



Golestan University



Journal of the Climate Change Research

Scientific Journal of Golestan University  
Vol. 5, No. 18, Summer 2024



## Modelling daily pan evaporation of Zayanderud dam station utilizing artificial intelligence and time series models

Ahmad Reza Ghasemi<sup>1\*</sup>, Marziyeh Esmaeilpour<sup>2</sup>

<sup>1</sup>Associate Professor, Department of Water Engineering, Shahrekord University, Shahrekord, Iran,  
Email: ar-ghasemi@sku.ac.ir

<sup>2</sup>Assistant Professor, Department of Geography and Urban Planning, Maragheh University, Maragheh, Iran,  
Email: s.esmaeilpour@gmail.com

### Article Info

**Article type:**  
Research Full Paper

Article history:  
Received: 2023-12-31  
Accepted: 2024-2-26

**Keywords:**  
Evaporation  
Missing data  
Simulation  
Extreme data

### ABSTRACT

The pan evaporation is used as a practical parameter in various fields, such as estimating water loss from lakes and dams, as well as estimating the Plant water requirement, especially in areas where there is no lysimeter information. Modeling this parameter can be useful in the estimation of missing data and long-term planning of water resources and agricultural development. In this research, by using an artificial intelligence model (Gene Expression Programming, GEP) and two time series models (Fourier and ARIMA), evaporation from the pan at Zayanderud dam station, was modeled in the period from 1344 to 1396 (53 years). The time series of pan evaporation on daily scale for the months of June (Khordad), July (Tir), August (Mordad), September (Shahrivar), and October (Mehr) as input to the Fourier and ARIMA models and 4 different patterns including the use of daily evaporation data 1 month, 2 months, 3 months and 4 months ago, were used as input for gene expression programming model. The results showed that the GEP model has acceptable results only in Mehr, and for other months the results are not acceptable in terms of statistical indicators. The daily evaporation estimated error was found to be 0.38 mm in Mehr. This error was acceptable based on the coefficient of determination of 0.84, the Nash-Sutcliffe coefficient of 0.83 and the Willmott's index of agreement of 0.95. Unlike the gene expression programming model, the Fourier model provided acceptable results in all the studied months. The error values were obtained between 1.02 and 0.7 mm per day in all the studied months, which is equivalent to 5.2 to 8.8 percent. Comparing the results of the above two models with the ARIMA model showed that, the error values of the ARIMA model in all months are higher (9.4 to 19.6%) than the Fourier model, and Gene Expression Programming model. Therefore, the best model for estimating pan evaporation in October (Mehr) is the Gene Expression Programming model and in the rest of the months, the Fourier model. The evaluation of the model in estimating the extremes daily evaporation data also showed that the highest and lowest accuracy is in October (Mehr) and June (Khordad), respectively. In total, according to the statistical indices, the ability of the Fourier model to estimate the daily evaporation in the Zayanderud dam station was proved. Therefore, this model can be recommended to estimate the daily evaporation and also to reconstruct the missing data in this station.

**Cite this article:** Ghasemi, A.R., Esmaeilpour, M. (2024). Modelling daily pan evaporation of Zayanderud dam station utilizing artificial intelligence and time series models. Journal of the Climate Change Research, 5 (18), 49-66.



©The author(s)  
Doi: 10.30488/CCR.2024.433405.1193

Publisher: Golestan University



نشریه پژوهش‌های تغییرات آب و هوایی



فصلنامه علمی دانشگاه گلستان

سال پنجم / شماره مسلسل هجدهم / تابستان ۱۴۰۳ / صفحات: ۶۶-۴۹



## مدل‌سازی تبخیر روزانه در ایستگاه سد زاینده‌رود با استفاده از مدل‌های هوش مصنوعی و سری‌زمانی

احمدرضا قاسمی<sup>۱\*</sup>، مرضیه اسمعیل‌پور<sup>۲</sup>

<sup>۱</sup> دانشیار گروه مهندسی آب دانشگاه شهرکرد، شهرکرد، ایران، رایانامه: ar-ghasemi@sku.ac.ir

<sup>۲</sup> استادیار، گروه جغرافیا و برنامه ریزی شهری، دانشگاه مراغه، مراغه، ایران، رایانامه: s.esmaeilpour@gmail.com

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله کامل علمی</p> <p>تاریخ دریافت: ۱۴۰۲/۱۰/۱۰ تاریخ پذیرش: ۱۴۰۲/۱۲/۷</p> <p>واژه‌های کلیدی: تبخیر داده‌های حدی داده‌های گمشده شبیه‌سازی</p>	<p>تبخیر از تشت به‌عنوان یک پارامتر کاربردی در زمینه‌های مختلف، مانند برآورد هدر رفت آب از دریاچه‌ها و مخازن سدها و همچنین برآورد نیاز آبی گیاهان به‌ویژه در مناطقی که اطلاعات لایسیمیتری وجود ندارد، کاربرد دارد. مدل‌سازی این پارامتر می‌تواند در زمینه بازسازی داده‌های گم‌شده و برنامه‌ریزی‌های درازمدت منابع آب و توسعه کشاورزی کارساز باشد. در این پژوهش با به کارگیری یک مدل هوش مصنوعی (برنامه‌ریزی بیان‌ژن) و دو مدل سری‌زمانی (فوریه و آریما)، تبخیر از تشت در ایستگاه سد زاینده‌رود در دوره زمانی ۱۳۴۴ تا ۱۳۹۶ (۵۳ سال) مدل‌سازی شد. سری‌زمانی داده‌های تبخیر از تشت در مقیاس روزانه برای ماه‌های گرم سال (خرداد، تیر، مرداد، شهریور و مهر)، به عنوان ورودی مدل‌های فوریه و آریما و ۴ الگوی مختلف شامل استفاده از داده‌های روزانه تبخیر ۱ ماه قبل، ۲ ماه قبل، ۳ ماه قبل و ۴ ماه قبل، به‌عنوان ورودی مدل برنامه‌ریزی بیان‌ژن استفاده شد. نتایج نشان داد که مدل برنامه‌ریزی بیان‌ژن تنها در ماه مهر نتایج قابل قبولی دارد و برای ماه‌های دیگر نتایج از نظر شاخص‌های آماری قابل قبول نمی‌باشد. میزان خطای برآورد تبخیر روزانه در ماه مهر ۰/۳۸ میلی‌متر بر روز (معادل ۲/۷ درصد) بدست آمد. این میزان خطا بر اساس ضریب تبیین ۰/۸۴ و ضریب نش- ساتکلیف (ضریب کارایی مدل) ۰/۸۳، قابل قبول ارزیابی شد. بر خلاف مدل برنامه‌نویسی بیان‌ژن، مدل فوریه در تمام ماه‌های مورد مطالعه نتایج قابل قبول ارائه داد. مقادیر خطای برآورد تبخیر روزانه در این روش بین ۱/۰۲ تا ۰/۷ میلی‌متر بر روز به‌دست آمد که معادل ۵/۲ تا ۸/۸ درصد است. مقایسه نتایج دو مدل فوق با نتایج مدل آریما نیز نشان داد مقادیر خطای مدل آریما در تمام ماه‌ها بیشتر (۹/۴ تا ۱۹/۶ درصد) از مدل‌های فوریه و برنامه‌ریزی بیان‌ژن است. بنابراین بهترین مدل برای برآورد تبخیر روزانه از تشت، در ماه مهر مدل برنامه‌ریزی بیان‌ژن و در بقیه ماه‌ها مدل فوریه می‌باشد. ارزیابی دقت و توانایی برآورد داده‌های حدی تبخیر روزانه نیز نشان داد، مدل فوریه در تخمین داده‌های حدی، دارای توانایی بالاتری نسبت به دو مدل دیگر است. بنابراین می‌توان این مدل را جهت برآورد تبخیر روزانه در ایستگاه سد زاینده‌رود و همچنین بازسازی داده‌های گم‌شده توصیه نمود.</p>

**استناد:** قاسمی، احمدرضا؛ اسمعیل‌پور، مرضیه. (۱۴۰۳). مدل‌سازی تبخیر روزانه در ایستگاه سد زاینده‌رود با استفاده از مدل‌های هوش مصنوعی و سری‌زمانی. نشریه پژوهش‌های تغییرات آب و هوایی، ۵ (۱۸)، ۶۶-۴۹.

Doi: 10.30488/CCR.2024.433405.1193

ناشر: دانشگاه گلستان

© نویسندگان.



## مقدمه

تبخیر از سطح آب یکی مهمترین عوامل هدر رفت آب از سطح دریاچه‌ها، سدها و مخازن آب است که هر ساله باعث از دست رفتن بخشی از هزینه‌های سنگین سدسازی و ذخیره آب در دنیا می‌شود. با تغییر الگوهای آب و هوایی در سال‌های اخیر، بحران آب به موضوعی بسیار مهم تبدیل شده که می‌تواند منجر به مشکلات اجتماعی، اقتصادی و سیاسی در جوامع گردد. علاوه بر این کاهش بارش منجر به افزایش رقابت برای منابع آب به‌ویژه برای کشاورزانی که به کشاورزی دیم وابسته هستند، خواهد شد (لی و همکاران، ۲۰۱۷: ۲۲۴). مدیریت مناسب منابع آب در بخش‌های ذخیره‌سازی، توزیع و مصرف، ضمن افزایش کارایی مصرف آب، موجب کاهش خسارات ناشی از تنش خشکی می‌شود (مرتضوی و همکاران، ۱۳۹۴: ۱۱۸).

تبخیر از تشت یکی از مؤلفه‌های تأثیرگذار در چرخه هیدرولوژی می‌باشد و برآورد آن در مقیاس‌های زمانی مختلف، به‌عنوان یکی از مهمترین پارامترهای جوی، از اهمیت ویژه‌ای در برنامه‌ریزی آبیاری، مدیریت آب در مزرعه و صرف‌جویی آب در کشاورزی برخوردار است (قهرمان و قره‌خانی، ۱۳۹۰: ۷۵). یکی از مسائل مهم در بحث تبخیر از سطح آزاد آب، ثبت مقدار تبخیر است (نجفوند و اسلامی، ۱۳۹۵: ۶۵). تشت تبخیر به علت سهولت نصب، اندازه‌گیری و تفسیر داده‌ها، در سراسر دنیا به‌عنوان روشی استاندارد برای تعیین تبخیر از دریاچه‌ها و مخازن استفاده می‌شود (گوپل، ۲۰۰۹). بالا بودن هزینه تعمیر و نگهداری و عدم امکان نصب آن‌ها در همه مناطق و مهمتر از همه برآوردهای کمتر از واقعیت تشت (شودهر، ۲۰۰۰)، باعث گسترش استفاده از روش‌های ریاضی و مدل‌های کامپیوتری جهت برآورد و تخمین تبخیر از سطح آب شده است. روش‌های هوش مصنوعی به‌دلیل دقت بالا در تخمین و مدل‌سازی داده‌های مختلف، در سال‌های اخیر به‌شدت مورد توجه پژوهشگران و محققان در زمینه‌های مختلف، از جمله موضوعات آب و هوایی قرار گرفته‌اند.

مدل‌سازی برنامه‌نویسی بیان‌ژن (GEP) یکی از این روش‌ها محسوب می‌شود.

مدل برنامه‌نویسی بیان ژن یکی از روش‌های محاسبات نرم محسوب می‌شود که به‌منظور حل مسائل مختلف از اصول انتخاب طبیعی داروین استفاده می‌کند. کوزا (۱۹۹۳: ۲۹۷) باتوسعه الگوریتم ژنتیک، مدل برنامه‌نویسی بیان‌ژن را ارائه داد، بنابراین می‌توان گفت که برنامه‌نویسی بیان‌ژن یک گونه خطی از برنامه‌نویسی ژنتیک است. برنامه‌نویسی بیان ژن شامل بخش‌های مجموعه توابع آماری و ریاضی، برازش تابع مناسب و پارامترهای کنترل و ارزیابی است (گندمی و همکاران، ۲۰۱۱: ۲۴۸). علاوه بر استفاده از روش‌های هوش مصنوعی در مدل‌سازی داده‌های اقلیمی، مدل‌های تحلیل یا تبدیل سری‌های زمانی نیز بصورت گسترده‌ای توسط محققان استفاده می‌شوند. مدل‌سازی سری زمانی به دلیل توانایی در تحلیل داده‌ها از کاربرد گسترده‌ای در پژوهش‌های هیدرولوژی و منابع آب برخوردار است. از آنجایی که متغیرهای تعیین‌کننده بیشتر سری‌های زمانی از عوامل طبیعی تأثیر می‌پذیرند، این سری‌ها عموماً دارای مؤلفه تناوبی هستند. برای نمونه می‌توان از سری‌های زمانی بارندگی، دما، دبی رودخانه‌ها و چشمه‌ها نام برد (فولادمنند، ۱۳۸۹: ۱۵۷). استفاده از مدل‌های سری زمانی در بیان خصوصیات پدیده تبخیر می‌تواند بسیار مفید واقع شود، چرا که مدل‌های جبری قادر به بیان ماهیت واقعی متغیرهای هیدرولوژیکی که ماهیت تصادفی نیز در آن‌ها مستتر است، نمی‌باشد. بنابراین لازم است ساختار مدل‌های سری زمانی با ساختار سری‌های هیدرولوژیکی سازگاری داشته باشد (سالاس، ۱۹۹۳).

مطالعه روش‌های برآورد و پیش‌بینی تبخیر از سطح آب می‌تواند در مدیریت منابع آب موجود و همچنین برنامه‌ریزی‌های آینده کمک کند. در طرح‌های توسعه کشاورزی، صنعت و ... باید موضوع تبخیر از سطح آب مخازن سطحی و میزان هدر رفت سالیانه آن مد نظر قرار گیرد. پژوهش‌های فراوانی در

عملکرد مدل مربوط به برآورد تبخیر در ماه اکتبر است (باقری و همکاران، ۱۴۰۲: ۱).

بذرافشان و همکاران (۱۳۹۶) نیز با استفاده از مدل ARIMA خطایی بین ۰/۶ تا ۰/۹ میلی‌متر بر روز را برای پیش‌بینی تبخیر ماهیانه از تشت در استان هرمزگان بدست آوردند. برخی پژوهش‌های جدید مانند نوذری و همکاران (۱۴۰۲) با مقایسه مدل‌های ARIMA، SARIMA، رگرسیون خطی و برنامه‌ریزی بیان‌ژن نشان دادند که مقدار خطای برآورد تبخیر برای این مدل‌ها به ترتیب برابر ۰/۹۱، ۰/۷۸، ۰/۷۲ و ۰/۸۷ میلی‌متر در روز است. پژوهشگران زیادی از مدل برنامه‌نویسی بیان‌ژن برای مطالعه فرآیندهای مختلف از جمله تخمین داده‌های گمشده (استوریکار و دئو، ۲۰۰۸: ۱۷۷)، برآورد دبی روزانه و پیش‌بینی جریان (شعیب و همکاران، ۲۰۱۵: ۳۲۶) و تخمین تبخیر و تعرق (یاسین و همکاران، ۲۰۱۶: ۱۱۰؛ سیاسر و همکاران، ۱۴۰۰: ۸۵؛ سبزواری و سعیدی‌نیا، ۱۴۰۰: ۲۳۷) سود جست‌ه‌اند.

همچنین مدل‌های سری‌زمانی مانند فوریه در پیش‌بینی بارندگی (تولارام و الهی، ۲۰۱۰: ۳۷۲)، مدل‌سازی دما (لاگواردیا، ۲۰۱۱: ۱۳۹۹)، مطالعه رژیم بارش (گاور و همکاران، ۲۰۲۳: ۴) و پیش‌بینی بارندگی (هایل و همکاران، ۲۰۲۰: ۲؛ مازیلین و سوهایلا، ۲۰۲۳: ۱) مورد استفاده قرار گرفته‌اند. استفاده فراوان از این مدل‌ها نشان می‌دهد که می‌دهند که مدل‌های مذکور توانایی لازم جهت شبیه‌سازی متغیرهای اقلیمی را دارا می‌باشند. از طرف دیگر با توجه به اینکه تغییرات متغیرهای اقلیمی بشدت تابع موقعیت جغرافیایی بوده و در مناطق مختلف متفاوت می‌باشند، نمی‌توان یک مدل را برای همه مناطق یا اقلیم‌ها توصیه نمود، و یا نتایج حاصل از یک مدل در یک منطقه را به مناطق دیگر تعمیم داد. بنابراین لازم است توانایی مدل‌های مختلف در مناطق گوناگون مورد ارزیابی قرار گیرد.

تبخیر از تشت یکی از متغیرهای اساسی در حوزه‌های مختلف کشاورزی و مدیریت منابع آبی بوده و مدل‌سازی آن در تعیین برنامه‌های مدیریتی این

ارتباط با برآورد و پیش‌بینی پارامترهای اقلیمی به کمک مدل‌های برنامه‌ریزی بیان‌ژن، فوریه و آریمای انجام گرفته است که در ادامه به برخی از مطالعات مرتبط با پژوهش حاضر اشاره می‌شود.

حقیقت‌جو و همکاران (۱۳۹۶: ۱۰۷) از مدل برنامه‌نویسی بیان‌ژن برای تخمین تبخیر روزانه در خراسان جنوبی استفاده و نشان دادند که مدل برنامه‌نویسی بیان‌ژن از دقت قابل قبولی در تخمین تبخیر روزانه از تشت برخوردار است. احمدپور و همکاران (۱۳۹۷: ۴۹) با به‌کارگیری مدل خطی SARIMA و مدل دوخطی (BiLinear)، تبخیر ماهانه در ایستگاه ایدنک واقع در حوضه مارون را برآورد و نشان دادند، مدل دو خطی از دقت بیشتری نسبت به مدل SARIMA در برآورد تبخیر برخوردار است.

میراندا و همکاران (۲۰۱۹: ۲) از مدل فوریه برای پیش‌بینی دما و رطوبت در مقیاس ساعتی استفاده کردند و نشان دادند این روش به‌خوبی می‌تواند ویژگی‌های اصلی متغیرهای دما و رطوبت را مدل‌سازی کند. مازیلین و سوهایلا (۲۰۲۳) نیز در پژوهشی به بررسی تغییرات و الگوهای بارندگی در سراسر شبه جزیره مالزی، شامل ۱۶ ایستگاه باران‌سنجی از سال ۱۹۹۹ تا ۲۰۱۹، با استفاده از روش تحلیل فوریه پرداختند. آنها نشان دادند الگوهای بارش در شبه جزیره مالزی را می‌توان به اندازه کافی توسط دو جزء اصلی اول داده‌ها توصیف کرد.

عمادی و همکاران (۱۴۰۱: ۳۵۴) در دو حوضه آبریز دریاچه ارومیه و گاوخونی تبخیر ماهانه را با استفاده از روش‌های شبکه عصبی مصنوعی، استنتاج فازی و برنامه‌نویسی بیان‌ژن مدل‌سازی کردند. نتایج آنها نشان داد که مدل برنامه‌نویسی بیان‌ژن در هر دو حوضه بهترین عملکرد را در میان سایر مدل‌ها داشته است. خانال و همکاران (۲۰۲۰) نیز از سه الگوریتم هموارسازی زمانی برای بهبود طبقه‌بندی شاخص NDVI در نیال استفاده و بهترین نتایج را برای مدل هموارسازی فوریه گزارش نمودند. در پژوهشی تبخیر از تشت در استان خوزستان با استفاده از روش برنامه‌نویسی بیان‌ژن مدل‌سازی و نشان داد شد که بهترین

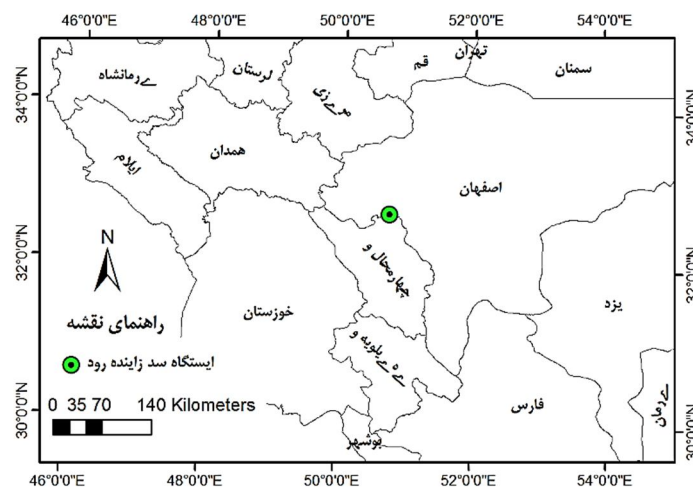
مدل‌ها در تخمین مقادیر حدی تبخیر روزانه نیز مورد ارزیابی قرار گرفت.

### مواد و روش‌ها

**منطقه مورد مطالعه:** تحقیق حاضر بر روی ایستگاه سد زاینده‌رود در استان اصفهان انجام شد. عرض جغرافیایی این ایستگاه ۳۲ درجه و ۴۳ دقیقه و ۴۸ ثانیه شمالی و طول جغرافیایی آن ۵۰ درجه و ۴۴ دقیقه و ۳۱ ثانیه شرقی می‌باشد. ارتفاع این ایستگاه از سطح دریا نیز ۲۱۷۳ متر می‌باشد.

**داده‌ها و روش تحلیل:** دریاچه سد زاینده رود با حجم ۱۴۷۰ میلیون متر مکعب و حجم مفید ۱۰۹۰ میلیون متر مکعب و مساحت ۵۴ کیلومتر مربع در جنوب غربی استان اصفهان و در طرفین مرز استان‌های اصفهان و چهارمحال و بختیاری در شرق رشته کوه زاگرس واقع شده است (شکل ۱). این سد با هدف تنظیم آب کشاورزی مورد نیاز ۱۰۰ هزار هکتار از اراضی کشاورزی پایین دست طراحی شده است. تخصیص آب این سد برای مصارف گوناگون در بخش وسیعی از مناطق خشک و نیمه خشک کشور، لزوم برنامه‌ریزی پایدار را بیش از پیش نشان می‌دهد. بنابراین، مدل‌سازی و شناخت الگوی تغییرات تبخیر به عنوان یکی از اصلی‌ترین عوامل هدررفت آب، می‌تواند در استفاده بهینه از آب این سد مفید باشد.

بخش‌ها، جزو ملزومات می‌باشد. علاوه بر این، وجود داده‌های گمشده در اطلاعات ایستگاه‌های هواشناسی نیز استفاده از این داده‌ها را با محدودیت روبرو می‌کند. در صورتی که بتوان مدل توانمندی برای برآورد مقادیر تبخیر از تشتک به دست آورد، می‌توان آن را جایگزین روش‌های مرسوم برای تخمین مقادیر گم شده نمود. سد زاینده رود بزرگترین سد تامین کننده آب در بخش مرکزی کشور است که آب شرب، صنعت، کشاورزی و محیط زیست منطقه‌ای وسیعی در استان‌های چهارمحال و بختیاری و اصفهان را تامین می‌کند. علاوه بر این با اجرای طرح انتقال آب به یزد بخشی از آب شرب این منطقه را نیز تامین می‌کند. با توجه به اهمیت حیاتی این سد در منطقه خشک و نیمه‌خشک مرکزی کشور، در این پژوهش میزان تبخیر از سطح آب در ایستگاه سد زاینده رود به کمک یک روش مبتنی بر هوش مصنوعی (برنامه‌نویسی بیان‌زن) و دو روش مبتنی بر سری‌زمانی (مدل تبدیل فوریه و مدل ARIMA) مورد مدل‌سازی قرار گرفت. مدل‌هایی که جهت مدل‌سازی داده‌ها از هوش انسان تقلید می‌کنند را اصطلاحاً مدل هوش مصنوعی گویند. هوش مصنوعی نوعی فناوری است که می‌تواند از تجربه خود بیاموزد، با ورودی‌های جدید سازگار شود. به عبارت دیگر در این تحقیق توانایی دو نوع روش مدل‌سازی متفاوت در مدل‌سازی تبخیر از تشت مورد مقایسه قرار گرفت. همچنین در این پژوهش توانایی



شکل ۱- موقعیت جغرافیایی ایستگاه سد زاینده رود

$$\hat{y}(t) = d_0 + d_1 \sin(\omega t) + d_2 \cos(\omega t) + d_3 \sin(2\omega t) + d_4 \cos(2\omega t) + \dots \quad (2)$$

این تابع بر اساس شرایط زیر تعریف می‌شود.

$$\beta_0(t) = 1, \beta_{2r-1}(t) = \sin(r\omega t), \beta_{2r}(t) = \cos(r\omega t) \quad (3)$$

که در آن،  $\omega$  فرکانس اساسی در دوره زمانی  $T$  است و از رابطه (۴) بدست می‌آید (مازیلن و سوهایلا، ۲۰۲۳).

$$\omega = \frac{2\pi}{T} \quad (4)$$

با به حداقل رساندن مجموع مجذور باقیمانده‌های مدل، ضرایب  $d_k$  با استفاده از روش حداقل مربعات محاسبه و مطابق رابطه (۵) محاسبه می‌شود.

$$SSE_i = \sum_{j=1}^T (x_j - y(t_j))^2, i = 1, 2, \dots, n \quad (5)$$

پارامتر هموارسازی  $\lambda$  بر اساس حداقل ارزیابی متقابل تعمیم یافته (GCV) انتخاب خواهد شد.

$$GCV(\lambda) = \left( \frac{n}{n - df(\lambda)} \right) \left( \frac{SSE}{n - df(\lambda)} \right) \quad (6)$$

که در آن  $df(\lambda)$  درجات آزادی عملگر هموارسازی است (مازیلن و سوهایلا، ۲۰۲۳).

ساختار ساده‌تر تبدیل داده‌ها به یک الگوی سینوسی را می‌توان مطابق زیر نمایش داد.

$$y_t = \bar{y} + C_1 \cos\left(\frac{2\pi t}{n} - \phi_1\right) \quad (7)$$

که در آن  $y_t$  مقدار متغیر مورد بررسی که می‌تواند مقدار تبخیر از تشت باشد.  $\bar{y}$  مقدار میانگین متغیر مورد مطالعه و  $C_1$ ، مقدار دامنه (Amplitude)،  $\phi_1$ ، زاویه فاز یا تغییر فاز (phase angle, or phase shift)،  $n$  تعداد داده‌ها و  $t$  زمان است. می‌توان معادله  $y_t$  را با استفاده از رابطه ۸ بصورت یک الگوی سینوس بازنویسی کرد.

$$\cos\left(\alpha - \frac{\pi}{2}\right) = \sin(\alpha) \quad (8)$$

توضیحات بیشتر در پژوهش‌هایی مانند تولارام و الهی

داده‌های مورد استفاده در این پژوهش، شامل داده‌های تبخیر از تشتک مربوط به این ایستگاه سد زاینده رود در یک دوره زمانی ۵۳ ساله از سال ۱۳۴۴ تا ۱۳۹۶ برای ۵ ماه فصل گرم (خرداد، تیر، مرداد، شهریور و مهر) می‌باشد که از شرکت مدیریت منابع آب ایران وابسته به وزارت نیرو تهیه گردید. داده‌های تبخیر از تشت در مقیاس روزانه و با واحد میلی‌متر برداشت شده‌اند. داده‌های ماه‌های دیگر سال دارای تعداد زیادی داده گمشده بودند که به دلیل اینکه هدف این پژوهش، مدل‌سازی روزانه داده‌ها می‌باشد، امکان بازسازی تعداد زیادی داده وجود نداشت. برای تحلیل و مدل‌سازی، داده‌ها به دو بخش آموزش (۶۰ درصد داده‌ها) و آزمون (۴۰ درصد داده‌ها) تقسیم شدند. برای آموزش و آزمون مدل ممکن از نسبت‌های مختلف داده‌ها استفاده شود که تابع تعداد داده‌های مورد استفاده است. در این تحقیق برای اینکه در بخش آزمون نیز تعداد داده کافی در اختیار باشد از نسبت ۶۰ و ۴۰ استفاده شد. در نهایت جهت مدل‌سازی داده‌های تبخیر از سه روش شامل، مدل برنامه‌ریزی بیان‌ژن، مدل‌سازی فوریه و مدل ARIMA استفاده شد.

**مدل‌سازی فوریه:** مدل هموارسازی فوریه در اصل تبدیل یک سری زمانی به مختصات فوریه آن است که بخشی از فرکانس‌های بالاتر را حذف می‌کند و سپس مختصات را به سیگنال تبدیل می‌کند. این سیگنال جدید اصطلاحاً یک سری هموار نامیده می‌شود. تابع تناوبی به عنوان مبنای فوریه نمایش داده می‌شود که می‌تواند به صورت رابطه (۱) نوشته شود.

$$y(t) = \sum_{K=1}^K d_K \beta_K(t) \quad (1)$$

که در آن،  $y(t)$  مقدار پارامتر مورد بررسی که در اینجا تبخیر از تشت است،  $d_k$  ضریب پایه،  $\beta_k$  تابع پایه و  $K$  نشان دهنده حداکثر اندازه توابع پایه است. توابع پایه تابع تناوبی که به عنوان مبنای فوریه نمایش داده می‌شود را می‌تواند به صورت رابطه (۲) نوشت.

$$Z_t = \sum_{i=1}^P (\phi_i Z_{t-i}) - \sum_{j=1}^q (\theta_j \varepsilon_{t-j}) + \varepsilon_t \quad (11)$$

مدل‌سازی برنامه‌نویسی بیان ژن (GEP):

برنامه‌ریزی بیان ژن، توسعه یافته روش‌های برنامه‌ریزی ژنتیک و الگوریتم ژنتیک است که در سال ۱۹۹۹ توسط فریرا بر اساس تئوری داروین ابداع شد (سبزواری و سعیدی‌نیا، ۱۴۰۰: ۲۴۲). برنامه‌نویسی بیان ژن یکی از روش‌های محاسبات نرم و هوش مصنوعی به‌منظور حل مسائل مختلف است. این مدل دارای ساختارهای درختی پیچیده‌ای است که با تغییر اندازه، شکل و ترکیب آنها، مانند یک ارگانسیم زنده، سازگار می‌شوند. برنامه‌نویسی بیان ژن یکی از اساسی‌ترین گام‌ها در زمینه بهینه‌سازی است که با مدل‌سازی فرآیندهای پیچیده طبیعی و مصنوعی با داشتن تعداد معینی بر داشت میدانی یا نقاط آزمون، مسئله مورد نظر را حل می‌کند. توضیحات تکمیلی در صادقی‌اقدم و همکاران (۱۳۹۶: ۹۰)؛ فرزادمهر و همکاران (۱۳۹۹: ۱۵۷)؛ سبزواری و سعیدی‌نیا (۱۴۰۰: ۲۴۲) ارائه شده است. در پژوهش حاضر برای مدل‌سازی تبخیر روزانه از تشت با استفاده از برنامه‌نویسی بیان ژن، ۴ الگو جهت بررسی کارایی این مدل در برآورد مقدار تبخیر از تشت، مورد بررسی قرار گرفت. این الگوها شامل استفاده از داده‌های روزانه تبخیر در ۱ ماه قبل، ۲ ماه قبل، ۳ ماه قبل و ۴ ماه قبل برای برآورد و پیش‌بینی تبخیر روزانه می‌باشند. برای برآورد یا شبیه‌سازی یک پارامتر از دو روش استفاده می‌شود. روش اول استفاده از پارامترهای موثر (مانند دما، رطوبت و ...) بر پارامتر هدف (تبخیر از تشت) است و روش دوم که بر اساس ساختارهای سری‌های زمانی استوار است، از داده‌های همان پارامتر در زمان‌های گذشته استفاده می‌کنند (برای نمونه از داده‌های تبخیر روزانه در طول ۴ ماه قبل (حدود ۱۲۰ داده) برای برآورد تبخیر در ماه حاضر). از هر دو روش در مطالعات متعددی استفاده

(۲۰۱۰)، لاگواردیا (۲۰۱۱)، میراندا و همکاران (۲۰۱۹) و مازیلین و سوهایلا (۲۰۲۳) ارائه شده است.

مدل ARIMA: مدل خودهمبستگی متوالی متغیرهای تصادفی (ARIMA<sup>۱</sup>) یکی از پرکاربردین مدل‌هایی است که جهت مدل‌سازی سری‌های زمانی مختلف مورد استفاده قرار می‌گیرد. این مدل با استفاده از ترکیبی از مدل‌های خودهمبستگی، میانگین متحرک و انتگرال، برای پیش‌بینی سری‌های زمانی به کار می‌رود. در این مدل با استفاده از توابع خود همبستگی و خودهمبستگی جزئی، مدل‌های سری‌زمانی مناسب به داده‌ها برازش داده می‌شود و با استفاده از رفتار این دو تابع، خواص سری‌های زمانی مانند ایستایی داده‌ها مورد بررسی قرار می‌گیرد (سلطانی و همکاران، ۱۳۹۶). این مدل در پژوهش‌های فراوانی مورد استفاده قرار گرفته‌اند (تیموی و تیموری، ۱۴۰۱؛ وو و همکاران، ۲۰۲۱؛ الارکون و همکاران، ۲۰۲۲: ۷).

شکل کلی مدل خود همبستگی AR به صورت رابطه ۹ می‌باشد (آزاد و همکاران، ۱۳۹۲).

$$Z_t = \sum_{i=1}^P (\phi_i Z_{t-i}) + \varepsilon_t \quad (9)$$

که در این رابطه  $p$  مرتبه مدل و  $\phi_i$  ضرایب مدل می‌باشند.  $\varepsilon$  نیز سری تصادفی نرمال و استاندارد می‌باشد.

از ترکیب مدل همبستگی (AR) و یک مدل میانگین متحرک (MA)، مدل ARMA(p,q) ایجاد می‌گردد. شکل عمومی مدل MA به صورت رابطه ۱۰ می‌باشد.

$$Z_t = \sum_{j=1}^q (\theta_j \varepsilon_{t-j}) + \varepsilon_t \quad (10)$$

که در این رابطه  $q$  مرتبه مدل و  $\theta_j$  ضرایب مدل می‌باشند. در نهایت مدل میانگین متحرک خودهمبسته ARMA(p,q) را می‌توان به صورت زیر نوشت.

<sup>1</sup> Autoregressive Integrated Moving Average

که در آنها،  $y_i$  متغیر مشاهداتی،  $\bar{y}$  میانگین متغیر،  $\hat{y}_i$  متغیر محاسبه شده توسط مدل و  $N$  تعداد داده‌ها است.

جدول ۱- وضعیت مدل بر اساس شاخص نش - ساتکلیف (N.S) (موریاسی و همکاران، ۲۰۰۷)

وضعیت مدل	نش-ساتکلیف
بسیار خوب	$1. \leq 0.0.75 < NSE$
خوب	$0. \leq 750.65 < NSE$
رضایت‌بخش	$0. \leq 650.50 < NSE$
غیر قابل قبول	$0. \leq 50 NSE$

### نتایج و بحث

**مدل‌سازی تبخیر روزانه با مدل برنامه‌ریزی بیان**  
 ژن: جدول ۲ آماره‌های محاسبه شده جهت ارزیابی نتایج مدل برنامه‌نویسی بیان ژن را در هر دو مرحله آزمون (با استفاده از ۶۰ درصد داده‌ها) و آموزش (با استفاده از ۴۰ درصد داده‌ها) نشان می‌دهد. در این جدول نتایج برآورد مقادیر تبخیر روزانه با استفاده از داده‌های ماه‌های قبل با تأخیرهای زمانی مختلف ارائه شده است. نتایج ارزیابی برآورد تبخیر روزانه در ماه مهر نشان می‌دهد کمترین میزان خطا به همراه مقادیر ضریب تبیین ( $R^2$ ) و ضریب کارایی مدل یا ضریب نش (N.S) معنی‌دار در مرحله آموزش مدل مربوط به برآورد تبخیر با استفاده از مقادیر ۴ ماه قبل (خرداد، تیر، مرداد و شهریور) می‌باشند. مقدار کارایی مدل در این الگو ۰/۸۹ و میزان خطای تخمین تبخیر روزانه تنها ۰/۲۰ میلی‌متر بدست آمده که بر اساس طبقه‌بندی (موریاسی و همکاران، ۲۰۰۷) که در جدول ۱ ارائه شده است، نشان دهنده عملکرد بسیار خوب مدل برنامه‌ریزی بیان ژن در برآورد تبخیر روزانه این ماه است. این الگو بهترین نتیجه را در بین تمام ماه‌ها و الگوهای مورد بررسی ارائه داده است. شکل ۲ مقادیر مشاهداتی تبخیر روزانه و مقادیر برآورد شده توسط مدل در ماه مهر را بر اساس الگوی استفاده از داده‌های روزانه ۴ ماه قبل نشان می‌دهد. مقادیر برآورد شده توسط مدل، بوسیله میانگین درازمدت تبخیر در هر روز از ماه در طول دوره آماری شبیه‌سازی شده است.

شده است. در این پژوهش از روش دوم جهت برآورد تبخیر از تشت استفاده شده است.

**معیارهای ارزیابی خطا:** در انتها نتایج مدل‌ها به کمک ۵ شاخص آماری جذر میانگین مربعات خطا (RMSE)، شاخص نش-ساتکلیف (N.S)، میانگین درصد مطلق خطای (MAPE)، ضریب تبیین ( $R^2$ ) و شاخص جرم باقیمانده (CRM) مورد ارزیابی قرار گرفت (معادله‌های ۷ تا ۱۱). ضریب تبیین و ضریب توافق ویلموت میزان انطباق نتایج مدل را با داده‌های واقعی ارزیابی می‌کنند و هرچه به ۱ نزدیکتر باشند نشان دهنده عملکرد بهتر مدل است. جذر میانگین مربعات خطا میزان خطای برآورد تبخیر توسط مدل را برحسب میلی‌متر بیان می‌کند. شاخص جرم مضاعف نیز بیان کنند بیش‌برآوردی و یا کم‌برآوردی مدل است و شاخص نش-ساتکلیف (N.S) نیز شاخص عملکرد مدل است که میزان موفقیت مدل در شبیه‌سازی داده‌ها را بیان می‌کند. حداکثر مقدار این شاخص ۱ و نشان‌دهنده موفقیت کامل مدل در شبیه‌سازی داده‌ها می‌باشد. مقادیر حاصل از این شاخص بر اساس جدول (۱) ارزیابی می‌شود (موریاسی و همکاران، ۲۰۰۷).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{N}} \quad (7)$$

$$NS = 1 - \frac{\sum_{i=1}^n (y - \hat{y})_i^2}{\sum_{i=1}^n (y - \bar{y})_i^2} \quad (8)$$

$$MAPE = \frac{1}{n} \sum \left| \frac{\hat{y} - y}{y} \right| \times 100 \quad (9)$$

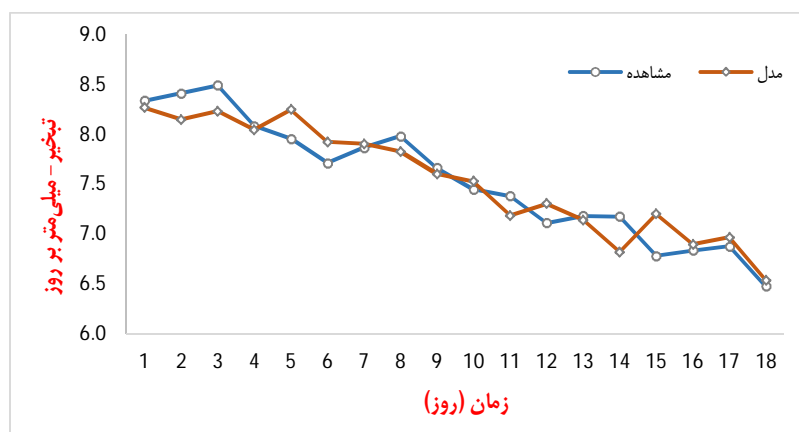
$$R^2 = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \bar{y})^2} \quad (10)$$

$$CRM = \frac{\sum y_i - \sum \hat{y}_i}{\sum y_i} \quad (11)$$

درصد و برای ماه تیر استفاده از داده‌های ماه قبل با خطای کمتر از یک درصد بدست آمد (جدول ۲). هر چند کمترین مقدار خطای برآورد مدل در مرحله آموزش در بین تمام الگوهای مورد بررسی در مرداد ماه به میزان ۰/۱۱ بدست آمده است، ولی ضریب تبیین و کارایی مدل در این حالت ۰/۶۳ است که نشان می‌دهد علی‌رغم اینکه میانگین خطاها در مدل پایین است ولی مدل در برآورد برخی داده‌ها دچار خطای زیادی شده است. میزان شاخص کارایی مدل (N.S=۰/۶۳) نیز نشان دهنده کارایی پایین مدل است. وضعیت مدل در این حالت بر اساس جدول ۱ (موریاسی و همکاران، ۲۰۰۷)، رضایت‌بخش می‌باشد. مقادیر شاخص CRM نیز در تمام الگوهای مرحله آموزش در ماه‌های مورد بررسی مانند ماه مهر صفر به‌دست آمده است که نشان دهنده عدم بیش‌برآوردی و یا کم‌برآوردی مدل است.

مقدار خطای برآورد تبخیر روزانه در این الگو نیز حدود ۲ درصد است. بعد از این الگو، استفاده از داده‌های ۳ ماه قبل (تیر، مرداد و شهریور) برای تخمین تبخیر روزانه ماه مهر با خطای ۰/۲۱ میلی‌متر (حدود ۲ درصد) و ضریب کارایی ۰/۸۸ بهترین نتیجه را نشان می‌دهد. در هر دو الگو میزان شاخص CRM صفر به‌دست آمده است که نشان می‌دهد مدل در تخمین تبخیر این ماه دچار بیش‌برآوردی و یا کم‌برآوردی نشده است.

نتایج مرحله آموزش در ماه‌های دیگر ضعیف‌تر از ماه مهر ولی از نظر شاخص‌های آماری مورد بررسی حداقل یک الگو با نتایج قابل قبول برای برآورد تبخیر روزانه را دارا می‌باشند. بهترین نتیجه برای ماه شهریور در الگوی استفاده از داده‌های ۳ ماه قبل، برای ماه مرداد استفاده از داده‌های ۲ ماه قبل با خطای ۱/۳



شکل ۲- مقادیر مشاهداتی تبخیر روزانه و مقادیر برآورد شده توسط مدل در ماه مهر در مرحله آموزش

همکاران، ۱۳۹۶). نتایج بدست آمده از بخش آزمون ملاک ارزیابی اصلی مدل و نشان دهنده توانایی یک مدل در برآورد و یا تخمین یک پارامتر است. نتایج بخش آزمون در ماه مهر (جدول ۲) نشان می‌دهد که تنها الگوی استفاده از داده‌های دو ماه قبل برای تخمین تبخیر نتایج قابل قبولی ارائه داده است. مقدار خطای مدل در این حالت ۰/۳۸ میلی‌متر بدست آمده است که معادل ۲/۷ درصد است. مقدار ضریب تبیین و ضریب نش در این الگو نیز به ترتیب ۰/۸۴ و ۰/۸۳

در مجموع مدل برنامه‌ریزی بیان‌ژن در مرحله آموزش نتایج قابل قبولی در تخمین تبخیر روزانه نشان داد ولی نتایج بخش آزمون ضعیف‌تر و برخی موارد غیر قابل قبول می‌باشند. زیرا بخش عمده‌ای از داده‌ها برای آموزش مدل‌ها استفاده می‌شود. به‌طور کلی تضعیف نتایج در بخش آزمون مدل‌ها قابل انتظار و تقریباً در تمامی پژوهش‌های مشابه مشاهده می‌گردد (مانند باقری و همکاران، ۱۴۰۲؛ امجد و همکاران، ۱۳۹۸؛ طاووسی و زهرایی، ۱۳۹۴؛ فیض‌اله‌پور و

است که نشان می‌دهد مدل ۸۴ درصد تغییرات داده‌های واقعی را مدل‌سازی کرده است.

شکل ۳ تغییرات مقادیر واقعی تبخیر روزانه ماه مهر را به همراه مقادیر برآورد شده توسط مدل را برای این الگو به همراه مقادیر خطا نشان می‌دهد. همانگونه که در شکل ۳ نشان داده شده است مقدار خطای برآورد تبخیر در روزهای مختلف ماه در این الگو در هر دو مرحله مدل‌سازی کمتر از ۱ میلی‌متر است. میانگین درصد خطای مطلق (MAPE) مدل در این ماه نیز در مرحله آموزش ۰/۲۹ میلی‌متر در روز (معادل ۳/۲ درصد) و در مرحله آزمون ۰/۳۸ میلی‌متر در روز و معادل ۲/۷ درصد بدست آمده است. مقایسه نتایج به دست آمده در پژوهش حاضر با

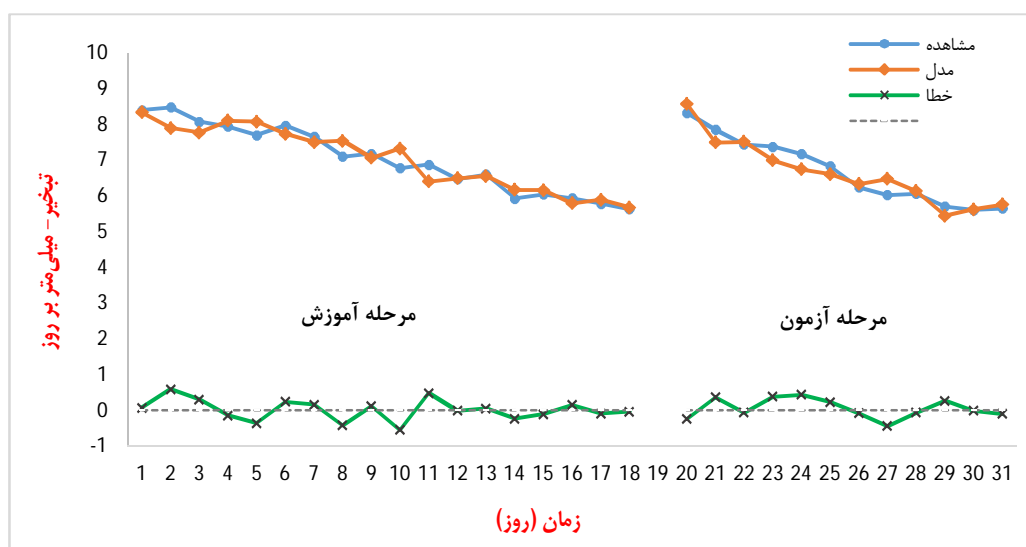
تحقیقات انجام شده در گذشته، نشان دهنده دقت بیشتر مدل برنامه‌ریزی بیان‌ژن در برآورد تبخیر از تشت در ایستگاه سد زاینده رود است. برای نمونه باقری و همکاران (۱۴۰۲) با استفاده از مدل برنامه‌نویسی بیان‌ژن بهترین نتیجه در برآورد تبخیر سه ایستگاه استان خوزستان، در ایستگاه دزفول و به مقدار ۰/۹۳ میلی‌متر گزارش کردند. حقیقت‌جو و همکاران (۱۳۹۶) نیز بهترین نتایج مدل برنامه‌ریزی بیان‌ژن در ایستگاه بشرویه در برآورد تبخیر از تشتک را با خطای ۱/۴۴ میلی‌متر گزارش کردند. محمدزاده و همکاران (۱۳۹۴) نیز میزان خطای شبیه‌سازی مدل برنامه‌ریزی بیان‌ژن برای تبخیر روزانه در استان خراسان جنوبی را بین ۱/۴ تا ۴/۳ گزارش کردند.

جدول ۲: مقادیر شاخص‌های آماری ارزیابی مدل برنامه‌نویسی بیان‌ژن

تاخیر زمانی ماهانه	مهر	R <sup>2</sup>	RMSE	N.S.	MAPE	CRM
1	آموزش	0.68	0.33	0.68	3.31	0.00
	آزمون	0.46	1.84	-1.43	7.77	0.19
2	آموزش	0.90	0.29	0.90	3.21	0.00
	آزمون	0.84	0.38	0.83	2.74	0.02
3	آموزش	0.88	0.21	0.88	2.12	0.00
	آزمون	0.56	0.35	0.18	3.23	-0.05
4	آموزش	0.89	0.20	0.89	2.14	0.00
	آزمون	0.63	0.51	-1.04	4.40	0.04
شهریور						
1	آموزش	0.34	0.31	0.32	2.47	0.00
	آزمون	0.27	1.00	0.04	6.55	-0.10
2	آموزش	0.68	0.22	0.89	1.53	0.00
	آزمون	0.13	0.72	0.01	4.27	-0.06
3	آموزش	0.77	0.18	0.77	1.31	0.00
	آزمون	0.41	0.65	0.11	3.91	-0.06
مرداد						
1	آموزش	0.51	0.12	0.50	0.84	0.00
	آزمون	0.06	0.52	0.00	2.49	-0.04
2	آموزش	0.63	0.11	0.63	0.64	0.00
	آزمون	0.27	0.33	0.10	1.47	-0.02
تیر						
1	آموزش	0.86	0.14	0.86	0.82	0.00
	آزمون	0.22	0.40	-0.16	1.51	0.01

شهریور)، غیر قابل قبول می‌باشند. همانگونه که در جدول ۲ نیز مشاهده می‌شود ضریب کارایی مدل یا نش- ساتکلیف (N.S) در تمام الگوهای مورد بررسی در این ماه‌ها کمتر ۰/۵ و بیان‌کننده عدم کارایی مدل در برآورد تبخیر روزانه است.

استفاده از داده‌های درازمدت روزانه در پژوهش حاضر (دوره آماری ۵۳ ساله)، تشخیص و اصلاح داده‌های پرت و استفاده از تنظیمات توصیه شده مدل بر اساس جدیدترین تحقیقات را می‌توان دلیل افزایش دقت در نتایج دانست. نتایج مدل برنامه‌ریزی بیان‌ژن در ماه‌های دیگر مورد بررسی (خرداد، تیر، مرداد،



شکل ۳- مقادیر مشاهداتی تبخیر روزانه و مقادیر برآورد شده توسط مدل در ماه مهر در مرحله آزمون

قابل قبولی ارائه نموده که نشان می‌دهد به‌خوبی توانایی مدل‌سازی تبخیر روزانه را دارد. مقادیر شاخص کارایی مدل (N.S) در تمام ماه‌ها حداقل ۰/۸۶ است که بر اساس طبقه‌بندی موربانی نشان دهنده توانایی بسیار خوب مدل در شبیه‌سازی داده‌های روزانه تبخیر است.

کمترین مقدار خطای مدل در ماه مهر و معادل ۰/۷ میلی‌متر در روز بدست آمد که بر اساس آماره MAPE که بیان‌کننده میانگین درصد خطای مطلق مدل است معادل ۸/۸ درصد است. میزان ضریب تبیین ( $R^2$ ) و کارایی مدل (N.S) در برآورد تبخیر روزانه نیز بیشتر از ماه‌های دیگر و ۰/۹ است (جدول ۳). به عبارت دیگر مدل فوریه ۹۰ درصد از تغییرات داده‌های واقعی تبخیر روزانه را شناسایی و شبیه‌سازی نموده است. در شکل ۴ الف میزان انطباق داده‌های مدل شده در مقابل داده‌های مشاهداتی (مقادیر تبخیر واقعی) در ماه مهر به‌خوبی دیده می‌شود.

برخلاف تحقیق حاضر، برخی تحقیقات نتایج مدل برنامه‌نویسی بیان‌ژن را در پیش‌بینی تبخیر روزانه از تشت، قابل قبول دانسته‌اند. مانند کیم و همکاران (۲۰۱۵)، که نشان دادند کارایی مدل برنامه‌نویسی بیان‌ژن در مقایسه با مدل رگرسیون خطی چندگانه در پیش‌بینی تبخیر روزانه از تشت در منطقه آب‌وهوای خشک در جنوب غربی ایران، برتری دارد. حقیقت‌جو و همکاران (۱۳۹۶) نیز نشان دادند که مدل برنامه‌ریزی بیان‌ژن کارایی بهتری نسبت به مدل عصبی فازی در برآورد روزانه تبخیر از تشتک استان خراسان جنوبی دارد.

**مدل‌سازی تبخیر روزانه به کمک مدل فوریه:**  
نتایج نهایی مدل‌سازی داده‌های تبخیر روزانه از تشت بر اساس مدل فوریه برای پنج ماه فصل گرم سال در جدول ۳ ارائه شده است. برخلاف مدل برنامه‌ریزی بیان‌ژن که تنها در ماه مهر نتایج قابل قبولی ارائه کرده بود، مدل فوریه در تمام ماه‌های مورد مطالعه نتایج

تغرق در سه ایستگاه در نیومکزیکو را بین ۰/۶۴ تا ۰/۹۲ میلی‌متر و در نبرسکا را ۱/۸۱ میلی‌متر گزارش نمودند. مقدار ضریب کارایی مدل (N.S) نیز در ماه‌های مورد مطالعه حداقل ۰/۸۶ بدست آمده است که نشان می‌دهد توانایی مدل فوریه در برآورد تبخیر روزانه در این ایستگاه، در وضعیت بسیار خوب (جدول ۱) قرار دارد. همچنین ضریب جرم باقیمانده (CRM) برای تمام ماه‌ها بسیار نزدیک به صفر و مثبت بدست آمده است که نشان می‌دهد مدل از نظر بیش‌برآوردی و کم‌برآوردی در حد متعادل قرار دارد و اصولاً بیش‌برآوردی و کم‌برآوردی قابل ملاحظه‌ای ندارد.

بیشترین خطای مدل فوریه نیز در ماه خرداد و در حدود یک میلی‌متر بر روز است (جدول ۳ و شکل ۴ ب). مقدار خطای مدل فوریه در ماه‌های دیگر نیز کمتر از یک میلی‌متر در روز به‌دست آمد. این خطا بر اساس شاخص MPAE بین ۵/۲ تا ۷/۵ درصد می‌باشد. بطور کلی تغییرات خطا در این مدل برای ماه‌های مورد مطالعه بین ۰/۷ تا ۱ میلی‌متر بر روز است که بر اساس شاخص MAPE در تمام ماه‌ها کمتر از ۹ درصد است. پژوهشگران دیگر نیز نتایج قابل قبول با خطای پایین از عملکرد مدل فوریه گزارش نموده‌اند. برای نمونه سویلو و همکاران (۲۰۱۲) با استفاده از روش فوریه مقدار خطا برای برآورد تبخیر و

جدول ۳: نتایج شاخص‌های آماری ارزیابی مدل فوریه

	خرداد	تیر	مرداد	شهریور	مهر
R <sup>2</sup>	0.87	0.89	0.89	0.86	0.9
RMSE	1.02	0.9	0.87	0.9	0.7
MAPE	7.5	5.2	5.4	6.5	8.8
N.S.	0.87	0.89	0.89	0.86	0.9
d	0.96	0.97	0.97	0.96	0.97
CRM	0	0	0	0	0

موقعیت جغرافیایی هستند، تفاوت نتایج مدل‌های یکسان در مناطق مختلف قابل توجیه است.

**نتایج مدل‌سازی تبخیر روزانه با مدل ARIMA:**  
نتایج حاصل از مدل آریمای در برآورد تبخیر روزانه ایستگاه سد زاینده‌رود در جدول ۴ ارائه شده است. مقایسه نتایج حاصل از این مدل با دو مدل قبل نشان می‌دهد مقادیر خطای این مدل، در تمام ماه‌ها بیشتر از مدل فوریه و مدل برنامه‌ریزی بیان‌ژن می‌باشند. مقادیر ضریب کارایی (N.S) مدل آریمای نیز نشان می‌دهد در تمام ماه‌ها مدل در وضعیت رضایت بخش قرار دارد. مقایسه نتایج سه مدل نشان می‌دهد بهترین نتیجه در ماه مهر با خطای ۰/۳۸ (۲/۷ درصد) مربوط به مدل برنامه‌ریزی بیان‌ژن است و در بقیه ماه بهترین نتیجه با خطایی بین ۰/۸۷ تا ۱/۰۲ میلی‌متر در روز (۵/۲ تا ۸/۸ درصد) مربوط به مدل فوریه می‌باشد. این در حالی است که مقادیر خطای مدل آریمای بین ۱/۴۹ تا ۱/۸۶ میلی‌متر بر روز (۹/۴ تا ۱۹/۶ درصد) بدست آمده است.

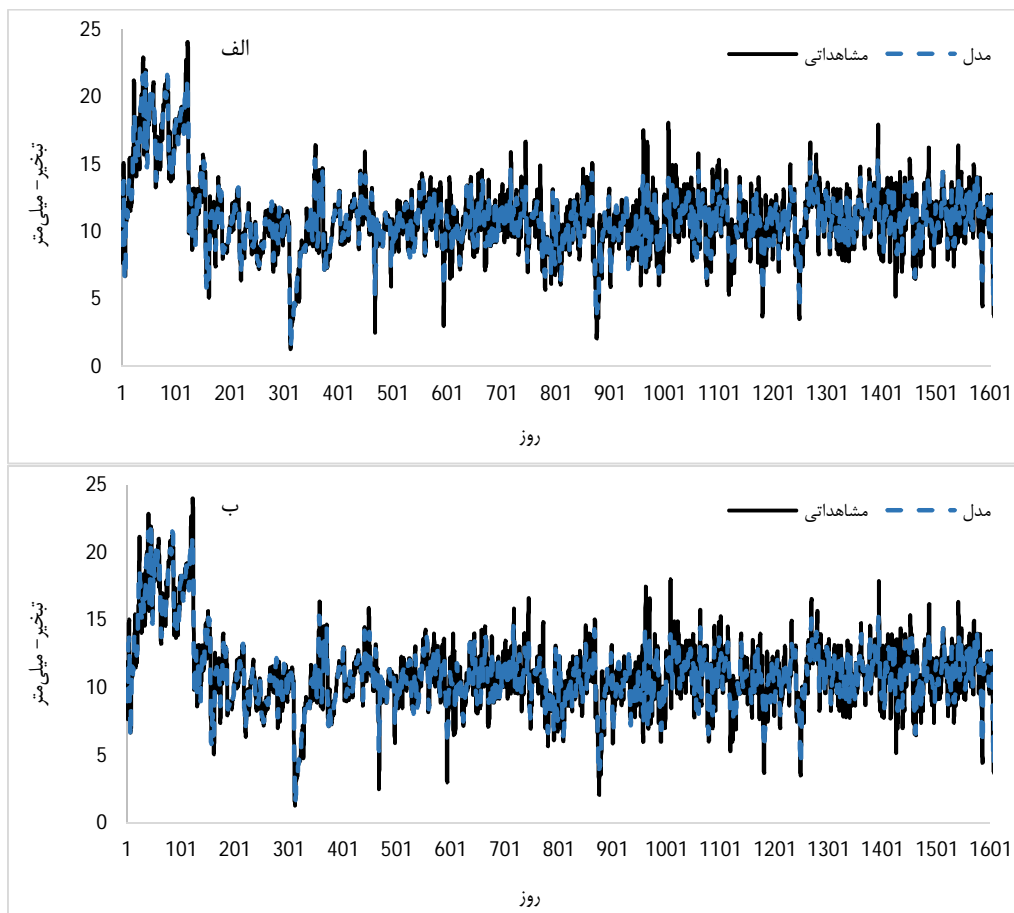
شکل ۴ انطباق مناسب نمودار داده‌های شبیه‌سازی شده با مدل فوریه را با داده‌های واقعی نشان می‌دهد که شاخص‌های آماری شرح داده شده در بالا، انطباق قابل قبول این دو سری را تایید می‌کند. همانگونه که در این شکل مشاهده می‌شود مدل توانسته مقادیر حدی تبخیر روزانه (مقادیر حداکثر و حداقل) را نیز تا حد بسیار زیادی بویژه در ماه مهر شبیه‌سازی کند. بالا بودن ضریب نش-ساتکلیف (N.S) نشان دهنده این توانایی است. مقایسه نتایج آماری حاصل از مدل‌های برنامه‌نویسی بیان‌ژن و فوریه حاکی از آن است که مدل فوریه با توانایی بیشتری قادر به تخمین و برآورد مقدار تبخیر روزانه در ایستگاه سد زاینده‌رود است. به‌عبارت دیگر می‌توان نتیجه گرفت که مدل فوریه توانایی بیشتری در تشخیص الگوی تغییرات غیرخطی حاکم بر داده‌های تبخیر روزانه، نسبت به مدل برنامه‌نویسی بیان‌ژن دارد. توانایی مدل‌ها در تشخیص الگوی تغییرات و شرایط غیرخطی حاکم بر داده‌های اقلیمی متفاوت است. با توجه به اینکه متغیرهای اقلیمی تابع

جدول ۴- نتایج شاخص‌های آماری ارزیابی مدل ARIMA

مدل	R <sup>2</sup>	RMSE	MAPE	N.S	d	CRM	
خرداد	ARIMA(0,1,5)	0.56	1.86	14.6	0.55	0.84	0.00
تبر	ARIMA(0,1,11)	0.65	1.58	9.4	0.65	0.89	0.00
مرداد	ARIMA(0,1,2)	0.68	1.49	9.5	0.64	0.89	0.00
شهریور	ARIMA(0,1,14)	0.56	1.61	11.9	0.53	0.83	0.00
مهر	ARIMA(1,0,1)	0.52	1.54	19.6	0.42	0.73	-0.05

در مدل برنامه‌ریزی بیان‌ژن گزارش نمودند و یا آزاد و همکاران (۱۳۹۴) نتایج شبکه عصبی مصنوعی در برآورد تبخیر-تعرق در ایستگاه ارومیه را دقیق‌تر از روش‌های سری زمانی (AR و ARMA) گزارش کردند.

برخی پژوهشگران دیگر نیز خطای مدل آریمادر برآورد و شبیه‌سازی برخی پارامترهای اقلیمی را بیشتر از مدل‌های پیچیده‌تر گزارش نموده‌اند. برای نمونه نوذری و همکاران (۱۴۰۲) با مقایسه مدل‌های ARIMA، رگرسیون خطی و برنامه‌ریزی بیان‌ژن در برآورد تبخیر از سد اکباتان، کمترین مقدار خطا را



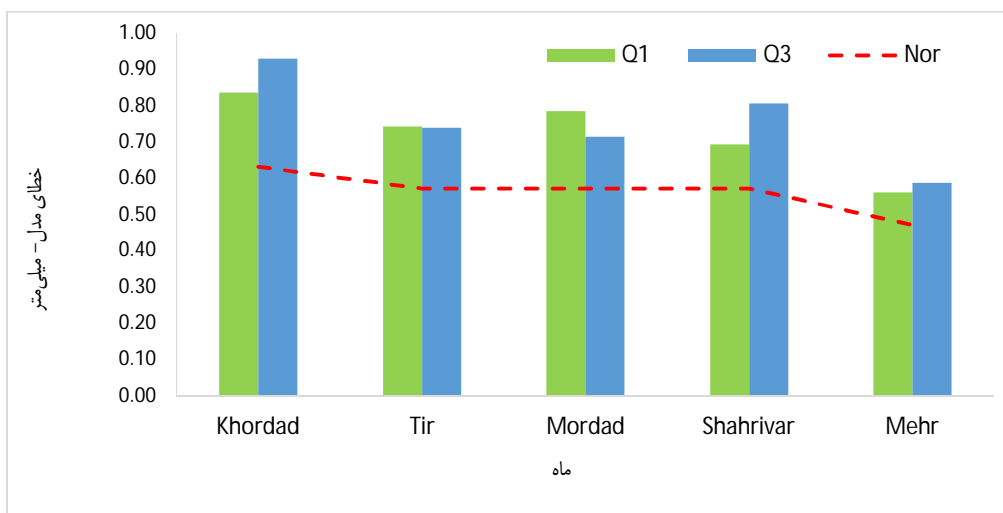
شکل ۴: تغییرات مقادیر مشاهداتی تبخیر روزانه از تشت و مقادیر برآورد شده توسط مدل فوریه در مهر (الف) و خرداد (ب)

بهترین نتایج هر دو مدل برنامه‌ریزی بیان‌ژن و فوریه را در این ماه نشان داد.

بیشترین تفاوت در تخمین داده‌های حدی نسبت به داده‌های غیر حدی نیز در خرداد بدست آمده است. به عبارت دیگر مدل در تخمین داده‌های حدی در این ماه نسبت به ماه‌های دیگر دچار خطای بیشتری شده است (شکل ۵). در تیرماه نیز مدل داده‌های حدی بالا و پایین را با دقت یکسان تخمین زده است ولی این داده‌ها را نسبت به داده‌های غیر حدی با خطای بیشتری برآورد کرده است. تخمین داده‌های حدی توسط مدل‌ها همواره یکی از مهمترین عوامل خطای مدل بوده است. عموماً مدل‌ها داده‌های غیر حدی را با دقت بیشتر و خطای کمتری تخمین می‌زنند.

### ارزیابی خطای مدل در تخمین مقادیر حدی

**تبخیر:** با توجه به اینکه بهترین نتایج بدست آمده برای برآورد تبخیر از تشت در ایستگاه سد زاینده‌رود در مدل فوریه به دست آمد، جهت بررسی دقت و توانایی این مدل در برآورد مقادیر حدی تبخیر روزانه، میانگین خطای مطلق در چارک اول (داده‌های حدی پایین) و چارک سوم (داده‌های حدی بالا) با مقدار میانگین خطای مطلق مقادیر داده‌های غیرحدی (داده‌های نرمال) مورد مقایسه قرار گرفت (شکل ۵). همانگونه که مشاهده می‌شود مدل فوریه مقادیر حدی تبخیر را در تمام ماه‌ها با خطای بیشتری نسبت به داده‌های غیرحدی تخمین زده است. کمترین تفاوت در تخمین داده‌های حدی با داده‌های معمولی در مهرماه مشاهده می‌شود. نتایج شاخص‌های آماری بررسی شده در بخش قبل نیز



شکل ۵: تغییرات میانگین مطلق خطای مدل در چارک اول (Q1)، چارک سوم (Q3) و داده‌های معمولی (Nor) مقادیر تبخیر

برنامه‌نویسی بیان‌ژن تنها در ماه مهر و الگوی استفاده از داده‌های روزانه دو ماه قبل برای تخمین تبخیر روزانه نتایج قابل قبولی ارائه داده است. مقدار ضریب تبیین و ضریب کارایی مدل (N.S) در این الگوی به ترتیب ۰/۸۴ و ۰/۸۳ بدست آمد، که نشان می‌دهد مدل برنامه‌نویسی بیان‌ژن توانسته است، ۸۴ درصد از تغییرات داده‌های واقعی را مدل‌سازی نماید. مقدار خطای برآورد تبخیر روزانه در این ماه کمتر از ۱

### جمع‌بندی و نتیجه‌گیری

پژوهش حاضر با هدف مدل‌سازی تبخیر روزانه از تشت در ایستگاه سد زاینده‌رود در ماه‌های گرم سال (خرداد، تیر، مرداد، شهریور و مهر) و در بازه زمانی سال‌های ۱۳۴۴ تا ۱۳۹۶ انجام گرفت. جهت مدل‌سازی از مدل برنامه‌ریزی بیان‌ژن به‌عنوان یک مدل هوش مصنوعی و دو مدل سری‌زمانی شامل مدل فوریه و آریمای استفاده گردید. نتایج نشان داد مدل

مدل)، توانایی این مدل در برآورد مقادیر حدی داده‌های تبخیر روزانه (چارک اول و سوم) نیز بررسی و نتایج نشان داد بیشترین دقت در تخمین این داده‌ها در مهرماه و کمترین دقت در خرداد ماه مشاهده می‌شود. قابل ذکر است تفاوت نتایج مدل‌ها می‌تواند ناشی از ساختار کلی آنها باشد. بطور کلی مدل‌ها برای شبیه‌سازی و پیش‌بینی داده‌ها ابتدا الگو تغییرات داده‌ها را شناسایی و بر اساس الگوی تشخیص داده شده اقدام به شبیه‌سازی یا پیش‌بینی می‌نمایند. به دلیل اینکه مدل‌های مختلف از ساختارها، روابط و تئوری‌های متفاوت در تشخیص الگوها استفاده می‌کنند، توانایی آنها در شبیه‌سازی داده‌ها متفاوت و در نتیجه، نتایج حاصله نیز متفاوت خواهد بود. در مجموع بر اساس نتایج آماری حاصل از بررسی سه مدل مورد بررسی، می‌توان استفاده از روش فوریه را برای برآورد میزان تبخیر و یا تخمین اعداد گمشده در ایستگاه سد زاینده‌رود توصیه نمود.

میلی‌متر به دست آمد. نتایج مدل برنامه‌نویسی بیان‌ژن بر اساس ضریب کارایی مدل ( $N.S < 0.5$ ) در ماه‌های دیگر غیر قابل قبول می‌باشند.

بر خلاف مدل برنامه‌نویسی بیان‌ژن، مدل فوریه در تمام ماه‌های مورد مطالعه نتایج قابل‌قبولی ارائه داد. مقادیر خطای برآورد تبخیر روزانه در ایستگاه سد زاینده رود در این روش در تمام ماه‌های مورد مطالعه بین حداکثر ۱/۰۲ تا حداقل ۰/۷ میلی‌متر در روز به دست آمد. این میزان خطا معادل ۵/۲ تا ۸/۸ درصد است. ارزیابی این مدل بر اساس شاخص جرم باقیمانده (CRM) نیز نشان داد مدل از نظر بیش‌برآوردی و کم‌برآوردی در حد متعادل قرار دارد و اصولاً بیش‌برآوردی و کم‌برآوردی قابل ملاحظه‌ای ندارد. مقایسه نتایج دو مدل فوق با مدل آریما نیز نشان داد مقادیر خطای این مدل، بیشتر از دو مدل فوریه و برنامه‌ریزی بیان‌ژن است. خطای مدل آریما در ماه‌های مورد بررسی بین ۹/۴ تا ۱۹/۶ درصد متغیر است. جهت ارزیابی دقیق‌تر مدل فوریه (به‌عنوان بهترین

## منابع

۱. احمدپور، عباس، فتحیان، حسین، و مقسمی، مهدی (۱۳۹۷). برآورد تبخیر ماهانه با مدل‌های سری زمانی SARIMA و BiLinear در ایستگاه ایدنک. *فصلنامه اکوبیولوژی تالاب*، ۳۷(۳)، ۴۹-۶۲.
۲. امجدی، حسین، ماشاءالله، علی، و سلطانی، محمد (۱۳۹۸). ارزیابی قابلیت مدل برنامه‌ریزی بیان‌ژن در شبیه‌سازی تابش خورشیدی رسیده به زمین. *مطالعه موردی اهواز، انرژی‌های تجدیدپذیر و نو*. ۲(۶)، ۹۳-۸۵.
۳. آزاد، نسرين، بهمنش، جواد، منتصری، مجید، و رضوردی‌نژاد، وحید (۱۳۹۴). مقایسه روش‌های سری زمانی و شبکه عصبی مصنوعی در پیش‌بینی تبخیر-تعرق مرجع، مطالعه موردی: ارومیه. *علوم و مهندسی آبیاری*، ۳۸(۴): ۷۵-۸۵.
۴. باقری، نیلوفر، قاسمی، احمدرضا، و نصر، محمدعلی (۱۴۰۲). مدل‌سازی تبخیر روزانه از تشت، با استفاده از برنامه‌نویسی بیان‌ژن و سری‌های زمانی در استان خوزستان. *پژوهش آب ایران*. ۴۹(۱)، ۱-۱۱.
۵. بذرافشان، ام‌البنین، چشم‌براه، عظیمه و حل‌ساز، ارشک. ۱۳۹۶. ارزیابی کارایی مدل‌های سری زمانی در پیش‌بینی تبخیر از تشت در نمونه‌های اقلیمی استان هرمزگان. *مهندسی و مدیریت آبخیز*، ۹(۳)، ۲۶۱-۲۵۰.
۶. تیموری، مریم، و تیموری، لیلا (۱۴۰۱). بررسی کاربرد مدل‌های سری زمانی در پیش‌بینی جریان ماهانه ایستگاه هیدرومتری آراز کوسه. *مطالعات علوم محیط زیست*، ۷(۲): ۴۷۹۹-۴۸۰۷.
۷. حقیقت‌جو، پرویز، محمدزاده، زهرا، و محمدرضاپور، ام‌البنین (۱۳۹۶). مقایسه روش‌های سیستم استنتاج عصبی - فازی و برنامه‌ریزی بیان‌ژن در برآورد تبخیر از تشتک. *مجله حفاظت منابع آب و خاک*، ۶(۴)، ۱۱۷-۱۰۷.
۸. سبزواری، یاسر، و مهری، سعیدی‌نیا (۱۴۰۰). ارزیابی مدل‌های تجربی و هوش مصنوعی در برآورد تبخیر-تعرق مرجع- مطالعه موردی: ایستگاه بروجرد. *نشریه علوم آب و خاک*، ۵(۲): ۲۵۳-۲۳۷.
۹. سلطانی، سمیه، صابری، عارف، و قیصوری، مرتضی. (۱۳۹۶). تعیین بهترین مدل سری زمانی در پیش‌بینی

۱۹. مرتضوی، مرتضی، توکلی، افشین، محمدی، محمد، و افصحی، کامران (۱۳۹۴). تاثیر کاربرد پلیمر سوپرجاذب بر صفات فیزیولوژیک و عملکرد گندم رقم آذری در شرایط دیم. نشریه زراعت، ۲۶(۱۰۶)، ۱۲۵-۱۱۸.
۲۰. نجفوند، مهدی، و اسلامی، حسین (۱۳۹۵). مقایسه روش‌های تجربی برآورد تبخیر از سطح آزاد آب. مطالعه موردی: سد تنظیمی دز. مجله مهندسی آب، ۴(۲)، ۶۵-۷۲.
۲۱. نوذری، حامد، آزادی، سعید، صدق‌نژاد، ناهید، و پویان‌فر، سجاد (۱۴۰۲). پیش‌بینی تبخیر ماهانه به کمک مدل‌های خطی و غیرخطی سری زمانی - مطالعه موردی: ایستگاه تبخیرسنجی سد اکباتان. نشریه هواشناسی کشاورزی، ۱۱(۱)، ۳۶-۴۷.
22. Alarcon, A., Cortes, D., Alvarez, J. Gonzalez, Y. (2022). Improving Monthly Rainfall Forecast in a Watershed by Combining Neural Networks and Autoregressive Model. *Environmental Processes*, 9(53): 1-26.
23. Gandomi, A., Alavi, A. Mirzahosseini M. & Moqhadas F. (2011). Nonlinear genetic-based models for prediction of flow number of asphalt mixtures. *Journal of Materials in Civil Engineering*, 23(4), 248-263.
24. Gaur, S., Singh, R., Bandyopadhyay, A., & Singh, R. (2023). Diagnosis of GCM-RCM-driven rainfall patterns under changing climate through the robust selection of multi-model ensemble and sub-ensembles. *Climatic Change*, 176(2): 1-30.
25. Goel, A. (2009). Application of SVMs Algorithms for Prediction of Evaporation in Reservoirs. World Environmental and Water Resources Congress, Missouri, United States.
26. Hael, M.A., Yongsheng, Y. & Saleh, B.I. (2020). Visualization of rainfall data using functional data analysis. *Applied Science*, 2(461). <https://doi.org/10.1007/s42452-020-2238-x>
27. Khanal, N., Matin, M., Uddin, K., Poortinga, A., Chishtie, F., Tenneson, K., & Saah, D. (2020). Comparison of Three Temporal Smoothing Algorithms to Improve Land Cover Classification: A بارندگی سالیانه ایستگاه‌های منتخب استان آذربایجان غربی. تحقیقات کاربردی علوم جغرافیایی. ۱۷(۴۴): ۱۰۵-۱۸۷. 20.1001.1.22287736.1396.17.44.5.1.۸۷-۱۰۵.
۱۰. سیاسر، هادی، امیر، سالاری، ام‌البنین، محمدرضاپور، و حلیمه، پیری (۱۴۰۰). برآورد روزانه تبخیر و تعرق مرجع در دشت سیستان با استفاده از الگوریتم‌های اکوسیستم بیابان، ۱۰(۳۲): ۸۵-۹۶.
۱۱. صادقی‌اقدم، فریبا، ندیری، عطاله، اصغری، اصغر، و آرمانفر، فریدون (۱۳۹۶). کاربرد برنامه‌ریزی بیان ژن برای ایجاد مدل ترکیبی فازی به منظور پیش‌بینی غلظت آرسنیک منابع آبی در حوضه آبریز سد سهند. پژوهش آب ایران. ۸۵: ۲۷-۹۶.
۱۲. طاووسی، تقی، و زهرایی، اکبر (۱۳۹۴). مدل‌سازی و پیش‌بینی تغییرات بیشینه دمای شیراز برای دوره اقلیمی منتهی به سال ۱۴۰۰ خورشیدی، مجله نیوار، ۳۹(۸۸)، ۴۳-۵۲.
۱۳. عمادی، علی، زمان‌زاد، سروین، زارعی، سهیلا، رشید، علی (۱۴۰۱). مدل‌سازی تبخیر ماهانه با استفاده از روش‌های منفرد و هیبریدی - موجک داده کاوی در حوضه‌های آبریز ایران با تنوع اقلیمی، نشریه آبیاری و آب ایران، ۳۷۳-۳۵۴، ۴۸(۴).
۱۴. فرزادمهر، مهنوش، دستورانی، مهدی، خاشعی، عباس، و جلالی، وحیدرضا (۱۳۹۹). برآورد هدایت هیدرولیکی اشباع خاک به روش برنامه ریزی بیان ژن و مقایسه آن با توابع انتقال. مدیریت حوزه آبخیز. ۱۱(۲۲): ۱۵۵-۱۶۴.
۱۵. فولادمند، حمیدرضا (۱۳۸۹). پیش‌بینی ماهانه تبخیر و تعرق پتانسیل گیاه مرجع در استان فارس. مجله دانش آب و خاک، ۲۰(۴)، ۱۶۹-۱۵۷.
۱۶. فیض‌اله پور، فرید، دلاور، مجی، و حسامی، مهدی (۱۳۹۶). ارزیابی و تحلیل عدم قطعیت برآورد تبخیر تعرق گیاه مرجع با استفاده از برنامه ریزی ژنتیک. نشریه دانش آب و خاک، ۴(۲۷)، ۱۴۷-۱۳۵.
۱۷. قهرمان، نوذر، و قره‌خانی، ابودر (۱۳۹۰). ارزیابی مدل‌های تصادفی سری زمانی در برآورد تبخیر از تشتک (مطالعه موردی: ایستگاه شیراز). نشریه پژوهش آب در کشاورزی، ۲۵(۱)، ۷۵-۸۱.
۱۸. محمدزاده، زهرا (۱۳۹۴). مقایسه روش‌های نروفازی و برنامه‌ریزی بیان ژن برای برآورد تبخیر از تشتک در خراسان جنوبی. پایان نامه کارشناسی ارشد، دانشگاه زابل.

- Pp. 1-72, In: Maidment D.R, *Handbook of Hydrology*, McGraw- Hill.
36. Shoaib, M., Shamseldin, A.S., Melville, B., & Muneer, M. (2015). Runoff forecasting using hybrid Wavelet Gene Expression Programming (WGEP) approach. *Journal of Hydrology*, 527: 326- 344.
  37. Soylu, M., Lenters, John, D., Istanbuluoglu, E., & Loheide, S. (2012). On evapotranspiration and shallow groundwater fluctuations: A Fourier-based improvement to the White method. *Papers in Natural Resources*. 613. <https://digitalcommons.unl.edu/natrespapers/613>
  38. Sudheer, K.P. (2000). Modeling hydrological processes using neural computing technique. PhD Thesis, Indian Institute of Technology, Delhi. India.
  39. Tularam, G., & Ilahee, M. (2010). Time Series Analysis of Rainfall and Temperature Interactions in Coastal Catchments. *Journal of Mathematics and Statistics*. 6 (3): 372-380.
  40. Ustoorikar, K., & Deo, M.C. (2008). Filling up gaps in wave data with genetic programming. *Marine Structures*, 21(2): 177-195.
  41. Wu, X., Zhou, J., Yu, H., Li, D., Xie, K., Chen, Y., Hu, J., Sun, H., & Xing, F. (2021). The development of a hybrid wavelet-ARIMA-LSTM model for precipitation amounts and drought analysis. *Atmosphere*. 12: <https://doi.org/10.3390/atmos12010074>.
  42. Yassin, M., Alazba, A., & Mattar, M. (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*. 163: 110-124.
  - Case Study from NEPAL. *Remote Sensing*, 12, 2888.
  28. Kim, S., Shiri, J., Singh, V., Kisi, O., & Landaras, G. (2015). Predicting daily pan evaporation by soft computing models with limited climatic data. *Hydrological Sciences Journal*, 60(6): 1120-1136.
  29. Koza, J.R. (1993). Hierarchical Automatic Function Definition in Genetic Programming. *Foundations of Genetic Algorithms*, 2, 297-318.
  30. Laguardia, G. (2011). Representing the precipitation regime by means of Fourier series. *International Journal of Climatology*, 31: 1398–1407.
  31. Li, L., Zhou, X., Li, Y., Gong, C., Lu, L., Fu, X., & Tao, W. (2017). Water absorption and water/fertilizer retention performance of vermiculite modified sulphoaluminate cementitious materials. *Construction and Building Materials*, 137, 224-234.
  32. Mazelan, N.A., & Suhaila J. (2023). Exploring rainfall variabilities using statistical functional data analysis. *Earth and Environmental Science*. 1167: 1-10. doi:10.1088/1755-1315/1167/1/012007.
  33. Miranda, A., Herrera, M., & Castano, V. (2019). Meteorological Temperature and Humidity Prediction from Fourier-Statistical Analysis of Hourly Data. *Advances in Meteorology*, 34(2), 2-13. <https://doi.org/10.1155/2019/4164097>
  34. Moriasi, D., Arnold, J., VanLiew, M.W., Bingner, R.L., Harmel, R.D., & Veith, T. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, 50(3), 885-905.
  35. Salas, J.D., (1993). Analysis and Modeling of Hydrologic Time Series.

