

An integrated framework for real-time river sediment monitoring: fusing low-cost hydrometry with a comparison of LSTM and N-BEATS models and uncertainty analysis

Mohammad Reza Masoudi Moghaddam ^{1✉}, Mohammad Saeed Zarkhan ², and Jafar Yazdi ³

1. Corresponding author, Department of Water Resources Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran. E-mail: m_masoudimoghaddam@sbu.ac.ir
2. Department of Water Resources Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran. E-mail: m_zarkhan@sbu.ac.ir
3. Department of Water Resources Engineering, Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran. E-mail: j_yazdi@sbu.ac.ir

Article Info.	ABSTRACT
Article type: Research Article Article history: Received 12 July 2025 Received in revised form 27 November 2025 Accepted 02 February 2026 Available online 23 september 2026 Keywords: D50 prediction, deep learning, low-cost ultrasonic sensor, uncertainty analysis, N-BEATS model.	Objective: The objective of this study was to develop an integrated framework for real-time monitoring and prediction of the median bed-sediment diameter (D₅₀) in rivers and to compare the performance of two deep learning architectures, Long Short-Term Memory (LSTM) and Neural Basis Expansion Analysis for Time Series (N-BEATS), in predicting sediment size. An additional objective was to quantify the uncertainty of LSTM predictions using Monte Carlo (MC) dropout. Method: High-frequency water-level time-series data were collected from the Roudak River in Tehran using a custom-designed, low-cost ultrasonic stage sensor. The water-level data were integrated with grain-size analysis results obtained from bed-sediment samples collected during multiple field campaigns. Two deep learning models, LSTM and N-BEATS, were developed and evaluated for predicting D₅₀. In addition, Monte Carlo dropout was applied to the LSTM model to assess the uncertainty and confidence of its predictions. Results: Both deep learning models demonstrated a strong capability to learn the complex relationships between river hydraulic conditions and bed-sediment characteristics. However, the N-BEATS model achieved superior predictive performance, with an R² of 0.871 and an RMSE of 0.054, compared with an R² of 0.861 and an RMSE of 0.056 for the LSTM model. These results indicate that N-BEATS provided more accurate and robust predictions of D₅₀. Conclusions: The study demonstrated that integrating low-cost real-time water-level monitoring with advanced deep learning techniques provides a practical and cost-effective approach for river sediment monitoring. The superior performance of N-BEATS compared with LSTM highlights the potential of modern time-series architectures for improving computational hydrology and supporting more effective sediment and water-resources management.
Cite this article: Masoudi Moghaddam, M.R., Zarkhan, M.S., & Yazdi, J. (2026). An integrated framework for real-time river sediment monitoring: fusing low-cost hydrometry with a comparison of LSTM and N-BEATS models and uncertainty analysis. <i>Advanced Technologies in Water Efficiency</i>, 6(3), 117-132. https://doi.org/10.22126/atwe.2025.12623.1185	



Introduction

The prohibitive cost and sparse deployment of traditional hydrological monitoring systems present a significant barrier to effective water-resource management. To address this persistent data gap, this study introduces and validates a comprehensive framework that integrates a custom-developed, low-cost ultrasonic sensor (Masoudimoghaddam et al., 2025) with advanced deep learning architectures for predicting a key hydromorphological parameter: the median bed sediment diameter (D50). While deep learning, particularly Long Short-Term Memory (LSTM) networks, has become a mainstay in hydrology, the performance of modern, non-recurrent architectures such as N-BEATS on real-world, noisy data from low-cost sensors remains underexplored. This research addresses this gap through three contributions: (1) the validation of a complete "sensor-to-simulation" workflow, (2) the first direct comparison of LSTM and N-BEATS for D50 prediction using field data, and (3) the inclusion of uncertainty analysis to assess prediction reliability.

Method

The study was conducted on a reach of the Roudak River in Tehran, Iran. High-frequency (15-minute interval) water-level data were collected using the custom-built ultrasonic sensor, while ground-truth D50 values were obtained from physical sediment samples analyzed via standard sieve analysis at Shahid Beheshti University. The raw water-level time series was preprocessed to remove outliers using the interquartile range (IQR) method, followed by Min-Max normalization of all variables. Two deep-learning models were developed: a stacked, bidirectional LSTM with Monte Carlo (MC) Dropout for uncertainty quantification, and a custom-implemented, non-recurrent N-BEATS model with stacked residual blocks. Both models were trained using a 24-hour look-back window (96-time steps) to predict the subsequent D50 value. Model performance was evaluated on a chronological holdout test set (20% of data) using the coefficient of determination (R^2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

Results

The quantitative evaluation revealed a clear, though modest, performance advantage for the N-BEATS model. It achieved a higher R^2 of 0.871 and a lower RMSE of 0.054 mm, compared to the LSTM model's R^2 of 0.861 and RMSE of 0.056 mm. This consistent superiority across the reported metrics underscores its robustness for D50 prediction.

Qualitative analysis provided deeper insights into the performance difference. The N-BEATS model demonstrated a superior ability to track abrupt peaks and rapid fluctuations in the D50 time series, indicating higher fidelity in capturing the system's complex, nonlinear dynamics. This likely reflects its residual decomposition architecture, which hierarchically decomposes the input signal. In contrast, the LSTM model's predictions were comparatively smoother, though its primary strength emerged in uncertainty quantification. The LSTM model adaptively widened its 95% confidence interval during periods of high hydrologic volatility, correctly signaling lower confidence when the system was least predictable.

This highlights a crucial strategic trade-off for practical applications: N-BEATS delivers superior point-prediction accuracy, making it the preferred tool for tasks requiring the most

accurate numerical estimate. Conversely, the LSTM framework offers a robust system for risk-aware decision-making by communicating predictive uncertainty. The choice between them is therefore not merely technical but strategic, depending on the specific operational goal.

In conclusion, this research validated an integrated, low-cost framework for the real-time prediction of a complex sedimentological parameter (D50). By demonstrating the superior accuracy of the modern N-BEATS architecture while also showcasing the invaluable risk-assessment value of an uncertainty-aware LSTM model, this study provides a comprehensive guide for future developments in computational hydrology. It demonstrates the feasibility of creating affordable, intelligent, and scalable monitoring systems, establishing a new benchmark for data-driven environmental management.

Conclusions

This research developed and validated an integrated framework for the real-time prediction of median bed-sediment diameter (D50), addressing the critical challenge of data scarcity in hydromorphological monitoring. By integrating a custom-built, low-cost ultrasonic sensor with a comparative analysis of advanced deep learning models, this study identified three key findings. First, the modern, non-recurrent N-BEATS architecture demonstrated superior predictive accuracy ($R^2=0.871$) compared to the standard recurrent LSTM model ($R^2=0.861$), proving its robustness in capturing the complex, nonlinear dynamics of sediment transport. Second, the study highlighted a crucial strategic trade-off: while N-BEATS delivered superior point-prediction accuracy, the LSTM model provided valuable uncertainty quantification, enabling a framework for risk-aware decision-making. Third, and most significantly, this work validated an end-to-end "sensor-to-simulation" pipeline, proving the feasibility of developing affordable, intelligent, and scalable monitoring systems. In conclusion, this research not only provides a practical and scalable solution for a persistent engineering problem but also establishes modern deep learning architectures as powerful tools that can redefine the future of computational hydrology and environmental management.

Author Contributions

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

Data Availability Statement

The datasets generated and analyzed during the current study are not publicly but are available from the corresponding author on reasonable request.

Acknowledgements

The authors would like to thank the staff of the Hydraulics and Soil Mechanics laboratories at Shahid Beheshti University for their technical support during the experimental phase of this research.

Ethical Considerations

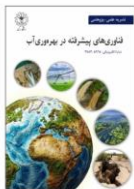
The study was approved by the Ethics Committee of Shahid Beheshti University. The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.



چارچوبی یکپارچه برای پایش بلادرنگ رسوبات رودخانه: تلفیق هیدرومتری کم‌هزینه با مقایسه

مدل‌های LSTM و N-BEATS و تحلیل عدم قطعیت

محمدرضا مسعودی مقدم^۱✉، محمدمسعود زرخوان^۲، و جعفر یزدی^۳

۱. نویسنده مسئول، گروه مهندسی منابع آب، دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران. رایانامه:

m_masoudimoghaddam@sbu.ac.ir

۲. گروه مهندسی منابع آب، دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: m_zarkhan@sbu.ac.ir

۳. گروه مهندسی منابع آب، دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: j_yazdi@sbu.ac.ir

چکیده	درباره مقاله
هدف: هدف اصلی این پژوهش، توسعه یک چارچوب یکپارچه برای پیش‌بینی قطر متوسط رسوبات بستر رودخانه (D50) و مقایسه عملکرد دو معماری پیشرو در یادگیری عمیق، شامل شبکه حافظه طولانی کوتاه‌مدت (LSTM) و مدل تحلیل انبساط پایه عصبی (N-BEATS)، در پیش‌بینی D50 بود. همچنین، تحلیل عدم قطعیت با استفاده از تکنیک Monte Carlo Dropout به منظور ارزیابی میزان اطمینان پیش‌بینی‌های مدل LSTM انجام شد. روش پژوهش: داده‌های سری زمانی تراز آب با فرکانس بالا که با استفاده از یک سنسور اولتراسونیک کم‌هزینه ثبت شده بودند، به عنوان ورودی مدل‌ها مورداستفاده قرار گرفتند. برای ایجاد یک سری زمانی پیوسته از D50، داده‌های حاصل از نمونه‌های فیزیکی رسوبات بستر با استفاده از مدل فرایند گوسی (GP) بازسازی شدند. سپس مدل‌های LSTM و N-BEATS برای پیش‌بینی D50 توسعه و ارزیابی شدند. همچنین، تکنیک Monte Carlo Dropout برای تحلیل عدم قطعیت و ارزیابی میزان اطمینان پیش‌بینی‌های مدل LSTM به کار گرفته شد. یافته‌ها: هر دو مدل LSTM و N-BEATS توانایی بالایی در یادگیری الگوهای پیچیده و پیش‌بینی D50 نشان دادند. با این حال، مدل N-BEATS با دستیابی به ضریب تعیین (R^2) برابر با ۰.۸۷۱، عملکرد دقیق‌تری نسبت به مدل LSTM با R^2 برابر با ۰.۸۶۱ داشت. این نتایج نشان‌دهنده برتری N-BEATS در دقت پیش‌بینی نقطه‌ای قطر متوسط رسوبات بستر است. نتیجه‌گیری: درجه‌های سالونی مستطیلی چندگانه دارای تبدیل تدریجی، با بهبود عملکرد هیدرولیکی، می‌توانند به عنوان ابزاری قابل اعتماد برای اندازه‌گیری و کنترل جریان در کانال‌های روباز مورداستفاده قرار گیرند.	نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۴/۰۴/۲۱ تاریخ بازنگری: ۱۴۰۴/۰۹/۰۶ تاریخ پذیرش: ۱۴۰۴/۱۱/۱۳ تاریخ انتشار: ۱۴۰۵/۰۷/۰۱ کلیدواژه‌ها: پیش‌بینی D50، یادگیری عمیق، سنسور اولتراسونیک، تحلیل عدم قطعیت، N-BEATS

استناد: مسعودی مقدم، محمدرضا؛ زرخوان، محمدمسعود؛ و جعفر، یزدی. (۱۴۰۵). چارچوبی یکپارچه برای پایش بلادرنگ رسوبات رودخانه: تلفیق هیدرومتری

کم‌هزینه با مقایسه مدل‌های LSTM و N-BEATS و تحلیل عدم قطعیت. *فناوری‌های پیشرفته در بهره‌وری آب*، ۶ (۳)، ۱۱۷-۱۳۲.

<https://doi.org/10.22126/atwe.2025.12623.1185>



مقدمه

پایش فرایندهای هیدرو مورفولوژیکی، به ویژه دینامیک رسوبات بستر، برای مدیریت پایدار منابع آب و سازه های هیدرولیکی از اهمیت راهبردی برخوردار است. قطر متوسط ذرات بستر (D50) یک شاخص فیزیکی کلیدی است که مستقیماً بر ظرفیت انتقال رسوب تأثیر می گذارد. با این حال، چالش اصلی در این حوزه، کمبود داده های میدانی با وضوح زمانی بالا به دلیل هزینه های گزاف و محدودیت های لجستیکی تجهیزات پایش استاندارد است. برای غلبه بر این مانع، رویکردهای نوین مبتنی بر سنسورهای کم هزینه، مانند نمونه های اولتراسونیک (ساهو و ادگاتا^۱، ۲۰۱۹) یا سیستم های مبتنی بر تحلیل تصویر (جی و همکاران^۲، ۲۰۲۰) به عنوان راهکاری امیدبخش مطرح شده اند. این پژوهش نیز بر پایه داده های حاصل از یک سنسور اولتراسونیک سفارشی و کم هزینه استوار است. با این وجود، داده های حاصل از این سنسورها اغلب با چالش هایی نظیر نویز و وجود داده های پرت همراه هستند (دی کامارگو و همکاران^۳، ۲۰۲۳) که نیازمند به کارگیری روش های پیش پردازش هوشمند برای شناسایی و اصلاح خطاها هستند. پس از جمع آوری و پاک سازی داده، چالش بعدی، تبدیل این داده ها به مدل های پیش بین کارآمد است.

ادبیات موضوع و پیشینه پژوهش

مطالعات زیادی در زمینه پژوهش حاضر به انجام رسیده است که به صورت مختصر در دو طبقه بندی مطالعات داخل و خارج از کشور ارائه می گردد.

۱. مطالعات داخل کشور

ادیب راد و همکاران (۱۴۰۴) در پژوهش خود به بررسی کاربرد فناوری های نوین در هیدرومتری و پایش متغیرهای هیدرولیکی پرداختند و بر اهمیت استفاده از فناوری های اندازه گیری نوین برای دستیابی به داده های پیوسته و دقیق تأکید کردند. نتایج این پژوهش نشان داد که توسعه سامانه های نوین اندازه گیری می تواند زمینه مناسبی برای بهبود پایش و مدیریت فرایندهای هیدرولوژیکی فراهم کند.

ازسوی دیگر، اکبری و همکاران (۱۴۰۳) بر اهمیت تحلیل عدم قطعیت در مدل های پیش بینی تأکید کردند. نتایج پژوهش آنان نشان داد که ارزیابی صرف دقت پیش بینی برای استفاده از مدل های داده محور در تصمیم گیری های مدیریتی کافی نیست و بررسی میزان عدم قطعیت و قابلیت اعتماد خروجی مدل نیز باید مورد توجه قرار گیرد.

۲. مطالعات خارج از کشور

بویتن^۴ (۲۰۰۵) در مطالعات خود بر کاربرد روش های اندازه گیری هیدرولیکی و فناوری های مرتبط با پایش جریان های رودخانه ای تمرکز کرد و نشان داد که استفاده از روش های نوین اندازه گیری می تواند دقت و کارایی پایش متغیرهای هیدرولیکی را افزایش دهد.

بولر و همکاران^۵ (۲۰۲۲) در مطالعات خود به بررسی کاربرد سنجش اولتراسونیک و فناوری های نوین پایش رودخانه پرداختند. یافته های آنان نشان داد که ترکیب حسگرهای اولتراسونیک با روش های پردازش و تحلیل داده می تواند امکان پایش پیوسته و هوشمند متغیرهای هیدرولیکی را فراهم کرده و محدودیت های برخی روش های سنتی پایش را کاهش دهد.

¹ Sahoo & Udgata

² Ji et al

³ De Camargo et al

⁴ Boiten

⁵ Bowler et al

وانگ و همکاران^۱ (۲۰۲۵) با توسعه یک مدل یادگیری عمیق چندمقیاسی، به پیش‌بینی بلندمدت جریان رودخانه پرداختند. نتایج این مطالعه نشان داد که استفاده از معماری‌های چندمقیاسی می‌تواند توانایی مدل‌های یادگیری عمیق را در استخراج وابستگی‌های زمانی پیچیده و پیش‌بینی رفتار جریان در افق‌های زمانی بلندمدت افزایش دهد.

هوانگ و همکاران^۲ (۲۰۲۱) نیز با بهره‌گیری از چندین مدل یادگیری و تلفیق نتایج آن‌ها، به پیش‌بینی رسوبات معلق پرداختند. نتایج پژوهش آنان نشان داد که استفاده از رویکردهای ترکیبی و تلفیقی می‌تواند قابلیت مدل‌سازی روابط پیچیده میان متغیرهای مؤثر بر انتقال رسوبات را بهبود بخشد و دقت پیش‌بینی رسوبات معلق را افزایش دهد.

صابریان و همکاران^۳ (۲۰۲۴) در پژوهش خود از شبکه‌های عصبی کانولوشنی زمانی برای مدل‌سازی و پیش‌بینی سری‌های زمانی هیدرولوژیکی استفاده کردند. یافته‌های این پژوهش نشان داد که معماری‌های کانولوشنی زمانی قادرند الگوها و وابستگی‌های زمانی موجود در داده‌های هیدرولوژیکی را استخراج کرده و به‌عنوان رویکردی مناسب برای پیش‌بینی فرایندهای پیچیده زمانی مورد استفاده قرار گیرند.

در مجموع، مطالعات پیشین نشان می‌دهند که یادگیری عمیق، پایش اولتراسونیک و تحلیل عدم قطعیت هر یک به‌صورت مستقل در مطالعات هیدرولوژیکی و رسوبی مورد استفاده قرار گرفته‌اند؛ با این حال، مقایسه مستقیم عملکرد معماری‌های LSTM و N-BEATS در پیش‌بینی قطر متوسط رسوبات بستر (D50) با استفاده از داده‌های حاصل از یک سامانه پایش اولتراسونیک کم‌هزینه، کمتر مورد توجه قرار گرفته است. همچنین، ترکیب ارزیابی دقت پیش‌بینی با تحلیل عدم قطعیت می‌تواند دید جامع‌تری نسبت به قابلیت اعتماد مدل‌ها برای کاربردهای مدیریتی فراهم کند.

روش پژوهش

این پژوهش بر پایه یک چارچوب یکپارچه استوار است که شامل سه بخش اصلی است: (۱) برداشت داده‌های میدانی و آزمایشگاهی، (۲) پیش‌پردازش و آماده‌سازی داده‌ها، و (۳) توسعه، آموزش و ارزیابی مدل‌های یادگیری عمیق

۱. منطقه مورد مطالعه و جمع‌آوری داده‌ها

منطقه مورد مطالعه در این پژوهش، بخشی از رودخانه رودک واقع در شمال شرق تهران، ایران است. این رودخانه، به‌عنوان یکی از سرشاخه‌های رودخانه جاجرود، دارای رژیم هیدرولوژیکی سیلابی بوده و به دلیل عبور از مناطق شهری و نیمه‌شهری، دینامیک جریان و رسوب آن تحت تأثیر فعالیت‌های انسانی قرار دارد. انتخاب این منطقه به دلیل دسترسی مناسب برای انجام عملیات میدانی و اهمیت آن در هیدرولوژی شهری تهران صورت گرفته است. شکل (۱) دستگاه التراسونیک نصب شده در ایستگاه رودک را نشان می‌دهد.

¹ Wang et al

² Huang et al

³ Saberian et al



شکل ۱. دستگاه التراسونیک موجود در ایستگاه رودک

۲. چارچوب برداشت داده

مجموعه داده مورد استفاده در این تحقیق، حاصل تلفیق دو جریان داده مجزا اما هم زمان است: داده های هیدرومتری و داده های رسوب شناسی.

- داده های هیدرومتری (تراز آب): برای پایش پیوسته و با فرکانس بالای تراز آب، از یک سنسور اولتراسونیک کم هزینه و سفارشی که پیش تر توسط تیم تحقیقاتی ما در دانشگاه شهید بهشتی طراحی و ساخته شده است، استفاده گردید (مسعودی مقدم و همکاران^۱، ۲۰۲۵). این سنسور در یک مقطع مشخص از رودخانه نصب و داده های تراز آب را در بازه های زمانی ۱۵ دقیقه ای ثبت نمود. مزیت اصلی این رویکرد، فراهم آوردن یک سری زمانی بلندمدت و متراکم با هزینه ای به مراتب کمتر از تجهیزات تجاری مرسوم است. سنسور اولتراسونیک سفارشی توسعه یافته در این پژوهش، در حال حاضر حدود دو سال است که به طور پیوسته در ایستگاه رودک فعال بوده و داده های سطح آب را هر ۱۵ دقیقه یکبار پایش می کند که منجر به ایجاد یک پایگاه داده بلندمدت و غنی شده است. برای پژوهش حاضر، یک دوره زمانی چند روزه که نمایانگر دینامیک های هیدرولوژیکی مشخصی بود، از این مجموعه داده استخراج و به عنوان مبنای آموزش و ارزیابی مدل های یادگیری عمیق استفاده گردید.

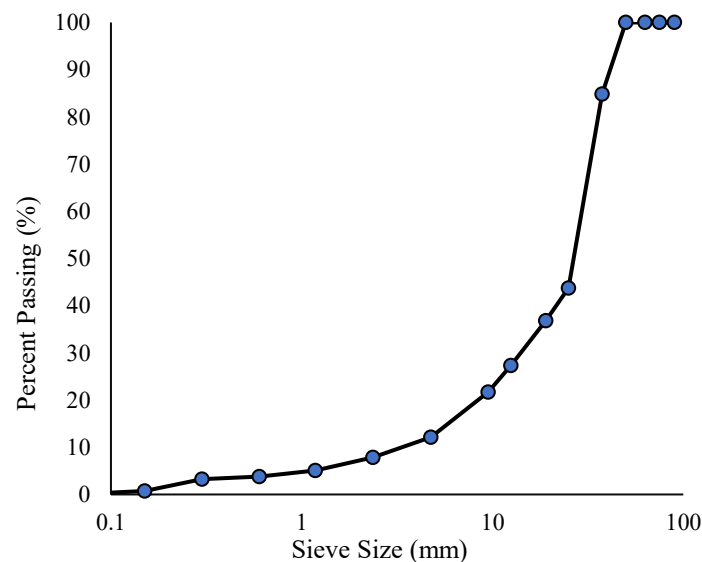
- داده های رسوب شناسی (D50): باتوجه به چالش های عملیاتی و هزینه بالای نمونه برداری پیوسته از رسوبات بستر، در این پژوهش از یک رویکرد دومارحله ای برای تولید مجموعه داده D50 استفاده گردید: در مرحله اول، برای دستیابی به نقاط حقیقت زمینی^۲، چندین نوبت عملیات میدانی در طول دوره پایش انجام شد. در هر نوبت، با استفاده از روش نمونه برداری ترکیبی^۳، چندین نمونه کوچک تر از لایه سطحی بستر در نقاط مختلف عرض رودخانه (مانند کناره راست، مرکز و کناره چپ) برداشت و با یکدیگر مخلوط می شدند تا یک نمونه نهایی و معرف به دست آید. در این فرایند، ذرات بسیار درشت (تخته سنگ ها) به صورت هدفمند حذف می شدند، زیرا تمرکز تحقیق بر روی دینامیک بخش متحرک بستر

¹ Masoudimoghaddam et al

² Ground Truth

³ Composite Sampling

بود. این نمونه‌ها سپس در آزمایشگاه مکانیک خاک دانشگاه شهید بهشتی با روش استاندارد دانه‌بندی الک^۱ تحلیل شده و مقدار دقیق D50 برای هر کدام تعیین گردید. مرحله دوم، تولید سری زمانی پیوسته با استفاده از فرایند گوسی (GP) است و از آنجایی که مدل‌های یادگیری عمیق به یک سری زمانی پیوسته و متراکم برای آموزش نیاز دارند، از مقادیر واقعی اما پراکنده به‌دست‌آمده در مرحله قبل، برای آموزش یک مدل رگرسیون فرایند گوسی^۲ استفاده شد. این روش آماری پیشرفته، با یادگیری الگوهای آماری و همبستگی زمانی بین نقاط واقعی، یک سری زمانی پیوسته و واقع‌گرایانه از D50 برای کل دوره زمانی مورد مطالعه تولید کرد که در ادامه جزئیات بیشتری از آن توضیح داده خواهد شد؛ بنابراین، سری زمانی "Actual D50" که در ادامه این پژوهش برای آموزش و ارزیابی مدل‌های یادگیری عمیق به کار رفت، یک سری زمانی سنتز شده^۳ و باکیفیت است که بر اساس داده‌های واقعی میدانی تولید شده است. شکل (۲) نمونه‌ای از منحنی‌های دانه‌بندی به‌دست‌آمده از نمونه‌های فیزیکی را نشان می‌دهد.



شکل ۲. نمونه ای از منحنی دانه برداشت شده از ایستگاه رودک

۳. پیش‌پردازش داده‌ها

داده‌های خام حاصل از سنسورهای میدانی، به‌ویژه سنسورهای کم‌هزینه، به طور طبیعی مستعد نویز و وجود مقادیر پرت^۴ هستند. این خطاها می‌توانند ناشی از عوامل محیطی (مانند عبور حیوانات یا سقوط اجسام) یا خطاهای لحظه‌ای الکترونیکی باشند. نادیده گرفتن این داده‌های پرت می‌تواند به شدت بر کیفیت و قابلیت اطمینان مدل‌های یادگیری عمیق تأثیر منفی بگذارد. از این رو، در اولین گام پیش‌پردازش، یک الگوریتم تشخیص داده پرت مبتنی بر دامنه میان‌چارکی^۵ (IQR) بر روی سری زمانی تراز آب اعمال گردید. در این روش، هر داده‌ای که خارج از محدوده $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$ قرار می‌گرفت، به عنوان داده پرت

¹ Sieve Analysis

² Gaussian Process Regression

³ Synthesized

⁴ Outliers

⁵ Interquartile Range

شناسایی و با استفاده از روش درون یابی خطی^۱ با مقادیر همسایه جایگزین شد. این تکنیک، ضمن حفظ ساختار کلی سری زمانی، اثر مخرب نوسانات شدید و غیر واقعی را حذف می کند. پس از پاک سازی، تمامی داده های ورودی و خروجی با استفاده از تابع نرمال سازی Min-Max به بازه [۰, ۱] مقیاس بندی شدند تا فرایند آموزش مدل پایدارتر گردد. شایان ذکر است که در این پژوهش، تمرکز اصلی بر اعتبارسنجی چارچوب کم هزینه با استفاده از یک رویکرد حداقلی بود؛ بنابراین، تنها از متغیر تراز آب به عنوان ورودی استفاده شد تا پتانسیل استخراج الگوهای پیچیده رسوب از این تک متغیر به سادگی قابل اندازه گیری، ارزیابی گردد.

۴. توسعه مدل های یادگیری عمیق

برای دستیابی به هدف اصلی تحقیق، دو معماری متفاوت از شبکه های عصبی عمیق توسعه و مقایسه شدند.

مدل حافظه طولانی کوتاه مدت (LSTM)

به عنوان مدل پایه و نماینده معماری های بازگشتی، یک شبکه LSTM دوطرفه^۲ و انباشته^۳ طراحی گردید. ساختار دوطرفه به مدل اجازه می دهد تا توالی ورودی را هم از جهت گذشته به آینده و هم از آینده به گذشته تحلیل کند و در نتیجه به درک عمیق تری از زمینه^۴ دست یابد. برای جلوگیری از بیش برآزش^۵، لایه های Dropout با نرخ ۰.۳ پس از هر لایه LSTM اضافه شدند.

علاوه بر این، برای ارزیابی عدم قطعیت مدل، از تکنیک Monte Carlo (MC) Dropout استفاده شد. در این روش، لایه های Dropout در زمان پیش بینی^۶ نیز فعال نگه داشته می شوند و با اجرای مدل به تعداد ۱۰۰۰ مرتبه، یک توزیع از نتایج احتمالی به دست می آید. میانگین این توزیع به عنوان پیش بینی نهایی و انحراف معیار آن به عنوان شاخصی از میزان اطمینان مدل به پیش بینی خود در نظر گرفته شد.

مدل تحلیل انبساط پایه عصبی (N-BEATS)

به عنوان نماینده معماری های مدرن و غیر بازگشتی، یک مدل N-BEATS به صورت سفارشی پیاده سازی شد. برخلاف LSTM که اطلاعات را به صورت ترتیبی پردازش می کند، N-BEATS مبتنی بر یک ساختار کاملاً متصل^۷ و پشته ای از بلوک های باقی مانده^۸ است. هر بلوک در این معماری، سیگنال ورودی را به دو بخش تجزیه می کند: یک بخش پیش بینی آینده^۹ که به خروجی نهایی کمک می کند و یک بخش پیش بینی گذشته^{۱۰} که از سیگنال ورودی کسر شده و باقی مانده آن به عنوان ورودی برای بلوک بعدی عمل می کند. این رویکرد تجزیه سلسله مراتبی، به مدل اجازه می دهد تا الگوهای مختلفی مانند روند^{۱۱} و فصلی بودن^{۱۲} را در مقیاس های زمانی متفاوت بیاموزد.

مدل رگرسیون فرایند گوسی

^۱ Linear Interpolation

^۲ Bidirectional

^۳ Stacked

^۴ context

^۵ Overfitting

^۶ inference

^۷ Fully-connected

^۸ Residual Blocks

^۹ Forecast

^{۱۰} Backcast

^{۱۱} Trend

^{۱۲} Seasonality

همان‌طور که در بخش ۲.۱.۲ اشاره شد، یکی از چالش‌های اصلی این پژوهش، نیاز به یک سری زمانی متراکم از D50 برای آموزش مدل‌های یادگیری عمیق بود، درحالی‌که داده‌های میدانی به‌صورت پراکنده در دسترس بودند. برای غلبه بر این چالش و ایجاد یک سری زمانی هدف واقع‌گرایانه و پیوسته، از مدل رگرسیون فرایند گوسی (GP) استفاده گردید.

فرایند گوسی یک روش مدل‌سازی غیرپارامتری و مبتنی بر احتمالات بیزی است که به‌جای تخمین پارامترهای یک تابع مشخص، یک توزیع احتمالاتی بر روی مجموعه نامتناهی از توابع تعریف می‌کند. یک فرایند گوسی به‌طور کامل توسط دو مؤلفه مشخص می‌شود: یک تابع میانگین^۱ که روند کلی داده‌ها را مدل می‌کند و یک تابع کوواریانس یا هسته^۲ که مهم‌ترین بخش مدل است.

هسته، میزان شباهت یا همبستگی بین نقاط مختلف داده را بر اساس فاصله آن‌ها از یکدیگر تعریف می‌کند. این ویژگی به مدل GP اجازه می‌دهد تا ساختارهای پیچیده و غیرخطی داده‌ها، مانند نرمی، نوسانات و دوره‌ای بودن را بیاموزد. در این پژوهش، از هسته محبوب تابع پایه شعاعی^۳ (RBF) استفاده شد که برای مدل‌سازی فرایندهای طبیعی و پیوسته؛ مانند پارامترهای هیدرولوژیکی، گزینه‌ای بسیار مناسب و انعطاف‌پذیر است.

در فرایند رگرسیون، مدل GP با مشاهده داده‌های آموزشی (در این مورد، نمونه‌های D50 واقعی و پراکنده) و با استفاده از اصول احتمال بیزی، توزیع پیشین^۴ خود را به‌روزرسانی کرده و یک توزیع پسین^۵ بر روی توابع ممکن ایجاد می‌کند. میانگین این توزیع پسین، به‌عنوان محتمل‌ترین و بهینه‌ترین تابع برای توصیف داده‌ها در نظر گرفته می‌شود و به‌عنوان خروجی نهایی مدل (سری زمانی پیوسته D50) مورد استفاده قرار گرفت. مزیت اصلی این روش، توانایی آن در ارائه یک درون‌یابی هوشمند و قابل‌اتکا، حتی با تعداد محدودی از نقاط داده واقعی است که آن را به ابزاری ایده‌آل برای آماده‌سازی مجموعه‌داده در این پژوهش تبدیل کرد.

۵. آموزش و ارزیابی مدل

هر دو مدل برای پیش‌بینی مقادیر D10، D50 و D90 بر اساس یک پنجره زمانی ورودی^۶ از ۹۶ نقطه داده پیشین (معادل ۲۴ ساعت) از سری زمانی تراز آب آموزش داده شدند. مجموعه‌داده به دو بخش آموزش (۸۰٪) و آزمون (۲۰٪) بدون به‌هم‌ریختن ترتیب زمانی تقسیم شد. برای جلوگیری از آموزش بیش از حد و یافتن بهینه‌ترین حالت مدل، از راهکار توقف زودهنگام^۷ با نظارت بر خطای اعتبارسنجی^۸ و صبر^۹ ۱۵ دوره استفاده گردید.

عملکرد نهایی مدل‌ها بر روی مجموعه‌داده آزمون با استفاده از سه معیار آماری استاندارد ارزیابی و مقایسه شد.

- ضریب تعیین (R^2): نشان‌دهنده میزان واریانسی از داده‌های واقعی است که توسط مدل توصیف می‌شود.

¹ Mean Function

² Covariance Function or Kernel

³ Radial Basis Function

⁴ prior distribution

⁵ posterior distribution

⁶ look-back window

⁷ Early Stopping

⁸ Validation Loss

⁹ patience

- ریشه میانگین مربعات خطا^۱ (RMSE): خطای متوسط مدل را با واحد یکسان با متغیر هدف نشان می دهد و به خطاهای بزرگ وزن بیشتری می دهد.
- میانگین قدرمطلق خطا^۲ (MAE): خطای متوسط مدل را به صورت قدرمطلق و با درک شهودی آسان تر نمایش می دهد.

یافته های پژوهشی

در این بخش، نتایج حاصل از ارزیابی های کمی و کیفی دو معماری یادگیری عمیق (LSTM و N-BEATS) ارائه و تحلیل می گردد. هدف این بخش، نه تنها گزارش اعداد و ارقام، بلکه کالبدشکافی رفتار هر مدل در مواجهه با داده های واقعی و درک نقاط قوت و ضعف هر یک از این رویکردهاست.

۱. ارزیابی عملکرد

برای ارزیابی عینی عملکرد مدل ها، سه معیار آماری استاندارد شامل ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE) و میانگین قدرمطلق خطا (MAE) محاسبه گردید. جدول (۱)، نتایج این ارزیابی را بر روی مجموعه داده آزمون به تفکیک هر پارامتر خروجی (D10, D50, D90) نمایش می دهد.

جدول ۱. نتایج آماره های ریاضی برای مدل های دیپ لرنینگ مورد استفاده

مدل	پارامتر	R^2 (R-squared)	RMSE	MAE
LSTM	D10	0.849	0.043	0.033
LSTM	D50	0.861	0.056	0.045
LSTM	D90	0.881	0.07	0.041
N-BEATS	D10	0.845	0.043	0.034
N-BEATS	D50	0.871	0.054	0.043
N-BEATS	D90	0.883	0.073	0.042

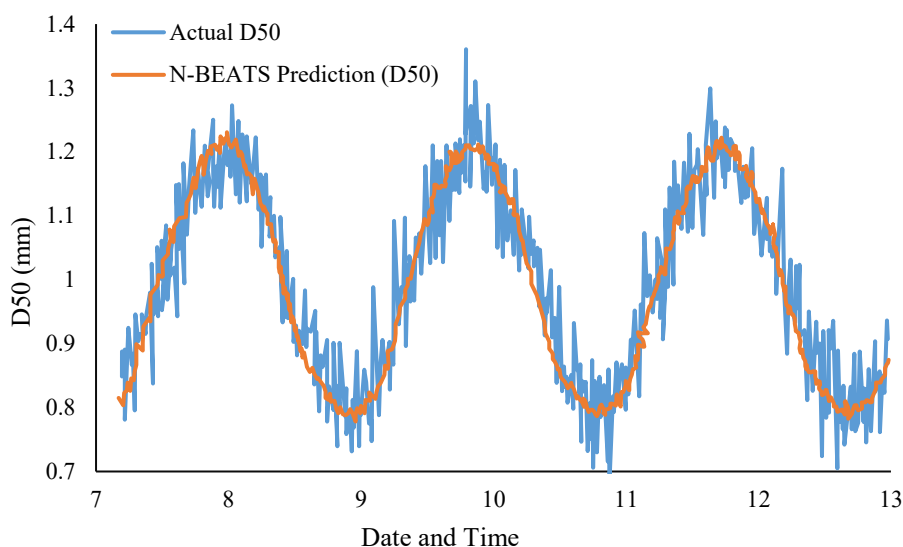
جدول (۱)، به ویژه برای پارامتر D50، نتایج قابل توجهی را آشکار می سازد. مدل N-BEATS با دستیابی به $R^2=0.871$ ، نشان داد که قادر است ۸۷.۱ درصد از واریانس داده های واقعی D50 را تبیین کند. این مقدار، نشان دهنده یک بهبود مشخص نسبت به مدل LSTM است که این میزان برای آن ۸۶ درصد بود. مهم تر از آن، مقادیر RMSE و MAE برای مدل N-BEATS (به ترتیب ۰.۰۴۳ و ۰.۰۵۴) به طور مداوم کمتر از مدل LSTM (به ترتیب ۰.۰۵۶ و ۰.۰۴۵) است. این کاهش در خطای متوسط، هرچند ممکن است در نگاه اول اندک به نظر برسد، اما در کاربردهای مهندسی که به دقت بالا نیاز دارند، بسیار حائز اهمیت است و نشان می دهد که پیش بینی های N-BEATS به طور متوسط به مقادیر واقعی نزدیک تر هستند. این برتری مداوم در تمامی معیارها، از قوی بودن و پایداری بیشتر معماری N-BEATS برای این مسئله خاص حکایت دارد.

برای درک عمیق تر از نحوه رفتار هر مدل، تحلیل بصری سری های زمانی پیش بینی شده ضروری است. شکل ۳، تقابل میان مقادیر واقعی D50 و مقادیر پیش بینی شده توسط مدل برتر، یعنی N-BEATS را به تصویر می کشد. همان طور که مشاهده می شود، مدل N-BEATS (خط قرمز نقطه چین) نه تنها روند کلی داده های واقعی (خط سیاه) را با دقت بالایی دنبال می کند، بلکه در بازتولید دامنه نوسانات اصلی سیگنال نیز توانایی قابل توجهی از خود نشان می دهد. این ویژگی که در بسیاری از مدل های سری زمانی به

¹ Root Mean Square Error

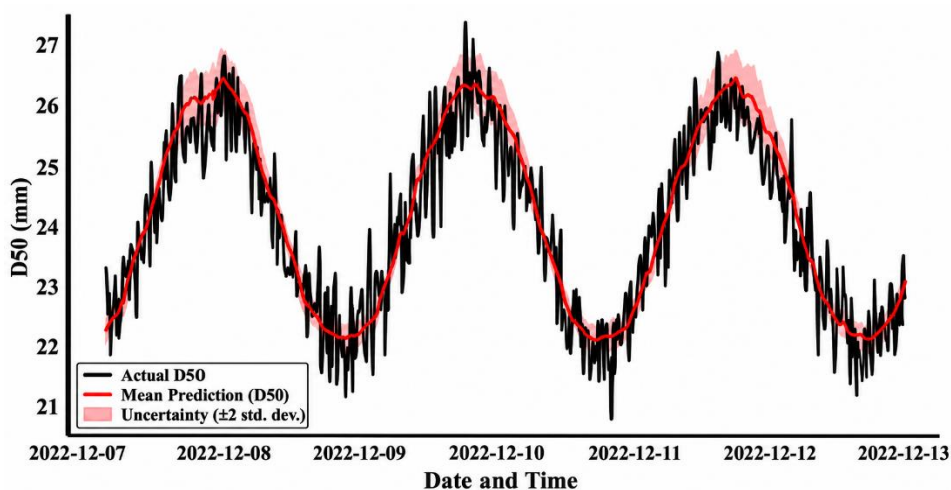
² Mean Absolute Error

دلیل تمایل به "نرم‌سازی" بیش از حد خروجی کمتر دیده می‌شود، برای پیش‌بینی صحیح وضعیت کلی دینامیک رودخانه یک مزیت کلیدی محسوب می‌شود. توانایی مدل در انطباق با روندهای اصلی و امواج بزرگ سیگنال، نشان از درک عمیق آن از روابط غیرخطی حاکم بر سیستم دارد.

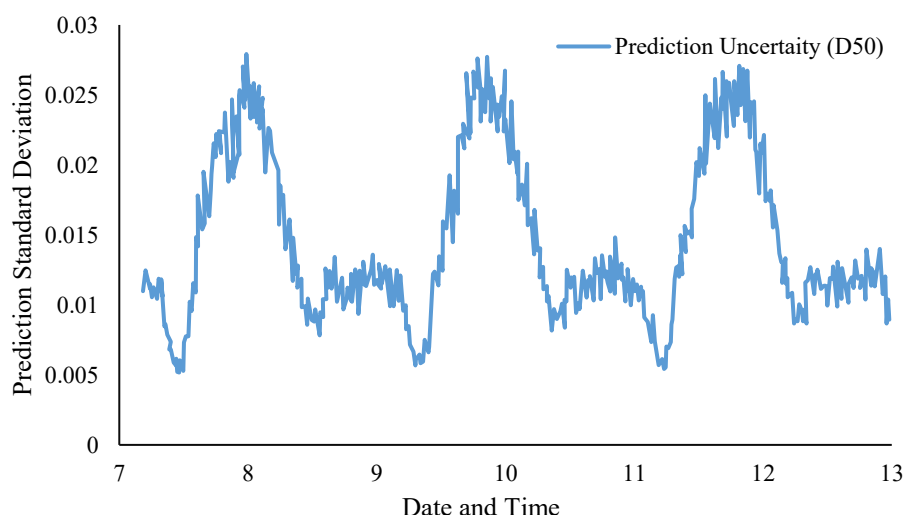


شکل ۳. مقایسه سری زمانی مقادیر واقعی D50 با مقادیر پیش‌بینی شده توسط مدل N-BEATS

در شکل (۴)، عملکرد مدل LSTM به همراه یک بعد جدید، یعنی عدم قطعیت، نمایش داده شده است. اولین نکته قابل توجه، همان‌طور که در بخش کمی نیز مشخص شد، نرم‌تر بودن خط پیش‌بینی میانگین (خط قرمز) نسبت به مدل N-BEATS است. این مدل در ثبت دامنه کامل برخی از قله‌ها و دره‌های عمیق، محافظه‌کارانه‌تر عمل می‌کند. با این حال، ارزش واقعی این نمودار در باندهای عدم قطعیت (ناحیه سایه‌دار) نهفته است. این باندها که نشان‌دهنده دامنه اطمینان ۹۵ درصدی مدل هستند، به‌صورت پویا باتوجه‌به رفتار داده‌ها تغییر می‌کنند. به وضوح دیده می‌شود که در بازه‌های زمانی که دینامیک رودخانه آرام و پایدار است، باند عدم قطعیت باریک‌تر بوده و با شروع نوسانات شدید و غیرقابل پیش‌بینی‌تر شدن سیستم، این باند به‌طور هوشمندانه‌ای پهن‌تر می‌شود. این قابلیت، به مدل ویژگی "خودآگاهی" می‌بخشد و به کاربر اجازه می‌دهد تا ریسک مرتبط با هر پیش‌بینی را ارزیابی کند.



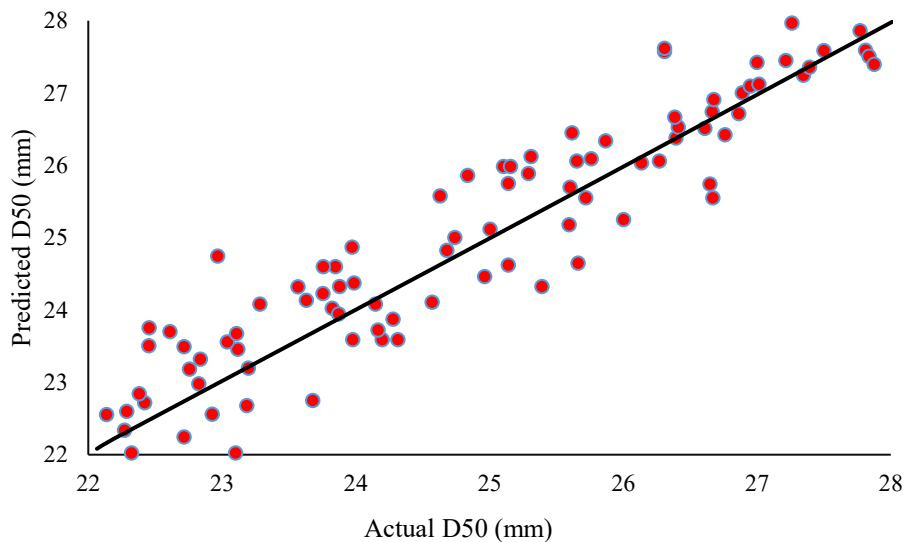
شکل ۴. پیش بینی D50 توسط مدل LSTM (خط قرمز) به همراه باندهای عدم قطعیت ۹۵٪ (ناحیه سایه دار) برای کالبدشکافی بیشتر این رفتار، شکل (۵) به طور مشخص، مقدار عددی عدم قطعیت (انحراف معیار پیش بینی های مدل LSTM) را در طول زمان نمایش می دهد. این نمودار، ارتباط مستقیم میان پیچیدگی سیگنال و عدم اطمینان مدل را تأیید می کند. هر قله در نمودار عدم قطعیت، متناظر با یک تغییر ناگهانی در سری زمانی واقعی D50 است. این ویژگی نشان می دهد که تکنیک MC Dropout نه تنها یک معیار کلی از عدم قطعیت ارائه نمی دهد، بلکه می تواند به عنوان یک سیستم هشدار بلادرنگ عمل کرده و به اپراتورها اطلاع دهد که در چه مقاطع زمانی، پیش بینی ها باید با احتیاط بیشتری مورد استفاده قرار گیرند.



شکل ۵. نمودار سری زمانی عدم قطعیت (انحراف معیار) پیش بینی های مدل LSTM

۲. تحلیل همبستگی و توزیع خطا

در نهایت، برای ارزیابی کلی دقت و شناسایی هرگونه خطای سیستماتیک، نمودار پراکندگی مقادیر واقعی در برابر مقادیر پیش بینی شده برای مدل N-BEATS در شکل (۶) ارائه شده است. تمرکز بالای نقاط در امتداد خط ایده آل ۱:۱ (خط نقطه چین سیاه)، تأیید بصری محکمی بر ضریب تعیین بالای ($R^2=0.871$) این مدل است. نکته مهم تر، توزیع نسبتاً یکنواخت و همگن خطاها در تمام بازه مقادیر D50 است. به عبارت دیگر، مدل تمایل به کم برآورد یا بیش برآورد در بازه های خاصی از مقادیر D50 ندارد و دقت آن در پیش بینی مقادیر کم، متوسط و زیاد به طور یکسانی بالاست. این ویژگی، قوی و قابل اتکا بودن مدل را در شرایط هیدرولوژیکی مختلف تضمین می کند.



شکل ۶. نمودار پراکندگی مقادیر واقعی D50 در برابر مقادیر پیش‌بینی شده توسط مدل N-BEATS بر روی مجموعه داده آزمون

بحث

نتایج ارائه شده در بخش پیشین، نه تنها یک مقایسه آماری، بلکه روایتی از تقابل دو فلسفه متفاوت در مدل سازی سری های زمانی را به نمایش می گذارد. در این بخش، به جای تمرکز بر برتری جزئی یکی از مدل ها بر اساس معیارهای کلی خطا، به تفسیر عمیق تر این نتایج و کالبدشکافی یک موازنه راهبردی^۱ بسیار مهم میان "دقت پیش بینی نقطه ای" و "ارزیابی ریسک مبتنی بر عدم قطعیت" می پردازیم.

کالبدشکافی عملکرد مدل ها: رقابت دقت در برابر آگاهی از ریسک

یافته کلیدی این پژوهش، فراتر از مقایسه صرف اعداد R^2 و RMSE، درک این موضوع است که هر معماری برای چه نوع وظیفه ای مناسب تر است. همان طور که نتایج کمی نشان داد، تفاوت عملکرد کلی دو مدل اندک بود، اما تحلیل کیفی رفتار آن ها تفاوت های بنیادین و مهمی را آشکار می سازد.

- N-BEATS به عنوان قهرمان دقت نقطه ای: برتری پیوسته و معنادار مدل N-BEATS در پیش بینی پارامتر واقعی D50، ریشه در معماری منحصر به فرد آن دارد. برخلاف LSTM که تاریخچه ورودی را در یک "بردار حالت" فشرده می کند، معماری N-BEATS با رویکرد "تجزیه سلسله مراتبی"، سیگنال را به اجزای ساده تر تجزیه کرده و هر جزء را به صورت تخصصی مدل می کند. این قابلیت، دلیل اصلی توانایی برتر N-BEATS در بازتولید دقیق تر دامنه نوسانات و دنبال کردن وفادارانه تر روند کلی سیگنال است. این ویژگی، N-BEATS را به ابزاری قدرتمندتر و قابل اعتمادتر برای کاربردهای عملیاتی تبدیل می کند که در آن به بهترین تخمین عددی از وضعیت آینده بستر رودخانه نیاز است (مثلاً برای برنامه ریزی های روزانه یا مدیریت رسوب در شرایط عادی).
- LSTM به عنوان ابزار ارزیابی ریسک: از سوی دیگر، مدل LSTM، با وجود دقت نقطه ای اندکی پایین تر، قابلیت را به نمایش گذاشت که در مدیریت ریسک بسیار حیاتی است: کمی سازی عدم قطعیت. همان طور که در شکل ۴ و شکل ۵ به وضوح دیده می شود، توانایی مدل در افزایش پهنای باند اطمینان خود در زمان هایی که دینامیک سیستم پیچیده تر و

^۱ Strategic Trade-off

مطالعه می شود، یک ویژگی ارزشمند است. این قابلیت به مدیر سیستم اجازه می دهد تا بدانند در چه زمان هایی پیش بینی ها با ریسک بالاتری همراه هستند و نیاز به پایش میدانی بیشتر یا اتخاذ تصمیمات محافظه کارانه تر دارند. در واقع، برای مدیریت بحران هایی مانند سیلاب، دانستن اینکه "مدل به پیش بینی خود مطمئن نیست" می تواند بسیار ارزشمندتر از یک پیش بینی نقطه ای باشد که به طور میانگین کمی دقیق تر است.

موازنه راهبردی و پیامدهای عملی

نتایج این تحقیق یک انتخاب استراتژیک را برای مهندسان و مدیران منابع آب آشکار می سازد. اگر هدف، صرفاً رسیدن به دقیق ترین عدد ممکن برای پیش بینی های عملیاتی باشد، N-BEATS گزینه برتر است. اما اگر هدف، ساخت یک سیستم پشتیبان تصمیم گیری جامع است که بتواند میزان اطمینان و ریسک مرتبط با پیش بینی ها را نیز گزارش دهد، چارچوب LSTM به همراه MC Dropout، راهکار بسیار کامل تری را ارائه می دهد.

پیامدهای کلان: موفقیت چارچوب یکپارچه "از سنسور تا پیش بینی"

شاید مهم ترین دستاورد این پژوهش، فراتر از مقایسه دو مدل، نمایش موفقیت یک چارچوب کامل و یکپارچه باشد. ما نشان دادیم که می توان از یک سنسور کم هزینه و سفارشی برای جمع آوری داده های با کیفیت و با فرکانس بالا استفاده کرد و سپس این داده ها را مستقیماً به معماری های پیشرفته یادگیری عمیق تزریق نمود تا به مدل های پیش بینی کارآمدی برای یک پارامتر بسیار پیچیده مانند D50 دست یافت. این موفقیت، مفهوم پایش هیدرولوژیکی را از یک فعالیت پرهزینه و محدود به چند ایستگاه خاص، به یک فرایند در دسترس، مقیاس پذیر و هوشمند تبدیل می کند. این چارچوب می تواند راه را برای ایجاد شبکه های پایش متراکم در حوضه های آبریز، به ویژه در مناطقی با محدودیت منابع مالی، هموار سازد و به این ترتیب، مدیریت بالادرنگ و پیشگیرانه منابع آب را به واقعیت نزدیک تر کند.

محدودیت ها و چشم انداز آینده

هر پژوهش علمی، گامی در یک مسیر بی پایان است و اذعان به محدودیت ها، زمینه را برای برداشتن گام های بعدی فراهم می کند. یک نکته روش شناختی مهم در این پژوهش، نحوه ایجاد یک سری زمانی متراکم D50 از روی نمونه های میدانی پراکنده بود. برای این منظور، از یک مدل رگرسیون فرایند گوسی (GP) به عنوان یک روش درون یابی هوشمند استفاده شد. این مدل، با یادگیری الگوهای آماری از نقاط نمونه برداری واقعی، توانست فواصل زمانی بین آن ها را با یک منحنی واقع گرایانه و پیوسته پر کند؛ بنابراین، سری زمانی هدف (Actual D50) که برای آموزش مدل های یادگیری عمیق استفاده شد، ترکیبی از داده های واقعی میدانی (به عنوان نقاط کنترل) و داده های درون یابی شده توسط GP بود. این رویکرد، راهکاری قدرتمند برای غلبه بر کمبود داده است، اما لازم است خواننده در نظر داشته باشد که عملکرد مدل های یادگیری عمیق، توانایی آن ها در یادگیری الگوهای این سری زمانی ترکیبی را منعکس می کند. اعتبارسنجی نهایی در برابر داده های میدانی کاملاً مستقل و با فرکانس بالا، گامی ارزشمند برای کارهای آتی خواهد بود. لازم به ذکر است که این یافته ها، هرچند امیدوارکننده، اما حاصل یک مطالعه موردی بر روی یک رودخانه خاص است و برای تأیید قابلیت تعمیم پذیری این چارچوب، انجام تحقیقات تکمیلی بر روی رودخانه هایی با رژیم های هیدرولوژیکی و مشخصات رسوبی متفاوت، امری ضروری است. علاوه بر این، مدل های توسعه یافته در این پژوهش تنها از یک متغیر ورودی (تراز آب) استفاده کردند. بی شک، افزودن متغیرهای تأثیرگذار دیگر مانند داده های بارندگی از بالادست حوضه، شیب رودخانه یا حتی پارامترهای کیفی آب مانند کدورت، می تواند به افزایش دقت و جامعیت مدل ها کمک شایانی نماید. اول، مجهز کردن مدل

N-BEATS به قابلیت تخمین عدم قطعیت با استفاده از روش‌هایی نظیر مدل‌سازی گروهی، تا بتوان دقت بالای این مدل را با قابلیت ارزیابی ریسک) مشابه آنچه در مدل LSTM ارائه شد (در یک چارچوب واحد ترکیب نمود. دوم، حرکت از پیش‌بینی یک پارامتر واحد (D50) به سمت پیش‌بینی کل منحنی دانه‌بندی رسوبات به‌عنوان خروجی مدل که درک بسیار کامل‌تری از ساختار بستر رودخانه ارائه خواهد داد. علاوه بر این، برای محک‌زدن جامع‌تر عملکرد، مطالعات آتی می‌توانند شامل مقایسه با مدل‌های کلاسیک سری زمانی مانند ARIMA نیز باشند تا میزان بهبود عملکرد حاصل از به‌کارگیری معماری‌های یادگیری عمیق در این چارچوب کم‌هزینه به‌صورت کمی مشخص گردد.

نتیجه‌گیری

این پژوهش باهدف غلبه بر چالش بنیادین کمبود داده‌های با فرکانس بالا در پایش رسوبات رودخانه‌ای، یک چارچوب یکپارچه و نوآورانه از سخت‌افزار تا نرم‌افزار را توسعه و اعتبارسنجی نمود. در این تحقیق، با تلفیق نمونه‌برداری‌های فیزیکی و دقیق میدانی به‌عنوان نقاط مبنا و به‌کارگیری یک مدل رگرسیون فرایند گوسی (GP) برای ایجاد یک سری زمانی پیوسته و واقع‌گرایانه، یک مجموعه داده باکیفیت از قطر متوسط رسوبات بستر (D50) تولید گردید. این سری زمانی سپس به‌عنوان حقیقت زمینی برای آموزش و مقایسه دو معماری پیشرفته یادگیری عمیق شامل شبکه بازگشتی LSTM و شبکه غیربازگشتی N-BEATS با استفاده از داده‌های تراز آب حاصل از یک سنسور اولتراسونیک کم‌هزینه، به کار گرفته شد. یافته‌های کلیدی این پژوهش، یک "موازنه راهبردی" مهم را برای کاربردهای عملی آشکار ساخت. از یک سو، معماری مدرن N-BEATS با دستیابی به دقت بالاتر، برتری خود را در پیش‌بینی نقطه‌ای به اثبات رساند و خود را به‌عنوان ابزاری قدرتمند برای تخمین‌های عددی در کاربردهای عملیاتی معرفی کرد. از سوی دیگر، چارچوب LSTM، با بهره‌گیری از تکنیک Monte Carlo Dropout، قابلیت حیاتی کمی‌سازی عدم قطعیت را به نمایش گذاشت. این ویژگی، LSTM را از یک مدل پیش‌بین صرف، به یک ابزار کارآمد برای تصمیم‌گیری‌های آگاه به ریسک تبدیل می‌کند که در شرایط بحرانی و مدیریت ریسک، ارزشی فراتر از دقت عددی بالاتر دارد. در نهایت، این مطالعه فراتر از مقایسه دو مدل، موفقیت یک راهکار کامل و مقرون‌به‌صرفه از "سنسور تا شبیه‌سازی" را به اثبات رساند. این پژوهش نشان می‌دهد که تلفیق هوشمندانه نمونه‌برداری میدانی هدفمند، مدل‌های آماری پیشرفته برای ایجاد سری‌های زمانی پیوسته، و معماری‌های نوین یادگیری عمیق، می‌تواند راه را برای توسعه نسل جدیدی از سیستم‌های پایش و پیش‌بینی هیدرولوژیکی هموار سازد. این چارچوب، یک نقشه راه عملی برای تولید اطلاعات ارزشمند از داده‌های میدانی گسسته ارائه می‌دهد و پتانسیل بالایی برای دموکراتیزه کردن پایش هوشمند محیط‌زیست، به‌ویژه در مناطق با چالش‌های داده‌ای، دارد.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

نویسندگان اصول اخلاقی را در انجام و انتشار این پژوهش علمی رعایت نموده‌اند و این موضوع مورد تأیید همه آنهاست.

تعارض منافع

بنا بر اظهار نویسندگان این مقاله تعارض منافع ندارد.

منابع

- ادیب راد، رزا، خسرو ی نژاد، بیتا، و ادیب راد، محمدحسین. (۱۴۰۴). کاربرد مدل هوش مصنوعی ماشین آموزش نیرومند با حذف داده پرت (ORELM) در پیش بینی دبی رودخانه. نشریه فناوری های پیشرفته در بهره وری آب، ۵(۳)، ۴۴-۵۵.
<https://doi.org/10.22126/atwe.2025.11918.1159>
- اکبری، علیرضا، پوررضا بیلندی، محسن، خاشعی سیوکی، عباس، کلانکی، مهدی، و میان آبادی، آمنه. (۱۴۰۳). بررسی عدم قطعیت پارامترهای مدل گیاهی DSSAT در برآورد عملکرد گیاه پنبه با استفاده از روش GLUE. نشریه فناوری های پیشرفته در بهره وری آب، ۴(۴)، ۱۹-۳۴.
<https://doi.org/10.22126/atwe.2024.10960.1135>
- پیروزمهر، صفورا، ایزدبخش، محمد علی، شعبانلو، سعید، رجبی، احمد، و یوسفوند، فریبرز. (۱۴۰۴). شبیه سازی نوسانات سطح آب زیرزمینی با استفاده از مدل های یادگیری ماشین رگرسیون خطی (SVR) و یادگیری افراطی (ELM). نشریه فناوری های پیشرفته در بهره وری آب، ۵(۳)، ۱-۱۹.
<https://doi.org/10.22126/atwe.2025.11658.1150>

References

- Adibrad, R., Khosravinezhad, B., & Adibrad, M.H. (2025). Application of outlier robust extreme learning machine (ORELM) in river flow prediction. *Advanced Technologies in Water Efficiency*, 5(3), 44-55. <https://doi.org/10.22126/atwe.2025.11918.1159> [In Persian]
- Akbari, M., Pourreza-Bilondi, M., Khashei-Siuki, A., Kalanki, E., & Mianabadi, A. (2024). Uncertainty analysis of DSSAT model parameters in estimating cotton yield using GLUE method. *Advanced Technologies in Water Efficiency*, 4(4), 19-34. <https://doi.org/10.22126/atwe.2024.10960.1135> [In Persian]
- Boiten, W. (2005). *Hydrometry*. Taylor & Francis Publications, London, United Kingdom. <https://doi.org/10.1201/9780203971093>
- Bowler, A. L., Bakalis, S., & Watson, N. J. (2020). Monitoring mixing processes using ultrasonic sensors and machine learning. *Sensors*, 20(7), 1813. <https://doi.org/10.3390/s20071813>
- Bowler, A. L., Pound, M. P., & Watson, N. J. (2022). A review of ultrasonic sensing and machine learning methods to monitor industrial processes. *Ultrasonics*, 124, 106776. <https://doi.org/10.1016/j.ultras.2022.106776>
- De Camargo, E. T., Spanhol, F. A., Slongo, J. S., da Silva, M. V. R., Pazinato, J., de Lima Lobo, A. V., ... & Martins, L. D. (2023). Low-cost water quality sensors for IoT: A systematic review. *Sensors*, 23(9), 4424. <https://doi.org/10.3390/s23094424>
- Ji, H. W., Yoo, S. S., Lee, B. J., Koo, D. D., & Kang, J. H. (2020). Measurement of wastewater discharge in sewer pipes using image analysis. *Water*, 12(6), 1771. <https://doi.org/10.3390/w12061771>
- Huang, C. C., Chang, M. J., Lin, G. F., Wu, M. C., & Wang, P. H. (2021). Real-time forecasting of suspended sediment concentrations in reservoirs by the optimal integration of multiple machine learning techniques. *Journal of Hydrology: Regional Studies*, 34, 100804. <https://doi.org/10.1016/j.ejrh.2021.100804>
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005-6022. <https://doi.org/10.5194/hess-22-6005-2018>
- Masoudimoghaddam, M., Yazdi, J., & Shahsavandi, M. (2025). A low-cost ultrasonic sensor for online monitoring of water levels in rivers and channels. *Flow Measurement and Instrumentation*, 102, 102777. <https://doi.org/10.1016/j.flowmeasinst.2024.102777>
- Pirouzmehr, M., Izadbakhsh, M. A., Shabanlou, S., Rajabi, A., & Yousefvand, F. (2025). Simulation of groundwater level fluctuations using support vector regression (SVR) and extreme learning machine (ELM) models. *Advanced Technologies in Water Efficiency*, 5(3), 1-19. <https://doi.org/10.22126/atwe.2025.11658.1150> [In Persian]
- Roushangar, K., Shahnazi, S., & Azamathulla, H. M. (2022). Sediment transport modeling through machine learning methods: review of current challenges and strategies. *River dynamics and flood hazards: studies on risk and mitigation*, 223-240. https://doi.org/10.1007/978-981-19-7100-6_13
- Saberian, M., Samadi, V., & Popescu, I. (2024). Probabilistic hierarchical interpolation and interpretable configuration for flood prediction. *Hydrology and Earth System Sciences Discussions*, 2024, 1-41. <https://doi.org/10.5194/hess-2024-261>
- Sahoo, A., & Udgata, S. (2019). A novel ANN based adaptive ultrasonic measurement system for accurate water level monitoring. *IEEE Transactions on Instrumentation and Measurement*, 69(6), 3359–3369. <https://doi.org/10.1109/TIM.2019.2939932>

- Wang, D., Li, Q., Liu, S., Pan, L., & Li, J. (2025). Long-term river flow forecasting: An integrated deep learning model with multi-scale feature extraction. *Expert Systems with Applications*, 281, 127387. <https://doi.org/10.1016/j.eswa.2025.127387>
- Ye, F., Jin, Z., Zhang, P., Xu, D. Z., Lan, L., & Sun, J. (2025). Research on the digital twin polder area system driven by integrating the Xin'anjiang model and the N-BEATS model. *International Journal of Intelligent Systems*, 2025, 8899669. <https://doi.org/10.1155/int/8899669>
- Yu, Y., Zhu, Y., Li, S., & Wan, D. (2014). Time series outlier detection based on sliding window prediction. *The Scientific World Journal*, 2014, 879736. <https://doi.org/10.1155/2014/879736>