - (۴) مولفه‌های کنترل: مقادیر مولفه‌های عددی و متغیرهای کیفی که برای کنترل اجرای برنامه‌ها استفاده می‌شوند.
 - (۵) شرط توقف اجرای برنامه: که معیاری برای حصول نتیجه و توقف اجرای برنامه می‌باشد.
- در مطالعه حاضر از برنامه GeneXpro Tools [6] برای توسعه و اجرای مدل‌های مبتنی بر برنامه‌ریزی ژنتیک استفاده به عمل آمد. برنامه مذکور بر اساس برنامه‌ریزی صریح ژنتیک (GEP) استوار است. برنامه‌ریزی بیان ژن ویرایش جدیدی از برنامه‌ریزی ژنتیک می‌باشد که به استنتاج برنامه‌های رایانه‌ای با اندازه‌ها و شکل‌های مختلف می‌پردازد. یکی از نقاط قوت GEP آن است که معیار تنوع ژنتیکی بسیار ساده بوده و لذا عملگرهای ژنتیک در سطح کروموزوم عمل می‌نمایند. همچنین یکی دیگر از نقاط قوت این روش، عبارت است از طبیعت منحصر به فرد چند ژنی بودن آن که زمینه ارزیابی مدل‌های پیچیده‌ای را که شامل چندین زیرمدل می‌باشند فراهم می‌آورد [6]. فرآیند مدل‌سازی جریان رودخانه چهل‌چای در تحقیق حاضر به شکل زیر صورت گرفت.
- گام اول، انتخاب تابع برازش مناسب می‌باشد که در این مطالعه تابع جذر میانگین مربعات خطا به عنوان تابع برازش انتخاب گردید. گام دوم، انتخاب مجموعه متغیرهای ورودی و مجموعه توابع به منظور تولید کروموزوم‌ها می‌باشد. در مسئله حاضر مجموعه ترمینال‌ها متشکل از مقادیر جریان رودخانه با تاخیرهای زمانی می‌باشد. در این مطالعه از چهار عملگر اصلی که شامل جمع، تفریق، ضرب و تقسیم بوده و نیز سایر توابع ریاضی استفاده گردید. گام سوم، شامل انتخاب ساختار و معماری کروموزوم‌ها می‌باشد. اندازه طول هر راس هشت و تعداد ژن‌ها سه عدد انتخاب شد. گام چهارم انتخاب تابع پیوندی است که در این مطالعه عمل جمع برای ایجاد پیوند بین زیرشاخه‌ها مورد استفاده قرار گرفت. در نهایت، در گام پنجم عملگرهای ژنتیک و نرخ هریک از آنها انتخاب می‌شود.

۳.۲. ارزیابی مدل

به منظور ارزیابی عملکرد مدل‌های به کار گرفته شده از سه معیار ضریب همبستگی (R)، جذر میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE) به شرح زیر استفاده گردید.

$$RMSE = \left[\frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)^2}{n} \right]^{0.5} \quad (1)$$

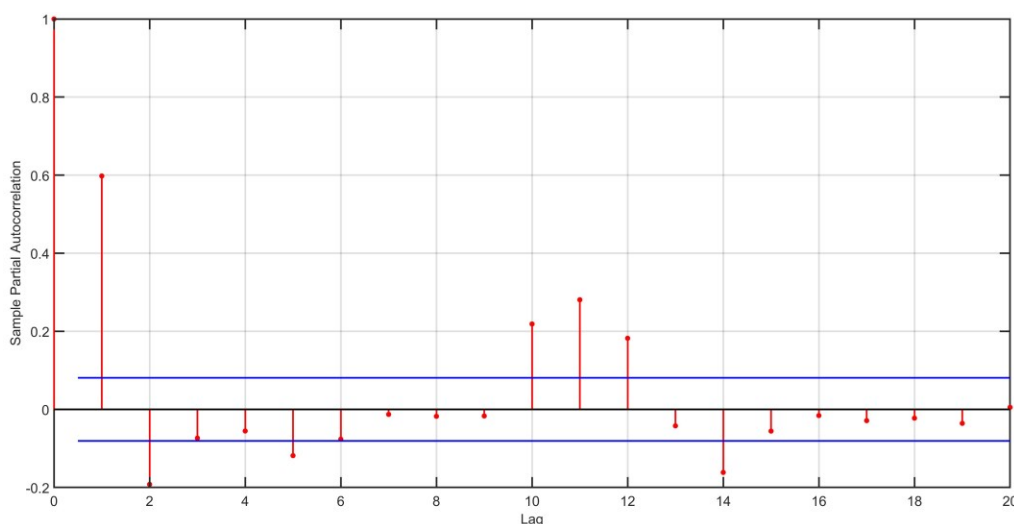
$$R = \sqrt{\frac{\left[\sum_{i=1}^n (Q_i - \bar{Q})(\hat{Q}_i - \bar{Q}) \right]^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2 \sum_{i=1}^n (\hat{Q}_i - \bar{Q})^2}} \quad (2)$$

$$MAE = \left| \frac{\sum_{i=1}^n (Q_i - \hat{Q}_i)}{n} \right| \quad (3)$$

که در روابط فوق Q_i ، $\hat{Q}_i$ ، $\bar{Q}$ به ترتیب دبی جریان مشاهداتی، دبی جریان محاسباتی، میانگین دبی جریان و n تعداد داده‌ها می‌باشد. مدلی که کمترین مقدار RMSE و MAE یا بیشترین مقدار ضریب همبستگی (R) را به خود اختصاص دهد به عنوان مدل مطلوب شناخته می‌شود.

۳. نتایج

در مطالعه حاضر برای مدل‌سازی جریان ماهانه رودخانه چهل‌چای با استفاده از مدل GEP، داده‌های ۴۱ سال (از مهر ۱۳۴۵ تا شهریور ۱۳۸۶) برای آموزش و ۱۰ سال (از مهر ۱۳۸۶ تا شهریور ۱۳۹۶) به عنوان داده‌های تست انتخاب شدند. به طور کلی ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد نیز برای تست در نظر گرفته شد. انتخاب جمعیت‌های اولیه تصادفی مختلف و تاثیرگذار در پدیده که در برنامه ریزی بیان ژن به عنوان داده‌های آموزشی از آن‌ها یاد می‌شود به منظور آموزش ماهیت سازوکار حاکم بر پدیده نه تنها سبب پیچیدگی الگو و افزایش حافظه درگیر خواهد شد، بلکه سبب کاهش دقت مدل نیز می‌شود (داننده مهر، ۱۳۸۹). لذا در الگوسازی جریان رودخانه‌ها نیز بایستی سعی نمود موثرترین داده‌های مشاهداتی را به عنوان داده‌های آموزشی انتخاب کرد. الگوهای مورد استفاده در مقیاس زمانی ماهانه در جدول ۱ آورده شده و جدول ۲ مشخصات برنامه‌ریزی بیان ژن به کار رفته را برای سری ماهانه نشان می‌دهد. با توجه به جدول ۱ برای پیش‌بینی جریان ماهانه رودخانه چهل‌چای، الگوهای ورودی براساس حافظه درونی داده‌ها و به صورت تاخیر پایه توسعه یافتند. برای انتخاب تاخیر بهینه نیز از نمودار PACF استفاده شد (شکل ۲). با توجه به شکل ۲ مشاهده می‌شود که استفاده از چهار تاخیر زمانی برای مدل‌سازی جریان ماهانه رودخانه چهل‌چای مناسب می‌باشد. بنابراین در جدول ۱ پیش‌بینی جریان ماهانه فقط به جریان یک ماه قبل، دو ماه قبل تا چهار ماه قبل وابسته است.



شکل ۲. نمودار PACF جریان ماهانه رودخانه چهل‌چای

جدول ۱. الگوهای ورودی مورد استفاده برای پیش‌بینی جریان رودخانه چهل‌چای در مقیاس زمانی ماهانه

الگوهای ورودی	ورودی‌ها	خروجی
P1	Qt-1	Qt
P2	Qt-1, Qt-2	Qt
P3	Qt-1, Qt-2, Qt-3	Qt
P4	Qt-1, Qt-2, Qt-3, Qt-4	Qt

جدول ۲. مشخصات برنامه‌ریزی بیان ژن به کار رفته برای سری ماهانه رودخانه چهل‌چای

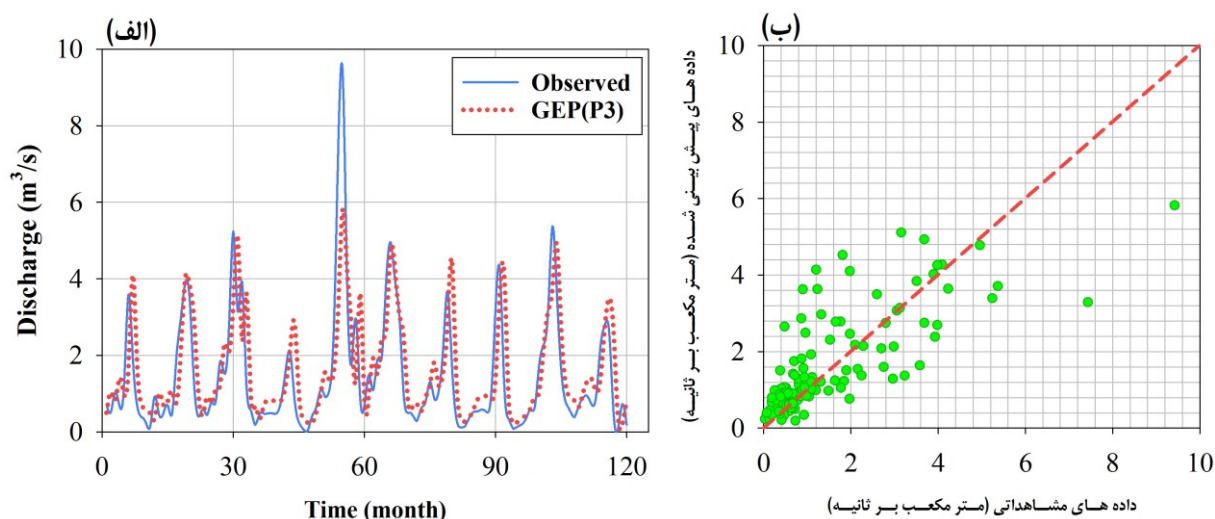
پارامتر	مقدار
اندازه سر	۸
تعداد کروموزوم‌ها	۳۰
تعداد ژن‌ها در هر کروموزوم	۳
معیار خطای تابع برازش	RMSE
تابع پیوند	جمع (+)

جدول ۳ نتایج ارزیابی الگوهای مختلف ورودی برنامه‌ریزی بیان ژن را برای مدل‌سازی جریان ماهانه رودخانه در مراحل آموزش و صحت-سنجی نشان می‌دهد. با توجه به این جدول می‌توان نتیجه گرفت که مدل شامل دبی سه ماه قبل به ترتیب با دارا بودن ضریب همبستگی و جذر میانگین مربعات خطای معادل با ۰/۷۶ و $1/0.3 (m^3/s)$ بیشترین دقت را در مرحله صحت‌سنجی به خود اختصاص داده و به عنوان بهترین الگو برای مدل برنامه‌ریزی بیان ژن انتخاب گردید. شکل (۳ الف و ب) نمودار مقادیر مشاهداتی و مقادیر پیش‌بینی شده حاصل از بهترین مدل برنامه‌ریزی بیان ژن (P3) برای سری ماهانه جریان رودخانه چهل‌چای، در مرحله صحت‌سنجی را نشان می‌دهد. با توجه به این شکل‌ها می‌توان نتیجه گرفت که برنامه‌ریزی بیان ژن در بهترین حالت خود به خوبی توانسته دبی جریان ماهانه رودخانه مورد مطالعه را پیش‌بینی کند. همان‌گونه که در شکل (۳ الف) مشاهده می‌شود، دقت نتایج مدل برنامه‌ریزی بیان ژن در پیش‌بینی دبی‌های کمتر، بیش از دبی‌های بزرگ‌تر می‌باشد.

جدول ۳. تحلیل‌های آماری نتایج برنامه‌ریزی بیان ژن برای الگوهای ورودی جریان ماهانه

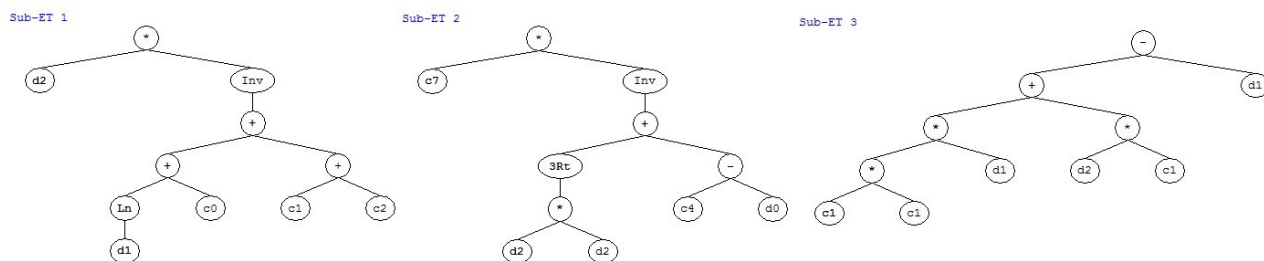
الگو	آموزش			صحت‌سنجی		
	R	MAE (m^3/s)	RMSE (m^3/s)	R	MAE (m^3/s)	RMSE (m^3/s)
P1	۰/۹۴	۱/۵۱	۰/۹۴	۰/۶۹	۱/۱۱	۰/۷۶
P2	۰/۸۷	۱/۵۲	۰/۸۷	۰/۶۸	۱/۰۷	۰/۷۵
P3	۰/۸۴	۱/۳۷	۰/۷۵	۰/۶۸	۱/۰۳	۰/۶۸
P4	۰/۸۹	۱/۵۱	۰/۶۸	۰/۷۳	۱/۱۰	۰/۷۴

یکی از قابلیت‌های مهم روش برنامه‌ریزی ژن، ارائه رابطه ریاضی برای متغیرهای هیدرولوژیکی مورد بررسی می‌باشد. در این مطالعه نیز رابطه ریاضی به دست آمده از برنامه‌ریزی ژن به همراه ساختار آن ارائه شده است. شکل ۴ ساختار درختی هریک از ژن‌ها را نشان می‌دهد. در این ساختار C0 و C1 ضرایب برنامه‌ریزی ژنتیک و d0، d1 و d2 به ترتیب معادل دبی یک، دو و سه ماه قبل می‌باشند. با استفاده از این ساختار درختی می‌توان شکل ریاضی جریان ماهانه رودخانه چهل‌چای را به دست آورد.



شکل ۳. نمودارهای الف) گرافیکی و ب) پراکندگی مقادیر مشاهداتی و پیش‌بینی شده حاصل از بهترین مدل برنامه‌ریزی ژن در مرحله

صحت‌سنجی



۴. نتیجه‌گیری

پیش‌بینی جریان رودخانه یکی از فاکتورهای مهم و تأثیرگذار در مدیریت منابع آب حوضه‌های آبخیز می‌باشد. تحلیل و بررسی این متغیر کاربردهای فراوانی داشته و می‌تواند در مدیریت سیلاب و پیش‌بینی آن کمک شایانی نماید. بنابراین، در این مطالعه جریان ماهانه رودخانه چهل‌چای در دوره آماری ۱۳۴۵ تا ۱۳۹۶ با استفاده از روش برنامه‌ریزی بیان ژن مدل‌سازی شد. به طور کلی نتایج حاصل از این مطالعه به شرح زیر ارائه می‌شود:

- برای سری ماهانه بهترین عملکرد برنامه‌ریزی بیان ژن با سه تاخیر زمانی در ورودی‌های مدل به دست آمد و ضریب همبستگی و میانگین مربعات مجذور خطا برای این الگو به ترتیب ۰/۷۷ و ۱/۰۳ متر مکعب بر ثانیه محاسبه گردید.
- با دقت بیشتر به شکل‌های ارائه شده برای روش GEP، می‌توان نتیجه گرفت که عملکرد روش مورد بررسی از نظر گرافیکی نزدیک به واقعیت می‌باشد اما از نقطه نظر تخمین دبی‌های کمینه و بیشینه، مشخص است که مدل GEP در پیش‌بینی دبی بیشینه خطای بیشتری داشته است اما در پیش‌بینی دبی متوسط و کمینه عملکرد آن به مراتب بهتر می‌باشد.

منابع

- [1] احمد نوحه گر، محبوبه معتمدنیا، آرش ملکیان، ۱۳۹۵. مدل‌سازی جریان روزانه رودخانه با استفاده از برنامه‌ریزی ژنتیک و شبکه عصبی (مطالعه موردی: حوضه آبخیز معرف امامه). پژوهش‌های جغرافیای طبیعی، ۴۸(۳)، ۳۶۷-۳۸۳.
- [2] فرشاد احمدی، یعقوب دین‌پژوه، احمد فاخری فرد، کیوان خلیلی، محمد علی قربانی، ۱۳۹۱. ارزیابی دقت برنامه‌ریزی ژنتیک و شبکه عصبی مصنوعی در پیش‌بینی جریان روزانه رودخانه (مطالعه موردی رودخانه باراندوزچای). نهمین کنفرانس بین‌المللی مهندسی رودخانه، ۳ تا ۵ بهمن، دانشگاه شهید چمران، اهواز.
- [3] محمد علی قربانی، جلال شیر، هانیه کاظمی. ۱۳۸۹. تخمین بیشینه، متوسط و کمینه دمای شهر تبریز با استفاده از روش‌های هوش مصنوعی. مجله دانش آب و خاک (دانشگاه تبریز). ۲۰(۱)، ۸۷-۱۰۴.
- [4] Aytek, A. and O., Kisi. 2008. A genetic programming approach to suspended sediment modeling. *Journal of Hydrology*, 351: 288-298.
- [5] Doğan, E., Işık, S., Toluk, T., & Sandalcı, M. (2007, March). Daily streamflow forecasting using artificial neural networks. In *International Congress on River Basin Management* (pp. 22-24).
- [6] Ferreira, C. 2001. Gene expression programming: A new adaptive algorithm for solving problems. *Complex Systems*, 13(2), 87-129.
- [7] Ferreira, C. 2002. Genetic representation and genetic neutrality in gene expression programming. *Advances in Complex Systems*, 5(4), 389-408.
- [8] Ferreira, C. 2004. Gene expression programming and the evolution of computer programs. *Recent developments in biologically inspired computing*, 82-103.
- [9] Ghorbani, M. A., Kisi, O., & Aalinezhad, M. 2010. A probe into the chaotic nature of daily streamflow time series by correlation dimension and largest Lyapunov methods. *Applied Mathematical Modelling*, 34(12), 4050-4057.
- [10] Govindaraju, R. S. 2000. Artificial neural network in hydrology. *Journal of hydrologic Engineering*, 5 (2): 115-123.
- [11] Khu, S. T., Liong, S. Y., Babovic, V., Madsen, H. and N., Muttil 2001. Genetic programming and its application in real-time runoff forecasting. *Journal of American Water Resources Association*, 37 (2): 439-451.
- [12] Koza, J. R. 1992. *Genetic Programming: on the programming of computers by means of natural selection*. Cambridge, MA: MIT Press.
- [13] Mehr, A. D. 2018. An improved gene expression programming model for streamflow forecasting in intermittent streams. *Journal of hydrology*. 563: 669-678.
- [14] Mehr, A. D., and Nourani, V. 2018. Season algorithm-multigene genetic programming: a new approach for rainfall-runoff modelling. *Water resources management*. 32:8. 2665-2679.