Natural Disasters

Homepage: https://disaster.ndri.ac.ir/

Application of deep transformer neural networks for short-term river discharge prediction using hydrological data; case study: the Gorganrud river

Mohammad Hossein Naseri[1], Hamid Ebadi[2], and Abbas Kiani[3]✉

1. M.Sc. Student, Department of Photogrammetry and Remote Sensing, Faculty of Surveying Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: m.naseri1@email.kntu.ac.ir
2. Professor, Faculty of Geodesy and Geomatics Engineering, K.N Toosi University of Technology, Tehran, Iran. E-mail: ebadi@kntu.ac.ir
3. Corresponding Author, Assistant professor, Faculty of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran. E-mail: a.kiani@nit.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 20 February 2025
Received in revised form 11 March 2025
Accepted 26 March 2025
Available online 29 March 2025

Keywords:
Discharge prediction,
Transformer,
Long Short-Term Memory (LSTM),
Flood,
Early warning.

ABSTRACT

Objective: This study presents a deep learning–based approach for short-term discharge prediction at the Aq Qala station on the Gorganrud River and evaluates the Transformer model's ability to reproduce short-term trends and critical conditions. The necessity of this research arises from the increasing frequency and intensity of flood events, which underscore the need for rapid, accurate forecasting tools to support water resource management decisions and early warning systems. To address this need, a Long Short-Term Memory (LSTM) neural network was used as a comparative benchmark to evaluate the proposed method's performance.

Method: Daily discharge, precipitation, and evaporation data were used as model inputs. After normalization and transformation into ten-day sequences, the data were fed into the model. The temporal segmentation included approximately ten years for training, one year for testing, and one year for evaluation (including a flood event) to assess model stability under both normal and extreme conditions. The Transformer model, with its attention mechanism, captures long-term dependencies and complex relationships among input variables. Its parallel processing capability enables faster, more stable training for time-series problems. Both models were trained for one-, two-, and three-day forecasting horizons and evaluated using the Root Mean Square Error (RMSE) and Nash–Sutcliffe Efficiency (NSE) criteria.

Results: For the one-day horizon, the Transformer model outperformed the LSTM, achieving an RMSE of 0.03 and an NSE of 0.92. As the forecasting horizon increased to two and three days, the accuracy of both models decreased, with RMSE values of approximately 0.064 and 0.068, respectively, reflecting the nonlinear and complex nature of hydrological systems. Quantitative and graphical comparisons indicate that the Transformer model performs better in reproducing overall discharge trends and capturing peak flows during critical periods.

Conclusion: The results demonstrate that the Transformer model is an effective tool for short-term river discharge prediction and can support early warning systems and decision-making in water resource management. This approach can serve as an auxiliary tool for water managers, rapid warning systems, and flood risk reduction planners.

Cite this article: Naseri, M. H., Ebadi, H., & Kiani, A. (2025). Application of deep transformer neural networks for short-term river discharge prediction using hydrological data; case study: the Gorganrud river. *Natural Disasters*, *1* (1), 83-94.

© The Author(s).

Publisher: Natural Disasters Research Institute (NDRI).

سوانح طبیعی
Homepage: https://disaster.ndri.ac.ir/

به‌کارگیری شبکه‌های عصبی عمیقِ ترنسفورمر در پیش‌بینی کوتاه‌مدت دبی رودخانه با استفاده از داده‌های هیدرولوژیکی؛ مطالعه موردی: رودخانه گرگان‌رود

محمدحسین ناصری[۱]، حمید عبادی[۲]، عباس کیانی[۳] ✉

۱. دانشجوی کارشناسی ارشد، گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران. رایانامه: m.naseri1@email.kntu.ac.ir

۲. استاد، گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران. رایانامه: ebadi@kntu.ac.ir

۳. نویسنده مسئول، استادیار، گروه نقشه‌برداری، دانشکده مهندسی عمران، دانشگاه صنعتی نوشیروانی بابل، مازندران، ایران. رایانامه: a.kiani@nit.ac.ir

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

تاریخ دریافت: ۱۴۰۳/۱۲/۰۲
تاریخ بازنگری: ۱۴۰۳/۱۲/۲۱
تاریخ پذیرش: ۱۴۰۴/۰۱/۰۶
تاریخ انتشار: ۱۴۰۴/۰۱/۰۹

کلیدواژه‌ها:
پیش‌بینی دبی،
ترنسفورمر،
حافظه طولانی کوتاه‌مدت،
سیلاب،
هشدار زودهنگام.

چکیده

هدف: هدف این پژوهش ارائه روشی مبتنی بر یادگیری عمیق برای پیش‌بینی کوتاه‌مدت دبی رودخانه گرگان‌رود در ایستگاه آق‌قلا و سنجش توانایی مدل ترنسفورمر در بازتولید روندهای کوتاه‌مدت و شرایط بحرانی است. ضرورت این تحقیق از افزایش فراوانی و شدت رخدادهای سیلابی نشأت می‌گیرد، زیرا این روند، نیاز به ابزارهای پیش‌بینی سریع و دقیق برای پشتیبانی تصمیم‌گیری مدیریت منابع آب و سامانه‌های هشدار را تشدید کرده است. برای پاسخ به این نیاز، شبکه عصبی حافظه طولانی کوتاه‌مدت در نقش مرجع مقایسه‌ای مورد استفاده قرار گرفت تا توانایی روش پیشنهادی سنجیده شود.

روش پژوهش: داده‌های روزانه دبی، بارش و تبخیر ورودی مدل در این پژوهش بودند که پس از نرمال‌سازی و تبدیل به توالی‌های ده‌روزه وارد مدل شدند. تقسیم‌بندی زمانی به‌طور تقریبی شامل ۱۰ سال برای آموزش، ۱ سال برای آزمایش و ۱ سال برای ارزیابی (شامل رخداد سیلاب) بود تا پایداری مدل در شرایط عادی و بحرانی سنجیده شود. مدل ترنسفورمر با مکانیزم توجه در شناسایی وابستگی‌های بلندمدت و روابط پیچیده بین متغیرهای ورودی را دارد و پردازش موازی موجب آموزش سریع‌تر و پایدارتر در مسائل سری‌زمانی می‌گردد. هر دو مدل برای پیش‌بینی افق‌های زمانی یک، دو و سه‌روزه آموزش و با معیارهای ریشه میانگین مربعات خطا و پارامتر نش–ساتکلیف ارزیابی شدند.

یافته‌ها: در افق یک‌روزه، مدل ترنسفورمر با ریشه میانگین مربعات خطا برابر با ۰/۰۳ و ضریب نش–ساتکلیف برابر با ۰/۹۲ عملکرد برتری نشان داد. با افزایش افق زمانی پیش‌بینی به دو و سه روز، دقت هر دو مدل کاهش یافت به‌طوری‌که ریشه میانگین مربعات خطا در حدود ۰/۰۶۴ و ۰/۰۶۸ گزارش شد که نشان‌دهنده پیچیدگی و غیرخطی بودن سیستم هیدرولوژیکی است. مقایسه کمی و نموداری نشان می‌دهد ترنسفورمر در بازتولید روند کلی و ردیابی اوج دبی در دوره بحرانی عملکرد مطلوبی دارد.

نتیجه‌گیری: نتایج پژوهش نشان می‌دهد که ترنسفورمر به‌عنوان یک ابزار مؤثر برای پیش‌بینی کوتاه‌مدت دبی و پشتیبانی از سامانه‌های هشدار زودهنگام و تصمیم‌گیری مدیریت منابع آب مناسب است. این روش می‌تواند ابزاری کمکی در تصمیم‌گیری برای مدیران منابع آب و سامانه‌های هشدار سریع و برنامه‌ریزان کاهش خطر سیلاب باشد.

استناد: ناصری، محمدحسین؛ عبادی، حمید؛ کیانی، عباس. (۱۴۰۴). به‌کارگیری شبکه‌های عصبی عمیقِ ترنسفورمر در پیش‌بینی کوتاه‌مدت دبی رودخانه با استفاده از داده‌های هیدرولوژیکی؛ مطالعه موردی: رودخانه گرگان‌رود. *سوانح طبیعی*، ۱ (۱)، ۸۳-۹۴.

CC BY NC
© نویسندگان.

ناشر: پژوهشکده سوانح طبیعی.

مقدمه

رودخانه‌ها از دیرباز منابع حیاتی تأمین آب شرب، کشاورزی و صنعت بوده‌اند؛ با این حال مخاطرات طبیعی رودخانه‌ها ازجمله سیل‌ها و خشکسالی‌های ناگهانی همواره تهدیدی جدی برای جوامع به‌شمار می‌روند (Eltahir & Wang, 1999). امروزه برای مقابله با نوسانات خطرناک جریان رودخانه‌ها، تغییرات اقلیمی، تشدید حوادث آب‌وهوایی و مدیریت بهینه منابع آب سازه‌هایی چون سد و شبکه‌های پایش ایجاد شده است. همچنین پیش‌بینی دقیق و به‌موقع جریان رودخانه‌ها برای اطلاع‌رسانی و تصمیم‌گیری درست برای کنترل سیلاب و برنامه‌ریزی منابع آب ضروری است (Cisterna-García et al., 2025). رودخانه‌ها سیستم‌های پیچیده‌ای‌اند که تحت تأثیر عوامل متنوع اقلیمی، هیدرولوژیکی و انسانی قرار دارند. پیش‌بینی دبی به علت ماهیت غیرخطی و وابستگی‌های زمانی پیچیده داده‌های هیدرولوژیکی به‌ویژه در شرایط بحرانی (مانند دوره‌های بارش شدید و رخدادهای سیلابی) با چالش‌هایی همراه است. همین امر سبب می‌شود توانایی مدل در تشخیص سریع روندهای افزایشی و اوج‌های جریان اهمیت زیادی داشته باشد (Yaseen et al., 2015).

هدف این پژوهش ارائه روشی مبتنی بر یادگیری عمیق برای پیش‌بینی کوتاه‌مدت دبی رودخانه گرگان‌رود در ایستگاه آق‌قلا و بررسی توانایی مدل ترنسفورمر[1] در بازتولید روندهای کوتاه‌مدت در شرایط غیرایستا و پرنوسان رودخانه‌ها (به‌خصوص هنگام سیلاب) است. مدل ترنسفورمر روش اصلی پژوهش است که به دلیل قابلیت مکانیزم توجه[2] در شناسایی وابستگی‌های بلندمدت و پردازش موازی داده‌های سری‌زمانی انتخاب شده و شبکه عصبی حافظه طولانی کوتاه‌مدت (LSTM)[3] به منظور مرجعی برای مقایسه عملکرد به‌کار گرفته شده است. داده‌ها به سه دسته آموزشی، آزمایشی و ارزیابی تقسیم شده‌اند. ارزیابی مدل شامل رخداد سیلاب بود تا پایداری مدل در شرایط عادی و بحرانی سنجیده شود.

پیشینه پژوهش

پژوهش‌های متعددی با هدف پیش‌بینی جریان رودخانه با بهره‌گیری از روش‌های مختلفی، از رویکردهای کلاسیک خطی تا الگوریتم‌های نوین یادگیری ماشین و یادگیری عمیق، انجام شده‌اند. روش‌های سری‌زمانی خطی مانند مدل خود همبسته یکپارچه میانگین متحرک (ARIMA)[4] به‌طور سنتی برای مدل‌سازی جریان به‌کار رفته که در مواجهه با رفتارهای غیرخطی محدودیت‌هایی نشان داده‌اند (Kilinc & Haznedar, 2022). برای مثال، پاتل و همکاران در یک مقایسه بین ARIMA، شبکه عصبی مصنوعی (ANN)[5] و ماشین بردار پشتیبان (SVM)[6] نشان دادند که مدل‌های ARIMA توانایی کمتری در ثبت دینامیک غیرخطی جریان نسبت به روش‌های یادگیری ماشین داشته است (Patel & Ramachandran, 2015). از این رو، استفاده از مدل‌های پیشرفته و دقیق در محیط‌های غیرخطی هیدرولوژیکی ضروری است. چنین مدل‌هایی می‌توانند تغییرات سریع بارش را با دقت بهتر ثبت کند و پیش‌بینی‌های قابل اعتمادتری برای هشدار سیلاب ارائه دهند (Chen et al., 2020).

در دهه‌های اخیر توجه به روش‌های غیرخطی افزایش‌یافته است چرا که بر اساس یافته‌های پژوهش‌ها، الگوریتم SVM به دلیل توانایی در ثبت روابط غیرخطی و شرایط داده‌ای محدود به‌طور گسترده در مدل‌سازی هیدرولوژیکی استفاده شده است (Yu et al., 2017). برای مثال، اسلام و همکاران با بهره‌گیری از داده‌های ماهواره‌ای نشان دادند که SVM در پیش‌بینی جریان رودخانه‌ای برتری معناداری نسبت به مدل SWAT[7] داشته است (Islam et al., 2023). یکی دیگر از روش‌های محبوب در چند سال اخیر، جنگل تصادفی (RF)[8] است. RF به‌عنوان یک الگوریتم ترکیبی از درخت‌های تصمیم، با ساخت و میانگین‌گیری نتایج تعداد زیادی از درخت‌های تصمیم، دقت پیش‌بینی را افزایش می‌دهد (Jesus et al., 2025). در مثالی دیگر، یو و همکاران

1. Transformer
2. attention
3. Long Short Term Memory
4. Autoregressive Integrated Moving Average
5. Artificial Neural Network
6. Support Vector Machine
7. Soil and Water Assessment Tool
8. Random Forest

نشان دادند که RF حتی در حوضه آبریزهای ناشی از ذوب برف نسبت به مدل‌های فیزیکی سنتی دقت بهتری دارد (Yu et al., 2017). با این حال، الگوریتم‌های یادگیری ماشین هنگام کار با مجموعه‌داده‌های بسیار حجیم هزینه محاسباتی بالایی دارد و به دقت تنظیم پارامترها حساس است (Mahmood et al., 2024).

از سویی روش‌های ANN نظیر پرسپترون چندلایه (MLP)[9] و مدل‌های مبتنی بر رگرسیون غیرخطی به‌خوبی می‌توانند روابط پیچیده داده‌های هیدرولوژیک را یاد بگیرند (Shamshirband et al., 2019). در بین مدل‌های پیشرفته یادگیری عمیق، LSTM به‌طور گسترده‌ای آزموده شده است. این مدل می‌تواند وابستگی‌های بلندمدت زمانی را ثبت کند. برای مثال لیسکچن و همکاران نشان دادند که LSTM به‌خوبی دینامیک غیرخطی را یاد می‌گیرد و در پیش‌بینی دبی رودخانه عملکرد بالایی دارد (Leščešen et al., 2025). در پژوهش متأخر آتشی و همکاران با استفاده از داده‌های ساعتی سطح آب در سه ایستگاه هیدرولوژی و افزودن متغیرهای اقلیمی شامل دما، رطوبت و تبخیر نشان دادند که ورود اطلاعات اقلیمی به‌طور قابل‌توجهی دقت پیش‌بینی‌های کوتاه‌مدت را افزایش می‌دهد اما با افزایش افق زمانی کاهش دقت رخ می‌دهد (Atashi & Gorji, 2025). یافته‌های تحقیقات نوین نشان می‌دهد مدل ترنسفورمر در پیش‌بینی جریان رودخانه قادر است بهتر از مدل‌های مرسوم LSTM و واحد بازگشتی دروازه‌ای (GRU)[10] عمل کند (Demiray & Demir, 2024). برای نمونه، کاستانگیا و همکاران با به‌کارگیری داده‌های تاریخی سطح آب در یک سیلاب بزرگ در جنوب‌شرقی اروپا، نشان دادند که مدل ترنسفورمر در بازتولید ساختار زمانی سیلاب‌ها عملکرد بهتری نسبت به LSTM دارد. این امر نمایانگر توانایی مکانیزم خودتوجهی در یادگیری الگوهای پیچیده زمانی است (Castangia et al., 2023).

علی‌رغم پیشرفت گسترده در روش‌های پیش‌بینی جریان، هر یک از روش‌های مذکور دارای مزایا و محدودیت‌های خاص خود است. روش‌های یادگیری عمیق معمولاً به انبوه داده‌ها و تنظیمات پارامتری نیاز دارند که می‌تواند منجر به مشکلاتی مانند بیش‌برازش شود. در حالی که روش‌های ترکیبی پیچیدگی محاسباتی بالاتری دارند. همچنین بسیاری از مطالعات گذشته بیشتر بر پیش‌بینی بلندمدت یا حوضه‌های دائمی تمرکز کرده‌اند و پژوهش‌های اندکی به موضوع پیش‌بینی کوتاه‌مدت در رودخانه‌های با جریان متغیر (مانند شرایط سیلابی) پرداخته‌اند. این کمبودها ضرورت معرفی مدلی را تأیید می‌کند که بتواند روابط غیرخطی و پیچیده را در شرایط بحرانی بیاموزد و دقت بالایی ارائه کند. هدف اصلی این پژوهش ارائه مدلی با دقت بالای پیش‌بینی است که بتواند در شرایط غیرایستا و پرنوسان رودخانه‌ها (به‌خصوص هنگام سیلاب) عملکرد بهتری داشته باشد. در این پژوهش از ساختار ترنسفورمر به منظور پیش‌بینی دبی رودخانه بهره‌گرفته‌ایم. گفتنی‌است که این روش اخیراً در مسائل مختلف زمانی جایگاه ویژه‌ای یافته است. مدل ترنسفورمر با مکانیزم توجه توانایی مناسبی در درک وابستگی‌های بلندمدت و ناهمگونی داده‌های سری‌زمانی دارد.

روش‌شناسی پژوهش

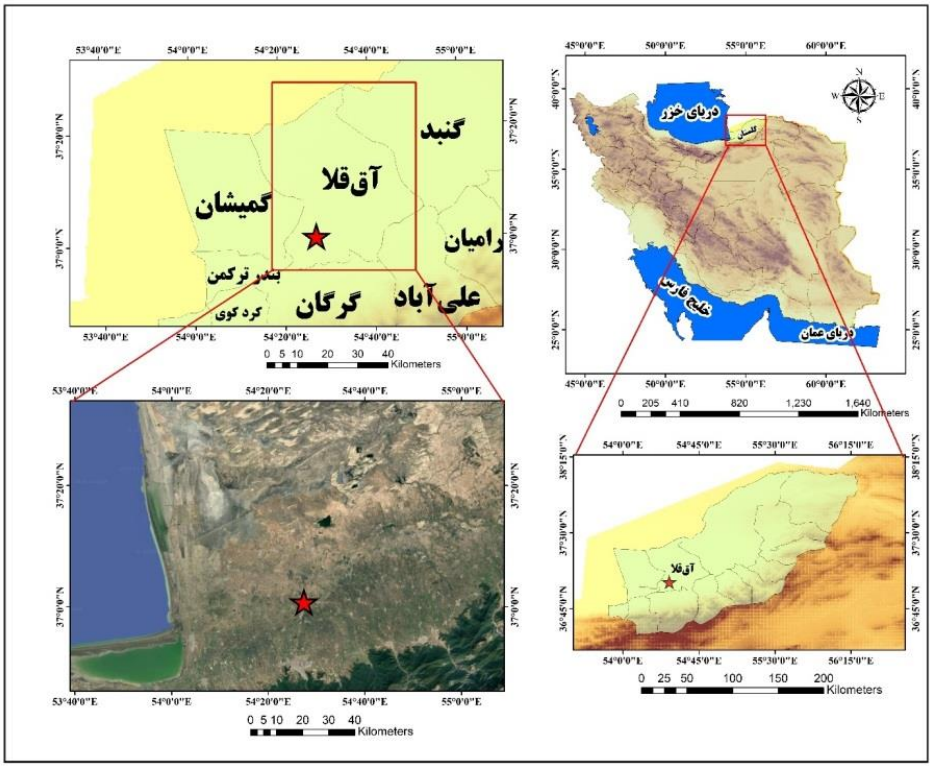
استان گلستان از مناطق آسیب‌پذیر سیلابی در ایران است. وقوع سیلاب‌های ویرانگر اخیر و به تبع تلفات و خسارات قابل‌توجه، ضرورت انجام پژوهش‌های کاربردی در این حوزه را افزایش می‌دهد. با توجه به عملکرد مناسب روش‌های یادگیری عمیق در مدل‌سازی سری‌های زمانی و توانایی آن‌ها در یادگیری روابط غیرخطی و وابستگی‌های بلندمدت، در این پژوهش از مدل‌های ترنسفورمر و LSTM برای پیش‌بینی دبی استفاده خواهد شد.

منطقه مورد مطالعه

شهرستان آق‌قلا در شمال استان گلستان و در عرض جغرافیایی $36°58'$ و طول $54°16'$ قرار دارد. این منطقه در پهنه‌ای هموار و کم‌ارتفاع از دشت گرگان قرار دارد که ویژگی‌های توپوگرافی آن، نظیر شیب اندک و اراضی پهن، سبب شده تا آق‌قلا یکی از نقاط حساس به وقوع سیلاب در کشور محسوب شود. مهم‌ترین جریان سطحی تأثیرگذار در این ناحیه، رودخانه گرگان‌رود است که از شرق استان سرچشمه گرفته و با گذر از شهرهایی نظیر گنبد کاووس و آق‌قلا، در نهایت به دریای خزر می‌ریزد. ویژگی‌های هیدرولوژیکی این رودخانه از جمله طول نسبتاً زیاد (حدود ۳۰۰ کیلومتر) و حوضه آبریز گسترده آن، باعث

9. Multi Layer Perceptron
10. Gated Recurrent Unit

می‌شود تا در دوره‌های بارندگی شدید، حجم قابل توجهی از رواناب به سمت مناطق پایین‌دست، از جمله آق‌قلا، حرکت کند. در طول دهه‌های اخیر، این شهرستان بارها شاهد وقوع سیلاب‌های قابل‌توجه بوده و یکی از شاخص‌ترین آن‌ها در فروردین‌ماه ۱۳۹۸ رخ داد. بارش‌های مداوم و بالاآمدن آب رودخانه منجر به سیلابی شد که در بخش‌های وسیعی از منطقه خسارت وارد کرد. موقعیت خاص جغرافیایی منطقه، عبور رودخانه گرگان‌رود و وقوع سیل‌های مخرب، اهمیت تحلیل زمانی سیلاب در این شهرستان را افزایش می‌دهد. در شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه نشان داده شده است.



شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه

داده‌ها

در این پژوهش، از داده‌های روزانه دبی، بارش و تبخیر رودخانه گرگان رود در ایستگاه آق‌قلا، واقع در استان گلستان، استفاده شده است. این داده‌ها بازه زمانی ۱۳ ساله از ابتدای سال ۱۳۸۶ تا پایان سال ۱۳۹۸ را پوشش می‌دهند و از سوی سازمان مدیریت منابع آب ایران (وابسته به وزارت نیرو) تأمین شده‌اند. ایستگاه مورد مطالعه با کد ۱۲۰۳۷ و مختصات جغرافیایی ۵۴٫۴۶۶۷ درجه طول شرقی و ۳۷٫۰۱۶۶ درجه عرض شمالی، در دشتی هموار و کم‌ارتفاع واقع شده است که از نظر هیدرولوژیکی تحت تأثیر مستقیم جریان‌های سطحی رودخانه گرگان‌رود قرار دارد.

مبانی تئوری تحقیق

پیش‌بینی دبی به مدل‌سازی توالی‌های زمانی غیرخطی نیاز دارد. ساختارهای بازگشتی حافظه‌دار مانند LSTM توانایی حفظ اطلاعات بلندمدت و مدل‌سازی وابستگی‌های زمانی را فراهم می‌کنند و از سوی دیگر معماری‌های مبتنی بر خودتوجهی (ترنسفورمر) با امکان وزن‌دهی دینامیک بین عناصر توالی و پردازش موازی، ظرفیت بالاتری در کشف روابط پیچیده نشان می‌دهند. بنابراین در این پژوهش، روش اصلی ترنسفورمر و مرجع مقایسه LSTM انتخاب شد. برای سنجش عملکرد، از ریشه میانگین مربعات خطا (RMSE)[11] به‌عنوان شاخص خطای عددی و از نش–ساتکلیف (NSE) به‌عنوان شاخص کیفیت بازتولید الگوهای زمانی استفاده شده است. در ادامه هریک از روش‌ها به تفصیل بررسی شده است.

11. Root Mean Square Error

شبکه عصبی حافظه طولانی کوتاه‌مدت

LSTM برای حل مشکلات شبکه عصبی بازگشتی (RNN)[۱۲] در پردازش دنباله‌های طولانی طراحی شده است. LSTM یک نوع خاص از RNN است که برای بهبود یادگیری در دنباله‌های طولانی طراحی شده است. نابودی گرادیان از مشکلات اصلی RNN است که می‌تواند یادگیری را در دنباله‌های طولانی به شدت مختل کند. LSTM با معرفی ساختار و مکانیسم‌های خاصی برای حافظه و کنترل جریان اطلاعات، این مشکلات را برطرف می‌کند. LSTM به طور خاص از واحدهای حافظه به نام سلول‌های حافظه استفاده می‌کنند. این سلول‌ها با استفاده از مکانیسم‌های مختلف مانند دروازه ورودی، دروازه فراموشی و دروازه خروجی اطلاعات را به صورت انتخابی ذخیره و بازیابی می‌کنند. دروازه ورودی تعیین می‌کند که چه مقدار از اطلاعات جدید باید وارد سلول حافظه شود. دروازه فراموشی تصمیم می‌گیرد که چه مقدار از اطلاعات قبلی باید حذف شود. دروازه خروجی تعیین می‌کند که چه مقدار از اطلاعات سلول حافظه باید به خروجی منتقل شود. این مکانیسم‌ها به مدل اجازه می‌دهند که اطلاعات را برای مدت زمان طولانی‌تری حفظ کنند و به این ترتیب مشکل ناپدید شدن و انفجار گرادیان را بهبود بخشند (Hochreiter, 1997).

شبکه عصبی ترنسفورمر

ترنسفورمر نخستین‌بار توسط واسوانی و همکاران در سال ۲۰۱۷ معرفی شد. این مدل بر پایه مکانیزم توجه طراحی شده و برخلاف مدل‌های RNN و LSTM به استفاده از ساختارهای تکراری برای پردازش داده‌های دنباله‌ای نیازی ندارد. هدف از ارائه این معماری، رفع محدودیت‌های موجود در مدل‌های بازگشتی از جمله ناپایداری گرادیان، کندی در آموزش و دشواری در یادگیری وابستگی‌های طولانی‌مدت بود. هسته اصلی ترنسفورمر را مکانیزم خودتوجهی[۱۳] تشکیل می‌دهد که به مدل این امکان را می‌دهد تا ضمن مقایسه هر عنصر ورودی با سایر عناصر موجود در توالی، اهمیت نسبی آن‌ها را هم مشخص کند. این ویژگی سبب می‌شود مدل بتواند وابستگی‌های بلندمدت را بدون نیاز به پردازش ترتیبی، به‌صورت موازی درک و پردازش کند که این مسئله موجب بهبود چشمگیر در سرعت آموزش و دقت مدل می‌شود. ساختار ترنسفورمر از دو بخش اصلی رمزگذار[۱۴] و رمزگشا[۱۵] تشکیل شده که هر یک شامل چندین لایه متشکل از بلوک‌های توجه و شبکه‌های عصبی پیش‌خور[۱۶] است. از آنجا که این مدل برخلاف RNN به‌صورت ذاتی اطلاعات مربوط به ترتیب عناصر را در خود ندارد، از بردارهای موقعیتی[۱۷] در نقش افزونه‌ای به ورودی استفاده می‌شود تا ترتیب زمانی یا مکانی داده‌ها به مدل منتقل شود. این ساختار موجب شده است تا ترنسفورمر به معماری‌ای قدرتمند در حوزه‌های مختلف یادگیری عمیق، از جمله ترجمه ماشینی، پردازش زبان طبیعی و همچنین تحلیل سری‌های زمانی در مسائل پیش‌بینی، تبدیل شود و در بسیاری از موارد عملکرد بهتری نسبت به معماری‌های پیشین از خود نشان دهد (Vaswani et al., 2017).

معیارهای ارزیابی

برای ارزیابی عملکرد مدل‌های پیشنهادی در پیش‌بینی دبی رودخانه، از دو معیار متداول آماری ریشه میانگین مربعات خطا (RMSE)[۱۸] و ضریب نش ساتکلیف (NSE)[۱۹] استفاده شده است. معیار RMSE که میزان خطای مطلق مدل در پیش‌بینی مقدار واقعی را نشان می‌دهد، در رابطه (۱) تعریف شده است و هرچه مقدار آن کمتر باشد، دقت مدل بالاتر خواهد بود (Chai & Draxler, 2014). در مقابل، معیار NSE از شاخص‌های مهم در ارزیابی مدل‌های هیدرولوژیکی، نشان‌دهنده قابلیت مدل در بازتولید الگوها و تغییرات زمانی موجود در داده‌های مشاهده‌شده و مقدار آن بین منفی بی‌نهایت تا ۱ متغیر است. طبق رابطه (۲) هرچه مقدار NSE به ۱ نزدیک‌تر باشد، عملکرد مدل مطلوب‌تر ارزیابی می‌شود (Knoben et al., 2019).

12. Recurrent Neural Network
13. Self Attention
14. Encoder
15. Decoder
16. Feed-Forward Neural Network
17. Positional Encoding
18. Root Mean Square Error
19. Nash Sutcliffe model efficiency coefficient

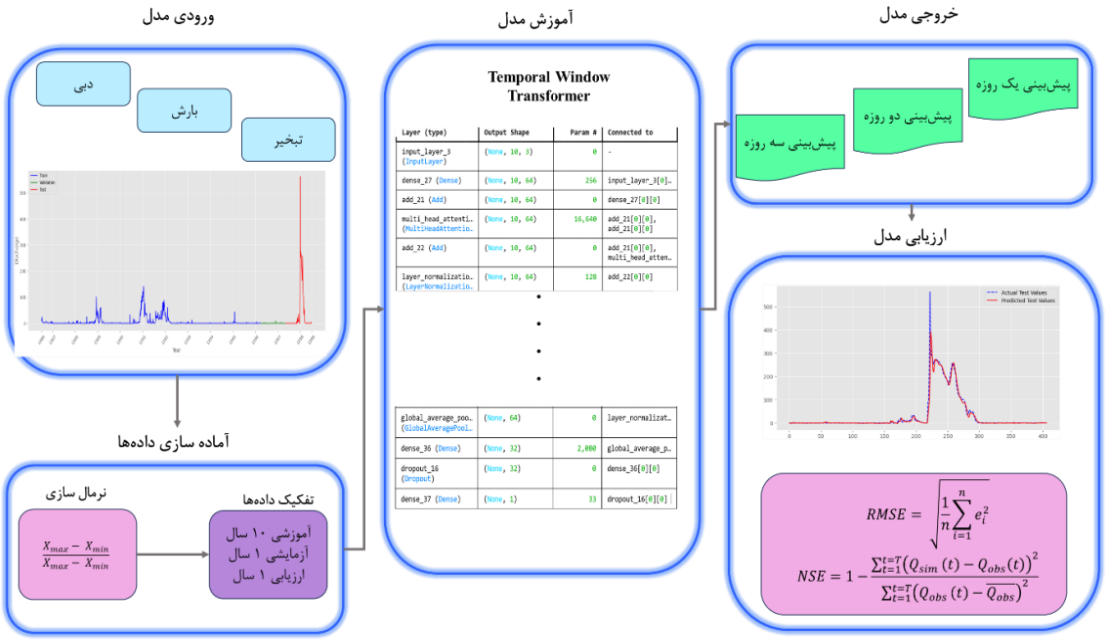
رابطه ۱) $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$ n تعداد نمونه‌ها و e_iخطای هر نمونه

رابطه ۲) $$NSE = 1 - \frac{\sum_{t=1}^{t=T}(Q_{sim}(t) - Q_{obs}(t))^2}{\sum_{t=1}^{t=T}(Q_{obs}(t) - \overline{Q_{obs}})^2}$$

$Q_{sim}(t)$ مقدار شبیه‌سازی شده و $Q_{obs}(t)$ مقدار مشاهده‌شده در زمان t، $\overline{Q_{obs}}$ میانگین مقدار مشاهده‌شده و T تعداد کل بازه‌های زمانی است.

روش پیشنهادی

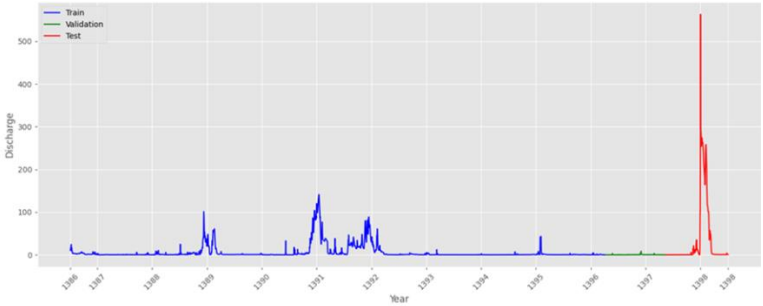
در شکل ۲ فلوچارت کلی روش پیشنهادی نمایش داده شده است. این فلوچارت از سه فاز اصلی تشکیل می‌شود: آماده‌سازی داده‌ها، آموزش مدل و ارزیابی. در ادامه هر مرحله به‌صورت کامل تشریح شده است.



شکل ۲. روند نمای کلی فرآیند پیش‌بینی دبی با استفاده از مدل ترنسفورمر

آماده‌سازی داده‌ها

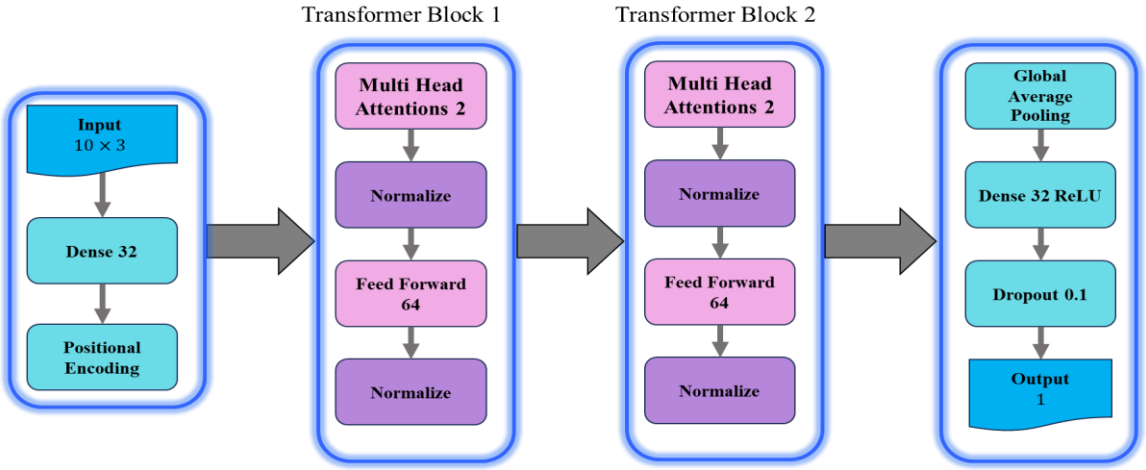
در این مرحله ابتدا داده‌های دبی، بارش و تبخیر به‌منظور افزایش کارایی مدل و تسهیل در هم‌مقیاس‌سازی با استفاده از نرمال‌سازی به بازه (۰ و ۱) منتقل شده‌اند. برای تقسیم داده‌ها به سه دسته آموزشی، آزمایشی و ارزیابی، داده‌های مربوط به سال‌های ۱۳۸۶ تا ۱۳۹۶ داده آموزشی، سال ۱۳۹۷ آزمایشی و سال ۱۳۹۸ برای ارزیابی نهایی در نظر گرفته شده‌اند. اهمیت انتخاب سال ۱۳۹۸ برای ارزیابی به سبب وقوع یکی از سیلاب‌های بزرگ در منطقه آق‌قلا در همان سال بود که شرایط آزمون مدل را چالش‌برانگیزتر کرده است. این تقسیم‌بندی به مدل امکان آموزش دیدن در شرایط عادی و آزموده‌شدن در مواجهه با رخدادهای شدید را می‌دهد. در شکل ۳ مجموع داده‌های دبی به تفکیک آموزشی، آزمایشی و ارزیابی مشاهده می‌شود.



شکل ۳. داده‌های دبی ایستگاه آق‌قلا، ۱۰ سال برای آموزش، یک سال آزمایش و یک سال جهت ارزیابی در نظر گرفته شد

آموزش مدل

آموزش مدل بر پایه ساختار ترنسفورمر با استفاده از توالی‌های ده‌روزه از سه متغیر دبی، بارش و تبخیر انجام شد. تبدیل اولیه ویژگی با یک لایه تراکم[۲۰] به بعد ۳۲، افزودن کدگذاری موقعیتی[۲۱]، دو بلاک رمزگذار شامل مکانیزم توجه چندسری[۲۲] با دو سر و ابعاد پیش‌خور ۶۴، نرمال‌سازی لایه‌ای و اتصال‌های باقیمانده پس از بلاک‌های رمزگذار برای فشرده‌سازی توالی استفاده شد و سپس یک لایه تراکم با ۳۲ واحد، Dropout با نرخ ۰٫۱ و در نهایت لایه خروجی خطی برابر ۱ برای تولید مقدار دبی پیش‌بینی‌شده قرار گرفت. تابع زیان انتخاب‌شده میانگین مربعات خطا (MSE) [۲۳] و بهینه‌ساز آدام[۲۴] با نرخ یادگیری (۰٫۰۰۰۱) بود. معیارهای پیگیری در زمان آموزش شامل میانگین قدرمطلق خطا (MAE) [۲۵] بود. آموزش برای حداکثر ۲۰۰ دوره[۲۶] با اندازه دسته[۲۷] ۱۲۸ انجام شد شکل ۴ ساختار مدل طراحی شده را نمایش می‌دهد.

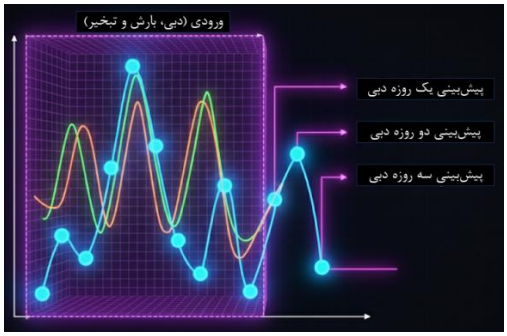


شکل ٤. ساختار کلی مدل ترنسفورمر طراحی شده

ارزیابی مدل

مطابق شکل ۵ مدل مورد نظر با استفاده از داده‌های ده روزه دبی، بارش و تبخیر، دبی یک الی سه روز آینده را پیش‌بینی می‌کند برای سنجش کمی عملکرد مدل از معیارهای RMSE و NSE استفاده شده است. علاوه بر محاسبات کمی، تطابق زمانی پیش‌بینی‌ها و مشاهدات رسم و به‌ویژه بازه مربوط به رخداد سیلاب سال ۱۳۹۸ با دقت بررسی شد تا توانایی مدل در دنبال‌کردن روندهای ناگهانی و اوج‌های جریان مشخص شود. برای اعتبارسنجی جامع‌تر، عملکرد مدل ترنسفورمر با مدل LSTM مقایسه شد. ضمن ارزیابی هر دو مدل با همان داده‌ها، پنجره زمانی و معیارها، نتایج مقایسه‌ای برای نشان‌دادن برتری‌ها و محدودیت‌های هر روش در شرایط عادی و بحران (سیلاب) ارائه شد.

20. Dense
21. positional encoding
22. Multi-Head Attention
23. Mean Square Error
24. Adam
25. Mean Absolute Error
26. Epoch
27. Batch Size



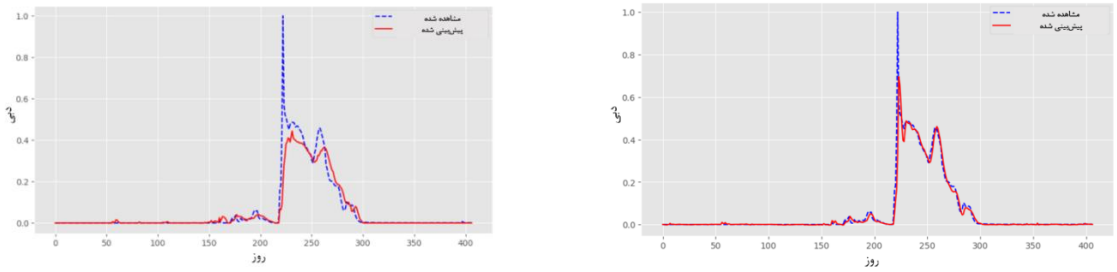
شکل ۵. پیش‌بینی دبی یک روزه، دو روزه و سه‌روزه با ورودی توالی‌های ده‌روزه دبی، بارش و تبخیر

یافته‌های پژوهش

در این بخش نتایج کمی و نموداری ارائه می‌شود. ابتدا نتایج پیش‌بینی زمانی برای هر یک از سه افق زمانی مورد مطالعه نمایش داده شده سپس مقایسه کمی عملکرد دو مدل آورده می‌شود.

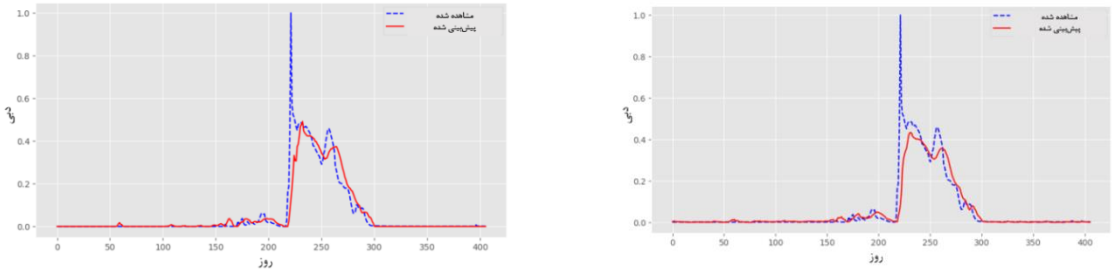
پیش‌بینی

در این بخش نتایج پیش‌بینی دبی برای سه افق زمانی یک‌روزه، دوروزه و سه‌روزه ارائه می‌شود. شکل ۶ نتایج پیش‌بینی یک‌روزه، شکل ۷ نتایج پیش‌بینی دوروزه و شکل ۸ نتایج پیش‌بینی سه‌روزه را نشان می‌دهند. هر یک از این شکل‌ها خروجی مدل ترنسفورمر و مدل LSTM را نمایش می‌دهد. منحنی‌های آبی‌رنگ بیانگر مقادیر مشاهده‌شده و منحنی‌های قرمزرنگ مقادیر پیش‌بینی‌شده می‌باشد.



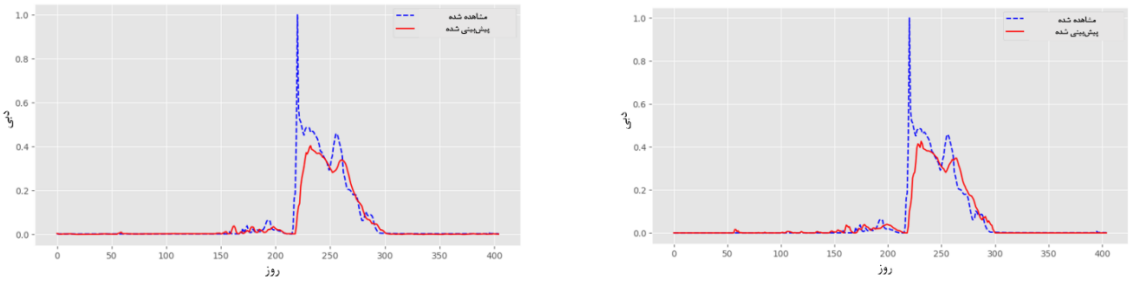
الف ب

شکل ٦. مقادیر دبی پیش‌بینی‌شده و مشاهده‌شده برای افق یک‌روزه الف: مدل ترنسفورمر ب: مدل LSTM



الف ب

شکل ۷. مقادیر دبی پیش‌بینی‌شده و مشاهده‌شده برای افق دوروزه الف: مدل ترنسفورمر ب: مدل LSTM



الف ب

شکل ۸. مقادیر دبی پیش‌بینی‌شده و مشاهده‌شده برای افق سه‌روزه الف: مدل ترنسفورمر ب: مدل LSTM

مقایسه کلی عملکرد مدل‌ترنسفورمر و LSTM بر اساس معیارهای NSE وRMSE

در این بخش مقایسه کلی عملکرد دو مدل ترنسفورمر و LSTM برای سه افق پیش‌بینی یک‌روزه، دوروزه و سه‌روزه ارائه می‌شود. هر شکل شامل سه جفت نمودار میله‌ای است. هر جفت مربوط به یک افق پیش‌بینی است شکل (۹) مربوط به معیار NSE و شکل (۱۰) معیار RMSE می‌باشد.

شکل ۹. مقایسه کلی عملکرد مدل‌ها بر اساس معیار NSE برای پیش‌بینی یک‌روزه، دوروزه و سه‌روزه

شکل ۱۰. مقایسه کلی عملکرد مدل‌ها بر اساس معیار RMSE برای افق‌های پیش‌بینی یک‌روزه، دوروزه و سه‌روزه

بحث

در این بخش نتایج کمی و نموداری حاصل از ارزیابی مدل‌های ترنسفورمر و LSTM برای پیش‌بینی دبی رودخانه گرگان‌رود در ایستگاه آق‌قلا تحلیل شده‌اند. مقایسه مدل‌ها در جدول ۱ نشان داده شده است.

جدول ۱. مقایسه مدل‌های ترنسفورمر و LSTM را برای پیش‌بینی یک روزه، دو روزه و سه روزه نشان می‌دهد

افق پیش‌بینی	مدل	معیار NSE	افت عملکرد نسبت به روز اول	برتری مدل	تحلیل کیفی عملکرد در سیلاب ۱۳۹۸
یک‌روزه	ترنسفورمر	۰/۹۲	–	ترنسفورمر	پاسخ سریع و تخمین نسبتاً دقیق اوج سیلاب
	LSTM	۰/۸۵	–		تخمین مناسب اما کم برآورد کردن اوج سیلاب
دوروزه	ترنسفورمر	۰/۷۹	کاهش ۱۳ درصد	LSTM	تأخیر در پیک سیلاب
	LSTM	۰/۸۱	کاهش ۴ درصد		حفظ نسبی مقدار اوج
سه‌روزه	ترنسفورمر	۰/۷۷	کاهش ۱۵ درصد	ترنسفورمر	حفظ روند کلی و دقت اوج کمتر
	LSTM	۰/۷۶	کاهش ۹ درصد		ناپایداری نسبت به زمان

مطابق جدول ۱ ترنسفورمر در افق یک‌روزه برتری مشهودی نسبت به LSTM دارد. افت عملکرد نسبت به افق یک‌روزه برای ترنسفورمر به‌ترتیب حدود ۱۳ درصد (دوروزه) و ۱۵ درصد (سه‌روزه) است، در حالی که برای LSTM افت دقت حدود ۴ درصد و ۹ درصد می‌باشد. در افق دوروزه اختلاف عملکرد اندک و به نفع LSTM است و در افق سه‌روزه عملکرد دو مدل تقریباً برابر می‌شود. این الگو نمایانگر کاهش دقت با افزایش افق زمانی پیش‌بینی و حساسیت بالاتر ترنسفورمر به طول افق زمانی در مقایسه با LSTM است. بنابراین می‌توان نتیجه گرفت مدل ترنسفورمر برای کاربردهای هشدار کوتاه‌مدت مناسب‌تر است. این برتری به‌احتمال زیاد ناشی از سازوکار توجه در ترنسفورمر است که روابط پیچیده بین دبی، بارش و تبخیر را بهتر می‌آموزد.

داده‌های ارزیابی (سال ۱۳۹۸) دارای شرایطی کاملاً متفاوت و بحرانی نسبت به داده‌های آموزشی هستند. دلیل انتخاب این سال، وقوع سیلاب گسترده و غیرمعمول در منطقه بود. در این سال، دبی رودخانه به‌طور قابل توجهی افزایش یافت که در طول دوره آموزش مدل، مشابه آن دیده نشده بود. اگرچه در سال ۱۳۹۱ نیز سیلابی رخ داده‌بود، شدت آن از نظر دبی به مراتب پایین‌تر از وضعیت سال ۱۳۹۸ بوده است. بنابراین مدل با داده‌ای مواجه شد که از منظر آماری و الگوهای زمانی، برایش ناشناخته بود و از نظر علمی می‌توان آن را وضعیت بحرانی دانست. مواجهه با چنین شرایط بحرانی باعث افزایش خطا در برخی نقاط اوج شد. از نظر تفسیر خطاها، می‌توان گفت اگرچه مدل در برخی موارد مقادیر اوج دبی را کمتر از مقدار واقعی تخمین زده است، اما در بازسازی روند وقوع سیلاب و افزایش ناگهانی جریان آب موفق بوده‌است. این موضوع از منظر کاربردی بسیار اهمیت دارد چرا که هدف نهایی در چنین مدل‌هایی، آگاهی‌بخشی پیش‌گیرانه و هشدار نسبت به وقوع سیلاب است، نه الزاماً تخمین دقیق عددی

دبی. در نهایت می‌توان نتیجه گرفت که مدل پیشنهادی، با وجود آموزش بر روی داده‌های فاقد شرایط شدید بحرانی، توانست در مواجهه با شرایط واقعی و دشوار سال ۱۳۹۸، رفتار رودخانه را با دقت قابل قبول پیش‌بینی نماید.

یافته‌های ما در بخش پیش‌بینی دبی با نتایج مطالعات مشابه مطابقت دارد. در سال‌های اخیر، کاربرد شبکه‌های عصبی عمیق به‌ویژه مدل‌های ترنسفورمر در پیش‌بینی سیلاب افزایش یافته است. برای نمونه لیو و همکاران یک مدل مبتنی بر ترنسفورمر برای پیش‌بینی ماهانه دبی رودخانه یانگ‌تسه توسعه دادند که توانست نقش داده‌های تاریخی جریان آب و الگوهای النینو نوسان جنوبی (ENSO)[۲۸]را به‌خوبی مدل‌سازی کند (Liu et al., 2022). کاستانگیا و همکاران نیز با ساخت یک مدل ترنسفورمر تغییرات روزانه سطح آب رودخانه در شبکه‌های آبی را پیش‌بینی و آن را با داده‌های سیلاب سال ۲۰۱۴ اروپا ارزیابی کردند (Castangia et al., 2023). ضمن اینکه برخی گزارش‌ها حاکی از برتری مدل ترنسفورمر نسبت به LSTM در پیش‌بینی‌های هیدرولوژیکی هستند (Li et al., 2019; Yin et al., 2022; Yin et al., 2023). نتایج حاصل از مدل پیشنهادی این پژوهش که بر پایه ترنسفورمر طراحی شده، با عملکرد گزارش‌شده در مطالعات مشابه سازگاری دارد و نشان می‌دهد که این دسته از مدل‌ها حتی در مواجهه با داده‌های بحرانی و غیرمعمول نیز توانایی بالایی در بازسازی روندهای دبی و پیش‌بینی وقوع سیلاب دارند. توان مدل در شناسایی روند افزایشی جریان و ارائه هشدارهای مؤثر، حاکی از ظرفیت بالای آن برای استفاده در سامانه‌های هشدار زودهنگام و برنامه‌ریزی منابع آب در شرایط ناپایدار اقلیمی است.

نتیجه‌گیری

در این پژوهش، عملکرد دو مدل یادگیری عمیق ترنسفورمر و LSTM برای پیش‌بینی کوتاه‌مدت دبی رودخانه در سه افق زمانی بررسی شد (تقسیم‌بندی حدوداً ۱۰ سال برای آموزش، ۱ سال برای آزمایش و ۱ سال برای ارزیابی شامل رخداد سیلاب). در پیش‌بینی یک‌روزه، مدل ترنسفورمر در مقدار NSE حدود ۷ درصد نسبت به LSTM عملکرد بهتری از خود نشان داد. در پیش‌بینی دوروزه، LSTM تقریبا ۲ درصد بهتر بود و در پیش‌بینی سه‌روزه، ترنسفورمر حدودا ۱ درصد بهتر بود. به‌طور کلی، ترنسفورمر در پیش‌بینی کوتاه‌مدت (یک‌روزه) برتری واضحی نشان می‌دهد. با افزایش افق زمانی، دقت هر دو مدل کاهش یافته و تفاوت عملکرد آن‌ها به‌طور قابل توجهی کمتر می‌شود. این الگو بیانگر طبیعت پیچیده و غیرخطی سامانه هیدرولوژیکی و حساسیت پیش‌بینی به افزایش افق زمانی است.

روش مبتنی بر ترنسفورمر می‌تواند یکی از ابزارهای مؤثر برای پیش‌بینی کوتاه‌مدت دبی و تصمیم‌گیری در سامانه‌های هشدار سریع باشد. با این حال، بهبود عملکرد در افق‌های بلندتر و افزایش مقاومت به داده‌های ناهنجار نیازمند کارهای تکمیلی است. پیشنهاد می‌شود در مطالعات آتی از مجموعه داده‌های طولانی‌تر، ویژگی‌های اقلیمی تکمیلی و مدل‌های هیبریدی برای بهبود عملکرد در افق‌های بلندتر بهره گرفته شود.

مشارکت نویسندگان

مشارکت نویسندگان در مقاله مستخرج از پایان‌نامه تقریباً به شکل زیر باشد:

نویسنده اول: تهیه و آماده‌سازی نمونه‌ها، انجام آزمایش و گردآوری داده‌ها، انجام محاسبات، تجزیه و تحلیل آماری داده‌ها، تحلیل و تفسیر اطلاعات و نتایج، تهیه پیشنویس مقاله

نویسنده دوم: استاد راهنمای پایان‌نامه، طراحی پژوهش، نظارت بر مراحل انجام پژوهش، بررسی و کنترل نتایج، اصلاح، بازبینی و نهایی‌سازی مقاله

نویسنده سوم: استاد مشاور پایان‌نامه، مشارکت در طراحی پژوهش، نظارت بر پژوهش، مطالعه و بازبینی مقاله

حامی مالی

مقاله حاضر با حمایت مالی معاونت پژوهشی دانشگاه صنعتی خواجه نصیرالدین طوسی انجام شد.

28. El Niño–Southern Oscillation

References

Atashi, V., & Gorji, H. T. (2025). Enhanced flood prediction using LSTM and climate parameters: multi-station analysis of snowmelt-induced flooding in the Red River of the North. *Journal of Hydroinformatics*, *27*(2), 245-260. https://doi.org/https://doi.org/۱۰/۲۱۶۶/hydro.2025.236

Castangia, M., Grajales, L. M. M., Aliberti, A., Rossi, C., Macii, A., Macii, E., & Patti, E. (2023). Transformer neural networks for interpretable flood forecasting. *Environmental Modelling & Software*, *160*, 105581. https://doi.org/https://doi.org/10.1016/j.envsoft.2022.105581

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250. https://doi.org/https://doi.org/10.5194/gmd-7-1247-2014

Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences*, *10*(17), 5776. https://doi.org/https://doi.org/10.3390/app10175776

Cisterna-García, A., González-Vidal, A., Martínez-Ibarra, A., Ye, Y., Guillén-Teruel, A., Bernal-Escobedo, L., & Skarmeta, A. F. (2025). Artificial intelligence for streamflow prediction in river basins: a use case in Mar Menor. *Scientific Reports*, *15*(1), 19481. https://doi.org/https://doi.org/10.1038/s41598-025-04524-0

Demiray, B. Z., & Demir, I. (2024). Towards Generalized Hydrological Forecasting using Transformer Models for 120-Hour Streamflow Prediction. *arXiv preprint arXiv:2406.07484* . https://doi.org/https://doi.org/10.48550/arXiv.2406.07484

Eltahir, E. A., & Wang, G. (1999). Nilometers, El Nino, and climate variability. *Geophysical Research Letters*, *26*(4), 489-492. https://doi.org/https://doi.org/10.1029/1999GL900013

Hochreiter, S. (1997). Long Short-term Memory. *Neural Computation MIT-Press*. https://doi.org/https://doi.org/10.1162/neco.1997.9.8.1735

Islam, K. I., Elias, E., Carroll, K. C., & Brown, C. (2023). Exploring random forest machine learning and remote sensing data for streamflow prediction: An alternative approach to a process-based hydrologic modeling in a snowmelt-driven watershed. *Remote Sensing*, *15*(16), 3999. https://doi.org/https://doi.org/10.3390/rs15163999

Jesus, G., Mardani, Z., Alves, E., & Oliveira, A. (2025). Deep Learning-Based River Flow Forecasting with MLPs: Comparative Exploratory Analysis Applied to the Tejo and the Mondego Rivers. *Sensors*, *25*(7), 2154. https://doi.org/https://doi.org/10.3390/s25072154

Kilinc, H. C., & Haznedar, B. (2022). A Hybrid Model for Streamflow Forecasting in the Basin of Euphrates. *Water*, *14*(1), 80. https://doi.org/https://doi.org/10.3390/w14010080

Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323-4331. https://doi.org/https://doi.org/10.5194/hess-23-4323-2019

Leščešen, I., Tanhapour, M., Pekárová, P., Miklánek, P., & Bajtek, Z. (2025). Long Short-Term Memory (LSTM) Networks for Accurate River Flow Forecasting: A Case Study on the Morava River Basin (Serbia). *Water*, *17*(6), 907. https://doi.org/https://doi.org/10.3390/w17060907

Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y.-X., & Yan, X. (2019). Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in neural information processing systems*, *32* .

Liu, C., Liu, D., & Mu, L. (2022). Improved transformer model for enhanced monthly streamflow predictions of the Yangtze River. *IEEE access*, *10*, 58240-58253. https://doi.org/https://doi.org/10.1109/ACCESS.2022.3178521

Mahmood, O. A., Sulaiman, S. O., & Al-Jumeily, D. (2024). Forecasting for Haditha reservoir inflow in the West of Iraq using Support Vector Machine (SVM). *PLoS One*, *19*(9 ,(e0308266. https://doi.org/https://doi.org/10.1371/journal.pone.0308266

Patel, S. S., & Ramachandran, P. (2015). A comparison of machine learning techniques for modeling river flow time series: the case of upper Cauvery river basin. *Water resources management*, *29*(2), 589-602. https://doi.org/https://link.springer.com/article/10.1007/s11269-014-0705-0

Shamshirband, S., Rabczuk, T., & Chau, K.-W. (2019). A survey of deep learning techniques: application in wind and solar energy resources. *IEEE access*, *7*-۱۶۴۶۵۰ ,164666- 164650 . https://doi.org/https://doi.org/10.1109/ACCESS.2019.2951750

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*. https://doi.org/https://doi.org/10.48550/arXiv.1706.03762

Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A., & Sayl, K. N. (2015). Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology* ,*530* ,829-844 . https://doi.org/https://doi.org/10.1016/j.jhydrol.2015.10.038

Yin, H., Guo, Z., Zhang, X., Chen, J., & Zhang, Y. (2022). RR-Former: Rainfall-runoff modeling based on Transformer. *Journal of Hydrology*, *609*, 127781. https://doi.org/https://doi.org/10.1016/j.jhydrol.2022.127781

Yin, H., Zhu, W., Zhang, X., Xing, Y., Xia, R., Liu, J., & Zhang, Y. (2023). Runoff predictions in new-gauged basins using two transformer-based models. *Journal of Hydrology*, *622*, 129684. https://doi.org/https://doi.org/10.1016/j.jhydrol.2023.129684

Yu, P.-S., Yang, T.-C., Chen, S.-Y., Kuo, C.-M., & Tseng, H.-W. (2017). Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *Journal of Hydrology*, *552*, 92-104. https://doi.org/https://doi.org/10.1016/j.jhydrol.2017.06.020