

Comparative performance of linear and deep learning models in monthly runoff prediction, a case study of the Qarasu Watershed

Nader Jandaghi^{1*} and Mehdi Alibegli²

¹ Associate Professor in Engineering Hydrology, Department of Rangeland and Watershed Management, Faculty of Agriculture and Natural Resources, Gonbad Kavous University, Iran

² PhD student in Computer Engineering, Gorgan Azad University, Iran

Received: 11 August 2025

Accepted: 04 January 2026

Extended abstract

Introduction

Water resources management in arid and semi-arid regions is one of the fundamental challenges of the present century. Population growth, climate change, and increasing demand have put significant pressure on surface water resources. In this context, monthly runoff prediction serves as a strategic tool for reservoir planning, flood control, and sustainable watershed management. However, the complexity of hydrological processes and the nonlinear relationships among climatic variables make such predictions challenging. In recent years, hybrid models have emerged as a novel approach in hydrology that enhances prediction accuracy and stability by integrating the strengths of different models. The present study aims to evaluate the performance of linear and deep learning models and to introduce an optimal model for enhancing water resources management.

Materials and methods

In this study, monthly runoff in the Qarasu watershed, with an area of 1,624 km², was modeled and predicted using data from three hydrometric stations (Pole-Touskestan, Naharkhoran, and Siyahab) over a common 36-year period (1985-2021). Three single models (SARIMA, BiLSTM, and GRU) and two hybrid models (BiLSTM-GRU and SARIMA-BiLSTM-GRU) were employed to model monthly runoff and forecast values for a 12-month horizon. In hybrid models, the SARIMA model is used to model the linear components of the time series, while BiLSTM and GRU models have the ability to identify and represent complex, non-linear patterns. Model performance was evaluated using RMSE, MAD, and MSE indices. The models were implemented using Python and commonly used libraries, including TensorFlow, Keras, numpy, pandas, matplotlib, scipy, and sklearn.

Results and discussion

The hybrid SARIMA-BiLSTM-GRU model, by integrating linear and nonlinear components, provided the most accurate monthly runoff predictions. The RMSE values of this model at the Pole-Touskestan, Naharkhoran, and Siyahab stations were estimated at 0.0295, 0.0173, and 0.1683 m³/s, respectively. The BiLSTM-GRU model ranked second, with RMSE values of 0.0326, 0.0226, and 0.3013 m³/s. Among the individual models, BiLSTM and GRU produced similar and relatively accurate results, while the linear SARIMA model, with RMSE values of 0.0851, 0.0230, and 0.3892 m³/s, showed the lowest performance. On average and based on the RMSE index, the SARIMA-BiLSTM-GRU hybrid model reduced prediction errors by 39.66% to 56.75% compared to the other models. The findings of this research confirm and reinforce the theoretical foundations regarding the superiority of complex hybrid models over simple, linear models.

Conclusions

This study demonstrated that hybrid models combining linear and deep learning approaches can significantly improve the accuracy and stability of monthly runoff predictions. The modeling and validation results indicated that the hybrid SARIMA-BiLSTM-GRU model, which effectively combines linear (SARIMA) and nonlinear (BiLSTM and GRU) components, was selected as the optimal model for this watershed, demonstrating the best performance in monthly runoff prediction. The results highlight the importance of applying hybrid approaches in water resources management and flood control. It is recommended to employ more advanced models and diverse input variables, such as precipitation and

* Corresponding author: nader.jandaghi@gmail.com

temperature, to further enhance the accuracy and reliability of predictions. Furthermore, given the observation of overfitting in the base LSTM and GRU models, it is recommended to employ more advanced architectures, such as models based on the Attention mechanism like the Transformer, to enable the model to intelligently focus on the key and influential time periods in the runoff process.

Keywords: Golestan Province, Hybrid models, Hydrological forecasting, SARIMA-BiLSTM-GRU, Water resources management

Cite this article: Jandaghi, N., Alibegli, M., 2026. Comparative performance of linear and deep learning models in monthly runoff prediction, a case study of the Qarasu Watershed. *Water. Eng. Manag.* 18(1), 147-164.

© 2026, The Author(s). Published by Soil Conservation and Watershed Management Research Institute (SCWMRI). This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)



مقایسه عملکرد مدل‌های خطی و یادگیری عمیق در پیش‌بینی رواناب ماهانه مطالعه موردی: حوزه آبخیز قره‌سو

نادر جندقی^{۱*} و مهدی علی‌بگلی^۲

^۱ دانشیار هیدرولوژی مهندسی، گروه مرتع و آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه گنبد کاووس، گنبد کاووس، ایران
^۲ دانشجوی دکتری مهندسی کامپیوتر، دانشگاه آزاد گرگان، گرگان، ایران

تاریخ پذیرش: ۱۴۰۴/۱۰/۱۴

تاریخ دریافت: ۱۴۰۴/۰۵/۲۰

چکیده مبسوط

مقدمه

مدیریت منابع آب در مناطق خشک و نیمه‌خشک از چالش‌های اساسی قرن حاضر است. رشد جمعیت، تغییرات اقلیمی و افزایش تقاضا فشار زیادی بر منابع آب سطحی وارد کرده است. در چنین شرایطی، پیش‌بینی رواناب ماهانه ابزاری راهبردی برای برنامه‌ریزی سدها، کنترل سیلاب و مدیریت پایدار حوزه‌های آبخیز محسوب می‌شود. با این حال، پیچیدگی فرایندهای هیدرولوژیکی و روابط غیرخطی متغیرهای اقلیمی، این پیش‌بینی را دشوار کرده است. در سال‌های اخیر، مدل‌های ترکیبی به‌عنوان رویکردی نوین در هیدرولوژی مطرح شده‌اند که با ادغام نقاط قوت مدل‌های مختلف، دقت و پایداری پیش‌بینی را بهبود می‌بخشند. هدف پژوهش حاضر، بررسی کارایی مدل‌های خطی و یادگیری عمیق و معرفی مدلی بهینه برای ارتقای مدیریت منابع آب است.

مواد و روش‌ها

در این پژوهش، رواناب ماهانه حوزه آبخیز قره‌سو با مساحت ۱۶۲۴ کیلومترمربع، با استفاده از داده‌های سه ایستگاه هیدرومتری پل توسکاستان، نهارخوران و سیاه‌آب طی یک دوره آماری ۳۶ ساله (۱۴۰۰-۱۳۶۴) مدل‌سازی و پیش‌بینی شد. برای این منظور، سه مدل منفرد (SARIMA، BiLSTM و GRU) و دو مدل ترکیبی (BiLSTM-GRU و SARIMA-BiLSTM-GRU) به کار گرفته شدند تا رواناب ماهانه را برای افق ۱۲ ماه آینده پیش‌بینی کنند. در مدل‌های ترکیبی، مدل SARIMA برای مدل‌سازی مؤلفه‌های خطی سری زمانی به کار می‌رود، در حالی که مدل‌های BiLSTM و GRU توانایی شناسایی و بازنمایی الگوهای پیچیده و غیرخطی را دارند. دقت مدل‌ها با شاخص‌های RMSE، MAD و MSE ارزیابی شد. پیاده‌سازی مدل‌ها با استفاده از Python و کتابخانه‌های پرکاربرد از جمله TensorFlow، Keras، numpy، pandas، matplotlib و sklearn انجام شد.

نتایج و بحث

مدل ترکیبی SARIMA-BiLSTM-GRU با تلفیق مؤلفه‌های خطی و غیرخطی دقیق‌ترین پیش‌بینی رواناب ماهانه را ارائه داد. مقادیر RMSE این مدل در ایستگاه‌های پل توسکاستان، نهارخوران و سیاه‌آب به ترتیب ۰/۰۲۹۵، ۰/۰۱۷۳ و ۰/۱۶۸۳ مترمکعب بر ثانیه برآورد شد. مدل BiLSTM-GRU با RMSE معادل ۰/۰۳۲۶، ۰/۰۲۲۶ و ۰/۰۳۰۱۳ گزینه دوم

* مسئول مکاتبات: nader.jandaghi@gmail.com

بود. در میان مدل‌های منفرد، BiLSTM و GRU نتایج مشابه و قابل قبول داشتند، اما مدل خطی SARIMA با RMSE برابر ۰/۰۸۵۱، ۰/۰۲۳۰ و ۰/۳۸۹۲ ضعیف‌ترین عملکرد را نشان داد. به‌طور میانگین و بر اساس شاخص RMSE، مدل ترکیبی SARIMA-BiLSTM-GRU توانست خطای پیش‌بینی را بین ۳۹/۶۶ درصد تا ۵۶/۷۵ درصد نسبت به سایر مدل‌ها کاهش دهد. یافته‌های این پژوهش، مبانی نظری مطرح‌شده در خصوص برتری مدل‌های ترکیبی پیچیده را نسبت به مدل‌های ساده و خطی تأیید و تقویت می‌کند.

نتیجه‌گیری

این پژوهش نشان داد که مدل‌های ترکیبی خطی و یادگیری عمیق دقت و پایداری پیش‌بینی رواناب ماهانه را به‌طور قابل توجهی بهبود می‌دهند. نتایج مدل‌سازی و اعتبارسنجی نشان داد که مدل ترکیبی SARIMA-BiLSTM-GRU با ترکیب مؤثر اجزای خطی (SARIMA) و غیرخطی (BiLSTM و GRU)، به‌عنوان مدل بهینه برای این حوضه انتخاب شد که دارای بهترین عملکرد در پیش‌بینی رواناب ماهانه بوده است. نتایج بر اهمیت به‌کارگیری رویکردهای ترکیبی در مدیریت منابع آب و کنترل سیلاب تأکید دارند. پیشنهاد می‌شود، استفاده از مدل‌های پیشرفته‌تر و ورودی‌های متنوع مانند بارش و دما برای افزایش دقت و قابلیت اعتماد پیش‌بینی‌ها مدنظر قرار گیرد. همچنین با توجه به مشاهده پدیده بیش‌برازش در مدل‌های پایه LSTM و GRU، پیشنهاد می‌شود از معماری‌های پیشرفته‌تری مانند مدل‌های مبتنی بر مکانیسم Attention نظیر Transformer استفاده شود تا مدل بتواند به‌طور هوشمند بر دوره‌های زمانی مؤثر و کلیدی در فرایند رواناب تمرکز کند.

واژه‌های کلیدی: استان گلستان، پیش‌بینی هیدرولوژیکی، مدل‌های ترکیبی، مدیریت منابع آب، SARIMA-BiLSTM-GRU

مقدمه

مدل‌های پیش‌بینی جریان رودخانه عمدتاً به دو دسته مبتنی بر فرایند و مبتنی بر داده تقسیم می‌شوند (Chen et al., 2023; Parisouj et al., 2022). مدل‌های فرایندی بر فهم ساز و کارهای تولید رواناب و انتقال در کانال‌ها متکی هستند و برای کالیبراسیون نیازمند داده‌های طولانی هیدرومتئورولوژیکی هستند (Beven, 2020; Yang et al., 2020)، اما پیچیدگی خطاها و عدم قطعیت‌ها دقت آنها را محدود می‌کند (Lin et al., 2021).

به‌کارگیری مدل‌های فرایندی به‌دلیل کمبود داده‌های ورودی بلندمدت و با وضوح مکانی لازم (مانند ویژگی‌های خاک، کاربری اراضی دقیق و داده‌های هواشناسی) و همچنین پیچیدگی ذاتی در کالیبراسیون ویژگی‌های فیزیکی با چالش جدی روبرو است. در مقابل، مدل‌های مبتنی بر داده با حداقل نیاز به داده‌های فیزیکی، توانایی مدل‌سازی روابط غیرخطی و انعطاف‌پذیری بیشتر، گزینه‌ای مناسب برای پیش‌بینی رواناب محسوب می‌شوند (Latif et al., 2023; Hauswirth et al., 2021; Hu et al., 2021).

مدیریت منابع آب به‌ویژه در مناطق نیمه‌خشک و خشک، یکی از مهم‌ترین چالش‌های قرن حاضر محسوب می‌شود. تغییرات اقلیمی، رشد جمعیت و افزایش نیاز به آب در بخش‌های کشاورزی، شرب و صنعت، فشار فزاینده‌ای بر منابع آب سطحی وارد کرده است (Nazeri et al., 2020; Heshmatpour et al., 2020). در این میان، پیش‌بینی دقیق رواناب ماهانه به‌عنوان یک ابزار کلیدی، نقش تعیین‌کننده‌ای در برنامه‌ریزی سدها، کنترل سیلاب، تخصیص منابع و مدیریت پایدار حوزه‌های آبخیز ایفا می‌کند (Jandaghi et al., 2020).

رواناب ماهانه حجم کل آب جاری شده در مقطع یک ایستگاه هیدرومتری طی یک ماه هیدرولوژیکی است. با این حال، ویژگی‌های پیچیده و غیرخطی فرایندهای هیدرولوژیکی، همراه با تأثیر متعدد اقلیمی و مکانی، این پیش‌بینی را به یکی از دشوارترین مسائل در هیدرولوژی تبدیل کرده است (Workneh and Jha, 2025).

2021. با این حال، مدل‌های خطی همچنان در برخی کاربردها مؤثر هستند، به‌ویژه برای پیش‌بینی تراز آب زیرزمینی ماهانه، که نمونه‌ای از آن استفاده از SARIMA در جنوب ایتالیا است (Boulariah et al., 2021).

در همین راستا، مدل‌های ترکیبی که قابلیت ادغام معماری‌ها و الگوریتم‌های مختلف را دارند، توجه روزافزونی در تحقیقات هیدرولوژیکی به خود جلب کرده‌اند. ترکیب مدل‌های یادگیری عمیق مختلف می‌تواند محدودیت‌های مدل‌های منفرد را جبران کرده و دقت، پایداری، قابلیت تعمیم و استحکام پیش‌بینی‌ها را افزایش دهد. برای نمونه، Sarkar et al., (2022) با ترکیب LSTM و GRU عملکرد برتری نسبت به مدل‌های منفرد ارائه دادند.

همچنین Ahmed et al., (2022) نیز با توسعه مدل ترکیبی BiLSTM-CNN و الگوریتم بهینه‌سازی کلونی مورچه^۵ نشان دادند که این رویکرد می‌تواند به‌طور قابل توجهی دقت پیش‌بینی رواناب را بهبود بخشد. مطالعات دیگر نیز بر برتری مدل‌های ترکیبی تأکید کرده‌اند؛ برای مثال، Luo and Gong, (2023) با معرفی مدل ARIMA-WOA-LSTM، نشان دادند که ادغام اجزای خطی^۶، غیرخطی^۷ و الگوریتم بهینه‌سازی نهنگ^۸ نسبت به مدل‌های منفرد و ساده‌تر، دقت و پایداری بالاتری ارائه می‌دهد. همچنین، Kilinc and Yurtsever (2022) با مدل GWO^۹-GRU و Fang et al., (2024) با مدل CNN-Transformer-LSTM، عملکرد برتر مدل‌های ترکیبی را در حوزه‌های مختلف رودخانه‌ای تأیید کردند. شایان ذکر است که همه مدل‌ها با ورودی‌های یکسان، تقسیم‌بندی آموزش/آزمون مشابه و معیارهای ارزیابی یکسان بررسی می‌شوند.

شواهد موجود نشان می‌دهد که استفاده از مدل‌های ترکیبی، به‌ویژه با ادغام روش‌های خطی و غیرخطی، به‌طور مؤثری می‌تواند دقت و پایداری پیش‌بینی‌های

در مطالعات هیدرولوژیکی، مدل‌های آماری خطی نظیر SARIMA^۱ به‌طور گسترده برای تحلیل و پیش‌بینی سری‌های زمانی رواناب به‌کار رفته‌اند. این مدل‌ها به دلیل سادگی، قابلیت تفسیر و توانایی در شناسایی الگوهای فصلی، همچنان مورد توجه هستند (Box et al., 2015; Wang et al., 2022). با این حال، محدودیت اصلی آنها در ناتوانی برای درک روابط غیرخطی و وابستگی‌های پیچیده میان داده‌های هیدرولوژیکی است که می‌تواند دقت پیش‌بینی را کاهش دهد (Luo and Gong, 2023).

ظهور روش‌های داده‌محور یادگیری عمیق دقت پیش‌بینی رواناب را بهبود بخشیده است. مدل‌هایی چون BiLSTM^۲ و GRU^۳ با سازوکارهای حافظه و مکانیسم گیت قادرند روابط غیرخطی و وابستگی‌های زمانی بلندمدت را به‌خوبی مدل‌سازی کنند و در مقایسه با روش‌های سنتی، عملکرد بهتری در پیش‌بینی سری‌های زمانی هیدرولوژیکی ارائه دهند (Mateus et al., 2021; Velarde et al., 2022).

در سال‌های اخیر، مدل‌های ترکیبی به‌عنوان رویکردی نوین در هیدرولوژی مطرح شده‌اند که با ادغام نقاط قوت مدل‌های مختلف، دقت و پایداری پیش‌بینی را بهبود می‌بخشند. برای مثال، BiLSTM-GRU وابستگی‌های زمانی و سرعت همگرایی را همزمان پوشش می‌دهد و مدل‌هایی مانند SARIMA-BiLSTM-GRU قادرند الگوهای فصلی و روابط غیرخطی رواناب را بازنمایی کنند (Luo and Gong, 2023; Workneh and Jha, 2025). امروزه، مدل‌های یادگیری عمیق و ترکیبی به‌طور گسترده در پیش‌بینی رواناب و دیگر متغیرهای هیدرولوژیکی و محیطی به‌کار گرفته شده‌اند. مدل‌های یادگیری عمیق، مانند LSTM، CNN-LSTM^۴ و C-LSTM، توانایی بالایی در شناسایی الگوهای غیرخطی و وابستگی‌های زمانی بلندمدت دارند و نشان داده‌اند که می‌توانند پیش‌بینی‌های دقیق و پایدار در جریان رودخانه و کیفیت آب ارائه دهند (Bai et al., 2020; Baek et al., 2020; Xiang et al., 2020).

⁵ Ant Colony Optimization (ACO)

⁶ Autoregressive integrated Moving Average (ARIMA)

⁷ Long Short-Term Memory (LSTM)

⁸ Whale Optimazation Algorithm (WOA)

⁹ Grey Wolf Optimazer

¹ Seasonal Autoregressive integrated Moving Average

² Bidirectional Long Short-Term Memory

³ Gated Recurrent Unit

⁴ Convolutional Neural Network-Long Short-Term Memory

معادل ۱۰۰ میلیون مترمکعب، در این حوزه جریان دارد (Ghezelsofli et al., 2022).

در این پژوهش، پس از تعیین موقعیت حوزه آبخیز قره‌سو، داده‌های رواناب ماهانه ایستگاه‌های هیدرومتری گردآوری شد. سه ایستگاه هیدرومتری پل توسکاستان، نهارخوران و سیاه‌آب به دلیل برخورداری از آمار قابل‌اعتماد انتخاب شدند. دوره آماری مشترک این ایستگاه‌ها ۳۶ سال (۱۳۶۴-۱۳۶۵ تا ۱۳۹۹-۱۴۰۰) بوده و داده‌ها از شرکت آب منطقه‌ای استان گلستان اخذ شد (جدول ۱).

در تعداد بسیار محدودی از داده‌های گم‌شده، بازسازی با استفاده از روش همبستگی خطی بین ایستگاه‌های مجاور و با ضریب همبستگی قابل قبول انجام شد. در این پژوهش، با توجه به اینکه مقادیر رواناب ماهانه وابسته به زمان هستند، در ابتدا این داده‌ها به صورت سری زمانی مرتب شد. از آنجا که صحت و همگنی داده‌ها نقش اساسی در تحلیل‌های هیدرولوژیکی دارد، پیش از انجام مراحل بعدی تحقیق، همگنی سری‌های زمانی با آزمون ران‌تست مورد بررسی قرار گرفت. سپس، رواناب ماهانه ایستگاه‌های هیدرومتری منتخب با استفاده از پنج مدل شبیه‌سازی شد که شامل سه مدل منفرد (SARIMA، GRU و BiLSTM) و دو مدل ترکیبی (BiLSTM-GRU و SARIMA-BiLSTM-GRU) است. در ذیل به معرفی مدل‌های مورد استفاده در این پژوهش پرداخته می‌شود.

مدل‌های مورد استفاده، مدل‌های باکس و جنکینز (SARIMA): مدل SARIMA یکی از انعطاف‌پذیرترین مدل‌ها برای پیش‌بینی سری‌های زمانی فصلی است و قادر به شناسایی الگوهای پیچیده داده‌های فصلی است. این مدل یک مدل خطی پویا و کلاسیک برای پیش‌بینی سری‌های زمانی است که توانایی بررسی روندهای عادی و فصلی را دارد و با استفاده از تفاضل‌گیری فصلی می‌توان مدل ARMA را بر روی داده‌ها برازش داد (Bowerman and O'Connell, 1993).

هر مدل SARIMA با شش مشخصه تعیین می‌شود که به صورت رابطه (۱) نمایش داده می‌شود (Box et al., 2015).

$$(p, d, q) \times (P, D, Q)_T \quad (1)$$

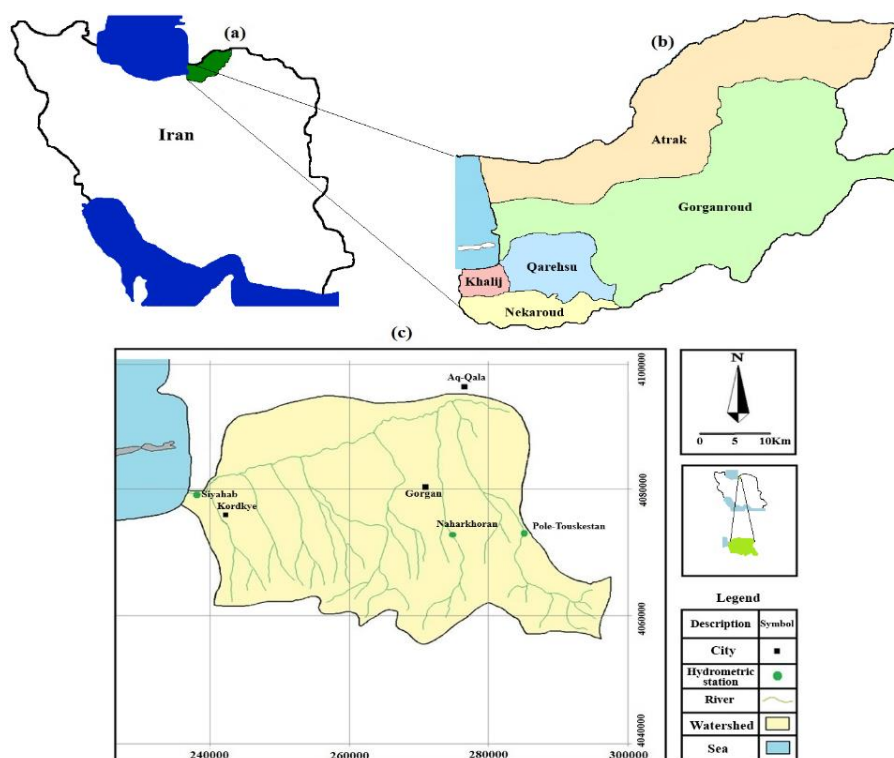
هیدرولوژیکی را ارتقا دهد. این مدل‌ها نسبت به مدل‌های منفرد یا سنتی، گزینه‌ای کارآمدتر برای مدیریت منابع آب، برنامه‌ریزی سدها، کنترل سیلاب و تخصیص بهینه منابع ارائه می‌دهند. به طور کلی، پیشرفت‌های اخیر در زمینه یادگیری عمیق و مدل‌های ترکیبی، افق‌های جدیدی برای مطالعات پیش‌بینی رواناب و تصمیم‌گیری‌های مدیریت منابع آب فراهم کرده است.

نوآوری اصلی این مطالعه در ترکیب متوالی سه‌گانه SARIMA، BiLSTM و GRU به صورت یک چارچوب یکپارچه برای مدل‌سازی رواناب ماهانه است. این معماری ترکیبی خاص که در مطالعات پیشین هیدرولوژی به‌ویژه برای داده‌های حوزه آبخیز قره‌سو بررسی نشده است، امکان استخراج مؤثر الگوهای خطی، غیرخطی و وابستگی‌های زمانی کوتاه و بلندمدت را فراهم کرده و عملکرد پیش‌بینی را به طور چشمگیری بهبود می‌بخشد.

در این پژوهش، با تمرکز بر حوزه آبخیز قره‌سو در استان گلستان، پنج مدل شامل سه مدل منفرد (SARIMA، GRU و BiLSTM) و دو مدل ترکیبی (BiLSTM-GRU و SARIMA-BiLSTM-GRU) برای پیش‌بینی رواناب ماهانه توسعه و مقایسه شده‌اند. انتخاب این حوزه به دلیل ویژگی‌های هیدرولوژیکی خاص، اهمیت آن در تأمین منابع آب استان و حساسیت آن به تغییرات اقلیمی و هیدرولوژیکی صورت گرفته است. هدف اصلی تحقیق، ارزیابی کارایی مدل‌های خطی و یادگیری عمیق در پیش‌بینی رواناب، شناسایی نقاط قوت و ضعف هر یک و معرفی یک مدل بهینه برای مدیریت منابع آب در مقیاس حوضه است.

مواد و روش‌ها

منطقه مورد مطالعه: حوزه آبخیز قره‌سو با مساحت ۱۶۲۴ کیلومتر مربع در شمال ایران و شمال غرب استان گلستان واقع شده است (شکل ۱). این حوزه از غرب به آبخیز خلیج گرگان، از شمال و شرق به آبخیز گرگانرود و دریای خزر و از جنوب به آبخیز نکارود محدود می‌شود. بر اساس میانگین بلندمدت تا سال ۱۳۹۸، حدود هشت درصد از منابع آب سطحی استان گلستان،



شکل ۱- موقعیت جغرافیایی حوزه آبخیز قره‌سو (c) در ایران (a) و استان گلستان (b)

Fig. 1. Geographical location of the Qarehsu watershed (c) in Iran (a) and Golestan Province (b)

جدول ۱- مشخصات ایستگاه‌های هیدرومتری منتخب در حوزه آبخیز قره‌سو

Table 1. Specifications of the selected hydrometric stations in the Qarehsu watershed

Station Name	Station Code	Longitude (UTM)	Latitude (UTM)	Elevation (m)
Pole-Touskestan	12-085	284272	4073485	465
Naharkhoran	12-043	274622	4071771	500
Siyahab	12-097	237291	4079672	-26

کرد که شامل تحلیل سری زمانی و تعیین ویژگی‌های بهینه برای هر دو مدل است (Hyndman, 2018). Zhang, 2003, Hyndman and Khandakar, 2008). برای ارزیابی کیفیت برازش مدل SARIMA معمولاً از آماره Q باکس-پیرس و فرم‌های اصلاح‌شده آن استفاده می‌شود که به صورت رابطه (۲) محاسبه می‌شود (Bowerman and O'Connell, 1993).

$$Q = n \sum_{i=1}^k r_i^2(Z_i, Z_i) \quad (2)$$

در رابطه (۲)، $r_i(Z_t, Z_t)$ تابع خودهمبستگی سری زمانی باقیمانده را در گام یا تاخیر i ام نشان می‌دهد. آماره Q دارای توزیع مجانبی $\chi^2(k)$ است. این آماره بر این اساس شکل گرفته که سری زمانی باقیمانده تا چه میزان تبدیل به نوفه سفیدشده که نشان‌دهنده برازش مناسب مدل است (Box et al., 2015). مدل نهایی از بین طیفی از مدل‌های با مرتبه‌های ممکن، بر پایه کمترین مقدار AIC و همراهی با نتایج قابل قبول

در مدل SARIMA، ویژگی‌های d و D به ترتیب نشان‌دهنده مرتبه تفاضلی کردن معمولی و فصلی سری زمانی، p و q به ترتیب مرتبه خودبازگشتی و میانگین متحرک معمولی و P و Q نیز مرتبه خودبازگشتی و میانگین متحرک فصلی با دوره تناوب T مدل را تعیین می‌کند. با استفاده از آزمون ADF، ایستایی سری بررسی شد. شیوه کلی مدل‌سازی SARIMA شامل حذف روند و روند فصلی در گام نخست و سپس مدل‌سازی بر روی مقادیر باقیمانده است.

برای حذف روند و نامابستگی فصلی، مراتب تفاضل‌گیری معمولی (d) و فصلی (D) با بررسی ACF تعیین شد. برای شناسایی مراتب (Q, P, q, p) پس از ایستاسازی، بر اساس الگوهای ACF و PACF تعیین شد. مدل نهایی بر اساس معیار اطلاعاتی آکائیک (AIC) انتخاب شد. برای شناسایی و انتخاب متغیرهای مناسب در مدل SARIMA، می‌توان از مدل ARIMA استفاده

ارائه می‌دهد (Cho, 2014). اگرچه GRU به‌طور کامل معادل LSTM نیست، اما به‌عنوان شکل ساده‌شده آن در بسیاری از کاربردها عملکردی قابل مقایسه و در برخی موارد کارا تر دارد.

علاوه بر این، ترکیب مدل‌های GRU با ساختارهای LSTM و BiLSTM در مطالعات اخیر نتایج دقیق‌تر و بهتری را در مدلسازی داده‌های دنباله‌ای و پیش‌بینی سری‌های زمانی نشان داده است (Jordan et al., 2021; Mateus et al., 2022; Velarde et al., 2021). پژوهش حاضر، استفاده از GRU به‌دلیل توانایی آن در یادگیری وابستگی‌های بلندمدت با هزینه محاسباتی کمتر نسبت به LSTM در دستور کار قرار گرفت.

مدل ترکیبی BiLSTM-GRU: مدل BiLSTM-GRU از ترکیب شبکه‌های عصبی BiLSTM و GRU ساخته شده و عملکرد بهتری نسبت به مدل‌های منفرد ارائه می‌دهد (Jordan et al., 2021; Mateus et al., 2021; Velarde et al., 2022). مراحل اصلی پیاده‌سازی شامل موارد زیر است.

پیش‌پردازش داده‌ها: حذف مقادیر پرت و داده‌های مشکل‌دار و بازسازی داده‌های گم‌شده.

نرمال‌سازی: مقیاس‌گذاری داده‌ها یا استانداردسازی برای بهبود همگرایی مدل.

ساخت توالی‌های ورودی و خروجی: داده‌های سری زمانی به پنجره‌های متحرک تقسیم می‌شوند تا مدل الگوهای زمانی را یاد گرفته و مقادیر آینده را پیش‌بینی کند.

تقسیم داده‌ها: بر اساس اصول رایج در مدلسازی داده‌های سری‌زمانی، داده‌ها بر اساس زمان به مجموعه‌های آموزشی (۷۰ درصد)، ارزیابی (۱۵ درصد) و آزمایش (۱۵ درصد) تقسیم می‌شوند. تمامی داده‌های مجموعه آموزشی مربوط به بازه زمانی قبل از داده‌های مجموعه ارزیابی و آزمایش بدون هم‌پوشانی زمانی هستند که هیچ نشت اطلاعاتی از آینده به گذشته رخ ندهد.

تنظیم طول توالی: طول توالی باید کافی باشد تا اطلاعات مهم را پوشش دهد و از مشکلاتی مانند کاهش سرعت آموزش جلوگیری شود.

آزمون‌های تشخیصی خودهمبستگی و نرمال بودن باقی‌مانده‌ها انتخاب شده است.

مدل شبکه‌های عصبی حافظه کوتاه‌مدت ماندگار

دوطرفه (BiLSTM): شبکه‌های عصبی حافظه کوتاه‌مدت ماندگار (LSTM) یکی از مدل‌های پیشرفته در یادگیری عمیق هستند که به‌طور ویژه برای مدلسازی داده‌های ترتیبی و سری‌های زمانی توسعه یافته‌اند. این مدل‌ها برای غلبه بر محدودیت شبکه‌های عصبی بازگشتی (RNN) طراحی شده‌اند که عمدتاً با مشکل ناپایداری گرادیان و فراموشی اطلاعات بلندمدت مواجه هستند. ساختار LSTM شامل سه دروازه اصلی دروازه ورودی، دروازه خروجی و دروازه فراموشی است. این دروازه‌ها با تنظیم جریان اطلاعات، امکان یادگیری وابستگی‌های کوتاه‌مدت و بلندمدت را فراهم می‌کنند (De la Fuente et al., 2024). پنجره زمانی در این مدل به صورت ۱+۱۲ انجام شد یعنی داده‌های هر ۱۲ ماه از ابتدا انتخاب و یک ماه بعدی به‌عنوان پیش‌بینی تقسیم شد. این پنجره زمانی لغزان به همین صورت دادگان می‌لغزد و دادگان تفکیک می‌کند.

مدل BiLSTM نسخه‌ای توسعه‌یافته از LSTM است که شامل دو لایه LSTM مجزا بوده و داده‌ها را هم از گذشته به آینده و هم از آینده به گذشته پردازش می‌کند. این ساختار امکان استخراج وابستگی‌های زمانی در هر دو جهت را فراهم می‌کند و موجب بهبود دقت پیش‌بینی در مسائل پیچیده مانند پردازش زبان طبیعی، تحلیل متن‌های طولانی و پیش‌بینی سری‌های زمانی می‌شود (Schuster and Paliwal, 1997; Nacer et al., 2023). به‌طور خلاصه، خروجی نهایی BiLSTM به‌صورت ترکیب خروجی‌های دو مسیر جلو و عقب محاسبه می‌شود (رابطه ۳).

$$[\overline{t^h}, \overline{t^h}] = t^h \quad (3)$$
 که در آن، $\overline{t^h}$ خروجی LSTM در جهت رو به جلو و $\overline{t^h}$ خروجی LSTM در جهت معکوس است.

مدل واحد بازگشتی دردار (GRU): مدل واحد بازگشتی دردار (GRU) نوعی شبکه عصبی بازگشتی (RNN) است که برای پردازش داده‌های ترتیبی طراحی شده است. این مدل به‌عنوان نسخه‌ای ساده‌تر و بهینه‌شده نسبت به LSTM معرفی شد و با ادغام گیت‌های ورودی و فراموشی، ساختاری فشرده‌تر

متفاوت تقسیم کرده و امکان جداسازی رفتارهای خطی و غیرخطی را فراهم می‌کند. در این پژوهش برای تفکیک مؤلفه‌های شبه‌خطی مانند روند و هارمونیک‌های ساده از مؤلفه‌های غیرخطی-غیرایستا مانند نوسانات پیچیده، آشوب، تغییرات دینامیکی در یک سری زمانی با استفاده از CEEMDAN، را می‌توان از یک الگوریتم ساخت یافته به‌دست آورد.

این الگوریتم بر اساس اصول تحلیل تجربی مود (EMD) و ویژگی‌های تحلیلی هیلبرت-هوانگ (HHT) استوار است. پس از اعمال تبدیل هیلبرت-هوانگ (HHT) بر روی هر IMF، از تحلیل طیف توان لحظه‌ای^۴ استفاده شده است. مؤلفه‌هایی که طیف توان آنها در طول زمان پایدار و دارای پهنای باند باریک است، به‌عنوان مؤلفه‌های خطی در نظر گرفته شده و برای مدل‌سازی با SARIMA انتخاب شدند. در مقابل، IMFهایی با طیف توان ناپایدار و پراکنده (دارای پهنای باند گسترده و تغییرات فرکانسی بالا) به‌عنوان مؤلفه‌های غیرخطی شناسایی و به مدل-BiLSTM-GRU وارد شدند (Sarkar et al., Li and Yang, 2023). (Luo and Gong, 2023, 2022).

در نهایت، پیش‌بینی‌های هر دو بخش با یکدیگر ترکیب شده و خروجی نهایی سری زمانی به دست می‌آید. این روش ترکیبی، از مزایای هر دو رویکرد کلاسیک و یادگیری عمیق بهره می‌برد و باعث افزایش دقت پیش‌بینی سری‌های زمانی پیچیده می‌شود. در این پژوهش از روش ترکیب وزنی خروجی‌ها برای ادغام مدل‌ها استفاده شده است. به این صورت که مدل‌های SARIMA و BiLSTM-GRU به‌طور مستقل آموزش دیده‌اند. سپس، پیش‌بینی نهایی از طریق مجموع وزنی دو خروجی محاسبه می‌شود. شایان ذکر است برای انجام این پژوهش از زبان برنامه‌نویسی پایتون، کتابخانه تانسورفلو، واسط برنامه‌نویسی Keras و کتابخانه‌های موردنیاز مانند sklearn، statsmodels، scipy، matplotlib، numpy، pandas، pyEMD و غیره استفاده شده است.

اعتبارسنجی مدل‌ها: در مرحله بعد، فرایند اعتبارسنجی مدل‌های مورد استفاده انجام شد. به‌منظور

بهینه‌سازی مدل: شامل جلوگیری از بیش‌برازش با روش‌هایی مانند توقف زودهنگام^۱ با مانیتورینگ خطای داده‌های اعتبارسنجی و توقف آموزش در صورت عدم بهبود پس از یک گام مشخص، میرایی یا لایه‌های Dropout با نرخ‌های بهینه‌شده پس از لایه‌های بازگشتی و کاملاً متصل، نرمال‌سازی دسته‌ای^۲ برای افزایش پایداری و سرعت همگرایی آموزش، تنظیم وزن‌ها^۳ و همچنین تعیین بهترین ابرمتغیرها (نرخ یادگیری، تعداد تکرار، اندازه بسته‌ها، نوع توابع فعال‌ساز و بهینه‌ساز) با استفاده از جستجوی شبکه‌ای یا الگوریتم‌های بهینه‌سازی انجام شد.

در پژوهش حاضر، مدل ترکیبی به‌صورت سری متوالی طراحی شده است. به این ترتیب که ابتدا لایه BiLSTM و سپس لایه GRU قرار گرفته‌اند. هدف از این معماری، استخراج عمیق‌تر ویژگی‌های وابسته به گذشته و آینده توسط BiLSTM و سپس پردازش مؤثرتر این ویژگی‌های استخراج‌شده و کاهش پیچیدگی محاسباتی در مرحله بعد توسط GRU است. خروجی نهایی مدل از آخرین حالت پنهان لایه GRU گرفته می‌شود.

مدل ترکیبی SARIMA-BiLSTM-GRU: ترکیب مدل‌های کلاسیک سری زمانی مانند SARIMA با مدل‌های یادگیری عمیق نظیر BiLSTM و GRU می‌تواند به شکل مؤثری دقت پیش‌بینی را افزایش دهد. در این رویکرد، مدل SARIMA برای مدل‌سازی مؤلفه‌های خطی سری زمانی به کار می‌رود، در حالی که مدل‌های BiLSTM و GRU توانایی شناسایی و بازنمایی الگوهای پیچیده و غیرخطی را دارند.

برای اجرای این روش، ابتدا سری زمانی با استفاده از روش تجزیه مود تجربی^۴ کامل همراه با نویز تطبیقی^۵ به مجموعه‌ای از توابع مود ذاتی (IMF) تفکیک می‌شود. این توابع، نوسان‌های موجود در یک موج یا سیگنال غیرایستا و غیرخطی را به‌صورت مولفه‌هایی با فرکانس‌های مختلف نشان می‌دهند. روش CEEMDAN که نسخه بهبود یافته روش EMD است، سیگنال پیچیده را به مؤلفه‌هایی با ویژگی‌های نوسانی

^۴ Empirical Mode Decomposition

^۵ CEEMDAN

^۶ Instantaneous Power Spectrum

^۱ Early Stopping

^۲ Batch Normalization

^۳ Weight Regularization

تمام ایستگاه‌ها بود که به دلیل بیرون‌زدگی میله‌ها از حدود اطمینان تشخیص داده شد. بنابراین، جهت ایستاسازی داده‌ها از تفاضل‌گیری فصلی بهره گرفته شد. تفاوت در مرتبه تفاضل‌گیری مستقیماً ناشی از ماهیت متفاوت سری‌های زمانی رواناب در هر ایستگاه بوده است. برای ایستگاه سیاه‌آب، آزمون‌های تشخیصی ADF و ACF نشان دادند که سری اصلی داده‌ها از نظر آماری مانا است و ضرورتی برای اعمال تفاضل‌گیری وجود نداشت.

پس از آن، مدل‌های مختلف SARIMA برای داده‌های رواناب ماهانه برازش و ارزیابی شدند (جدول ۲). معیار پذیرش مدل، بزرگ‌تر بودن مقدار p-value از ۰/۰۵ در آزمون مانایی باقی‌مانده‌ها بود. همان‌گونه که در جدول ۲ مشاهده می‌شود، کلیه مدل‌ها این شرط را برآورده کردند و در نتیجه، می‌توان اظهار داشت که مدل‌های انتخاب‌شده توانسته‌اند سری‌های زمانی رواناب را به‌طور مناسب برازش نمایند.

سپس مدل SARIMA برای پیش‌بینی رواناب ماهانه در دوره ۱۲ ماهه آتی در ایستگاه‌های هیدرومتری منتخب به‌کار گرفته شد (شکل ۳). نتایج نشان داد مدل SARIMA در هر سه ایستگاه توانسته الگوی کلی رواناب ماهانه را شبیه‌سازی کند، هرچند دقت آن در ایستگاه نهارخوران بیشتر و در ایستگاه پل توسکاستان در برخی ماه‌ها با خطای بیش‌برآورد همراه بوده است.

ارزیابی دقت و قابلیت اعتماد مدل‌ها، از سه شاخص ارزیابی رایج و معتبر شامل ریشه میانگین مربعات خطا (RMSE)، میانگین قدر مطلق خطا (MAD) و میانگین مربعات خطا (MSE) استفاده شد (رابطه‌های ۴ تا ۶). این معیارها به‌دلیل کاربرد گسترده در ارزیابی عملکرد مدل‌های شبیه‌سازی و پیش‌بینی، امکان مقایسه نتایج را فراهم می‌کنند (Moriasi et al., 2007).

$$MSE = \frac{\sum_{i=1}^m (Y_t - \hat{Y}_t)^2}{m} \quad (4)$$

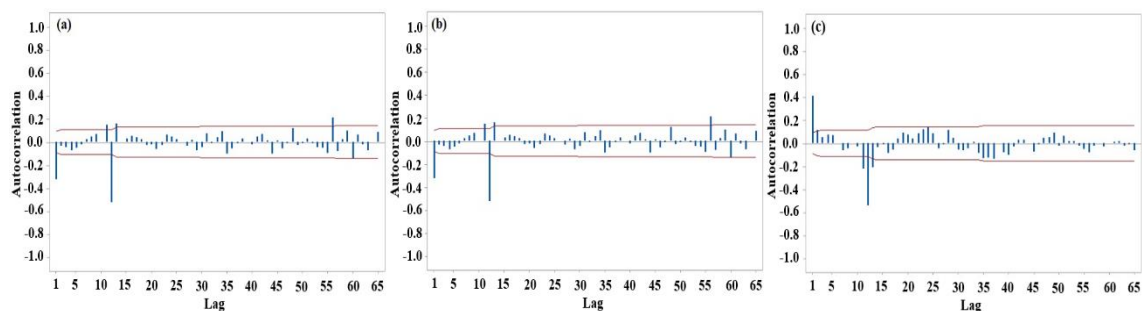
$$RMSE = \sqrt{\frac{\sum_{i=1}^m (Y_t - \hat{Y}_t)^2}{m}} \quad (5)$$

$$MAD = \frac{1}{m} \sum_{i=1}^m |Y_t - \hat{Y}_t| \quad (6)$$

در روابط فوق، \hat{Y}_t مقادیر برازش‌شده و یا پیش‌بینی‌شده توسط مدل، Y_t مقدار سری زمانی مشاهده‌شده در زمان t و m تعداد مقادیر برازش‌شده یا مقادیر پیش‌بینی است.

نتایج و بحث

نتایج مدل‌سازی خطی SARIMA: بررسی سری‌های زمانی رواناب ماهانه نشان داد که در ایستگاه‌های هیدرومتری پل توسکاستان و نهارخوران روندی نزولی وجود دارد، در حالی که در ایستگاه سیاه‌آب روند معنی‌داری مشاهده نشد. به‌منظور حذف روند، از تفاضل‌گیری مرتبه اول استفاده شد. نتایج تحلیل خودهمبستگی (شکل ۲) بیانگر وجود خودهمبستگی معنی‌دار و همچنین الگوی فصلی با تناوب ۱۲ ماهه در



شکل ۲- نمودارهای خودهمبستگی مقادیر رواناب ماهانه در ایستگاه‌های هیدرومتری پل توسکاستان (a)، نهارخوران (b) و سیاه‌آب (c)
 Fig. 2. Autocorrelation plots of monthly runoff at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b), and Siahah (d)

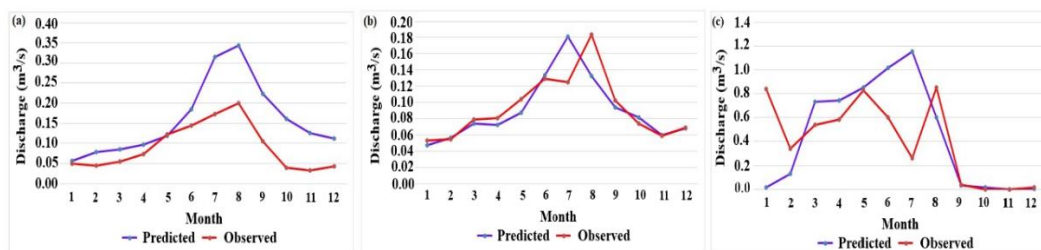
جدول ۲- نتایج مدل‌سازی SARIMA برای سری‌های زمانی رواناب ماهانه در ایستگاه‌های هیدرومتری منتخب

Table 2. SARIMA modeling results for monthly streamflow time series at selected hydrometric stations

Station Name	SARIMA Model	p-value
Pole-Touskestan	$(2,1,2) \times (1,0,1)_{12}$	0.635
Naharkhoran	$(1,1,1) \times (1,0,2)_{12}$	0.752
Siyahab	$(0,0,2) \times (1,0,1)_{12}$	0.506

کاهش می‌یابد. همچنین مدل روند کلی تغییرات را در ایستگاه سیاه‌آب بازتولید کرده اما در برخی ماه‌ها اختلاف قابل توجهی میان مقادیر مشاهده‌ای و برآوردی وجود دارد. این موضوع می‌تواند به دلیل غیرخطی بودن یا تاثیرپذیری داده‌ها از عوامل محلی مانند برداشت بی‌رویه آب از رودخانه باشد.

ایستگاه‌های واقع در سرشاخه مانند نهارخوران و پل توسکاستان کمتر تحت تأثیر تجمع جریان‌ها و نوسانات پیچیده قرار دارند، در حالی که ایستگاه سیاه‌آب در خروجی حوضه واقع شده و رواناب آن تحت تأثیر تجمع جریان‌ها، تغییرات مکانی بارش و برداشت آب کشاورزی در پایین‌دست است که می‌تواند دقت پیش‌بینی مدل



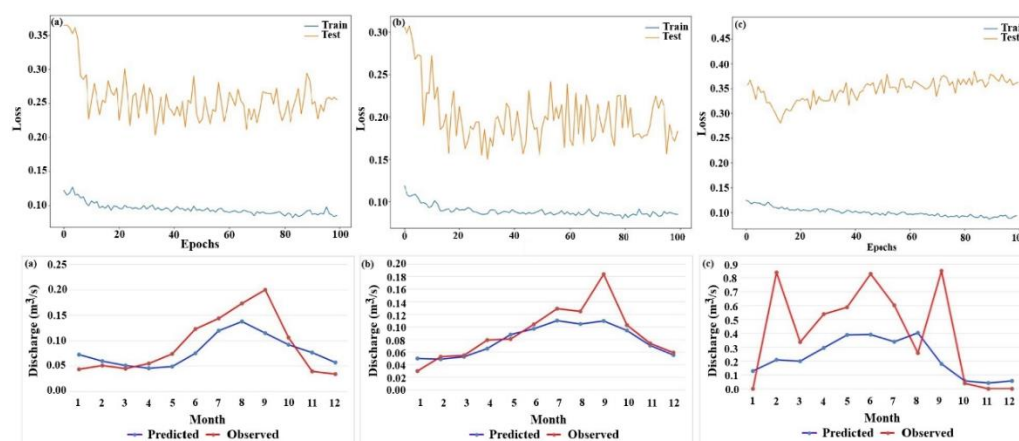
شکل ۳- مقایسه مقادیر رواناب ماهانه مشاهده‌ای و برآورد شده با استفاده از مدل SARIMA در ایستگاه‌های هیدرومتری:

پل توسکاستان (a)، نهارخوران (b) و سیاه‌آب (c)

Fig. 3. Comparison of observed and predicted monthly runoff values using the SARIMA model at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b) and Siahab (c)

در ایستگاه نهارخوران نیز هرچند هر دو تابع کاهش یافته و نزدیک به هم قرار می‌گیرند، اما همچنان نشانه‌هایی از بیش‌برازش دیده می‌شود. در ایستگاه سیاه‌آب نیز اتلاف آموزش کاهش یافته ولی منحنی اعتبارسنجی با نوسان همراه است که باز هم حاکی از بیش‌برازش است. به‌طور کلی می‌توان نتیجه گرفت که مدل BiLSTM در اکثر ایستگاه‌ها دچار بیش‌برازش شده و تعمیم‌پذیری محدودی دارد.

نتایج مدل غیرخطی BiLSTM: تفاوت بین منحنی یادگیری داده‌های Train و Test نشان‌دهنده سطح تعمیم‌پذیری مدل است. بررسی منحنی یادگیری مدل BiLSTM نشان داد، در ایستگاه پل توسکاستان، تابع اتلاف آموزش^۱ سریعاً کاهش یافته اما تابع اعتبارسنجی^۲ پس از کاهش اولیه دچار نوسانات شدید شده و در سطحی نسبتاً بالا ثابت می‌ماند که بیانگر بیش‌برازش است (شکل ۴، قسمت بالا).



شکل ۴- منحنی‌های یادگیری و مقایسه مقادیر رواناب ماهانه مشاهده‌ای و برآورد شده با استفاده از مدل BiLSTM در ایستگاه‌های

هیدرومتری: پل توسکاستان (a)، نهارخوران (b) و سیاه‌آب (c)

Fig. 4. Learning Curves and comparison of observed and predicted monthly runoff values using the BiLSTM model at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b) and Siahab (c)

² Test Loss

¹ Train Loss

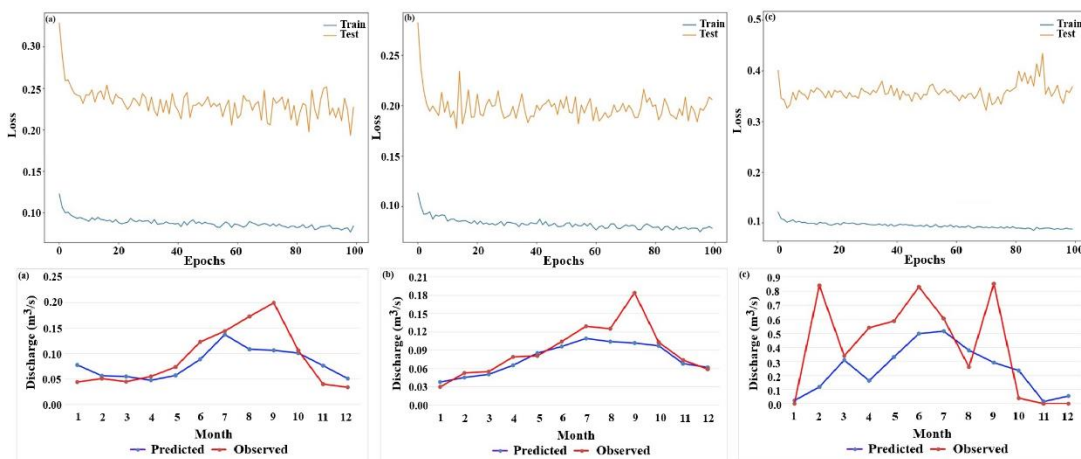
نشان‌دهنده‌ی بیش‌برازش است؛ یعنی مدل روی داده‌های آموزشی خوب یاد گرفته، اما در تعمیم به داده‌های جدید دقت کمتری دارد. در ایستگاه سیاه‌آب (شکل ۵c) شدت بیش‌برازش بیشتر است، زیرا مقدار اتلاف آموزش نه تنها کاهش نیافته بلکه روند افزایشی و نوسانی دارد. مقایسه نتایج GRU و BiLSTM نشان داد که الگوهای کلی هر دو مدل تقریباً مشابه هستند، اما شدت بیش‌برازش در مدل GRU اندکی بیشتر ارزیابی شد.

نتایج پیش‌بینی رواناب ماهانه با استفاده از مدل GRU برای ۱۲ ماه آینده (شکل ۵، قسمت پایین) نشان داد که در هر سه ایستگاه، مدل توانسته روند کلی تغییرات ماهانه رواناب را به‌طور نسبی بازتولید کند، اگرچه همانند مدل BiLSTM، در برخی ماه‌ها و به‌ویژه در دوره‌های پیک جریان، اختلاف‌هایی بین مقادیر مشاهده‌شده و پیش‌بینی‌شده مشاهده شد. علاوه بر این، عملکرد مدل در ایستگاه‌های بالادست (پل توسکستان و نهارخوران) بهتر از ایستگاه سیاه‌آب در خروجی حوزه بود. به‌طور کلی، نتایج نشان می‌دهد که مدل GRU توانایی مناسبی در بازتولید الگوی رواناب دارد، اما به‌دلیل بروز بیش‌برازش در برخی ایستگاه‌ها، دقت پیش‌بینی آن محدود بوده و تا حدی وابسته به ویژگی‌های مکانی و داده‌های ورودی است.

شکل ۴، در قسمت پایین مقادیر مشاهده‌ای و پیش‌بینی‌شده رواناب ماهانه طی ۱۲ ماه آینده را نشان می‌دهد. نتایج بیانگر آن است که مدل در ایستگاه‌های پل توسکستان و نهارخوران توانسته روند کلی تغییرات رواناب را نسبتاً خوب بازتولید کند، هرچند در برخی ماه‌ها با کم‌برآوردی همراه بوده است. در ایستگاه سیاه‌آب اختلاف بیشتری میان مقادیر مشاهده‌ای و برآوردی دیده می‌شود که نشان‌دهنده دقت کمتر مدل در خروجی حوزه است.

این موضوع بیانگر آن است که داده‌های این ایستگاه احتمالاً تحت تأثیر عوامل پیچیده‌ای مانند تغییرات کاربری اراضی، فعالیت‌های شدید انسانی، وابستگی زیاد به جریان‌های و با وجود نوفه‌های غیرسیستماتیک قوی قرار دارند که در ورودی‌های مدل لحاظ نشده‌اند. به‌طور کلی، مدل BiLSTM در پیش‌بینی الگوهای زمانی رواناب موفق عمل کرده اما در برخی ایستگاه‌ها نیاز به بهبود دقت دارد.

نتایج مدل GRU: بررسی منحنی‌های یادگیری مدل GRU در ایستگاه‌های منتخب حوزه آبخیز قره‌سو (شکل ۵، قسمت بالا) نشان داد که در هر سه نمودار، مقدار تابع اتلاف آموزش سریع کاهش یافته و در سطح پایینی تثبیت شده است. در مقابل، اتلاف اعتبارسنجی نسبتاً بالا باقی مانده و نوسانات زیادی دارد. این الگو



شکل ۵- منحنی‌های یادگیری و مقایسه مقادیر رواناب ماهانه مشاهده‌ای و برآورد شده با استفاده از مدل GRU در ایستگاه‌های هیدرومتری:

پل توسکستان (a)، نهارخوران (b) و سیاه‌آب (c)

Fig. 5. Learning Curves and comparison of observed and predicted monthly runoff values using the GRU model at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b) and Siahab (c)

کم‌برآوردی همراه بوده و این خطا در ایستگاه پایین‌دست (سیاه‌آب) محسوس‌تر است.

نتایج مدل ترکیبی SARIMA-BiLSTM-GRU

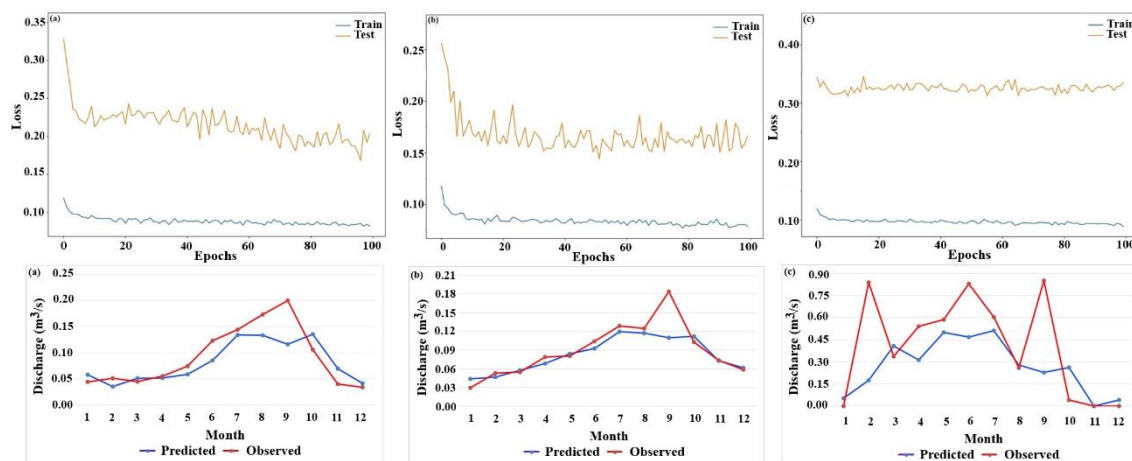
شکل ۷ (قسمت بالا) منحنی‌های یادگیری مدل ترکیبی SARIMA-BiLSTM-GRU را برای ایستگاه‌های هیدرومتری منتخب در حوزه آبخیز قره‌سو نشان می‌دهد. کاهش سریع تابع اتلاف در دوره‌های ابتدایی بیانگر توانایی بالای مدل در شناسایی الگوهای اصلی داده‌ها و بهبود تدریجی عملکرد آن است.

در تمامی ایستگاه‌ها، مقادیر اتلاف آموزش و اعتبارسنجی از همان ابتدا بسیار به یکدیگر نزدیک بوده و پس از چند دوره به سطحی پایین و نسبتاً پایدار همگرا می‌شوند. این همگرایی و نزدیکی خطوط بیانگر یادگیری کامل الگوهای داده‌ای بدون بروز بیش‌برازش است. به بیان دیگر، مدل علاوه بر دقت بالا در داده‌های آموزشی، توانایی تعمیم‌پذیری مناسبی نسبت به داده‌های جدید نیز داشته است. این نتایج نشان می‌دهد که مدل ترکیبی SARIMA-BiLSTM-GRU قادر به شناسایی مؤثر ویژگی‌های غیرخطی سری‌های زمانی هیدرولوژیکی بوده و بر روی داده‌های ناشناخته نیز عملکرد مطلوبی دارد.

نتایج مدل ترکیبی BiLSTM-GRU: شکل ۶

(قسمت بالا) منحنی‌های یادگیری مدل ترکیبی BiLSTM-GRU را برای ایستگاه‌های هیدرومتری منتخب در حوزه آبخیز قره‌سو نشان می‌دهد. نتایج نشان می‌دهد در ایستگاه‌های هیدرومتری بالادست (پل توسکاستان و نهارخوران) مدل به سرعت همگرا شده و دقت آموزش بالا بوده است، هرچند نوسانات تابع اتلاف اعتبارسنجی بیانگر بیش‌برازش نسبی است. در ایستگاه سیاه‌آب، با وجود پایین بودن خطای آموزش، خطای اعتبارسنجی همراه با نوسان بوده که نشان‌دهنده بیش‌برازش خفیف است. در مجموع، رفتار مدل ترکیبی مشابه مدل‌های منفرد BiLSTM و GRU است و بهبود چشمگیری در الگوی یادگیری دیده نمی‌شود.

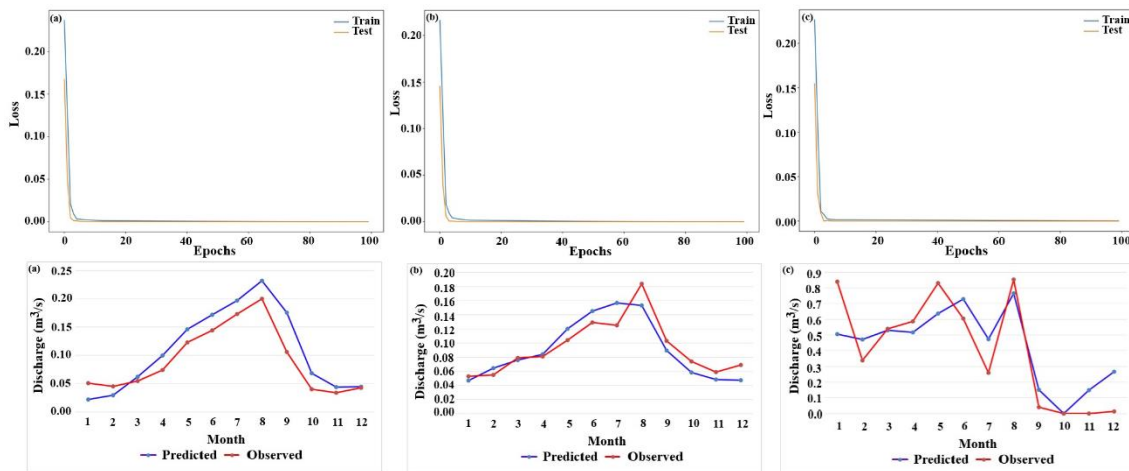
شکل ۶ (قسمت پایین) نتایج پیش‌بینی رواناب ماهانه را طی ۱۲ ماه آینده نشان می‌دهد. نتایج نشان داد که مدل ترکیبی BiLSTM-GRU در ایستگاه‌های بالادست (ناهارخوران و پل توسکاستان) عملکرد بهتری نسبت به ایستگاه سیاه‌آب واقع در خروجی حوزه داشته و روند تغییرات را نسبتاً خوب بازنمایی کرده است. با این حال، مشابه مدل‌های منفرد، در بیشتر ماه‌ها با



شکل ۶- منحنی‌های یادگیری و مقایسه مقادیر رواناب ماهانه مشاهده‌ای و برآورد شده با استفاده از مدل BiLSTM-GRU در ایستگاه‌های

هیدرومتری: پل توسکاستان (a)، نهارخوران (b) و سیاه‌آب (c)

Fig. 5. Learning Curves and comparison of observed and predicted monthly runoff values using the BiLSTM-GRU model at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b) and Siahhab (c)



شکل ۷- منحنی‌های یادگیری و مقایسه مقادیر رواناب ماهانه مشاهده‌ای و برآورد شده با استفاده از مدل SARIMA-BiLSTM-GRU در ایستگاه‌های هیدرومتری: پل توسکاستان (a)، نهارخوران (b) و سیاه‌آب (c)

Fig. 7. Learning Curves and comparison of observed and predicted monthly runoff values using the SARIMA-BiLSTM-GRU model at hydrometric stations: Pole-Touskestan (a), Naharkhoran (b) and Siahab (c)

مدل ترکیبی BiLSTM-GRU به‌عنوان دومین مدل برتر در پیش‌بینی رواناب ماهانه معرفی شد که مقدار RMSE آن در ایستگاه‌های پل توسکاستان، نهارخوران و سیاه‌آب به ترتیب 0.326 ، 0.226 و 0.313 مترمکعب بر ثانیه بوده است. به‌طور کلی، مدل‌های ترکیبی در پیش‌بینی رواناب ماهانه عملکرد بهتری نسبت به مدل‌های ساده داشتند.

پس از مدل‌های ترکیبی، دو مدل منفرد BiLSTM و GRU به ترتیب بیشترین دقت را در پیش‌بینی رواناب ماهانه داشتند و عملکرد آنها بسیار نزدیک به یکدیگر بود. هرچند مدل ترکیبی BiLSTM-GRU بهبود اندکی در نتایج ایجاد کرد، اما از نظر دقت تفاوت قابل توجهی با این دو مدل ساده نشان نداد. در نهایت، مدل خطی SARIMA به‌عنوان ضعیف‌ترین روش در مدل‌سازی رواناب ماهانه شناسایی شد.

مقدار RMSE آن در ایستگاه‌های هیدرومتری پل توسکاستان، نهارخوران و سیاه‌آب به ترتیب 0.0851 ، 0.230 و 0.3892 مترمکعب بر ثانیه بوده است. عملکرد قابل توجه مدل ساده SARIMA در ایستگاه نهارخوران ($RMSE=0.0230$) که بسیار نزدیک به مدل ترکیبی برتر ($RMSE=0.0173$) است نشان می‌دهد که در نقاط با شرایط هیدرولوژیک کنترل‌شده‌تر و خطی‌تر مثل نهارخوران، مدل‌های ساده‌تر به دلیل غالب بودن مؤلفه‌های خطی و فصلی و همچنین کمتر بودن اثر مؤلفه‌های غیرخطی پیچیده می‌توانند بسیار رقابتی

شکل ۷ (قسمت پایین) مقادیر رواناب ماهانه مشاهده‌ای و پیش‌بینی‌شده توسط مدل SARIMA-BiLSTM-GRU را طی ۱۲ ماه آینده نشان می‌دهد. نتایج حاکی از آن است که در ایستگاه نهارخوران، همپوشانی بیشتری میان مقادیر مشاهده‌ای و برآوردی نسبت به سایر ایستگاه‌ها وجود دارد که نشان‌دهنده عملکرد دقیق‌تر مدل در این ایستگاه است. همچنین بررسی‌ها نشان داد که اگرچه مدل ترکیبی توانسته روند تغییرات داده‌ها را به‌طور مناسب شناسایی کند، اما در ایستگاه پل توسکاستان پیش‌بینی‌ها عمدتاً با بیش‌برآوردی همراه بوده است.

نتایج اعتبارسنجی مدل‌ها: بر اساس مقایسه پنج مدل مورد استفاده برای پیش‌بینی رواناب ماهانه، مدل ترکیبی SARIMA-BiLSTM-GRU به‌عنوان مناسب‌ترین مدل انتخاب شد که در هر سه ایستگاه هیدرومتری کمترین خطای محاسباتی را داشته است (جدول ۳). مقدار RMSE این مدل در ایستگاه‌های هیدرومتری پل توسکاستان، نهارخوران و سیاه‌آب به ترتیب 0.295 ، 0.173 و 0.1683 مترمکعب بر ثانیه بوده است. در این مدل، SARIMA برای شبیه‌سازی مؤلفه‌های خطی و شبکه‌های BiLSTM و GRU برای شناسایی الگوهای پیچیده و غیرخطی به کار گرفته شده‌اند که در نهایت موجب بهبود پیش‌بینی رواناب شده است.

هستند. این نتایج بر اهمیت مقیاس و ویژگی‌های فیزیکی خاص محل در انتخاب رویکرد بهینه مدل‌سازی تأکید دارد.

عمل کنند. در مقابل، در مقیاس کل حوضه (سیاه‌آب) به دلیل افزایش پیچیدگی سیستم، حتی مدل‌های پیشرفته‌تر نیز با چالش بیشتری در رفتار رواناب مواجه

جدول ۳- نتایج اعتبارسنجی مدل‌های در ایستگاه‌های هیدرومتری منتخب

Table 3. Validation results of the models at selected hydrometric stations

Station Name	standard metric	Models				
		SARIMA	BiLSTM	GRU	BiLSTM-GRU	SARIMA-BiLSTM-GRU
Pole-Touskestan	MAD	0.0686	0.0889	0.0275	0.0887	0.0245
	RMSE	0.0851	0.0355	0.0380	0.0326	0.0295
	MSE	0.0072	0.0018	0.0014	0.0009	0.0009
Naharkhoran	MAD	0.0140	0.0888	0.0153	0.0891	0.0148
	RMSE	0.0230	0.0243	0.0260	0.0226	0.0173
	MSE	0.0005	0.0009	0.0007	0.0005	0.0003
Siyahab	MAD	0.2492	0.2445	0.2308	0.3191	0.1403
	RMSE	0.3892	0.3153	0.3183	0.3013	0.1683
	MSE	0.1514	0.1678	0.1013	0.0908	0.0283

مدل ترکیبی SARIMA-BiLSTM-GRU با تلفیق قابلیت‌های مدل خطی و شبکه‌های عمیق غیرخطی، عملکردی برتر و پایدارتر ارائه می‌دهد. همچنین، نتایج این تحقیق با تکمیل این مبانی نشان می‌دهد که در حوضه‌هایی با رفتار هیدرولوژیک خطی‌شده‌تر (مانند ایستگاه نه‌خوران)، مدل‌های ساده‌تری مانند SARIMA نیز می‌توانند به تنهایی به عملکردی رقابتی دست یابند، که این موضوع اهمیت انطباق ساختار مدل با ویژگی‌های خاص حوضه را برجسته می‌سازد.

نتیجه‌گیری

در سال‌های اخیر، مدل‌های یادگیری عمیق و ترکیبی به‌طور گسترده‌ای در پیش‌بینی رواناب و سایر متغیرهای هیدرولوژیکی و محیطی به کار گرفته شده‌اند. این مدل‌ها به دلیل توانایی در شبیه‌سازی روابط پیچیده و غیرخطی در داده‌های سری زمانی، توجه بسیاری را جلب کرده‌اند. در این راستا، مدل‌های ترکیبی به‌ویژه با ادغام روش‌های خطی و غیرخطی، به‌منظور بهبود دقت، پایداری و تعمیم‌پذیری پیش‌بینی‌ها، مورد توجه قرار گرفته‌اند.

در این پژوهش، پنج مدل شامل سه مدل منفرد (SARIMA، GRU و BiLSTM) و دو مدل ترکیبی (BiLSTM-GRU و SARIMA-BiLSTM-GRU) برای پیش‌بینی رواناب ماهانه در سه ایستگاه هیدرومتری حوزه آبخیز قره‌سو مدل‌سازی شد. نتایج مدل‌سازی و اعتبارسنجی نشان داد که مدل ترکیبی SARIMA-BiLSTM-GRU با ترکیب مؤثر اجزای

به‌طور کلی، مدل‌های ترکیبی SARIMA-BiLSTM-GRU و BiLSTM-GRU بهترین دقت را برای پیش‌بینی رواناب ماهانه داشتند. در حالی که مدل‌های GRU و BiLSTM به‌عنوان جایگزین‌های ساده و کارا نیز قابل استفاده هستند. به‌طور میانگین و بر اساس شاخص RMSE، مدل SARIMA-BiLSTM-GRU توانست خطای پیش‌بینی را به شکل قابل توجهی کاهش دهد؛ به‌گونه‌ای که در مقایسه با مدل‌های SARIMA، BiLSTM، GRU و BiLSTM-GRU به ترتیب ۳۹/۶۶ درصد، ۴۲/۶۶ درصد، ۴۳/۷۴ درصد و ۵۶/۷۵ درصد عملکرد بهتری نشان داد.

نتایج این مطالعه با پژوهش‌های پیشین در زمینه پیش‌بینی رواناب و سایر متغیرهای محیطی در ایران و جهان همخوانی دارد (Bai et al., Xiang et al., 2020؛ Luo and Gong, 2023؛ Ahmed et al., 2022؛ 2021؛ Sarkar et al., 2022؛ Kilinc and Yurtsever, 2022؛ Fang et al., 2024). Workneh and Jha, (2025). می‌دهد که ترکیب مدل‌های خطی و یادگیری عمیق می‌تواند رویکردی مؤثر برای ارتقای دقت، پایداری و قابلیت اعتماد در پیش‌بینی رواناب و مدیریت منابع آب در مقیاس حوزه آبخیز باشد.

یافته‌های این پژوهش، مبانی نظری مطرح‌شده در خصوص برتری مدل‌های ترکیبی پیچیده را تأیید و تقویت می‌کند. مشابه نتایج مطالعاتی مانند Ahmed et al., (2022) و Luo and Gong (2023) که بر مزیت ادغام مؤلفه‌های خطی و غیرخطی و الگوریتم‌های بهینه‌سازی تأکید داشتند، این مطالعه نیز نشان داد که

رواناب توسعه یافتند و در نظر نگرفتن متغیرهای هواشناسی مؤثر (مانند بارش و دما) می‌تواند از جامعیت آن بکاهد. دوم، مدل‌های پایه یادگیری عمیق با چالش بیش‌برازش مواجه بودند که با معماری ترکیبی ارائه‌شده تلاش شد مهار گردد. سوم، کارایی مدل بهینه به شرایط هیدرولوژیک و داده‌های حوضه مورد مطالعه (قره‌سو) وابسته است و تعمیم آن به سایر حوضه‌ها نیازمند واسنجی مجدد است. در نهایت، این مطالعه عدم قطعیت ذاتی در پیش‌بینی‌های حاصل را به صورت کمی بررسی نکرده است.

با توجه به مشاهده پدیده بیش‌برازش در مدل‌های پایه LSTM و GRU، پیشنهاد می‌شود از معماری‌های پیشرفته‌تری مانند مدل‌های مبتنی بر مکانیسم Attention نظیر Transformer استفاده شود تا مدل بتواند به‌طور هوشمند بر دوره‌های زمانی مؤثر و کلیدی در فرایند رواناب تمرکز کند. همچنین برای غلبه بر پیچیدگی و دقت کمتر مدل‌ها در ایستگاه خروجی حوضه، ادغام داده‌های تأثیرگذار اضافی مانند نوسانات سطح آب زیرزمینی، و داده‌های مرتبط با فعالیت‌های انسانی (مانند الگوی برداشت آب) در کنار متغیرهای هواشناسی ضروری به نظر می‌رسد. تحلیل عدم قطعیت و حساسیت مدل‌ها نسبت به تغییرات اقلیمی و ورودی‌ها، به همراه ادغام مدل‌های آماری و یادگیری ماشین، می‌تواند به ارائه پیش‌بینی‌های دقیق‌تر و پایدارتر کمک کند و زمینه استفاده عملی در مدیریت منابع آب، کنترل سیلاب و برنامه‌ریزی سدها را فراهم آورد.

تشکر و قدردانی

نویسندگان قدردان پشتیبانی همه‌جانبه و فراهم آوردن زیرساخت‌های پژوهشی توسط دانشگاه گنبد کاووس هستند که بدون آن، انجام این مطالعه میسر نمی‌شد.

تعارض منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافی در خصوص نگارش و انتشار مطالب و نتایج این پژوهش ندارند.

خطی (SARIMA) و غیرخطی (GRU و BiLSTM)، به‌عنوان مدل بهینه برای این حوضه انتخاب شد که دارای بهترین عملکرد در پیش‌بینی رواناب ماهانه بوده است. این مدل با در نظر گرفتن هر دو جنبه خطی و غیرخطی داده‌ها، دقت پیش‌بینی را به‌طور قابل توجهی افزایش می‌دهد. با این وجود، مدل بهینه ترکیبی در ایستگاه پل توسکاستان با بیش‌برازش مواجه بوده است. احتمالاً این موضوع با رویدادهای حدی یا تغییرات رژیم هیدرولوژیکی مرتبط است که در تاریخچه داده‌های آموزشی مدل به اندازه کافی نمایندگی نشده و تنها بر پایه الگوهای گذشته آموزش دیده است. در مقایسه، مدل‌های BiLSTM-GRU، BiLSTM و GRU به ترتیب در جایگاه‌های بعدی قرار دارند. بیش‌برازش مشاهده شده در دو مدل BiLSTM و GRU احتمالاً به پیچیدگی ذاتی این معماری‌ها در مواجهه با داده‌های سری‌زمانی با نویز و وابستگی‌های غیرخطی مرتبط است. این مدل‌ها تمایل دارند نه تنها الگوهای کلی، بلکه نوسانات خاص و احتمالاً نویز موجود در داده‌های آموزشی رواناب گذشته را نیز بیاموزند، که منجر به کاهش کارایی در داده‌های آزمون می‌شود.

مدل SARIMA، اگرچه در شبیه‌سازی روندهای خطی داده‌ها موفق بوده است، اما در مقایسه با مدل‌های ترکیبی، دقت کمتری در پیش‌بینی رواناب ماهانه نشان داد. عملکرد ضعیف‌تر کلیه مدل‌ها در ایستگاه سیاه‌آب واقع در خروجی حوضه ناشی از پیچیدگی هیدرولوژیکی افزوده در این مقیاس است که عمدتاً شامل ترکیب غیرخطی پاسخ زیرحوضه‌های ناهمگن، تأثیر بیشتر عوامل انسانی و مدیریتی (مانند برداشت آب و تغییر کاربری) و افزایش اثرات انتقال و تأخیر در مسیرهای طولانی‌تر است. این عوامل باعث کاهش وضوح روابط خطی و افزایش عدم قطعیت در مدلسازی می‌شوند. به طور کلی، استفاده از مدل‌های ترکیبی یادگیری عمیق و خطی برای پیش‌بینی رواناب، راهکاری مؤثر برای افزایش دقت و پایداری پیش‌بینی‌ها و مدیریت منابع آب در مقیاس حوزه آبخیز محسوب می‌شود.

این مطالعه با وجود نتایج رضایتبخش، دارای محدودیت‌هایی است که زمینه را برای پژوهش‌های آتی فراهم می‌کند. نخست، مدل‌ها تنها بر پایه سری زمانی

منابع مورد استفاده

- Ahmed, A.M., Deo, R.C., Ghahramani, A., Feng, Q., Raj, N., Yin, Z., Yang, L., 2022. New double decomposition deep learning methods for river water level forecasting. *Stoten* 831, 154722.
- Baek, S.S., Pyo, J., Chun, J.A., 2020. Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. *Water* 1(12), 3399.
- Bai, Y., Bezak, N., Zeng, B., Li, C., Sapač, K., Zhang, J., 2021. Daily runoff forecasting using a cascade long short-term memory model that considers different variables. *Water Resour. Manag.* 35(4), 1167-118.
- Beven, K., 2020. Deep learning, hydrological processes and the uniqueness of place. *Hydrol. Process.* 34, 3608–3613.
- Boulariah, O., Mikhailov, P.A., Longobardi, A., Elizariiev, A.N., Aksenov, S.G., 2021. Assessment of prediction performances of stochastic models: Monthly groundwater level prediction in Southern Italy. *J. Groundw. Sci. Eng.* 9(2), 161-170.
- Bowerman, B.L., O'Connell, R., 1993. *Forecasting and time series: An applied approach*, Third edition, Amazon Publication.
- Box, G.E., Jenkins, G.M., Reinsel, G.C., 2015. *Time series analysis: forecasting and control* (Vol. 734). John Wiley & Sons.
- Chen, S., Huang, J., Huang -Chuan, J.R., 2023. Improving daily streamflow simulations for data-scarce watersheds using the coupled SWAT-LSTM approach. *J. Hydrol.* 622, 129734.
- Cho, K., 2014. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *arXiv preprint arXiv:1406.1078*.
- De la Fuente, L.A., Ehsani, M.R., Gupta, H.V., Condon, L.E., 2024. Toward interpretable LSTM-based modeling of hydrological systems. *Hydrol. Earth Syst. Sci.* 28, 945-971.
- Fang, J., Yang, L., Wen, X., Li, W., Yu, H., Zhou, T., 2024. A deep learning-based hybrid approach for multi-time-ahead streamflow prediction in an arid region of Northwest China. *Hydrol. Res.* 55, 180–204.
- Ghezelsofli, H., Jandaghi, N., Ghareh Mahmoodlu, M., Azimmohseni, M., Seyedian, M., 2022. Modeling and forecasting of monthly runoff in the time domain (case study: River basin Gharasou). *Environ. Erosion Res. J.* 12(3), 165-188 (in Persian).
- Hauswirth, S.M., Bierkens, M.F., Beijk, V., Wanders, N., 2021. The potential of data driven approaches for quantifying hydrological extremes. *Adv. Water Resour.* 155, 104017.
- Heshmatpour, A., Jandaghi, N., Pasand, S., Ghareh Mahmoodlu, M., 2020. Drought effects on surface water quality in Golestan Province for Irrigation Purposes, Case study: Gorganroud River. *Phys. Geograph. Quart.* 13(48), 75-88 (in Persian).
- Hu, X., Shi, L., Lin, G., Lin, L., 2021. Comparison of physical-based, data-driven and hybrid modeling approaches for evapotranspiration estimation. *J. Hydrol.* 601, 126592.
- Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. *J. Stat. Softw.* 22, 1-27.
- Hyndman, R.J., 2018. *Forecasting: principles and practice*. OTexts.
- Jandaghi, N., Azimmohseni, M., Ghareh Mahmoodlu, M., 2021. Rainfall-runoff process modeling using time series transfer function. *Environ. Erosion Res. J.* 11(2), 111-128 (in Persian).
- Jordan, I.D., Sokół, P.A., Park, I.M., 2021. Gated recurrent units viewed through the lens of continuous time dynamical systems. *Front. Comput. Neurosci.* 678158, 15. <https://doi.org/10.3389/fncom.2021.678158>
- Kilinc, H.C., Yurtsever, A., 2022. Short-term streamflow forecasting using hybrid deep learning model based on Grey Wolf Algorithm for hydrological time series. *Sustain.* 14, 3352.
- Latif, S.D., Ahmed, A.N., 2023. Streamflow prediction utilizing deep learning and machine learning algorithms for sustainable water supply management. *Water Resour. Manag.* 37, 3227–3241.
- Li, G., Yang, N., 2023. A hybrid SARIMA-LSTM model for air temperature forecasting. *Adv. Theory Simul.* 6(2), 2200502.
- Lin, Y., Wang, D., Wang, G., Qiu, J., Long, K., Du, Y., Xie, H., Wei, Z., Shanguan, W., Dai, Y., 2021. A hybrid deep learning algorithm and its application to streamflow prediction. *J. Hydrol.* 601, 126636.
- Luo, J., Gong, Y., 2023. Air pollutant prediction based on ARIMA-WOA-LSTM model. *Atmos. Pollut. Res.* 14, 1-13.
- Mateus, B.C., Mendes, M., Farinha, J.T., Assis, R., Cardoso, A.M., 2021. Comparing LSTM and GRU models to predict the condition of a pulp paper press. *Energ.* 14(21), 6958.
- Moriassi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50(3), 885-900.

- Nacer, S.M., Nadia, B., Abdelghani, R., Mohamed, B., 2023. A novel method for bearing fault diagnosis based on BiLSTM neural networks. *Int. J. Adv. Manuf. Technol.* 125(3), 1477-1492.
- Nazeri, Y., Jandaghi, N., Ghareh Mahmoodlu, M., Azimmohseni, M., 2023. Relationship between rainfall and groundwater level using time-lagged regression. *J. Water Soil Conserv.* 30(2), 27-49 (in Persian).
- Parisouj, P., Mokari, E., Mohebzadeh, H., Goharnejad, H., Jun, C., Oh, J., Bateni, S.M., 2022. Physics-Informed Data-Driven model for predicting streamflow: A case study of the Voshmgir Basin, Iran. *Appl. Sci.* 12, 7464.
- Sarkar, N., Gupta, R., Keserwani, P.K., 2022. Air Quality Index prediction using an effective hybrid deep learning model. *Environ. Pollut.* 315, 120404. <https://doi.org/10.1016/j.envpol.2022.120404>
- Schuster, M., Paliwal, K.K., 1997. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* 45(11), 2673-2681.
- Velarde, G., Brañez, P., Bueno, A., Heredia, R., Lopez-Ledezma, M., 2022. An Open Source and Reproducible Implementation of LSTM and GRU Networks for Time Series Forecasting. *Eng. Proc.* 18(1), 30.
- Wang, X., Kang, Y., Hyndman, R.J., Li, F., 2022. Distributed ARIMA models for ultra-long time series. *Int. J. Forecast.* <https://doi.org/10.1016/j.ijforecast.2022.05.001>.
- Workneh, H., Jha, M., 2025. Utilizing hybrid deep learning models for streamflow prediction. *Water* 17(13), 1913. <https://doi.org/10.3390/w17131913>
- Xiang, Z., Yan, J., Demir, I., 2020. A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resour. Res.* 56(1), e2019WR025326.
- Yang, S., Yang, D., Chen, J., Santisirisomboon, J., Lu, W., Zhao, B., 2020. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *J. Hydrol.* 590, 125206.
- Zhang, G.P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomput.* 50, 159-175.