Archive of SID

ارزیابی مدل‌های ژنتیکی جهت مدل‌سازی جریان رودخانه

صابر علیدادی ده کهنه [۱]، اباذر سُلگی [۲]، مهرنوش شهنی دارابی [۳]، حیدر زارعی [۴،۱].

تاریخ ارسال:۱۳۹۶/۰۹/۲۳
تاریخ پذیرش:۱۳۹۷/۰۳/۰۴

چکیده

پیش‌بینی جریان رودخانه‌ها به عنوان یکی از منابع اصلی تأمین کننده آب بشر، همواره یکی از موضوعات مهم مورد بحث در هیدرولوژی و منابع آب بوده است. بدین جهت، مدل‌های مختلفی برای مدل‌سازی و پیش‌بینی جریان رودخانه‌ها مورد استفاده قرار گرفته است. در این مطالعه، به ارزیابی دو مدل ژنتیکی به نام‌های برنامه‌ریزی ژنتیک و برنامه‌ریزی بیان ژن پرداخته شده است. برای این منظور، با استفاده از داده‌های روزانه جریان، دما، بارش و تبخیر در ایستگاه تله‌زنگ اقدام به مدل‌سازی جریان رودخانه دز شده است. نتایج نشان داد که مدل برنامه‌ریزی بیان ژن با ضریب تبیین ۰/۸۶ و جذر میانگین مربعات خطای ۰/۰۰۳۰(مترمکعب در ثانیه) نسبت به مدل برنامه‌ریزی ژنتیک با ضریب تبیین ۰/۸۵ و جذر میانگین مربعات خطای ۰/۰۰۳۷(مترمکعب در ثانیه) دارای عملکرد بهتری می‌باشد. علاوه بر این، سرعت اجرای مدل برنامه‌ریزی بیان ژن نسبت به مدل برنامه‌ریزی ژنتیک بیشتر بوده و در زمان کوتاهی قادر به ارائه نتایج می‌باشد. با افزایش تعداد داده‌های ورودی مدل برنامه‌ریزی ژنتیک کند شده و گاهی قادر به ارائه نتایج نمی‌باشد درحالی‌که مدل برنامه‌ریزی بیان ژن این قابلیت را دارد که با تعداد ورودی‌ها و داده‌های بیشتر، نیز عمل مدل‌سازی را انجام دهد. به‌طورکلی نتایج نشان داد که مدل برنامه‌ریزی بیان ژن برای مدل‌سازی و پیش‌بینی جریان رودخانه قابلیت خوبی دارد.

کلیدواژه: مدل GEP، مدل GP، مدل‌سازی جریان رودخانه، رودخانه دز.

[۱] تلفن تماس و ایمیل نویسنده مسئول: ۰۹۱۶۳۰۳۳۰۵۵، zareih@scu.ac.ir

۱- صابر علیدادی ده کهنه-آدرس : اهواز سازمان آب و برق خوزستان، معاونت آبرسانی Alidadis@yahoo.com

[۲]- اباذر سُلگی- آدرس: نهاوند- شهر فیروزان، دانشجوی دکتری مهندسی منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهید چمران اهواز. اهواز. تلفن تماس و ایمیل: ۰۹۱۸۹۴۹۹۵۷۹، A-solgi@phdstu.scu.ac.ir

۳- مهرنوش شهنی دارابی- آدرس: اهواز-گلستان -خیابان اصفهان بین آذر و دی- ساختمان نگین، کارشناس کنترل کیفیت شرکت آب غدیر خوزستان و دانشجوی دکتری مهندسی منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهید چمران اهواز. shehnidarabi.mehrnoush@gmail.com

٤- حیدر زارعی- استادیار گروه هیدرولوژی و منابع آب دانشکده مهندسی علوم آب، دانشگاه شهید چمران اهواز. اهواز. تلفن تماس و ایمیل: ۰۹۱۶۳۰۳۳۰۵۵، zareih@scu.ac.ir

www.SID.ir

مقدمه

یکی از مهم‌ترین مسائل در مهندسی منابع آب، پیش‌بینی میزان دبی رودخانه می‌باشد که از نظر برنامه-ریزی منابع آب دارای اهمیت است. بهره‌برداری بهینه و مناسب از منابع آب یکی از وظایف مهم مدیران منابع آب می‌باشد. از طرفی دیگر عدم تخمین مناسب رواناب حاصل از بارش در حوضه‌های آبریز، مدیریت بهینه منابع آبی و سیلاب‌ها به ویژه مدیریت بهره‌برداری مخازن سدها را دچار مشکل می‌نماید. در این میان مقوله شبیه‌سازی به عنوان راهکاری مناسب جهت تخمین رواناب، نمود پیدا می‌کند. تاکنون مدل‌ها و روش‌های مختلفی برای تخمین و شبیه‌سازی فرآیند بارش رواناب ارائه شده است که با استفاده از آن‌ها شدت جریان رودخانه‌ها پیش‌بینی می‌شود. برای پیش-بینی جریان رودخانه، استفاده از مدل‌های آماری، هیدرولیکی و هیدرولوژیکی دارای سابقه‌ای طولانی هستند. تجربه نشان داده که این مدل‌ها، متنوع بوده و اعمال کلیه آن‌ها در مدل‌های طراحی شده بسیار مشکل می‌باشد. علاوه بر این، نبود قطعیت و غیرخطی بودن قوی روابط، بین متغیرها، مسأله را پیچیده می-کند. همچنین مدل‌های فیزیکی و مفهومی هیدرولوژیکی به دلیل نیاز به اطلاعات فراوان، پارامترهای گوناگون و واسنجی وقت‌گیر، کمتر مورد توجه قرار گرفته‌اند. برنامه‌ریزی ژنتیک(GP) و برنامه-ریزی بیان ژن[۲](GEP) از جدیدترین الگوریتم‌های فراکاوشی هستند که به دلیل دارا بودن دقت کافی، مورد توجه پژوهشگران قرار گرفته‌اند. از برتری عمده این مدل‌ها این است که می‌توانند برای شرایطی که، الف) ارتباط موجود بین متغیرهای مسأله به خوبی شناخته شده نباشد، یا صحت و سقم شناخت فعلی از رابطه نام برده با تردید همراه باشد، ب) پیدا کردن حل نهایی مسأله مورد بررسی مشکل باشد، ج) حل ریاضی مرسوم وجود نداشته باشد (یا حل تحلیلی را ایجاب می‌کند)، د) راه‌حل تقریبی قابل قبول باشد، ه) تعداد داده‌هایی که باید به وسیله کامپیوتر مورد آزمون، دسته-بندی و جمع‌بندی قرار گیرند زیاد باشد (مانند داده‌های ماهواره‌ای)، به کار برده شوند (بنژاف و همکاران، ۱۹۹۸). بنابراین، در سال‌های اخیر،کاربرد مدل‌های هوشمند از جمله برنامه‌ریزی ژنتیک و برنامه‌ریزی بیان-ژن به عنوان مدل‌های جدید و ابزاری توانمند در فرآیندهای هیدرولوژیکی نظیر پیش‌بینی، افزایش یافته است، زیرا با استفاده از آن‌ها، عدم قطعیت‌های موجود را به همراه درک روابط ذاتی میان داده‌ها که ریاضیات کلاسیک قادر به حل آن‌ها نمی‌باشد، می‌توان مدل-سازی نموددر زمینه‌ی استفاده از مدل برنامه‌ریزی ژنتیک برای پیش‌بینی جریان رودخانه می‌توان به مطالعات زیر اشاره نمود:

قربانی و همکاران (۲۰۱۰) عملکرد سه روش برنامه ریزی ژنتیک، شبکه عصبی مصنوعی و نروفازی را در روندیابی سیلاب رودخانه قزل‌ایرماق ترکیه مورد ارزیابی قرار دادند. نتایج این مطالعه نشان داد که از بین سه روش مذکور مدل برنامه‌ریزی ژنتیک با دقت بیشتری هیدروگراف خروجی را شبیه‌سازی کرده است. ظهیری و عظمت اله (۲۰۱۴) برای پیش‌بینی دبی جریان در مقاطع مرکب از دو روش برنامه‌ریزی ژنتیک خطی و مدل درختی M5 استفاده کردند. نتایج نشان داد که هر چند هر دو مدل از دقت بالایی برای پیش بینی جریان برخوردار بودند، اما دقت روش برنامه‌ریزی ژنتیک خطی بالاتر از مدل درختی M5 بود.

فربود فام و همکاران (۲۰۰۹) با استفاده از برنامه-ریزی ژنتیک، پیش‌بینی جریان روزانه رودخانه لیقوان را مورد بررسی قرار دادند. نتایج این تحقیق نشان داد که برنامه‌ریزی ژنتیک دارای دقت بسیار بالایی نسبت به روش شبکه‌های عصبی مصنوعی و مدل‌های سری زمانی می‌باشد.

برنامه‌ریزی بیان ژن شاخه‌ای از الگوریتم‌های تکاملی است که توانایی مدل‌سازی فرآیندهای غیرخطی و پویا را دارد. در زمینه استفاده از برنامه‌ریزی بیان ژن برای پیش‌بینی پارامترهای مختلف هیدرولوژیکی میتوان به مطالعات زیر اشاره کرد:

[1]- Genetic Programming

[2]-Gene Expression Programming

www.SID.ir

Archive of SID

-مقایسه برنامه‌ریزی بیان ژن با سیستم استنتاج فازی عصبی تطبیقی(ANFIS) جهت پیش‌بینی کوتاه مدت نوسانات عمق آب توسط شیری وکایزی (۲۰۱۱) صورت پذیرفته است. نتایج نشان داد که هر دو مدل می‌توانند به عنوان یک ابزار دقیق برای پیش‌بینی نوسانات عمق آب استفاده شوند. اما مدل GEP دارای فرمول ساختاری ساده و راحت‌تری نسبت به مدل ANFIS می‌باشد.

مدل‌سازی تعرق مرجع روزانه با استفاده از GEP در شمال اسپانیا برای یک دوره ۵ ساله (۱۹۹۹-۲۰۰۳) توسط شیری و همکاران (۲۰۱۲) صورت پذیرفته است. نتایج روشGEP با روش‌های ANFIS و پریستلی-تیلور[2]و مدل هارگریوز سامانی[3]مقایسه شد. نتایج نشان داد که مدل GEP بهتر از سایر مدل‌ها بوده و بعد از آن مدل ANFIS عملکرد بهتری داشته است.

کاوه‌کار و همکاران (۲۰۱۳) به بررسی شبیه‌سازی نوسانات تراز آب دریاچه ارومیه با استفاده از مدل‌های برنامه‌ریزی بیان ژن و شبکه عصبی مصنوعی پرداختند. نتایج حاکی از دقت مطلوب مدل برنامه‌ریزی بیان‌ژن در شبیه‌سازی نوسانات سطح آب بوده است.

احمدی و همکاران (۲۰۱۵) به مقایسه روش‌های برنامه‌ریزی ژنتیک و ماشین بردار پشتیبان برای پیش-بینی جریان روزانه رودخانه باراندوزچای پرداختند. مقایسه نتایج دو مدل نشان داد که گرچه دقت روش برنامه‌ریزی ژنتیک نسبت به روش ماشین بردار پشتیبان اندکی بیشتر بود، اما روش ماشین بردار پشتیبان به مراتب ساده‌تر از روش برنامه‌ریزی ژنتیک می‌باشد و این روش می‌تواند به عنوان یک روش کاربردی برای پیش‌بینی جریان روزانه به کار رود.

با توجه به نام مشابه مدل برنامه‌ریزی بیان ژن با برنامه‌ریزی ژنتیک، ممکن است برخی‌ها تصور کنند که این دو مدل تفاوتی ندارد، لذا در این مطالعه سعی شده است با بیان نمودن تفاوت‌ها و شناسایی این دو مدل، از آن‌ها برای پیش‌بینی جریان رودخانه دز استفاده شود.

در بخش دوم به بررسی مواد و روش‌های استفاده شده در این مطالعه پرداخته خواهد شد. نتایج و بحث در بخش سوم ارائه شده است. نتیجه گیری و تقدیر و تشکر نیز به ترتیب در بخش‌های سوم و چهارم این مطالعه گنجانده شده است.

مواد و روش‌ها

منطقه مورد مطالعه

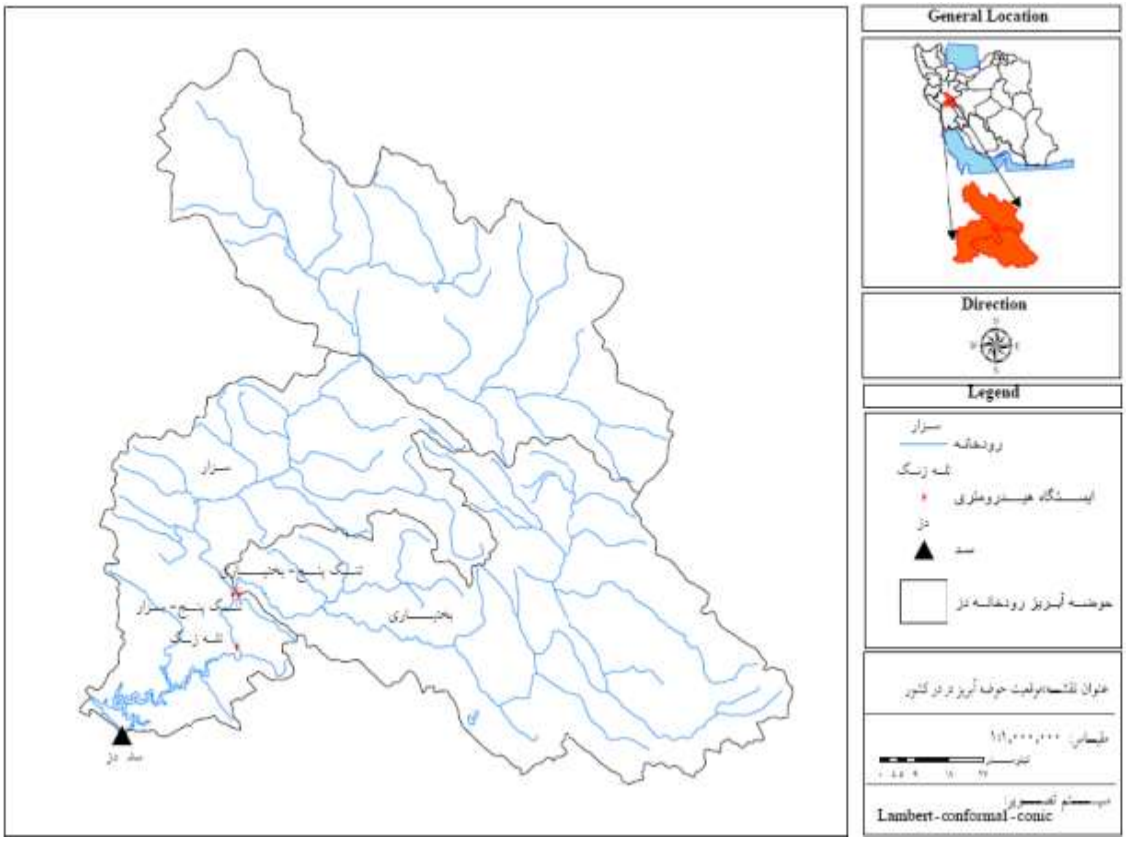
حوضه آبریز دز بخشی از ارتفاعات زاگرس میانی را در بر می‌گیرد. حوضه دز از غرب به حوضه کرخه، از شمال به حوضه قره‌چای و زاینده‌رود و از شرق و جنوب به حوضه کارون محدود است. این حوضه از لحاظ تقسیم‌بندی کلی حوضه‌های ایران، جزئی از حوضه آبریز خلیج‌فارس می‌باشد. این حوضه در محدوده بین ′۳۵-°۳۲ تا ′۰۷ °۳۴ عرض جغرافیایی شمالی و ′۲۰ -°۴۸ تا ′۲۰-°۵۰ طول جغرافیایی شرقی، در جنوب غربی ایران واقع شده است. شکل (۱) موقعیت عمومی حوضه آبریز رودخانه دز، شبکه زهکشی و ایستگاه هیدرومتری تله زنگ را نشان می‌دهد.

در این تحقیق به منظور مدل‌سازی آبدهی ایستگاه ورودی به مخزن سد (تله زنگ) از داده‌های خود ایستگاه استفاده شده است و به همین دلیل مدل‌های استفاده شده در این تحقیق خودهمبسته نامیده می‌شوند. به منظور مدل‌سازی آمار آبدهی ایستگاه هدف (ایستگاه تله زنگ در محل ورودی به مخزن سد دز) در مقیاس روزانه، طول دوره آماری آبدهی این ایستگاه از سال آبی ۱۳۷۴-۱۳۷۳ تا سال آبی ۱۳۹۴-۱۳۹۳ انتخاب شده است. در واقع آمار مورد استفاده شامل ۷۲۰۸ داده می-باشد که از مهرماه سال ۱۳۹۳ شروع شده و به اسفند ۱۳۹۳ ختم می‌شود. در جدول (۱) پارامترهای مورد استفاده برای مدل‌سازی ارائه شده است.

[1]- Adaptive Neuro Fuzzy Inference System
[2]- Priestley–Taylor
[3]- Hargreaves–Samani

www.SID.ir

Archive of SID



شکل۱- موقعیت ایستگاه و منطقه مورد مطالعه.

جدول ۱- پارامترهای روزانه مورد استفاده در این تحقیق.

نام ایستگاه	پارامترهای استفاده شده	میانگین	حداقل	حداکثر
تله‌زنگ	دما(°C)	۲۴/۵۳	۰/۰۰	۴۳/۰۰
	بارش(mm)	۲/۱۲	۰/۰۰	۱۵۷/۰۰
	جریان(m^3/s)	۲۰۰/۵۷	۳۴/۰۰	۷۲۶۸/۰۰
	تبخیر (mm)	۷/۷۵	۰/۳	۸۳/۱۰

که در این روابط X داده مورد نظر، $\bar{X}$ میانگین داده‌ها، Xmax حداکثر داده‌ها، Xmin حداقل داده‌ها و y داده استاندارد شده می‌باشد. در این تحقیق از ۷۵ درصد داده‌ها برای مرحله آموزش و ۲۵ درصد برای مرحله تست استفاده شده است.

برنامه‌ریزی ژنتیک(GP)

برنامه‌ریزی ژنتیک به عنوان پیشرفتی از الگوریتم ژنتیک و به عنوان توسعه‌ای جدید، در الگوریتم‌های

به منظور هماهنگ‌تر شدن بهتر داده‌ها و افزایش سرعت اجرای مدل‌ها، داده‌ها با استفاده از روش استانداردسازی به اعدادی بین صفر تا یک تبدیل شدند. با توجه به پیشنهاد سلگی (۱۳۹۳) از رابطه زیر برای نرمال سازی (استانداردسازی) استفاده شده است.

$$y = 0.5 + (0.5 \times (\frac{x-\bar{x}}{x_{max}-x_{min}})) \qquad (۱)$$

[1]-Genetic Programming

www.SID.ir

Archive of SID

استفاده می‌شود، از اهمیت مضاعفی برخوردار خواهد بود، چون ارائه داده‌های ورودی غیر مرتبط، سبب کاهش دقت الگو و ایجاد الگوهای پیچیده‌تری می‌شود که تفسیر آن‌ها با دشواری های بیشتری مواجه است. برنامه‌ریزی ژنتیک تناوبی از راه‌حل‌های با طول ثابت را از طریق ابداع ساختارهای غیرخطی (نمودار درخت) با اندازه‌ها و اشکال متفاوت ارایه می‌کند. الفبای مورد استفاده در ایجاد چنین ساختاری متنوع‌تر از صفر و یک، افراد در الگوریتم ژنتیک است و سیستم نمایش تطبیق‌پذیر و غنی‌تری را ایجاد می‌کند. به رغم فقدان کروموزوم (ژنوم) ساده و مستقل، ساختارهای غیرخطی از افراد GP نیز مانند کروموزم‌های خطی الگوریتم ژنتیک[3](GA)، تکثیرکننده‌های عادی با نقش دوگانه‌ی ژنوتیپ-فنوتیپ هستند. قابل توجه است که نمودار درختی GP، به مولکول‌های پروتئین در استفاده از الفبای غنی و نمایش مرتبه‌ای بی‌نظیر و پیچیده‌شان شباهت دارند. در حقیقت نمودار درختی قادر به نمایش عوامل متنوع زیادی هستند.

مشکل چنین تکثیرکننده‌های پیچیده این است که در آن‌ها تولیدمثل با اصلاحات، به شدت در ضوابط تکاملی محبوس شده است. چون به‌سازی‌ها واقعاً باید روی خود نمودار درختی رخ دهد و در نتیجه تنها دامنه-ی محدودی از بهسازی‌ها امکان‌پذیر است، این عملگر-های ژنتیکی GP در سطح درختی هستند که به‌سازی-ها یا تبادلات شاخه‌های خاص بین درختان را اداره می-کنند (فریرا، ۲۰۰۶). در GP عملگرها مستقیماً روی نمودار درختی عمل می‌کنند و اگر چه این موضوع به ظاهر دارای مزایایی است، اما تکنیک GP را به مقدار زیادی محدود می‌کند (همان‌طوری‌که در طبیعت نیز محدودیت‌هایی برای پیوند و هرس وجود دارد).

اجزای تشکیل دهنده برنامه‌ریزی ژنتیک

به‌طورکلی، الگوریتم‌های ژنتیکی و به عنوان تصمیمی از آن برنامه‌ریزی ژنتیک از اجزاء زیر تشکیل تکاملی ایجاد شده است. برنامه‌ریزی ژنتیک که به طور وسیعی در کاربرد هوش مـصنوعی در حـل مسائل مهندسی استفاده می‌شود در سال ۱۹۸۵ توسط کرامر[1] ابداع و سپس توسط کوزا[2] گسترش بیشتری یافت. برنامه‌ریزی ژنتیک از روش‌های جدید بین روش‌های الگوریتم گردشی می‌باشد که بـه دلیـل دارا بودن دقت کافی، به عنوان یک روش کاربردی مطرح می‌شـود. الگوریتم‌های گردشی اقدام به تعریف یک تابع هدف در قالب معیارهای کیفی نمـوده و سپس تابع هدف را برای اندازه‌گیری و مقایسه روش‌های مختلف حل کرده و در یک فرایند گام به گام تصحیح ساختار داده‌ها به کار گرفتـه و در نهایت، روش حل مناسب را ارائه می‌نمایند. مبنای تمامی ایـن روش‌ها بر اساس نظریه تکاملی داروین استوار است.

با توجه به پیشرفت‌های اخیر در فناوری رایانه و الگوریتم‌های جدید ریاضی، تکنیک‌های مدل‌سازی داده محور به عنوان روش‌های جدید برای شبیه‌سازی و پیش بینی پدیده‌هایمختلف طبیعی و مصنوعی توسعه یافته-اند. این تکنیک‌ها در مدل‌سازی مورد توجه خاص محققین می‌باشند که به طور گسترده برای مدل‌سازی-های پیچیده از آن‌ها استفاده شده است. مدل‌های داده محور مانند برنامه‌ریزی ژنتیک نیاز به شناخت جامع فیزیک فرآیند مورد مطالعه ندارند. توانایی این تکنیک ها برای مدل‌سازی و پیش‌بینی فرآیندهای پیچیده مختلف در منابع گوناگون مورد بررسی قرار گرفته است. در برنامه‌ریزی ژنتیک، ابتدا بلوک‌های موجود که شامل متغیرهای ورودی و هدف و نیز تابع ارتباط دهنده آن‌ها می‌باشند، تعریف گردیده و سپس ساختار مناسب الگو و ضرایب آن تعیین می‌شود. این روش شامل یک معادله ارتباط دهنده بین متغیرهای ورودی و خروجی بوده، لذا قادر به انتخاب خودکار متغیرهای مناسب الگو و حذف متغیرهای غیر مرتبط است که این امر سبب کاهش ابعاد متغیرهای ورودی خواهد شد (گلابی و همکاران، ۱۳۹۲). انتخاب ورودی‌های مناسب، یکی از مهم‌ترین مواردی است که بایستی در این روش مورد توجه قرار گیرد. این امر در شرایطی که از داده‌های ورودی مختلف

[1]-Cramer
[2]-Koza
[3]-Genetic Algorithm

www.SID.ir

Archive of SID

استفاده شود، به‌طوری‌که راه‌حل‌های مناسب تنها با حرکت این بلوک‌های ساختمانی در اطراف، می‌تواند کشف شود. در آخر می‌توان گفت به سبب نقش دوگانه نمودار درختی (ژنوتیپ و فنوتیپ)، برنامه‌ریزی ژنتیک نیز همانند الگوریتم ژنتیک برای بیان ساده و اولیه ناتوان است. در همه‌ی موارد تمام نمودار درختی به عنوان راه حل است نه بیشتر و نه کمتر (فریرا، ۲۰۰۶).

برنامه‌ریزی بیان ژن[10]

برنامه‌ریزی بیان ژن که در ادامه سیر تکاملی مدل های هوشمند به وجود آمده است جزء روش‌های الگوریتم گردشی محسوب می‌شود که مبنای تمامی آنها بر اساس نظریه تکامل داروین استوار است (داننده مهر و همکاران، ۱۳۸۹). مزیت برنامه‌ریزی بیان ژن نسبت به مدل‌های دیگر از جمله شبکه عصبی مصنوعی این است که در برنامه‌ریزی بیان ژن، ابتدا ساختار (متغیرهای ورودی، هدف و مجموع توابع) تعریف شده و سپس ساختار بهینه مدل و ضرایب طی فرایند آموزش تعیین می‌شوند، در حالی که در شبکه‌های عصبی مصنوعی، ابتدا باید ساختار تعیین شده، فقط ضرایب مدل طی فرآیند آموزش حاصل می‌شوند. همچنین این الگوریتم به طور خودکار می‌تواند متغیرهای ورودی که در مدل بیشترین تأثیر را دارند، انتخاب کند. در این روش 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 برنامه ریزی ژنتیک ترکیب می‌شوند. اولین مرحله در الگوریتم مدل، تولید جمعیت اولیه از راه حل‌هاست. این امر می تواند به وسیله فرآیند تصادفی و یا در نظر گرفتن اطلاعات ورودی درباره مسأله انجام شود. سپس کروموزوم‌ها به صورت بیان درختی اظهار شده و توسط تابع برازش ارزیابی می‌گردند. در صورت دستیابی به راه‌حل مطلوب و یا رسیدن نسل‌ها به تعداد معین، تکامل متوقف شده و

[6]- Mutation Operator
[7]-Combination Operator
[8]-Permutation Operator
[9]-Transposition or Inversion Operator
[10] -Wavelet Transform

می‌شوند. درخت تجزیه[1] جمعیت[2] تابع برازندگی[3] و عملگرهای ژنتیکی. که عملگرهای ژنتیکی خود به عملگر تولید مثل، عملگر انتخاب[4] عملگر تلاقی[5] عملگر جهش[6] عملگر ترکیب[7] و عملگر جایگشت[8] تقسیم می‌شود. جهت آشنایی با این مباحث به منابع (فریرا، ۲۰۰۱؛ فریرا، ۲۰۰۶) مراجعه شود.

ساختار کلی برنامه‌ریزی ژنتیک

ساختار کلی برنامه‌ریزی ژنتیک که در آن فرآیند گام به گام به صورت فلوچارت مشخص شده، در شکل (۲) نشان داده شده است. شرط توقف در این مدل به دو صورت با تعداد تکرار ثابت و یا با دقت مشخص قابل تعریف می‌باشد.

به طور خلاصه، عملگرها در برنامه‌ریزی ژنتیک بیشتر به یک ریاضیدان هوشیار شباهت دارند تا به یک روش کور طبیعی. اما در سیستم‌های انطباقی، روش کور طبیعی مؤثرتر است، چون سیستم‌هایی مانند GP در ضوابط تکاملی به شدت محدود می‌شوند. برای مثال اجرای سایر عملگرها در GP، از قبیل جهش نقطه‌ای ساده در عین حال با تکرار زیاد بی حاصل است. به دلیل اینکه جهش‌های بیشتر، ساختار های نامعتبر و نادرستی را با توجه به قواعد نتیجه می‌دهد. بدیهی است که اجرای سایر عملگرها از قبیل ترانهش یا وارون‌سازی[9] نیز مشکلات مشابهی را تولید می‌کند و فضای تحقیقی در GP به طور وسیعی کشف نشده باقی می‌ماند (فریرا، ۲۰۰۶).

اگرچه کوزا عملگرهای انتخاب، تلاقی و جهش را به عنوان عملگرهای اصلی GP شرح داد، اما عملگر ترکیب، تنها عملگر مورد استفاده در بیشتر کاربردهای GP است و در نتیجه مواد ژنتیکی جدیدی در منبع ژنتیکی افراد GP وارد نمی‌شود. در حقیقت جمعیت بزرگی از نمودار درختی باید به منظور آماده‌سازی جمعیت اولیه (با تمام بلوک‌های ساختمانی لازم)

[1]-Parse Tree
[2]-Population
[3]-Fitness Function
[4]-Selection Operator
[5]-Crossover Operator

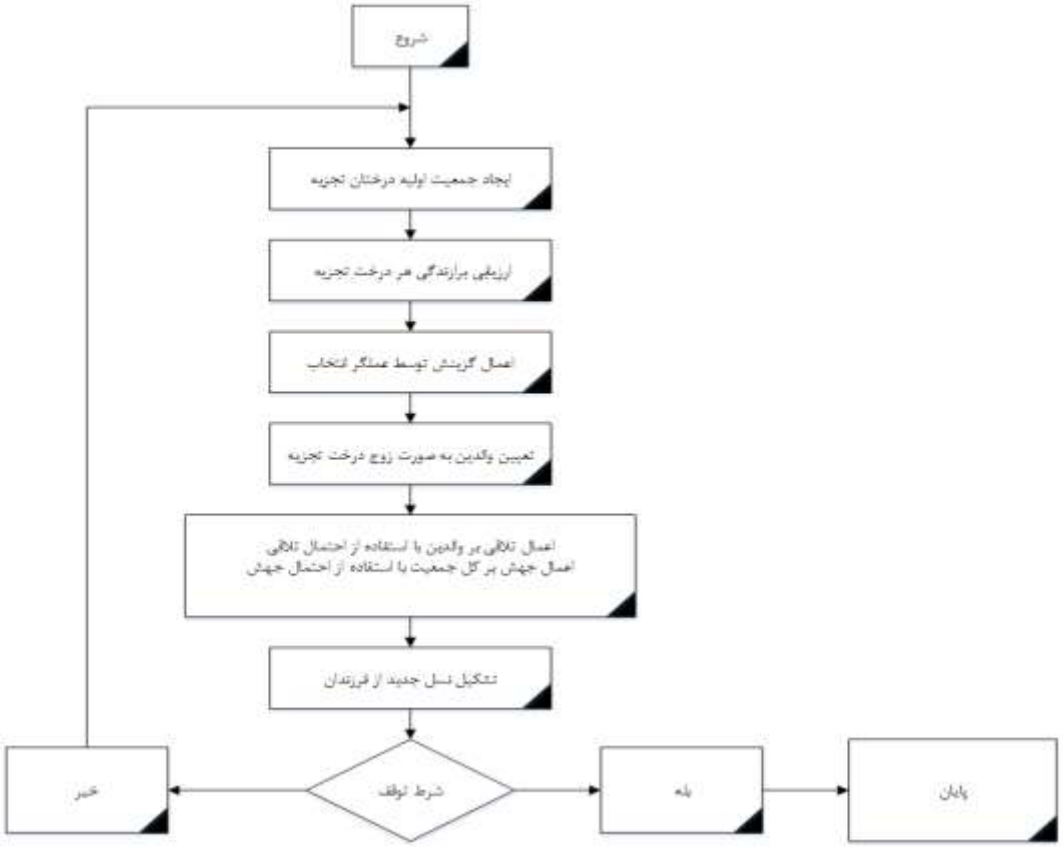
www.SID.ir

Archive of SID

بهترین راه حل ارائه می‌شود. اگر شرایط توقف یافت نشود، نخبه‌گزینی انجام می‌شود و باقی راه‌حل‌ها به فرآیندی گزینشی واگذار می‌شوند. این فرآیند برای چندین نسل تکرار می‌شود و با پیش رفتن نسل به جلو کیفیت جمعیت نیز به طور نسبی بهبود می‌یابد (فریرا، ۲۰۰۶).

برنامه‌ریزی بیان ژن نیز همانند الگوریتم ژنتیک و برنامه‌ریزی ژنتیک، یک الگوریتم ژنتیکی است به طوری که از جمعیتی از افراد استفاده می‌کند که آن‌ها را مطابق برازندگی انتخاب می‌کند و تغییرات ژنتیکی را با استفاده از یک یا چند عملگر ژنتیکی اعمال می‌نماید. تفاوت اساسی بین این سه الگوریتم، مربوط به ماهیت افراد آن‌ها می‌باشد به طوری‌که در الگوریتم ژنتیک، افراد رشته‌های خطی با طول ثابت (کروموزم‌ها) و در برنامه‌ریزی ژنتیک، نهاده‌های غیر خطی با اندازه‌ها و اشکال متفاوت (درختان تجزیه) می‌باشند، در حالی که در برنامه‌ریزی بیان‌ژن، افراد به صورت رشته‌های خطی با طول ثابت (ژنوم یا کروموزوم‌ها) کد گذاری شده و سپس به شکل نهادهای غیر خطی با اندازه‌ها و اشکال متفاوت (یعنی نمایش دیاگرام ساده یا بیان درختی) بیان می‌شوند (فریرا، ۲۰۰۱).

جهت کسب اطلاعات بیشتر با این مدل به منابع (فریرا، ۲۰۰۱؛ شیری و کایزی، ۲۰۱۱) مراجعه شود



شکل ۲- فلوچارت مدل برنامه‌ریزی ژنتیک.

www.SID.ir

Archive of SID

ساختارهای مختلف برای مدل‌سازی

برای اجرای مدل‌های برنامه‌ریزی ژنتیک و برنامه ریزی بیان ژن با توجه به پارامترهای ورودی ۱۰ نوع ترکیب مختلف مطابق جدول (۲) مورد بررسی قرار گرفت
.

جدول ۲- ترکیب¬های مختلف مدل¬های GP و GEP.

ترکیب	ورودی	خروجی
۱	Q_{t-2}, Q_{t-1}, Q_t	Q_{t+1}
۲	P_t, Q_t	Q_{t+1}
۳	$P_{t-1}, P_t, Q_{t-1}, Q_t$	Q_{t+1}
۴	$P_{t-2}, P_{t-1}, P_t, Q_{t-2}, Q_{t-1}, Q_t$	Q_{t+1}
۵	T_t, P_t, Q_t	Q_{t+1}
۶	$T_{t-1}, T_t, P_{t-1}, P_t, Q_{t-1}, Q_t$	Q_{t+1}
۷	$T_{t-2}, T_{t-1}, T_t, P_{t-2}, P_{t-1}, P_t, Q_{t-2}, Q_{t-1}, Q_t$	Q_{t+1}
۸	E_t, T_t, P_t, Q_t	Q_{t+1}
۹	$E_{t-1}, E_t, T_{t-1}, T_t, P_{t-1}, P_t, Q_{t-1}, Q_t$	Q_{t+1}
۱۰	$E_{t-2}, E_{t-1}, E_t, T_{t-2}, T_{t-1}, T_t, P_{t-2},, QP_{t-1_{t-2}}, Q_{t-1},$	Q_{t+1}

پارامترهای E_t, T_t, P_t, Q_t به ترتیب جریان، بارش، دما و تبخیر رودخانه دز در دوره زمانی روزانه و $E_{t-2}, E_{t-1}, E_t, T_{t-2}, T_{t-1}, T_t, P_{t-2}, P_{t-1}, P_t, Q_{t-2}, Q_{t-1}, Q_t$ و به ترتیب جریان، بارش، دما و تبخیر در دوره های زمانی گذشته و Q_{t+1} جریان در دوره آتی می‌باشد.

معیارهای ارزیابی مدل‌ها

برای ارزیابی مدل‌ها از سه معیار، ضریب تبیین (R2)، جذرمیانگین مربعات خطا (RMSE) و شاخص کلی عملکرد مدل (OI) استفاده شد. در زیر روابط این معیارها ارائه شده است.

$$\mathrm{RMSE} = \sqrt{\frac{\sum(\mathrm{Q_{obs}} - \mathrm{Q_{pre}})^2}{\mathrm{n}}} \quad (۲)$$

$$\mathrm{R^2} = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{Q_{obs}} - \mathrm{Q_{pre}})^2}{\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{Q_{obs}} - \overline{\mathrm{Q}})^2} \quad (۳)$$

$$\mathrm{OI} = \frac{1}{2}\left(2 - \frac{RMSE}{Q_{max} - Q_{min}} + \frac{\sum_{i=1}^{N}(Q_{obs} - Q_{pre})^2}{\sum_{i=1}^{N}(Q_{obs} - \bar{Q})^2}\right) \quad (۴)$$

در روابط فوق، n : تعداد داده‌ها، $\mathrm{Q_{obs}}$: داده‌های مشاهداتی، $\overline{\mathrm{Q}}$: میانگین داده‌های مشاهداتی، $\mathrm{Q_{pre}}$: داده‌های محاسباتی، Q_{max}: حداکثر داده‌ها، Q_{min}: حداقل داده‌ها می‌باشد. ضریب $\mathrm{R^2}$ میزان انطباق داده-هایی که توسط مدل به وجود می‌آید و داده‌های واقعی رانشان می‌دهد. RMSE جذر میانگین مربع خطاهای داده‌های محاسباتی و مشاهداتی را بیان می‌کند. واضح است که هر چه مقدار این عدد کمتر باشد به تبع آموزش و شبیه‌سازی داده‌ها بهتر صورت گرفته است. OI شاخص کلی عملکرد مدل است. هر چه RMSE کمتر باشد، پیش‌بینی دقیق‌تر است. OI بین مقادیر ∞- تا یک تغییر می‌کند. هر چه مقدار شاخص کلی عملکرد مدل به یک نزدیک‌تر باشد نشان‌دهنده یک تناسب کامل بین مقادیر تجربی و پیش‌بینی می‌باشد (مَتَر و المود، ۲۰۱۵).

نتایج و بحث

برای اجرای مدل برنامه‌ریزی ژنتیک از نرم‌افزار GPdotNET(v4.0beta3) و برای اجرای مدل برنامه

www.SID.ir

Archive of SID

ریزی بیان ژن از نرم‌افزار GeneXproTools (v5.0) استفاده شده است.

در جدول (۳) ارائه شده است. در ترکیب شماره ۱ که فقط از داده‌های جریان با تأخیرهای زمانی استفاده شده است، ضریب تبیین ۰/۸۵۴ و جذر میانگین مربعات خطای ۰/۰۰۳۷ بوده است. در ترکیب ۲ که از داده‌های بارش و جریان استفاده شده است مدل دارای ضریب تبیین ۰/۷۳۴ و جذرمیانگین مربعات خطای ۰/۰۳۰۰ بوده است. در ترکیب ۳ که از داده‌های بارش و جریانبا یک گام تأخیر زمانی استفاده شد عملکرد مدل نسبت به ترکیب ۲ دارای عملکرد بهتری داشته است. در ترکیب ۴ که از داده‌های بارش و جریان با دو گام زمانی تأخیر استفاده شد نشان داد که عملکرد مدل نسبت به ترکیب ۲ و ۳ دارای عملکرد بهتری بوده است. این مطلب به این دلیل می‌باشد که پارامتر جریان دارای تأثیر بیشتری بر خروجی است و گام‌های زمانی تأخیر این پارامتر نیز دارای تأثیر بیشتری از گام‌های زمانی تأخیر بارش می‌باشد. در ترکیب ۵ که از پارامترهای بارش، دما و جریان استفاده شد عملکرد مدل نسبت به ترکیب ۴ کمی کاهش یافت. در ترکیب ۶ که از داده‌های بارش، دما و جریان با یک گام تأخیر زمانی استفاده شد عملکرد مدل نسبت به ترکیب ۵ بهتر شد. در ترکیب ۷ که از داده های بارش، دما و جریان با دو گام زمانی تأخیر استفاده شد عملکرد مدل کاهش یافت و این مطلب را می‌رساند که داده دما را بیشتر از یک گام زمانی نباید تأخیر داد.

در ترکیب ۸ که از پارامترهای بارش، دما، تبخیر و جریان استفاده شد، عملکرد مدل خوب بوده است. در ترکیب‌های ۹ و ۱۰ که از داده‌های بارش، دما، تبخیر و جریان به ترتیب با یک و دو گام زمانی تأخیر استفاده شد به ترتیب عملکرد مدل افزایش و کاهش یافت. این مطلب نشان می‌دهد که استفاده از داده‌های تبخیر و دما با بیش از یک گام زمانی نه تنها عملکرد مدل را افزایش نمی‌دهند بلکه باعث کاهش عملکرد مدل می‌شوند. بر اساس نتایج جدول (۳) ترکیب ۱ دارای بهترین

نتایج مدل برنامه‌ریزی ژنتیک

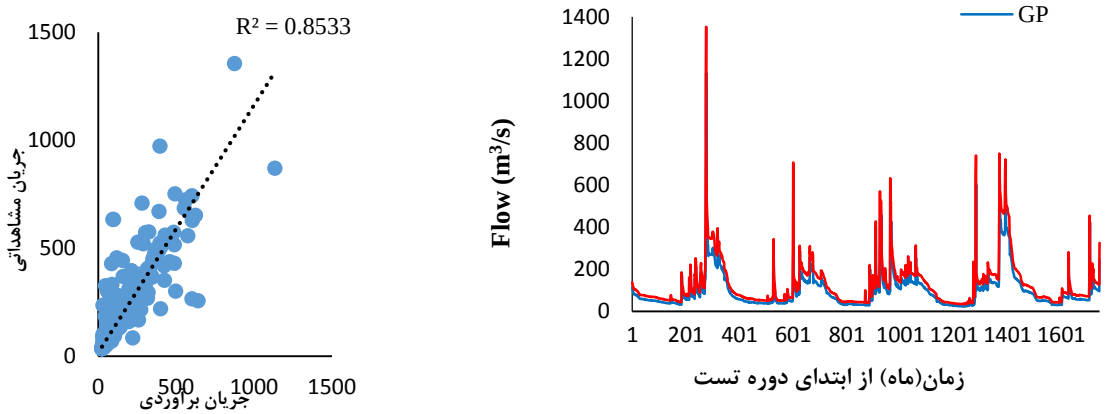
ترکیب‌های مختلف برای مدل برنامه‌ریزی ژنتیک مورد بررسی قرار گرفت. نتایج حاصل از اجرای این مدل عملکرد برای مدل برنامه‌ریزی ژنتیک بوده است و بعد از آن ترکیب‌های ۵ و ۹ دارای بهترین عملکرد بوده‌اند یعنی اینکه پارامتر جریان بیشترین تأثیر را بر خروجی مدل داشته است. شکل (۳) مقایسه مقادیر مشاهداتی را با مقادیر برآوردی برای ساختار برتر مدل برنامه‌ریزی ژنتیک در مرحله تست نشان می‌دهند. با توجه به این شکل عملکرد مدل مناسب بوده است. در شکل (۴) ساختار درختی برای ترکیب برتر مدل برنامه‌ریزی ژنتیک نشان داده است.

www.SID.ir

Archive of SID

جدول ۳- بهترین ساختار مدل برنامه‌ریزی ژنتیک.

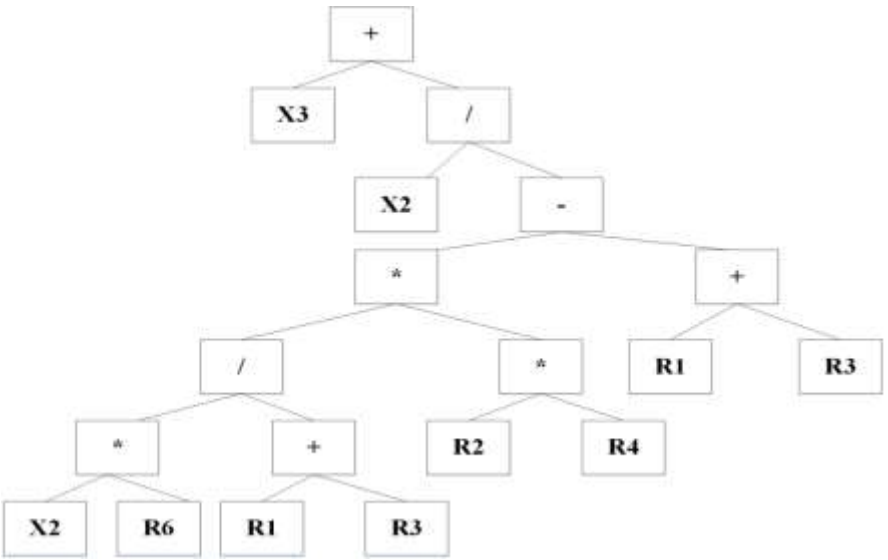
ترکیب	ضریب تبیین		جذر میانگین مربعات خطا	
	آموزش	تست	آموزش	تست
۱	۰/۸۷۸	۰/۸۵۴	۰/۰۰۹۸	۰/۰۰۳۷
۲	۰/۸۰۶	۰/۷۳۴	۰/۰۳۷۲	۰/۰۳۰۰
۳	۰/۸۱۶	۰/۷۵۶	۰/۰۳۲۳	۰/۰۲۳۹
۴	۰/۸۶۵	۰/۸۳۵	۰/۰۱۰۱	۰/۰۱۱۵
۵	۰/۸۵۱	۰/۸۲۵	۰/۰۲۶۹	۰/۰۱۵۴
۶	۰/۸۷۹	۰/۸۳۲	۰/۰۰۹۶	۰/۰۱۲۷
۷	۰/۸۲۳	۰/۷۵۲	۰/۰۲۹۲	۰/۰۲۶۴
۸	۰/۸۲۳	۰/۸۲۱	۰/۰۳۰۰	۰/۰۱۶۸
۹	۰/۸۵۸	۰/۸۴۸	۰/۰۱۱۱	۰/۰۱۰۴
۱۰	۰/۸۷۱	۰/۸۳۰	۰/۰۱۰۲	۰/۰۱۳۶



شکل ۳- مقایسه نتایج مدل برنامه‌ریزی ژنتیک با مقادیر مشاهده‌ای، مرحله تست.

www.SID.ir

Archive of SID



شکل۴- ساختار درختی ترکیب برتر مدل برنامه‌ریزی ژنتیک.

یکی از قابلیت مدل برنامه‌ریزی ژنتیک استخراج روابط بین پارامترهای ورودی و خروجی می‌باشد. این روابط برای ترکیب‌های مختلف اجرا شده در این پژوهش استخراج شده و در جدول (۴) به تفکیک ارائه شده است. با استفاده از این روابط به راحتی می‌توان برای دوره‌های آتی اقدام به پیش‌بینی نمود.

www.SID.ir

Archive of SID

جدول ۴- روابط استخراجی از اجرای مدل برنامه‌ریزی ژنتیک

ترکیب	رابطه استخراجی
1	$(Q_t + (\frac{Q_{t-1}}{\left(\left(\frac{5.39805 Q_{t-1}}{13.18345}\right) \times (25.68799)\right) - (13.18345)}))$
2	$(((\frac{Q_t}{1.783221 \times (P_t + Q_t) + 5.98407}) \times (5.75671 - (Q_t - 0.302375))$
3	$(((P_{t-1} + (3.52472 - \left((6.88405 + P_{t-1}) \times \left(\frac{Q_{t-1}}{3.84679}\right)\right)))/3.52472) \times Q_{t)}$
4	$((\frac{Q_t}{1.04559})/1.04559)$
5	$((\frac{(6.77062 - Q_t) + (\frac{(0.26802 - Q_t) - 1.47579}{(6.97579 - 2.97049)})}{6.77062}) \times Q_t)$
6	$(((Q_{t-1} + Q_t) \times \left(\frac{Q_{t-1}}{(0.76514 + \left((P_t \times Q_{t-1}) - (4.99808 + Q_{t-1})\right))}\right)) + Q_t)$
7	$(Q_t - \left((T_{t-1} \times Q_t) \times Q_{t-1}\right))$
8	$((((11.56449) \times (9.53029 Q_t))/(9.06262 + (Q_t + E_t)))/(0.96189 \times (3.28603 + (9.06262 + (R5 \times (Q_t + E_t))))))$
9	$(((Q_{t-1}/(P_{t-1} + ((2.97081 + Q_t) - 6.57088)))/2.97081) + Q_t)$
10	$(Q_t - (Q_t \times (Q_t \times Q_{t-1})))$

نتایج مدل برنامه‌ریزی بیان ژن

نتایج حاصل از بررسی ترکیب‌های مختلف مدل برنامه‌ریزی بیان ژن در جدول (۵) ارائه شده است. ترکیب شماره ۱ دارای ضریب تبیین ۰/۷۹۷ و جذر میانگین مربعات خطای ۰/۰۰۴۳ بوده است. ترکیب ۲ دارای ضریب تبیین ۰/۷۹۹ و جذر میانگین مربعات خطای ۰/۰۰۹۹ بوده است یعنی اینکه علاوه بر پارامتر جریان که در ترکیب ۱ استفاده شده بود اضافه نمودن پارامتر بارش باعث بهتر شدن عملکرد شده است. نتایج ترکیب شماره ۳ نشان می‌دهد که تأخیر با یک گام زمانی نسبت به ترکیب ۲ باعث افزایش عملکرد مدل شده است. در ترکیب ۴ که تأخیر با دو گام زمانی استفاده شده است عملکرد مدل نسبت به ترکیب ۲ و ۳ بهتر شده است. یعنی استفاده کردن تأخیر با دو گام زمانی نتیجه بهتری نسبت به تأخیر با یک گام زمانی داشته است. در ترکیب ۵ با اضافه نمودن پارامتر دما عملکرد مدل اندکی کاهش یافت. در ترکیب ۶ با اضافه نمودن یک گام زمانی تأخیر عملکرد مدل نسبت به ترکیب ۵ اندکی بهتر شد. نتایج ترکیب ۷ نشان می‌دهد که استفاده کردن از دوگام زمانی تأخیر با استفاده از داده دما باعث کاهش عملکرد مدل نسبت به ترکیب ۶ شده است. در ترکیب ۸ که از چهار پارامتر دما، بارش، تبخیر و جریان استفاده شد عملکرد مدل مناسب بود ولی در ترکیب‌های ۹ و ۱۰ با افزایش تأخیرهای زمانی عملکرد مدل کاهش یافت.

با توجه به نتایج ترکیب‌های مختلف می‌توان به این نتیجه رسید که در صورتی که از خود پارامترها بدون تأخیر زمانی استفاده شود استفاده از پارامترهای دما، بارش، تبخیر و جریان مناسب می‌باشد در صورتی که دسترسی به این داده‌ها وجود نداشته باشد می‌توان از داده‌های بارش، دما و جریان استفاده نمود و در صورت نبود داده دما، استفاده از داده بارش و جریان برای این کار مناسب خواهد بود.

به طور کلی با توجه به نتایج حاصل از ترکیب‌های مختلف در این پژوهش، ترکیب ۴ دارای بهترین عملکرد بوده است. و این به این دلیل می‌باشد که پارامتر جریان

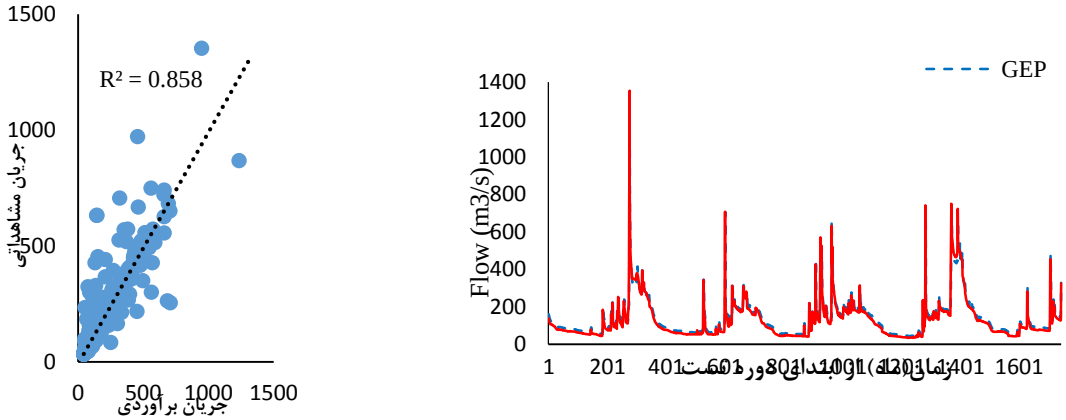
www.SID.ir

Archive of SID

دارای تأثیر بیشتری بر خروجی می‌باشد و گام‌های زمانی تأخیر این پارامتر نیز دارای تأثیر بیشتری از گام-های زمانی تأخیر دیگر پارامترها می‌باشد. شکل (۵) مقایسه مقادیر مشاهداتی را با مقادیر برآوردی برای ساختار برتر مدل برنامه‌ریزی بیان ژن در مرحله تستنشان می‌دهند. با توجه به این شکل عملکرد مدل مناسب بوده است.

جدول ۵- بهترین ساختار مدل برنامه‌ریزی بیان ژن.

ترکیب	ضریب تبیین		جذر میانگین مربعات خطا	
	آموزش	تست	آموزش	تست
۱	۰/۸۲۷	۰/۷۹۷	۰/۰۰۹۸	۰/۰۰۴۳
۲	۰/۸۸۵	۰/۷۹۹	۰/۰۰۹۱	۰/۰۰۹۹
۳	۰/۸۷۸	۰/۸۵۲	۰/۰۰۹۴	۰/۰۰۳۱
۴	۰/۹۰۳	۰/۸۵۷	۰/۰۰۸۱	۰/۰۰۳۰
۵	۰/۸۶۵	۰/۸۵۲	۰/۰۰۹۶	۰/۰۰۳۰
۶	۰/۸۸۵	۰/۸۵۶	۰/۰۰۹۳	۰/۰۰۳۲
۷	۰/۹۰۱	۰/۸۵۴	۰/۰۰۸۸	۰/۰۰۳۰
۸	۰/۸۶۵	۰/۸۵۲	۰/۰۱۰۰	۰/۰۰۳۰
۹	۰/۸۳۰	۰/۷۶۰	۰/۰۱۰۵	۰/۰۱۰۳
۱۰	۰/۷۶۷	۰/۷۱۴	۰/۰۱۲۲	۰/۰۰۸۹

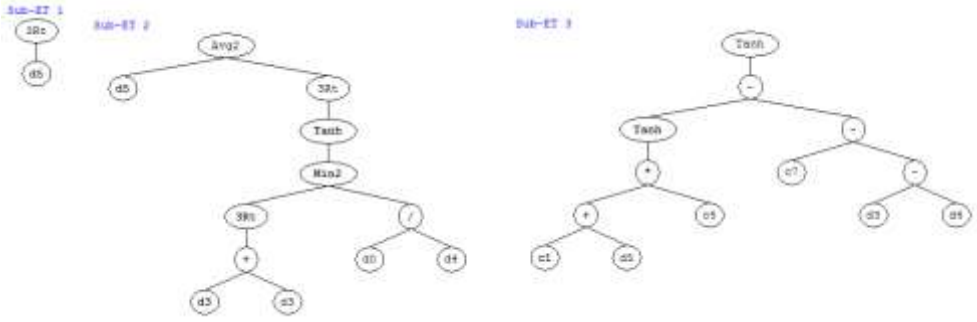


شکل ۵- مقایسه نتایج مدل برنامه‌ریزی بیان ژن با مقادیر مشاهده‌ای، مرحله تست.

در شکل (۶) ساختار درختی برای ترکیب برتر مدل برنامه‌ریزی بیان ژن نشان داده است. ساختار درختی کمک می‌کند که در هر مرحله جمعیت اولیه به صورت ساختار خطی ساده بیان شود و تمام تغییرات تنها بر روی ساختارهای ساده انجام گیرد، در نتیجه نیازی به ساختارهای نسبتاً پیچیده برای گسترش در هر مرحله نخواهد بود. در این شکل مقادیر C مربوط به ضرایب برنامه‌ریزی بیان ژن و مقادیر d مربوط به پارامترهای ورودی می‌باشد که با تأخیرهای مختلف ارائه شده است

www.SID.ir

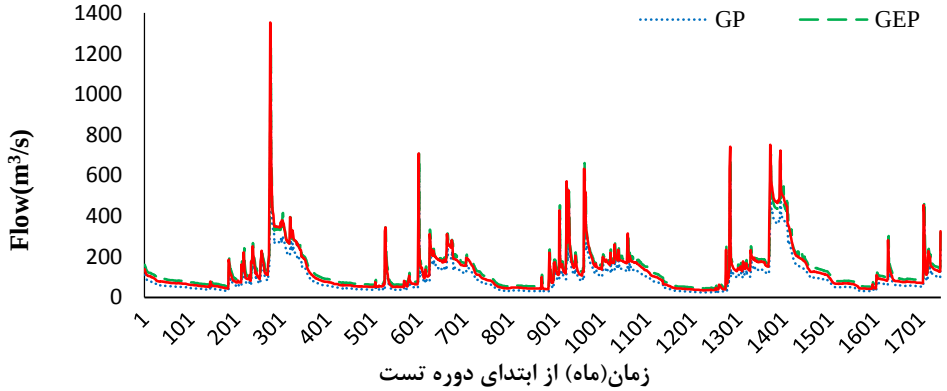
Archive of SID



شکل ۶- ساختار درختی ترکیب برتر مدل برنامه‌ریزی بیان ژن.

مقایسه نتایج مدل‌های استفاده شده

در این قسمت به منظور این‌که مقایسه‌ی مناسبی بین عملکرد دو مدل صورت گیرد، مقایسه مدل‌ها به دو صورت گرافیکی و کمی صورت گرفته است. در شکل (۷) عملکرد دو مدل با مقادیر مشاهداتی مورد مقایسه قرار گرفته است. همان‌طوری‌که در شکل (۷) دیده می‌شود مقادیر برآوردی با مدل برنامه‌ریزی بیان ژن به مقادیر مشاهداتی نزدیک‌تر می‌باشد به همین دلیل این مدل دارای عملکرد بهتری نسبت به مدل برنامه ریزی ژنتیک می‌باشد. با مقایسه مقادیر حداکثری و حداقلی جریان نیز این نتیجه حاصل می‌شود که عملکرد مدل برنامه‌ریزی بیان ژن بهتر از مدل برنامه‌ریزی ژنتیک بوده است. برای مقایسه کمی بین دو مدل از سه معیار ارزیابی به نام‌های ضریب تبیین، جذر میانگین مربعات خطا و شاخص کلی عملکرد مدل استفاده شد. نتایج این بررسی در جدول (۶) ارائه شده است.



شکل ۷- مقایسه مدل‌های استفاده شده در این تحقیق.

www.SID.ir

Archive of SID

جدول۶- مقایسه مدل‌های استفاده شده در این تحقیق.

مدل	شاخص کلی عملکرد مدل		جذرمیانگین مربعات خطا		ضریب تبیین	
	آموزش	تست	آموزش	تست	آموزش	تست
GP	۰/۸۷۴۷	۰/۸۶۲۹	۰/۰۰۹۸	۰/۰۰۳۷	۰/۸۷۷۹	۰/۸۵۳۳
GEP	۰/۹۱۴۸	۰/۹۰۸۷	۰/۰۰۸۱	۰/۰۰۳۰	۰/۹۰۱۶	۰/۸۵۸۰

با توجه به جدول (۶) و مقایسه مدل‌های برنامه ریزی ژنتیک و برنامه‌ریزی بیان‌ژن مشاهده می‌شود که براساس ضریب تبیین و معیار OI که هرچه مقادیر به یک نزدیک‌تر باشند عملکرد بهتری دارند، مدل برنامه ریزی بیان‌ژن دارای ضریب تبیین و معیار OI بیشتری بوده و لذا دارای عملکرد بهتری می‌باشد. همچنین با توجه به پارامتر جذر میانگین مربعات خطا، مشاهده می‌شود که مدل برنامه‌ریزی بیان‌ژن دارای جذر میانگین مربعات خطای کمتری بوده و در نتیجه عملکرد بهتری داشته است.

نتیجه‌گیری

در این تحقیق از دو مدل ژنتیکی برنامه‌ریزی ژنتیک و برنامه‌ریزی بیان‌ژن جهت مدل‌سازی جریان روزانه رودخانه دز استفاده شده است. هدف از این مطالعه نشان دادن تفاوت‌های بین دو مدل و همچنین بررسی توانایی آن‌ها برای مدل‌سازی جریان رودخانه دز بوده است. برای این کار از داده‌های روزانه‌ی دما، بارش، تبخیر و جریان رودخانه در ایستگاه تله زنگ استفاده شد. نتایج نشان داد که مدل برنامه‌ریزی بیان ژن نسبت به مدل برنامه‌ریزی ژنتیک دارای عملکرد بهتری می‌باشد. بررسی ساختارهای مختلف نشان داد که از بین پارامترهای استفاده شده در این تحقیق پارامتر جریان دارای بیشترین اهمیت و سپس پارامترهای بارش، دما و تبخیر در مرحله بعدی اهمیت قرار دارند. به طورکلی نتایج نشان داد که مدل برنامه‌ریزی بیان ژن برای مدل‌سازی و پیش‌بینی جریان رودخانه قابلیت خوبی دارد.

تقدیر و تشکر

نویسندگان این اثر، از سازمان آب و برق خوزستان و شرکت آب غدیر خوزستان به دلیل همکاری‌های لازم و در اختیار قرار دادن اطلاعات این مورد نیاز این مطالعه، کمال تقدیر و تشکر را بعمل می‌آورند.

منابع

احمدی، ف.، ف. رادمنش و ر. میرعباسی‌نجف‌آبادی. ۱۳۹۳. مقایسه روش‌های برنامه‌ریزی ژنتیک و ماشین بردار پشتیبان در پیش‌بینی جریان روزانه رودخانه (مطالعه موردی: رودخانه باراندوز چای). نشریه آب و خاک (علوم و صنایع کشاورزی). (۶) ۲۸: ۱۱۷۱-۱۱۶۲.

داننده مهر، ع. و م.ر. مجدزاده طباطبائی. ۱۳۸۹. بررسی تأثیر توالی دبی روزانه در پیش‌بینی جریان رودخانه‌ها با استفاده از برنامه ریزی ژنتیک. نشریه آب و خاک (علوم و صنایع کشاورزی)، (۲) ۲۴ : ۳۳۳ -۳۲۵.

سلگی، ا. ۱۳۹۳. پیش‌بینی جریان رودخانه با مدل ترکیبی شبکه عصبی مصنوعی-موجک و مقایسه آن با روش‌های سیستم استنتاج فازی عصبی تطبیقی و شبکه عصبی مصنوعی(مطالعه موردی: رودخانه گاماسیاب نهاوند). پایان‌نامه کارشناسی ارشد در رشته مهندسی منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهید چمران اهواز.

فربودنام، ن.، م.ع. قربانی، و م.ت. اعلمی. ۱۳۸۸. پیش‌بینی جریان رودخانه با استفاده از برنامه‌ریزی ژنتیک (مطالعه موردی: حوضه آبریزرودخانه لیقوان). مجله دانش آب و خاک، ۱۹ (۱) : ۱۲۳-۱۰۷.

کاوه‌کار، ش.، م.ع. قربانی، ا. اشرف‌زاده و ص. دربندی. ۱۳۹۲. شبیه‌سازی نوسانات تراز آب با استفاده از برنامه‌ریزی

www.SID.ir

Archive of SID

بیان ژن. نشریه مهندسی عمران و محیط زیست. ۴۳ (۳)، پیای۷۲: ۷۵-۶۹.

گلابی، م.ر.، ع.م. آخوندعلی، و ف. رادمنش. ۱۳۹۲. مقایسه عملکرد الگوریتم‌های مختلف شبکه عصبی مصنوعی در مدل‌سازی بارندگی فصلی (مطالعه موردی: ایستگاه‌های منتخب استان خوزستان). نشریه تحقیقات کاربردی علوم جغرافیایی. ۱۳(۳۰): ۱۶۹-۱۵۱.

Banzhaf, W., P. Nordin, P.E. Keller and F.D. Francone. 1998. Genetic programming. Kaufmann M, editor. San Francisco. CA. 512 p.

Ferreira, C. 2006. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence (Studies in Computational Intelligence). ed n, editor. Springer-Verlag New York, Inc. Secaucus, NJ, USA.

Ferreira, C. 2001. Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst, 13:87-129.

Ghorbani, M.A., O. Kisi and M.A. Aalinezhad. 2010. A probe into the chaotic nature of daily stream flow time series by correlation dimension and largest Lyapunov methods. Applied Mathematical Modelling, 34(12):4050-4057.

Shiri, J. and Ö. Kişi. 2011. Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Computers & Geosciences, 37(10):1692-701.

Shiri, J., Ö. Kişi, G. Landeras, J.J. López, A.H. Nazemi and L.C.P.M. Stuyt. 2012. Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). Journal of Hydrology, 414–415:302-316.

Mattar, M.A. and A.I. Alamoud. 2015. Artificial neural networks for estimating the hydraulic performance of labyrinth-channel emitters. Computers and Electronics in Agriculture, 114:189-201.

Zahiri, A. and H.M. Azamathulla. 2014. Comparison between linear genetic programming and M5 tree models to predict flow discharge in compound channels. Neural Comput & Applic, 24:413-420.

www.SID.ir

Archive of SID

فصلنامه علمی پژوهشی مهندسی آبیاری و آب ایران
سال نهم شماره سی و پنجم بهار ۱۳۹۸

Genetic Models Assessment in Order to Model River Flow

Saber Alidadi Deh kohneh[1], Abazar Solgi[2], Mehrnoush Shehni Darabi[3], Hiedar Zarei[4,1]

Abstract

Predicting rivers flow, as one of the main resources of water for human beings, has always been one of the important issues discussed in hydrology and water resources. So, different models have been used for modeling and predicting rivers flow.

In this study, two genetic models, Gene Expression Programming (GEP) and genetic programming (GP) were evaluated. Flow, temperature, precipitation, and evaporation data were used to model the daily flow in Tale Zang station. The results showed that GEP model, with R2=0.86 and RMSE=0.0030 m3/s, had the better performance than GP model, with R2=0.85 and RMSE=0.0037 m3/s. Moreover, the speed of GEP was more than GP and had ability to present results in a short time. By increasing the number of data, the speed of GP decreased and sometimes it couldn't present results, while GEP had the ability to work with more data and model the river flow.

Key words: Generally, the results showed that GEP model had good ability for modeling and predicting rivers flow.

.

[1] Corresponding Author Email: zareih@scu.ac.ir.
1- Water Supply Deputy of Khuzestan Water and Power Authority.
2- Ph.D. Student, Department of Water Resources Engineering, Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Iran.
3- Quality Control Expert, Water Transfer and Utilization of Ghadir Company of Khuzestan and Ph.D. Student, Department of Water Resources Engineering, Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran.
4- Assistant professor, Department of Hydrology and Water Resources, Faculty of Water Sciences Engineering, Shahid Chamran University of Ahvaz, Iran.

www.SID.ir