

مقاله پژوهشی

کاربرد الگوریتم پیش‌پردازش‌کننده تجزیه مد تجربی گروهی برای بهبود دقت پیش‌بینی جریان روزانه رودخانه با استفاده از روش‌های هوشمند مصنوعی

رضا مستوری^{۱*}، محسن نجارچی^۲ و سید علیرضا ترابی^۳

چکیده

با توجه به اهمیت پیش‌بینی جریان رودخانه در مدیریت منابع آب، از روش‌های مختلفی برای مدل‌سازی کردن جریان رودخانه‌ها استفاده می‌شود تا با به‌کارگیری مدلی مطمئن در مدیریت خشکسالی و سیلاب، خسارت‌های ناشی از آنها به حداقل برسد. در این مطالعه نیز برای پیش‌بینی سری زمانی جریان روزانه ایستگاه گچسر برای حوضه آبریز بخش مرکزی ایران با توجه به ویژگی‌های غیرخطی مقیاس‌های زمانی چندگانه، مدل ترکیبی درخت تصمیم‌گیری (MT) برپایه الگوریتم پیش‌پردازش‌کننده (EEMD) در دوره آماری ۱۳۶۳ تا ۱۳۹۱ پیشنهاد شد. در ادامه، برای تعیین بهترین تعداد تأخیر زمانی از دبی جریان رودخانه، تکنیک‌های ACF و PACF به کار رفت که مقدار بهینه سه تأخیر زمانی به‌عنوان متغیرهای ورودی در نظر گرفته شد. نتایج به‌دست‌آمده از روش ترکیبی پیشنهادی با دیگر روش‌های هوشمند مصنوعی مانند مدل درخت (MT) و سیستم استنتاج عصبی-فازی (ANFIS) با استفاده از شاخص‌های آماری خطا بررسی و مقایسه شد. نتایج نشان داد بر دقت روش EEMD-MT نسبت به روش MT، حدود ۷/۵ درصد افزوده شد و خطای مدل‌سازی نیز برای این ایستگاه در ترم شاخص خطای RMSE، ۶/۸ درصد در مرحله آزمایش کاهش یافت؛ بنابراین، با پیش‌پردازش کردن داده‌های هواشناسی و برطرف کردن توزیع تصادفی، ناپایداری و روند غیرخطی داده‌های ورودی به مدل، دقت پیش‌بینی مدل‌های پیشنهادی در تخمین دبی روزانه جریان رودخانه افزایش یافته و برای مطالعات آتی، روش مناسب و دقیقی است.

واژه‌های کلیدی: پیش‌بینی جریان رودخانه، روش‌های هوشمند مصنوعی، الگوریتم تجزیه مد تجربی گروهی، ایستگاه‌های هیدرومتری

ارجاع: مستوری ر.، نجارچی م. و ترابی س. ع.، ۱۴۰۱. کاربرد الگوریتم پیش‌پردازش‌کننده تجزیه مد تجربی گروهی برای بهبود دقت پیش‌بینی جریان روزانه رودخانه با استفاده از روش‌های هوشمند مصنوعی. مجله پژوهش آب ایران. ۴۵-۱: ۱۴-۱. <https://dx.doi.org/10.22034/iwrj.2022.10048.2353>

۱- استادیار گروه مهندسی و مدیریت منابع آب، دانشگاه آزاد اسلامی واحد اراک.

۲- استادیار گروه مهندسی و مدیریت منابع آب، دانشگاه آزاد اسلامی واحد اراک.

۳- دانشجوی دکتری گروه مهندسی و مدیریت منابع آب، دانشگاه آزاد اسلامی واحد اراک.

* نویسنده مسئول: r-mastouri@iau-arak.ac.ir

تاریخ پذیرش: ۱۴۰۰/۰۸/۱۳

تاریخ دریافت: ۱۴۰۰/۰۵/۱۰

مقدمه

اهمیت و جایگاه آب در دوام و بقای حیات بر کسی پوشیده نیست؛ از این رو، پژوهشگران و متخصصان، همواره به مسئله آب توجه کرده اند. کشور ایران با برخورداری از موقعیت جغرافیایی و اقلیم خاص، سهم اندکی از بارش های آسمانی به عنوان منبع اصلی آب دارد؛ به گونه ای که با داشتن میانگین متوسط بارندگی سالانه ۲۵۰ میلی متر، یعنی حدود یک سوم مقدار متوسط بارندگی جهان، جزء اقلیم های خشک و نیمه خشک به شمار می آید که آب و کنترل و مدیریت آن، اهمیت زیادی دارد (طاهری و همکاران، ۱۳۸۹). با توجه به این مسئله و اهمیت و حساسیت موضوع مهار آب های سطحی، به ویژه در کشور ما- که بیشتر رودخانه های مناطق مختلف، فصلی است- و در کنار کمبود منابع آبی که در پهنه وسیعی از کشور وجود دارد، نیاز به شناسایی و مدل سازی رفتار رودها و جریان های آبی برای برنامه ریزی های بلندمدت و استفاده بیشتر و بهتر از توانایی های آنها بسیار احساس می شود. تازه تأسیس بودن بیشتر ایستگاه های هیدرومتری، نقص های موجود در آمار بیشتر این ایستگاه ها، قرارگرفتن بیشتر رودها در مناطق خشک، وضعیت بحرانی برداشت آب های زیرزمینی و لزوم توجه بیشتر به آب های سطحی، علل بیشتر و دقیق تری است که به مقوله پیش بینی در حوضه های آبریز کشورمان جلوه مهم تری می دهد. پیش بینی مؤلفه های مؤثر در تحلیل سامانه های منابع آب، از جمله جریان رودخانه، یکی از موارد اساسی در طراحی بهره برداری و مطالعات مربوط به این منابع به شمار می آید (مرشدلو و عبدالله زاده، ۱۳۸۶). روابط بارش باران و رواناب، یکی از پیچیده ترین پدیده های هیدرولوژیکی است که به علت متغیرهای مختلف فیزیکی و هیدرولوژیکی، شامل تغییرات فزاینده و زمان بندی فوق العاده ویژگی های آبخیز و الگوهای بارندگی و تعداد متغیرهای دخیل در مدل سازی فرایندهای فیزیکی درک می شود؛ بنابراین، درک و پیش بینی فرایندهای تولید رواناب و انتقال آن به نقطه خروجی حوضه آبریز، یکی از اساسی ترین مباحث علم هیدرولوژی است (جلالی، ۱۳۸۸؛ نبی زاده و همکاران، ۲۰۱۲). همانگونه که پیش تر گفته شد، در دهه های اخیر، پیش بینی و مدل سازی جریان رودخانه برای برنامه ریزی و مدیریت منابع آب، اهمیت زیادی دارد. خاصیت غیرخطی بودن فرایند بارش - رواناب و

پیچیده بودن مدل های فیزیکی باعث شده است پژوهشگران به مدل های هوشمند مانند شبکه های عصبی مصنوعی^۱ (ANN)، سیستم تطبیقی استنتاج عصبی فازی^۲ (ANFIS)، ماشین بردار پشتیبان^۳ (SVM)، الگوریتم مدل درخت (MT)^۴، برنامه نویسی بیان ژن^۵ (GEP)، رگرسیون چندجمله ای تکاملی (EPR)^۶ و روش دسته بندی گروهی داده ها^۷ (GMDH) روی بیاورند. علاوه بر این، از مزایای روش های هوشمند این است که از آنها برای شبیه سازی مسائل غیرخطی و پیچیده می توان استفاده کرد. در بحث بارش - رواناب، گفتنی است روابط حاصل از روش های هوش مصنوعی از فرایند هیدرولوژیکی مبتنی بر شاخصه فیزیکی در حوضه آبریز ارزیابی نمی شود؛ بلکه از تجزیه و تحلیل مقادیر ورودی و خروجی به دست می آید که به طور سری زمانی است؛ به عنوان مثال در ایران، کارآموز و همکاران (۱۳۸۳)، کاربرد شبکه عصبی مصنوعی را در مدل سازی هیدرواقلمی پیش بینی جریان ورودی به مخزن بررسی کردند. برای پیش بینی بلندمدت جریان رودخانه، دو رویکرد مختلف پیشنهاد شده است. در رویکرد اول از مدلی برای ارتباط متغیرهای هیدرواقلمی و جریان رودخانه استفاده شده است؛ به گونه ای که از متغیرهای بارش، دمای هوا، رواناب و بودجه برفی مشاهده شده در زمان حال به عنوان ورودی های مدل استفاده می شود. در رویکرد دوم از ترکیب دو مدل پیش بینی هواشناسی برای پیش بینی متغیرهای بارش و دمای هوا و یک مدل هیدرولوژیکی و از شبکه های انتشار خطا به عقب و آدلاین تطبیقی برای مدل سازی برای پیش بینی جریان ماهانه رودخانه زاینده رود در محل ورودی به سد زاینده رود استفاده شد.

نتایج به دست آمده از دو رویکرد مذکور، نشان دهنده برتری عملکرد شبکه عصبی مصنوعی نسبت به رویکرد اول است. رضایی و همکاران (۱۳۸۶)، دبی اوج حوضه آبریز سفیدرود (ناحیه غیرخزری) را با استفاده از شبکه عصبی مصنوعی، مدل سازی کردند. این پژوهش بر جمع آوری و انتخاب

1- Artificial Neural Networks

2- Adaptive Neuro-Fuzzy Inference System

3- Support Vector Machine

4- Model Tree

5- Gene Expression Programming

6- Evolutionary Polynomial Regression

7- Group Method of Data Handling

پیشنهادی کمک کند. نتایج نشان داد روش W-ANN، شاخص آماری خطای ریشه مربعات خطا را به ترتیب، ۷۴ و ۴۳ درصد برای پیش‌بینی جریان روزانه آب رودخانه کاهش داد و بر مقدار ضریب همبستگی برای هر یک از ایستگاه‌ها به میزان ۴۷ و ۱۱ درصد افزود.

وَنگ و همکاران (۲۰۱۵) براساس اهمیت پیش‌بینی رواناب روزانه در امر مدیریت منابع آب، یک روش هوشمند مصنوعی ترکیبی مبتنی بر تجزیه گروهای داده‌های پیش‌بینی کردند. در این پژوهش، روش ANN با روش EEMD^۲ ترکیب شد تا پیش‌بینی مناسب متوسط و بلندمدتی از رواناب داشته باشد. آنها برای ارزیابی روش پیشنهادی خود از دو مخزن ذخیره آب در کشور چین بهره بردند و عملکرد روش پیشنهادی خود را با شاخص‌های آماری خطا مانند ضریب همبستگی و ریشه میانگین مربعات خطا بررسی کردند. نتایج، افزایش دقت و کاهش خطا در روش پیشنهادی را نسبت به روش ANN نشان داد.

چَنگ و همکاران (۲۰۱۵) برای پیش‌بینی رواناب روزانه مخزن در استان Southeast Guizhou واقع در کشور چین از تکنیک ANN-QPSO استفاده کردند که به‌عنوان یک روش ترکیبی - بهینه‌سازی شناخته می‌شود. نتایج شاخص‌های آماری خطا نشان داد مقادیر R و NSE برای روش ANN-QPSO نسبت به روش ANN در مرحله آموزش تا مقدار ۳/۹۷ و ۱۳/۴۱ بهبود یافت و برای مرحله آزمایش تا ۵/۴۲ و ۱۷/۴۶ افزایش داشت. پراساد و همکاران (۲۰۱۷) با استفاده از رویکردی نوین، دبی جریان ماهانه رودخانه را بررسی و پیش‌بینی کردند که متغیری حیاتی برای برنامه‌ریزی منابع آب و تخصیص منابع آب برای امورات کشاورزی است. آنها روشی ترکیبی مبتنی بر نظریه موجک و شبکه عصبی مصنوعی برای هدف پژوهش خود در نظر گرفتند. در ادامه، تکنیک IIS^۳ برای انتخاب بهترین ویژگی‌ها به‌عنوان متغیرهای ورودی انتخاب و با روش W-ANN ترکیب شد؛

سپس، مدل ترکیبی پیشنهادی با روش مدل درخت مقایسه شد. آنها برای ارزیابی روش ترکیبی خود، سه ایستگاه هیدرولوژی از مناطق کشاورزی کشور استرالیا را انتخاب کردند. رضایی - بالف و کیسی (۲۰۱۸) با استفاده

هیدروگراف‌هایی با منشأ بارندگی در ۱۲ زیرحوضه با زمان تمرکز برابر و یا کمتر از ۲۴ ساعت مبتنی بوده است. نتایج نشان داد مدل شبکه عصبی، کارایی زیادی در پیش‌بینی دبی حداکثر خروجی از حوضه آبریز سفیدرود دارد. فتحیان و هرمزی‌نژاد (۱۳۹۰)، متغیرهای کمی و کیفی جریان در رودخانه کارون را با استفاده از شبکه‌های عصبی مصنوعی پیش‌بینی کردند. در این پژوهش، با استفاده از مدل شبکه عصبی مصنوعی پرسپترون چندلایه، شبکه پیش‌رونده، شبکه تابع پایه شعاعی، سه روز آینده مقادیر کمی و کیفی جریان در رودخانه کارون، بازه ملاثانی - فارسیات، شامل دبی جریان، اشل، کل املاح محلول در آب و هدایت الکتریکی پیش‌بینی شد. برای تعیین داده‌های ورودی مؤثر در مدل‌های شبکه عصبی مصنوعی و تعداد نرون‌ها در لایه میانی هر یک از مدل‌ها از قابلیت بهینه‌سازی الگوریتم ژنتیک استفاده شد. نتایج نشان داد دقت مدل‌های پیشنهادی در برآورد و پیش‌بینی متغیرهای کمی و کیفی جریان در رودخانه کارون، به‌ترتیب برابر با ۹۴/۵ درصد، ۸۰/۵ درصد و ۸۶ درصد است. برزگر و همکاران (۱۳۹۸) با استفاده از مدل‌های هوشمند مصنوعی مرکب نظارت‌شده، نوسان‌های ماهانه سطح آب دریاچه ارومیه را پیش‌بینی کردند. در این پژوهش، روش‌های هوشمند مصنوعی برای پیش‌بینی سطح آب یک، دو و سه ماه آینده دریاچه ارومیه به کار رفت و درنهایت، یک مدل هوش مصنوعی مرکب نظارت‌شده برای رسیدن به عملکردی بهتر از مدل‌های منفرد به‌کاررفته استفاده شد. نتایج معیارهای عملکرد مدل هوش مصنوعی مرکب نظارت‌شده نشان داد مدل‌سازی یک ماه آینده سطح آب با مقادیر $RMSE$ ، R^2 و MAE به‌ترتیب، برابر با ۰/۹۸۹۶، ۰/۵۴۷ و ۰/۰۴۲۱ متر در مقایسه با مدل بردار پشتیبان، عملکرد بهتری داشته است که این عملکرد برای پیش‌بینی‌های دو و سه ماه آینده سطح آب دریاچه نیز صادق است.

در خارج از ایران نیز کیسی (۲۰۰۹)، روشی ترکیبی برای پیش‌بینی روزانه رواناب واقع در منطقه Thrace در کشور ترکیه ارائه داد. او برای افزایش مدل شبکه عصبی مصنوعی، داده‌های هیدرولوژیکی را با استفاده از نظریه موجک (WT^۱) تجزیه کرد تا روندهای ناپایدار و نویزی از داده‌های هیدرولوژیکی خارج شود و به افزایش دقت مدل

2- Ensemble empirical mode decomposition

3- Iterative Input Selection

1- Wavelet transform

از تکنیک‌های نوین داده- مبنا مانند EPR و ¹OP-ELM، جریان روزانه آب رودخانه تجن را با استفاده از سه ایستگاه هیدرومتریک واقع در ایران پیش‌بینی کردند. همچنین، آنها روش MLPNN را برای مقایسه روش‌های پیشنهادی خود در نظر گرفتند. نتایج نشان داد روش EPR، عملکرد بهتری داشته است و مقادیر پیک دبی جریان را نسبت به دو روش OP_ELM و MLPNN بهتر شبیه‌سازی کرده است. از دیگر نتایج پژوهش آنها این بود که برای حوضه تجن- ایران، رابطه‌ای برای تخمین میزان دبی جریان ارائه دادند.

در پژوهش حاضر، پس از جمع‌آوری اطلاعات از سازمان‌های مرتبط، ابتدا برای تشخیص میزان اثرگذاری مقادیر پیشین هر یک از متغیرهای هیدرولوژیکی (روزهای گذشته)، که به‌عنوان متغیرهای ورودی و خروجی بر پیش‌بینی مدل‌های داده- مبنا معرفی می‌شود، از تکنیک‌های ²ACF و ³PACF استفاده می‌شود. در این روش‌ها، تأثیر گام‌های زمانی پیشین متغیرها ارزیابی می‌شود. در گام بعدی، برای حذف و اصلاح ترند و نویز، که ممکن است در متغیرهای ورودی مدل‌های پیشنهادی وجود داشته باشد، روش‌های پیش‌پردازش داده‌ها مانند نظریه موجک⁴ و روش EEMD برای پردازش داده‌ها به کار می‌رود. پس از مدل‌سازی، نتایج حاصل از مقادیر پیش‌بینی مدل‌های پیشنهادی مانند ANFIS و MT، بررسی و میزان اثرگذاری هر یک از روش‌های پیش‌پردازش‌کننده بر عملکرد مدل با استفاده از شاخص‌های آماری خطا شناسایی می‌شود.

مواد و روش‌ها

منطقه مورد مطالعه

سد امیر کبیر، که سد کرج نیز نامیده می‌شود، در شمال شهر کرج و در ۲۵ کیلومتری جاده کرج- چالوس قرار دارد. این سد، اولین سد چندمنظوره کشور ایران و یکی از منابع تأمین آب شهر تهران است (عزیزی، ۱۳۹۵). سد بتنی امیر کبیر با گنجایشی بالغ بر ۲۰۵ میلیون متر مکعب، ارتفاع ۱۸۰ متر از پی، عمق ۱۷ متر از پی، متوسط آب سالانه ۵۰۰ متر مکعب بر ثانیه و نیز

دریاچه‌ای به طول ۱۴ کیلومتر و حجم مفید ۱۹۵ میلیون متر مکعب، اولین سد چندمنظوره ایران است که در شمال شهر کرج و در ۲۵ کیلومتری محور کرج- چالوس واقع شده است. دریاچه این سد، که از رودخانه کرج تغذیه می‌شود، یکی از مهم‌ترین زیستگاه‌های طبیعی ماهی قزل‌آلای خال قرمز و تنها زیستگاه ماهی کورگونوس است. از دیگر اهداف این سد، از تأمین آب لوله‌کشی شهر تهران و گسترش کشاورزی و آبرسانی به ۵۰ هزار هکتار زمین‌های کشاورزی کرج می‌توان نام برد. نیروگاه برق آبی این سد نیز با بیشتر از ۴۶ سال فعالیت به شبکه برق کشور متصل است و توانایی تولید ۹۰ مگاوات برق را دارد؛ هرچند این مقدار تولید برق با توجه به حوضچه آبریز و دبی خروجی‌اش، بسیار کم دانسته شده است و از نقص در مهندسی سد به‌عنوان عامل این مشکل نام برده می‌شود. انباشته‌شدن آب رودخانه کرج در پشت سد امیر کبیر، علاوه بر کارایی آن در تولید انرژی برقی و تأمین بخشی از آب آشامیدنی شهر تهران، باعث به وجود آمدن دریاچه زیبایی شده است که موجب استفاده‌های گردشگری، تفریحی و ورزشی از سد کرج شده است و یکی از بهترین استراحتگاه‌ها و جاذبه‌های طبیعی کرج را به وجود آورده است. طول این دریاچه در زمان تکمیل ظرفیت آب پشت سد به مسافتی بالغ بر ۵۵۰۰ متر می‌رسد. این موضوع نشان می‌دهد در صورت لحاظ کردن مسائل زیست‌محیطی و تکیه بر مصوبات شورای عالی امنیت ملی، به این دریاچه و روستای واریان به‌عنوان دهکده توریستی ساحلی مشرف به آن همچون فرصتی در بخش اکوتوریست می‌توان نگاه کرد (قاسمی زیارانی، ۱۳۸۴). شکل ۱، نمایی از سد کرج و موقعیت ایستگاه‌های سینوپتیک را در این منطقه نشان می‌دهد.

رودخانه کرج از کوه‌های جنوبی البرز سرچشمه می‌گیرد. این رودخانه در حد فاصل بیلقان تا دیزین واقع شده است و با طول تقریبی ۷۵ کیلومتر از ارتفاعات دیزین از کوهی به نام کلونستک در بلندی‌های خرسنگ سرچشمه می‌گیرد و پس از طی مسیری از رودخانه ولایت‌رود به رود کرج می‌پیوندد. این رودخانه در تمام طول مسیر خود از شیب تند دره‌ها و صخره‌ها عبور می‌کند و پس از گذرکردن از حاشیه شرقی شهر کرج و سیراب‌کردن دشت‌های کرج و شهریار به جاجرود می‌پیوندد. آب رودخانه کرج پس از پیمودن ۲۰ کیلومتر رو به غرب به

1- Optimally Pruned Extreme Learning Machine

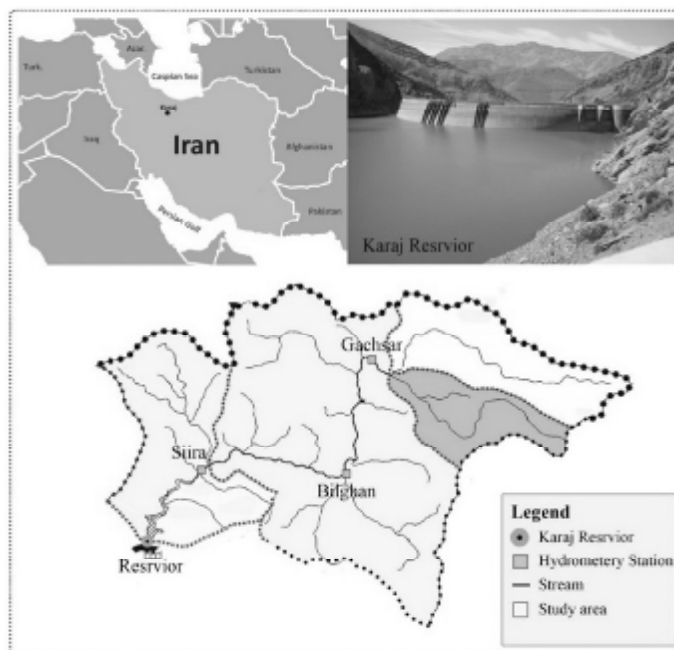
2- Auto-Correlation Function

3- Partial Auto-Correlation Function

4- Wavelet Theory

بهتری ارائه می‌دهد، در این مطالعه، ایستگاه‌هایی انتخاب شده است که این ویژگی را داشته باشد (قاسمی، ۱۳۸۴). همچنین، جدول ۱، نشان‌دهنده برخی از ایستگاه‌ها در این منطقه است که در این پژوهش از ایستگاه گچسر برای توسعه مدل‌های پیشنهادی استفاده می‌شود.

وارنکه‌رود می‌رسد و به آب دیزین می‌پیوندد. این آب در گچسر با آب علفچره‌های سمانوزک و کندوان در گچسر یکی می‌شود و به سمت جنوب جریان پیدا می‌کند. منطقه مورد مطالعه از مناطق کوهپایه‌ای و کوهستانی ایران است که در دامنه رشته‌کوه‌های البرز واقع شده است. از آنجا که در بررسی تغییرات اقلیمی، آمارهای درازمدت، نتایج



شکل ۱- موقعیت سد و رودخانه کرج

جدول ۱- مشخصات ایستگاه‌های هیدرومتری و سینوپتیک

ایستگاه	سال‌های آماری	طول جغرافیایی	عرض جغرافیایی	ارتفاع (متر)
گته ده	۱۳۹۰-۱۳۴۵	E ۵۱° ۸'	E ۳۶° ۱۰'	۲۳۰۸
بیلقان	۱۳۹۰-۱۳۴۵	E ۵۱° ۴'	E ۳۶° ۱۰'	۲۱۲۸
سیرا	۱۳۹۰-۱۳۴۶	E ۵۱° ۹'	E ۳۶° ۱'	۲۰۵۴
گچسر	۱۳۹۰-۱۳۵۹	E ۵۱° ۱۹'	E ۳۶° ۶'	۲۲۵۲
نسا	۱۳۹۰-۱۳۴۷	E ۵۱° ۱۹'	E ۳۶° ۴'	۱۸۹۸
سد امیرکبیر	۱۳۹۰-۱۳۴۷	E ۵۱° ۵'	E ۳۵° ۵۷'	۱۶۹۷

ذاتی (IMF^1) بسط می‌دهد. روش تجزیه مد تجربی را نوردن ای- هوانگ و همکاران در سال ۱۹۹۸ در زمان فعالیت هوانگ در ناسا ابداع کردند. ابداع روش EEMD، نتیجه مطالعه تبدیل هیلبرت و بهبود یافته روش EMD نوردن ای- هوانگ و همکاران در سال ۱۹۹۸ است (وو و هوانگ، ۲۰۰۹). نتایج تجزیه EEMD، داده‌های لازم را برای تکمیل تبدیل هیلبرت فراهم می‌کند. در موارد

الگوریتم تجزیه مد تجربی (EEMD)

تجزیه مد تجربی (EEMD) یک روش تجزیه اطلاعات برای استخراج سیگنال‌ها در اطلاعات تولیدشده توسط عملیات غیرخطی و غیرایستاست که شامل جداسازی محلی و کامل برگرفته از اطلاعات از یک سیگنال در نوسان‌های سریع و آهسته است. EEMD یک سری زمانی را به تعداد محدودی از حالت‌های نوسانی به نام توابع مد

1- Intrinsic Mode Function

زیادی، EEMD با تبدیل هیلبرت، ترکیب و به نام تبدیل هیلبرت- هوانگ (HHT) شناخته می شود. پس از مطالعه مفهوم فرکانس آئی در تبدیل هیلبرت، هوانگ، مفهوم تازهای از تابع مد ذاتی (IMF) را ارائه کرد که مؤلفه اساسی هر مجموعه داده پیچیده است؛ سپس آنها روش تازهای به نام تجزیه مد تجربی ایجاد کردند تا یک سیگنال را به مجموعه ای از توابع مد ذاتی تجزیه کند (رضایی بلف و همکاران، ۲۰۱۹). برخلاف روش های تجزیه سنتی، مانند تبدیل فوری و موجک، این روش جدید، داده - وفقی است. مزیت اصلی EEMD بر روش های دیگر، امکان تجزیه خودکار به صورت کاملاً داده - وفقی است و به یک مجموعه توابع مفروض، همانند تبدیل موجک نیاز ندارد. الگوریتم EEMD می کوشد تقریباً هر سیگنالی را به مجموعه ای متناهی از توابع تجزیه کند و تبدیل هیلبرت، مقادیر فرکانس لحظه ای فیزیکی را به دست می آورد. این توابع، توابع مد ذاتی نامیده می شود. الگوریتم از یک روند غربال تکرارشونده استفاده می کند که با موفقیت، میانگین محلی را از یک سیگنال کم می کند. این روند غربال گری در ادامه بیان می شود (تانگ و همکاران، ۲۰۲۰):

1. اکستریم محلی سیگنال (ماکزیمم، مینیمم) تعیین می شود.
2. مقدار ماکزیمم به یک تابع درون یابی متصل می شود و پوش بالایی سیگنال را می سازد.
3. مقدار مینیمم به یک تابع درون یابی متصل می شود و پوش پایینی سیگنال را می سازد.
4. میانگین محلی به عنوان نصف اختلاف بین پوش بالایی و پایینی محاسبه می شود.
5. میانگین محلی از سیگنال کم می شود.
6. این روند بر باقیمانده تکرار می شود.

روند غربال گری تا زمانی تکرار می شود که سیگنال به تعریف یک IMF برسد؛ سپس IMF، از سیگنال اصلی، کسر و روند غربال بر باقیمانده تکرار می شود. این کار تا زمانی ادامه می یابد که باقیمانده، یک تابع یکنواخت شود. آخرین IMF استخراج شده، کمترین مؤلفه فرکانسی است. با اعمال EEMD، یک سیگنال به مجموعه ای از توابع تک مؤلف با نام توابع مد ذاتی (IMFs) تجزیه می شود. یک تابع تک مؤلفه، تابعی نوسانی نزدیک به رایج ترین و ابتدایی ترین تابع هارمونیک را نشان می دهد (علی و همکاران، ۲۰۲۰).

ایده اصلی در پس زمینه IMF، جداسازی اطلاعات به بخش میانگین محلی با تغییرات آرام و بخش نوسانی متقارن با تغییرات سریع است که بخش آخر، همان IMF است و میانگین محلی با عنوان باقیمانده تعریف می شود. این باقیمانده به عنوان ورودی برای تجزیه بیشتر تا جایی ادامه می یابد که هیچ نوسان دیگری وجود نداشته باشد. به عنوان حالتی نوسانی، دامنه و فرکانس یک IMF با زمان به طریقی تغییر می کند که به طور محلی، متقارن و تک مؤلفه باشد؛ بنابراین، فرکانس لحظه ای به صورتی معنادار محاسبه می شود (وو و هوانگ، ۲۰۰۹)؛ از این رو، IMF ها فرکانس هایی در محدوده بالاترین تا پایین ترین سیگنال را در بر می گیرد که به عنوان سیگنال های مدوله شده دامنه و فرکانس (AM-FM) موجود است که AM پوش را حمل می کند و FM تغییر دامنه ثابت در فرکانس است و با استفاده از روشی غربالی محاسبه می شود. برای انجام دادن آن، یک IMF باید دو شرط را ارضا کند (هوانگ و همکاران، ۱۹۹۸):

(۱) تعداد اکستریم (Ne ماکزیمم و مینیمم محلی) و تعداد عبور از صفرها Nz باید برابر باشد یا حداکثر یک اختلاف داشته باشد.

$$(N_z - 1) \leq N_e \leq (N_z + 1) \quad (1)$$

(۲) در هر نقطه، مقدار میانگین پوشی که با ماکزیمم محلی و پوشی که با مینیمم محلی به وجود می آید، صفر است.

مدل درخت تصمیم (MT)

الگوریتم M5 رایج ترین طبقه بندی استفاده شده در خانواده مدل تصمیم گیری درختی است که کوینلان در سال ۱۹۹۲ برای پیش بینی داده های پیوسته ارائه کرد. مدل درختی M5 یک الگوریتم پیش بینی عددی است و گره های درخت با خواص بیشینه خطاهای مورد انتظاری برگزیده می شود که تابعی از انحراف استاندارد متغیرهای خروجی است. مدل M5، مدلی درختی برای پیش بینی صفات عددی پیوسته است که در آن، توابع رگرسیونی خطی در برگ های این درخت ظاهر می شود (کوینلان، ۱۹۹۲). این مدل، برخلاف مدل های درخت تصمیم معمول، که کلاس یا رده های گسسته را به عنوان خروجی ارائه می کند، یک مدل خطی چندمتغیره را برای داده ها در هر گره از مدل درختی می سازد. تشکیل ساختار مدل های درخت تصمیم گیری، شامل مراحل ایجاد درخت و هرس

پنج مرحله دارد: فازسازی متغیرهای ورودی، کاربرد عملگرهای فازی در بخش مقدم، استنباط از مقدم به تالی، جمع نتایج تالی‌ها و فازیزدایی (نادری و همکاران، ۲۰۱۳).

سیستم استنتاج فازی- عصبی تطبیقی را جانگ در سال ۱۹۹۳ ساخت و به‌عنوان تقریب‌زننده‌ای جامع به کار رفت. اگر سیستم استنتاج فازی برای مسئله‌ای به کار گرفته شود که مجموعه‌ای از داده‌های ورودی و خروجی دارد، الزاماً نیاز نیست ساختار فازی از پیش تعیین‌شده‌ای برای حل مسئله به کار رود و با بررسی اجمالی اطلاعات، توابع عضویت مناسب را نمی‌توان تشخیص و به ورودی‌ها اختصاص داد. متغیرهای مربوط به توابع عضویت باید به‌گونه‌ای انتخاب شود که بیشترین هماهنگی را با داده‌های ورودی- خروجی داشته باشد. برای این کار از تکنیک سازگاری عصبی می‌توان استفاده کرد که با دستورهای موجود در جعبه ابزار فازی کار می‌کند. روش آموزش سازگاری عصبی، مشابه شبکه‌های عصبی کار می‌کند. درحقیقت، سیستم استنتاج فازی- عصبی تطبیقی، سیستم‌های فازی است که برای بهبود متغیرهای خود از روش آموزش تطبیقی استفاده می‌کند یا به عبارت دیگر، از شبکه‌های عصبی برای آموزش مدل بهره می‌گیرد (نادری و یوسفزاده، ۲۰۱۷). برای آموزش مدل در انفیس از دو روش انتشار بازگشتی خطا و روش ترکیبی (ترکیب انتشار بازگشتی خطا با کمترین مربعات) استفاده می‌شود که به‌تازگی در جعبه ابزار متلب گنجانده شده است. در شکل ۲ ساختار انفیس نشان داده شده است.

همانگونه که در شکل ۲ مشاهده می‌شود، انفیس از شبکه عصبی ۵ لایه‌ای تشکیل شده است که لایه اول شامل ورودی‌ها، لایه دوم شامل ترم‌های کلامی ورودی‌ها یا به عبارتی، توابع عضویت، لایه سوم نشان‌دهنده بخش مقدم قوانین فازی، لایه چهارم شامل بخش تالی قوانین فازی و لایه پنجم نشان‌دهنده مرحله فازیزدایی است. یکی از تفاوت‌های شبکه‌های عصبی معمولی با انفیس این است که در شبکه‌های عصبی، معمولاً گره‌ها، توابع یکسان دارد و اتصال کاملی بین تمامی گره‌ها در لایه‌های مجاور برقرار است.

کردن آن است (ویتن و فرانک، ۲۰۰۵) که در مرحله تشکیل درخت، از الگوریتمی استنتاجی یا معیار تقسیم (انشعاب) برای تولید یک درخت تصمیم استفاده می‌شود. معیار تقسیم برای الگوریتم مدل M5، انحراف معیار مقادیر هر کلاسی است که به‌عنوان کمیتی از خطا به یک گره می‌رسد و کاهش مورد انتظار در این خطا را به‌عنوان نتیجه آزمون هر صفت در آن گره محاسبه می‌کند. کاهش معیار (SDR^۱) از معادله (۲) محاسبه می‌شود (کوبیلان، ۱۹۹۲):

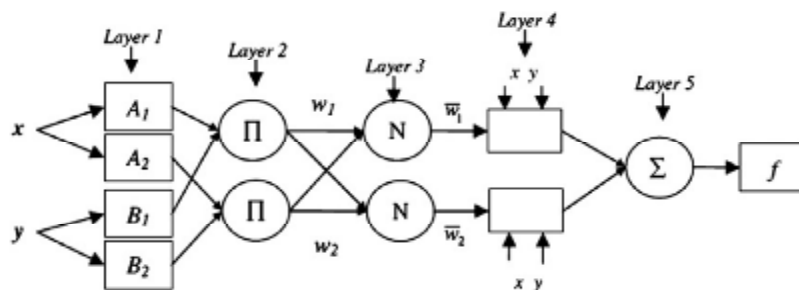
$$SDR = sd(T) - \sum_i \frac{T_i}{|T|} \times sd(T_i) \quad (2)$$

$$sd(T) = \sqrt{\frac{1}{N} \left(\sum_{i=1}^N y_i^2 - \frac{1}{N} \left(\sum_{i=1}^N y_i \right)^2 \right)} \quad (3)$$

که در آن، T نشان‌دهنده مجموعه‌ای از نمونه‌هاست که به گره می‌رسد، Sd نشان‌دهنده انحراف معیار، y_i مقدار عددی ویژگی هدف نمونه i ، N داده‌ها و T_i نشان‌دهنده زیرمجموعه‌ای از نمونه‌هاست که i امین نتیجه تست پتانسیلی را دارد. به‌علت فرایند انشعاب، داده‌های قرارگرفته در گره‌های فرزند، انحراف معیار کمتری نسبت به گره مادر دارد و بنابراین، خالص‌تر است. پس از به حداکثر رساندن تمامی انشعاب‌های ممکن، M5 صفتی را انتخاب می‌کند که کاهش مورد انتظار را بیشینه کند. این تقسیم، بیشتر ساختار شبه‌درختی بزرگی را تشکیل می‌دهد که باعث بیش‌برازش می‌شود. برای غلبه بر مسئله بیش‌برازش، درخت تشکیل‌شده باید هرس شود. این کار با جایگزینی درختی فرعی با یک برگ انجام می‌شود؛ بنابراین، مرحله دوم در طراحی مدل درختی، شامل هرس کردن درخت رشدیافته و جایگزینی درختان فرعی با توابع رگرسیونی خطی است (سولوماتین و دولا، ۲۰۰۳).

سیستم استنتاج فازی- عصبی تطبیقی (ANFIS)

در این قسمت، اصول منطق فازی، نظریه مجموعه‌های فازی و سیستم‌های فازی ممدانی و سوگنو توضیح داده شده است. نظریه منطق فازی را لطفی‌زاده در سال ۱۹۶۵ مطرح کرد (لطفی‌زاده، ۱۹۷۸). سیستم استنتاج فازی از جمع مواردی ساخته می‌شود که تاکنون بیان شد. دو نوع سیستم استنتاج فازی به نام‌های ممدانی و سوگنو وجود دارد. این دو نوع سیستم در روش مشخص کردن خروجی‌ها با یکدیگر تفاوت دارد. فرایند استنتاج فازی،



شکل ۲- ساختار سیستم استنتاج فازی - عصبی تطبیقی

باشد، یعنی تغییرات دو متغیر به‌طور هم‌جهت اتفاق می‌افتد؛ پس با افزایش در هر متغیر، متغیر دیگر نیز افزایش می‌یابد و برعکس، اگر مقدار R منفی شد، یعنی دو متغیر در جهت عکس یکدیگر عمل می‌کند؛ پس با افزایش مقدار یک متغیر، مقادیر متغیر دیگر کاهش می‌یابد و برعکس، اگر مقدار به‌دست‌آمده، صفر شد، نشان می‌دهد هیچ رابطه‌ای بین دو متغیر وجود ندارد و اگر +۱ شد، همبستگی مثبت کامل و اگر -۱ شد، همبستگی کامل و منفی است. همچنین، RMSE برای اندازه‌گیری دقت پیش‌بینی مدل استفاده می‌شود که با جذرگرفتن از خطاها، یک مقدار مثبت را نشان می‌دهد. محدوده RMSE از صفر تا یک عدد مثبت بزرگ ادامه می‌یابد که واحد آن از جنس متغیر خروجی مشاهداتی است. مقادیر بالا در متغیر R و کم در RMSE نشان‌دهنده عملکرد مطلوب مدل است.

جدول ۲- روابط شاخص‌های ارزیابی خطا برای پیش‌بینی دبی جریان رودخانه

$R = \frac{\sum_{i=1}^M (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^M (O_i - \bar{O})^2 \sum_{i=1}^M (P_i - \bar{P})^2}}$ $RMSE = \sqrt{\frac{\sum_{i=1}^M (P_i - O_i)^2}{M}}$ $RSE = \frac{\sum_{i=1}^M (P_i - O_i)^2}{\sum_{i=1}^M (\bar{O} - O_i)^2}$ $MAE = \frac{\sum_{i=1}^M P_i - O_i }{M}$ $RAE = \frac{\sum_{i=1}^M P_i - O_i }{\sum_{i=1}^M O_i - \bar{O} }$
--

همچنین، MAE مقدار میانگین مثبت خطاها در یک مجموعه از مقادیر پیش‌بینی‌شده را محاسبه می‌کند و واحد آن از جنس متغیر خروجی مشاهداتی است. مقدار

مدل انفیس، بسیار پیچیده‌تر از سیستم‌های استنتاج فازی است که تا به حال از آنها صحبت شد و برای تمام حالت‌های سیستم استنتاج فازی در دسترس نیست؛ به‌ویژه که فقط سیستم‌های سوگنو را پوشش می‌دهد. به‌طور کلی، محدودیت‌های انفیس به‌صورت ذیل است (صدرممتازی و همکاران، ۲۰۱۳):

- باید از نوع سیستم سوگنو مرتبه صفر یا یک باشد؛
- قوانین، یک خروجی داشته باشد و آن نیز از روش فازی‌زدایی میانگین‌گیری وزن‌ها به دست آمده باشد؛
- تمامی توابع عضویت خروجی باید یکسان باشد و همگی خطی یا ثابت باشد و
- قوانین مختلف، توابع عضویت خروجی یکسانی نداشته باشد؛ به عبارت دیگر، تعداد توابع عضویت خروجی با تعداد قوانین برابر باشد.

معیارهای ارزیابی خطا برای مقایسه عملکرد مدل‌ها

برای مقایسه عملکرد مدل‌های ارائه‌شده و روابط تجربی در مراحل آموزش و آزمایش، شاخص‌های آماری شامل ضریب همبستگی (R)، ریشه میانگین مربعات خطا^۱ (RMSE)، خطای استاندارد نسبی^۲ (RSE)، میانگین خطای مطلق^۳ (MAE) و ریشه مطلق خطا^۴ (RAE) در جدول ۲ نشان داده شده است.

O_i مقدارهای مشاهده‌ای، P_i مقدارهای پیش‌بینی، $\bar{O}$ میانگین داده‌های مشاهده‌ای، $\bar{P}$ میانگین داده‌های پیش‌بینی و M تعداد داده‌هاست. مقدار اندازه‌گیری‌شده R، میزان همبستگی بین دو متغیر وابسته را نشان می‌دهد که مقدار آن بین +۱ و -۱ است. اگر مقدار به‌دست‌آمده مثبت

1- Root Mean Square Error

2- Relative Standard Error

3- Mean Absolute Error

4- Root Absolute Error

که قضاوت اولیه‌ای برای تشخیص وابستگی دبی جریان به چند روز پیش است. همانگونه که گفته شد، نمودار PACF فقط دید اولیه‌ای را درباره تعداد روزهای هم‌بسته ارائه می‌دهد. شکل ۴، نمودار PACF را برای ایستگاه گچسر به نشان می‌گذارد. بر مبنای این شکل، به نظر می‌رسد سری زمانی تا سه روز قبل برای متغیر دبی جریان وابستگی وجود دارد. گفتنی است، ستون اول مربوط به روز اول است و به صورت مشخص، این ستون در نظر گرفته نمی‌شود.

پس از اینکه متغیرهای ورودی به مدل با استفاده از تابع PACF برای هر دو ایستگاه مد نظر انتخاب شد، هر یک از مجموعه داده‌ها برای ایستگاه گچسر به دو بخش آموزش مدل و صحت‌سنجی مدل تقسیم می‌شود. بنا بر مشاهدات علمی، که براساس تکنیک سعی و خطا به دست آمده است، بهترین تقسیم داده‌ها براساس درصد، اینگونه است که ۷۰-۸۰ درصد داده‌ها برای آموزش مدل و بقیه داده‌ها برای مرحله صحت‌سنجی انتخاب می‌شود. در این پژوهش، داده‌ها برای مرحله آموزش، ۷۵ درصد داده‌ها را تشکیل داده و ۲۵ درصد باقیمانده برای مرحله آزمایش یا صحت‌سنجی در نظر گرفته شد. معادله (۴)، نشان‌دهنده متغیرهای ورودی با تأخیر زمانی به دست آمده از دبی جریان رودخانه‌های مد نظر در این پژوهش، یعنی گچسر و متغیر خروجی مدل است. در این معادلات، t نشان‌دهنده زمان، $t-1$ مقدار دبی جریان یک روز قبل تر و Qg دبی جریان برای ایستگاه گچسر است.

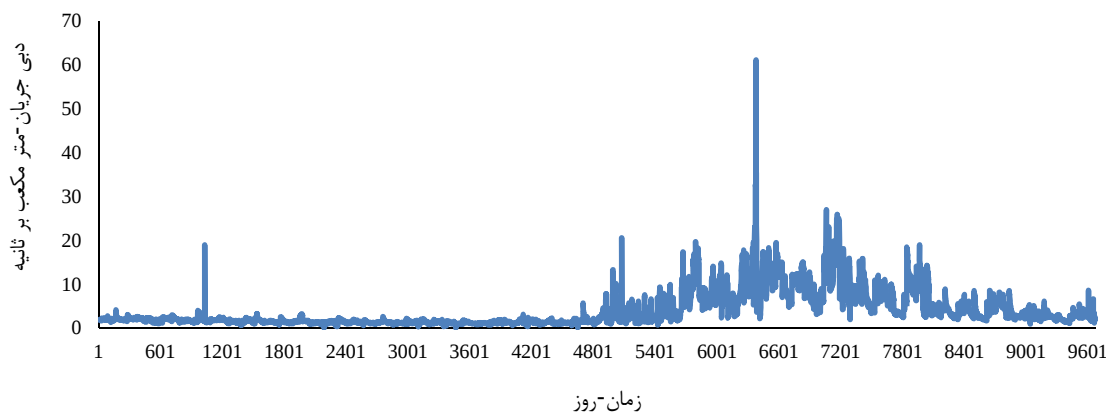
$$Qg(t) = f(Qg(t-1), Qg(t-2), Qg(t-3), Qg(t-4)) \quad (4)$$

این شاخص آماری، معمولاً مشابه مقدار RMSE است. شاخص خطای آماری RSE، مدل‌هایی را مقایسه می‌کند که مقادیر خطای اندازه‌گیری شده آنها واحدهای متفاوتی دارد. همچنین، مفهوم شاخص خطای RAE به مقدار متغیر آماری RSE شباهت بسیاری دارد.

نتایج و بحث

در این بخش، نتایج استخراج شده از مدل‌های پیشنهادی مستقل ANFIS، MT و مدل ترکیبی EEMD-MT تجزیه و تحلیل می‌شود؛ از این رو، دبی جریان به صورت روزانه برای ایستگاه گچسر در شکل ۳ نشان داده شده است. برای تجزیه و تحلیل بیشتر مقادیر دبی جریان، برخی از متغیرهای آماری برای کل مجموعه داده‌های دو ایستگاه مانند انحراف معیار، میانگین، ماکزیمم، مینیمم، مختصات جغرافیایی ایستگاه‌ها و ارتفاع از سطح دریا در جدول ۳ آمده است.

برای پیش‌بینی دبی جریان رودخانه با استفاده از روش‌های هوشمند مصنوعی، مهم‌ترین و اساسی‌ترین گامی که به طور مستقیم بر عملکرد مدل‌ها تأثیرگذار است، انتخاب مناسب و بهینه متغیرهای ورودی به مدل است که هم از پیچیدگی بیهوده مدل برای پیش‌بینی دبی جریان و هم از اتلاف بیهوده زمان در اجرای برنامه جلوگیری می‌کند؛ از این رو، در این بخش، برای ایستگاه مد نظر (گچسر)، داده‌های هیدرولوژیکی مربوط به دبی باید بررسی شود تا بهترین تعداد متغیر ورودی با به حساب آوردن تأخیر سری زمانی معرفی شود. برای تعیین مناسب تأخیر تعداد روزهای قبل به عنوان تأخیر زمانی (t ، n) از تابع خودهمبستگی جزئی PACF استفاده می‌شود



شکل ۳- سری زمانی دبی جریان در ایستگاه هیدرومتری گچسر

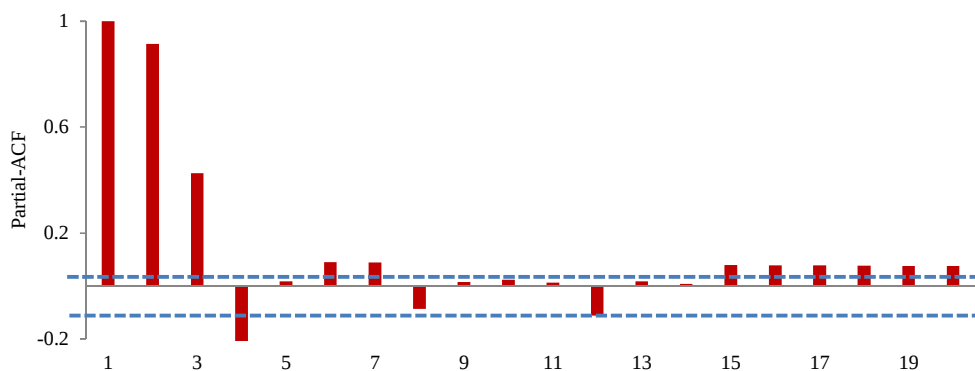
جدول ۳- متغیرهای آماری مجموعه داده‌ها برای ایستگاه گچسر

ایستگاه	مختصات جغرافیایی	ارتفاع (m)	مینیمم (m^3/s)	ماکزیمم (m^3/s)	میانگین (m^3/s)	انحراف معیار
گچسر	$36^{\circ}06'N$ $51^{\circ}19'E$	۲۲۵۸	۰	۶۱	۳/۹۳	۳/۹۹

مرحله آموزش، ضعیف‌ترین پیش‌بینی را با توجه به شاخص ریشه میانگین کمترین مربعات خطا ($RMSE=6.74$) و ($MAE=5.51$) در مقایسه با دیگر روش‌های مستقل و ترکیبی برای پیش‌بینی دبی روزانه جریان داشت. علت این ضعف می‌تواند این مسئله باشد که هر مدل، متغیرهایی درونی دارد که تأثیر آنها بر عملکرد مدل‌ها بسیار زیاد است و روش انفیس نیز از این قاعده مستثنا نیست و باید متغیرهای این مدل را با سعی و خطا به دست آورد؛ بنابراین، این مورد می‌تواند عدم قطعیت زیادی برای نتایج به دست آمده از روش انفیس برای پیش‌بینی دبی جریان رودخانه به همراه داشته باشد. در این ایستگاه، روش مدل درخت با ترکیب روش تجزیه مد تجربی گروهی، بهترین دقت ($R=0.96$) را در پیش‌بینی دبی روزانه جریان رودخانه کرج داشته است.

برای مدل‌سازی با روش‌های مستقل ANFIS، MT و مدل ترکیبی EEMD-MT داده‌های ورودی و خروجی به مدل معرفی می‌شود و برنامه برای پیش‌بینی دبی جریان شروع به کار می‌کند؛ اما برای مدل‌های هوشمند مصنوعی برپایه روش‌های پیش‌پردازش‌کننده، هر یک از این متغیرها با استفاده از تکنیک پیش‌پردازش‌کننده EEMD، پیش از وارد شدن به مدل، باید به چند سیگنال فرکانس بالا و پایین (زیرسری‌های تجزیه شده) تبدیل شود تا نویز اضافی به عنوان خطا در داده‌ها و روند از داده‌های سری زمانی حذف شود. برای استخراج سیگنال دبی جریان رودخانه از الگوریتم تجزیه مد تجربی در این پژوهش از کد موجود در برنامه Matlab استفاده شد.

جدول ۴ نیز مقادیر شاخص‌های آماری خطا را برای بررسی چگونگی عملکرد و دقت هر یک از روش‌های پیشنهادی ارائه می‌دهد. در این ایستگاه، روش انفیس در



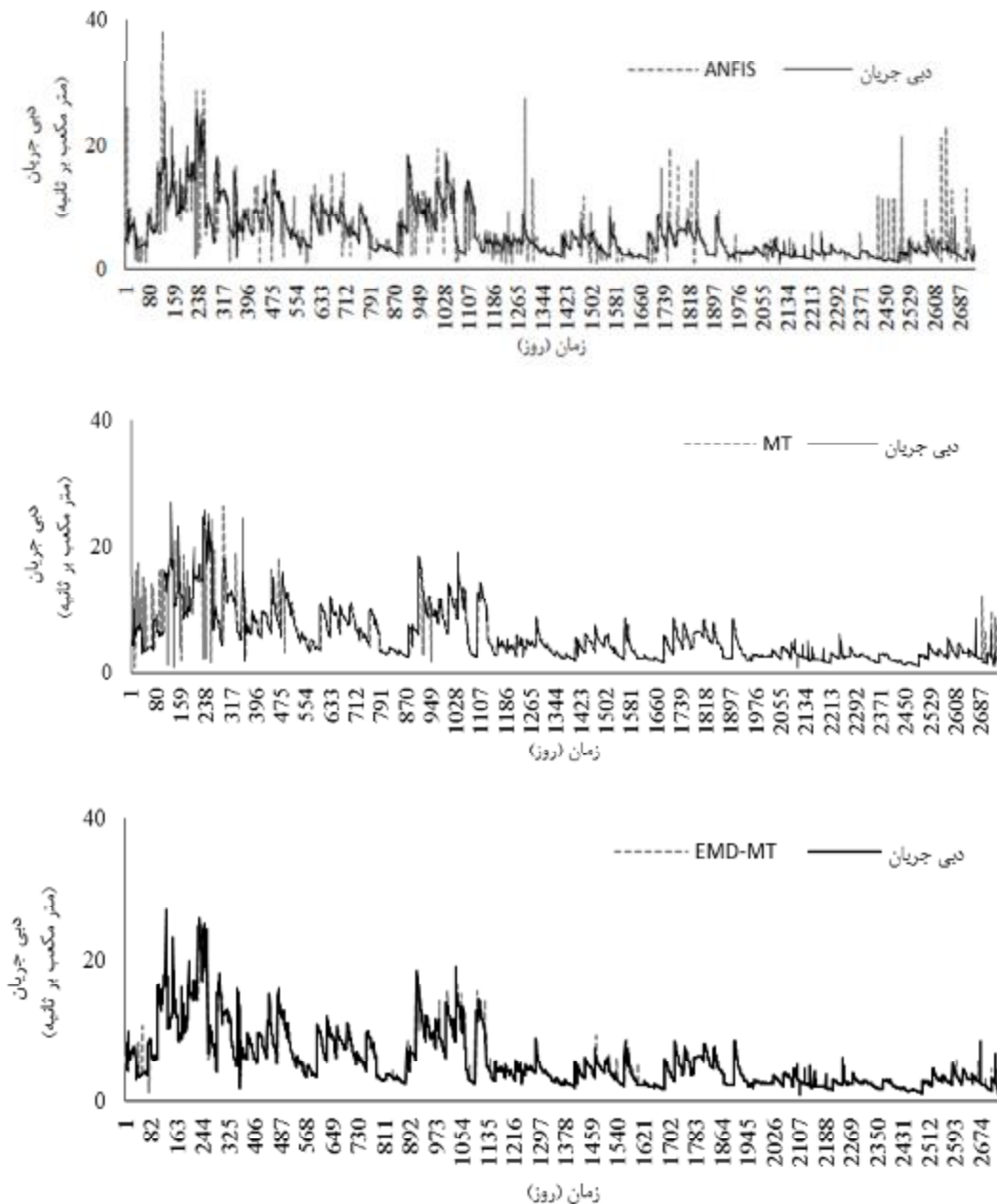
شکل ۴- نمودار PACF برای ایستگاه گچسر

جدول ۴- نتایج عملکرد روش‌های مستقل و ترکیبی برای ایستگاه گچسر

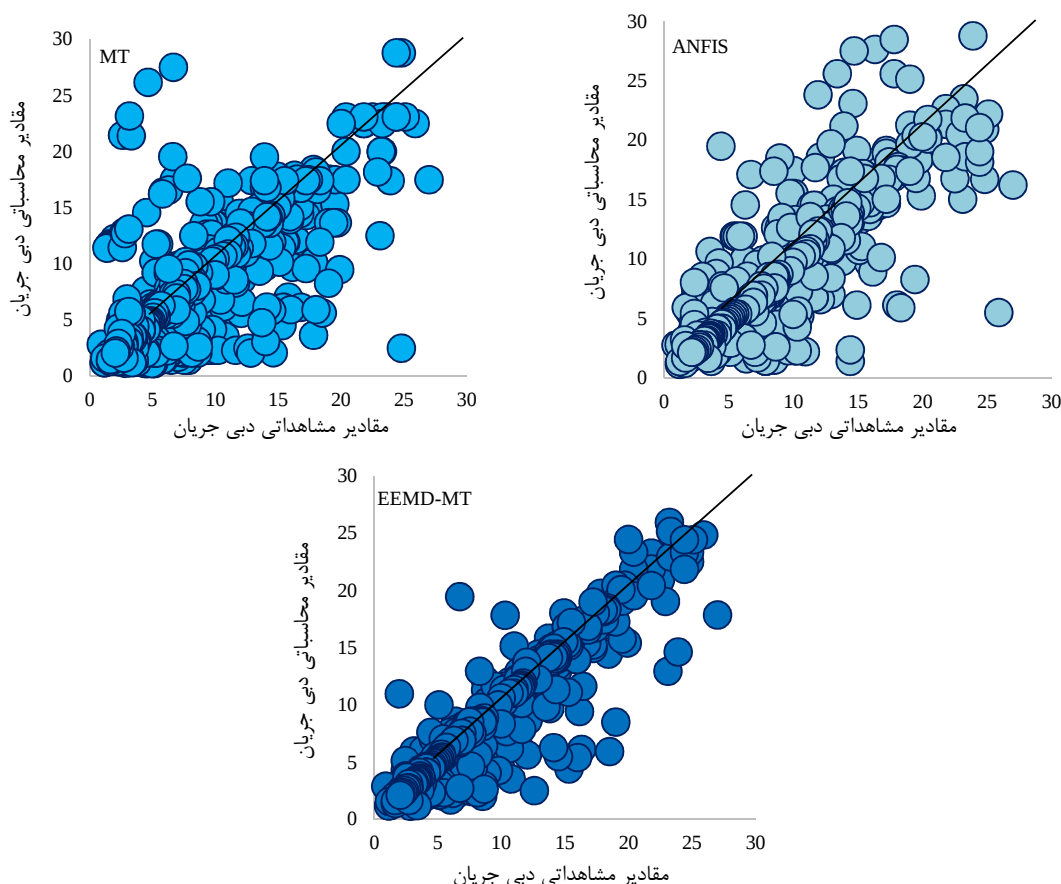
مدل‌ها	R	RMSE(m^3/s)	RSE(m^3/s)	MAE(m^3/s)	RAE(m^3/s)
مرحله آموزش مدل‌های مستقل و ترکیبی					
ANFIS	۰/۸۱	۶/۷۴	۲/۷۹	۵/۵۱	۴/۸۹
MT	۰/۸۶	۴/۸۵	۱/۷۶	۲/۴۸	۳/۸۵
EEMD-MT	۰/۹۶	۱/۲۱	۰/۲۳۶	۱/۰۸	۰/۵۶
مرحله آزمایش مدل‌های مستقل و ترکیبی					
ANFIS	۰/۸۰	۲/۹۵	۱/۱۶	۱/۲۵	۳/۴۲
MT	۰/۷۸	۳/۳۶۸	۰/۵۹	۱/۱۲۹	۲/۶۸۹
EEMD-MT	۰/۹۱	۱/۶۴	۰/۱۸۹	۰/۶۸	۰/۸۸۶

جریان، شکل‌های ۵ و ۶، نشان‌دهنده چگونگی پیش‌بینی مدل‌های هوشمند مصنوعی برای دوره مد نظر است. در پایان، براساس نتایج به‌دست‌آمده، اینگونه به نظر می‌رسد که روش پیش‌پردازش‌کننده تجزیه مد تجربی در ترکیب با روش مستقل مدل درخت، کارآمد بود و در بیشتر پیش‌بینی‌ها دقت قابل قبول‌تری ارائه داد.

در روش EEMD-MT، شاخص خطای مطلق میانگین، برابر ۱/۰۸ و شاخص خطای مطلق نسبی، ۰/۵۶ گزارش شده است که با داشتن کمترین مقادیر، نشان‌دهنده قدرت زیاد آموزش در این ایستگاه بود. برای تجزیه و تحلیل بیشتر دقت و عملکرد هر یک از روش‌های مستقل و ترکیبی در پیش‌بینی دبی روزانه



شکل ۵- نمودار سری زمانی دبی مشاهداتی و محاسباتی برای هر یک از مدل‌های پیشنهادی در ایستگاه گچسر



شکل ۶- نمودار پراکندگی دبی مشاهداتی و محاسباتی برای مدل‌های مستقل و ترکیبی در ایستگاه گچسر

نتیجه‌گیری

هدف این پژوهش، استفاده از تکنیک‌های داده‌کاوی در تخمین دبی جریان رودخانه در مقیاس روزانه مبتنی بر الگوریتم‌های تطبیقی هوشمند با استفاده از پایگاه داده‌های منابع آب بود. خلاصه نتایج این پژوهش در ذیل بیان شده است.

- در این پژوهش، از روش‌های هوشمند مصنوعی مانند مدل درخت و انفیس برای پیش‌بینی دبی جریان روزانه رودخانه گچسر کرج استفاده شد. برای تعیین و پیش‌بینی مناسب تعداد روزهای قبل به‌عنوان تأخیر زمانی $t-n$ از تابع خودهمبستگی جزئی PACF استفاده می‌شود که قضاوت اولیه ای برای تشخیص وابستگی هر متغیر هیدرولوژیکی مانند دبی جریان به چند روز پیش است. براساس این تابع، برای حوضه گچسر نیز تأخیر زمانی به‌عنوان متغیر ورودی مدل به میزان ۳ روز در نظر گرفته شد. در این سناریو، داده‌های روزانه دبی رودخانه ایستگاه گچسر به دو

قسمت آموزش (۷۵ درصد داده‌ها) و آزمایش (۲۵

درصد داده‌ها) تقسیم شد.

- برای مدل‌های هوشمند مصنوعی برپایه روش پیش‌پردازش‌کننده، هر یک از متغیرهای هواشناسی با استفاده از تکنیک پیش‌پردازش‌کننده EEMD، پیش از وارد شدن به مدل، باید به چند سیگنال فرکانس بالا و پایین (زیرسری‌های تجزیه‌شده) تبدیل شود تا نویز اضافی به‌عنوان خطا در داده‌ها و روند از داده‌های سری زمانی حذف شود. نتایج نشان داد ترکیب هر یک از مدل‌های هوشمند مصنوعی با رویکرد EEMD، عملکرد قابل قبولی از پیش‌بینی دبی روزانه جریان رودخانه‌ها داشت.

- نتایج مربوط به پیش‌بینی هر یک از مدل‌های پیشنهادی در این پژوهش برای ایستگاه گچسر نشان داد روش مدل ANFIS در مرحله آموزش، ضعیف‌ترین پیش‌بینی را با توجه به شاخص ریشه میانگین مربعات خطا ($RMSE=6.74$) و

خاک (علوم و فنون کشاورزی و منابع طبیعی).
۱۱(۱): ۳۹-۲۵.

۴. طاهری ح، افقری ن، صفوی ح.ر، و سرایی م.ح. ۱۳۸۹. تخمین پارامترهای هواشناسی با استفاده از تکنیک‌های داده‌کاوی در پایگاه داده‌های منابع آب، پنجمین کنگره ملی مهندسی عمران، مشهد، دانشگاه فردوسی مشهد، ۸ ص.

۵. عزیزی ح، ر. ۱۳۹۵. بررسی آثار تغییر اقلیم بر منابع آب حوضه آبریز رودخانه و سد کرج، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد شهر قدس، ۱۱۸ ص.

۶. فتحیان ح، و هرمزی‌نژاد ا. ۱۳۹۰. پیش‌بینی پارامترهای کمی و کیفی جریان در رودخانه کارون (بازه ملاتانی- فارسیات) با استفاده از شبکه‌های عصبی مصنوعی. مجله تالاب، دانشگاه آزاد اسلامی واحد اهواز، ۲(۸): ۴۳-۲۹.

۷. قاسمی زیارانی ا. ۱۳۸۴. ساماندهی حوضه آبخیز سد کرج (به‌منظور جلوگیری از آلودگی آب‌های حوضه)، پایان‌نامه کارشناسی ارشد، دانشکده محیط زیست دانشگاه تهران، ۱۵۳ ص.

۸. کارآموز م، عراقی‌نژاد ش، و رضوی س.س. ۱۳۸۳. مدل‌سازی هیدرواقليمی پیش‌بینی جریان ورودی به مخزن: کاربرد شبکه‌های عصبی مصنوعی، اولین کنفرانس مدیریت منابع آب، تهران، انجمن علوم و مهندسی منابع آب ایران، دانشکده فنی دانشگاه تهران، ۱۳ ص.

۹. مرشدلو ح، و عبدالله‌زاده بارفروش ا. ۱۳۸۶. داده‌کاوی پویا با استفاده از عامل هوشمند، اولین کنفرانس داده‌کاوی ایران، تهران، دانشگاه صنعتی امیرکبیر، مؤسسه پژوهشی داده‌پردازان گیتا، ۱۲ ص.

10. Ali M. Prasad R. Xiang Y. and Yaseen Z. M. (2020). Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. Journal of Hydrology. 584: 124647.

11. Cheng C. T. Niu W. J. Feng Z. K. Shen J. J. and Chau K. W. 2015. Daily reservoir runoff forecasting method using artificial neural network based on quantum-behaved particle swarm optimization. Water. 7(8): 4232-4246.

(MAE=5.51) در مقایسه با دیگر روش مستقل (MT) و ترکیبی (EEMD-MT) برای پیش‌بینی دبی روزانه جریان داشت. در این ایستگاه در مرحله آزمایش، روش مدل درخت با ترکیب روش تجزیه مد تجربی گروهی، بهترین دقت ($R=0.91$) را در پیش‌بینی دبی روزانه جریان رودخانه کرج داشت.

• همانگونه که متصور است، پیش‌بینی مقادیر زیادی از دبی جریان، که گاهی ممکن است به ایجاد سیلاب در مناطق شهری (مانند شهر ساری که رودخانه تجن از میان این شهر رد می‌شود) منجر شود، اهمیت زیادی دارد؛ از این‌رو، دو روش مدل درخت و مارس برای مقادیر بالای ۸۰ متر مکعب بر ثانیه، عملکرد ضعیفی در پیش‌بینی دبی جریان روزانه داشته که با به‌کارگیری روش ترکیبی، این موضوع حاصل شد. هیدروگراف، مقادیر مشاهده‌ای دبی جریان روزانه را برای هر یک از روش‌های هوشمند پیش‌بین نشان می‌دهد. در این هیدروگراف‌ها، این نکته مهم است که روش‌های مستقل ANFIS و MT برای پیش‌بینی مقادیر بالای $80 (m^3/s)$ روش‌های مطمئنی نیست و به‌طور عمده، در پیش‌بینی روزانه دبی جریان رودخانه گچسر و حتی رودخانه‌های موجود با مشخصات همین رودخانه، ضعیف است.

منابع

۱. برزگر ر، اصغری‌مقدم ا، و فیجانی ا. ۱۳۹۸. پیش‌بینی نوسانات ماهانه سطح آب دریاچه ارومیه با استفاده از مدل هوش مصنوعی مرکب نظارت‌شده. نشریه دانش آب و خاک، ۲۹(۱): ۱۶۵-۱۷۷.

۲. جلالی م، ۱۳۸۸. شبیه‌سازی جریان رودخانه با استفاده از شبکه عصبی مصنوعی (مطالعه موردی: حوضه قره‌آغاچ استان فارس)، پایان‌نامه کارشناسی ارشد، دانشکده منابع طبیعی، دانشگاه مازندران، ۱۲۶ ص.

۳. رضایی ع، مهدوی م، لوکس ک، فیض‌نیا س، و مهدیان م.ح. ۱۳۸۶. مدل‌سازی منطقه‌ای دبی‌های اوج در زیرحوضه‌های آبخیز سد سفیدرود با استفاده از شبکه‌های عصبی مصنوعی، علوم آب و

- neural network and ANFIS. *Construction and Building Materials*. 42: 205-216.
23. Tang J. Gao F. Liu F. and Chen X. 2020. A denoising scheme-based traffic flow prediction model: Combination of ensemble empirical mode decomposition and fuzzy C-Means neural network. *IEEE Access*. 8: 11546-11559.
 24. Wang W. C. Chau K.W. Qiu L. and Chen Y. B. 2015. Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental research*. 139: 46-54.
 25. Wu Z. and Huang N. E. 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*. 1(01): 1-41.
 26. Witten I.H. and Frank E. 2005. *Data mining: practical machine learning tools and techniques with Java implementations*, San Francisco: Morgan Kaufmann.
 27. Zadeh L. A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy sets and systems*. 1(1): 3-28.
 12. Huang N. E. Shen Z. Long S. R. Wu M. C. Shih H. H. Zheng Q. Yen N. C. Tung C. C. and Liu H. H. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*. 454: 903-995.
 13. Kişi Ö. 2009. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering*. 14(8): 773-782.
 14. Nabizadeh M. Mosaedi A. and Dehghani A.A. 2012. Intelligent Estimation of Stream Flow by Adaptive Neuro-Fuzzy Inference System. *Journal of Water and Irrigation Management*. 2: 69-80.
 15. Nadiri A. A. Fijani E. Tsai F. T. C. and Asghari Moghaddam A. 2013. Supervised committee machine with artificial intelligence for prediction of fluoride concentration. *Journal of Hydroinformatics*. 15(4): 1474-1490.
 16. Nadiri A. A. and Yousefzadeh S. (2017). A Comparison of the Performance of Artificial Neural Network, Fuzzy Logic and Adaptive Neuro-Fuzzy Inference Systems Models in the Estimation of Aquifer Hydraulic Conductivity. A Case Study: Maraghe-Bonab Aquifer. *Hydrogeomorphology*. 3(10): 21-40.
 17. Prasad R. Deo R. C. Li Y. and Maraseni T. 2017. Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmospheric research*. 197: 42-63.
 18. Quinlan J. R. 1992. Learning with continuous classes. In: Adams, Sterling, editors, *Proceedings of AI'92*. World Scientific. 1992: 343-348.
 19. Rezaie-Balf M. and Kisi O. 2018. New formulation for forecasting streamflow: evolutionary polynomial regression vs. extreme learning machine. *Hydrology Research*. 49(3): 939-953.
 20. Rezaie-Balf M. Fani Nowbandegani S. Samadi S. Z. Fallah H. and Alaghmand S. 2019. An Ensemble Decomposition-Based Artificial Intelligence Approach for Daily Streamflow Prediction. *Water*. 11(4): 709.
 21. Solomatine D. P. and Dulal K. 2003. Model tree as an alternative to neural network in rainfall-runoff modeling. *Hydr. Sci. J.* 48(3): 399-411.
 22. Sadrmohtazi A. Sobhani J. and Mirgozar M. A. 2013. Modeling compressive strength of EPS lightweight concrete using regression,