



پیش‌بینی جریان با استفاده از رویکرد تلفیقی موجک- برنامه‌ریزی بیان ژن و ارزیابی تأثیر پارامترهای هواشناسی بر کارایی آن

فربا آذرپیرا^{۱*}، سجاد شهابی^{۲*}

۱. دانشجوی کارشناسی ارشد مهندسی و مدیریت منابع آب، گروه مهندسی آب، دانشکده مهندسی عمران و نقشه‌برداری، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران.

۲. استادیار، گروه مهندسی آب، دانشکده مهندسی عمران و نقشه‌برداری، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران.

تاریخ پذیرش مقاله: ۱۳۹۹/۰۵/۰۹

تاریخ دریافت مقاله: ۱۳۹۹/۰۱/۰۹

چکیده

پیش‌بینی جریان رودخانه برای برنامه‌ریزی و مدیریت سیستم‌های منابع آب بسیار حائز اهمیت بوده و میزان دقت انجام این فرایند در درستی پیاده‌سازی اهداف برنامه‌ریزی شده نقش بی‌بدین دارد. از طرفی، محاسبات نرم قابلیت بالایی در مدل‌سازی فرایندهای هیدرولوژیک دارد. از این‌رو، در این پژوهش، مدل ترکیبی موجک- برنامه‌ریزی بیان ژن در مقایسه با رویکرد منفرد آن توسعه داده شد، تا جریان روزانه رودخانه خشک‌رود واقع در استان گیلان، را پیش‌بینی کند. بدین منظور، علاوه بر فرایند پیش‌پردازش داده‌های هیدرومتری، تأثیر پارامترهای هواشناسی در عملکرد و کارایی مدل نیز بررسی شد. هم‌چنین، پیش‌پردازش با ویژگی‌های مختلف و پیش‌بینی برای چهار زمان یک، دو، سه و شش روز انجام شد. برای ارزیابی عملکرد مدل‌ها از شاخص‌های آماری ضریب همبستگی (R)، شاخص توافق (I_a)، ضریب ناش-سانتلیف (NSE)، خطای مطلق میانگین (MAE)، جذر متوسط مربعات خطا (RMSE) و معیار جریان حدی (PFC) استفاده شد. نتایج نشان می‌دهد که استفاده از موجک برای پیش‌پردازش داده‌ها موجب ارتقای قابل‌ملاحظه کارایی مدل ترکیبی در مقایسه با مدل منفرد شده‌است، به گونه‌ای که ضریب همبستگی داده‌های ارزیابی برای زمان پیش‌بینی سه روز از ۰/۲۷ به ۰/۸۰ افزایش و هم‌زمان خطای مطلق میانگین از ۱/۴ به ۰/۸۰ مترمکعب در ثانیه کاهش یافته است. از طرف دیگر، پارامترهای هواشناسی موجب شده‌اند که مقادیر حدی در سری‌زمانی داده‌های دبی رودخانه به‌خوبی پیش‌بینی شده و کارایی مدل در پیش‌بینی مقادیر مذکور به‌مقدار قابل‌ملاحظه‌ای ارتقا یابد. نتایج حاصل از این مطالعه بیانگر آن است که رویکرد ترکیبی موجک- برنامه‌ریزی بیان ژن در ترکیب با پارامترهای هواشناسی، می‌تواند با کارایی بالایی در پیش‌بینی جریان به‌کار رود.

کلیدواژه‌ها: پیش‌پردازش، دبی روزانه، رودخانه خشک‌رود، سری‌زمانی.

Streamflow Forecasting Using Wavelet-Gene Expression Programming Hybrid Approach and Assessing the Effects of Meteorological Parameters on its Capability

Fariba Azarpira¹, Sajad Shahabi^{2*}

1. M.Sc. Student, Department of Water Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran.

2. Assistant Professor, Department of Water Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran.

Received: March 28, 2020

Accepted: July 30, 2020

Abstract

Forecasting stream flow is very important for planning and managing water resources systems and its precision has a significant role in accurately implementing the planned objectives. Besides, soft computing has a significant ability in modelling hydrologic processes. Therefore, in the present study, the Hybrid Wavelet-Gene Expression Programming model has been developed in comparison with its singular approach so that it forecasts the daily streamflow of Khoshkroud river located in Guilan province. For this purpose, in addition to the process of pre-processing hydrometric data, the effect of meteorological parameters on the model's performance and efficiency has been studied. Also, pre-processing was performed with different properties and for four durations of one, two, three and six days. Correlation coefficient (R), index of agreement (I_a), Nash-Sutcliffe coefficient (NSE), mean absolute error (MAE), root-mean-square error (RMSE) and peak flow criteria (PFC) statistical indices were used to assess the models' performances. The results show that using wavelet transform to pre-process hydrometric data will significantly improve the efficiency of the hybrid model in comparison with the singular model, such that the correlation coefficient of the validation data for three days has increased from 0.27 to 0.80 and similarly, the mean absolute error has decreased from 1.4 to 0.80 m³/Sec. On the other hand, meteorological parameters have caused the extreme values in the river's flow rate time series to be well modelled and their efficiency in the extreme values to be significantly increased. The results obtained from this research express that the hybrid model alongside the meteorological parameters can be successfully and efficiently used in flow forecasting.

Keywords: Daily discharge, Khoshkroud river, Preprocessing, Time series.

مقدمه

پیش‌بینی جریان از مطالعات پایه هیدرولوژیکی است که بسیاری از مباحث منابع آب، به آن وابسته است. از این‌رو، فرایند مدل‌سازی آن در زمان‌های مختلف نیازمند دقت قابل‌قبولی است. به دلیل ماهیت غیرخطی و پیچیده تخمین جریان آتی، این فرایند همواره در طول دهه‌های متمادی مورد توجه پژوهش‌گران مختلف بوده و بر این اساس، مدل‌های فیزیکی و تجربی متعددی جهت پیش‌بینی جریان با استفاده از داده‌های مختلف ارائه شده است (۱۴، ۱۹، ۲۵ و ۳۳)، که عملکرد آن‌ها وابسته به سری‌های زمانی است. سری‌های زمانی هیدرولوژیکی به‌طور معمول غیرایستا و حاوی اطلاعات اضافی هستند که در تحلیل این فرایند کاربرد نداشته و تنها منجر به پیچیدگی و وقت‌گیر شدن محاسبات می‌شوند، بنابراین استفاده از روش‌های پیش‌پردازش داده هم‌چون سری‌فوری، تبدیل موجک (۳۱ و ۱۲)، تجزیه حالت تجربی^۱ (EEMD) (۳۴) و ۳۹ و تجزیه حالت متغیر^۲ (VMD) (۳۶) می‌تواند به شناسایی دقیق‌تر ویژگی‌های سری زمانی و بهینه‌تر شدن عملکرد مدل‌ها، کمک کند (۱۷). با توسعه رویکردهای محاسبات نرم در دو دهه اخیر، پژوهش‌های زیادی جهت بررسی قابلیت آن در پیش‌بینی جریان انجام شده است (۳۷) که از جمله می‌توان به موارد زیر اشاره کرد.

پیش‌بینی جریان روزانه حوضه کروکاوک^۳ در ترکیه با مدل‌های برنامه‌ریزی بیان‌ژن^۴ (GEP)، شبکه عصبی مصنوعی^۵ (ANN) و سیستم استنتاج فازی-عصبی تطبیقی^۶ (ANFIS) توسط کیسی و همکاران (۱۶) نشان داد که مدل GEP نسبت به دو روش دیگر دارای برتری می‌باشد. مطالعه شعیب و همکاران (۳۲) در خصوص پیش‌بینی جریان چهار حوضه با اقلیم متفاوت در ترکیه با روش موجک برنامه‌ریزی بیان ژن^۷ (WGEP) و با استفاده از موجک‌های مادر و سطوح تجزیه مختلف، قابلیت

بالای این مدل ترکیبی را ثابت نمود. هم‌چنین می‌توان به مطالعات نورانی و همکاران (۲۴)، مهر و نورانی (۲۱) و هادی و تامبول (۱۳) اشاره کرد که در آن‌ها از GP و برنامه‌ریزی ژنتیک چند ژنی^۸ (MGGP) به‌عنوان تحلیل‌گر حساسیت در روش‌های موجک-برنامه‌ریزی ژنتیک-شبکه عصبی (WGPNN)، برنامه‌ریزی ژنتیک چند ژنی-میانگین متحرک سریع-نمودار پرتو^۹ (Pareto-Optimal MA-MGGP) و موجک پیوسته^{۱۰}-برنامه‌ریزی ژنتیک چند ژنی-شبکه عصبی (CWT-MGGP-ANN) استفاده شده و کارایی مدل‌های ترکیبی بالاتر از مدل‌های منفرد گزارش شده است. نتایج مطالعه شهابی و همکاران (۳۰) نیز قابلیت بالای مدل ترکیبی موجک برنامه‌ریزی ژنتیک را در پیش‌بینی سری‌های زمانی نشان داده است. در پژوهشی دیگر، چمنی و روشنگر (۲) مدل تلفیقی تجزیه مد تجربی یکپارچه کامل-گاوسی را در پیش‌بینی زمانی و مکانی رودخانه آرکانزاس واقع در ایالات متحده به‌کار بردند، که نتایج حاصل نشان می‌دهد که استفاده از روش CEEMD باعث بهبود عملکرد و کارایی مدل‌های مورد استفاده شده است.

درخصوص مطالعات صورت‌گرفته در داخل کشور هم می‌توان به این موارد اشاره کرد. پژوهش نبی‌زاده و همکاران (۴) در پیش‌بینی جریان روزانه رودخانه لیکوان با استفاده از سامانه استنباط فازی-عصبی تطبیقی، کارایی قابل‌قبول این رویکرد را نشان داد. آبابایی و همکاران (۱) جهت شبیه‌سازی جریان روزانه ورودی به سد طالقان، از چهار مدل شبکه عصبی و دو مدل همراشتین-واینر (ترکیب اطلاعات) استفاده کردند. نتایج نشان دادند که مدل‌های ترکیب اطلاعات از دقت بالاتری برخوردار هستند و اثرات تغییر اقلیم، برخلاف مدل‌های منفرد، بر عملکرد آن‌ها تأثیرگذار نیست. شریف‌آذری و عراقی‌نژاد (۳) جهت شبیه‌سازی سری‌های زمانی هیدرولوژیکی

پیش‌بینی جریان با استفاده از رویکرد تلفیقی موجک- برنامه‌ریزی بیان ژن و ارزیابی تأثیر پارامترهای هواشناسی بر کارایی آن

پیش‌پردازش کرده، تا تأثیر انتخاب این دو پارامتر (موجک مادر و سطح تجزیه) در میزان درستی عمل پیش‌پردازش و همچنین، نحوه عملکرد شبکه GEP در برابر این تغییرات مشخص شود. علاوه بر آن، تأثیر پارامترهای هواشناسی بر عملکرد و کارایی رویکردهای WGEP و GEP، به‌خصوص در مقادیر حدی بررسی شده است.

مواد و روش‌ها

برنامه‌ریزی بیان ژن (GEP)

برنامه‌ریزی بیان ژن برگرفته از نظریه داروین و تکامل‌یافته برنامه‌ریزی ژنتیک است. نوعی شبکه داده‌کاو است که ساختاری جعبه‌سیاهی نداشته (مزیت و برتری آن نسبت به شبکه عصبی) و می‌تواند رابطه بین پارامترها را به‌صورت معادله ریاضی نمایش دهد. همچنین، با شرایط مختلف سازگار است (۲۲). کروموزوم‌های این شبکه ساختاری درختی دارند که از گره‌ها و ترمینال‌ها تشکیل شده‌اند. گره‌ها شامل توابع ریاضی و ترمینال‌ها شامل پارامترها و ثابت‌ها هستند که برای تشکیل جمعیت اولیه به‌صورت تصادفی انتخاب می‌شوند. اجرای این شبکه نیازمند تعیین ساختار کروموزوم‌ها، تابع برازش، نوع توابع ریاضی و میزان عملگرهای ژنتیکی است که در حین اجرای آن، درصدی از کروموزوم‌ها می‌توانند با عملگرهای ژنتیکی (پیوند و جهش) ساختار جدیدی را به‌خود اختصاص بدهند و نسل جدیدی را تولید کنند. ایجاد جمعیت جدید تا جایی ادامه خواهد یافت که خطای مدل و یا تعداد تکرارهای مجاز آن به‌میزان از پیش تعیین‌شده برسد (۱۱).

تبدیل موجک^{۱۴}

تبدیل موجک یکی از مباحث مهم ریاضی است، که در دو دهه اخیر در هیدرولوژی و مطالعه سری‌های زمانی

رودخانه مهاباد، مدل اولیه خودگردان‌ساز نزدیک‌ترین همسایه را توسعه دادند. مدل پیشنهادی، علاوه بر عملکرد مطلوب‌تر نسبت به مدل کلاسیک آن، از قابلیت تولید اعداد خارج از محدوده ثبت‌شده در سری‌زمانی و شبیه‌سازی چندمتغیره برخوردار است. مطالعات روانسالار و همکاران (۲۶) به‌منظور پیش‌بینی جریان ماهانه رودخانه بهشهر با استفاده از روش‌های موجک برنامه‌ریزی خطی ژنتیک^{۱۱} (WLGP) و موجک شبکه عصبی^{۱۲} (WANN) و همچنین عبدالمهدی و همکاران (۵) برای پیش‌بینی جریان روزانه رودخانه‌های بهشت‌آباد و جنگان ایران، با استفاده از شبکه‌های GEP و WANN، نشان دادند که پیش‌پردازش داده‌های ورودی سبب ارتقای عملکرد مدل‌ها شده است. نتایج مدل ترکیبی GEP-GA توسط مهر (۲۰) نیز نشان دادند که بهینه‌سازی خارجی می‌تواند بر عملکرد شبکه GEP مؤثر باشد. به‌طوری‌که، این مدل نتایج مطلوب‌تری نسبت به شبکه‌های GEP، رگرسیون خطی چندگانه^{۱۳} (MLR) و GEP-LR ارائه داده است.

در پژوهش حاضر، جریان روزانه رودخانه خشک‌رود که یکی از رودخانه‌های دائمی استان گیلان است و از نظر اقتصادی و آبی‌پروری جایگاه ویژه‌ای در منطقه دارد، با استفاده از رویکرد برنامه‌ریزی بیان ژن و همچنین رویکرد ترکیبی موجک- برنامه‌ریزی بیان ژن (WGEP) پیش‌بینی شده است. جهت مدل‌سازی از داده‌های هیدرومتری (دبی) و پارامترهای هواشناسی مؤثر در چرخه هیدرولوژی (بارش، دمای میانگین هوای روزانه، فشار و نقطه شبنم) استفاده شده است. به این صورت که در هر دو مدل منفرد و ترکیبی، سناریوهایی با استفاده از داده‌های هیدرومتری تنها و سناریوهایی با اثر توأم داده‌های هیدرومتری و هواشناسی تشکیل شده است. در مدل ترکیبی، تبدیل موجک با استفاده از موجک‌های مادر مختلف و چندین سطح تجزیه، داده‌های هیدرومتری را

که در آن، Ψ تبدیل موجک گسسته، $f(t)$ سری زمانی، j سطح تجزیه و k طول از مبدأ می باشد.

رویکرد ترکیبی موجک- برنامه ریزی بیان ژن

به مدلی ترکیبی گفته می شود که نتایج یک شبکه یا الگوریتم بر عملکرد سایر شبکه ها اثرگذار باشد و یا قابلیت روش های متعدد با هم ترکیب شوند. برنامه ریزی بیان ژن نوعی شبکه داده کاوی است که نوع، تعداد و ویژگی پارامترهای وارد شده به آن بر عملکرد و دقت شبکه مؤثر است. بنابراین استفاده از تبدیل موجک برای تحلیل و پیش پردازش داده ها می تواند بر عملکرد برنامه ریزی بیان ژن تأثیر گذاشته و یک مدل ترکیبی را ایجاد کند. در مدل ترکیبی مدنظر انتخاب نوع موجک مادر و سطح تجزیه تبدیل موجک از عوامل مهم و اثرگذار در نتیجه شبکه است. بنابراین با انتخاب بهترین شبیه ترین موجک مادر به سیگنال اصلی و مناسب ترین سطح تجزیه که حاوی فرکانس های مؤثر سیگنال اصلی است، سری تجزیه می شود و داده های استخراج شده از تبدیل موجک به برنامه ریزی بیان ژن اعمال می شوند و مدلی جهت شبیه سازی جریان ارائه می گردد. شکل (۱) فلوجارت این مدل ترکیبی را نشان می دهد.

به شکل گسترده ای مورد توجه قرار گرفته است (۲۹). در ادامه، تبدیل موجک ها به اختصار شرح داده می شود.

تبدیل موجک برای غلبه بر نواقص سری فوری در تجزیه سری های زمانی ارائه گردیده است، پس قادر به ثبت زمان و فرکانس به طور هم زمان و تغییر طول پنجره های نمونه برداری است، به طوری که میانگین فرکانس های هر پنجره صفر باشد. این تبدیل، فرایند پیش پردازش را با تصویر و گسسته سازی سری زمانی روی توابع پایه شناخته شده انجام داده و ویژگی های آن را بر اساس ضوابط و شکل تابع پایه استخراج کرده و در قالب ضرایبی که مجموع آن ها برابر با مقدار اولیه سری در همان زمان است، نمایش می دهد (۹ و ۱۸). لذا این تبدیل به سه نوع پارامتر مهم مقیاس^{۱۵} جهت مشخص کردن طول بازه های تصویر سازی، ضریب انتقال^{۱۶} (سطح تجزیه داده ها) و موجک مادر^{۱۷} ($\Psi(t)$) با ویژگی زیر نیاز دارد:

$$\int_{-\infty}^{+\infty} \Psi(t) dt = 0 \quad (1)$$

که با تابعی به شکل رابطه (۲)، فرایند تجزیه سری زمانی را انجام دهد:

$$\Psi = 2^{-j/2} \int_1^j \Psi(2^{-j/2} - k) f(t) dt \quad (2)$$

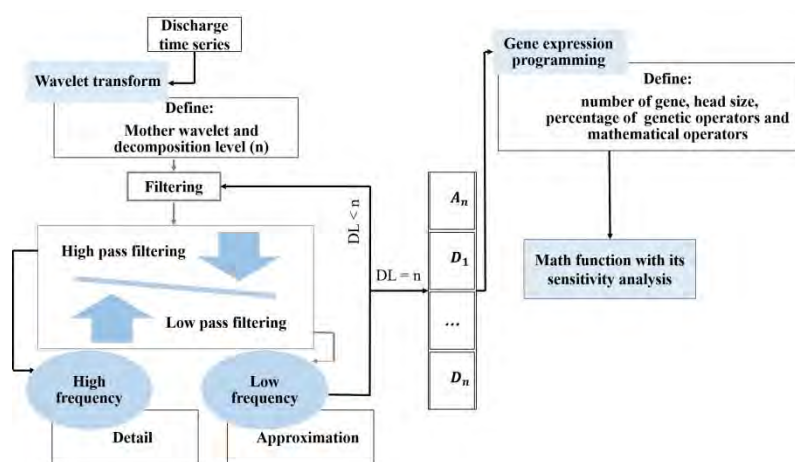


Figure 1. The overall flowchart of proposed hybrid WGEP model

سناریوهای ورودی مدل

در این مقاله، شبیه‌سازی جریان روزانه رودخانه خشک‌رود با مدل ترکیبی موجک- برنامه‌ریزی بیان ژن (WGEP) و مدل منفرد مورد توجه قرار گرفته است. بدین منظور از سری‌های زمانی ۲۷ ساله هیدرولوژیکی و هواشناسی روزانه آن (سال آبی ۷۰-۱۳۶۹ تا سال آبی ۹۵-۱۳۹۵)، استفاده شده است. فرایند مدل‌سازی برای بررسی موارد مختلفی انجام شده است، که عبارتند از:

- بررسی تأثیر پیش‌پردازش سری‌های زمانی جریان بر عملکرد مدل GEP

- بررسی نقش سطوح تجزیه مختلف در عملکرد مدل WGEP

- بررسی تأثیر داده‌های هواشناسی (دمای میانگین هوای روزانه، بارش، فشار و نقطه شبنم) بر ارتقای عملکرد مدل‌های GEP و WGEP

- بررسی قابلیت دو مدل مذکور، با استفاده از هر دو نوع ورودی، در زمان شبیه‌سازی‌های کوتاه‌مدت و بلندمدت.

لازم به ذکر است که پارامترهای هواشناسی فوق‌پس از بررسی گستره بزرگی از پارامترها، به‌عنوان مؤثرترین پارامترها بر مدل‌های پیشنهادی انتخاب شده‌اند. بر همین اساس مدل‌های متعددی با سناریوهای ارائه‌شده در جدول (۱) ساخته شده است تا مدل بهینه برای شبیه‌سازی جریان رودخانه خشک‌رود در زمان‌های مختلف ارائه گردد.

در توضیح جدول (۱) می‌توان گفت، جهت مدل‌سازی

با GEP، از دو ساختار ورودی داده (هیدرولوژیکی (H) تنها و هیدرولوژیکی و هواشناسی (H&M) به صورت توأمان) و چهار زمان پیش‌بینی یک، دو، سه و شش روز (L1، L2، L3 و L6) استفاده شده و حداکثر تأثیر داده‌های سه روز متوالی بر عملکرد مدل مورد ارزیابی قرار گرفته است. بدین ترتیب که مدل‌های G(0) داده‌های روز جاری یا زمان فعلی (t)، G(1) داده‌های امروز و یک روز قبل (t، t-1) و G(2) داده‌های امروز، یک و دو روز قبل (t، t-1، t-2) را شامل می‌شوند. در مدل WGEP نیز، مهم‌ترین مرحله در پیش‌پردازش داده توسط تبدیل موجک، انتخاب موجک مادر است. در میان موجک‌های مادر مختلفی که تاکنون شناخته شده، موجک هار^{۱۸} اولین مورد است. هار، ساختاری گسسته و ساده دارد. به دلیل اهمیت آن، به‌ویژه از نقطه نظر تاریخی و به‌عنوان یک مبنای مناسب برای ارزیابی و مقایسه، در این مقاله جهت تجزیه داده‌ها استفاده شده است. علاوه بر این، موجک دوشی^{۱۹} یکی از موجک‌های مهم است که در قالب مرتبه‌های مختلف وجود داشته و با افزایش مرتبه، طول بازه‌های پشتیبانی و تعداد نمونه‌گیری‌های آن افزایش می‌یابد. در این تحقیق به دلیل شباهت ساختار موجک دوشی مرتبه هفتم (db7) با سری زمانی مورد بررسی، از آن به‌عنوان دومین موجک مادر استفاده شده است.

با انتخاب موجک مادر، سری‌های زمانی جریان روزانه در سطوح سه تا شش تجزیه شده که نتیجه آن، تعداد {سطح تجزیه + ۱} زیرسری است.

Table 1. Various input scenarios of models

Hydrological (H) / Hydrological & Meteorological (H&M) time series							
Lead time (day)		GEP		Mother Wavelet (MW)	WGEP		Input time
Time (day)	Symbol	Time (day)	Symbol		Decomposition level	Symbol	
1	L1		G (0)	haar	3		t
2	L2	t	G (1)		4	MW ³	t, t-1
3	L3	t, t-1	G (2)	db7	5	MW ⁴	
6	L6	t, t-1, t-2			6	MW ⁵	
						MW ⁶	

است که مقداری بین ۱ و $-\infty$ دارد اما بازه‌ی [۰،۱] می‌تواند مبنای سنجش صحت و دقت مدل موردنظر قرار گرفته و مقادیر کم‌تر از صفر دلیل بر ناکارآمدی مدل است (۲۸).

$$NSE = 1 - \frac{\sum_1^n (R_i^f - R_i^o)^2}{\sum_1^n (R_i^o - R_m^o)^2} \quad (5)$$

خطای مطلق میانگین^{۳۳} (MAE) برای همه خطاها صرف‌نظر از علامت آن‌ها وزن یکسانی را در نظر می‌گیرد اما جذر متوسط مربعات خطا^{۳۴} (RMSE)، خطاها را به‌توان دو رسانده و برای خطاهای بیش‌تر وزن بیش‌تری را در نظر می‌گیرد. با این وجود هر دو شاخص، واحدی یکسان با پارامترهای قابل اندازه‌گیری دارند و مقدار کم‌تر آن‌ها بیانگر عملکرد بهتر مدل است. روابط (۶) و (۷) نحوه محاسبه‌ی این شاخص‌ها را نشان می‌دهند:

$$MAE = \frac{\sum_1^n |R_i^f - R_i^o|}{n} \quad (6)$$

$$RMSE = \sqrt{\frac{(R_i^f - R_i^o)^2}{n}} \quad (7)$$

معیار جریان حدی^{۳۵} (PFC) شاخصی است جهت ارزیابی مقادیر حدی که هم‌بعد با پارامتر مورد ارزیابی است و مقدار صفر مبنای تحلیل آن است (۲۷).

$$PFC = \frac{(\sum_1^{NP} (R_i^o - R_i^f)^2 \times R_i^{o2})^{0.25}}{(\sum_1^{NP} (R_i^o - R_i^f)^2)^{0.5}} \quad (8)$$

NP تعداد دبی‌های بیش از یک‌سوم دبی میانگین بوده و سایر پارامترها، پیش‌تر تعریف شده‌اند.

منطقه مطالعاتی

منطقه مطالعاتی این تحقیق، رودخانه خشک‌رود واقع در شهرستان رودسر از توابع استان گیلان و از رودخانه‌های حوضه دریای کاسپین است (شکل (۲)). این رودخانه، یکی از ۵۲ رودخانه اصلی و دائمی استان گیلان می‌باشد که به‌واسطه پرورش آبزیان و صنعت گردشگری و جریان در پایین‌دست منطقه گردشگری جواهردشت تا دریا، دارای اهمیت استراتژیک در استان می‌باشد.

پس از پیش‌پردازش سری دبی و آماده‌سازی آن با خصوصیات مختلف، یعنی تصویرسازی سری دبی توسط موجک مادر هار در چهار سطح تجزیه (haar^۳-haar^۶) و هم‌چنین موجک مادر دوبشی هفتم (db7^۳-db7^۶)، زیرسری‌های حاصل از هر ویژگی موجک (در مجموع هشت ویژگی حاصل می‌شود)، یک‌بار به تنهایی و بار دیگر همراه با داده‌های هواشناسی (دمای میانگین هوای روزانه، بارش، فشار و نقطه شبنم)، با تأخیرهای (t) و (t-1) (WG(0), WG(1)) به شبکه GEP اعمال شده است. در نهایت نتایج عملکرد همه مدل‌های ساخته‌شده با شاخص‌های آماری ارزیابی شده و مورد بحث و بررسی قرار گرفته‌اند.

ارزیابی مدل‌ها

در این مقاله از شاخص‌های آماری زیر جهت ارزیابی دقت و عملکرد مدل استفاده شده است.

ضریب همبستگی^{۳۰} (R) که مقداری بین -۱ و +۱ دارد. این شاخص وابستگی خطی بین دو پارامتر را بیان می‌کند و ۱ نشان‌گر بهترین ارتباط مستقیم بین دو پارامتر است (۷).

$$R = \frac{\sum_1^n (R_i^f - R_m^f)(R_i^o - R_m^o)}{\sqrt{\sum_1^n (R_i^f - R_m^f)^2 \sum_1^n (R_i^o - R_m^o)^2}} \quad (3)$$

که در آن R_m^f و R_m^o به‌ترتیب مقادیر میانگین محاسباتی و مشاهداتی، R_i^f و R_i^o به‌ترتیب، مقادیر محاسباتی و مشاهداتی اُم و n طول سری‌زمانی و یا همان تعداد مشاهدات است. شاخص توافقی^{۳۱} (I_a) یک شاخص بی‌بعد و بیانگر دقت مقادیر محاسباتی بوده و در بازه [۰،۱] متغیر است. مبنای تحلیل آن یک (۳۵) و از رابطه زیر حساب می‌شود:

$$I_a = 1 - \frac{\sum_1^n (R_i^f - R_i^o)^2}{\sum_1^n (|R_i^f - R_m^o| + |R_i^o - R_m^o|)^2} \quad (4)$$

ضریب ناش^{۳۲} (NSE)، معیاری نرمال‌شده و بدون بعد

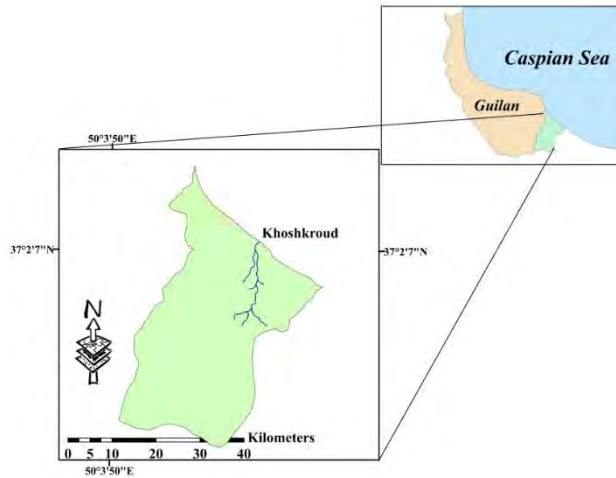


Figure 2. The location of Khoshkroud river

سطح تجزیه به زیرسری‌های مربوطه شکسته شود، که این کار با توجه به سناریوهای ارائه‌شده در جدول (۱) صورت پذیرفته است. شکل (۴)، سری زمانی دبی روزانه رودخانه خشک‌رود را با استفاده از موجک مادر دوشی مرتبه هفتم و تا پنج سطح تجزیه نشان می‌دهد، که چنانچه ملاحظه می‌شود شامل یک تقریب و پنج جزئیات خواهد بود.

در ادامه نتایج حاصل از مدل‌های منفرد و ترکیبی و برای تمامی سناریوهای ورودی در جدول (۴) آورده شده است. نتایج ارائه‌شده در این جدول نشان می‌دهد که با افزایش زمان پیش‌بینی، برای هر دو مدل ترکیبی و منفرد، قابلیت مدل کاهش یافته است، که این امر نتایج حاصل از مطالعات آداموسکی (۶) و ین و همکاران (۳۸) را تأیید می‌کند. در مدل منفرد که از داده‌های بدون پیش‌پردازش به‌عنوان ورودی مدل استفاده می‌شود، اختلاف کارایی آن به مراتب بیشتر است. چنین عملکردی به دلیل افزایش عدم قطعیت در زمان‌های طولانی‌تر، دور از انتظار نیست. در مدل منفرد، در زمان پیش‌بینی یک روز که از پیچیدگی کم‌تری برخوردار است، مقدار ضریب همبستگی ۰/۵۸ است که در زمان پیش‌بینی شش روز به ۰/۱۸ می‌رسد.

مساحت حوضه آبریز خشک‌رود حدود ۱۰۰ کیلومترمربع بوده و حداکثر ارتفاع آن ۳۶۰۸ متر می‌باشد. همچنین طول شاخه اصلی رودخانه، تقریباً ۲۵ کیلومتر است. برای مدل‌سازی توسط WGEP از داده‌های دبی روزانه ایستگاه هیدرومتری باجی‌گواهر واقع بر روی این رودخانه و داده‌های ایستگاه سینوپتیک رامسر که نزدیک‌ترین ایستگاه هواشناسی با طول داده مناسب می‌باشد، استفاده شده است. بدین منظور داده‌ها به دو بخش آموزش و ارزیابی و با نسبت ۷ به ۳ تقسیم شده‌اند. مشخصات اصلی این دو ایستگاه و سری‌های زمانی، در جدول (۲) آورده شده است.

در ادامه مشخصات آماری تمامی پارامترهای هیدرومتری و هواشناسی مورد استفاده، در جدول (۳) آورده شده است.

شکل (۳) مقادیر دبی ایستگاه هیدرومتری باجی‌گواهر را به تفکیک فازهای آموزش و ارزیابی و در برابر زمان نشان می‌دهد.

نتایج و بحث

برای استفاده از رویکرد ترکیبی WGEP لازم است تا نخست سری زمانی براساس نوع موجک مادر انتخابی و

Table 2. The characteristics of Bajiguabar hydrometric and Ramsar synoptic stations

Station	Longitude, latitude		Zone	Start	End
	X	Y			
Bajiguabar (hydrometric)	449671	4095317	39S	1369-70	1395-96
Ramsar (synoptic)	471490	4083827	39S	1369-70	1395-96

Table 3. The statistical parameters of daily time series

Parameter	Dew point (°C)	Pressure (mbar)	Precipitation (mm)	Average temperature A (°C)	Discharge (m ³ /S)
Statistical index					
Minimum	-4.1	993.60	0	-1.5	0
Maximum	26.18	1043.95	340.2	30.5	143
Average	13.12	1015.18	3.4	16.73	2.48
Standard deviation	6.67	6.92	13.55	7.19	4.07
Skewness	-0.11	0.37	9.77	0	13.57

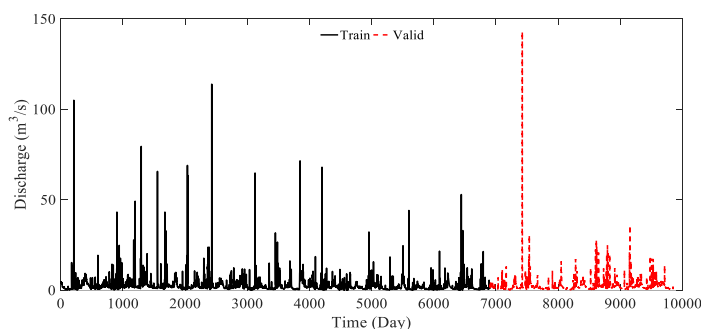


Figure 3. Training and testing phase vs. time

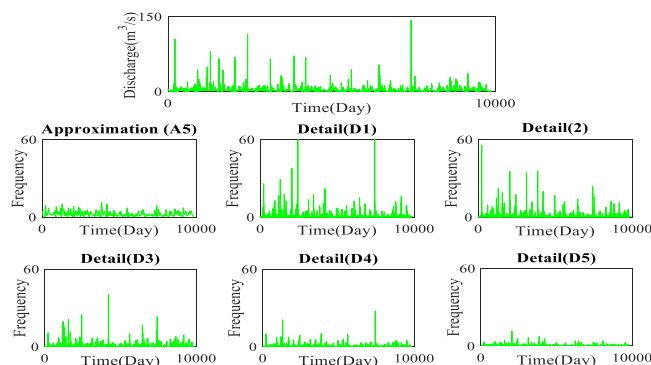


Figure 4. Wavelet decomposition of the observed Khoshkroud daily discharge time series with five levels using db7

همکاران (۱۰) و هادی و تامبول (۱۳) نیز، به ترتیب، در پیش‌بینی جریان روزانه و ماهانه، با استفاده از سناریوهای مختلف، به این نتیجه دست یافته‌اند. اما ایجاد تأخیر در داده‌ها به ارتقای عملکرد مدل کمک شایانی نکرده است. زیرا امکان وجود تأثیر مشابه تأخیرها روی شرایط مدل‌سازی وجود دارد (۱۵). هم‌چنین می‌تواند ناشی از افزایش هر چه بیشتر اطلاعات خام و خطادار به شبکه GEP و آموزش نادرست آن و یا قابلیت شبکه باشد (۸).

این مقادیر، دقت ناکافی مدل و همبستگی ضعیف مقادیر محاسباتی با مقادیر مشاهداتی را نشان می‌دهند. همان‌طورکه در شکل‌های (۵) تا (۸) ملاحظه می‌شود، داده‌های هواشناسی با افزایش زمان پیش‌بینی در بهبود عملکرد مدل تأثیر بیشتری دارند. در واقع GEP با حضور پارامترهای مؤثر در چرخه هیدرولوژی، در یافتن الگوهای پیچیده‌تر توانمندتر شده است. زیرا نقص اطلاعات ورودی خود را با مشخصات هواشناسی تا حدودی رفع کرده است. دیوپ و

Table 4. Performance and error indices of the GEP and WGEP models of one, two, three and six-day lead times

Lead time (day)	Input type	Scenario	Training results				Testing results					
			I _a	R	MAE (m ³ /S)	RMSE (m ³ /S)	PFC (m ³ /S)	I _a	R	MAE (m ³ /S)	RMSE (m ³ /S)	PFC (m ³ /S)
1	H	G (0)	0.68	0.57	0.82	3.44	1.64	0.68	0.58	0.69	3.05	2.98
		haar ⁶ (1)	0.83	0.74	0.81	3.04	0.42	0.94	0.88	0.65	1.83	0.19
		Db7 ³ (1)	0.93	0.88	0.66	1.99	0.34	0.95	0.92	0.52	1.49	0.52
	H&M	G (0)	0.69	0.58	0.82	3.45	1.64	0.69	0.57	0.69	3.08	2.94
		haar ³ (1)	0.83	0.74	0.72	3	0.43	0.91	0.87	0.56	1.92	0.63
		Db7 ⁶ (1)	0.93	0.88	0.63	1.97	0.31	0.96	0.93	0.49	1.39	0.43
2	H	G (1)	0.4	0.37	1.34	3.90	1.7	0.38	0.34	1.19	3.53	3.27
		haar ³ (1)	0.75	0.64	0.95	3.23	0.46	0.82	0.74	0.82	2.51	0.7
		Db7 ⁴ (1)	0.85	0.77	0.89	2.67	0.41	0.87	0.82	0.72	2.2	0.74
	H&M	G (0)	0.46	0.37	1.08	3.95	1.73	0.43	0.33	0.92	3.60	3.27
		haar ⁴ (1)	0.73	0.62	0.99	3.29	0.47	0.82	0.75	0.84	2.49	0.74
		Db7 ³ (1)	0.85	0.77	0.96	2.69	0.4	0.88	0.82	0.78	2.16	0.71
3	H	G (0)	0.25	0.30	1.56	4.01	1.73	0.24	0.27	1.40	3.61	3.25
		haar ³ (0)	0.66	0.54	1.09	3.55	0.48	0.72	0.64	0.92	2.89	0.84
		Db7 ⁴ (1)	0.80	0.71	1.05	2.95	0.41	0.86	0.77	0.86	2.37	0.64
	H&M	G (0)	0.42	0.28	1.31	4.17	1.73	0.37	0.22	1.13	3.89	3.26
		haar ⁶ (0)	0.63	0.53	1.16	3.56	0.5	0.85	0.78	0.98	2.37	0.54
		Db7 ⁴ (1)	0.83	0.72	1.1	2.92	0.38	0.86	0.76	0.95	2.42	0.62
6	H	G (0)	0.24	0.23	1.63	4.08	1.73	0.24	0.20	1.47	3.69	3.27
		haar ⁴ (0)	0.53	0.44	1.27	0.78	0.51	0.51	0.45	1.06	3.35	0.95
		Db7 ⁵ (1)	0.70	0.58	1.39	3.43	0.47	0.72	0.61	1.17	2.97	0.85
	H&M	G (2)	0.33	0.21	1.64	4.35	1.73	0.33	0.18	1.45	3.89	3.27
		haar ⁶ (1)	0.55	0.45	1.45	3.75	0.5	0.61	0.53	1.24	3.18	0.93
		Db7 ⁴ (1)	0.69	0.58	1.35	3.44	0.47	0.70	0.60	1.07	2.98	0.88

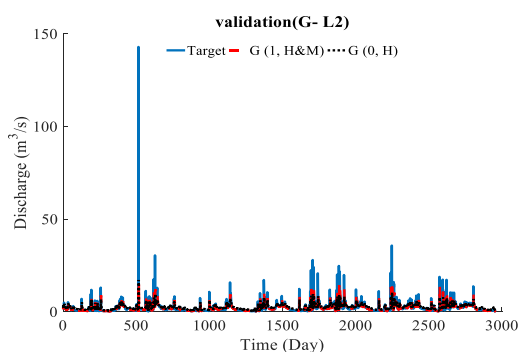


Figure 5. Observed and forecasted discharge by GEP model at one-day lead time for validating data

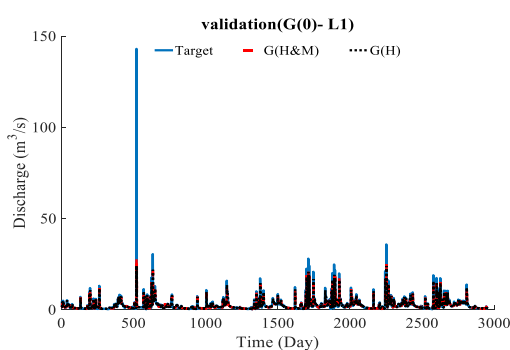


Figure 6. Observed and forecasted discharge by GEP at two-day lead time for validating data

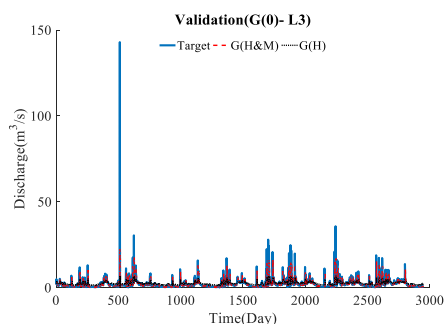


Figure 7. Observed and forecasted discharge by GEP model at three-day lead time for validating data

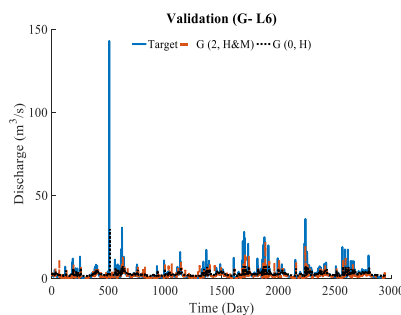


Figure 8. Observed and forecasted discharge by GEP at six-day lead time for validating data

ترکیب‌های مختلف دبی یا فاکتورهای هواشناسی تنها یک ترکیب نسبت به حالت دبی تنها بهبود قابل‌ملاحظه‌ای داشته است.

پیش‌پردازش داده‌ها سبب کاهش خطا و عدم قطعیت‌ها می‌شود (۲۳). همین امر موجب می‌شود که ایجاد تأخیر و افزودن داده‌های مناسب به فرایند مدل‌سازی کمک کند. علاوه بر این همان‌طور که در شکل‌های (۹) تا (۱۲) مشخص است، با افزایش زمان پیش‌بینی، تأثیر داده‌های هواشناسی برای موجک db7 کم‌تر می‌شود (عکس GEP و haar-GEP) و تنها حضور آن‌ها در شبیه‌سازی برخی مقادیر بیشینه مهم خواهد بود. به‌طوری‌که PFC برای زمان پیش‌بینی یک روز از ۰/۵۲ به ۰/۴۳ کاهش یافته است (۱۷ درصد بهبود یافته است). چنین نتایجی حاکی از درستی انتخاب دوشی مرتبه هفتم است که توانسته فرکانس‌های کارآمد را استخراج کند، به‌طوری‌که تنها در شبیه‌سازی مقادیر زیاد به سایر اطلاعات محیطی احتیاج پیدا کند. بنابراین در مدل db7-GEP به دلیل آن‌که مناسب‌ترین داده‌ها به شبکه داده‌کاو اعمال می‌شود از اطلاعات کمکی (بارش، دمای میانگین هوا، نقطه شبنم و فشار هوا) در تشخیص ساختارهای جدیدتر استفاده می‌کند و با کمک آن‌ها ضعف شبکه در آموزش‌پذیری و یا جامعیت داده‌های آموزشی را جبران می‌کند. از این‌رو، طبیعی است که در اثر کاهش قابلیت مدل با افزایش زمان پیش‌بینی، توانایی آن در استفاده از این داده‌ها در شبیه‌سازی بهتر، کم‌تر شود. درحالی‌که در مدل‌های GEP و haar-GEP وجود خطا و نقص داده سبب شده که ارتباط بین پارامترها به درستی مشخص نشود و با افزایش عدم قطعیت در زمان‌های پیش‌بینی بلندتر، میزان درستی کاهش می‌یابد. به همین دلیل، داده‌های هواشناسی به‌عنوان اطلاعات

با توجه به نتایج حاصل از مدل‌های ترکیبی WGEP که در جدول (۴) مشاهده می‌شود، مدل با موجک مادر db7 به مراتب عملکرد بهتری نسبت به مدل با موجک مادر هار دارد. از طرف دیگر db7 در مقایسه با GEP، به‌گونه‌ای برتری دارد که نتایج زمان ششم آن حتی از نتایج زمان پیش‌بینی یک روز GEP هم بهتر هستند. همان‌طور که ملاحظه می‌شود، دوشی مرتبه هفتم از کارایی بالاتری نسبت به هار برخوردار بوده و شباهت بیشتری با سری زمانی دبی دارد، پس به‌طور قطع باید نتایج بهتری را ارائه دهد. اما بیش‌ترین اختلاف عملکرد آن‌ها مربوط به زمان پیش‌بینی یک روز با داده H&M است که RMSE دوشی هفتم حدود ۲۸ درصد نسبت به هار کاهش یافته است. کم‌ترین اختلاف RMSE مربوط به زمان پیش‌بینی سه روز، با داده H&M است که هار دو درصد نسبت به دوشی مرتبه هفتم بهتر عمل کرده است. بر این اساس، مشاهده می‌شود که در زمان‌های پیش‌بینی طولانی‌تر، داده‌های هواشناسی در ارتقا عملکرد مدل haar-GEP نسبت به db7-GEP بسیار مؤثر بوده و اختلاف RMSE آن‌ها را کم‌تر می‌کنند. این رویکرد، به یکی دیگر از مزیت‌های داده‌های هواشناسی اشاره دارد که توانسته‌اند عدم تطابق موجک مادر با ساختار سری را جبران کنند. هم‌چنین ایجاد تأخیر در داده‌ها و شبیه‌سازی داده‌های دو روز متوالی، در اکثر سطوح و زمان‌ها، سبب بهبود نتایج شده است. با بررسی کلیه نتایج، مشخص می‌شود که بالای ۸۰ درصد عملکرد مدل‌های با داده‌های هیدرومتری، به دلیل اضافه‌شدن داده‌های هواشناسی ارتقا یافته‌اند. هرچند در بیش‌تر مدل‌ها، میزان بهبود جزئی است، اما در برخی سطوح به میزانی است که قابلیت آن‌ها را به‌طور قابل‌توجهی ارتقا داده است. هادی و تامبول (۱۳) نیز نشان داده‌اند که از میان

پیش‌بینی جریان با استفاده از رویکرد تلفیقی موجک- برنامه‌ریزی بیان ژن و ارزیابی تأثیر پارامترهای هواشناسی بر کارایی آن

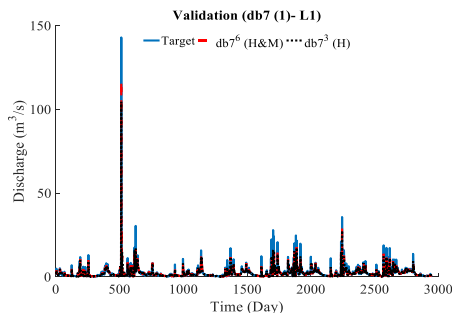


Figure 9. Observed and forecasted discharge by WGEP at one-day lead time for validating data

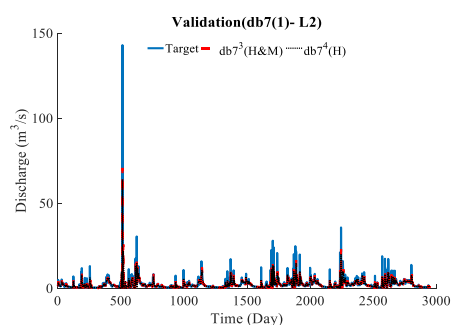


Figure 10. Observed and forecasted discharge by WGEP at two-day lead time for validating data

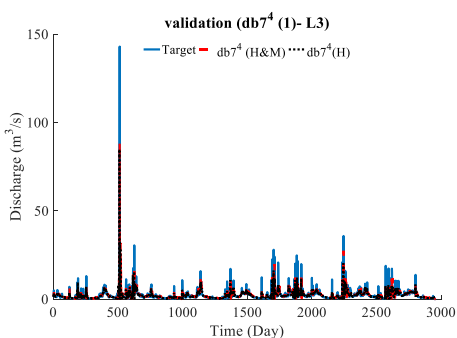


Figure 11. Observed and forecasted discharge by WGEP at three-day lead time for validating data

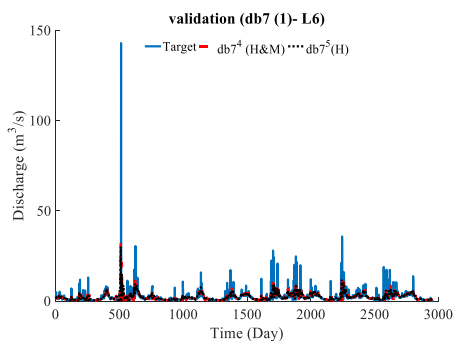


Figure 12. Observed and forecasted discharge by GEP at six-day lead time for validating data

اضافه، در زمان‌های پیش‌بینی بالاتر نقش بیشتری دارند، در واقع به نسبت عدم قطعیت‌های حاکم در مدل ایفای نقش می‌کنند. به‌طوری‌که، در haar-GEP برای زمان پیش‌بینی سه روز، با افزودن داده‌های هواشناسی مقدار RMSE، ۱۸ درصد بهبود یافته است. در مدل GEP نیز برای همین زمان پیش‌بینی، علی‌رغم این‌که نتایج (R, RMSE, MAE) با داده H بهتر است اما طبق شکل (۷)، مدل با داده H&M مشابه‌تر با مقادیر مشاهداتی عمل کرده است. همچنین، شاخص I_a که مقدار درستی و دقت مدل را بررسی می‌کند، نشان می‌دهد که مدل G(H&M)، ۵۴ درصد نسبت به G(H) بهبود یافته است. صرف‌نظر از شاخص‌های آماری مورد استفاده، ملاحظه شکل‌های (۹) تا (۱۲) به‌خوبی تأثیر پارامترهای هواشناسی منتخب در ارتقای عملکرد مدل، به‌خصوص در مقادیر حدی (به‌طور مشخص مقادیر سیلاب‌های سالانه) را نشان می‌دهد، که این امر با توجه به اهمیت بالای مقادیر حدی در مدیریت و برنامه‌ریزی منابع آب و این نکته که حتی ارتقای جزئی در پیش‌بینی این مقادیر حائز اهمیت است، می‌تواند بسیار تأثیرگذار باشد.

شکل‌های (۱۳) تا (۱۶)، مقادیر شاخص‌های مختلف ارزیابی و خطا را در برابر سطوح تجزیه و برای مدل‌های ترکیبی مختلف با و بدون پارامترهای هواشناسی نشان می‌دهد. همان‌گونه که در این شکل‌ها مشخص است، در محدوده بهینه سطوح تجزیه (سطح ۳ تا ۶)، تأثیرات سطح تجزیه، بر کارایی مدل چندان قابل‌ملاحظه نیست، ولی با این وجود شاخص‌های مختلف ارزیابی در زمان‌های پیش‌بینی کم‌تر، حساسیت بیشتری به سطوح تجزیه دارند، درحالی‌که با افزایش زمان پیش‌بینی حساسیت مدل‌های ترکیبی به میزان سطح تجزیه کم‌تر می‌شود (دقت کم‌تر، حساسیت به تغییرات جانبی کم‌تر).

کارایی مدل‌های ترکیبی به دلیل کاهش خطا و آماده‌سازی داده ورودی از مدل‌های منفرد بیش‌تر است، به‌گونه‌ای‌که به‌طور مثال میزان جذر متوسط مربعات خطا برای مدل منفرد در زمان پیش‌بینی یک روز بیش از دو برابر آن در مدل ترکیبی با استفاده از دوشمی مرتبه هفتم به‌دست آمده است (RMSE(G-L1)= 2.22RMSE(db7- GEP- L1)). هم‌چنین اختلاف کارایی مدل‌های ترکیبی در مقایسه با مدل‌های منفرد همواره با افزایش زمان پیش‌بینی با شیب تندی افزایش می‌یابد، به‌طوری‌که مدل‌های منفرد در زمان‌های پیش‌بینی بالا، فاقد کارایی هستند، اما مدل‌های ترکیبی دارای قابلیت اعتماد بالایی می‌باشند.

در مدل GEP، تأثیر افزودن تأخیر به ورودی‌ها، برای زمان پیش‌بینی‌های مختلف، تقریباً به یک شکل است و با افزایش تعداد تأخیر، کارایی مدل کم‌تر می‌شود، زیرا افزایش تعداد پارامترهای خطا دار اعمال شده به آن، موجب ارائه مدل با خطای بیش‌تری نیز می‌شود، در نقطه مقابل، مدل WGEP، به دلیل حذف داده‌های اضافی، با افزایش اطلاعات ورودی به نتایج بهتری دست یافته است.

پیش‌پردازش داده‌ها سبب بهبود عملکرد مدل می‌شود، اما در صورتی‌که موجک مادر به درستی انتخاب نشود، مدل آن‌طور که باید ارتقا نمی‌یابد. زیرا عدم شباهت و نزدیکی ویژگی‌های تابع پایه به سری زمانی منجر به از بین رفتن برخی خصوصیات آن می‌شود. از طرفی پهنای باند نمونه‌برداری موجک هرچه عریض‌تر باشد، ویژگی‌های بیش‌تری را می‌تواند از سری استخراج کند و مجموعه کامل‌تری را برای آموزش شبکه GEP فراهم نماید. علاوه بر موجک مادر، انتخاب سطح تجزیه‌ای که کارآمدترین ویژگی‌ها را نمایش دهد، حائز اهمیت است و افزایش سطح تجزیه و اعمال جزئیات بیش‌تر به شبکه GEP، الزاماً سبب بهبود نتایج نمی‌شود.

برای هر دو مدل ترکیبی و منفرد، با افزایش زمان

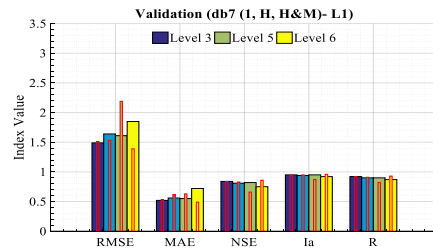


Figure 13. Performance and error indices of WGEP model for various decomposition levels at one-day lead time (db7)

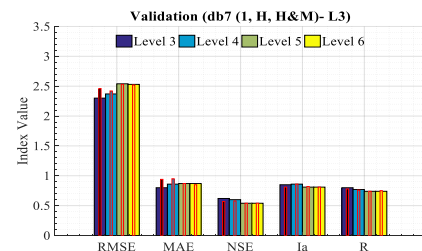


Figure 14. Performance and error indices of WGEP model for various decomposition levels at three-day lead time (db7)

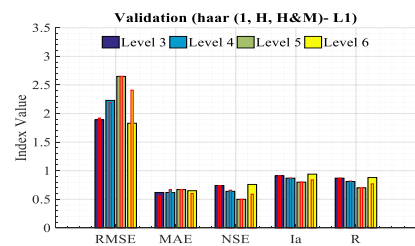


Figure 15. Performance and error indices of WGEP model for various decomposition levels at one-day lead time (haar)

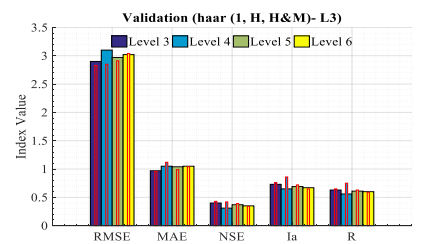


Figure 16. Performance and error indices of WGEP model for various decomposition levels at three-day lead time (haar)

مهم‌ترین نتایج حاصل از اجرای بیش از یک‌صد مدل برای سناریوهای مختلف این تحقیق را می‌توان به شرح زیر خلاصه کرد.

میان رویکردهای متفاوت مورد توجه، بیش‌ترین تأثیر را فرایند پیش‌پردازش به‌دلیل استخراج فرکانس‌های مؤثر بر نتایج داشته است، به‌گونه‌ای که پیش‌پردازش داده‌ها با استفاده از تبدیل موجک توانسته به مقدار قابل‌ملاحظه‌ای کارایی و قابلیت مدل را افزایش دهد، این افزایش کارایی، به‌ویژه در زمان‌های پیش‌بینی بالاتر به‌وضوح مشاهده می‌شود. زیرا با کاهش خطاها توانسته تا حدودی از عدم‌قطعیت‌های آن‌ها بکاهد. از طرفی به‌رغم این‌که ورود داده‌های هواشناسی می‌تواند عملکرد مدل را بهبود بخشد، ولی افزایش قابلیت و کارایی مدل بدین سبب به‌میزان جزئی می‌باشد. چون‌که به‌عنوان اطلاعات اضافه کمی از نقص داده می‌کاهد. با این وجود تأثیر ورود پارامترهای هواشناسی در کارایی مدل به‌منظور پیش‌بینی مقادیر حدی بیشینه و کمینه، قابل توجه است.

پی‌نوشت‌ها

1. Ensemble Empirical Mode Decomposition
2. Variational Mode Decomposition
3. Kurukavak
4. Gene Expression Programming
5. Artificial Neural Network
6. Adaptive Neuro-Fuzzy Inference System
7. Wavelet-Genetic Expression Programming
8. Multi Gene Genetic Programming
9. Pareto-Optimal Moving Average- Multi Gene Genetic Programming
10. Continuous Wavelet Transform
11. Wavelet Linear Genetic Programming
12. Wavelet Artificial Neural Network
13. Multiple Linear Regression
14. Wavelet Transform
15. Scale
16. Transmission
17. Mother Wavelet
18. haar
19. Daubechies
20. Correlation Coefficient
21. Index of Agreement
22. Nash-Sutcliffe Model Efficiency Coefficient
23. Mean Absolute Error
24. Root Mean Square Error
25. Peak Flow Criteria

پیش‌بینی، از دقت عملکرد مدل کاسته می‌شود و این می‌تواند ناشی از اطلاعات کم و محدودی باشد که برای زمان‌های بلندتر به مدل اعمال می‌شود که یکی از راه‌کارهای ارتقای آن‌ها، استفاده از خروجی سایر زمان‌های پیشین به‌عنوان ورودی زمان‌های آتی است، لیکن میزان این کاهش کارایی مدل در رویکرد منفرد به مراتب بیش‌تر است.

مطابق با شکل‌های (۵) تا (۱۲)، داده‌های هواشناسی در بهبود عملکرد مدل GEP مؤثرند، اما اثرگذاری آن‌ها کم‌تر از پیش‌پردازش داده است. این‌گونه‌که، در زمان پیش‌بینی یک روز، شاخص PFC برای $G(H\&M)$ ، $db7(H)$ و $db7(H\&M)$ به‌ترتیب ۱۳۴ درصد، ۸۴ درصد و ۸۶ درصد نسبت به $G(H)$ کاهش یافته است. این اختلاف می‌تواند ناشی از تمایز نقش دو عامل در مدل‌سازی باشد. داده‌های هواشناسی اطلاعات جانبی غیرثابتی هستند که در تغییر حالت آن مؤثرند، اما پیش‌پردازش سبب اصلاح اطلاعات اصلی و اولیه شبکه می‌شود.

مدل‌های ترکیبی به‌رغم کارایی بالایی که دارند، در پیش‌بینی مقادیر حدی جریان رودخانه دارای ضعف می‌باشند، لیکن این ضعف با افزودن پارامترهای هواشناسی (دمای میانگین هوای روزانه، بارش، فشار و نقطه شبنم) براساس شاخص PFC از سه تا ۱۷ درصد بهبود یافته است.

نتیجه‌گیری

در این مقاله پیش‌بینی جریان روزانه رودخانه خشک‌رود واقع در شرق استان گیلان و در زمان‌های پیش‌بینی یک، دو، سه و شش روز، بررسی گردید. به‌این منظور از روش برنامه‌ریزی بیان ژن و رویکرد ترکیبی موجک- برنامه‌ریزی بیان ژن استفاده شده است. در هر دو رویکرد از داده‌های هیدرومتری و هواشناسی استفاده شده و تأثیر پارامترهای هواشناسی و هم‌چنین فرایند پیش‌پردازش بر دقت و کارایی مدل ارزیابی شده است. نتایج این مطالعه نشان می‌دهد که از

منابع

10. Diop, L., Bodian, A., Djaman, K., Yaseen, Z. M., Deo, R. C., El-Shafie, A., & Brown, L. C. (2018). The influence of climatic inputs on stream-flow pattern forecasting: case study of Upper Senegal River. *Environmental earth sciences*, 77(5), 182.
11. Ferreira, C. (2001). Gene expression programming: a new adaptive algorithm for solving problems. arxiv preprint cs/0102027.
12. Freire, P. K. D. M. M., Santos, C. A. G., & da Silva, G. B. L. (2019). Analysis of the use of discrete wavelet transforms coupled with ANN for short-term streamflow forecasting. *Applied Soft Computing*, 80, 494-505.
13. Hadi, S. J., & Tombul, M. (2018). Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination. *Journal of Hydrology*, 561, 674-687.
14. Khairuddin, N., Aris, A. Z., Elshafie, A., Sheikhy Narany, T., Ishak, M. Y., & Isa, N. M. (2019). Efficient forecasting model technique for river stream flow in tropical environment. *Urban Water Journal*, 16(3), 183-192.
15. Kim, K. J., Kim, Y. O., & Kang, T. H. (2017). Application of time-lagged ensemble approach with auto-regressive processors to reduce uncertainties in peak discharge and timing. *Journal of Hydrology: Regional Studies*, 9, 140-148.
16. Kisi, O., Shiri, J., & Tombul, M. (2013). Modeling rainfall-runoff process using soft computing techniques. *Computers & Geosciences*, 51, 108-117.
17. Li, F. F., Wang, Z. Y., & Qiu, J. (2019). Long-term streamflow forecasting using artificial neural network based on preprocessing technique. *Journal of Forecasting*, 38(3), 192-206.
18. Maheswaran, R., & Khosa, R. (2012). Comparative study of different wavelets for hydrologic forecasting. *Computers & Geosciences*, 46, 284-295.
19. Mehdizadeh, S., Fathian, F., & Adamowski, J. F. (2019). Hybrid artificial intelligence-time series models for monthly streamflow modeling. *Applied Soft Computing*, 80, 873-887.
20. Mehr, A. D. (2018). An improved gene expression programming model for streamflow forecasting in intermittent streams. *Journal of hydrology*, 563, 669-678.
21. Mehr, A. D., & Nourani, V. (2017). A Pareto-optimal moving average-multigene genetic programming model for rainfall-runoff modelling. *Environmental modelling & software*, 92, 239-251.
1. آبابایی، ب.، سهرابی، ت.، و میرزایی اصلی، ف. (۱۳۹۲). شبیه‌سازی جریان روزانه ورودی به سد طالقان با استفاده از مدل‌های همراشتین- واینر. مدیریت آب و آبیاری. ۳ (۱): ۱۲-۱.
۲. چمنی، م.، و روشنگر، ک. (۱۳۹۸). ارزیابی مدل تلفیقی تجزیه مد تجربی یکپارچه کامل- گاوسی در پیش‌بینی زمانی و مکانی دبی رودخانه. مدیریت آب و آبیاری. ۹ (۲): ۲۷۷-۲۸۹.
۳. شریف‌آذری، س.، و عراقی‌نژاد، ش. (۱۳۹۲). توسعه مدل ناپارامتری شبیه‌ساز داده‌های ماهانه هیدرولوژیکی. مدیریت آب و آبیاری. ۳ (۱): ۸۱-۹۵.
۴. نبی‌زاده، م.، مساعدی، ا.، و دهقانی، ا.ا. (۱۳۹۱). تخمین هوشمند دبی روزانه با بهره‌گیری از سامانه استنباط فازی- عصبی تطبیقی. مدیریت آب و آبیاری. ۲ (۱): ۸۳-۹۵.
5. Abdollahi, S., Raeisi, J., Khalilianpour, M., Ahmadi, F., & Kisi, O. (2017). Daily mean streamflow prediction in perennial and non-perennial rivers using four data driven techniques. *Water Resources Management*, 31(15), 4855-4874.
6. Adamowski, J. F. (2008). River flow forecasting using wavelet and cross-wavelet transform models. *Hydrological Processes: An International Journal*, 22(25), 4877-4891.
7. Anctil, F., & Ramos, MH. (2019). Verification metrics for hydrological ensemble forecasts. *Handbook of Hydrometeorological Ensemble Forecasting*; Springer: Berlin/Heidelberg, Germany, 893-922.
8. Banihabib, M. E., & Mousavi-Mirkalaei, P. (2019). Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region. *Sustainable Cities and Society*, 48, 101585.
9. Boggess, A., Narcowich, FJ., Donoho, DL., & Donoho, PL. (2002). A first course in wavelets with Fourier analysis. *Physics Today*, 55(5), 63.

22. Mehr, A. D., Nourani, V., Kahya, E., Hrnjica, B., Sattar, A. M., & Yaseen, Z. M. (2018). Genetic programming in water resources engineering: a state-of-the-art review. *Journal of hydrology*, 566, 643-667.
23. Nourani, V., Baghanam, A. H., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet-artificial intelligence models in hydrology: a review. *Journal of Hydrology*, 514, 358-377.
24. Nourani, V., Komasi, M., & Alami, MT. (2012). Hybrid wavelet-genetic programming approach to optimize ANN modeling of rainfall-runoff process. *Journal of Hydrologic Engineering*, 17(6), 724-41.
25. Phukoetphim, P., Shamseldin, A. Y., & Adams, K. (2012). Multimodel Approach Using Neural Networks and Symbolic Regression to Combine the Estimated Discharges of Rainfall-Runoff Models. *Journal of Hydraulic Engineering*, 17 (9), 975-985.
26. Ravansalar, M., Rajaei, T., & Kisi, O. (2017). Wavelet-linear genetic programming: a new approach for modeling monthly streamflow. *Journal of Hydrology*, 549, 461-475.
27. Rezaie-Balf, M., Kim, S., Fallah, H., & Alaghmand, S. (2019). Daily river flow forecasting using ensemble empirical mode decomposition based heuristic regression models: Application on the perennial rivers in Iran and South Korea. *Journal of Hydrology*, 572, 470-485.
28. Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, 480, 33-45.
29. Sang, Y. F. (2013). A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric research*, 122, 8-15.
30. Shahabi, S., Khanjani, M. J., & Kermani, M. R. H. (2017). Significant wave height modelling using a hybrid Wavelet-genetic Programming approach. *KSCE Journal of Civil Engineering*, 21(1), 1-10.
31. Shiri, J., & Kisi O. (2010). Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model. *Journal of Hydrology*, 394(3-4), 486-93.
32. Shoaib, M., Shamseldin, AY., Melville, BW., & Khan, MM. (2015). Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach. *Journal of Hydrology*, 527, 326-44.
33. Tikhamarine, Y., Souag-Gamane, D., Ahmed, A. N., Kisi, O., & El-Shafie, A. (2020). Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. *Journal of Hydrology*, 582, 124435.
34. Wang, WC., Chau, KW., Xu DM., & Chen, XY. (2015). Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resources Management*, 29(8), 2655-75.
35. Willmott, CJ. (1981). On the validation of models. *Physical geography*, 2(2), 184-94.
36. Xie, T., Zhang, G., Hou, J., Xie, J., Lv, M., & Liu, F. (2019). Hybrid Forecasting Model for Non-Stationary Daily Runoff Series: A Case Study in the Han River Basin, China. *Journal of Hydrology*, 123915.
37. Yaseen, ZM., El-Shafie, A., Jaafar, O., Afan, HA., & Sayl, KN. (2015). Artificial intelligence based models for stream-flow forecasting: 2000-2015. *Journal of Hydrology*, 530, 829-44.
38. Yin, Z., Feng, Q., Wen, X., Deo, R. C., Yang, L., Si, J., & He, Z. (2018). Design and evaluation of SVR, MARS and M5Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. *Stochastic Environmental Research and Risk Assessment*, 32(9), 2457-2476.
39. Zhang, X., Tuo, W., & Song, C. (2019). Application of MEEMD-ARIMA combining model for annual runoff prediction in the Lower Yellow River. *Journal of Water and Climate Change*.