



Development of a nonlinear hybrid model for forecasting monthly reservoirs inflow Based on hydro-climate parameters and basin vegetation cover (Case study: Dez dam)

Pouria Zahirpour¹ | Saman Javadi^{2*} | Farimah Sadat Jamali³ | Ali Mohammadi⁴

1. Department of Water Engineering, Faculty of Agriculture Technology (Aburaihan), College of Agriculture and Natural Resources, University of Tehran, Tehran, Iran. pouryazahirpor@gmail.com

2. Department of Water Engineering, Faculty of Agriculture Technology (Aburaihan), College of Agriculture and Natural Resources, University of Tehran, Tehran, Iran. javadis@ut.ac.ir

3. Department of Physical Geography, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran. f_jamali@sbu.ac.ir

4. Department of Water Engineering, Faculty of Agriculture Technology (Aburaihan), College of Agriculture and Natural Resources, University of Tehran, Tehran, Iran. ali.mohammadi@ut.ac.ir

article history

Received December 29, 2023

Revised February 04, 2024

Accepted March 08, 2024

Abstract

Dez dam has an essential role in controlling the floods and water supply in the plains of Khouzestan. Therefore, accurate and reliable prediction of the flow input to the reservoir is crucial to make decisions about the exploitation and management of this reservoir. In this research, a novel WAVELET-ARIMA-NARX model was developed to predict the monthly inflow to Dez dam reservoir. The WAVELET method was used to break down the time series into sub-series and better analyzing, the ARIMA model was used to model the linear component of the analyzed series, and the NARX model was used to model the error resulting from the WAVELET-ARIMA model. The NDVI index was also used to measure the accuracy change of the WAVELET-ARIMA-NARX hybrid model. The results showed that the prediction performance of the WAVELET-ARIMA-NARX hybrid model has improved significantly compared to the ARIMA model. In such a way according to the RMSE criterion, the prediction accuracy in the verification and training stages has decreased by 74% and 82%, compared to the ARIMA and WAVELET-ARIMA model, respectively. Also, the NDVI parameter along with the average temperature and rainfall of the basin as an input of the NARX model has increased the accuracy of the model. In the model with 10 neurons in the hidden layer, compared to the two-parameter model of rainfall and NDVI with 15 neurons, in evaluation section, the MAE values have decreased from 27.2 to 18.5 and the RMSE from 0.45 to 0.26. These values indicate the importance and impact of simultaneous consideration of three parameters on forecasting accuracy.

Keywords

WAVELET theory, Seasonal Index, Artificial Neural Network, Hybrid Model, Dez dam



توسعه مدل هیبریدی غیرخطی برای پیش‌بینی جریان ماهانه ورودی مخازن سدها براساس پارامترهای هیدروکلیما و پوشش گیاهی حوضه (مطالعه موردی: سد دز) پوریا ظهیرپور^۱ | سامان جوادی^{۲*} | فریماه سادات جمالی^۳ | علی محمدی^۴

۱. گروه مهندسی آب، دانشکده فناوری کشاورزی (ابوریحان)، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، تهران، ایران. pouryazahirpor@gmail.com
۲. گروه مهندسی آب، دانشکده فناوری کشاورزی (ابوریحان)، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، تهران، ایران. javadis@ut.ac.ir
۳. گروه جغرافیای طبیعی، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. f.jamali@sbu.ac.ir
۴. گروه مهندسی آب، دانشکده فناوری کشاورزی (ابوریحان)، دانشکدگان کشاورزی و منابع طبیعی، دانشگاه تهران، تهران، ایران. ali.mohammadi@ut.ac.ir

تاریخ‌های مقاله

تاریخ دریافت: ۱۴۰۲/۱۰/۰۸

تاریخ بازنگری: ۱۴۰۲/۱۱/۱۵

تاریخ پذیرش: ۱۴۰۲/۱۲/۱۸

چکیده

سد دز نقش اساسی در کنترل سیلاب‌های رودخانه دز ایفا کرده و از طرفی تأمین‌کننده آب مورد نیاز کشاورزی دشت‌های خوزستان است. بنابراین پیش‌بینی دقیق و قابل اعتماد ورودی جریان به مخزن، یک مرجع حیاتی برای تصمیم‌گیری در مورد بهره‌برداری و مدیریت این مخزن است. در این تحقیق مدل WAVELET-ARIMA-NARX برای پیش‌بینی جریان ماهانه ورودی مخزن سد دز توسعه یافت. از روش WAVELET برای تجزیه سری زمانی به زیر سری‌ها و تحلیل بهتر سری زمانی، از مدل ARIMA به منظور مدل‌سازی مؤلفه خطی سری‌های تجزیه شده و از مدل NARX جهت مدل‌سازی خطای حاصل از مدل WAVELET-ARIMA استفاده شد. از شاخص NDVI حوضه نیز برای سنجش تغییر دقت مدل هیبریدی WAVELET-ARIMA-NARX بهره برده شد. نتایج نشان داد عملکرد پیش‌بینی مدل هیبریدی WAVELET-ARIMA-NARX نسبت به مدل ARIMA بهبود قابل توجهی یافته است، به نحوی که با توجه به معیار RMSE، دقت پیش‌بینی در مراحل صحت‌سنجی و آموزش، نسبت به مدل ARIMA و WAVELET-ARIMA به ترتیب ۷۴ و ۸۲ درصد کاهش یافته است. همچنین پارامتر NDVI به همراه دما و بارش متوسط حوضه به عنوان ورودی مدل NARX دقت مدل را افزایش داده به این صورت که در این مدل با ۱۰ نرون در لایه پنهان، در مقایسه با مدل دو پارامتری بارش و NDVI با تعداد ۱۵ نرون، در بخش ارزیابی مدل مقادیر MAE از ۲۷/۲ به ۱۸/۶ و RMSE از ۴۵/۰ به ۲۶/۰ رسیده است که این مقادیر اهمیت و تأثیر در نظرگر فتن هم‌زمان سه پارامتر بر دقت پیش‌بینی را نشان می‌دهد.

کلیدواژگان

تئوری موجک، شاخص فصلی، شبکه عصبی مصنوعی، مدل هیبریدی، سد دز

مقدمه

در کشور ایران که در اقلیم خشک و نیمه خشک قرار دارد، همواره موضوعات متأثر از خشکسالی مانند تأمین آب، چالش‌های فراوانی به دنبال داشته است [۱] که می‌توانند توسعه پایدار کشاورزی در کشور را تحت تأثیر قرار دهند [۲]. در این راستا، سد های مخزنی در تأمین نیازهای شرب، کشاورزی و صنعت از اهمیت ویژه‌ای برخوردارند. در این راستا، برآورد مقدار آورد رودخانه به مخزن سد از موضوعات بااهمیت محسوب می‌شود. مخازن به عنوان یک زیرساخت مهم برای مدیریت و بهره‌برداری از منابع آب، مزایای بسیاری از جمله کنترل سیل، کاهش خشکسالی، آبیاری کشاورزی، تولید برق آبی و تأمین منابع آب را برای ما به ارمغان می‌آورد [۳]. بنابراین پیش‌بینی دقیق جریان ورودی مخزن برای برنامه‌ریزی و مدیریت منابع آب موجود ضروری است. پیش‌بینی جریان ورودی را می‌توان برای بهره‌برداری و مدیریت مخزن اعمال کرد. این بر توسعه مدلی تأکید دارد که دقیق باشد و بتواند به راحتی برای پیش‌بینی جریان ورودی مخزن عملیاتی اعمال شود [۴]. پیش‌بینی بلندمدت جریان به مدیران و برنامه‌ریزان آب کمک می‌کند تا برای رهاسازی یا ذخیره آب مخازن در زمان‌های حساس تصمیم بهتری اتخاذ کنند. هدف از پیش‌بینی بلندمدت جریان، تخمین حجم آورد در یک نقطه خاص در دوره زمانی مشخص در آینده است.

بررسی داده‌ها به منظور شناسایی و تعیین ماهیت یک پدیده، قدم اول در مدل‌سازی اصولی است که با استفاده از آن می‌توان در گام بعدی به نوعی پیش‌بینی صحیح و قابل اتکا دست یافت. لذا آگاهی مناسب و عمیق از این دانش، بازخورد مطلوب تصمیمات را در آینده را تضمین می‌کند. آنالیز داده‌های هیدرولوژیک می‌تواند اطلاعات آماری بسیار مهم گذشته و آینده را مورد بررسی قرار دهد. با توجه به تغییرات شدید جریان ورودی، پیش‌بینی جریان را بسیار دشوار می‌کند. بنابراین، یک روش توانمند برای کمک به درک ماهیت چنین پدیده‌ای ضروری است [۵]. مدل‌های پیش‌بینی متعددی طی سال‌های اخیر ایجاد و استفاده شده‌اند. سری‌های زمانی پیش‌بینی به تکنیک‌های خطی و غیرخطی تقسیم می‌شوند. در میان آن‌ها، میانگین متحرک یکپارچه اتورگرسیو (ARIMA) و شبکه عصبی مصنوعی (ANN) به طور گسترده در پژوهش‌های صورت‌گرفته استفاده شده‌اند [۶ و ۷].

مدل میانگین متحرک یکپارچه اتورگرسیو (ARIMA) از پذیرفته شده‌ترین مدل‌های خطی است [۸]. ARIMA به دلیل دقت قابل توجه پیش‌بینی و انعطاف‌پذیری در مدل کردن الگوی بسیاری از انواع سری‌های زمانی، به خوبی عمل می‌کند. با این وجود، مهم‌ترین محدودیت، شکل خطی ظاهری داده‌های مرتبط است که آن‌ها را برای مدل‌سازی سری‌های زمانی غیرخطی چندوجهی نامناسب می‌کند [۹]. یکی از مدل‌های رایج پیش‌بینی غیرخطی شبکه عصبی مصنوعی (ANN) است که از عملکرد مغز انسان الهام گرفته شده است [۱۰]. آن‌ها به طور فزاینده‌ای برای مسائل مختلف هیدرولوژیک در دهه گذشته به کار گرفته شده‌اند [۱۱]. با این حال، برخی از مشکلات هیدرولوژیک مانند جریان ورودی با تغییرات شدید فصلی، فقط خطی یا غیرخطی نیستند و ترکیبی از دو مؤلفه خطی و غیرخطی هستند. تبدیل موجک (WT) تأثیرات مفید معناداری بر عملکرد مدل‌سازی دارد. رویکرد موجک یک تحلیل طیفی وابسته به زمان است که سری‌های زمانی را در فضای فرکانس زمانی تجزیه می‌کند تا توصیفی در مقیاس زمانی از فرایندها ارائه دهد. با این حال، برتری اصلی تبدیل موجک توانایی آن‌ها در به دست آوردن هم‌زمان اطلاعات در مورد زمان، مکان و فرکانس یک سیگنال است. تبدیل موجک گسسته تنها از زیرمجموعه‌ای از مقیاس‌ها و موقعیت‌ها برای انجام محاسبات استفاده می‌کند و برای تحلیل سری‌های زمانی راحت‌تر است [۱۲]. انواع روش‌های پیش‌بینی برای زیرمجموعه‌های ایجاد شده با تبدیل موجک گسسته یا پیوسته اعمال می‌شود. بعداً این زیرمجموعه‌های پیش‌بینی شده با بازسازی موجک برای تکمیل پیش‌بینی ترکیب می‌شوند [۱۳]. در

سال های اخیر، WT به یک روش رایج در مطالعات هیدرولوژیک تبدیل شده است [۱۴ و ۱۵]. تجزیه و تحلیل موجک (WA) یک روش قوی تر برای تجزیه و تحلیل سری های زمانی است. WA در هیدرولوژی، می تواند روشی مؤثر برای تجزیه و تحلیل سری های هیدرولوژیک ارائه دهد [۱۶]. در تحقیقات علمی اگرچه تبدیل موجک به طور گسترده برای آنالیز سری های زمانی هیدرولوژیک به کار گرفته شده است، اما مسائل کلیدی بسیاری در مورد آن وجود دارد، که عمدتاً شامل دو مسئله اصلی به عنوان انتخاب موجک و انتخاب سطح تجزیه است. از آنجا که هر روش دارای معایب خاص خود و شرایط کاربرد محدود است، نتایج دقیق تجزیه و تحلیل سری های هیدرولوژیک معمولاً با استفاده از هیچ روشی به تنهایی قابل دستیابی نیست. اگرچه تبدیل موجک نسبت به سایر روش های متداول بهتر عمل می کند، اما همان طور که پیش تر گفته شد، مسائل حل نشده زیادی دارد. برای غلبه بر این مشکل می توان از روش موجک همراه با مدل های پیش بینی متداول استفاده کرد [۱۷]. بهبود تکنیک های پیش بینی برای افزایش دقت در پیش بینی بلندمدت جریان روزانه برای عملکرد بهینه مخازن ضروری است. از این رو تحقیقات مختلفی در این راستا انجام شده است.

در پژوهش نیک منس و همکاران (۱۳۹۴) به ارائه یک مدل هوشمند تلفیقی مبتنی بر شبکه عصبی مصنوعی و تبدیلات موجک پرداخته شد که برای شبیه سازی آبدهی متوسط ماهانه در رودخانه کر و ایستگاه پل خان مورد استفاده قرار گرفت. نتایج این پژوهش نشان داد مدل تلفیقی شبکه عصبی مصنوعی و تبدیل موجک با ۲ درجه تجزیه سازی برای مناسب ترین ساختار، بهترین نتایج را ارائه می کند [۱۸].

Seo و همکاران (۲۰۱۵) دقت مدل های WANN و WANFIS را برای موجک های مادر مختلف، از جمله موجک های Symmlet، Daubechies و Coiflet ارزیابی کردند. نتایج پژوهش آن ها نشان داد عملکرد مدل به مجموعه های ورودی و موجک های مادر بستگی دارد و تجزیه موجک با استفاده از موجک مادر db10، می تواند کارایی مدل های ANN و ANFIS را بیشتر بهبود بخشد [۱۹]. Iwok و همکاران (۲۰۱۶) عملکرد بین اتورگرسیو میانگین متحرک (ARIMA) - مدل فوریه (Fourier) و مدل موجک را مقایسه کردند. نتایج بررسی ها حاکی از آن بود که مدل موجک بهتر از مدل ARIMA-Fourier عمل می کند [۲۰]. Banihabib و همکاران (۲۰۱۷) با ارائه یک مدل هیبریدی DARIMA-NARX عملکرد آن را مورد ارزیابی قرار دادند و تأثیر داده های بارندگی و شاخص اقلیمی (NAO) را روی مدل DARIMA-NARX بررسی کردند. مقایسه مدل ها نشان داد هر دو مدل ARIMA و DARIMA-NARX قادر به پیش بینی جریان روزانه بلندمدت برای سال بعد بودند [۲۱].

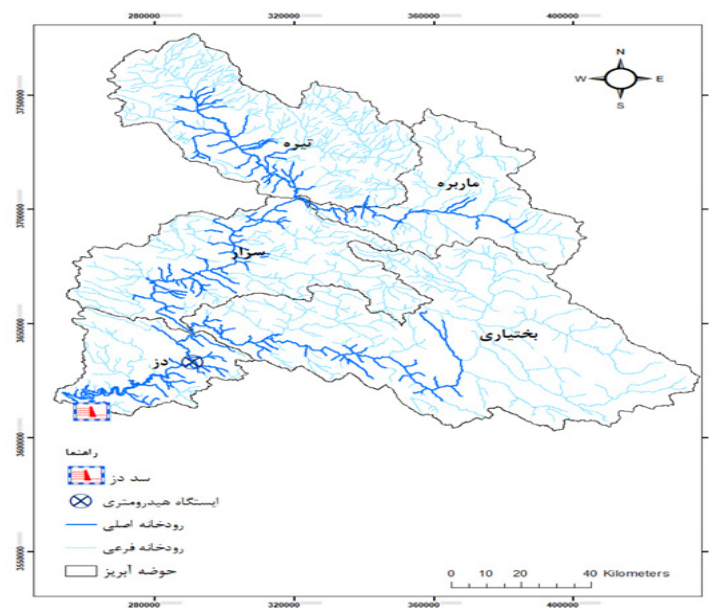
در پژوهش های صورت گرفته از مدل های متعددی جهت پیش بینی جریان ورودی به مخازن استفاده شده است و هر یک از این مدل ها دارای ضعف ها و عملکردهای محدودی هستند. سری های زمانی پیش بینی به تکنیک های خطی و غیرخطی تقسیم می شوند و برخی از مسائل هیدرولوژیک مانند جریان ورودی با تغییرات شدید فصلی روبه رو بوده، که صرفاً خطی یا غیرخطی نیستند و لذا ترکیبی از دو مؤلفه خطی و غیر خطی هستند. بنابراین در این پژوهش جهت مدل کردن مؤلفه خطی آبدهی از مدل ARIMA و شبکه عصبی مصنوعی (ANN) برای مدل کردن مؤلفه غیر خطی استفاده شده است. دلیل انتخاب مدل های نامبرده موفقیت عملکرد پیش بینی با توجه به داده های مورد استفاده در تحقیقات پیشین است. در این تحقیق از موجک WAVELET برای تجزیه سری زمانی به زیرسری ها و تحلیل بهتر سری زمانی، از مدل ARIMA به منظور مدل سازی مؤلفه خطی سری های تجزیه شده و از مدل NARX برای مدل سازی خطای حاصل از مدل WAVELET-ARIMA و از شاخص NDVI حوضه جهت افزایش دقت مدل هیبریدی WAVELET-ARIMA-NARX استفاده شد. شایان یادآوری است که محدودیت

پژوهش حاضر، مقیاس زمانی (ماهانه) بوده که اگر این مقیاس در بازه زمانی روزانه انجام می‌شد بر دقت کار افزوده می‌شد و همچنین، به دلیل گسسته بودن سری‌های زمانی هیدرولوژیک و همچنین زمان کم در دسترس در این پژوهش از موجک گسسته استفاده شده است. از مهم‌ترین نوآوری‌های صورت‌گرفته در این پژوهش می‌توان به: ۱. استفاده از متد موجک جهت پیش‌پردازش داده‌ها و آنالیز بهتر سری زمانی جهت مدل‌سازی کارآمدتر؛ ۲. استفاده از پارامترهای پیش‌بینی‌کننده دما، بارش و شاخص نرمال شده اخلاف پوشش گیاهی (NDVI) برای افزایش دقت مدل هیبریدی و ۳. بهره‌گیری از توابع موجک مختلف در تجزیه سری زمانی اشاره کرد.

مواد و روش‌ها

منطقه مطالعه شده

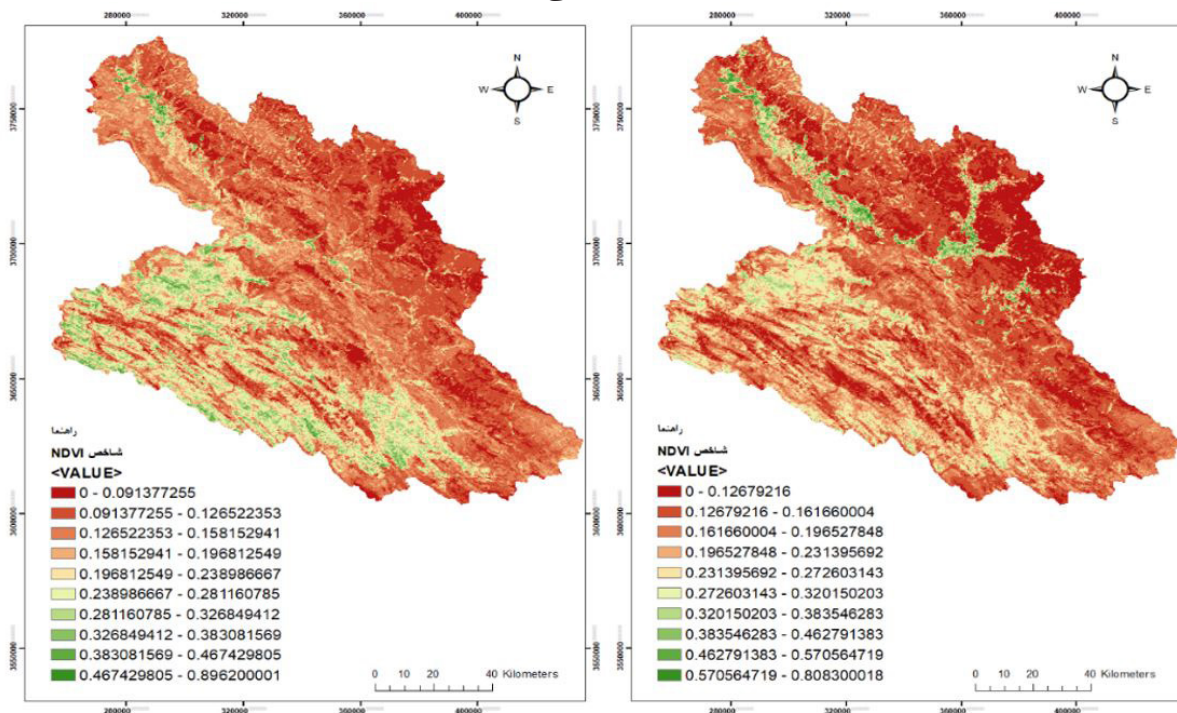
حوضه آبریز دز بخشی از ارتفاعات زاگرس میانی است. این حوضه در محدوده عرض جغرافیایی بین ۳۲ درجه و ۳۵ دقیقه تا ۳۴ درجه و ۷ دقیقه شمالی و طول جغرافیایی بین ۴۸ درجه و ۲۰ دقیقه تا ۵۰ درجه و ۲۰ دقیقه شرقی، در جنوب غربی ایران واقع شده است (شکل ۱). حوضه دز از غرب به حوضه کرخه، از شمال به حوضه قره‌چای و زاینده‌رود و از شرق و جنوب به حوضه کارون محدود می‌شود. رودخانه دز از ارتفاعات جبال زاگرس سرچشمه می‌گیرد و از دو شاخه اصلی به نام‌های سزار و بختیاری تشکیل می‌شود که پس از هم پیوستن رودخانه دز را تشکیل می‌دهند. رودخانه دز با طول ۳۳۶ کیلومتر و آبدهی متوسط ۲۴۱ متر مکعب در ثانیه از استان لرستان سرچشمه گرفته و با مساحت ۲۳۲۲۹ کیلومتر مربع باکد حوضه ۲۳۳ جزء حوضه بزرگ کارون است که بعد از پیوستن رودخانه‌های کوچک و بزرگ بسیاری در استان لرستان و خوزستان حوضه آبریز دز را تشکیل می‌دهند [۲۲].



شکل ۱. نقشه حوضه آبریز سد دز و موقعیت سد دز

داده‌های مورد استفاده در این پژوهش آبدهی ماهانه رودخانه دز، دما و بارش است که از دفتر مطالعات شرکت مدیریت منابع آب ایران اخذ شد. طول دوره آماری ۱۸ سال که از سال آبی ۱۳۷۹-۱۳۸۰ تا سال ۱۳۹۶-۱۳۹۷ و داده‌ها به صورت ماهانه است. از داده‌های سال آبی ۱۳۹۶-۱۳۹۷ در مرحله صحت‌سنجی مدل‌ها استفاده شد. در این پژوهش جهت افزایش دقت مدل هیبریدی، پارامتر NDVI مورد بررسی قرار داده شد. به این منظور از محصول

سنجنده MODIS با عنوان استفاده MOD13Q1 استفاده شد. تصاویر این محصول که دارای وضوح ۲۵۰ متر و فاصله تصویربرداری ۱۶ روزه است از سایت www.ladsweb.modaps.eosdis.nasa.gov دانلود شد. علت انتخاب این محصول سنجنده MODIS دقت زیاد آن نسبت به محصولات دیگر آن بود. با توجه به طول دوره آماری (۱۸ سال) و محصول MOD13Q1 با فاصله زمانی ۱۶ روز، ۴۱۴ تصویر اخذ شد. تصاویر با استفاده از نرم افزار ARCGIS تحلیل و NDVI هر یک از تصاویر که مربوط به حوضه دز است محاسبه شد. در مرحله بعد یک سری زمانی ۱۶ روزه که تغییرات NDVI طی ۱۸ سال را نشان می دهد به سری زمانی ماهانه طی ۱۸ سال تبدیل شد. شکل ۲، NDVI حوضه آبریز دز در مهرماه ۱۳۷۹ و شهریورماه ۱۳۹۷ را نشان می دهد.



شکل ۲. نقشه NDVI محاسبه شده حوضه آبریز دز (سمت راست: شهریور ۱۳۹۷ و سمت چپ: مهر ۱۳۷۹)

نرمال سازی داده ها

وقتی داده های سری زمانی دارای توزیع نرمال و طول دوره آماری کافی باشند مدل های خطی کمترین خطا را در پیش بینی جریان رودخانه دارند. از جمله روش های نرمال سازی می توان به لگاریتم گیری، باکس-کاکس، جذگیری و نمایی اشاره کرد [۲۳]. در این پژوهش داده های آبدی ماهانه رودخانه دز را با استفاده از لگاریتم گیری نرمال شد.

ایستایی داده ها

برای بررسی وضعیت ایستایی داده ها می توان از نمودار تابع خودهمبستگی نگار (ACF) استفاده کرد. اگر همبستگی بین یک سری با سری های تشکیل شده از گام های مختلف زمانی آن به دست آورده شده (r_k) و ضرایب این همبستگی را در مقابل گام های تأخیر زمانی ترسیم شود، نمودار حاصل را خودهمبستگی نگار می نامند که از آن برای تشخیص ایستایی داده ها و الگوی احتمالی مدل داده ها استفاده می شود [۲۴]. تابع خودهمبستگی در رابطه ۱ با تأخیر K نشان داده شده است:

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad -1 \leq r_k \leq 1 \quad (1)$$

اگر در نمودار خود همبستگی نگار، τk ها با آهنگی مناسب میرا شوند و در این محدوده قرار بگیرند، آن سری زمانی ایستا در نظر گرفته می شود که در آن n داده ها است.

آزمون های ایستایی

آزمون ریشه واحد augmented Dickey-Fuller

آزمون ریشه واحد Dickey – Fuller از طریق تخمین OLS یا مدل های رگرسیونی با وجود یک روند خطی به وجود آمده است. مدل که یک مدل خود همبسته با تأخیر یک است عبارت است از رابطه $[25]$:

$$x_t = \rho \cdot x_{t-1} + \varepsilon_t \quad t = 1, 2, \dots, N \quad (2)$$

که در آن ε_t سری تصادفی نرمال و استاندارد و مستقل با میانگین صفر و واریانس σ^2 است، در صورتی که با استفاده از روش حداکثر درستنمایی مقدار $\hat{\rho}$ به صورت رابطه ۳ و ۴ محاسبه می شود $[25]$.

$$\hat{\rho} = (\sum_{t=2}^N X_{t-1}^2)^{-1} \cdot \sum_{t=2}^N X_t \cdot X_{t-1} \quad (3)$$

$$\hat{t} = \frac{\hat{\rho} - 1}{\hat{\sigma}_p} \quad (4)$$

که در آن خطای استاندارد OLS برای ضریب است. Dickey و Fuller حد توزیع آزمون t را در شرایط فرض و مورد بررسی قرار دارند که عبارت اند از (رابطه ۵) $[25]$:

$$\begin{cases} H_0: \rho = 1 \\ H_1: |\rho| < 1 \end{cases} \quad (5)$$

در صورت احراز شرایط فرض سری نایستا بوده و در شرایط فرض سری ایستا خواهد بود.

مدل های تصادفی

مدل مرکب اتورگرسیو. میانگین متحرک (ARMA)

مدل ARMA از ترکیب دو فرایند رگرسیو و میانگین متحرک به دست می آید و شامل p جمله از اتورگرسیو و q جمله از میانگین متحرک است و به صورت رابطه ۶ است $[26]$:

$$(1 - \alpha_1 B - \dots - \alpha_p B^p)(x_t - \mu) = (1 + \beta_1 B + \dots + \beta_q B^q)z_t \quad (6)$$

که در آن، مقدار متغیر در زمان t ، بیانگر فرایند تصادفی محض با میانگین صفر و واریانس و پارامتر میانگین متحرک در گام تأخیر q و ضریب مدل اتورگرسیو در گام تأخیر p است.

مدل مرکب خودرگرسیو میانگین متحرک انباشته (ARIMA)

شرط بهره گیری از مدل های خودرگرسیو میانگین متحرک (ARMA) ایستا بودن سری های زمانی است. برای ایستا کردن سری های زمانی اغلب از روش تفاضل استفاده می شود. جزء یکپارچگی در مدل خودرگرسیو میانگین متحرک انباشته بیانگر میزان تفاضل مورد نیاز برای ایستا کردن سری زمانی است. مدل های خودرگرسیو میانگین متحرک انباشته را معمولاً به صورت «ARIMA(p,d,q)» نشان داده می شوند در آن p مرتبه مورد نیاز برای خودرگرسیو، d میزان تفاضل مورد نیاز برای ایستا کردن سری زمانی و q میزان مرتبه مورد نیاز برای میانگین متحرک است. اگر یک

فرایند «ARIMA(p, 1, q)» در نظر گرفته شود [۲۷] (رابطه ۷):

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (7)$$

به طوری که در آن y_t و e_t به ترتیب مقادیر واقعی و باقیمانده‌های مدل در دوره t ، $\theta_i (i = 1, 2, \dots, q)$ و $\phi_j (j = 1, 2, \dots, p)$ اعداد صحیح و بیانگر مرتبه مدل هستند.

تبدیل موجک (WT)

تبدیل موجک (Wavelet Transform) یکی از تبدیلات مهم ریاضی است که در حوزه‌های مختلف علوم کاربرد دارد. ایده اصلی تبدیل موجک این است که بر ضعف‌ها و محدودیت‌های موجود در تبدیل فوریه غلبه کند. این تبدیل را به خلاف تبدیل فوریه، می‌توان در مورد سیگنال‌های غیر ایستا و سیستم‌های دینامیک نیز مورد استفاده قرار داد. یکی از حوزه‌های علمی که روش‌های موجک در آن کاربردهای زیادی پیدا کرده‌اند، تحلیل سری‌های زمانی گسسته است.

تبدیل موجک پیوسته (CWT)

تبدیل موجک یک روش تجزیه مؤثر است. این روش یک روش تحلیلی از یک سیگنال را هم در زمان و هم در فرکانس ارائه می‌کند و به نظر می‌رسد موفق‌تر از تبدیل‌های فوریه معمولی است که تحلیل فرکانس زمانی را برای متغیرهای شامل سیگنال‌های غیر ثابت ارائه نمی‌دهند. مقیاس زمانی تبدیل موجک یک سیگنال زمانی پیوسته $x(t)$ ، به صورت رابطه ۸ تعریف می‌شود [۲۸].

$$W(s, \tau) = |s|^{-1/2} \int_{-\infty}^{\infty} \psi^* \left(\frac{t-\tau}{s} \right) x(t) dt \quad (8)$$

$$\tau \in R, s \in R, s \neq 0$$

که در آن $\psi(t)$ تابع موجک یا موجک مادر است. s مقیاس یا ضریب فرکانس، τ ضریب زمان، R دامنه اعداد حقیقی است. رابطه ۸ توضیح می‌دهد که تبدیل موجک تجزیه سری زمانی در چندین مقیاس و وضوح مختلف است. سری اصلی را می‌توان از طریق تبدیل معکوس بازسازی کرد. تبدیل موجک پیوسته در تمامی مقیاس‌ها می‌تواند عمل کند، اما در عین حال به تعداد زیاد داده و زمان محاسباتی بالایی نیاز دارد. بنابراین، به دلیل گسسته بودن سری‌های زمانی هیدرولوژیک و همچنین زمان کم در دسترس در این پژوهش از موجک گسسته استفاده شده است.

تبدیل موجک گسسته (DWT)

برای کاربردهای عملی در هیدرولوژی، محققان به جای سیگنال زمان پیوسته به سیگنال زمان گسسته دسترسی دارند [۲۴]. گسسته‌سازی رابطه ۸ بر اساس قانون دوزنقه‌ای شاید ساده‌ترین گسسته‌سازی تبدیل موجک پیوسته (CWT) باشد. این تبدیل ضرایب $2N$ را از یک مجموعه داده به طول N تولید می‌کند. از این رو، اطلاعات غیر ضروری در ضرایب محبوس می‌شود که ممکن است ویژگی مطلوبی باشد یا نباشد برای غلبه بر این مشکل، تبدیل موجک گسسته (DWT) که قدرت دو مقیاس لگاریتمی را ارائه می‌دهد، می‌تواند در کاربردهای عملی استفاده شود. با توجه به پیچیدگی محاسباتی CWT ترجیح دادن DWT بر CWT می‌تواند مفید واقع شود [۲۴]. به منظور غلبه بر افزونگی اطلاعاتی که گفته شد فاصله بندی لگاریتمی یکنواخت می‌تواند برای گسسته‌سازی مقیاس s با رزولوشن متناظر در موقعیت τ استفاده شود. این عملیات، N ضریب تبدیل برای تکمیل فرایند تشریح یک سری زمانی به

طول N را مقدور می‌کند. چنین موجک گسسته‌ای به فرم زیر است (رابطه ۹):

$$\psi_{m,n}(t) = \frac{1}{\sqrt{s^m}} \psi\left(\frac{t-n\tau \cdot s^m}{s^m}\right) \quad (9)$$

در اینجا m و n اعداد صحیحی هستند که به ترتیب عوامل مقیاس و انتقال موجک را کنترل می‌کنند؛ مقدار S_0 باید بزرگ‌تر از یک و τ_0 باید بزرگ‌تر از صفر باشد. برای یک سری زمانی گسسته X_t تبدیل موجک به صورت رابطه ۱۰ به دست می‌آید:

$$W_{j,k} = 2^{-j/2} \sum_{t=0}^{N-1} \psi(2^{-j}t - k)x_t \quad (10)$$

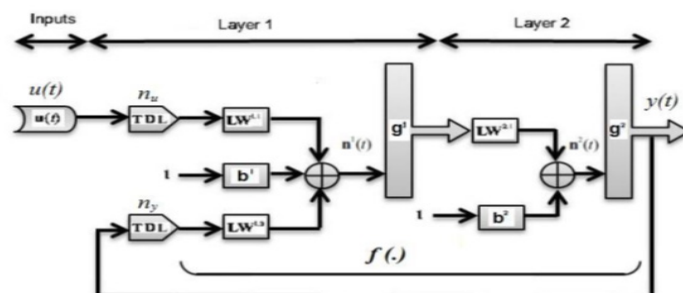
در اینجا ضریب موجک برای موجک گسسته در مقیاس و مکان است. برای انواع مختلف تبدیل موجک، مصالحه بین فشردگی (Compact) و صاف بودن (Smooth) با یکدیگر تفاوت دارند. این ویژگی به این معناست که می‌توان نوع خاصی از تبدیل موجک را انتخاب کرد که با ویژگی (Features) مورد نظر برای استخراج از سیگنال تناسب بیشتری داشته باشد. DWT یک سیگنال را به موجک پدر و مادر تبدیل می‌کند. موجک‌های پدر، مؤلفه‌های با مقیاس بالا و فرکانس پایین را نشان می‌دهند (مؤلفه‌های تقریبی (A)). موجک‌های مادر نمایشی از اجزای با مقیاس پایین و فرکانس بالا (جزئیات (D)) هستند. بنابراین، DWT این امکان را می‌دهد تا رفتارهای سری‌های مختلف را در مقیاس‌های زمانی مختلف به طور مستقل مطالعه کرد.

شبکه عصبی مصنوعی NARX

شبکه‌های عصبی مصنوعی به عنوان زیرمجموعه‌ای از تکنیک‌های هوش مصنوعی از عناصر عملیاتی ساده‌ای ساخته می‌شوند که به صورت موازی در کنار هم عمل می‌کنند. این عناصر ساده از سیستم‌های عصبی زیستی الهام گرفته شده‌اند. در طبیعت عملکرد شبکه‌های عصبی از طریق نحوه اتصال بین اجزای تعیین می‌شود. شبکه اتورگرسیو غیر خطی با ورودی‌های برون زایک شبکه پویای با اتصالات پس‌خوردی به لایه‌های مختلف شبکه است. مدل NARX براساس مدل ARX است که معمولاً در مدل‌سازی سری زمانی استفاده می‌شود. معادله مدل NARX به صورت رابطه ۱۱ است [۲۹]:

$$y(t) = f(y(t-1), y(t-r), \dots, y(t-n_y), u(t-1), u(t-r), \dots, u(t-n_u)) \quad (11)$$

که در آن f یک نگاشت انجام گرفته به وسیله شبکه عصبی مصنوعی و y پیش‌بینی است. اعداد n_y و n_u به ترتیب وقفه‌های ورودی برون‌زا و مقادیر واقعی هدف بوده، که به وسیله سیستم شبکه عصبی تعیین می‌شوند. همچنین می‌توان مدل NARX را برای تقریب تابع f با استفاده از یک شبکه عصبی پیش‌خور اجرا کرد. این شبکه در مدل‌سازی سری‌های زمانی و همچنین در مدل‌سازی سیستم‌های پویای غیر خطی استفاده می‌شود. نمودار شبکه حاصل که در آن از یک شبکه پیش‌خور دو لایه برای تقریب تابع استفاده شده در شکل ۳ آورده شده است.



شکل ۳. شبکه پیش‌خور دو لایه

مدل هیبریدی WAVELET-ARIMA-NARX

با هدف بهبود دقت پیش‌بینی، در این پژوهش ایده در نظر گرفتن سری زمانی را به عنوان ترکیب یک مؤلفه ثابت خطی (SL)، و یک ساختار غیرخطی (SN) در نظر گرفته می‌شود (رابطه ۱۲).

$$S(t) = S_L(t) + S_N(t) \quad (12)$$

ایده این است که سری زمانی فصل‌زدایی شده توسط موجک تجزیه می‌شود. با توجه به اینکه سری زمانی آبدهی رودخانه جزء سری‌های هیدرولوژیک بوده و ماهیتی گسسته دارد، بنابراین از تبدیل موجک گسسته برای تجزیه سری‌سزمانی در این پژوهش استفاده می‌شود. انتخاب تابع موجک مناسب به نام موجک مادر و سطح مناسب تجزیه موجک یک موضوع بسیار مهم است، زیرا مطالعه‌ای وجود ندارد که بهترین عملکرد مدل را برای یک تابع موجک خاص یا سطح تجزیه نشان دهد. تجزیه و تحلیل‌های دقیق در مورد عملکرد موجک‌های مادر مختلف در شبیه‌سازی‌ها هیدرولوژیک به این نتیجه رسیده‌اند که تعیین موجک مادر ایده‌آل نیازمند آزمایش آن با سعی و خطاست [۳۰]. ابتدا از طریق ARMIA مدل‌سازی می‌شود. این روش WAVELET-ARIMA است که جزء خطی سری‌سزمانی را مدل می‌کند. پس از مدل‌سازی سری‌سزمانی آبدهی ماهانه رودخانه دز توسط مدل WAVELET-ARIMA، باقی‌مانده مدل WAVELET-ARIMA (مدل خطی) که شامل قسمت غیرخطی سری‌های زمانی است از رابطه ۱۳ به دست می‌آید.

$$e(t) = S(t) - \hat{S}_L(t) \quad (13)$$

درخور یادآوری است که قسمت غیرخطی که توسط رابطه ۱۳ محاسبه شده، خطایی است که توسط مدل WAVELET-ARIMA در پیش‌بینی ایجاد شده یا به عبارتی مدل نامبرده نتوانسته الگوی آن را استخراج کند. بنابراین برای استخراج الگوی غیرخطی یا همان خطای ایجاد شده توسط مدل WAVELET-ARIMA در این پژوهش از مدل شبکه عصبی NARX استفاده شد. در مدل NARX جهت افزایش دقت پیش‌بینی مدل هیبریدی از ورودی‌های دما، بارش و شاخص پوشش گیاهی (NDVI) استفاده شد. با بررسی‌های انجام شده، در بین شاخص دیگری مانند SAVI، PVI و غیره؛ این شاخص به دلیل سادگی محاسبه و در دسترس بودن پارامترهای ورودی، به عنوان پرکاربردترین شاخص در اغلب مطالعات استفاده شده که عملکرد قابل قبول و اطمینانی را از خود نشان داده است. شایان ذکر است که در این پژوهش ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد دیگر برای آزمون در نظر گرفته شد.

تعیین اینکه کدام الگوریتم آموزشی در مورد یک مسئله خاص سریع‌تر عمل می‌کند کار بسیار دشواری است،

زیرا این مسئله به فاکتورهای زیادی بستگی دارد از جمله پیچیدگی مسئله، تعداد داده‌های مجموعه آموزشی، تعداد وزن‌ها و بایاس‌ها در شبکه و مقدار خطای هدف. در این پژوهش از تابع آموزش لونبرگ-مارکوارت در شبکه عصبی مصنوعی NARX استفاده شد، زیرا عموماً الگوریتم لونبرگ-مارکوارت در اکثر مسائل کارایی بالایی را از خود نشان داده و از دقت و سرعت بالایی برخوردار است و در بسیاری از موارد این الگوریتم به حداقل خطا رسیده است. معیار ارزیابی

در پژوهش‌های مدل‌سازی، یکی از بخش‌های مهم بررسی عملکرد مدل‌ها است که روش‌های گوناگونی برای این مهم ارائه شده است [۳۱]. در این پژوهش به منظور ارزیابی عملکرد مدل‌های مورد استفاده برای پیش‌بینی آبدهی رودخانه از معیارهای مجذور متوسط خطا (RMSE) و میانگین خطای مطلق (MAE) استفاده شده است (روابط ۱۴ و ۱۵) [۳۲]:

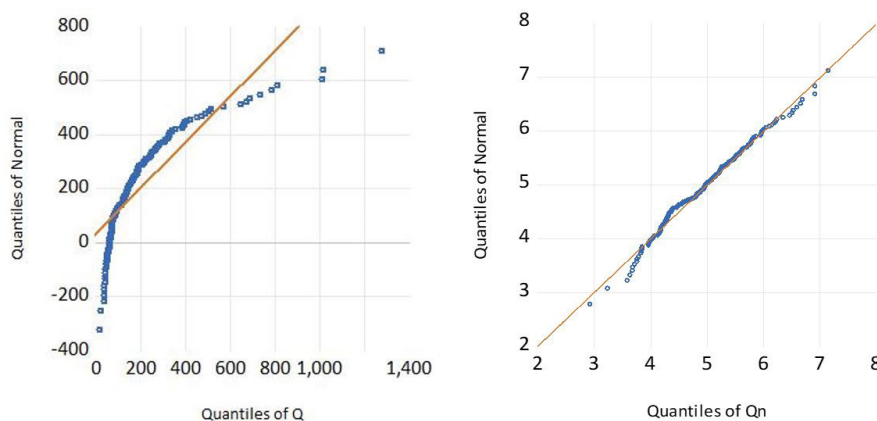
$$MAE = \frac{1}{N} \sum_{i=1}^N |O_i - P_i| \quad (14)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2} \quad (15)$$

و به ترتیب آبدهی مشاهداتی و آبدهی پیش‌بینی شده و N تعداد داده‌های مشاهداتی است.

نتایج و بحث

برای بررسی وضعیت نرمال بودن داده‌ها از نمودار Q-Q بهره گرفته شد (شکل ۴). با توجه به نمودار رسم شده، داده‌ها غیر نرمال هستند. بنابراین، باید با استفاده از روش مناسب اقدام به نرمال‌سازی داده‌ها کرد. در این پژوهش از روش لگاریتم‌گیری بر مبنای عدد طبیعی، داده‌ها را تبدیل کرده و وضعیت داده‌ها کنترل شد که داده‌ها تقریباً به خط نرمال نزدیک شده و می‌توان فرض نرمال بودن داده‌ها را پذیرفت.

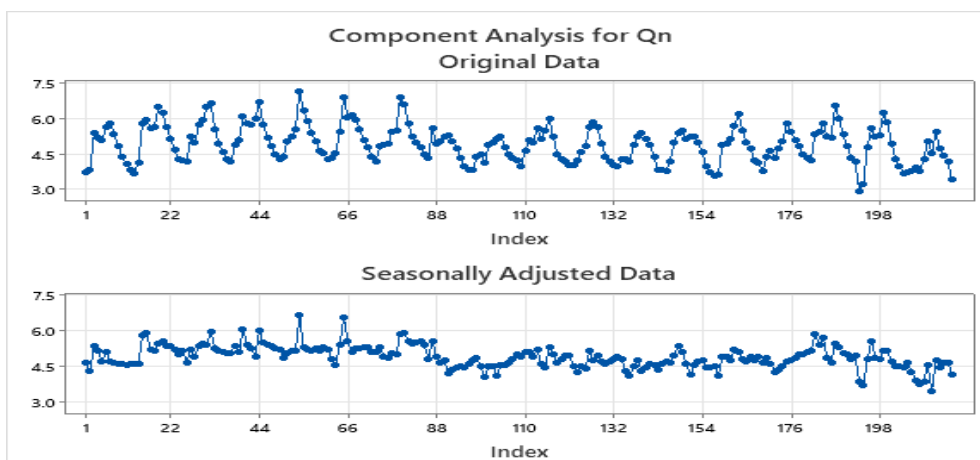


شکل ۴. وضعیت داده‌های سری زمانی قبل و بعد از نرمال‌سازی

حذف مؤلفه فصلی داده‌ها

در این پژوهش برای حذف مؤلفه فصلی داده‌ها از تعدیلات فصلی استفاده شد. به این منظور، با استفاده از نرم‌افزار Minitab داده‌ها را فصل‌زدایی کرده که در شکل ۵ نشان داده شده است. با توجه به اینکه داده‌ها دارای مقیاس ماهانه هستند و در تأخیرهای ۱۲، ۲۴، ۳۶ و... دارای مقادیر همبستگی زیادی است، بنابراین $m=12$ در نظر گرفته

شد و برای تعیین شاخص فصلی و حذف آن از میانگین متحرک استفاده شد. شاخص فصلی محاسبه شده از داده های ماهانه کم شد تا اثرات فصلی حذف شود. نمودار داده های فصل زدایی شده در شکل ۵ ارائه شده است.



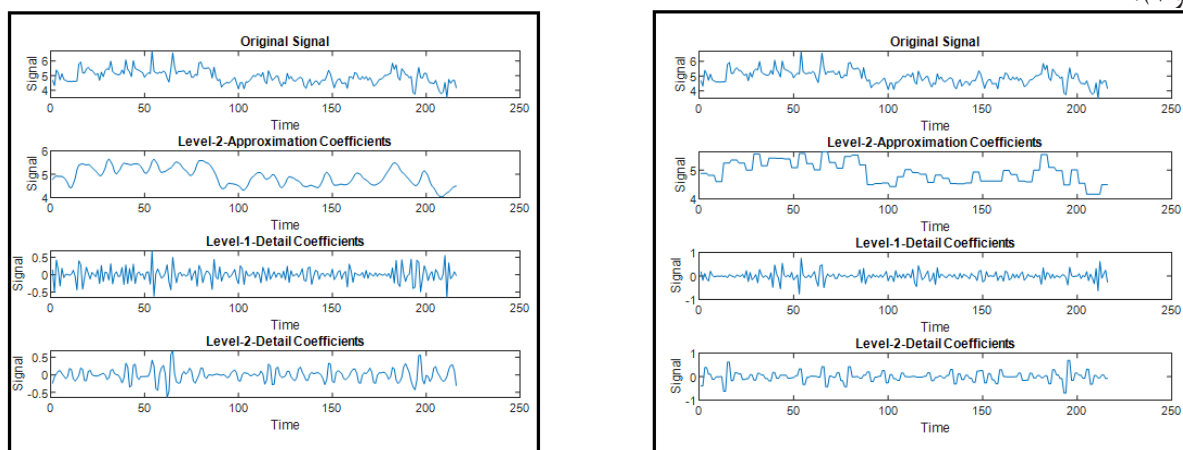
شکل ۵. نمودار سری فصل زدایی شده آبدهی رودخانه دز

تجزیه با استفاده از متد موجک گسسته (DWT)

با توجه به اینکه سری های هیدرولوژیک مشاهده شده در طبیعت معمولاً سیگنال های گسسته هستند، بنابراین اغلب از تبدیل موجک گسسته (DWT) استفاده می شود. در این پژوهش برای مشخص کردن تعداد بهینه سطح تجزیه از رابطه ۱۶ استفاده می شود [۳۰].

$$L = \text{int}[\log(N)] \quad (16)$$

L تعداد سطح تجزیه، N تعداد داده های آبدهی است. بنابراین با توجه به اینکه N برابر ۲۰۴ است تعداد سطوح تجزیه ۲ در نظر گرفته شد. تابع موجک 4DB و DB1 جهت تجزیه سری زمانی انتخاب شد. دلیل انتخاب این توابع موجک موفقیت عملکرد آن در تجزیه سری زمانی آبدهی رودخانه در تحقیقات قبلی بود [۳۳]. داده های فصل زدایی شده با استفاده از نرم افزار MATLAB و استفاده از تابع یاد شده به دو سطح تجزیه می شود (شکل های ۷ و ۶):



شکل ۶. سری زمانی تجزیه شده با موجک db1؛ شکل ۷. سری زمانی تجزیه شده با موجک db4

بررسی وضعیت ایستایی زیرسری‌های تجزیه شده

در این پژوهش جهت بررسی ایستایی سری‌های تجزیه شده از آزمون دیکی فولر تعمیم یافته (ADF) استفاده شد. نتایج آزمون سه سری آینده به شرح جدول‌های ۱ و ۲ است:

جدول ۱. نتایج آزمون adf سری تجزیه شده با تابع موجک db4

سری زمانی	D1	D2	A2
حد بحرانی	-۳/۴۶	-۳/۴۶	-۳/۴۶
t آماره	۱۰/۰۶	-۹/۷۷	-۱/۳۱
P-VALUE	۰/۰۰	۰/۰۰	۰/۶۳

جدول ۲. نتایج آزمون adf سری تجزیه شده با تابع db1

سری زمانی	D1	D2	A2
حد بحرانی	-۳/۴۶	-۳/۴۶	-۳/۰۶
t آماره	-۹/۳۲	-۸/۰۸	-۳/۴۶
P-VALUE	۰/۰۰	۰/۰۰	-۰/۰۳

با توجه به نتایج آزمون ADF مشاهده می‌شود در سری‌های 1D و 2D در توابع db1 و db4 ایستا هستند و نیازی به ایستا کردن ندارند اما سری 2A در دو تابع با توجه به P-VALUE و آماره آزمون مشاهده می‌شود این سری‌ها دارای روند و غیر ایستا هستند، بنابراین با استفاده از تفاضل‌گیری مرتبه اول اقدام به ایستاسازی سری‌ها شده است.

انتخاب مدل برتر براساس معیار AIC

در گذشته پارامترهای مدل ARIMA با استفاده از نمودارهای خود همبستگی (ACF) و خود همبستگی جزئی (PACF) انتخاب و تعیین می‌شدند اما اخیراً از معیار AIC جهت تعیین عملکرد نسبی پیش‌بینی استفاده می‌شود که توانسته است کارایی بالایی از خود نشان دهد [۳۴].

سری زمانی آینده رودخانه دز با توجه به ماهانه بودن، ابتدا توسط تعدیلات فصلی و میانگین متحرک فصل زدایی شده و داده‌های حاصل توسط توابع موجک db1 و db4 تجزیه شد. این سری‌ها با استفاده از مدل ARIMA مدل‌سازی شد و برای هر سری تجزیه شده برای ۱۲ ماه آینده پیش‌بینی صورت گرفت. داده‌های پیش‌بینی شده برای سری‌های تجزیه شده با استفاده از موجک معکوس بازسازی شد تا ۱۲ ماه پیش‌بینی شده توسط WAVELET-ARIMA حاصل شود. جدول ۳ نتایج ۱۰ مدل برتر سری آینده براساس معیار AIC را نشان می‌دهد. بر این اساس، مدل آریمای (۱،۱،۱) دارای کمترین مقدار AIC یعنی ۰/۸۰۴۳۷ و مدل ARIMA (۵،۱،۳) دارای بیشترین مقدار AIC یعنی ۰/۸۱۵۷۷ است.

جدول ۳. نتایج ۱۰ مدل برتر سری آینده براساس معیار AIC

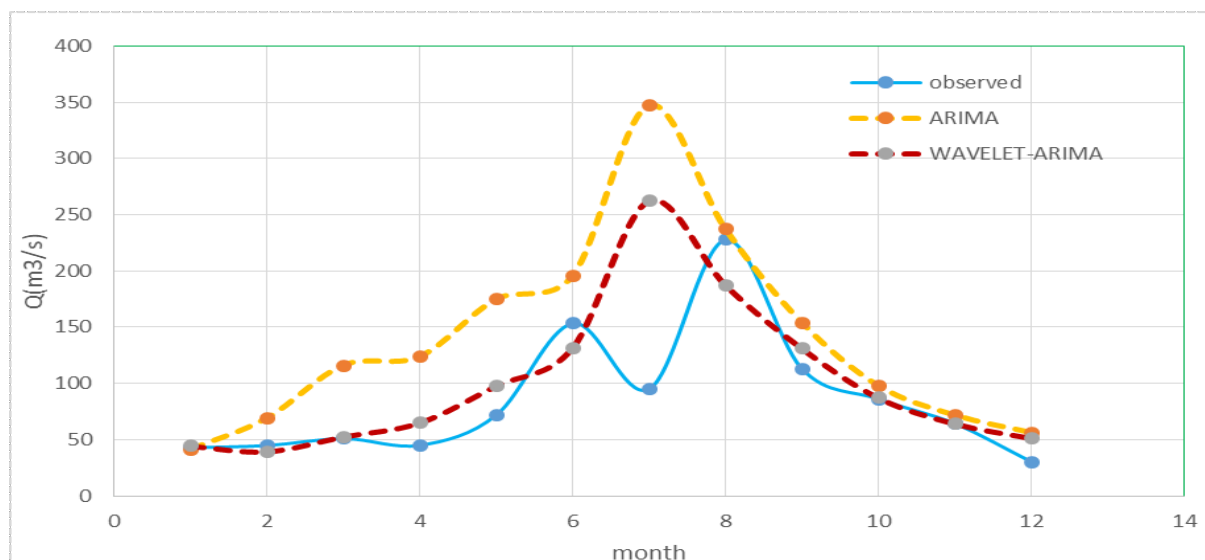
ردیف	مدل ARIMA	AIC	ردیف	مدل ARIMA	AIC
۱	(۱،۱،۱)	۰/۸۰۴۳۷	۶	(۰،۱،۳)	۰/۸۱۰۷۴
۲	(۰،۱،۲)	۰/۸۰۴۴۶	۷	(۲،۱،۱)	۰/۸۱۱۱۶
۳	(۱،۱،۳)	۰/۸۰۷۵۶	۸	(۳،۱،۱)	۰/۸۱۲۲۸
۴	(۵،۱،۶)	۰/۸۰۸۸۱	۹	(۲،۱،۶)	۰/۸۱۵۷۶
۵	(۱،۱،۲)	۰/۸۰۹۲۰	۱۰	(۵،۱،۳)	۰/۸۱۵۷۷

با استفاده از معیارهای ارزیابی RMSE و MAE عملکرد WAVELET-ARIMA و همچنین، تأثیر انتخاب نوع تابع موجک بر مدل هیبریدی بررسی شد. نتایج حاصل از مدل سازی در جدول ۴ قابل مشاهده است:

جدول ۴. نتایج ارزیابی عملکرد پیش بینی مدل هیبریدی WAVELET-ARIMA بر اساس نوع تابع موجک مورد استفاده

معیار ارزیابی			آموزش			صحت سنجی	
مدل	MAE	RMSE	RMSE/Qtrain	MAE	RMSE	RMSE/Qtest	
ARIMA	۶۰/۵۴	۱۳۱/۳۷	۰/۶۸	۵۵/۰۸	۸۶/۲۹	۱/۰۲	
(WAVELET-ARIMA)(db1)	۲۷/۱۱	۷۰/۵۷	۰/۳۶	۴۲/۸۷	۷۱/۱۶	۰/۸۳	
(WAVELET-ARIMA)(db4)	۱۵/۱۹	۳۸/۷۶	۰/۲۰	۲۶/۹۹	۵۱/۵۹	۰/۶۰	

با توجه به جدول ۳ ملاحظه می شود مدل WAVELET-ARIMA نسبت به مدل ARIMA عملکرد بهتری در آموزش و صحت سنجی داشته به طوری که RMSE در آموزش از ۱۳۱ به ۳۸ و در بخش صحت سنجی از ۸۶ به ۵۶ و معیار MAE در آموزش از ۶۰ به ۱۵ و در صحت سنجی از ۵۵ به ۲۶ کاهش یافته است. همچنین، می توان نتیجه گرفت که نوع تابع موجک در تجزیه سری زمانی مؤثر است، زیرا سری زمانی که با تابع موجک db4 تجزیه شده، در مدل پیش بینی خطای کمتری نسبت به سری زمانی که با تابع موجک db1 تجزیه شده است را دارد. مقادیر رسم شده در شکل ۸ مقدار آبدهی مشاهداتی و پیش بینی شده با مدل ARIMA-WAVELET و ARIMA را نشان می دهد. بر این اساس، مدل ARIMA موفق به مدل کردن نقاط دبی پیک نشده که ناشی از ضعف مدل ARIMA در پیش بینی مؤلفه غیرخطی است. با مشاهده نمودار رسم شده مدل ARIMA-WAVELET پیش بینی نقاط اوج دبی تا حدودی بهبود یافته که نشان دهنده تأثیر آنالیز موجک بر دقت پیش بینی سری زمانی آبدهی است.



شکل ۸. مقایسه آبدی مشاهده‌ای و پیش‌بینی شده توسط مدل WAVELET-ARIMA و مدل ARIMA در مرحله صحت‌سنجی

مدل سازی مدل WAVELET-ARIMA-NARX

در این تحقیق برای افزایش دقت مدل WAVELET-ARIMA از شبکه عصبی مصنوعی NARX استفاده شده است. مدل WAVELET-ARIMA مؤلفه خطی در سری زمانی را مدل می‌کند، اما در مدل سازی مؤلفه غیر خطی موفقیت چندانی نداشته است. اگر داده‌های مشاهده‌ای دبی رودخانه از مقدار پیش‌بینی شده توسط مدل WAVELET-ARIMA کم شود، مقدار خطا به دست می‌آید که نشان می‌دهد مدل نتوانسته این مقادیر را مدل کند که همان مؤلفه غیر خطی است. اگر نتایج پیش‌بینی شده توسط شبکه عصبی NARX و مدل WAVELET-ARIMA با هم جمع زده شود، مدل هیبریدی WAVELET-ARIMA-NARX حاصل می‌شود. پارامترهای ورودی شبکه عصبی NARX عبارت است از: پارامترهای هیدروکلیما (دما، بارش) و شاخص نرمال شده تفاوت پوشش گیاهی (NDVI) حوضه آبریز سد دز و مقدار پیش‌بینی شده توسط مدل WAVELET-ARIMA است. در واقع هدف شبکه عصبی NARX مدل کردن خطای حاصل از WAVELET-ARIMA است. در جدول ۵ حالت‌های مختلف پارامترهای ورودی به مدل شبکه عصبی NARX ارائه شده است.

جدول ۵. پارامترهای ورودی به شبکه NARX

علامت	A	B	C	D	E	F	G	H
پارامتر ورودی به مدل	$Qf(t)$	$Qf(t), P(t)$	$Qf(t), T(t)$	$Qf(t), N(t)$	$Qf(t), P(t), T(t)$	$Qf(t), P(t), N(t)$	$Qf(t), T(t), N(t)$	$Qf(t), P(t), T(t), N(t)$

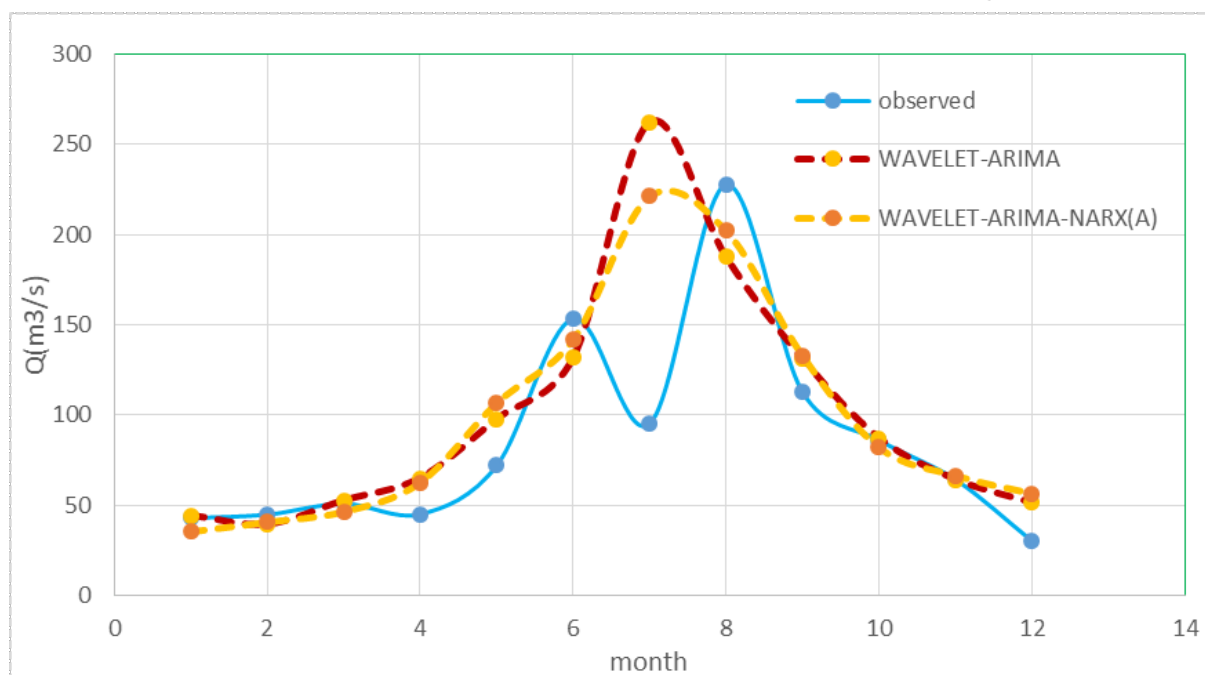
$Qf(t)$ مقدار پیش‌بینی شده با مدل WAVELET-ARIMA، $P(t)$ بارش متوسط حوضه آبریز، $T(t)$ دمای متوسط حوضه آبریز و $N(t)$ شاخص نرمال شده تفاوت پوشش گیاهی (NDVI) است. در این پژوهش حالت A به عنوان حالت مبنا در نظر گرفته شد و با افزایش سایر پارامترها دقت مدل مورد بررسی قرار گرفت. تعداد نرون‌های شبکه ۱ تا ۲۰ و تعداد تأخیر لایه خروجی ۱ تا ۲۰ ماه و الگوریتم آموزش شبکه NARX لونیبرگ-مارکوارت تعیین شد. با توجه به حالت‌های مختلف ورودی، برای هر ورودی ۴۰۰ پیش‌بینی (در مجموع ۳۶۰۰ پیش‌بینی) صورت گرفت و مدل برتر هر ورودی شناسایی شد. در جدول ۶ نتایج حاصل از مدل سازی مدل WAVELET-ARIMA-NARX

در مرحله آموزش و صحت سنجی ارائه شده است.

جدول ۶. نتایج حاصل از مدل سازی مدل WAVELET-ARIMA-NARX در مرحله آموزش و صحت سنجی

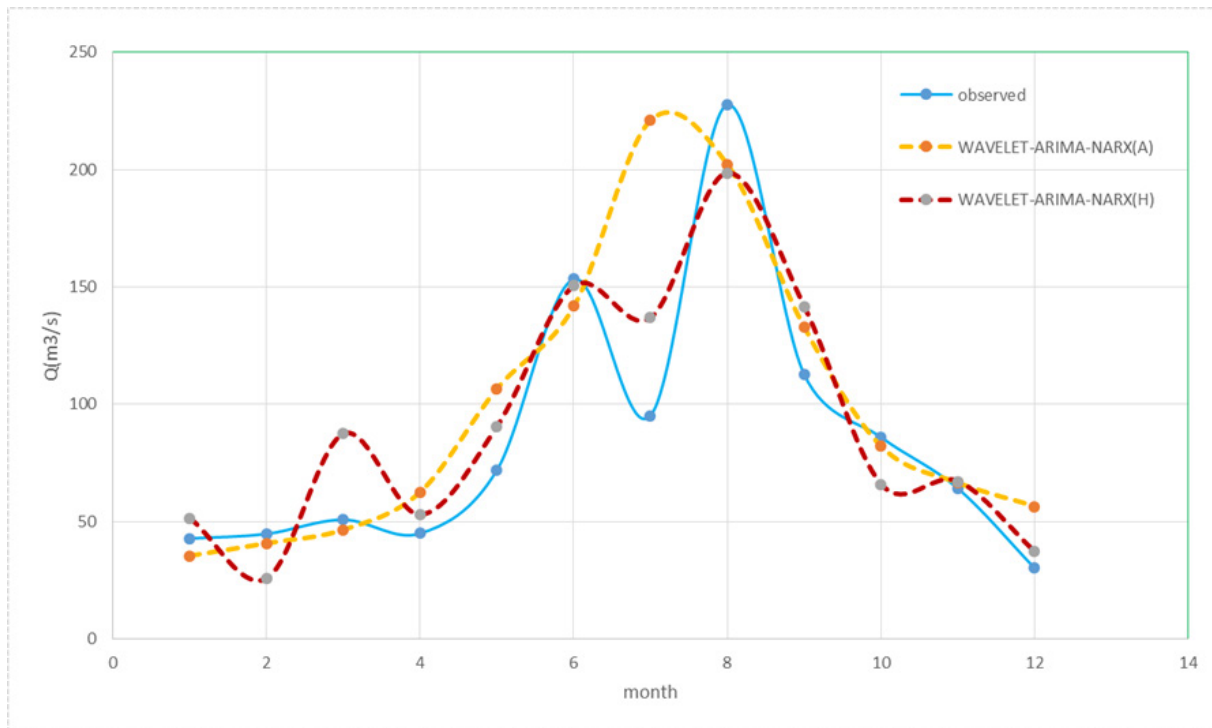
ارزیابی مدل			آموزش مدل			ساختار شبکه		مدل
RMSE/Qtest	RMSE	MAE	RMSE/Qtrain	RMSE	MAE	تأخیر خروجی	نرون لایه پنهان	
۰/۴۷	۴۰/۲۱	۲۳/۶۷	۰/۱۵	۲۸/۶۹	۱۵/۹۳	۷	۲۰	A
۰/۳۳	۲۸/۰۳	۲۴/۳۳	۰/۱۴	۲۸/۰۴	۲۱/۲۳	۱۶	۱۷	B
۰/۴۲	۳۵/۸۰	۲۳/۱۹	۰/۱۰	۱۸/۸۲	۷/۴۷	۱۶	۱۴	C
۰/۴۸	۴۱/۶۹	۳۴/۲۵	۰/۲۲	۴۳/۰۹	۲۴/۵۱	۱۲	۱۵	D
۰/۲۹	۲۵/۱۹	۱۸/۷۴	۰/۰۹	۱۷/۲۰	۸/۶۷	۱۷	۱۰	E
۰/۴۵	۳۸/۴۳	۲۷/۱۶	۰/۲۳	۴۵/۲۵	۲۷/۵۷	۱۳	۱۵	F
۰/۳۱	۲۶/۵۲	۲۱/۹۳	۰/۱۳	۲۶/۵۱	۱۶/۵۴	۱۵	۱۲	G
۰/۲۶	۲۲/۵۰	۱۸/۶۲	۰/۱۲	۲۳/۲۴	۷/۷۴	۱۵	۱۰	H

در حالت های A, B, C و E پیش بینی بدون تأثیر پارامتر NDVI صورت گرفته است و در حالت D, F, G و H با توجه به یکی از اهداف این پژوهش که تأثیر NDVI در دقت پیش بینی مدل هیبریدی بوده است، این پارامتر به مدل اضافه شد. با توجه به شکل ۹ مشاهده می شود که دقت مدل هیبریدی نسبت به WAVELET-ARIMA بهبود نسبی یافته، به طوری که مقدار پیش بینی شده دبی اوج مدل WAVELET-ARIMA-NARX(A) کاهش نسبی داشته و به مقدار مشاهداتی نزدیک شده است. نتایج معیارهای ارزیابی پژوهش در جدول ۵ نیز حاکی از این موضوع است، به طوری که مقدار RMSE در مرحله صحت سنجی از ۵۱ به ۴۱ و MAE از ۲۶ به ۲۳ کاهش یافته است، اما همچنان در پیش بینی دبی اوج دارای تأخیر یک ماهه است.



شکل ۹. مقایسه آبدهی مشاهداتی و پیش بینی شده توسط مدل WAVELET-ARIMA-NARX(A) و مدل WAVELET-ARIMA در مرحله صحت سنجی

باتوجه به نتایج جدول ۵ مشاهده می شود استفاده از سه پارامتر دما و بارش و NDVI تحت عنوان «پیش بینی H»، به صورت هم زمان بیشترین تأثیر را در دقت پیش بینی داشته و نتایج پیش بینی بهبود یافته است. استفاده از پارامتر NDVI به صورت منفرد (D) یا به همراه بارندگی (F) در دقت پیش بینی نسبت به حالت مبنای تغییر چندان نداشته است، اما استفاده از شاخص NDVI به همراه دما و بارش به صورت هم زمان به عنوان ورودی به مدل، دقت پیش بینی به صورت قابل ملاحظه ای افزایش پیدا کرده است. نمودار مدل پیش بینی H که شامل ورودی های دما، بارش و NDVI است در شکل ۱۰ ارائه شده که حاکی از دقت مناسب این رویکرد است.



شکل ۱۰. مقایسه آبدهی مشاهده‌ای و پیش‌بینی شده توسط مدل WAVELET-ARIMA-NARX(H) و مدل WAVELET-ARIMA-NARX(A) در مرحله صحت‌سنجی

نتیجه‌گیری

آنالیز داده‌های هیدرولوژیک می‌تواند اطلاعات آماری بسیار مهم گذشته و آینده را مورد بررسی و پیش‌بینی قرار دهد. با توجه به اینکه مدل‌های مورد استفاده در پیش‌بینی سری‌های هیدرولوژیک دارای محدودیت هستند، لذا در این تحقیق یک مدل هیبریدی توسعه داده شد که ضعف‌های مدل‌های منفرد را پوشش دهد. همچنین در این پژوهش از یک سری پارامترهای مؤثر جهت افزایش دقت پیش‌بینی جریان ورودی بهره‌برده شد. نتایج اجرای مدل WAVELET-ARIMA که مؤلفه خطی سری زمانی را پوشش می‌دهد، حاکی از افزایش دقت مدل WAVELET-ARIMA نسبت به ARIMA است. باقی‌مانده‌های مدل WAVELET-ARIMA که خطای ناشی از مدل بوده و جزء مؤلفه غیر خطی سری زمانی است، با استفاده از شبکه عصبی NARX مدل شد. در شبکه عصبی NARX از سه پارامتر دما، بارش و شاخص نرمال شده اختلاف پوشش گیاهی (NDVI) استفاده شد. نتایج نشان داد استفاده از شاخص پوشش گیاهی، دما و بارش به صورت هم‌زمان در مدل WAVELET-ARIMA-NARX به عنوان ورودی دقت مدل را افزایش خواهد داد. به بیانی دیگر، استفاده از سه پارامتر نامبرده به صورت هم‌زمان با

هم بیشترین تأثیر روی مقادیر پیش بینی شده را داشته و خطای مدل WAVELET-ARIMA-NARX به کمترین مقدار نسبت به حالت های دیگر می رسد. از طرف دیگر پارامترهای دما و بارش با هم به عنوان ورودی باعث کاهش خطای پیش بینی می شود.

بر اساس نتایج، دقت مدل هیبریدی توسعه داده شده نسبت به مدل WAVELET-ARIMA بهبود نسبی یافته، به طوری که مقدار پیش بینی شده دبی اوج مدل WAVELET-ARIMA-NARX(A) کاهش نسبی داشته و به مقدار مشاهداتی نزدیک شده است. نتایج معیارهای ارزیابی نتایج پژوهش بیانگر این موضوع است، به طوری که مقدار RMSE در مرحله صحت سنجی از ۵۱ به ۴۱ و MAE از ۲۶ به ۲۳ کاهش یافته است. اما همچنان در پیش بینی دبی اوج دارای تأخیر یک ماهه است. بنابراین با توجه به فرضیات پژوهش در استفاده از متد موجک به عنوان روش پیش پردازش داده ها، تأثیر نوع تابع موجک بر دقت پیش بینی و استفاده از پارامتر NDVI به عنوان یک پارامتر بهبود دهنده نتایج پیش بینی، این پژوهش نشان داد آنالیز موجک باعث افزایش دقت پیش بینی می شود و نوع تابع موجک بر نتایج پیش بینی مؤثر است و شاخص NDVI به همراه بارندگی و دما باعث افزایش دقت مدل می شود.

منابع

- [1] Banihabib ME, Jurik L, Khorasani MM, Shahdany SM, Mohammadi A, Pokrývková J. Optimizing Embedded Water Trades to Conserve Lakes in Arid and Semiarid Regions. *Polish Journal of Environmental Studies*. 2021 Jan 1;30(5):4413-23.
- [2] Yousefi H, Mohammadi A, Mirzaaghabeik M, Noorollahi Y. Virtual water evaluation for grains products in Iran Case study: pea and bean. *J water land dev*. 2017 Dec 1;35(1):275-80.
- [3] Ward PJ, de Ruyter MC, Mård J, Schröter K, Van Loon A, Veldkamp T, von Uexkull N, Wanders N, AghaKouchak A, Arnbjerg-Nielsen K, Capewell L. The need to integrate flood and drought disaster risk reduction strategies. *Water Security*. 2020 Dec 1;11:100070.
- [4] Herbert ZC, Asghar Z, Oroza CA. Long-term reservoir inflow forecasts: enhanced water supply and inflow volume accuracy using deep learning. *Journal of Hydrology*. 2021 Oct 1;601:126676.
- [5] Moeeni H, Bonakdari H. Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model. *Stochastic environmental research and risk assessment*. 2017 Oct;31(8):1997-2010.
- [6] Kassem AA, Raheem AM, Khidir KM. Daily streamflow prediction for khazir river basin using ARIMA and ANN models. *Zanco Journal of Pure and Applied Sciences*. 2020;32(3):30-9.
- [7] Allawi MF, Hussain IR, Salman MI, El-Shafie A. Monthly inflow forecasting utilizing advanced artificial intelligence methods: a case study of Haditha Dam in Iraq. *Stochastic Environmental Research and Risk Assessment*. 2021 Nov;35(11):2391-410.
- [8] Alzahrani SI, Aljamaan IA, Al-Fakih EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. *Journal of infection and public health*. 2020 Jul 1;13(7):914-9.
- [9] Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003 Jan 1;50:159-75.
- [10] Banihabib ME, Arabi A, Salha AA. A dynamic artificial neural network for assessment of land-use change impact on warning lead-time of flood. *International Journal of Hydrology Science and Technology*. 2015;5(2):163-78.
- [11] Xu A, Chang H, Xu Y, Li R, Li X, Zhao Y. Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. *Waste Management*. 2021 Apr 1;124:385-402.
- [12] Ebrahimi H, Rajaei T. Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global and Planetary Change*. 2017 Jan 1;148:181-91..
- [13] Wang Z, Lou Y. Hydrological time series forecast model based on wavelet de-noising and ARIMA-LSTM. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) 2019 Mar 15 (pp. 1697-1701). IEEE.
- [14] Shoaib M, Shamseldin AY, Melville BW, Khan MM. Runoff forecasting using hybrid wavelet gene expression programming (WGEP) approach. *Journal of Hydrology*. 2015 Aug 1;527:326-44.
- [15] Agarwal A, Maheswaran R, Sehgal V, Khosa R, Sivakumar B, Bernhofer C. Hydrologic regionalization using wavelet-based multiscale entropy method. *Journal of Hydrology*. 2016 Jul 1;538:22-32.
- [16] Rodríguez-Murillo JC, Filella M. Significance and causality in continuous wavelet and wavelet coherence spectra applied to hydrological time series. *Hydrology*. 2020 Nov 2;7(4):82.
- [17] Sang YF. A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric research*. 2013 Mar 1;122:8-15.
- [18] Nikmanesh, Mohammad Reza. Predicting average monthly discharge using an integrated model of artificial neural network and wavelet transformations (Case study: Ker River - Pol Khan Station). *Journal of Water and Soil Conservation Research*, 2015; 22 (3): 231-239. [in Persian].
- [19] Seo Y, Kim S, Kisi O, Singh VP. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *Journal of Hydrology*. 2015 Jan 1;520:224-43.
- [20] Iwok IA, Udoh GM. A Comparative study between the ARIMA-Fourier Model and the Wavelet model.

- American Journal of Scientific and Industrial Research. 2016;7(6):137-44.
- [21] Banihabib ME, Ahmadian A, Jamali FS. Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data. *GeoResJ*. 2017 Jun 1;13:9-16
- [22] Adib, Arash, Gorjizadeh, Ali. Drought assessment and monitoring using drought indicators; Case study of Dez catchment. *Iranian Journal of Irrigation and Water Engineering*, 2016; 7 (2): 173-185. [in Persian].
- [23] Cryer JD, Chan KS. *Time series analysis: with applications in R*. New York: Springer; 2008 Apr 4.
- [24] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*. 1989 Jul;11(7):674-93.
- [25] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*. 1979 Jun 1;74(366a):427-31.
- [26] Said SE, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*. 1984 Dec 1;71(3):599-607.
- [27] Zhang X, Wu X, Zhu G, Lu X, Wang K. A seasonal ARIMA model based on the gravitational search algorithm (GSA) for runoff prediction. *Water Supply*. 2022 Aug 1;22(8):6959-77.
- [28] Daubechies I. The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*. 1990 Sep;36(5):961-1005.
- [29] Shao Y, Zhao J, Xu J, Fu A, Li M. Application of rainfall-runoff simulation based on the NARX dynamic neural network model. *Water*. 2022 Jun 29;14(13):2082.
- [30] Nury AH, Hasan K, Alam MJ. Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University-Science*. 2017 Jan 1;29(1):47-61.
- [31] Abbaszadeh H, Daneshfaraz R, Sume V, Abraham J. Experimental investigation and application of soft computing models for predicting flow energy loss in arc-shaped constrictions. *AQUA—Water Infrastructure, Ecosystems and Society*. 2024 Feb 23;jws2024010.
- [32] Hoque MJ, Bayezid M, Sharan AR, Kabir MU, Tareque T. Prediction of Strength Properties of Soft Soil Considering Simple Soil Parameters. *Open Journal of Civil Engineering*. 2023 Jul 31;13(3):479-96.
- [33] Kumar P, Foufoula-Georgiou E. A multicomponent decomposition of spatial rainfall fields: 1. Segregation of large-and small-scale features using wavelet transforms. *Water Resources Research*. 1993 Aug;29(8):2515-32.
- [34] Cheng Y, Zhang H, Liu Z, Chen L, Wang P. Hybrid algorithm for short-term forecasting of PM_{2.5} in China. *Atmospheric environment*. 2019 Mar 1; 200:264-79.