نشریه پژوهش‌های حفاظت آب و خاک
جلد بیست و دوم، شماره اول، ۱۳۹٤
http://jwsc.gau.ac.ir

مقایسه مدل‌های غیرخطی سری زمانی و برنامه‌ریزی ژنتیک در پیش‌بینی جریان روزانه رودخانه‌ها (مطالعه موردی: رودخانه باراندوزچای ارومیه)

[*]فرشاد احمدی[۱]، یعقوب دین‌پژوه[۲]، احمد فاخری‌فرد[۳]، کیوان خلیلی[۴] و صابره دربندی[۵]

[۱]دانشجوی کارشناسی‌ارشد گروه مهندسی آب، دانشگاه تبریز، [۲]دانشیار گروه مهندسی آب، دانشگاه تبریز، [۳]استاد گروه مهندسی آب، دانشگاه تبریز، [۴]استادیار گروه مهندسی آب، دانشگاه ارومیه، [۵]استادیار گروه مهندسی آب، دانشگاه تبریز

تاریخ دریافت: ۹۲/٤/۳۰ ؛ تاریخ پذیرش: ۹۲/۱۱/۲۷

چکیده

در این مطالعه برای پیش‌بینی جریان روزانه رودخانه باراندوزچای ارومیه در دوره آماری ۸۸-۱۳۵۲، از مدل غیرخطی سری زمانی دوخطی و روش برنامه‌ریزی ژنتیک استفاده شده و نتایج براساس شاخص‌های آماری جذر میانگین مربعات خطا و ضریب همبستگی مورد مقایسه قرار گرفته است. در این مطالعه مدل دوخطی BL(1,11,1,1) با داشتن کم‌ترین مقدار معیار اکایکه اصلاح‌شده به‌عنوان مدل مناسب سری روزانه انتخاب و پس از انجام آزمون نکویی برازش الگوی نام‌برده برای پیش‌بینی جریان روزانه رودخانه باراندوزچای به‌کار برده شد. نتایج نشان داد که این الگو از توانایی لازم برای مدل‌سازی جریان روزانه برخوردار است. ضریب همبستگی و جذر میانگین مربعات خطا برای مدل BL(1,11,1,1) در مرحله صحت‌سنجی به‌ترتیب معادل ۰/۹۰۲ و ۳/۵۲۰ مترمکعب بر ثانیه به‌دست آمد. به‌منظور مدل‌سازی جریان روزانه رودخانه با برنامه‌ریزی ژنتیک از حافظه‌های دبی یک روز قبل، دو روز قبل تا چهار روز قبل استفاده شد. نتایج نشان داد که تا حافظه سه روز قبل دقت مدل‌سازی رو به بهبود بوده و بعد از آن کاهش می‌یابد. ضریب همبستگی و جذر میانگین مربعات خطا در بهترین حالت برنامه‌ریزی ژنتیک در مرحله صحت‌سنجی به‌ترتیب ۰/۹۲۸ و ۲/۸٦۳ مترمکعب بر ثانیه می‌باشد. به‌طورکلی با مقایسه نتایج به‌دست آمده از دو الگوی نام‌برده می‌توان نتیجه گرفت که روش برنامه‌ریزی ژنتیک با ۲۲/۹ درصد خطای کم‌تر نسبت به مدل دوخطی، جریان روزانه رودخانه باراندوزچای را پیش‌بینی می‌کند.

واژه‌های کلیدی: آبدهی، برنامه‌ریزی ژنتیک، پیش‌بینی، دوخطی، صحت‌سنجی

[*] مسئول مکاتبه: farshad.paper@yahoo.com

مقدمه

نیاز روزافزون به آب سبب گردیده است که برنامه‌ریزی‌های مدیریتی به‌منظور کنترل مصرف آب در آینده از اهمیت بیش‌تری برخوردار باشد. با پیش‌بینی نمودن جریان رودخانه‌ها علاوه‌بر مدیریت بهره‌برداری از منابع آب، می‌توان حوادث طبیعی مانند سیل و خشکسالی را نیز پیش‌بینی و مهار نمود. همچنین می‌توان از نتایج یک مدل شبیه‌سازی به‌منظور بررسی صحت داده‌ها و یا اصلاح و تکمیل آن‌ها استفاده کرد. مطابق تقسیم‌بندی گووینداراجو (۲۰۰۰) مدل‌هایی که امروزه در هیدرولوژی مورد استفاده قرار می‌گیرند، شامل مدل‌های ریاضی- فیزیکی، مدل‌های ژئومورفولوژیکی و مدل‌های تجربی می‌باشند. مدل‌های سری زمانی به‌عنوان یک مدل ریاضی- فیزیکی توانایی زیادی در مدل‌سازی پدیده‌های خطی و غیرخطی دارند. مدل سری زمانی از دو بخش اصلی، شامل مؤلفه تصادفی و مؤلفه جبری مدل تشکیل شده که مؤلفه جبری مدل با استفاده از ارقام مشاهداتی و مؤلفه تصادفی با استفاده از روش‌های مختلف استوکاستیک[1] به‌دست می‌آید. بنابراین ساختار مدل‌های سری زمانی می‌تواند با ساختار سری‌های هیدرولوژیکی در صورت انتخاب درست مدل و محاسبات صحیح آن سازگاری و مطابقت ویژه‌ای داشته باشد (سالاس، ۱۹۹۳). بیش‌تر مدل‌های سری زمانی رایج در هیدرولوژی و منابع آب تحلیل مدل‌های خطی است (ناوه و همکاران، ۲۰۱۲). مدل‌های غیرخطی بیش‌تر در علوم مرتبط با آمار، اقتصاد و ریاضیات مورد بحث قرار گرفته و توسعه یافته‌اند و کم‌تر در منابع آب استفاده شده است. در حالی‌که بسیاری از فرآیندهای مربوط به سیستم‌های طبیعی نسبت به زمان غیرخطی بوده اگرچه جنبه‌های خاصی از این سیستم‌ها ممکن است نسبت به جنبه‌های دیگر به فرآیند خطی نزدیک‌تر باشند (تسونیس، ۲۰۰۱). وانگ و همکاران (۲۰۰٦) با استفاده از آزمون BDS و روش همبستگی نمایی، غیرخطی بودن فرآیند جریان رودخانه را در مقیاس‌های زمانی مختلف مورد بررسی قرار داده و به این نتیجه رسیدند که هرچه مقیاس زمانی کوچک‌تر می‌شود، غیرخطی بودن افزایش می‌یابد، به‌طوری‌که دبی‌های روزانه ماهیت غیرخطی بیش‌تری نسبت به دبی‌های ماهانه نشان دادند. خلیلی و همکاران (۲۰۱۱ و ۲۰۱۳) با استفاده از آزمون BDS، میزان غیرخطی بودن سری‌های سالانه، ماهانه و روزانه رودخانه‌های شهرچای و باراندوزچای ارومیه را مورد بررسی قرار دادند. نتایج نشان داد که سری‌های سالانه هر دو رودخانه، خطی بوده و سری‌های ماهانه و روزانه، غیرخطی می‌باشند و با کوچک‌تر شدن مقیاس زمانی، شدت غیرخطی بودن جریان رودخانه افزایش شدیدی را نشان می‌دهد.

1- Stochastic

با توجه به پژوهش‌های صورت گرفته در این مطالعه از مدل غیرخطی دوخطی[1] برای مدل‌سازی جریان روزانه رودخانه باراندوزچای که خاصیت غیرخطی دارد (خلیلی و همکاران، ۲۰۱۳)، استفاده گردید. مدل دوخطی توسط گرانگر و آندرسون (۱۹۷۸) معرفی شد و پژوهش‌های فراوانی پس از آن بر روی این مدل به‌عمل آمد. سوبارائو و گابر (۱۹۸٤) بر روی برخی خصوصیات و کاربردهای این مدل مباحثی انجام داده و نمایش ماتریسی و فضایی آن را ارایه نمودند. دای و بیلارد (۲۰۰۳) مشکل تخمین پارامترهای مدل فضایی دوخطی را مورد توجه قرار داده و با فرض ایستا بودن سری، روشی به‌نام تخمین شرطی حداکثر درست نمایی با استفاده از الگوریتم عددی نیوتن رافسون ارایه نمودند. لیفشیتس (۲۰۰٦) منشأ تغییرات مدل دوخطی در سری‌های زمانی با ماهیت غیرخطی ضعیف را مورد بررسی قرار داد و به روابط و معادلات ریاضی حاکم بر این شرایط دست یافت. در ادامه به برخی از مهم‌ترین مطالعات مربوط به مدل‌سازی غیرخطی سری‌های زمانی اشاره می‌گردد.

وانگ و همکاران (۲۰۰۵) از ترکیب مدل‌های ARMA[2] و GARCH[3] برای برازش واریانس و میانگین روزانه جریان رودخانه زرد در چین استفاده کردند. نتایج این مطالعه نشان داد که مدل ARMA-GARCH نتایج بسیار سودمندی در مدل‌سازی سری روزانه جریان رودخانه ارایه می‌کند. کایادو (۲۰۰۷) عملکرد مدل‌های تک‌پارامتری سری‌های زمانی را در پیش‌بینی میزان آب مصرفی در مقیاس‌های روزانه و هفتگی اسپانیا از سال ۲۰۰۱ تا ۲۰۰٦ مورد بررسی قرار داد. در این پژوهش مدل‌های ARIMA و GARCH بر روی سری داده‌های مشاهداتی برازش داده شده و کارآیی این مدل‌ها مورد ارزیابی و تأیید قرار گرفت. ناوه و همکاران (۲۰۱۲) از مدل دوخطی برای پیش‌بینی جریان ماهانه رودخانه‌های شهرچای و باراندوزچای ارومیه استفاده کردند. نتایج حاصل نشان‌دهنده دقت بالای این مدل در مقایسه با مدل‌های خطی می‌باشد. همچنین کارایی مدل غیرخطی دوخطی در پیش‌بینی جریان رودخانه مورد تأیید قرار گرفت.

برنامه‌ریزی ژنتیک[4] یک تکنیک خودکار می‌باشد که راه‌حل مسأله را با استفاده از برنامه‌نویسی کامپیوتری ارایه می‌کند. الگوریتم‌های تکاملی که برنامه‌ریزی ژنتیک نیز عضوی از آن‌ها می‌باشد توانایی مدل‌سازی فرایندهای کاملاً غیرخطی و پویا را دارند. بنا به اهمیت موضوع، تاکنون پژوهشگران مختلفی در سراسر جهان اقدام به مطالعه جریان رودخانه‌ها نموده‌اند. خو و همکاران (۲۰۰۱) رواناب

1- Bilinear Model
2- Autoregressive Moving Average
3- Generalized Autoregressive Conditional Heteroskedasticity (GARCH)
4- Genetic Programming

ساعتی حوضه اُرگوال فرانسه را با استفاده از روش برنامه‌ریزی ژنتیک، پیش‌بینی نمودند. بررسی نتایج نشان داد که برنامه‌ریزی ژنتیک توانایی پیش‌بینی دقیق رواناب را در تمام فواصل زمانی ملاحظه شده و به‌ویژه فواصلی که کوتاه‌تر از زمان تمرکز حوضه هستند را داراست. همچنین نتایج با مطالعه انجام شده دیگر که در آن روش‌های خود همبستگی و فیلترکالمن استفاده شده بود، مقایسه و مشخص گردید که برنامه‌ریزی ژنتیک ابزار بهتری برای پیش‌بینی رواناب است. همچنین بر طبق پژوهش‌های آیتک و آسه (۲۰۰۸) روش برنامه‌ریزی ژنتیک یک روش مناسب و علمی در پیش‌بینی روابط بارش- رواناب می‌باشد. گوون (۲۰۰۹) برای پیش‌بینی دبی جریان روزانه رودخانه شویل‌کیل در ایالات متحده از روش برنامه‌ریزی ژنتیک خطی استفاده کرد و نتایج حاصل را با دو الگو از روش‌های شبکه عصبی مورد مقایسه قرارداد. وی نشان داد که هر دو روش نتایج قابل‌قبولی داشته است ولی روش برنامه‌ریزی ژنتیک خطی از دقت بالاتری نسبت به روش‌های شبکه عصبی برخوردار است. قربانی و همکاران (۲۰۱۰) عملکرد سه روش برنامه‌ریزی ژنتیک، شبکه عصبی مصنوعی[1] و نروفازی[2] را در روندیابی سیلاب رودخانه قزل‌ایرماق ترکیه مورد ارزیابی قرار دادند. نتایج مطالعه آن‌ها نشان داد از بین سه روش نام‌برده مدل برنامه‌ریزی ژنتیک با دقت بیش‌تری هیدروگراف خروجی را شبیه‌سازی می‌کند.

با بررسی منابع موجود مشخص شد که تاکنون از مدل غیرخطی دوخطی برای مدل‌سازی جریان روزانه رودخانه استفاده نشده است و به‌طورکلی در مهندسی رودخانه کاربرد بسیار محدودی داشته است. همچنین خلیلی و همکاران (۲۰۱۱ و ۲۰۱۳) و ناوه و همکاران (۲۰۱۲) نشان دادند که جریان روزانه رودخانه باراندوزچای خاصیت غیرخطی داشته و کارایی مدل غیرخطی دوخطی در پیش‌بینی جریان ماهانه این رودخانه تأیید شده است. در این مطالعه و در ادامه پژوهش‌های صورت گرفته جریان روزانه رودخانه باراندوزچای با استفاده از الگوی دوخطی مدل‌سازی شده و نتایج به‌دست آمده از آن با مدل برنامه‌ریزی ژنتیک مقایسه می‌شود تا توانایی مدل دوخطی در مدل‌سازی متغیرهای غیرخطی مورد ارزیابی قرار گیرد.

مواد و روش‌ها

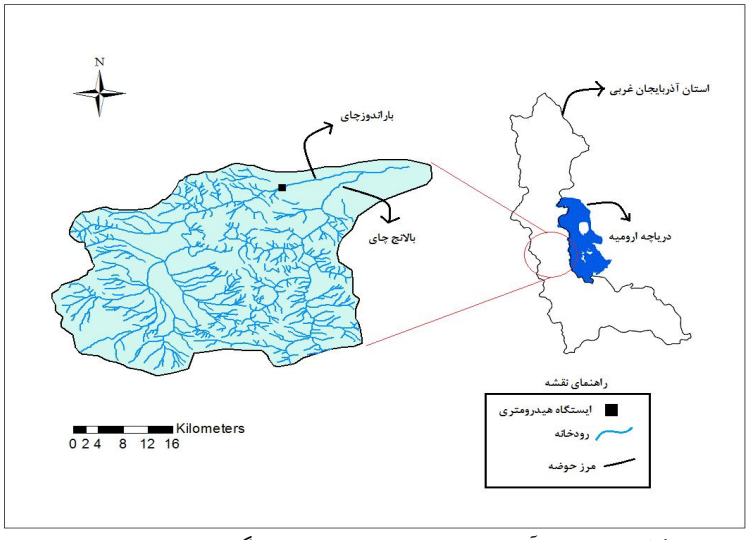
منطقه مورد مطالعه و داده‌های مورد نیاز: در این مطالعه از داده‌های دبی جریان روزانه ایستگاه هیدرومتری دیزج واقع در رودخانه باراندوزچای ارومیه در طول دوره آماری ۸۸-۱۳۵۲ استفاده

1- Artificial Neural Network (ANN)
2- Neuro-Fuzzy

گردید. مساحت حوضه باراندوزچای در ایستگاه هیدرومتری دیزج ٦٦٠/٧١ کیلومترمربع است. این حوضه در شمال‌غرب کشور بین دریاچه ارومیه و مرز ایران و کشورهای عراق و ترکیه واقع شده است. موقعیت جغرافیایی این حوضه از ٤٤ درجه ٤٥ دقیقه طول شرقی و ٣٧ درجه ٦ دقیقه تا ٣٧ درجه ٢٩ دقیقه عرض شمالی است. طول آبراهه اصلی حوضه ٧٥ کیلومتر بوده، حداکثر و حداقل ارتفاع این حوضه به‌ترتیب ٣٥٠٠ و ١٢٥٠ متر ازسطح آزاد آب‌ها می‌باشد. در جدول ١ مشخصات جغرافیایی و آماری ایستگاه هیدرومتری دیزج آورده شده است. شکل ١ نیز حوضه آبریز باراندوزچای و موقعیت ایستگاه هیدرومتری دیزج را نشان می‌دهد.

جدول ١- مشخصات آماری سری جریان روزانه رودخانه باراندوزچای.

ایستگاه هیدرومتری	دوره آماری	طول جغرافیایی	عرض جغرافیایی	تعداد داده‌ها	مشخصات آماری دبی روزانه	مترمکب بر ثانیه
					میانگین	٧/٩٤
					واریانس	٨٦/١٥
دیزج	١٣٥٢-٨٨	٤٠-٤٥	٢٣-٣٧	١٣٥٠٤	انحراف معیار	٩/٢٨
					مقدار ماکزیمم	١٦٨/٥٩
					مقدار مینیمم	٠/٠٠



شکل ١- حوضه آبریز باراندوزچای و موقعیت ایستگاه هیدرومتری دیزج.

مدل دوخطی: با توجه به این‌که مدل ARMA رایج‌ترین مدل کاربردی در پیش‌بینی سری‌های زمانی هیدرولوژیکی است و حالت خاصی از مدل غیرخطی دوخطی محسوب می‌شود، در ابتدا شکل کلی این مدل به‌صورت زیر ارایه می‌گردد:

$$ARMA(p,q):\ \ Z_t = \sum_{i=1}^{P}(\varphi_i.Z_{t-i}) - \sum_{j=1}^{q}(\theta_j.\varepsilon_{t-j}) + \varepsilon_t \qquad (1)$$

که در آن، Z_t: سری زمانی نرمال و استاندارد، φ_j: ضرایب مدل خودهمبستگی (AR) و ε_t: سری تصادفی و q و p مرتبه مدل می‌باشند.

مدل‌های خطی سری زمانی در واقع بسط مرتبه اول سری‌های تیلور[1] می‌باشند. ایده اصلی مدل دوخطی نیز غیرخطی بودن بسط مرتبه دوم سری تیلور می‌باشد. شکل کلی الگوی دوخطی به‌صورت زیر است (تسای، ۲۰۰۲).

$$Z_t = \sum_{i=1}^{P}(\varphi_i.Z_{t-i}) - \sum_{j=1}^{q}(\theta_j.\varepsilon_{t-j}) + \sum_{i=0}^{r}\sum_{j=1}^{s}(\beta_{ij}.Z_{t-i-j}.\varepsilon_{t-j}) + \varepsilon_t \qquad (2)$$

که در آن، Z_t: سری زمانی مورد نظر و p، q، r و s اعداد صحیح مثبت می‌باشند که رسته یا مرتبه مدل دوخطی را نشان می‌دهند. مدل بالا در برخی منابع مانند فان و یائو (۲۰۰۳) به‌صورت مدل $BL(p,q,r,s)$ نیز نمایش داده می‌شود. مدل دوخطی در واقع مدل ARMA (خطی) بسط داده شده می‌باشد که عبارت غیرخطی $\sum_{i=0}^{r}\sum_{j=1}^{s}(\beta_{ij}.Z_{t-i-j}.\varepsilon_{t-j})$ به‌سمت راست آن اضافه شده است (آینکاران، ۲۰۰٤). در این عبارت حاصل‌ضرب دو متغیر ε_{t-j} و Z_{t-i-j} که هر دو نبست به زمان متغیر هستند، باعث شده معادله از حالت خطی خارج شده و مدل غیرخطی شود. برازش یک مدل سری زمانی غیرخطی دوخطی شامل دو جنبه مهم است: یکی تعیین مرتبه مدل شامل (p,q,r,s) و دیگری تخمین ضرایب φ,θ,β می‌باشد. تعیین رسته مدل با استفاده از روش‌های شناخته شده مانند معیار آکایکه اصلاح شده (AICC) انجام می‌گیرد. مقدار AICC برای مرتبه‌های مختلف از رابطه زیر محاسبه شده و هر مرتبه که مقدار آکایکه کم‌تری داشته باشد، به‌عنوان رسته مدل انتخاب خواهد شد (آکایکه، ۱۹۷٤).

$$AICC(p,q) = nLn(\hat{\sigma}_\varepsilon^2) + \frac{2(p+q+r+s)}{(n-p-q-r-s)} \qquad (3)$$

1- Taylor Series

که در آن، n: تعداد داده‌ها، p: مرتبه مدل در بخش خود همبسته، q: مرتبه مدل در بخش میانگین متحرک و $\hat{\sigma}_{\varepsilon}^{۲}$: واریانس خطاها می‌باشد. البته هنوز عملکرد این روش‌ها در مدل دوخطی بررسی نشده و ناشناخته می‌باشد که این موضوع به‌دلیل نبود ارایه تئوری تخمین حداکثر درست‌نمایی برای مدل‌های بی‌لینیر می‌باشد (فان و یائو، ۲۰۰۳). با داشتن مرتبه مدل (p,q,r,s) بی‌لینیر می‌توان از روش استاندارد تابع لایکلیهود[1] برای تخمین ضرایب مدل استفاده کرد. فرض می‌کنیم $Z_{۱},...,Z_{t}$ سری داده‌های مشاهداتی باشند که به حالت ایستا تبدیل شده‌اند. فرض می‌کنیم $\theta=(\theta_{۱}^{\tau},\theta_{۲}^{\tau})^{\tau}$ باشد که در آن:

$$\theta_{۱}=(b_{۱},...,b_{p},a_{۱},...,a_{q})^{\tau} \qquad \theta_{۲}=(c_{۱۱},...,c_{۱s},c_{۲۱},...,c_{rs})^{\tau} \qquad (٤)$$

می‌توان تابع لگاریتم لایکلیهود (احتمالی)[2] را به‌صورت زیر نوشت (فان و یائو، ۲۰۰۳):

$$l(\theta,\sigma^{۲})=-\frac{N-p'}{۲}\log\sigma^{۲}-\frac{۱}{۲\sigma^{۲}}\sum_{t=p'+۱}^{T}\hat{\varepsilon}_{t}(\theta)^{۲} \qquad (٥)$$

که در آن، $p'=\max\{p,r\}$ و مقادیر $\hat{\varepsilon}_{p'}(\theta),\hat{\varepsilon}_{p'+۱}(\theta),...$ با در نظر گرفتن شرایط $q'=\max\{q,s\}$ و با استفاده از مدل عمومی دوخطی قابل محاسبه خواهند بود.

در این مطالعه برای بررسی نکویی برازش مدل‌های خطی و غیرخطی منتخب، از آزمون پورت مانتئو استفاده گردید. بدین منظور ابتدا سری باقی‌مانده مدل به‌دست آمده و برای تأخیر مشخص ضرایب خودهمبستگی سری محاسبه گردید. سپس با استفاده از رابطه زیر آماره پورت مانتئو (Q) محاسبه شد (سالاس، ۱۹۹۳):

$$Q=n\sum_{K=۱}^{L}r_{k}^{۲} \qquad (٦)$$

که در آن، n: تعداد داده، L: تعداد تاخیر، r_{k}: ضریب خود همبستگی مرتبه K ام و ε_{t}: سری باقی‌مانده مدل می‌باشد. آماره آزمون پورت مانتئو دارای توزیع کی‌دو با درجه آزادی $(L-p-q)$ است و در صورتی که $Q\le\chi_{(۱-\alpha),(L-p-q)}^{۲}$ باشد، مقادیر ε_{t} ناهمبسته بوده و مدل مانند آن کفایت لازم را دارد. سطح معنی‌داری نیز $\alpha=۰/۵$ در نظر گرفته شد.

1- Likelihood Function
2- Conditional Log Likelihood Function

در این مطالعه برای مدل دوخطی درجه آزادی به‌صورت $(L-p-q-r-s)$ و سری باقی‌مانده آن (ε_t) با معلوم بودن مرتبه و ضرایب مدل و با در نظر گرفتن مقدار اولیه صفر برای $\varepsilon_{t-1},\varepsilon_{t-2},...,\varepsilon_{t-q}$ به شرح زیر محاسبه گردید:

$$\varepsilon_t = Z_t - \left\{ \sum_{i=1}^{p} (\phi_i . Z_{t-i}) - \sum_{j=1}^{q} (\theta_i . \varepsilon_{t-i}) + \sum_{i=1}^{m} \sum_{j=1}^{s} (\beta_{ij} . Z_{t-i} . \varepsilon_{t-j}) \right\} \quad (۷)$$

برنامه‌ریزی ژنتیک: روش برنامه‌ریزی ژنتیک 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 مسأله در یک فرآیند گام به گام تصحیح ساختار داده‌ها به‌کار می‌گیرند و در نهایت، جواب مناسب را ارایه می‌نمایند. روش برنامه‌ریزی ژنتیک جدیدترین شیوه از بین روش‌های الگوریتم تکاملی می‌باشد که به‌دلیل دارا بودن دقت کافی، از کاربرد بیش‌تری برخوردار است (آلوییسی، ۲۰۰۵). تفاوت اساسی موجود بین الگوریتم ژنتیک و برنامه‌ریزی ژنتیک به طبیعت هر یک از افراد برمی‌گردد، به‌نحوی‌که افراد در الگوریتم ژنتیک، ردیف‌های خطی با طول ثابت می‌باشند (کروموزم‌ها) ولی در برنامه‌ریزی ژنتیک، همان شاخه‌های مجزا می‌باشند. همچنین در برنامه‌ریزی ژنتیک بر ساختار درختی مجموعه‌ها تأکید می‌شود ولی الگوریتم ژنتیک، براساس سیستم ارقام دودویی عمل می‌نماید. فرآیند گام به گام حل یک مسأله با استفاده از برنامه‌ریزی ژنتیک متشکل از ۵ مرحله به‌شرح زیر می‌باشد (قربانی و همکاران، ۲۰۱۰):

۱) انتخاب مجموعه ترمینال: که همان متغیرهای مستقل مسأله 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 مسأله خاص تا چه اندازه می‌باشد. ٤) مؤلفه‌های کنترل: مقادیر مؤلفه‌های 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 این مطالعه از برنامه GeneXpro Tools (فریرا، ۲۰۰۱) برای توسعه و اجرای مدل‌های مبتنی بر برنامه‌ریزی ژنتیک استفاده به‌عمل آمد. برنامه نام‌برده براساس 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 بنابراین 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 (فریرا، ۲۰۰۱). فرآیند مدل‌سازی جریان رودخانه باراندوزچای در این پژوهش به شکل زیر صورت گرفت.

گام اول، انتخاب تابع برازش مناسب می‌باشد که در این مطالعه تابع جذر میانگین مربعات خطا به‌عنوان تابع برازش انتخاب گردید. گام دوم، انتخاب مجموعه متغیرهای ورودی و مجموعه توابع به‌منظور تولید کروموزوم‌ها می‌باشد. در این مسأله مجموعه ترمینال‌ها متشکل از مقادیر جریان رودخانه با تاخیرهای زمانی می‌باشد. در این مطالعه از چهار عملگر اصلی که شامل $\{+,-,\times,\div\}$ و نیز توابع ریاضی $\{X^2, X^3, \sqrt[3]{X}, \sqrt{X}, Log(X)\}$ استفاده به‌عمل آمد. گام سوم، شامل انتخاب ساختار و معماری کروموزوم‌ها می‌باشد. اندازه طول هر راس ۷ و تعداد ژن‌ها ۳ عدد انتخاب شد. گام چهارم انخاب تابع پیوندی است که در این مطالعه عمل جمع برای برای ایجاد پیوند بین زیرشاخه‌ها مورد استفاده قرار گرفت. در نهایت، در گام پنجم عملگرهای ژنتیک و نرخ هر یک از آن‌ها انتخاب می‌شود.

اصولاً وارد کردن داده‌ها به‌صورت خام باعث کاهش سرعت و دقت برنامه‌ریزی ژنتیک می‌شود. برای پرهیز از چنین شرایطی و همچنین به‌منظور یکسان‌سازی ارزش داده‌ها، عمل بی‌بعدسازی صورت می‌گیرد. در این پژوهش همه داده‌های ورودی قبل از معرفی به شبکه با استفاده از رابطه ۸ بی‌بعد شده‌اند (داننده‌مهر و طباطبایی، ۲۰۱۰).

$$Q_n = ۰/۱ + ۰/۸ * \left(\frac{Q_. - Q_{\min}}{Q_{\max} - Q_{\min}} \right) \qquad (۸)$$

که در آن، Q_n: مقدار دبی بی‌بعد روزانه، $Q_.$: دبی روزانه مشاهده شده، $Q_{\max}$، $Q_{\min}$ به‌ترتیب مقادیر حداکثر و حداقل دبی‌های روزانه می‌باشند. رابطه ۷ نشان می‌دهد که دبی‌های مشاهداتی در بازه ۰/۹-۰/۱ بی‌بعد می‌شوند.

ارزیابی مدل: در این پژوهش برای ارزیابی مدل‌های مورد نظر از نمایه‌های ضریب همبستگی و جذر میانگین مربعات خطا استفاده می‌شود (رابطه‌های ۹ و ۱۰). در صورتی که مقدار ضریب همبستگی بالا و خطا کم باشد، می‌توان نتایج دقیق‌تر و قابل اعتمادتری به‌دست آورد.

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \overline{Q})^2}} \quad (۹)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{n}} \quad (۱۰)$$

که در آن‌ها، Q_i: مقدار مشاهده شده در گام زمانی i ام، $\hat{Q}_i$: مقدار محاسبه شده در همان زمان، n: تعداد داده‌ها و $\overline{Q}$: میانگین مقادیر مشاهداتی می‌باشد.

نتایج و بحث

نتایج مدل‌سازی غیرخطی سری روزانه جریان رودخانه باراندوزچای: در این قسمت از هر دو مدل خطی و غیرخطی استفاده شده و نتایج حاصل مورد مقایسه قرار گرفته است. در این مطالعه برای تخمین پارامترها از ۸۵ درصد داده (دوره آماری ۸۳-۱۳۵۲) استفاده گردید و ۱۵ درصد (دروه آماری ۸۸-۱۳۸٤) نیز برای تست مدل انتخاب شد. برای سری روزانه جریان رودخانه، مدل‌های ARMA(1,11) و BL(1,11,1,1) با داشتن کم‌ترین مقدار آکایکه اصلاح شده معادل ۳٦۷۳/۵۲ و ۳٦۷۵/۱۲ به‌عنوان مناسب‌ترین مدل خطی و غیرخطی انتخاب شدند. قبل از استفاده از مدل‌های منتخب برای برازش و پیش‌بینی جریان روزانه رودخانه مورد مطالعه، آزمون نکویی برازش انجام شد. جدول ۲ نتایج آزمون نکویی برازش سری روزانه جریان را با استفاده از روش پورت مانتئو نشان می‌دهد. با توجه به این جدول می‌توان نتیجه گرفت که مقدار آماره پورت مانتئو (Q) برای مدل‌های منتخب کم‌تر از مقدار جدول کای‌اسکور در سطح ۵ درصد می‌باشد. بنابراین، نتایج به‌دست آمده از این روش صحت و کفایت مدل‌های برازشی ARMA(1,11) و BL(1,11,1,1) را تأیید می‌نماید. با توجه به تأیید مدل‌ها، پارامترهای هر دو مدل محاسبه شده و از آن‌ها برای تولید آمار از سال آبی ۸۸-۱۳۸٤ (به‌مدت ۵ سال) استفاده گردید. شکل ۲ نمودار پراکندگی مقادیر مشاهداتی در مقابل مقادیر محاسباتی به‌دست آمده از مدل‌های ARMA(1,11) و BL(1,11,1,1) را برای سری روزانه جریان رودخانه مورد مطالعه نشان می‌دهد. در شکل ۳ نیز مقادیر مشاهداتی با مقادیر پیش‌بینی شده به‌دست

آمده از مدل BL(1,11,1,1) نسبت به زمان ترسیم شده است. نتایج ارزیابی عملکرد مدل‌های بسط داده شده در جدول ۳ ارایه گردیده است. با توجه به این جدول مشاهده می‌شود که مدل ARMA(1,11) نتوانسته به خوبی دبی جریان رودخانه باراندوزچای را مدل‌سازی کند که این امر به‌دلیل خاصیت غیرخطی مقیاس روزانه می‌باشد. دقت بالای مدل غیرخطی ناشی از تطابق ماهیت جریان رودخانه با مدل منتخب می‌باشد. به‌طوری‌که ضریب همبستگی و مقدار خطا در مدل BL(1,11,1,1) به‌ترتیب ۰/۹۰۲ و ۳/۵۲۰ مترمکعب بر ثانیه محاسبه گردید. جدول ٤ نیز رابطه مدل‌های خطی و غیرخطی جریان روزانه رودخانه باراندوزچای را نشان می‌دهد.

جدول ۲- نتایج آزمون پورت مانتئو مدل‌های خطی و غیرخطی برازشی جریان روزانه رودخانه باراندوزچای.

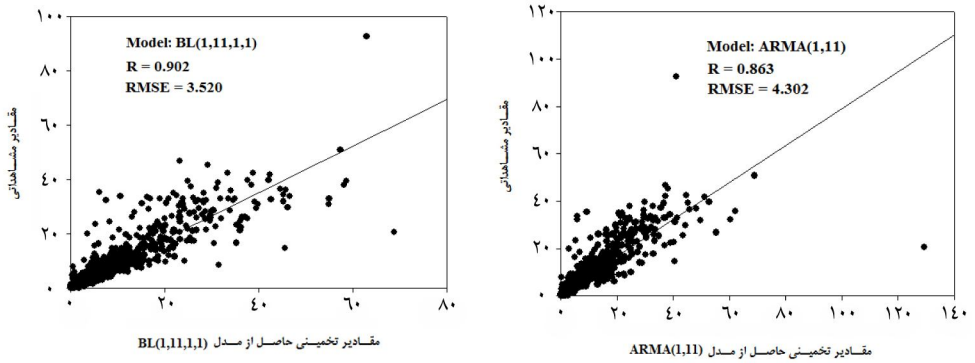
مدل	تعداد داده	تاخیر	درجه آزادی	سطح معنی‌داری	آزمون Q	مقدار جدول	نتیجه آزمون
ARMA(1,11)	۱۱٦۷۹	۳۵۰	۳٤۵	۵ درصد	۳۱۲/۵۳	۳۵۰/٦۰	مورد قبول
BL(1,11,1,1)	۱۱٦۷۹	۳۵۰	۳٤۳	۵ درصد	۳۳٤/۱۹	۳۵۰/٦۰	مورد قبول

جدول ۳- تحلیل آماری دقت مدل‌سازی خطی و غیرخطی جریان روزانه رودخانه باراندوزچای در مرحله صحت‌سنجی (۸۸-۱۳۸٤).

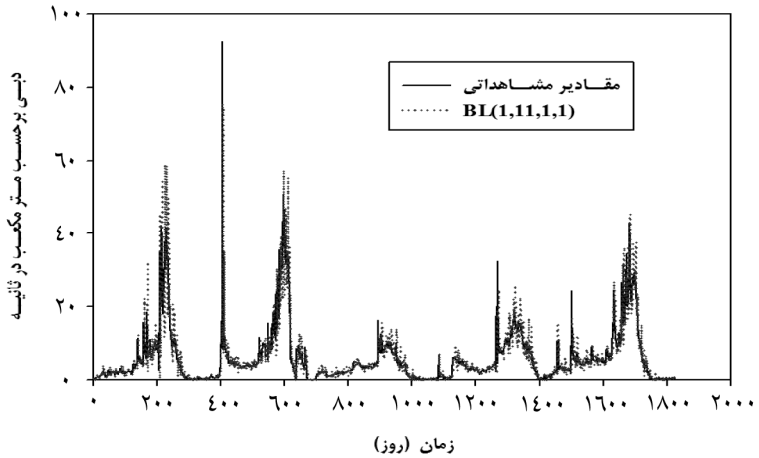
مدل برازشی	RMSE (مترمکعب بر ثانیه)	ضریب همبستگی
ARMA(1,11)	٤/۳۰۲	۰/۸٦۳
BL(1,11,1,1)	۳/۵۲۰	۰/۹۰۲

جدول ٤- رابطه خطی و غیرخطی سری روزانه جریان رودخانه باراندوزچای.

رابطه ریاضی مدل	مدل	مقیاس زمانی
$Z_t = ۰/۹۹۱۹Z_{t-۱} - ۰/۰٤۸۳۷\varepsilon_{t-۱} - ۰/۱۵۳٦\varepsilon_{t-۲} - ۰/۰۸۹۹\varepsilon_{t-۳}$ $- ۰/۰۷۷۲\varepsilon_{t-٤} - ۰/۰۲۰۵\varepsilon_{t-۵} - ۰/۰۵۲٤\varepsilon_{t-٦} - ۰/۰٤۸۷\varepsilon_{t-۷}$ $- ۰/۰۱۵۷\varepsilon_{t-۸} - ۰/۰۰٤۵\varepsilon_{t-۹} - ۰/۰۳٤۲\varepsilon_{t-۱۰} - ۰/۰۵۰۲\varepsilon_{t-۱۱} + \varepsilon_t$	ARMA(1,11)	روزانه
$Z_t = ۰/۹۹۱۹Z_{t-۱} - ۰/۰٤۸۳۷\varepsilon_{t-۱} - ۰/۱۵۳٦\varepsilon_{t-۲} - ۰/۰۸۹۹\varepsilon_{t-۳}$ $- ۰/۰۷۷۲\varepsilon_{t-٤} - ۰/۰۲۰۵\varepsilon_{t-۵} - ۰/۰۵۲٤\varepsilon_{t-٦} - ۰/۰٤۸۷\varepsilon_{t-۷}$ $- ۰/۰۱۵۷\varepsilon_{t-۸} - ۰/۰۰٤۵\varepsilon_{t-۹} - ۰/۰۳٤۲\varepsilon_{t-۱۰} - ۰/۰۵۰۲\varepsilon_{t-۱۱} + \varepsilon_t$ $+ (۰/۰۵ \times Z_{t-۱} \times \varepsilon_{t-۱})$	BL(1,11,1,1)	



شکل ۲- نمودار پراکندگی مقادیر مشاهداتی و تخمینی به‌دست آمده از مدل‌های خطی و غیرخطی برای سری روزانه رودخانه باراندوزچای در مرحله صحت‌سنجی (۸۸-۱۳۸٤).



شکل ۳- مقایسه آمار واقعی با مقادیر پیش‌بینی شده به‌دست آمده از مدل BL(1,11,1,1) برای سری روزانه در مرحله صحت‌سنجی (۸۸-۱۳۸٤).

نتایج مدل‌سازی جریان روزانه رودخانه باراندوزچای با استفاده از برنامه‌ریزی ژنتیک: در این مطالعه برای مدل‌سازی جریان روزانه رودخانه باراندوزچای با استفاده از برنامه‌ریزی ژنتیک، داده‌های ۳۳ سال (از ۱ مهر ۱۳۵۲ تا ۳۱ شهریور ۱۳۸۳) برای آموزش و ۵ سال (از ۱ مهر ۱۳۸٤ تا ۳۱ شهریور ۱۳۸۸) به‌عنوان داده‌های تست انتخاب شدند. به‌طورکلی ۸۵ درصد داده‌ها برای آموزش و ۱۵ درصد

نیز برای تست در نظر گرفته شد تا با مدل‌سازی سری زمانی جریان رودخانه مطابقت داشته باشد. همان‌گونه که پیش‌تر نیز ذکر شد، انتخاب جمعیت‌های اولیه تصادفی مختلف و تأثیرگذار در پدیده (که در برنامه‌ریزی ژنتیک به‌عنوان داده‌های آموزشی از آن‌ها یاد می‌شود) به‌منظور آموزش ماهیت سازوکار حاکم بر پدیده نه تنها سبب پیچیدگی الگو و افزایش حافظه درگیر خواهد شد، بلکه سبب کاهش دقت الگو نیز می‌شود. بنابراین در الگوسازی جریان رودخانه‌ها نیز باید سعی نمود مؤثرترین داده‌های مشاهداتی را به‌عنوان داده‌های آموزشی انتخاب کرد. الگوهای مورد استفاده در مقیاس زمانی روزانه در جدول ۵ آورده شده و جدول ٦ مشخصات برنامه‌ریزی ژنتیک به‌کار رفته را برای سری روزانه نشان می‌دهد. جدول ۷ مقایسه شاخص‌های آماری مربوط به نتایج به‌دست آمده از کاربرد روش برنامه‌ریزی ژنتیک در مدل‌سازی جریان روزانه رودخانه را برای الگوهای مختلف ورودی نشان می‌دهد. با توجه به این جدول مشاهده می‌شود که مدل با دبی سه روز قبل کم‌ترین خطا و بیش‌ترین دقت را دارا بوده و به‌عنوان بهترین الگوی ساخته شده برای برنامه‌ریزی ژنتیک انتخاب گردید. همچنین در مقایسه با مدل‌های خطی و غیرخطی سری زمانی روش برنامه‌ریزی ژنتیک نتیجه بهتری ارایه کرده است. در جدول ۸ رابطه ریاضی مستخرج از برنامه‌ریزی ژنتیک در الگوی سوم، برای سری روزانه جریان رودخانه باراندوزچای ارایه شده است. شکل‌های ٤ و۵ نمودارهای مقادیر مشاهداتی و مقادیر پیش‌بینی شده به‌دست آمده از بهترین مدل برنامه‌ریزی ژنتیک را برای سری روزانه جریان رودخانه باراندوزچای در مرحله صحت‌سنجی (دوره آماری ۸۸-۱۳۸٤) نشان می‌دهد. با توجه به این شکل‌ها می‌توان نتیجه گرفت که برنامه‌ریزی ژنتیک در بهترین حالت خود به خوبی توانسته دبی جریان روزانه رودخانه مورد مطالعه را پیش‌بینی کند.

مقدار این متغیرها همیشه بین صفر و ۱ قرار دارد که مقادیر کوچک‌تر آن‌ها بیانگر وجود اختلاف‌های کم‌تر (اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده در طول دوره آماری) در تولید داده توسط مدل است (جدول ۱). نتایج t-test مدل نیز در آزمون پیش‌بینی همه متغیرها در ۱۲ ماه سال بیانگر نبود اختلاف معنی‌دار مقادیر مشاهده شده با مقادیر پیش‌بینی شده در سطح احتمال ۵ درصد است (جدول ۲). با توجه به این‌که اختلاف معنی‌داری بین مقادیر شبیه‌سازی شده و مشاهده شده در طی بازه زمانی ۲۰۰۹-۱۹٦۱ مشاهده نمی‌شود و مقادیر بیش‌تر حول محور یک یه یک و در بازه ۲۰ درصد حدود اطمینان متمرکز هستند (شکل‌های ۱ و ۲)، از این‌رو، توانایی مدل در شبیه‌سازی و تولید داده‌های اقلیمی برای ۲۰ سال آینده مورد تأیید قرار گرفت.

جدول ٥- الگوهای ورودی مورد استفاده در مقیاس زمانی روزانه رودخانه باراندوزچای.

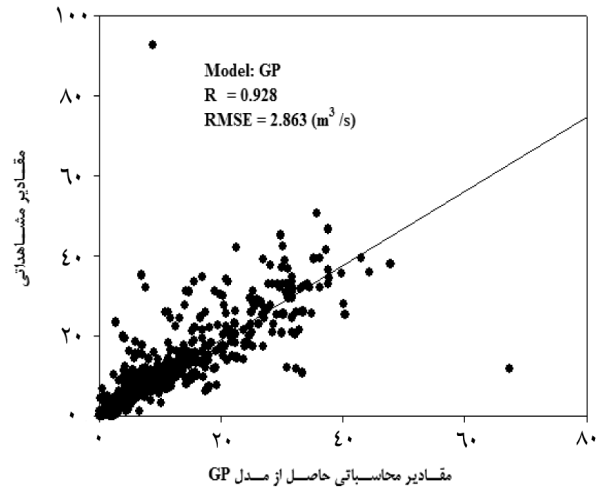
ردیف	الگوی ورودی روزانه
۱	$Q(t) = f\{Q(t-۱)\}$
۲	$Q(t) = f\{Q(t-۱), Q(t-۲)\}$
۳	$Q(t) = f\{Q(t-۱), Q(t-۲), Q(t-۳)\}$
٤	$Q(t) = f\{Q(t-۱), Q(t-۲), Q(t-۳), Q(t-٤)\}$

جدول ٦- مشخصات برنامه‌ریزی ژنتیک به‌کار رفته شده در این پژوهش.

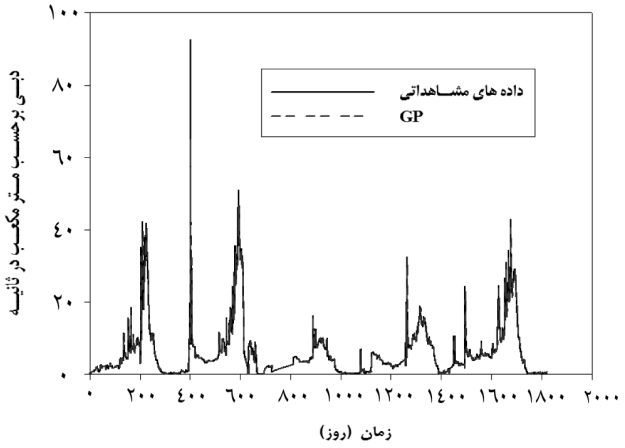
ردیف	معیار مورد نظر	مقدار عددی
۱	تولید بدون بهبود	۱٥۰
۲	تولید از زمان شروع	۲۰۰
۳	حداکثر مقدار اجرا	۱٥۰
٤	اندازه حداکثر برنامه	٥۱۲

جدول ۷- تحلیل‌های آماری نتایج برنامه‌ریزی ژنتیک برای الگوهای مختلف ورودی جریان روزانه رودخانه باراندوزچای.

الگوهای ورودی	مرحله صحت‌سنجی		
	ضریب همبستگی	RMSE (مترمکعب بر ثانیه)	
		داده‌های نرمال‌شده	داده‌های اصلی
۱	۰/۹۲٤	۰/۰۱۳۹	۲/۹٤۹
۲	۰/۹۲۳	۰/۰۱٤۱	۲/۹۸۷
۳	۰/۹۲۸	۰/۰۱۳٥	۲/۸٦۳
٤	۰/۹۱۷	۰/۰۱٤۷	۳/۰۹۹



شکل ٤- الف- نمودار پراکندگی مقادیر مشاهداتی و محاسباتی در بهترین حالت برنامه‌ریزی ژنتیک (۸۸-۱۳۸٤).



شکل ٥- نمودار مقادیر مشاهداتی و محاسباتی به‌دست آمده از مدل GP در مرحله صحت سنجی (۸۸-۱۳۸٤).

جدول ۸- رابطه ریاضی مستخرج از برنامه‌ریزی ژنتیک برای جریان روزانه رودخانه باراندوزچای در مرحله آموزش (۸۳-۱۳۵۲).

رابطه ریاضی دبی جریان ماهانه رودخانه باراندوزچای	مقیاس زمانی
$Q_{t-1}+\left[\left(\sqrt{Q_{t-1}}-Q_{t-1}\right)\times\left(Q_{t-2}\times Q_{t-1}\right)\right]-\left(Q_{t-1}\right)^{3}$	روزانه

نتیجه‌گیری کلی

تعیین جریان رودخانه‌ها یکی از مؤلفه‌های مهم و تأثیرگذار در مدیریت منابع آب حوضه آبریز می‌باشد. مطالعه این پدیده هیدرولوژیکی کاربردهای فراوانی از جمله در زمینه پیش‌بینی سیلاب را داراست. از طرفی به‌نظر می‌رسد فرآیند جریان رودخانه در مقطع زمانی کوتاه‌مدت به‌ویژه روزانه غیرخطی باشد. منبع اصلی غیرخطی بودن فرآیند جریان‌های کوتاه‌مدت روزانه از غیرخطی بودن فرآیندهای کوتاه‌مدت اقلیمی مانند بارش و درجه حرارت منشعب شده و غیرخطی بودن واکنش بارش رواناب نیز باعث پیچیده‌تر شدن فرآیند رودخانه می‌گردد (خلیلی و همکاران، ۲۰۱۱).

یکی از کاربردهای این پژوهش معرفی و استفاده از مدل غیرخطی دوخطی در منابع آب می‌باشد. این مدل با توجه به نتایج به‌دست آمده در این مطالعه با سری زمانی جریان رودخانه هم‌خوانی داشته و باعث بهبود نتایج نسبت به مدل‌های خطی می‌شود. استفاده از این مدل به‌دلیل شباهت زیاد به ساختار مدل‌های سری زمانی رایج ARMA و سادگی ساختار آن نسبت با سایر مدل‌های غیرخطی سری زمانی می‌تواند در هیدرولوژی و منابع آب به سهولت مورد استفاده قرار گیرد. نتایج این مطالعه نیز نشان داد که مدل دوخطی با دقت بالا و خطای کم‌تری نسبت به مدل خطی توانسته جریان روزانه رودخانه باراندوزچای ارومیه را مدل‌سازی نماید. نتایج حاصله با پژوهش ناوه و همکاران (۲۰۱۲) همخوانی دارد.

علاوه‌بر مدل دوخطی از برنامه‌ریزی ژنتیک نیز برای مدل‌سازی جریان روزانه رودخانه باراندوزچای استفاده گردید. نتایج این مطالعه در مجموع روش برنامه‌ریزی ژنتیک را به‌عنوان یک روش صریح و دقیق برای پیش‌بینی جریان رودخانه‌ها پیشنهاد می‌کند. این پیشنهاد با نتایج به‌دست آمده از مطالعات گون (۲۰۰۹) که به پیش‌بینی جریان روزانه رودخانه به کمک برنامه‌ریزی ژنتیک پرداخته است مطابقت کامل دارد. همچنین فربودنام و همکاران (۲۰۰۹) و داننده مهر و طباطبایی (۲۰۱۰) نیز نشان دادند که برنامه‌ریزی ژنتیک از دقت بیش‌تری نسبت به شبکه عصبی مصنوعی برخوردار است و نتیجه این پژوهش نیز برتری روش GP را نسبت به مدل غیرخطی دوخطی نشان می‌دهد.

منابع

1.Ainkaran, P. 2004. Analysis of some linear and nonlinear time series models. A thesis submitted in fulfillment of the requirements for the degree of Master of Science, School of Mathematics and Statistics, University of Sydney.

2.Akaike, H. 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control. 19: 716-723.

3.Alvisi, S., Mascellani, G., Franchini, M., and Bardossy, A. 2005. Water level forecasting through fuzzy logic and artificial neural network approaches. J. Hydrol. Earth. Sys. Sci. 2: 1107-1145.

4.Aytek, A., and Asce, M. 2008. An application of artificial intelligence for rainfall runoff modeling. J. Hydrol. Earth. Syst. Sci. 117: 2. 145-155.

5.Caiado, J. 2007. Forecasting water consumption in Spain using univariate time series models. Munich Personal RePEc Archive, no: 6610.

6.Dai, Y., and Billard, L. 2003. Maximum likelihood estimation in space time bilinear models. J. Time Series Anal. 24: 1. 25-44.

7.Dananademehr, A., and Tabatabai, M.R.M. 2010. Prediction of Daily discharge trend of river flow based on genetic programming. J. Water Soil. 24: 2. 325-333. (In Persian)

8.Fan, J., and Yao, Q. 2003. Nonlinear time series, nonparametric and parametric methods. Springer-Verlag, NewYork, Inc.

9.Farbodfam, N., Ghorbani, M.A., and Aalami, M.T. 2009. Forecasting river flow using genetic programming (Case study: Lighwan watershed). J. Agric. Sci. 19: 1. 107-123. (In Persian)

10.Ferreira, C. 2001. Gene expression programming: A new adaptive algorithm for solving problems. Complex Systems. 13: 2. 87-129.

11.Ghorbani, M.A., Shiri, J., and Kazemi, H. 2010. Estimation of Maximum, Mean and Minimum Air Temperature in Tabriz City Using Artificial Intelligent Methods. J. Agric. Sci. 20: 3. 87-104. (In Persian)

12.Ghorbani, M.A., Kisi, O., and 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: 4050-4057.

13.Govindaraju, R.S. 2000. Artificial neural network in hydrology. J. Hydrol. Engin. 5: 2. 115-123.

14.Granger, C.W.J., and Andersen, A.P. 1978. An Introduction to Bilinear Time Series Models. Vandenhoek and Ruprecht: Gottingen.

15.Guven, A. 2009. Linear genetic programming for time-series modeling of daily flow rate. J. Earth. Syst. Sci. 118: 2. 157-173.

16.Khalili, K., Fakheri Fard, A., Dinpashoh, Y., and Ghorbani, M.A. 2011. Nonlinearity and testing for stream flow processes by BDS test (Case study: Shaharchai river). J. Agric. Sci. 22: 2. 25-37. (In Persian)

17. Khalili, K., Ahmadi, F., Abghari, H., and Besharat, S. 2013. Stationary and nonlinearity analysis of river flow at different time scales (Case study: Barandouz-Chai River). 9th International River Engineering Conference. Ahwaz. Iran.
18. Khu, S.T., Liong, S.Y., Babovic, V., Madsen, H., and Muttil, N. 2001. Genetic programming and its application in real-time runoff forming. J. Amer. Water Resour. Assoc. 37: 2. 439-451.
19. Lifshits, M.A. 2006. Invariance principle in a bilinear model with weak nonlinearity. J. Math. Sci. 137: 1. 4541-4545.
20. Naveh, H., Khalili, K., Alami, M.T., and Behmanesh, J. 2012. Forecasting River flow By Bilinear Nonlinear Time Series Model (Case Study: Barandoz-Chay & Shahar-Chai Rivers). J. Water Soil. 26: 5. 1299-1307. (In Persian)
21. Salas, J.D. 1993. Analysis and modeling of hydrologic time series. Handbook of hydrology. 19: 1-72.
22. Subba Rao T., and Gabr, M. 1984. An Introduction to Bispectral Analysis and Bilinear Time Series Models. Lecture Notes in Statistics, 24, Springer-Verlag: New York.
23. Tsay, R.S. 2002. Analysis of financial time series. University of Chicago, A Wiley-Interscience Publication, John Wiley f Sons; Inc.
24. Tsonis, A.A. 2001. Probing the linearity and nonlinearity in the transitions of the atmospheric circulation. Nonlinear Processes Geophysics. 8: 341-345.
25. Wang, W., Van, G., Pieter, H.A.J.M., Vrijling, J.K., and Ma, J. 2005. Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes. Nonlin. Processes Geophys. 12: 55-66.
26. Wang, W., Vrijling, J.K., Pieter, H.A.J.M., Van, G., and Ma, J. 2006. Testing for nonlinearity streamflow processes at different timescales. J. Hydrol. 322: 1. 247-268.

Gorgan University of Agricultural Sciences and Natural Resources

J. of Water and Soil Conservation, Vol. 22(1), 2015
http://jwsc.gau.ac.ir

Comparing nonlinear time series models and genetic programming for daily river flow forecasting (Case study: Barandouz-Chai River)

***F. Ahmadi[1], Y. Dinpashoh[2], A. Fakheri Fard[3], K. Khalili[4] and S. Darbandi[5]**

[1]M.Sc. Student, Dept. of Water Engineering, Tabriz University, [2]Associate Prof., Dept. of Water Engineering, Tabriz University, [3]Professor, Dept. of Water Engineering, Tabriz University, [4]Assistant Prof., Dept. of Water Engineering, Urmia University, [5]Assistant Prof., Dept. of Water Engineering, Tabriz University

Received: 07/21/2013; Accepted: 02/16/2014

Abstract

In this study daily river flows of Barandouz-Chai in the period of 1973 to 2009 have been forecasted using nonlinear time series and Genetic programming models and results have been analyzed with root mean square error and regression coefficient. After evaluating models by goodness of fit, nonlinear BL(1,11,1,1) with the minimum AICC was selected for modeling daily river flows forecasting. For BL(1,11,1,1) model, regression coefficient and RMSE have been obtained as 0.902 and 3.52 (m^3/s), respectively. Genetic programming has been used for modeling river flows with consideration to memory of one, two, three and four days before. Results showed until three days memory, accuracy of model is acceptable but after that accuracy of model decreased. Regression coefficient and RMSE of Genetic programming in the testing phase calculated equal to 0.928 and 2.863 (m^3/s) respectively. Results showed, Genetic programming with 22.9 percent less error was better than bilinear time series model in Barandouz-chai river flow forecasting.

Keywords: Discharge, Bilinear model, Forecasting, Genetic programming, Testing

* Corresponding Author; Email: farshad.paper@yahoo.com