Iranian Journal of Soil and Water Research
Homepage: http://ijswr.ut.ac.ir
ISNN: 2423-7833

Comparison of Box-Jenkins Models with Machine Learning Algorithms for streamflow yield modeling (Case study: Taleghan watershed)

Sina Nazari Cheragh Tape[1] | Mohamad Ansari Ghojghar[2]✉ | Arash Malekian[3] | Mehrnoosh Ghadimi[4]

1. Department of Reclamation of Arid and Mountainous Regions Engineering, Faculty of Natural Resources, University of Tehran, Karaj, Iran. E-mail: sina.nazari.ch@ut.ac.ir
2. Corresponding Author, Department of Reclamation of Arid and Mountainous Regions Engineering, Faculty of Natural Resources, University of Tehran, Karaj, Iran. E-mail: ansari.ghojghar@ut.ac.ir
3. Department of Reclamation of Arid and Mountainous Regions Engineering, Faculty of Natural Resources, University of Tehran, Karaj, Iran. E-mail: malekian@ut.ac.ir
4. Natural Geography Group, Faculty of Geography, University of Tehran, Tehran, Iran. Email: ghadimi@ut.ac.ir

Article Info

Article type: Research Article

Article history:

Received: Oct. 11, 2025

Revised: Dec. 10, 2025

Accepted: Dec. 15, 2025

Published online: Jan. 2026

Keywords:
Streamflow yield forecasting,
ARIMA model,
ELM model,
SARIMA model,
XGBoost model.

ABSTRACT

This study compared the performance of ARIMA, SARIMA, ELM and XGBoost models for modeling and forecasting monthly streamflow yield in the Taleghan watershed. The dataset included monthly mean discharge time series from five hydrometric stations (Joestan, Mehran Joestan, Dehdar, Gatehdeh, and Alizan Joestan) over a 30-year period from the 1989–1990 to 2018–2019 water years. Data were split 80% for training and 20% for testing, with models trained and evaluated using four input combinations of one to four prior months. Performance was assessed via root mean square error (RMSE), mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE), and correlation coefficient (R). Results showed machine learning models outperformed classical ones, especially at Joestan station. XGBoost achieved the best results with NSE of 0.978 (training) and 0.961 (testing) using four-month inputs. Increasing prior months improved accuracy; for XGBoost testing, four months raised NSE by 2.1% (0.94 to 0.96) and cut RMSE by 2.6% (0.16 to 0.15). Machine learning models offer effective tools for streamflow yield forecasting and water resources management. Future work could focus on hybrid models and climatic data integration.

Cite this article: Nazari Cheragh Tape, S., Ansari Ghojghar, M., Malekian, A., Ghadimi, M., (2026) Comparison of Box-Jenkins Models with Machine Learning Algorithms for streamflow yield modeling (Case study: Taleghan watershed), *Iranian Journal of Soil and Water Research*, 56 (11), 3009-3026. https://doi.org/10.22059/ijswr.2025.403797.670025

Publisher: University of Tehran Press.
DOI: https://doi.org/10.22059/ijswr.2025.403797.670025

EXTENDED ABSTRACT

Introduction

Streamflow yield is one of the most important hydrological variables for water resources management in arid and semi-arid regions such as Iran. Accurate forecasting of monthly streamflow yield is essential for optimal reservoir operation, drought mitigation, agricultural water allocation, and sustainable water supply planning. Given Iran's climate with low and highly irregular precipitation, reliable monthly streamflow yield forecasting is particularly critical. Major challenges in monthly streamflow yield forecasting include high temporal variability, non-stationarity, and pronounced nonlinear behavior of hydrological time series. Although classical statistical models such as ARIMA and SARIMA have long been applied, they frequently fail to capture the complex patterns inherent in monthly streamflow yield data. In recent years, machine learning and data-driven techniques have demonstrated significantly superior performance in modeling such intricate relationships. Despite these advances, issues persist, including the scarcity of long-term, high-quality monthly streamflow yield records, complex interactions between climate drivers and catchment characteristics, and the requirement for robust generalization to unseen conditions. Consequently, comparative studies and the development of hybrid approaches remain vital for further improvement. Given the pivotal role of accurate monthly streamflow yield forecasting in operational hydrology and long-term water resources planning in Iran, investment in advanced data-driven modeling has become a strategic necessity.

Materials and Methods

This study was conducted using long-term monthly mean discharge time series recorded at five hydrometric stations (Joestan, Mehran-Joestan, Dehdar, Gate-Deh, and Alizan-Joestan) within the Taleghan watershed, northern Iran. The dataset covers a continuous 30-year period from the water year 1368 to 1398 (1989–2019). Missing values in the monthly time series were reconstructed using linear regression and inverse distance weighting (IDW). Homogeneity of the series was verified using the Run Test at a 95% confidence level, while long-term persistence was confirmed by Hurst exponent values greater than 0.5, indicating the suitability of the data for time-series forecasting.

Four models were employed and compared:

Two classical approaches: ARIMA (for non-seasonal patterns) and SARIMA (for seasonal patterns)

Two advanced machine learning models: Extreme Learning Machine (ELM) and XGBoost

All models were developed using lagged monthly streamflow yield values (1 to 4 previous months) as predictors. Performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Nash–Sutcliffe Efficiency (NSE), and Pearson correlation coefficient (R). The dataset was divided into 80% training and 20% testing subsets, and k-fold cross-validation was applied to ensure model robustness and prevent overfitting. Stationarity was examined using the Augmented Dickey–Fuller test, and differencing was applied where necessary. All analyses and modeling were performed in R using the packages forecast, xgboost, elmNNRcpp, and hydroGOF.

Results and Discussion

The performance of the four models was systematically evaluated across the five stations using the monthly mean discharge time series. The classical ARIMA and SARIMA models provided reasonable results for relatively smooth series but exhibited clear limitations in representing sharp peaks and strong nonlinearity typically observed in monthly streamflow yield data. In contrast, the machine learning models consistently outperformed the statistical models at all stations.

Among all tested configurations, XGBoost achieved the highest accuracy. At Joestan station, using four lagged months as inputs, XGBoost yielded an NSE of 0.978 in the training period and 0.961 in the testing period. Increasing the number of lagged inputs from one to four months systematically improved forecasting accuracy; for XGBoost, the NSE in the testing phase increased from 0.940 to 0.960 (a 2.1% improvement), while RMSE decreased from 0.16 to 0.15 m^3/s (a 6.2% reduction).

The Extreme Learning Machine (ELM) also performed strongly and offered considerably faster training times, making it a practical alternative under computational constraints. Overall, XGBoost proved to be the most robust and accurate model for monthly streamflow yield forecasting in the Taleghan watershed, demonstrating excellent generalization and stability.

Conclusion

The results strongly recommend the adoption of XGBoost as the preferred tool for operational monthly streamflow yield forecasting in semi-arid basins similar to Taleghan. ELM serves as an efficient high-speed alternative, whereas traditional ARIMA and SARIMA models remain suitable only for preliminary analyses of less complex monthly streamflow yield series. This study confirms the superior capability of advanced

machine learning techniques in monthly streamflow yield time-series forecasting and provides a solid foundation for their operational implementation in water resources management in Iran.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank the reviewers and editor for their critical comments that helped to improve the paper. The authors gratefully acknowledge the support and facilities provided by the Department of Reclamation of arid and mountainous regions, Faculty of Natural Resources, University of Tehran, Iran.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Conflict of interest

The author declares no conflict of interest.

مجله تحقیقات آب و خاک ایران، دوره ۵۶، شماره ۱۱ شاپا: ۷۸۳۳-۲۴۲۳
Homepage: http://ijswr.ut.ac.ir

مقایسه مدل‌های باکس جنکینز با الگوریتم‌های یادگیری ماشین به منظور مدل‌سازی آبدهی جریان(مطالعه موردی: حوزه آبخیز طالقان)

سینا نظری چراغ تپه[۱] | محمد انصاری قوجقار[۲]✉ | آرش ملکیان[۳] | مهرنوش قدیمی[۴]

۱.گروه مهندسی احیاء مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: sina.nazari.ch@ut.ac.ir

۲.نویسنده مسئول، گروه مهندسی احیاء مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: ansari.ghojghar@ut.ac.ir

۳. گروه مهندسی احیاء مناطق خشک و کوهستانی، دانشکده منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانامه: malekian@ut.ac.ir

۴. گروه جغرافیای طبیعی، دانشکده جغرافیا، دانشگاه تهران، تهران، ایران. رایانامه: ghadimi@ut.ac.ir

اطلاعات مقاله

نوع مقاله: مقالهٔ پژوهشی

تاریخ دریافت: ۱۴۰۴/۴/۲۹
تاریخ بازنگری: ۱۴۰۴/۶/۲۹
تاریخ پذیرش: ۱۴۰۴/۷/۷
تاریخ انتشار: بهمن ۱۴۰۴

واژه‌های کلیدی:
پیش‌بینی آبدهی جریان،
مدل ARIMA،
مدل ELM،
مدل SARIMA،
مدل XGBoost.

چکیده

این پژوهش با هدف مقایسه عملکرد مدل¬های ARIMA، SARIMA، ELM و XGBoost در مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه در حوزه آبخیز طالقان انجام شد. داده‌ها شامل سری‌های زمانی دبی متوسط ماهانه پنج ایستگاه هیدرومتری شامل جوستان، مهران جوستان، دهدر، گته‌ده و علیزان جوستان در دوره سی ساله آبی از ابتدای سال آبی ۱۳۶۸ تا اواخر سال آبی ۱۳۹۸ بود. داده‌ها به نسبت ۸۰٪ آموزش و ۲۰٪ آزمون تقسیم شدند و مدل‌ها با چهار ترکیب ورودی مختلف شامل ۱ تا ۴ ماه گذشته آموزش دیده و ارزیابی گردیدند. عملکرد مدل‌ها با معیارهای ریشه میانگین مربعات خطا، خطای مطلق میانگین، ضریب نش‌ساتکلیف و ضریب همبستگی سنجیده شد. نتایج نشان داد در اکثر ایستگاه‌ها و به ویژه در ایستگاه جوستان مدل‌های یادگیری ماشین به طور قابل توجهی دقیق‌تر از مدل‌های کلاسیک عمل کردند. مدل XGBoost با ضریب نش‌ساتکلیف ۰/۹۷۸در مجموعه آموزش و ۰/۹۶۱ در مجموعه آزمون برای ایستگاه جوستان با ترکیب چهار بهترین عملکرد را داشت. افزایش ماه‌های ورودی دقت پیش‌بینی را بهبود بخشید. به عنوان مثال با افزایش تعداد ماه‌های گذشته به عنوان ورودی از ۱ به ۴ دقت مدل XGBoost در مرحله آزمون با بهبود ۲,۱ درصدی ضریب نش‌ساتکلیف از ۰/۹۴ به ۰/۹۶ و کاهش ۶/۲ درصدی ریشه میانگین مربعات خطا از ۰/۱۶ به ۰/۱۵ به طور معناداری ارتقا یافت. استفاده از مدل‌های یادگیری ماشین ابزاری مؤثر برای مدل‌سازی، پیش‌بینی آبدهی جریان و مدیریت منابع آب است. تحقیقات آتی می‌تواند بر توسعه مدل‌های ترکیبی و ادغام داده‌های اقلیمی متمرکز شود.

استناد: لرمحمد حسنی؛ محبوبه، طالب نژاد؛ رضوان، نوشادی؛ مسعود، (۱۴۰۴) ارزیابی بهره‌وری اقتصادی آب در کشت کینوا با کاربرد آب شور و کم‌آبیاری در سیستم آبیاری قطره‌ای نواری در استان فارس، *مجله تحقیقات آب و خاک ایران*، ۵۶ (۱۱)، ۳۰۲۶-۳۰۰۹.
https://doi.org/10.22059/ijswr.2025.398660.669982

DOI: https://doi.org/10.22059/ijswr.2025.398660.669982

مقدمه

مدل‌سازی و پیش‌بینی دقیق آبدهی جریان به دلیل عدم قطعیت‌های هواشناسی–هیدرولوژیکی دشوار است (Moradian et al., 2024) روش‌های پیش‌بینی آبدهی جریان به دو دسته مدل‌های مفهومی و مدل‌های آماری تقسیم می‌شوند؛ مدل‌های مفهومی به داده‌ها و دانش گسترده نیاز دارند، بنابراین مدل‌های آماری به دلیل سادگی، پرکاربردتر هستند (سیدیان و همکاران، ۱۳۹۳). مدل‌سازی آبدهی جریان از طریق شبیه‌سازی فرآیندهای هیدرولوژیکی و هیدرولیکی، پیش‌بینی دقیق، ارزیابی ریسک، کاهش خسارت، برنامه‌ریزی واکنش اضطراری و پهنه‌بندی مناطق در معرض خطر را ممکن می‌سازد (Kumar, Sharma, et al., 2023).

دقت و قابلیت اطمینان مدل‌های آبدهی جریان به کیفیت و در دسترس بودن داده‌ها بستگی دارد. مدل‌سازی آبدهی جریان با چالش‌های متعددی در ابعاد داده‌ای، فرآیندی، مدل‌سازی و تعامل انسان و آب مواجه است. کمبود و عدم دسترسی به داده‌های مشاهداتی، کیفیت پایین داده‌ها، نبود اطلاعات کافی درباره تأثیرات انسانی، از مهم‌ترین موانع داده‌ای هستند. بنابراین، پیشرفت در مدل‌سازی آبدهی جریان نیازمند رویکردی چندبعدی است که همزمان به بهبود دسترسی به داده‌ها، توسعه مدل‌های ترکیبی و افزایش تعامل میان پژوهشگران، مدیران و جامعه توجه داشته باشد تا بتوان پاسخ‌های مؤثرتری به ریسک‌های ناشی از آن ارائه داد (Burnner et al., 2021; Kumar, Sharma et al., 2023)

در سال‌های اخیر، رویکرد ترکیبی استفاده از مدل‌های آماری سنتی و روش‌های نوین یادگیری ماشین و هوش مصنوعی به یک روند جهانی در مدل‌سازی و پیش‌بینی پدیده‌های هیدرولوژیکی، تبدیل شده است. تجربه‌های گسترده پژوهشی نشان داده‌اند که مدل‌های آماری مانند رگرسیون، ARIMA[1] و دیگر مدل‌های سری زمانی در شناسایی روندهای خطی و الگوهای ساده مؤثر هستند، اما برای روابط غیرخطی پیچیده و داده‌های متغیر زمانی کارایی کافی ندارند. از سوی دیگر، مدل‌های یادگیری ماشین مانند شبکه‌های عصبی بازگشتی (RNN)[2]، شبکه‌های عصبی حافظه‌دار بلند–کوتاه مدت (LSTM)[3]، و سایر الگوریتم‌های پیشرفته، توانایی شناسایی الگوهای غیرخطی و ویژگی‌های پنهان در داده‌های هیدرولوژیکی را به طرز قابل توجهی دارند.

در ایران، چالش‌های قابل توجهی در زمینه داده‌های هیدرولوژیکی برای پیش‌بینی آبدهی جریان وجود دارد. این مشکلات عمدتاً شامل کمبود داده‌های بلندمدت و با کیفیت، ناسازگاری و پراکندگی داده‌ها و محدودیت‌های زیر ساختی مانند فقدان تجهیزات مدرن و نگهداری ضعیف ایستگاه‌های نظارتی است. کمبود نیروی انسانی متخصص و زیر ساخت‌های ناکافی برای پایش خودکار، کیفیت و پوشش داده‌ها را به شدت محدود کرده است. این مسائل منجر به وجود داده‌های ناقص، پرت و پر از نویز می‌شود که قابلیت اعتماد به داده‌ها را کاهش می‌دهد. مشکلات کیفیت داده‌ها و عدم قطعیت‌های ناشی از آن‌ها مستقیماً بر دقت پژوهش‌های علمی و اثر بخشی مدیریت منابع آب و پاسخ به مخاطرات تأثیر می‌گذارد. برای رفع این چالش‌ها، نیاز به اصلاح روش‌های جمع‌آوری داده، توسعه پایگاه‌های داده جدید و بهبود زیر ساخت‌های نظارتی است (Yu, et al., 2025)

تفسیرپذیری در یادگیری ماشین به معنای توانایی توضیح و تصمیم‌گیری مدل به زبانی قابل فهم برای انسان است. این ویژگی در کاربردهای حساس به‌دلیل افزایش اعتماد، شفافیت، امکان بررسی صحت مدل و رعایت الزامات اخلاقی و قانونی بسیار حیاتی است. مدل‌های آماری سنتی مانند رگرسیون خطی و درخت تصمیم به‌دلیل ساختار ساده و روابط شفاف، تفسیرپذیری بسیار بالایی دارند و تصمیم‌گیرندگان به‌راحتی می‌توانند تأثیر هر متغیر و مسیر تصمیم را ببینند. در مقابل، شبکه‌های عصبی عمیق و مدل‌های پیچیده یادگیری ماشین معمولاً مانند جعبه سیاه عمل می‌کنند و توضیح دقیق دلیل پیش‌بینی‌هایشان دشوار است. به همین دلیل، در بسیاری از کاربردهای عملی و حساس، مدل‌های سنتی به‌خاطر تفسیرپذیری بالا همچنان ترجیح داده می‌شوند (Doshi-Velez & Kim, 2017).

مدل‌سازی داده‌محور[4] رویکردی است که به جای تکیه بر اصول و معادلات فیزیکی، بر پایه تحلیل و استفاده از داده‌های مشاهده شده برای شناخت و پیش‌بینی رفتار سیستم‌ها بنا شده است. این نوع مدل‌ها با استخراج الگوها و روابط موجود در داده‌ها، امکان پیش‌بینی متغیرهای مورد نظر را فراهم می‌کنند و برای سیستم‌های پیچیده که شناخت کامل فرآیندهای فیزیکی آنها دشوار است، بسیار مفید هستند. مدل‌های داده محور اغلب شامل تکنیک‌هایی همچون شبکه‌های عصبی، ماشین‌های بردار پشتیبان و الگوریتم‌های تکاملی می‌باشند که

1 AutoRegressive Integrated Moving Average
2 Recurrent Neural Network
3 Long Short-Term Memory
4 Data-Driven Modeling

توانایی مدل‌سازی روابط غیرخطی و پیچیده را دارند. مزیت اصلی این مدل‌ها، عدم نیاز به دانش عمیق فیزیکی و قابلیت کاربرد در شرایط متنوع داده‌ای است، اما چالش‌هایی مانند جعبه سیاه بودن و وابستگی شدید به کیفیت داده‌ها نیز دارند (Elshorbagy, et al., 2010).

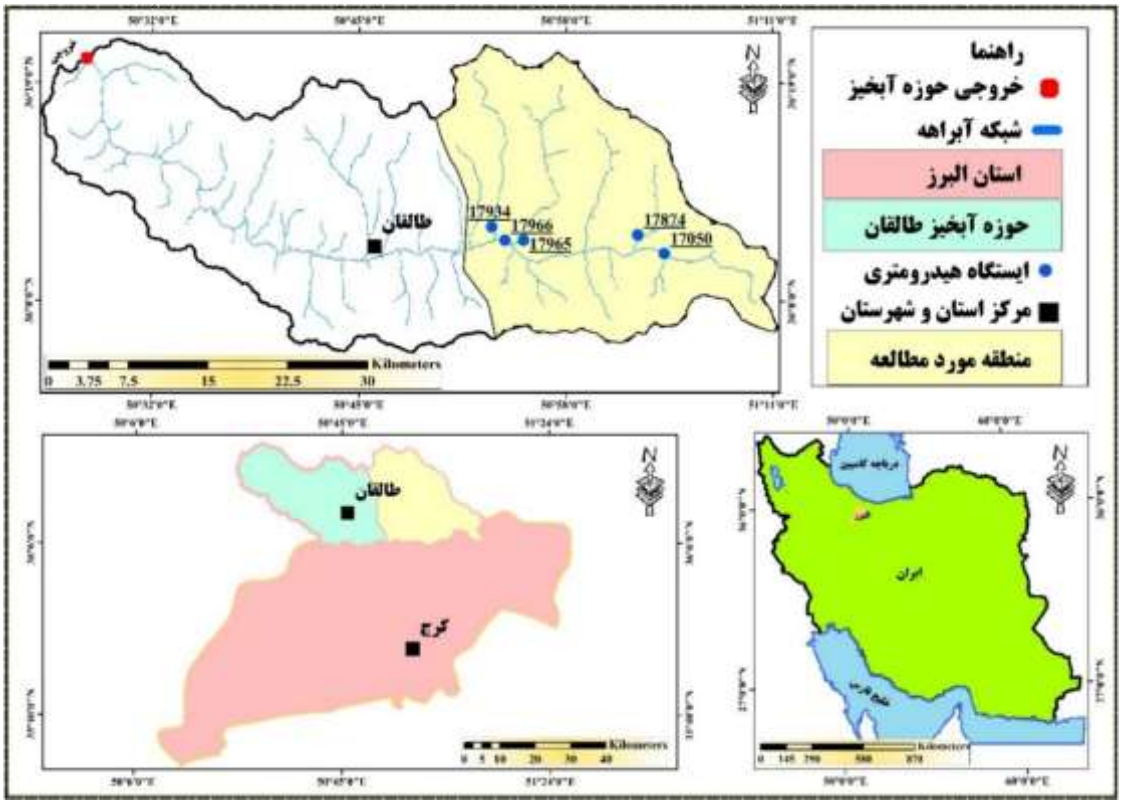
در پژوهشی ولی‌پور و همکاران (۱۳۹۱)، به پیش‌بینی ورودی ماهانه مخزن سد دز با مقایسه مدل‌های ARMA، ARIMA و شبکه عصبی مصنوعی خودرگرسیون پرداختند و از داده‌های دبی ماهانه ایستگاه هیدرومتری تله زنگ ۴۴ سال از ۱۳۸۶-۱۳۳۹ در حوزه آبخیز سد دز واقع در جنوب غرب ایران استفاده کردند و مدل‌ها را با معیارهای RMSE، MBE و خطای نسبی ارزیابی نمودند. نتایج نشان داد شبکه عصبی پویا با تابع فعالیت سیگموئید و ۱۷ نورون در لایه پنهان به عنوان مدل بهینه، قابلیت پیش‌بینی ۵ ساله را فراهم می‌کند، در حالی که مدل ARIMA (۴،۱،۱) برای پیش‌بینی ۱۲ ماهه مناسب‌تر بود. علی‌احمدی و همکاران (۱۴۰۰) در تحقیقی به پیش‌بینی دبی ماهانه رودخانه‌های سیستان و پریان در حوزه آبخیز هیرمند در استان سیستان و بلوچستان ایران با استفاده از مدل‌های سری زمانی SARIMA پرداختند. داده‌های مورد استفاده، دبی جریان ماهانه از سازمان آب منطقه‌ای طی دوره ۱۸ ساله از ۱۳۹۷-۱۳۷۹ خورشیدی بود. ارزیابی مدل‌ها با معیارهای آکائیک (AIC) و شوارتز بیزین (SBC) نشان داد SARIMA (۲،۱،۰) (۱،۱،۱) برای رودخانه سیستان و (۱،۱،۱) (۱،۱،۱)SARIMA برای رودخانه پریان با کمترین مقادیر معیارها بهینه‌ترین مدل‌ها بوده‌اند. نتایج حاکی از روند کاهشی معنادار دبی در هر دو رودخانه بود. باربولسکو و ژن (۲۰۲۴) در مطالعه‌ای به پیش‌بینی دبی ماهانه رودخانه بوزائو در رومانی با استفاده از سه مدل یادگیری ماشین (LSTM، ELM و BPNN) پرداختند. داده‌های مورد استفاده شامل سری‌های زمانی ماهانه دبی آب در بازه ۵۶ ساله (ژانویه ۱۹۵۵ تا دسامبر ۲۰۱۰) بود که به سه زیر دوره تقسیم شد: کل دوره (S)، پیش از احداث سد سیرو (۱۹۸۳-۱۹۵۵، S1) و پس از آن (۲۰۱۰-۱۹۸۴، S2). ارزیابی مدل‌ها با معیارهای MAE، MSE و R^2 نشان داد ELM با کمترین خطای مطلق برای S2 و LSTM با بالاترین ضریب تعیین برای S2 بهترین عملکرد را دارند، ضمن آن‌که احداث سد منجر به تغییر معنادار رژیم جریان رودخانه شد.

نوآوری این پژوهش در مقایسه سیستماتیک و جامع مدل‌های کلاسیک سری زمانی باکس-جنکینز (ARIMA و SARIMA) با الگوریتم‌های یادگیری ماشین (XGBoost و ELM) برای مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه در حوزه آبخیز طالقان نهفته است. برخلاف مطالعات پیشین که اغلب بر یکی از این رویکردها تمرکز داشته‌اند، این تحقیق برای نخستین بار عملکرد این مدل‌ها را در چهار ترکیب ورودی متفاوت (استفاده از داده‌های ۱ تا ۴ ماه گذشته) ارزیابی کرده و نشان می‌دهد که مدل‌های یادگیری ماشین، با بهره‌گیری از قابلیت‌های یادگیری غیرخطی و انطباق‌پذیری بالا، به طور معناداری دقت پیش‌بینی را افزایش می‌دهند. علاوه بر این، پژوهش حاضر آشکار می‌سازد که افزایش تعداد ماه‌های گذشته به عنوان ورودی، کارایی مدل‌ها را بهبود می‌بخشد و امکان ترکیب بهینه این رویکردها را برای مدیریت پایدار منابع آب در مناطق کوهستانی و خشک پیشنهاد می‌کند. این دستاوردها نه تنها شکاف موجود در ادبیات مرتبط با پیش‌بینی آبدهی جریان در شرایط داده‌های محدود هیدرولوژیکی را پر می‌کنند، بلکه ابزارهای عملی برای سیاست‌گذاری‌های محیطی ارائه می‌دهند.

روش‌شناسی پژوهش

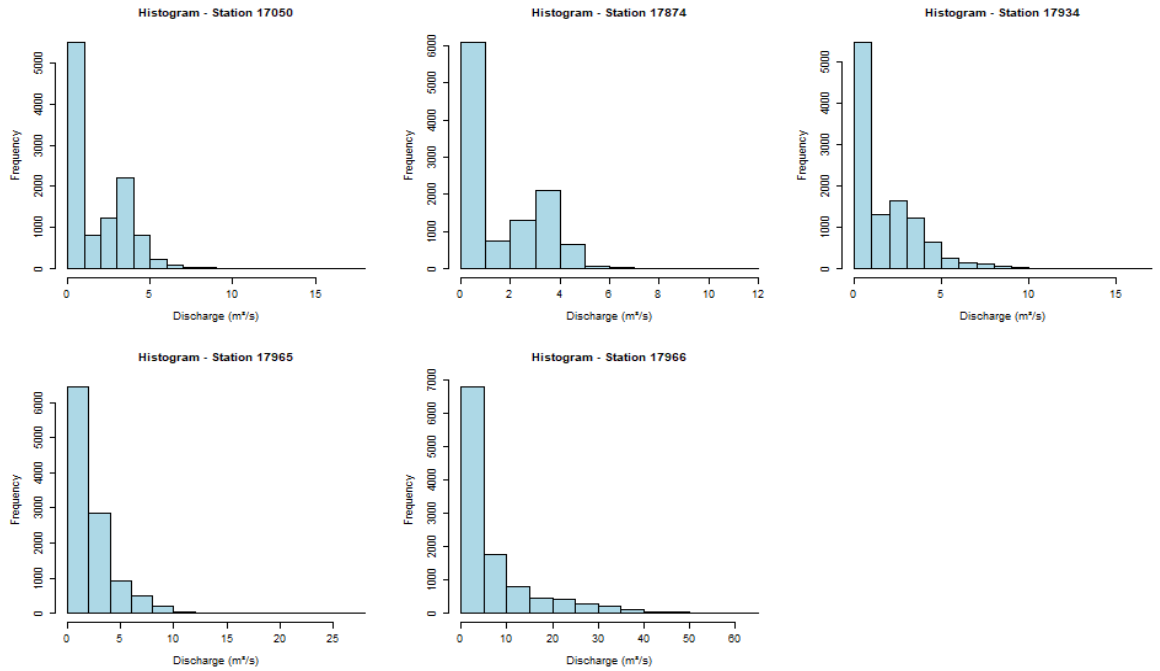
منطقه مورد مطالعه

شکل ۱ منطقه مورد مطالعه و توزیع ایستگاه‌های استفاده شده در این پژوهش را نمایش می‌دهد. حوزه آبخیز طالقان در دامنه‌های جنوبی رشته کوه البرز و در شمال غرب شهر کرج قرار گرفته و به عنوان یکی از زیر حوضه‌های مهم حوزه آبخیز سفیدرود شناخته می‌شود. این حوضه، بین عرض‌های جغرافیایی ۳۶ درجه و ۲۰ دقیقه و ۴۸ ثانیه تا ۳۶ درجه و ۵ دقیقه و ۲۳ ثانیه شمالی و طول‌های جغرافیایی ۵۰ درجه و ۳۹ دقیقه و ۳۵ ثانیه تا ۵۱ درجه و ۱۱ دقیقه و ۷ ثانیه شرقی واقع شده است. مساحت حوزه آبخیز حدودا ۱۲۴۳ کیلومتر مربع است و فاصله آن تا مرکز استان البرز (کرج) ۱۰۵ کیلومتر و تا شمال غرب تهران ۱۲۰ کیلومتر است. این شهرستان از دو بخش، چهار دهستان و هفتاد و پنج روستا تشکیل شده است. اقلیم این منطقه بر اساس طبقه‌بندی دومارتن، از نوع مدیترانه‌ای تا مرطوب بوده و ارتفاع آن از سطح دریا بین ۱۲۶۰ تا ۴۱۲۵ متر متغیر است. رودخانه طالقان، که از کوه‌های کندوان و کهار بزرگ سرچشمه می‌گیرد، به عنوان مهم‌ترین رودخانه این شهرستان شناخته می‌شود. به دلیل وجود ارتفاعات و شیب‌های منتهی به سد طالقان، ۶۳ درصد از مساحت شهرستان در زمره اراضی حفاظتی قرار دارد و نقش مهمی در تأمین آب منطقه ایفا می‌کند (ابراهیمی و همکاران، ۱۳۹۹).



شکل ۱ - منطقه و ایستگاه‌های مورد مطالعه

داده‌های مورد استفاده شامل سری‌های زمانی دبی متوسط ماهانه طی دوره ۳۰ ساله آبی از ابتدای سال آبی ۱۳۶۸ تا انتهای سال آبی ۱۳۹۸ و از پنج ایستگاه هیدرومتری (دهدر، گته‌ده، جوستان، علیزان جوستان و مهران جوستان) واقع در حوزه آبخیز طالقان است که به عنوان داده‌های ورودی مدل مورد استفاده قرار گرفت و در چهار ترکیب ورودی مختلف (۱ به معنی استفاده فقط از داده ماه قبل تا ۴ به معنی استفاده از داده‌های چهار ماه گذشته) و دو دسته داده آموزش و آزمون با نسبت ۸۰:۲۰ مدل‌سازی شده‌اند. در این پژوهش از داده‌های دبی متوسط ماهانه استفاده گردید که جدول ۱ مشخصات این ایستگاه‌ها و شکل ۲ هیستوگرام‌های توزیع دبی در پنج ایستگاه هیدرومتری را نشان می‌دهند.



شکل ۲ - هیستوگرام‌های توزیع دبی در پنج ایستگاه مورد مطالعه

جدول ۱ - مشخصات ایستگاه‌های مورد مطالعه

شماره ردیف	کد ایستگاه	نام ایستگاه	ارتفاع از سطح دریا (متر)	انحراف معیار	میانگین	نام رودخانه
۱	۱۷۰۵۰	گته‌ده	۲۳۵۳	۱/۸۳۴	۱/۸۵۸	شاهرود
۲	۱۷۸۷۴	دهدر	۲۲۶۸	۱/۵۱۵	۱/۵۸۶	دهدر
۳	۱۷۹۶۵	مهران جوستان	۲۰۴۹	۱/۸۶۱	۱/۸۸۴	شاهرود
۴	۱۷۹۶۶	جوستان	۱۹۷۹	۲/۰۰۹	۲/۳۲۳	شاهرود
۵	۱۷۹۳۴	علیزان جوستان	۱۹۵۳	۸/۵۲۲	۷/۶۵۵	شاهرود

مدل میانگین متحرک خود همبسته یکپارچه[1](ARIMA)

مدل ARIMA برای اولین بار در سال ۱۹۷۰ میلادی توسط باکس و جنکینز معرفی شد. این مدل یک روش آماری پرکاربرد برای تحلیل و پیش‌بینی سری‌های زمانی غیر ایستا است. این مدل می‌تواند به صورت خودرگرسیونی[2](AR)، میانگین متحرک[3](MA)، خودرگرسیونی میانگین متحرک[4](ARMA) یا خودرگرسیونی میانگین متحرک تلفیقی (ARIMA) و به این صورت ARIMA (p.d,q) تعریف شود که در آن P نشان‌دهنده مرتبه خودرگرسیونی، q نشان‌دهنده مرتبه میانگین متحرک، p,q نشان‌دهنده مرتبه خودرگرسیون میانگین متحرک (Chattopadhyay., 2010) و d تعداد دفعات تفاضل‌گیری لازم برای ساکن کردن سری زمانی است.

مدل میانگین متحرک خود همبسته یکپارچه فصلی[5](SARIMA)

مدل SARIMA یک روش آماری شناخته شده است که به طور گسترده برای پیش‌بینی داده‌های سری زمانی استفاده می‌شود، به ویژه زمانی که داده‌ها الگوهای فصلی یا چرخه‌ای مشخصی را نشان می‌دهند. این مدل بر اساس مفاهیم اصلی مدل ARIMA ساخته شده، اما با افزودن مولفه‌های فصلی که نوسانات و الگوهای تکرارشونده در دوره‌های فصلی خاص را شبیه‌سازی می‌کند (Panicker et al., 2024).

مدل بهینه سازی حدی گرادیان[6](XGBoost)

این روش در سال ۲۰۱۶ توسط Chen و Guestrin معرفی شد و زیر شاخه‌ای از رویکرد Gradient Boosting می‌باشد. این الگوریتم مبتنی بر ایده تقویت است که تمام پیش‌بینی‌های یادگیرندگان ضعیف را برای ایجاد یک یادگیرنده قوی از طریق استراتژی‌های آموزش ترکیب می‌کند و بوستینگ تکنیکی است که به صورت متوالی ترکیبی از الگوریتم‌ها را می‌سازد، به طوری که هر الگوریتم جدید برای کاهش خطای الگوریتم قبلی طراحی می‌شود. و یکی از پرکاربردترین مدل‌ها در مسائل هیدرولوژی (مانند پیش‌بینی جریان رودخانه) می‌باشد هدف XGBoost جلوگیری از تطبیق بیش از حد و همچنین بهینه‌سازی منابع محاسباتی است و برای مدیریت داده‌های پیچیده و مقابله با خطاهای پیش‌بینی نمونه‌های دشوار طراحی شده است(Kumar, kedam et al., 2023) . این امر با ساده‌سازی توابع هدف به دست می‌آید که امکان ترکیب شرایط پیش‌بینی و منظم‌سازی را فراهم می‌کند، اما سرعت محاسباتی بهینه را حفظ می‌کند. همچنین، محاسبات موازی به طور خودکار در مرحله آموزش برای توابع در XGBoost اجرا می‌شود. یادگیرنده اول، ابتدا به کل فضای داده‌های ورودی برازش داده می‌شود. این فرایند برازش برای چند بار تکرار می‌شود تا زمانی که معیار توافق برآورده شود. پیش‌بینی نهایی مدل از مجموع پیش‌بینی هر یاد گیرنده به دست می‌آید (Chen & Guestrin, 2016; Szczepanek, 2022)

مدل ماشین یادگیری حدی[7](ELM)

این مدل توسط Huang et al., (2004) به عنوان یک الگوریتم یادگیری ماشین جدید برای شبکه‌های عصبی پیش خور[8]با یک لایه

1 Autoregressive Integrated Moving Average
2 AutoRegressive
3 Moving Average
4 AutoRegressive Moving Average
5 Seasonal Autoregressive Integrated Moving Average
6 Extreme Gradient Boosting
7 Extreme Learning Machine
8 Feedforward: Feedforward Neural Networks

پنهان (SLFNs) در کنفرانس بین المللی هوش مصنوعی ارائه شد که وزن‌های ورودی را به صورت تصادفی انتخاب و وزن‌های خروجی را به صورت تحلیلی با استفاده از معکوس تعمیم یافته مور–پنروز محاسبه می‌کند. هدف این مدل حل مشکل کندی یادگیری در الگوریتم‌های سنتی مانند پس‌انتشار[2]می‌باشد(Huang et al., 2004; Jia et al., 2024) . ELM یک معماری یادگیری سریع با نورون‌های تصادفی است که نیاز به تنظیم دستی ندارد (Van Thieu et al., 2024) و برای سری‌های زمانی با تغییرات ناگهانی (مانند سیلاب) و غیرخطی مناسب است (Barbulescu & Zhen, 2024) .

مدل‌های پیش‌بینی

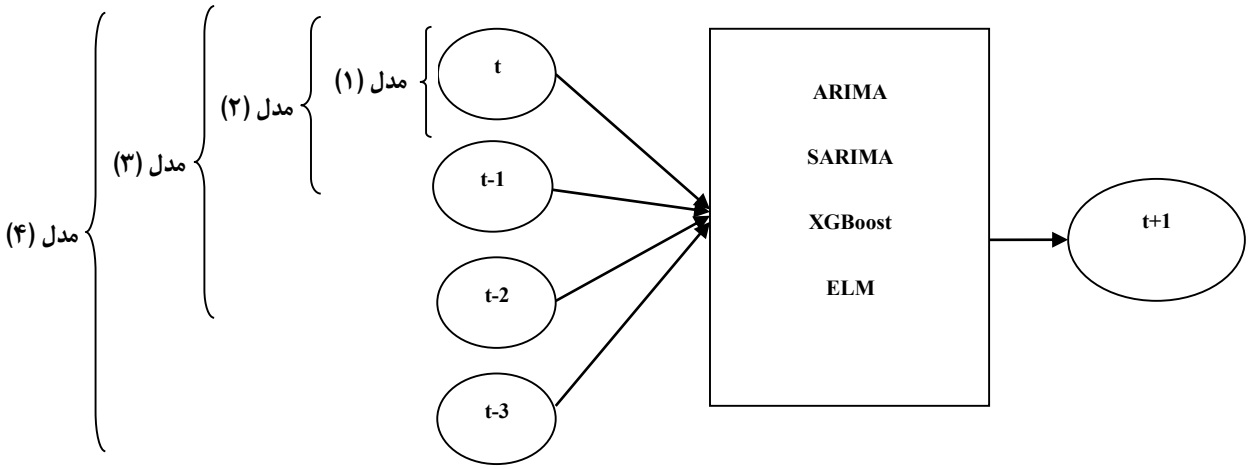
برای پیش‌بینی و تحلیل رفتار آبدهی جریان ماهانه حوضه، دو مدل پیشرفته و دو مدل سنتی مورد مطالعه و آزمایش قرار گرفتند. سپس، با استفاده از شاخص‌های ارزیابی دقت، مناسب‌ترین مدل برای هر پیش‌بینی انتخاب شد. در این فرآیند، تعداد ماه‌های گذشته به عنوان عامل اصلی برای پیش‌بینی ماه آتی در نظر گرفته شد. در رابطه اول، پیش‌بینی دبی متوسط ماهانه برای ماه بعدی با تکیه بر سری زمانی دبی با یک ماه تأخیر انجام شد، به این صورت که برای برآورد در زمان $(t+1)$، از مقدار دبی در زمان (t) استفاده گردید. در رابطه دوم، پیش‌بینی ماه آتی بر اساس داده‌های دبی تا دو ماه پیشین صورت گرفت و به همین ترتیب، رابطه‌های سوم و چهارم با در نظر گرفتن داده‌ها تا سه و چهار ماه قبل پیش رفتند رابطه‌های ۱ تا ۴ برای هر چهار مدل استفاده شده در این پژوهش مورد استفاده قرار گرفتند و به عنوان توضیحات تکمیلی ارایه می‌شوند. شکل ۳ چارچوب‌های به کار رفته برای گزینش ماه‌های پیشین در پیش‌بینی ماه آینده را نشان می‌دهد. مدل‌های به کار رفته در این تحقیق شامل XGBoost، SARIMA، ARIMA و ELM است.

رابطه ۱) $$Q_{(t+1)} = f(Q_{(t)})$$

رابطه ۲) $$Q_{(t+1)} = f(Q_{(t)}, Q_{(t-1)})$$

رابطه ۳) $$Q_{(t+1)} = f(Q_{(t)}, Q_{(t-1)}, Q_{(t-2)})$$

رابطه ۴) $$Q_{(t+1)} = f(Q_{(t)}, Q_{(t-1)}, Q_{(t-2)}, Q_{(t-3)})$$



شکل ۳ – چارچوب‌های به کار رفته برای پیش‌بینی دبی حداکثر ماهانه

معیارهای ارزیابی مدل

در این پژوهش برای مقایسه نتایج و ارزیابی مدل‌ها از چهار معیار اصلی استفاده گردید که هر کدام جنبه خاصی از دقت یا عملکرد مدل را می‌سنجنند.

(RMSE) ریشه میانگین مربعات خطا: میانگین ریشه‌دار خطاهای مربعی بین مقدار پیش‌بینی شده و واقعی را نشان می‌دهد. مقدار کمتر به معنی دقت بالاتر مدل است و این معیار با رابطه ۵ محاسبه می‌شود.

رابطه ۵) $$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - f_i)^2}$$

1 SLFNs: Single Layer Feedforward Neural Networks
2 Backpropagation

در روابط ۵ تا ۸، f_i مقادیر پیش‌بینی شده در گام زمانی iام، o_i مقادیر مشاهداتی در گام زمانی iام، $\overline{f}$ میانگین مقادیر پیش‌بینی شده، n تعداد داده‌ها و $\overline{o}$ میانگین مقادیر مشاهداتی می‌باشد.

(MAE) میانگین قدر مطلق خطا: میانگین قدر مطلق اختلاف بین مقادیر پیش‌بینی شده و واقعی است. عدد کوچکتر نشان دهنده خطای کمتر و مدل بهتر است و با رابطه ۶ محاسبه می‌شود.

رابطه ۶)
$$MAE = \frac{1}{n}\sum_{i=1}^{n}|o_i - f_i|$$

(NS) شاخص کارایی نش‌ساتکلیف: معیاری برای سنجش کارایی مدل‌های پیش‌بینی است که عدد آن می‌تواند بین منفی بی‌نهایت و ۱ باشد. عدد نزدیک به ۱ نشان می‌دهد مدل عملکرد بسیار خوبی دارد و با رابطه ۷ محاسبه می‌شود.

رابطه ۷)
$$NS = 1 - \frac{\sum_{i=1}^{n}(o_i - f_i)^2}{\sum_{i=1}^{n}(\overline{o} - o_i)^2}$$

(R) ضریب همبستگی: معیاری آماری است که شدت و جهت رابطه خطی بین دو متغیر کمی را نشان می‌دهد. مقدار این ضریب در بازه [۱, ۱-] قرار دارد. عدد نزدیک به ۱ نشان دهنده همبستگی مثبت کامل (رابطه مستقیم قوی)، عدد نزدیک به ۱- نشان دهنده همبستگی منفی کامل (رابطه معکوس قوی) و عدد صفر بدین معنی است که هیچ رابطه خطی بین دو متغیر وجود ندارد. در ارزیابی مدل‌های پیش‌بینی، ضریب همبستگی بین مقادیر واقعی و پیش‌بینی شده بیانگر میزان تطابق خطی مدل با داده‌های واقعی است. مقدار بیشتر این ضریب بیانگر کیفیت و دقت بالاتر مدل است. با کمک رابطه ۸ ضریب همبستگی محاسبه می‌شود.

رابطه ۸)
$$R = \frac{\frac{1}{n}\sum_{i=1}^{n}(o_i - \overline{o})(f_i - \overline{f})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - \overline{o})^2}\sqrt{\sum_{i=1}^{n}(f_i - \overline{f})^2}}$$

نتایج

آماده‌سازی داده‌ها یکی از مراحل بنیادین و کلیدی در فرآیند تحلیل داده‌های هیدرولوژیکی است که به منظور تضمین دقت، قابلیت اعتماد و کارایی تحلیل‌های آماری و مدل‌سازی انجام می‌شود. کیفیت داده‌های ورودی تأثیر مستقیمی بر عملکرد مدل‌های پیش‌بینی و نتایج نهایی پژوهش دارد. در این مطالعه، داده‌های خام از شرکت مدیریت منابع آب ایران جمع‌آوری شدند. برای ارتقای کیفیت داده‌ها و آماده‌سازی آن‌ها برای تحلیل‌های بعدی، مراحل زیر با دقت اجرا شدند: به منظور بازسازی مقادیر ناموجود و تکمیل سری زمانی ایستگاه‌های مورد بررسی، از دو رویکرد پیشرفته بهره گرفته شد: رگرسیون خطی و روش وزن‌دهی معکوس فاصله. پس از ارزیابی و نظارت بر کیفیت آمار ایستگاه‌ها و بر طرف کردن کمبودهای آماری، یکنواختی داده‌ها با به کارگیری آزمون Run Test سنجیده شد و تصادفی بودن داده‌ها در سطح اطمینان ۹۵ درصد تأیید گردید. برای ارزیابی کفایت دوره آماری، از ضریب هرست استفاده شد. با توجه به جدول ۲ ضرایب هرست برای پنج ایستگاه هیدرومتری مورد مطالعه نشان داد که داده‌های دبی متوسط ماهانه در دوره آماری ۳۰ ساله (۱۳۹۸-۱۳۶۹) برای پنج ایستگاه هیدرومتری، با میانگین ضریب هرست حدود ۰/۷۵ و مقادیر عمدتاً بالاتر از ۰/۷، دارای حافظه بلند مدت قوی و کفایت لازم برای تحلیل‌های هیدرولوژیکی هستند.

جدول ۲ - ضرایب هرست برای پنج ایستگاه هیدرومتری مورد مطالعه

نام ایستگاه	دبی جریان ماهانه (k)
جوستان (۱۷۹۶۶)	۰/۸۱
مهران جوستان (۱۷۹۶۵)	۰/۶۶
علیزان جوستان (۱۷۹۳۴)	۰/۷۴
گته‌ده(۱۷۰۵۰)	۰/۸۳
دهدر (۱۷۸۷۴)	۰/۷۵

قبل از شروع مدل‌سازی یا پیش‌بینی سری‌های زمانی، انجام تحلیل توصیفی مقدماتی از داده‌ها ضروری است. برای ارزیابی ایستایی

یا ناایستایی داده‌های آموزشی از آزمون Dickey-Fuller استفاده شد (کاکرین[1] ۲۰۰۵) که نتایج آن در جدول ۳ قابل مشاهده می‌باشد.

جدول ۳ – آزمون Dickey-Fuller

نام ایستگاه	P – Value
جوستان (۱۷۹۶۶)	۰/۰۱۳
مهران جوستان (۱۷۹۶۵)	۰/۰۲۲
علیزان جوستان (۱۷۹۳۴)	۰/۰۱۶
گته‌ده (۱۷۰۵۰)	۰/۰۱۱
دهدر (۱۷۸۷۴)	۰/۰۳۷

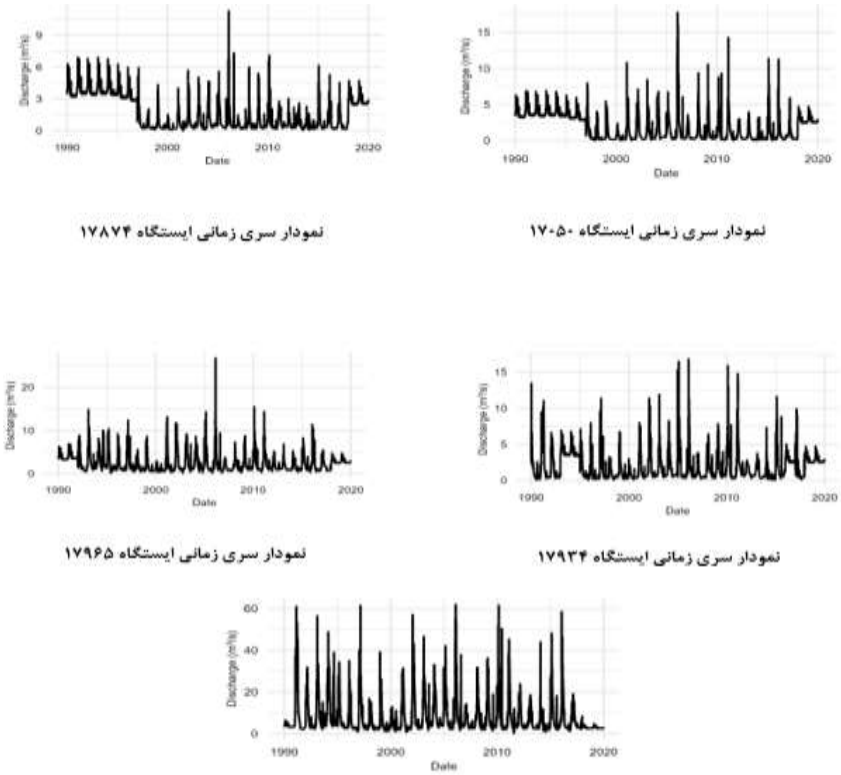
در این آزمون، سطح معناداری p-value کمتر از ۰/۰۵ تعیین شد. فرض صفر (H_0) این آزمون ناایستا بودن سری زمانی را مورد بررسی قرار می‌دهد. نتایج این آزمون نشان داد که هیچ‌کدام از ایستگاه‌های مورد مطالعه ناایستا نیستند.

مشخصات مدل‌های باکس جنکینز در جدول ۴ ارایه شده است.

جدول ۴ - مشخصات پارامترهای مدل‌های باکس جنکینز

شماره ایستگاه	$\varnothing_1$	$\varnothing_2$	θ_1	θ_2	Q	λ^2	مدل بهینه	مدل بهینه
۱۷۰۵۰	۰/۴۰۳۹	۰/۲۱۳۴	۰/۳۴۰۳	۰/۳۸۱۲	۳/۱۲	۱۳	**ARIMA**(۴،۰،۱)	**SARIMA** $(۳،۲،۰)(۱،۰،۲)_{۱۲}$
۱۷۸۷۴	۰/۴۲۵۶	۰/۲۲۱۷	۰/۳۲۷۸	۰/۳۹۱۲	۲/۷۳	۱۱	**ARIMA** (۵،۱،۳)	**SARIMA** $(۱،۰،۳)(۱،۰،۱)_{۱۲}$
۱۷۹۶۵	۰/۴۷۳۸	۰/۲۰۳۹	۰/۲۷۸۱	۰/۴۱۲۳	۴/۳۲	۱۲	**ARIMA** (۰،۰،۲)	**SARIMA** $(۲،۳،۲)(۳،۲،۰)_{۱۲}$
۱۷۹۶۶	۰/۵۱۲۴	۰/۲۴۶۳	۰/۳۹۴۱	۰/۴۳۱۸	۳/۱۲	۱۴	**ARIMA** (۱،۱،۳)	**SARIMA** $(۱،۳،۳)(۲،۰،۳)_{۱۲}$
۱۷۹۳۴	۰/۵۰۹۱	۰/۲۵۶۶	۰/۳۶۴۱	۰/۳۹۶۴	۲/۹۶	۱۹	**ARIMA** (۳،۳،۰)	**SARIMA** $(۳،۳،۰)(۳،۰،۲)_{۱۲}$

نمودار سری زمانی پنج ایستگاه هیدرومتری در شکل ۴ نمایش داده شده است.



نمودار سری زمانی ایستگاه ۱۷۰۵۰ — نمودار سری زمانی ایستگاه ۱۷۸۷۴

نمودار سری زمانی ایستگاه ۱۷۹۳۴ — نمودار سری زمانی ایستگاه ۱۷۹۶۵

نمودار سری زمانی ایستگاه ۱۷۹۶۶

شکل ۴ – نمودار سری زمانی ایستگاه‌های هیدرومتری

جدول ۵ نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل ARIMA را نمایش می‌دهد.

1 Cochrane

عملکرد کلی مدل خودرگرسیون یکپارچه میانگین متحرک در تمام ایستگاه‌ها نشان دهنده توانایی متوسط تا خوب در مدل‌سازی آبدهی جریان ماهانه حوضه است، که با توجه به طبیعت غیرخطی، فصلی و پرنوسان داده‌های هیدرولوژیکی در حوزه آبخیز طالقان، نتیجه قابل قبولی محسوب می‌شود. در مجموعه داده آموزش، میانگین ضریب نش‌ساتکلیف در ترکیب ۴ که از داده‌های چهار ماه گذشته استفاده می‌کند برابر با ۰٫۷۵۲ است، که بر اساس طبقه‌بندی استاندارد نش‌ساتکلیف NS بزرگتر از ۰٫۷۵ نشان دهنده مدل بسیار خوب و NS بین ۰٫۶۵ تا ۰٫۷۵ مدل خوب تلقی می‌شود، بیانگر تناسب مناسب مدل با داده‌های آموزشی است و نشان می‌دهد که مدل حدود ۷۵ درصد از واریانس داده‌ها را توضیح می‌دهد، که این مقدار برای سری‌های زمانی با نوسانات بالا ارزشمند است. ریشه میانگین مربعات خطا و میانگین مطلق خطا نیز در این ترکیب پایین‌ترین مقادیر را دارند (میانگین ریشه میانگین مربعات خطا ۱٫۳۳۸ متر مکعب بر ثانیه و میانگین مطلق خطا ۱٫۲۵۲ متر مکعب بر ثانیه)، که بیانگر دقت بالای پیش‌بینی در فاز آموزش است، زیرا ریشه میانگین مربعات خطا خطاهای بزرگ را بیشتر مجازات می‌کند و میانگین مطلق خطا انحراف متوسط را اندازه‌گیری می‌کند و مقادیر پایین هر دو نشان دهنده نزدیکی بیشتر پیش‌بینی‌ها به مقادیر واقعی است. ضریب همبستگی نیز در محدوده ۰٫۷۰۷ تا ۰٫۷۱۳ قرار دارد، که همبستگی خطی قوی را تأیید می‌کند، اما کمتر از ۰٫۸ بودن آن می‌تواند به دلیل وجود الگوهای غیرخطی یا فصلی بودن در داده‌ها باشد، که مدل خودرگرسیون یکپارچه میانگین متحرک به تنهایی ممکن است آن‌ها را کامل پوشش ندهد. با کاهش تعداد داده‌های گذشته از ۴ به ۱ عملکرد مدل کمی کاهش می‌یابد برای مثال ضریب نش‌ساتکلیف در ترکیب ۱ به میانگین ۰٫۷۲۹ می‌رسد، که کاهش حدود ۳ درصدی را نشان می‌دهد و بیانگر حساسیت مدل به حجم اطلاعات ورودی است، زیرا داده‌های گذشته کمتر ممکن است الگوهای بلند مدت را کمتر پوشش دهد و منجر به افزایش خطا شود. در مجموعه آزمون این الگو حفظ می‌شود، اما با کاهش نسبی عملکرد میانگین ضریب نش‌ساتکلیف در ترکیب ۴ برابر ۰٫۷۲۶، که تفاوت حدود ۵-۳ درصدی بین آموزش و آزمون را نشان می‌دهد. این تفاوت می‌تواند به دلیل وجود نویز یا تغییرات غیر منتظره در داده‌های آزمون باشد اما کلیت آن بیانگر تعمیم‌پذیری خوب مدل خودرگرسیون یکپارچه میانگین متحرک در حوزه آبخیز طالقان است زیرا کاهش کمتر از ۵ درصد در معیارها معمولاً به عنوان نشانه پایداری مدل در ادبیات هیدرولوژی تلقی می‌شود و نشان می‌دهد که مدل بیش از حد به داده‌های آموزش وابسته نیست. محاسبه درصد تغییرات در معیارها مانند کاهش ۳٫۶ درصدی NS از ترکیب ۴ به ۱ در آموزش نشان دهنده اهمیت انتخاب ترکیب ورودی مناسب برای بهینه‌سازی مدل است.

جدول ۵ - نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل ARIMA

ایستگاه	شماره ترکیب	مجوعه داده آموزش				مجموعه داده آزمون			
		NS	RMSE	MAE	R	NS	RMSE	MAE	R
جوستان (۱۷۹۶۶)	۴	۰/۷۵۶	۲/۶۹۳	۲/۴۸۶	۰/۷۱۳	۰/۷۲۹	۲/۷۲۳	۲/۵۳۱	۰/۶۸۷
	۳	۰/۷۴۹	۲/۶۹۸	۲/۴۹۳	۰/۷۰۹	۰/۷۲۵	۲/۷۲۷	۲/۵۳۷	۰/۶۸۳
	۲	۰/۷۴۱	۲/۷۰۳	۲/۴۹۷	۰/۷۰۷	۰/۷۱۷	۲/۷۳۱	۲/۵۴۴	۰/۶۸۱
	۱	۰/۷۳۳	۲/۷۱۲	۲/۵۰۳	۰/۷۰۴	۰/۷۱۱	۲/۷۴۲	۲/۵۴۹	۰/۶۷۸
مهران جوستان (۱۷۹۶۵)	۴	۰/۷۵۴	۰/۷۹۸	۰/۷۱۶	۰/۷۱۱	۰/۷۲۸	۰/۸۲۳	۰/۷۴۲	۰/۶۸۴
	۳	۰/۷۴۸	۰/۷۹۹	۰/۷۲۴	۰/۷۰۸	۰/۷۲۴	۰/۸۲۹	۰/۷۴۸	۰/۶۸۲
	۲	۰/۷۳۹	۰/۸۰۲	۰/۷۳۱	۰/۷۰۵	۰/۷۱۵	۰/۸۳۳	۰/۷۵۲	۰/۶۷۹
	۱	۰/۷۳۱	۰/۸۱۳	۰/۷۳۹	۰/۷۰۳	۰/۷۰۷	۰/۸۳۷	۰/۷۵۴	۰/۶۷۵
علیزان جوستان (۱۷۹۳۴)	۴	۰/۷۵۲	۰/۶۱۸	۰/۵۸۳	۰/۷۰۹	۰/۷۲۶	۰/۶۳۹	۰/۵۹۷	۰/۶۸۴
	۳	۰/۷۴۶	۰/۶۲۳	۰/۵۸۶	۰/۷۰۶	۰/۷۲۲	۰/۶۴۲	۰/۵۹۹	۰/۶۸۱
	۲	۰/۷۳۷	۰/۶۲۷	۰/۵۸۹	۰/۷۰۴	۰/۷۱۳	۰/۶۴۷	۰/۶۰۱	۰/۶۷۷
	۱	۰/۷۲۸	۰/۶۳۱	۰/۵۹۵	۰/۷۰۱	۰/۷۰۴	۰/۶۵۱	۰/۶۰۷	۰/۶۷۱
گته‌ده (۱۷۰۵۰)	۴	۰/۷۵۱	۰/۶۳۳	۰/۵۹۵	۰/۷۰۷	۰/۷۲۴	۰/۶۵۴	۰/۶۱۱	۰/۶۸۳
	۳	۰/۷۴۵	۰/۶۳۷	۰/۵۹۷	۰/۷۰۲	۰/۷۱۹	۰/۶۵۷	۰/۶۱۲	۰/۶۷۵
	۲	۰/۷۳۴	۰/۶۴۱	۰/۵۹۸	۰/۶۹۸	۰/۷۱۲	۰/۶۶۲	۰/۶۱۱	۰/۶۶۸
	۱	۰/۷۲۷	۰/۶۴۲	۰/۶۰۳	۰/۶۹۶	۰/۷۰۲	/۶۶۵	۰/۶۱۴	۰/۶۶۳
دهدر (۱۷۸۷۴)	۴	۰/۷۴۹	۰/۵۰۷	۰/۴۷۱	۰/۷۰۵	۰/۷۲۲	۰/۵۲۵	۰/۴۹۳	۰/۶۷۹
	۳	۰/۷۴۴	۰/۵۱۱	۰/۴۷۴	۰/۶۹۹	۰/۷۱۷	۰/۵۲۷	۰/۴۹۴	۰/۶۷۳
	۲	۰/۷۳۵	۰/۵۱۴	۰/۴۷۷	۰/۶۹۵	۰/۷۱۱	۰/۵۲۹	۰/۴۹۴	۰/۶۶۴
	۱	۰/۷۲۶	۰/۵۱۹	۰/۴۸۱	۰/۶۹۳	۰/۷۰۱	۰/۵۳۱	۰/۴۹۹	۰/۶۶۱

عملکرد کلی مدل SARIMA در مدل‌سازی آبدهی جریان ماهانه حوضه در حوزه آبخیز طالقان، حاکی از توانایی برجسته این مدل

در شناسایی و مدل‌سازی الگوهای فصلی و دوره‌ای موجود در داده‌ها است، که این ویژگی آن را به ابزاری قدرتمند برای تحلیل سری‌های زمانی هیدرولوژیکی با نوسانات فصلی تبدیل می‌کند. جدول ۶ نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل SARIMA را نمایش می‌دهد و در مجموعه داده آموزش، میانگین ضریب نش‌ساتکلیف در ترکیب ۴ برابر با ۰٫۸۲۵ است، که بر اساس معیارهای هیدرولوژیکی (NS بزرگتر از ۰٫۷۵ نشان‌دهنده مدل بسیار خوب) بیانگر تناسب عالی با داده‌های آموزشی است. این مقدار نشان می‌دهد که مدل توانسته حدود ۸۲٫۵ درصد از واریانس موجود در داده‌های دبی را توضیح دهد، که دستاوردی قابل توجه در مقایسه با مدل‌های ساده‌تر است. ریشه میانگین مربعات خطا و میانگین مطلق خطا نیز در این ترکیب به ترتیب میانگین ۰٫۷۷۳ و ۰٫۷۳۲ را نشان می‌دهند، که مقادیر بسیار پایینی هستند و دقت بالای پیش‌بینی را تأیید می‌کنند. ریشه میانگین مربعات خطا با تأکید بر کاهش تأثیر خطاهای بزرگ و میانگین مطلق خطا با تمرکز بر انحراف متوسط، هر دو حاکی از نزدیکی پیش‌بینی‌ها به مقادیر واقعی هستند، که نشان دهنده موفقیت مدل در فاز آموزش است. ضریب همبستگی نیز در محدوده ۰٫۷۸۸ تا ۰٫۷۹۳ قرار دارد، که همبستگی خطی قوی را نشان می‌دهد، اما مقادیر کمتر از ۰٫۸ ممکن است به دلیل وجود نوسانات غیرخطی، نویز محیطی یا اثرات اقلیمی غیرقابل پیش‌بینی باشد که حتی مدل فصلی نیز نمی‌تواند آن‌ها را به طور کامل حذف کند.

جدول ۶ - نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل SARIMA

ایستگاه	شماره ترکیب	مجموعه داده آموزش				مجموعه داده آزمون			
		NS	RMSE	MAE	R	NS	RMSE	MAE	R
جوستان	۴	۰/۸۲۵	۱/۶۴۸	۱/۵۵۴	۰/۷۹۳	۰/۸۰۷	۱/۶۷۲	۱/۵۸۶	۰/۷۷۳
(۱۷۹۶۶)	۳	۰/۸۱۷	۱/۶۵۱	۱/۵۶۳	۰/۷۸۶	۰/۸۰۲	۱/۶۷۷	۱/۵۹۳	۰/۷۶۶
	۲	۰/۸۱۲	۱/۶۵۹	۱/۵۷۱	۰/۷۸۱	۰/۷۹۴	۱/۶۸۴	۱/۵۹۴	۰/۷۶۱
	۱	۰/۸۰۶	۱/۶۶۳	۱/۵۷۳	۰/۷۷۴	۰/۷۸۶	۱/۶۸۷	۱/۵۹۹	۰/۷۵۳
مهران	۴	۰/۸۲۴	۰/۴۹۸	۰/۴۶۳	۰/۷۹۳	۰/۸۰۵	۰/۵۰۶	۰/۴۸۴	۰/۷۷۲
جوستان	۳	۰/۸۱۶	۰/۴۹۹	۰/۴۶۵	۰/۷۸۵	۰/۸۰۱	۰/۵۱۱	۰/۴۸۷	۰/۷۶۴
(۱۷۹۶۵)	۲	۰/۸۱۱	۰/۵۰۱	۰/۴۶۵	۰/۷۷۹	۰/۷۹۳	۰/۵۱۴	۰/۴۸۷	۰/۷۵۹
	۱	۰/۸۰۳	۰/۵۰۳	۰/۴۷۳	۰/۷۷۳	۰/۷۸۵	۰/۵۱۹	۰/۴۹۱	۰/۷۵۲
علیزان	۴	۰/۸۲۴	۰/۳۹۲	۰/۳۶۵	۰/۷۹۲	۰/۸۰۴	۰/۴۱۳	۰/۳۸۴	۰/۷۷۱
جوستان	۳	۰/۸۱۴	۰/۳۹۸	۰/۳۶۷	۰/۷۸۴	۰/۷۹۹	۰/۴۱۷	۰/۳۸۷	۰/۷۶۲
(۱۷۹۳۴)	۲	۰/۸۰۹	۰/۴۰۱	۰/۳۶۹	۰/۷۷۷	۰/۷۹۱	۰/۴۲۱	۰/۳۸۸	۰/۷۵۸
	۱	۰/۸۰۲	۰/۴۰۳	۰/۳۷۱	۰/۷۷۲	۰/۷۸۱	۰/۴۲۴	۰/۳۹۳	۰/۷۵۱
گته‌ده	۴	۰/۸۲۳	۰/۴۰۴	۰/۳۷۲	۰/۷۹۱	۰/۸۰۳	۰/۴۲۷	۰/۳۹۳	۰/۷۶۸
(۱۷۰۵۰)	۳	۰/۸۱۲	۰/۴۰۷	۰/۳۷۶	۰/۷۸۳	۰/۷۹۴	۰/۴۳۱	۰/۳۹۴	۰/۷۶۱
	۲	۰/۸۰۷	۰/۴۰۹	۰/۳۷۹	۰/۷۷۴	۰/۷۸۹	۰/۴۳۳	۰/۳۹۷	۰/۷۵۶
	۱	۰/۸۰۱	۰/۴۱۱	۰/۳۸۲	۰/۷۷۱	۰/۷۷۸	۰/۴۳۵	۰/۳۹۹	۰/۷۴۸
دهدر	۴	۰/۸۱۸	۰/۳۲۴	۰/۳۰۹	۰/۷۸۸	۰/۸۰۳	۰/۳۳۷	۰/۳۱۷	۰/۷۶۵
(۱۷۸۷۴)	۳	۰/۸۱۱	۰/۳۲۷	۰/۳۱۱	۰/۷۸۲	۰/۷۹۳	۰/۳۳۹	۰/۳۱۸	۰/۷۵۹
	۲	۰/۸۰۵	۰/۳۲۹	۰/۳۱۴	۰/۷۷۲	۰/۷۸۸	۰/۳۴۱	۰/۳۱۸	۰/۷۵۵
	۱	۰/۷۹۴	۰/۳۳۱	۰/۳۱۶	۰/۷۶۸	۰/۷۷۵	۰/۳۴۳	۰/۳۲۴	۰/۷۴۴

با کاهش تعداد داده‌های گذشته از ۴ به ۱، عملکرد مدل به طور ملموسی کاهش می‌یابد، میانگین ضریب نش‌ساتکلیف در ترکیب ۱ به ۰٫۸۰۳ می‌رسد، که کاهش حدود ۲٫۷ درصدی را نشان می‌دهد و بیانگر وابستگی قوی مدل به اطلاعات گسترده‌تر از دوره‌های گذشته است. این کاهش می‌تواند ناشی از ناتوانی مدل در ثبت الگوهای فصلی و دوره‌ای با داده‌های محدودتر باشد به ویژه در منطقه‌ای مانند طالقان که فصول مختلف تأثیرات متفاوتی بر دبی دارند (مثلاً بارش‌های سنگین زمستانی و جریان‌های پایه تابستانی). در مجموعه داده آزمون، میانگین ضریب نش‌ساتکلیف در ترکیب ۴ برابر ۰٫۸۰۵ است، که کاهش ۲٫۴ درصدی نسبت به آموزش را نشان می‌دهد اما همچنان در محدوده مدل‌های بسیار خوب (NS بزرگتر از ۰٫۷۵) قرار دارد. تعمیم‌پذیری قابل قبول مدل را تأیید می‌کند زیرا کاهش کمتر از ۳ درصد معمولاً به عنوان نشانه پایداری مدل در ادبیات هیدرولوژی تلقی می‌شود. همچنین محاسبه ضریب تغییرات برای RMSE در ترکیب ۴

تقریباً ۰٫۳۸ در کل ایستگاه‌ها نشان دهنده پراکندگی نسبی خطاها است که با افزایش واریانس دبی به طور قابل توجهی افزایش می‌یابد و این امر لزوم توجه به ویژگی‌های هیدرولوژیکی خاص هر ایستگاه را خاطر نشان می‌کند. محاسبات تکمیلی مانند میانگین هندسی RMSE (حدود ۰٫۵۹) نیز نشان می‌دهد که مدل در ایستگاه‌های با دبی پایین‌تر عملکرد بهتری دارد.

عملکرد کلی مدل ELM در جدول ۷ در مدل‌سازی آبدهی جریان ماهانه حوضه در حوزه آبخیز طالقان، نشان دهنده توانایی فوق العاده این مدل در شناسایی و پیش‌بینی الگوهای پیچیده و غیرخطی موجود در داده‌ها است که این ویژگی آن را از روش‌های کلاسیک متمایز می‌کند و به دلیل ساختار یادگیری سریع و انعطاف‌پذیرش آن را به ابزاری ایده آل برای تحلیل‌های هیدرولوژیکی تبدیل می‌کند. با بررسی جدول ۷ در مجموعه داده آموزش، میانگین ضریب نش‌ساتکلیف در ترکیب ۴ برابر با ۰٫۹۲۷ است که بر اساس معیارهای هیدرولوژیکی بیانگر تطابق استثنایی با داده‌های آموزشی است. این مقدار نشان می‌دهد که مدل توانسته حدود ۹۲٫۷ درصد از واریانس موجود در داده‌های دبی را توضیح دهد که دستاوردی برجسته در مقایسه با مدل‌های سنتی است و بیانگر قدرت بالای مدل ELM در پردازش داده‌های هیدرولوژیکی با نوسانات فصلی، نویز محیطی و تغییرات اقلیمی در طالقان است. ریشه میانگین مربعات خطا و میانگین مطلق خطا نیز در این ترکیب به ترتیب میانگین ۰٫۳۸۴ و ۰٫۳۵۳ را نشان می‌دهند که مقادیر بسیار پایینی هستند و دقت بی‌نظیر پیش‌بینی را تأیید می‌کنند. ریشه میانگین مربعات خطا با تأکید بر کاهش تأثیر خطاهای بزرگ و میانگین مطلق خطا با تمرکز بر انحراف متوسط، هر دو حاکی از نزدیکی چشمگیر پیش‌بینی‌ها به مقادیر واقعی هستند که نشان دهنده موفقیت مدل در فاز آموزش است. ضریب همبستگی نیز در محدوده ۰٫۸۸۱ تا ۰٫۸۹۷ قرار دارد که همبستگی خطی بسیار قوی را نشان می‌دهد و بیانگر توانایی مدل در بازتولید الگوهای دبی با دقت بالا است اگرچه مقادیر کمتر از ۰٫۹ ممکن است به دلیل وجود نویز یا اثرات غیرخطی پیچیده باشد که حتی مدل ELM نیز نمی‌تواند آن‌ها را به طور کامل حذف کند.

جدول ۷ - نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل ELM

ایستگاه	شماره ترکیب	مجموعه داده آموزش				مجموعه داده آزمون			
		NS	RMSE	MAE	R	NS	RMSE	MAE	R
جوستان	۴	۰/۹۲۷	۰/۸۲۸	۰/۸۰۱	۰/۸۹۹	۰/۹۱۲	۰/۸۴۱	۰/۸۲۱	۰/۸۸۲
(۱۷۹۶۶)	۳	۰/۹۲۱	۰/۸۳۱	۰/۸۰۴	۰/۸۹۳	۰/۹۰۴	۰/۸۴۳	۰/۸۲۲	۰/۸۷۶
	۲	۰/۹۱۸	۰/۸۳۴	۰/۸۰۸	۰/۸۸۶	۰/۹۰۱	۰/۸۴۵	۰/۸۲۴	۰/۸۷۱
	۱	۰/۹۰۹	۰/۸۳۸	۰/۸۰۹	۰/۸۸۱	۰/۸۹۶	۰/۸۵۳	۰/۸۲۹	۰/۸۶۵
مهران	۴	۰/۹۲۵	۰/۲۳۷	۰/۲۰۴	۰/۸۹۷	۰/۹۰۹	۰/۲۵۳	۰/۲۲۱	۰/۸۸۱
جوستان	۳	۰/۹۱۹	۰/۲۳۹	۰/۲۰۹	۰/۸۹۲	۰/۹۰۳	۰/۲۵۷	۰/۲۲۴	۰/۸۷۵
(۱۷۹۶۵)	۲	۰/۹۱۶	۰/۲۴۱	۰/۲۱۳	۰/۸۸۵	۰/۸۹۹	۰/۲۶۱	۰/۲۲۷	۰/۸۶۹
	۱	۰/۹۰۷	۰/۲۴۴	۰/۲۱۷	۰/۸۷۹	۰/۸۹۳	۰/۲۶۴	۰/۲۳۱	۰/۸۶۳
علیزان	۴	۰/۹۲۳	۰/۱۹۸	۰/۱۸۶	۰/۸۹۶	۰/۹۰۴	۰/۲۰۶	۰/۱۹۱	۰/۸۸۱
جوستان	۳	۰/۹۱۸	۰/۱۹۸	۰/۱۸۷	۰/۸۹۱	۰/۹۰۱	۰/۲۰۹	۰/۱۹۲	۰/۸۷۴
(۱۷۹۳۴)	۲	۰/۹۱۵	۰/۱۹۹	۰/۱۸۸	۰/۸۸۴	۰/۸۹۶	۰/۲۱۲	۰/۱۹۲	۰/۸۶۸
	۱	۰/۹۰۶	۰/۲۰۱	۰/۱۹۳	۰/۸۷۷	۰/۸۹۱	۰/۲۱۴	۰/۱۹۴	۰/۸۶۱
گته‌ده	۴	۰/۹۲۱	۰/۲۰۴	۰/۱۹۴	۰/۸۹۴	۰/۹۰۳	۰/۲۱۶	۰/۱۹۵	۰/۸۷۹
(۱۷۰۵۰)	۳	۰/۹۱۶	۰/۲۰۷	۰/۱۹۴	۰/۸۸۹	۰/۸۹۹	۰/۲۱۷	۰/۱۹۵	۰/۸۷۳
	۲	۰/۹۱۴	۰/۲۰۹	۰/۱۹۶	۰/۸۸۳	۰/۸۹۵	۰/۲۱۹	۰/۱۹۷	۰/۸۶۶
	۱	۰/۹۰۳	۰/۲۱۱	۰/۱۹۷	۰/۸۷۶	۰/۸۸۹	۰/۲۲۱	۰/۱۹۹	۰/۸۵۹
دهدر	۴	۰/۹۱۴	۰/۱۵۳	۰/۱۳۱	۰/۸۹۱	۰/۹۰۳	۰/۱۶۵	۰/۱۴۳	۰/۸۷۷
(۱۷۸۷۴)	۳	۰/۹۱۲	۰/۱۵۵	۰/۱۳۴	۰/۸۸۷	۰/۸۹۷	۰/۱۶۷	۰/۱۴۴	۰/۸۷۱
	۲	۰/۹۰۹	۰/۱۵۷	۰/۱۳۵	۰/۸۸۱	۰/۸۹۴	۰/۱۶۹	۰/۱۴۷	۰/۸۶۵
	۱	۰/۹۰۱	۰/۱۵۹	۰/۱۳۷	۰/۸۷۳	۰/۸۸۵	۰/۱۷۳	۰/۱۵۱	۰/۸۵۷

عملکرد کلی مدل XGBoost در جدول ۸ در مدل‌سازی آبدهی جریان ماهانه حوضه در حوزه آبخیز طالقان، نشان دهنده توانایی بی‌نظیر این مدل در شناسایی و پیش‌بینی الگوهای پیچیده و غیرخطی موجود در داده‌ها است، که آن را به یکی از قدرتمندترین ابزارها در تحلیل سری‌های زمانی هیدرولوژیکی تبدیل می‌کند. با بررسی جدول ۸ در مجموعه داده آموزش، میانگین ضریب نش‌ساتکلیف در ترکیب

۴ برابر با ۰،۹۷۸ است که بیانگر تطابق استثنایی با داده‌های آموزشی و توضیح حدود ۹۷،۸ درصد از واریانس موجود در دبی است. این مقدار همراه با ریشه میانگین مربعات خطا (۰،۱۴۱) و میانگین مطلق خطا (۰،۱۲۴) نشان‌دهنده دقت بی‌سابقه‌ای است که مدل در فاز آموزش به دست آورده است. ضریب همبستگی (۰،۹۵۱ تا ۰،۹۶۶) نیز همبستگی خطی بسیار قوی را تأیید می‌کند اگرچه تفاوت اندک با ۱ ممکن است به نویز محیطی یا اثرات اقلیمی غیرقابل پیش‌بینی مرتبط باشد. با کاهش تعداد داده‌های گذشته به ترکیب ۱، ضریب نش‌ساتکلیف به ۰،۹۶۳ کاهش می‌یابد (کاهش ۱،۵ درصد) که نشان دهنده اهمیت داده‌های بلندمدت برای حفظ دقت مدل است. در مجموعه داده آزمون میانگین ضریب نش‌ساتکلیف در ترکیب ۴ برابر ۰،۹۵۷ است که کاهش ۲،۱٪ نسبت به آموزش را نشان می‌دهد و تعمیم‌پذیری استثنایی مدل را تأیید می‌کند. این کاهش اندک (کمتر از ۲،۵ درصد) بیانگر پایداری بالای مدل در برابر داده‌های جدید است.

جدول ۸ - نتایج آماری داده‌های ورودی جهت مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه حوضه با مدل XGBOOST

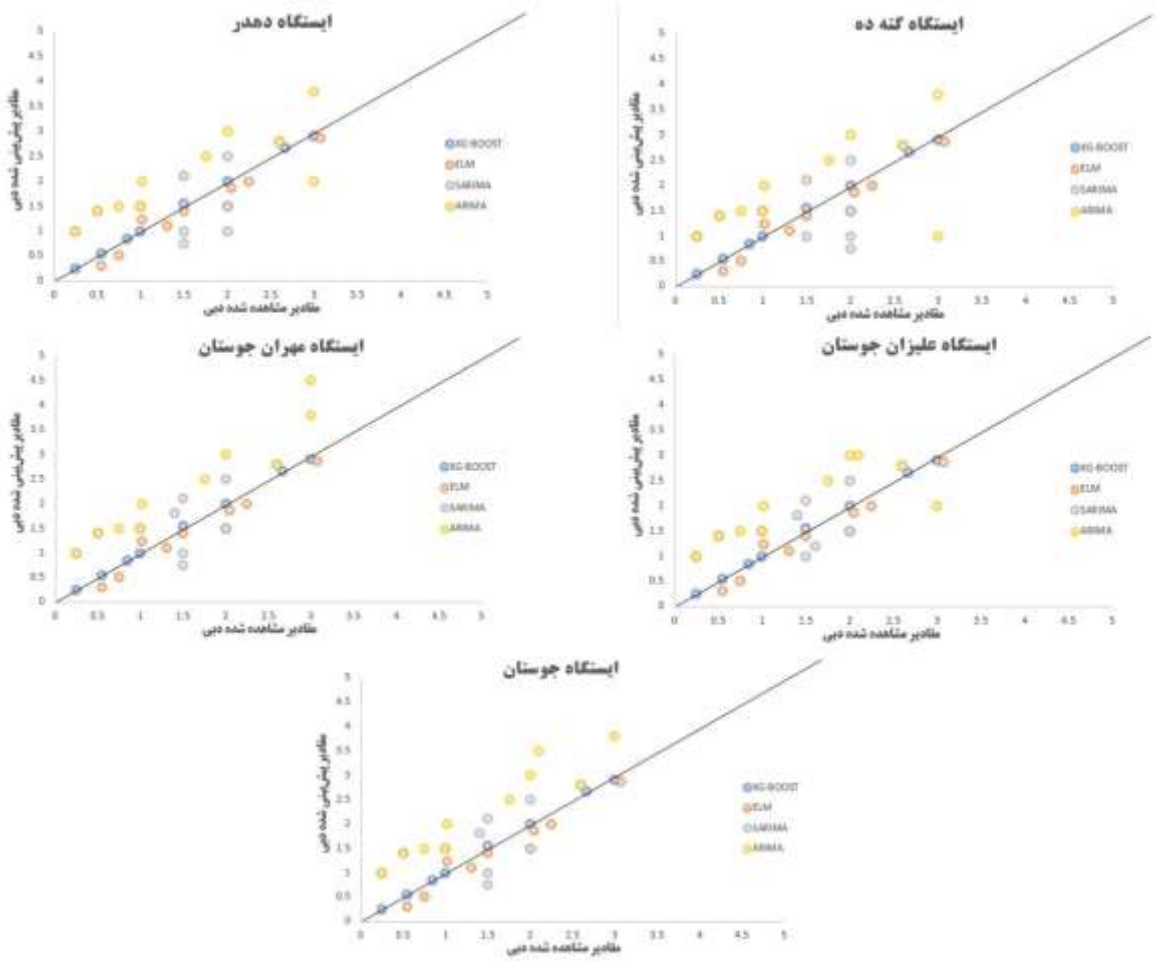
ایستگاه	شماره ترکیب	مجوعه داده آموزش				مجموعه داده آزمون			
		NS	RMSE	MAE	R	NS	RMSE	MAE	R
جوستان	۴	۰/۹۷۸	۰/۳۶۲	۰/۳۴۸	۰/۹۶۶	۰/۹۶۱	۰/۳۷۶	۰/۳۵۴	۰/۹۴۷
(۱۷۹۶۶)	۳	۰/۹۷۳	۰/۳۶۴	۰/۳۴۸	۰/۹۶۲	۰/۹۵۷	۰/۳۷۸	۰/۳۵۵	۰/۹۴۱
	۲	۰/۹۶۹	۰/۳۶۷	۰/۳۴۹	۰/۹۵۵	۰/۹۵۱	۰/۳۷۹	۰/۳۵۸	۰/۹۳۵
	۱	۰/۹۶۳	۰/۳۷۱	۰/۳۵۱	۰/۹۵۱	۰/۹۴۶	۰/۳۸۳	۰/۳۶۱	۰/۹۳۱
مهران	۴	۰/۹۷۷	۰/۱۰۴	۰/۰۸۷	۰/۹۶۳	۰/۹۵۹	۰/۱۱۵	۰/۰۹۴	۰/۹۴۷
جوستان	۳	۰/۹۷۲	۰/۱۰۴	۰/۰۸۹	۰/۹۶۱	۰/۹۵۴	۰/۱۱۷	۰/۰۹۷	۰/۹۳۹
(۱۷۹۶۵)	۲	۰/۹۶۸	۰/۱۰۸	۰/۰۹۱	۰/۹۵۴	۰/۹۴۸	۰/۱۱۷	۰/۰۹۹	۰/۹۳۴
	۱	۰/۹۶۱	۰/۱۱۲	۰/۰۹۳	۰/۹۴۸	۰/۹۴۴	۰/۱۲۱	۰/۱۰۳	۰/۹۲۹
علیزان	۴	۰/۹۷۵	۰/۰۸۵	۰/۰۶۹	۰/۹۶۲	۰/۹۵۷	۰/۰۹۳	/۰۷۷	۰/۹۴۶
جوستان	۳	۰/۹۷۱	۰/۰۸۶	۰/۰۶۹	۰/۹۵۹	۰/۹۵۳	۰/۰۹۵	۰/۰۷۷	۰/۹۳۷
(۱۷۹۳۴)	۲	۰/۹۶۶	۰/۰۸۶	۰/۰۷۱	۰/۹۵۳	۰/۹۴۷	۰/۰۹۷	۰/۰۷۹	۰/۹۳۳
	۱	۰/۹۵۹	۰/۰۸۹	۰/۰۷۳	۰/۹۴۷	۰/۹۴۳	۰/۱۰۱	۰/۰۸۱	۰/۹۲۷
گتەده	۴	۰/۹۷۴	۰/۰۹۱	۰/۰۷۴	۰/۹۶۱	۰/۹۵۵	۰/۱۰۳	۰/۰۸۲	۰/۹۴۴
(۱۷۰۵۰)	۳	۰/۹۶۹	۰/۰۹۲	۰/۰۷۴	۰/۹۵۸	۰/۹۵۱	۰/۱۰۴	۰/۰۸۲	۰/۹۳۶
	۲	۰/۹۶۵	۰/۰۹۲	۰/۰۷۷	۰/۹۵۲	۰/۹۴۵	۰/۱۰۶	۰/۰۸۴	۰/۹۳۲
	۱	۰/۹۵۸	۰/۰۹۶	۰/۰۷۹	۰/۹۴۴	۰/۹۴۱	۰/۱۰۸	۰/۰۸۷	۰/۹۲۴
دهدر	۴	۰/۹۷۳	۰/۰۶۳	۰/۰۴۳	۰/۹۵۸	۰/۹۵۴	۰/۰۷۵	۰/۰۵۴	۰/۹۴۱
(۱۷۸۷۴)	۳	۰/۹۶۸	۰/۰۶۳	۰/۰۴۴	۰/۹۵۵	۰/۹۴۹	۰/۰۷۶	۰/۰۵۴	۰/۹۳۵
	۲	۰/۹۶۴	۰/۰۶۵	۰/۰۴۶	۰/۹۵۱	۰/۹۴۳	۰/۰۷۷	۰/۰۵۶	۰/۹۲۸
	۱	۰/۹۵۶	۰/۰۶۷	۰/۰۴۷	۰/۹۴۳	۰/۹۳۹	۰/۰۷۹	۰/۰۵۹	۰/۹۲۳

بحث

در این بخش بر اساس شکل ۵ عملکرد مدل‌های (XGBoost، ELM، SARIMA و ARIMA) در پنج ایستگاه هیدرومتری مورد مطالعه (دهدر، گتەده، جوستان، مهران جوستان و علیزان جوستان) در پیش بینی آبدهی جریان ماهانه حوضه از طریق نمودارهای پراکندگی مقایسه با نیمساز ارزیابی شد. محور افقی هر نمودار مقادیر واقعی مشاهده شده دبی را نشان می‌دهد در حالی که محور عمودی به مقادیر پیش‌بینی شده اختصاص یافته است و خط مورب (y=x) به عنوان معیار تطابق کامل عمل می‌کند. در تمامی ایستگاه‌ها مدل XGBoost با کمترین پراکندگی و انحراف نقاط حول خط مورب بهترین عملکرد را از خود نشان داد که این امر حاکی از RMSE پایین، همبستگی قوی و دقت بالا در کل دامنه مقادیر است به ویژه در شرایط حدی[1] و نوسانات فصلی. در مقابل، مدل ARIMA با بیشترین فاصله از خط مورب به خصوص در مقادیر بالا و پایین، ضعیف‌ترین مدل شناخته شد و نشان دهنده انحراف سیستماتیک[2](مانند بیش برآورد[3] یا کم برآورد[4]) و واریانس بالا بود در حالی که مدل ELM و مدل SARIMA عملکرد متوسطی ارائه کردند مدل SARIMA اغلب در داده‌های سری

1 Extreme conditions
2 Systematic Bias
3 over estimation
4 under-estimation

زمانی و فصلی برتری نسبی داشت اما همچنان از مدل‌های مبتنی بر یادگیری ماشین عقب ماند.



شکل ۵- نمودار مقایسه با نیمساز

با وجود برتری مدل‌های یادگیری ماشین بر مدل‌های کلاسیک در مدل‌سازی و پیش‌بینی آبدهی جریان ماهانه در حوزه آبخیز طالقان، نتایج این مطالعه با محدودیت‌هایی روبرو است. نخست، تمرکز صرف بر داده‌های دبی متوسط ماهانه بدون ادغام متغیرهای کلیدی هیدرواقلیمی نظیر شدت بارش لحظه‌ای، شاخص‌های ذوب برف یا دمای سطحی، دقت مدل‌ها را در سناریوهای پویا کاهش می‌دهد. دوم، حجم نمونه محدود داده‌های ماهانه تلاش برای کاهش نویز، واریانس مدل‌های غیرخطی مانند ELM را افزایش می‌دهد، به ویژه با توجه به مقداردهی تصادفی وزن‌ها که ثبات نتایج را در تکرارهای مستقل تضعیف می‌کند. سوم، افق پیش‌بینی کوتاه‌مدت (یک ماهه) بدون ارزیابی بلندمدت یا شبیه‌سازی سناریوهای تغییر اقلیم، مانند افزایش شدت رویدادهای شدید تحت تأثیر گرم شدن جهانی در ایران، قابلیت پیش‌بینی در شرایط آینده را محدود می‌سازد. در نهایت، تمرکز ایستگاه محور بر پنج هیدرومتر بدون تحلیل مکانی–زمانی یکپارچه، اثرات هتروژنیتی حوزه (مانند تفاوت‌های توپوگرافی بالادست و پایین‌دست) را به میزان کافی پوشش نمی‌دهد.

نتایج پژوهش حاضر با مطالعه یانگ (۲۰۰۲) که به توسعه مدل‌سازی مکانیکی مبتنی بر داده (DBM) برای پیش‌بینی سیلاب بلادرنگ با تأکید بر کارایی پارامتریک و تفسیرپذیری هیدرولوژیکی پرداخت و از مدل تابع انتقال (TF) غیر خطی مبتنی بر جریان به عنوان جانشین ذخیره حوزه و فیلتر کالمن تطبیقی برای پیش‌بینی چند گامه استفاده کرد همخوانی دارد زیرا XGBoost نیز تعاملات غیرایستا را با تنظیم‌سازی منظم مدیریت می‌کند و RMSE را تا ۷۰٪ نسبت به ARIMA کاهش می‌دهد یافته‌های کومار و همکاران (۲۰۲۳) مبنی بر برتری CatBoost در رود نارمادا (هند) با پژوهش حاضر همسو است زیرا XGBoost مشابه CatBoost در محیط‌های داده محدود برتر عمل می‌کند، اما نیاز به ادغام متغیرهای اقلیمی (مانند بارش) برای تعمیم‌پذیری بیشتر ضروری است. این مقایسه برتری رویکرد داده‌محور پژوهش حاضر را در حوضه‌های نیمه‌خشک ایران تأیید می‌کند.

نتیجه‌گیری

این پژوهش با هدف مقایسه کارایی مدل‌های آماری کلاسیک ARIMA و SARIMA و الگوریتم‌های یادگیری ماشین ELM و XGBoost در مدل‌سازی و پیش بینی آبدهی جریان ماهانه حوزه آبخیز طالقان انجام شد که بر پایه سری‌های زمانی دبی متوسط ماهانه به عنوان ورودی‌های مدل‌سازی در بازه ۳۰ ساله آبی از اوایل سال آبی ۱۳۶۸ تا انتهای سال آبی ۱۳۹۸ استوار بود. یافته‌ها نشان داد که الگوریتم‌های یادگیری ماشین به ویژه مدل XGBoost با بهره‌گیری از ساختارهای پویای تصمیم‌گیری و توانایی مدیریت روابط غیرخطی، دقت بالاتری در پیش‌بینی آبدهی جریان ماهانه ارائه می‌دهند. این برتری به ویژه در ایستگاه‌هایی مانند جوستان مشهود است جایی که مدل‌های آماری به دلیل نیاز به ایستایی داده‌ها با چالش‌هایی مواجه شدند. تحلیل عملکرد مدل‌ها حاکی از آن است که ضریب نش‌ساتکلیف در الگوریتم یادگیری ماشین XGBoost برای ایستگاه دهدر به ۰/۹۷۳ در فاز آموزش و ۰/۹۵۴ در فاز آزمون رسید، که نشان دهنده توانایی این مدل در تعمیم‌پذیری است. در مقابل، مدل‌های سنتی ARIMA و SARIMA هر چند در ثبت الگوهای فصلی با میانگین دبی ۳/۴۲۹ متر مکعب بر ثانیه در بهار برای ایستگاه دهدر موفق عمل کردند، اما در برابر نوسانات غیرخطی و نویزهای داده‌ای محدودیت‌هایی نشان دادند که با افزایش RMSE تا ۲/۷۲۳ در ایستگاه جوستان مشهود است. بررسی تأثیر تعداد ماه‌های ورودی نیز نشان داد که ترکیب چهار ماه با کاهش میانگین مطلق خطا تا ۰/۰۱۳۶ در مدل XGBoost بهترین تعادل را بین دقت و پایداری مدل‌ها ایجاد می‌کند. این نتایج تأکید دارند که پیش پردازش داده‌ها و انتخاب متغیرهای مناسب، مانند داده‌های فصلی، نقشی کلیدی در بهبود کارایی مدل‌ها ایفا می‌کند. در نهایت، برتری مدل XGBoost به دلیل انعطاف‌پذیری در تنظیم پارامترها و مدل ELM به دلیل سرعت محاسباتی بالا، راه را برای توسعه ابزارهای پیش‌بینی آبدهی جریان ماهانه هموار می‌کند.

"هیچ گونه تعارض منافعی بین نویسندگان وجود ندارد"

منابع

سیدیان، سیدمرتضی؛ سلیمانی، مریم و کاشانی، مجتبی (۱۳۹۳). پیش‌بینی دبی جریان رودخانه با استفاده از داده‌کاوی و سری زمانی. *اکوهیدرولوژی*، ۱(۲)، ۱۶۷–۱۷۹.

ابراهیمی، پیام؛ سلاجقه، علی؛ محسنی ساروی، محسن؛ ملکیان، آرش و سعدالدین، امیر (۱۳۹۹). پیش‌بینی سلامت محیطی با استفاده از برنامه‌ریزی بیان ژن و شبکه بیزین در حوزهٔ آبخیز طالقان. مرتع و آبخیزداری، *مجله منابع طبیعی ایران*، ۷۳(۱)، ۱–۱۸.

علی‌احمدی، ندا؛ مرادی، ابراهیم؛ حسینی، سیدمهدی و سردارشهرکی، علی (۱۴۰۰). پیش‌بینی دبی رودخانه هیرمند با استفاده از تکنیک سری زمانی (SARIMA). *نشریه علمی پژوهشی مهندسی آبیاری و آب ایران*، ۱۲(۴۵)، ۱۷۲–۱۹۱.

REFERENCES

Ali Ahmadi, N., Moradi, E., Hosseini, S. M., and Sardar Shahraki, A. (2021). Forecasting the discharge of Hirmand River using time series technique (SARIMA). Iranian Journal of Irrigation and Water Engineering, 12(45), 172-191. (in persian)

Bărbulescu, A., & Zhen, L. (2024). Forecasting the river water discharge by artificial intelligence methods. *Water (Switzerland)*, 16(9). https://doi.org/10.3390/w16091248

Bărbulescu, A., & Zhen, L. (2024). Forecasting the River Water Discharge by Artificial Intelligence Methods. Water (Switzerland), 16(9). https://doi.org/10.3390/w16091248

Brunner, M. I., Slater, L., Tallaksen, L. M., & Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. Wiley Interdisciplinary Reviews: *Water*, 8(3), 1–32. https://doi.org/10.1002/wat2.1520

Chattopadhyay, S., & Chattopadhyay, G. (2010). Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN. Comptes Rendus Geoscience, 342(2), 100–107. https://doi.org/10.1016/j.crte.2009.10.016

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17 August, 785–794. https://doi.org/10.1145/2939672.2939785

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

Ebrahimi, P., Salamat, A., Mohseni Saravi, M., Malekian, A., and Sa'adoddin, A. (2020). Prediction of environmental health using gene expression programming and Bayesian network in Taleghan watershed.

Journal of Rangeland and Watershed Management, Iranian Journal of Natural Resources, 73(1). (in persian)

Ebtehaj, I., & Bonakdari, H. (2022). A reliable hybrid outlier robust non-tuned rapid machine learning model for multi-step ahead flood forecasting in Quebec, Canada. *Journal of Hydrology*, 614(PB), 128592. https://doi.org/10.1016/j.jhydrol.2022.128592

Elshorbagy, A., Corzo, G., Srinivasulu, S., & Solomatine, D. P. (2010). Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology. *Hydrology and Earth System Sciences*, 14(10), 1931–1941. https://doi.org/10.5194/hess-14-1931-2010

Huang, G. Bin, Zhu, Q. Y., & Siew, C. K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. IEEE International Conference on Neural Networks - Conference Proceedings, 2(August 2004), 985–990. https://doi.org/10.1109/IJCNN.2004.1380068

Kumar, V., Kedam, N., Sharma, K. V., Mehta, D. J., & Caloiero, T. (2023). Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. Water (Switzerland), 15(14).Jia, W., Chen, M., Yao, H., Wang, Y., Wang, S., & Ni, X. (2024). Improving sub-daily runoff forecast based on the multi-objective optimized extreme learning machine for reservoir operation. *Water Resources Management*, 38(15), 6173–6189. https://doi.org/10.1007/s11269-024-03953-2

Kumar, V., Sharma, K. V., Caloiero, T., Mehta, D. J., & Singh, K. (2023). Comprehensive overview of flood modeling approaches: A review of recent advances. *Hydrology*, 10(7). https://doi.org/10.3390/hydrology10070141

Moradian, S., AghaKouchak, A., Gharbia, S., Broderick, C., & Olbert, A. I. (2024). Forecasting of compound ocean-fluvial floods using machine learning. *Journal of Environmental Management*, 364(April), 121295. https://doi.org/10.1016/j.jenvman.2024.121295

Panicker, N. K. K., & Valarmathi, J. (2025). Time series prediction of aerosol optical depth across the northern Indian region: Integrating PSO-optimized SARIMA-SVR based on MODIS data. *Acta Geophysica*, 73, 2097–2126. https://doi.org/10.1007/s11600-024-01472-7

Seidian, S. M., Soleimani, M., and Kashani, M. (2014). Forecasting river flow discharge using data mining and time series. Ecohydrology, 1(2), 167-179. (in persian)

Szczepanek, R. (2022). Daily streamflow forecasting in mountainous catchment using XGBoost, LightGBM and CatBoost. *Hydrology*, 9(12). https://doi.org/10.3390/hydrology9120226

Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of Hydrology, 476, 433–441. https://doi.org/10.1016/j.jhydrol.2012.11.017

Van Thieu, N., Nguyen, N. H., Sherif, M., El-Shafie, A., & Ahmed, A. N. (2024). Integrated metaheuristic algorithms with extreme learning machine models for river streamflow prediction. Scientific Reports, 14(1), 1–26. https://doi.org/10.1038/s41598-024-63908-w

Young, P. C. (2002). The Key to Success in Environmental and Biologic Systems Analysis Advances in Real-Time Flood Forecasting. doi: 10.1098/rsta.2002.1008. PMID: 12804258.

Yu, J., Li, Y., Huang, X., & Ye, X. (2025). Data quality and uncertainty issues in flood prediction: A systematic review. *International Journal of Digital Earth*, 8947, 1–30. https://doi.org/10.1080/17538947.2025.2495738