

Bài báo khoa học

So sánh đánh giá hiệu suất dự toán của ba mô hình CNN, LSTM và CNN-LSTM dự đoán lưu lượng dòng chảy bằng thuật toán học máy

Nguyễn Thị Thùy Linh^{1*}, Nguyễn Văn An²

¹ Trường Đại học Sư Phạm, Đại học Đà Nẵng; linhpknguyen@gmail.com

² Trường Đại học Sư Phạm, Đại học Đà Nẵng; nvan@ued.udn.vn

*Tác giả liên hệ: linhpknguyen@gmail.com; Tel: +84-372251803

Ban Biên tập nhận bài: 25/9/2024; Ngày phản biện xong: 6/11/2024; Ngày đăng bài: 25/3/2025

Tóm tắt: Trước sự gia tăng các tác động của biến đổi khí hậu, dự đoán lưu lượng dòng chảy là công cụ thiết yếu trong quản lý tài nguyên nước và ứng phó với thiên tai. Việc dự báo chính xác dòng chảy là một vấn đề rất phức tạp thu hút sự quan tâm của nhiều nhà nghiên cứu trong và ngoài nước. Trong nghiên cứu này, mạng nơ ron tích chập (*Convolutional Neural Network - CNN*) được kết hợp mạng nơ ron trí nhớ dài ngắn hạn (*Long Short Term Memory - LSTM*) tạo một mô hình mới là CNN-LSTM dùng để dự đoán lưu lượng dòng chảy. Với CNN trích dẫn đặc điểm thời gian và LSTM dự đoán lưu lượng. Mục tiêu chính của bài báo này là so sánh hiệu suất dự đoán của ba mô hình: CNN, LSTM và CNN-LSTM nhằm xác định mô hình nào có khả năng dự đoán lưu lượng dòng chảy tốt nhất. Kết quả thử nghiệm mô hình, CNN-LSTM có giá trị R^2 ($R^2_{CNN} = 0,950$, $R^2_{LSTM} = 0,956$, $R^2_{CNN-LSTM} = 0,960$) và NSE ($NSE_{CNN} = 0,948$, $NSE_{LSTM} = 0,953$, $NSE_{CNN-LSTM} = 0,958$) cao nhất cho thấy mô hình này dự đoán dòng chảy với độ chính xác cao hơn hai mô hình còn lại. Với sai số RMSE thấp nhất ($RMSE_{CNN} = 422,375$, $RMSE_{LSTM} = 402,139$, $RMSE_{CNN-LSTM} = 379,384$) mô hình CNN-LSTM vượt trội hơn tất cả mô hình AI thông thường. Do đó CNN-LSTM có giá trị thực tế lớn trong dự báo lưu lượng dòng chảy.

Từ khóa: Dự đoán dòng chảy; CNN, LSTM, CNN-LSTM.

1. Đặt vấn đề

Dự báo dòng chảy đóng vai trò quan trọng trong việc quản lý tối ưu và vận hành hiệu quả tài nguyên nước. Do đó, chủ đề này đã nhận được sự chú ý từ nhiều nhà nghiên cứu, thúc đẩy sự phát triển của hàng loạt mô hình dự báo trong những thập kỷ qua. Trong số các mô hình dự báo, các kỹ thuật dựa trên mô hình thống kê và mô hình dựa trên dữ liệu đang trở nên phổ biến do tính đơn giản và độ tin cậy cao của chúng [1].

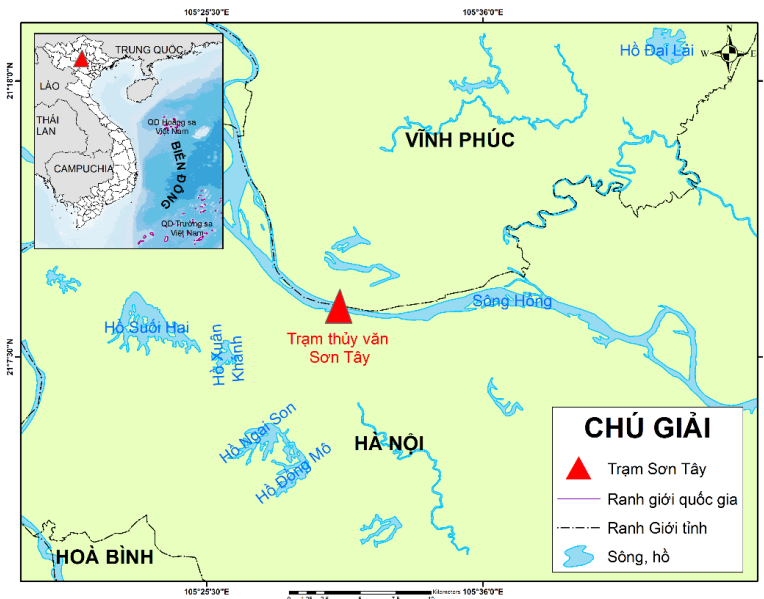
Các mô hình dựa trên dữ liệu này có thể được phân loại thành mô hình chuỗi thời gian và mô hình trí tuệ nhân tạo (AI). Nhiều nhà nghiên cứu đã triển khai các mô hình chuỗi thời gian trong việc dự báo lưu lượng dòng chảy, bao gồm mô hình tự hồi quy (AR), trung bình trượt (MA), mô hình trung bình trượt tự hồi quy (ARMA), và mô hình trung bình trượt tích hợp tự hồi quy (ARIMA) [2]. Tuy nhiên, do các giả thuyết tuyến tính của các mô hình này, chúng không phù hợp để dự báo dòng chảy có đặc tính phi tuyến tính và không cố định. Do đó, các mô hình trí tuệ nhân tạo (AI) với khả năng lập bản đồ phi tuyến tính đã được ứng dụng vào dự báo dòng chảy, bao gồm máy vectơ hỗ trợ hồi quy (SVR) [3], hệ thống suy luận mờ (FIS) [4], hồi quy Bayesian (BR) [5] và mạng lưới thần kinh nhân tạo (ANN) [6]. Tuy nhiên, phần lớn các mô hình này không thể biểu diễn đầy đủ thông tin theo bản chất phi tuyến tính và không cố định của dòng chảy [7]. Các mô hình học sâu, như mạng niềm tin sâu (DBN)

và mạng thần kinh tái phát (RNN), có thể khắc phục nhược điểm này nhờ khả năng học sâu hơn. Tuy nhiên, các mô hình học sâu này hoàn toàn dựa vào dữ liệu được quan sát trong quá khứ. Do đó có thể xảy ra trường hợp thông tin dòng chảy từ thời điểm trong quá khứ có thể không còn phù hợp so với thời điểm hiện tại. Do đó, việc sử dụng mô hình học sâu có khả năng tự động “ghi nhớ” hoặc “quên” thông tin trước đó sẽ có thể nâng cao độ chính xác trong dự báo dòng chảy [8]. Long Short Term Memory (LSTM) một trong những mô hình học sâu, có khả năng giải quyết nhiệm vụ này. LSTM đã được triển khai hiệu quả trong một số lĩnh vực, chẳng hạn như chẩn đoán tai nạn [9], dự đoán giá điện [10], dự báo độ sâu mực nước ngầm [11] và nhiều lĩnh vực khác. Trong nghiên cứu này, các tác giả đã sử dụng mô hình học sâu dựa trên sự tích hợp của mạng nơ-ron tích chập (Convolutional Neural Network - CNN) và mạng nơ-ron trí nhớ dài ngắn hạn (Long Short Term Memory - LSTM) để dự đoán lưu lượng dòng chảy. Trong đó, mô hình CNN được áp dụng để trích xuất các đặc điểm nội tại của chuỗi thời gian dòng chảy, mô hình LSTM sử dụng các đặc điểm được trích xuất bởi CNN để dự đoán dòng chảy. Mục đích áp dụng mô hình CNN-LSTM để dự đoán dòng chảy là nhằm tận dụng khả năng xử lý phi tuyến tính của CNN, từ đó đạt được độ chính xác cao trong dự báo dòng chảy ngắn hạn. Hơn nữa, trong mô hình CNN-LSTM, CNN được sử dụng để loại bỏ nhiễu và xem xét mối tương quan giữa các biến có độ trễ của dòng chảy. Mô hình LSTM sau đó xử lý thông tin thời gian và ánh xạ chuỗi thời gian vào các không gian có thể tách rời để đưa ra dự đoán chính xác hơn. Bên cạnh mô hình CNN-LSTM, nghiên cứu này còn tiến hành so sánh hiệu suất của mô hình với các mô hình trí tuệ nhân tạo khác để kiểm tra khả năng dự đoán lưu lượng dòng chảy của các mô hình học sâu và AI truyền thống.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Dữ liệu nghiên cứu

Trong nghiên cứu này nhóm nghiên cứu đã sử dụng dữ liệu thủy văn và khí tượng tại trạm Sơn Tây, Hà Nội để thử nghiệm mô hình dự đoán lưu lượng dòng chảy. Trong đó, dữ liệu lưu lượng được thu thập liên tục hàng ngày từ năm 1961 đến năm 2022. Ngoài ra, nghiên cứu này còn khai thác dữ liệu về nhiệt độ, lượng mưa được thu thập tại cùng trạm khí tượng nhằm mục đích cải thiện độ chính xác kết quả dự báo. Nội dung chi tiết dữ liệu nghiên cứu được sử dụng được thể hiện Bảng 1. Lý do chọn trạm thủy văn Sơn Tây để thử nghiệm mô hình vì trạm thủy văn Sơn Tây là trạm nằm ở vị trí giao nhau các nhánh của sông Hồng. Dự báo được lưu lượng dòng chảy tại đây sẽ giúp cảnh báo lũ và quản lý tài nguyên nước cho khu vực hạ lưu. Sơ đồ nghiên cứu và vị trí trạm thủy văn được thể hiện ở Hình 1.



Hình 1. Sơ đồ khu vực nghiên cứu và vị trí trạm thủy văn Sơn Tây, Hà Nội.

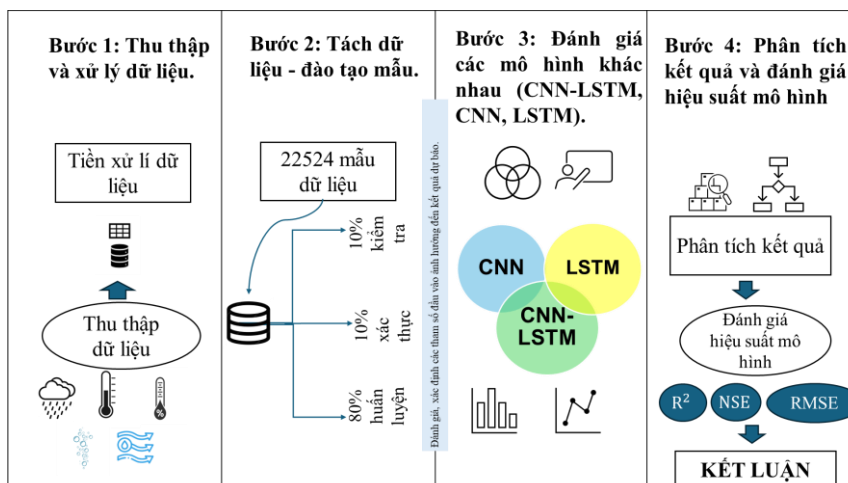
Bảng 1. Dữ liệu khí tượng thủy văn tại Trạm Sơn Tây, Hà Nội.

Loại số liệu	Số liệu	Thời gian
Thủy văn	Lưu lượng	01/01/1961 -31/12/2022
Khí tượng	Nhiệt độ	01/01/1961 -31/12/2022

Loại số liệu	Số liệu	Thời gian
Khí tượng	Lượng mưa	01/01/1961 -31/12/2022
Khí tượng	Độ ẩm	01/01/1961 -31/12/2022
Khí tượng	Bốc hơi	01/01/1961 -31/12/2022

2.2. Quy trình nghiên cứu

Hình 2 trình bày quy trình nghiên cứu mô hình dự báo lưu lượng dòng chảy tại trạm thủy văn Sơn Tây, Hà Nội.



Hình 2. Quy trình nghiên cứu mô hình dự báo lưu lượng dòng chảy tại trạm thủy văn Sơn Tây, Hà Nội.

Đầu tiên, Bước 1 dữ liệu cho nghiên cứu thu thập trạm thủy văn Sơn Tây, Hà Nội. Dữ liệu thực nghiệm bao gồm dữ liệu về lưu lượng dòng chảy và các yếu tố khí tượng thủy văn như nhiệt độ, độ ẩm, lượng mưa, bốc hơi trong quá khứ. Có tổng cộng 22524 mẫu dữ liệu được thu thập từ 01/01/1961 -31/12/2022. Cần chuyển đổi dữ liệu dạng thô sang dạng dữ liệu phù hợp cho mô hình, để làm được điều đó cần một chuỗi chức năng được sử dụng. Sau đó, chia tập dữ liệu thành tập con để huấn luyện, xác thực và kiểm tra mô hình. Đánh giá, xác định các tham số đầu vào ảnh hưởng đến kết quả dự báo Bước 2. Mô hình sẽ liên tục học hỏi từ bộ dữ liệu huấn luyện trước đó, từ đó dần dần cải thiện khả năng mô hình trong việc dự đoán. Khi hoàn tất việc huấn luyện, mô hình sẽ được xác thực và kiểm tra. Để đánh giá hiệu quả hoạt động của mô hình CNN-LSTM, các mô hình CNN và LSTM độc lập được chọn để so sánh hiệu quả Bước 3. Cuối cùng Bước 4 phân tích kết quả các mô hình trong việc dự đoán lưu lượng dòng chảy. Ba sai số được dùng làm tiêu chí đánh giá: Hệ số xác định (R^2), hệ số Nash (NSE) và sai số bình phương trung bình gốc (RMSE). Và cuối cùng đưa ra kết luận về hiệu quả mô hình.

2.3. Phương pháp nghiên cứu

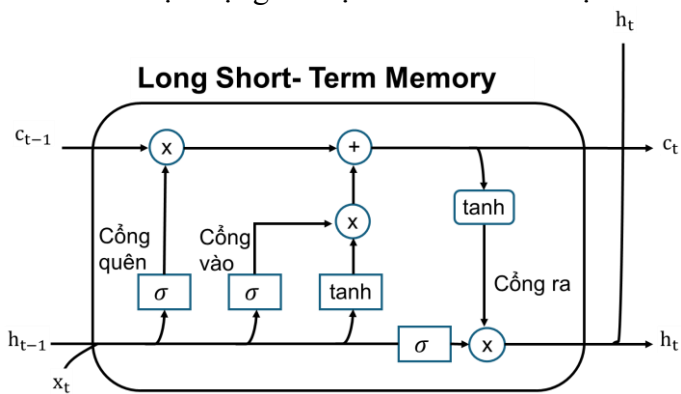
2.3.1. Mạng nơ ron tích chập (Convolutional Neural Network - CNN)

CNN là một mô hình mạng được đề xuất bởi Lecun và các cộng sự vào năm 1998 [12]. Đây là một loại ngôn ngữ mạng nơ-ron chuyển tiếp, có khả năng cao trong việc xử lý hình ảnh và ngôn ngữ tự nhiên. Nó có thể được sử dụng một cách hiệu quả để dự đoán chuỗi thời gian [13]. CNN chủ yếu gồm 2 phần: lớp tích chập và lớp gộp. Trong nghiên cứu này, loại mạng CNN một chiều (Conv1D) được sử dụng. Trong Conv1D các hạt nhân tích chập đi theo một hướng. Dữ liệu đầu ra của Conv1D là hai chiều [14]. Conv1D có khả năng trích xuất tính năng mạnh mẽ, đặc biệt là trong các ứng dụng liên quan đến dữ liệu tuần tự hoặc chuỗi thời gian [15]. Lớp chập của CNN gồm nhiều hạt nhân tích chập và công thức tính toán [16]. Đa số các hạt nhân tích chập và công thức tính toán thể hiện trong công thức. Shu và cộng sự (2021) đã khám phá tính ứng dụng mạng nơ-ron tích chập - CNN để dự báo lưu lượng dòng

chảy hằng thẳng của một con sông, so sánh hiệu suất của CNN với mạng nơ-ron nhân tạo và học máy cực đại. Kết quả cho thấy CNN có hiệu suất vượt trội hơn [17]. Nghiên cứu [17] đã đề xuất mạng nơ-ron tích chập (CNN) để dự báo lũ và hạn hán ở cả vùng khô cằn và nhiệt đới. Kết quả cho thấy CNN nổi bật hơn so với các mô hình khác trong dự báo lũ và có thể xử lý tốt hơn nhiều đặc điểm trong đầu vào. Tuy nhiên, CNN yêu cầu khối lượng dữ liệu lớn để thực hiện trong quá trình huấn luyện và gặp khó khăn trong việc xử lý các mối quan hệ dài hạn trong dữ liệu.

2.3.2. Trí nhớ ngắn hạn dài (Long Short Term Memory - LSTM)

Mạng nơ-ron tái phát - RNN là loại mạng thần kinh nhân tạo mạnh mẽ và sử dụng dữ liệu chuỗi thời gian quá khứ và hiện tại để dự đoán dữ liệu trong tương lai ở một khoảng thời gian xác định. Tuy nhiên, một nhược điểm của RNN là chỉ ghi nhớ những thông tin gần đây mà không thể nhớ lại thông tin từ xa hơn [15], khắc phục những nhược điểm đó mô hình LSTM ra đời. Đây là một biến thể của mạng nơ-ron tái phát RNN, được áp dụng để giải quyết các bài toán có sự phụ thuộc dài hạn (Long-term dependency). Mô hình LSTM lần đầu tiên được giới thiệu vào năm 1997, sau đó mô hình này đã trải qua nhiều lần cải tiến và phổ biến rộng rãi [13]. Mạng nơ-ron LSTM có cấu trúc dạng chuỗi, gồm nhiều mô-đun được lặp lại, có 4 lớp tương tác với nhau theo một cách đặc biệt. Cách thức hoạt động của LSTM là ghi nhớ những thông tin quan trọng liên quan đến việc dự đoán và loại bỏ những dữ liệu không cần thiết [16]. Mỗi mô-đun LSTM bao gồm có trạng thái tế bào (cell state) và các cổng (gate). Các cổng đều có nhiệm vụ sàng lọc thông tin với những mục đích khác nhau (Cổng quên, cổng đầu vào và cổng đầu ra) [13]. Cổng quên (Forget gate) có chức năng loại bỏ những thông tin không cần thiết khỏi trạng thái tế bào bên trong. Cổng đầu vào (Input gate) giúp sàng lọc những thông tin cần thiết để thêm vào trạng thái tế bào. Cổng đầu ra (Output gate) xác định những thông tin nào từ các trạng thái tế bào bên trong sẽ được sử dụng làm đầu ra [18]. Các cổng giúp cập nhật và kiểm soát luồng thông tin qua các khối bộ nhớ [15]. Sơ đồ cấu trúc LSTM chi tiết được thể hiện trong Hình 3. Nghiên cứu [19] đã áp dụng thành công mô hình LSTM để mô tả đặc tính lượng mưa và dòng chảy trong lưu vực lớn và phức tạp. LSTM thể hiện hiệu suất dự đoán vượt trội hơn các mô hình truyền thống, tuy nhiên, bên cạnh đó một trong những hạn chế của mô hình này là nó cần một lượng dữ liệu lớn để có thể hiệu chỉnh chính xác. Năm 2018, nghiên cứu [20] đã sử dụng mô hình ANN và LSTM để mô phỏng lượng mưa và dòng chảy, mô hình LSTM được chứng minh có hiệu suất tốt hơn trong dự đoán chuỗi thời gian thủy văn. So với các nghiên cứu trước đây trong các mô hình dự đoán lượng mưa-dòng chảy, kết quả mô hình LSTM có giá trị R² và NSE cao hơn. Tuy nhiên, cần nhiều nghiên cứu hơn trong việc mô hình hóa dự đoán thủy văn bằng cách sử dụng học máy sâu.

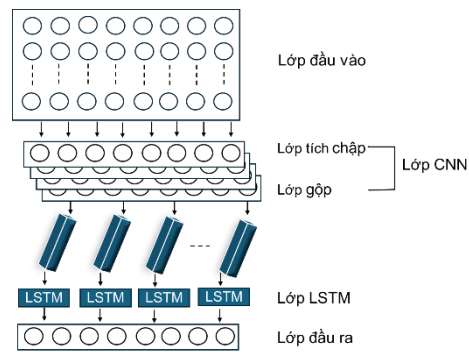


Hình 3. Sơ đồ cấu trúc LSTM [21].

2.3.3. Mạng nơ-ron CNN-LSTM

CNN có khả năng tập trung vào các đặc trưng nổi bật được áp dụng trong xử lý hình ảnh như nhận dạng đối tượng, phân loại hình ảnh và phát hiện gương mặt, vì vậy nó thường được sử dụng trong các kỹ thuật nhận diện và trích xuất đặc trưng [13]. LSTM có đặc điểm mở rộng theo chuỗi thời gian nên được sử dụng rộng rãi trong các bài toán dự đoán chuỗi thời gian. Dựa trên các đặc điểm trên của CNN và LSTM, một mô hình dự đoán lưu lượng dòng chảy được ra đời là mô hình CNN-LSTM. Sơ đồ cấu trúc mô hình được trình bày ở Hình 4.

Cấu trúc chính bao gồm các lớp CNN và LSTM: lớp đầu vào, lớp tích chập, lớp gộp, lớp LSTM và lớp kết nối [22]. CNN là một trong những cấu trúc mạng tiêu biểu nhất trong Deep Learning. So với các mô hình Deep Learning khác, CNN có khả năng khai thác thông tin và tính năng mạnh mẽ hơn [23]. Kết hợp CNN và LSTM có thể trích xuất hiệu quả đặc điểm hơn đặc điểm không gian thời gian từ dữ liệu đầu vào [23]. Mô hình CNN-LSTM kết hợp cho ra kết quả dự đoán có độ chính xác cao hơn so với việc sử dụng riêng biệt mô hình CNN và LSTM độc lập. Nghiên cứu này đã sử dụng bộ dữ liệu khí tượng thủy văn tại trạm thủy văn Sơn Tây để chứng minh rằng mô hình kết hợp CNN-LSTM là hệ thống dự báo dòng chảy hiệu quả và có khả năng áp dụng thực tiễn trong việc dự đoán lưu lượng dòng chảy.



Hình 4. Sơ đồ cấu trúc CNN-LSTM [24].

2.4. Các chỉ số đánh giá độ chính xác

Hiện nay nhiều nghiên cứu sử dụng ba trị số chính để đánh giá độ chính xác của mô hình là sai số bình phương trung bình gốc (*Root Mean Square Error - RMSE*) và hệ số Nash (*Nash Sutcliffe Efficiency - NSE*) và hệ số xác định (R^2) [25, 26]. Các trị số này thường được áp dụng để so sánh giữa các giá trị thực tế đo được và các giá trị dự báo trong các mô hình dự báo thủy văn.

Hệ số xác định R^2 cho phép đo lường mức độ chắc chắn của mối quan hệ giữa biến phản ứng và biến dự báo trong mô hình [27]. Nó khác với chỉ số Nash và RMSE ở đặc điểm thang đo kết quả không phụ thuộc vào thang đo đầu vào [28]. Hệ số xác định R^2 mô tả lượng phương sai quan sát được giải thích bởi mô hình [20]. Công thức tính được trình bày tại (1). Trong hầu hết các trường hợp, giá trị R^2 nằm trong khoảng từ [0,1]. Giá trị bằng 0 biểu thị hiệu suất thấp nhất chỉ ra không có sự tương quan và giá trị càng gần 1 có nghĩa là dự đoán càng chính xác [29].

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

Hệ số Nash (*Nash Sutcliffe Efficiency - NSE*) là một thước đo phổ biến và đáng tin cậy được sử dụng trong đánh giá hiệu suất chất lượng mô hình thủy văn [29]. NSE đo lường khả năng của mô hình trong việc dự đoán các giá trị thực tế so với giá trị trung bình được quan sát đồng thời cho biết tỷ lệ phương sai ban đầu được mô hình giải thích [20]. Mô hình dự báo có độ chính xác cao khi có hệ số Nash gần bằng 1 [30]. Phương trình toán học của NSE như sau (2):

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

RMSE được sử dụng như một chỉ số thống kê tiêu chuẩn để đánh giá hiệu suất mô hình trong các nghiên cứu liên quan đến khí tượng, chất lượng không khí và khí hậu [31]. Giá trị RMSE thường được áp dụng để đo lường độ lệch chuẩn giữa các giá trị dự đoán và giá trị quan sát thực tế. Các trị số được tính toán theo công thức dưới đây (3).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

Trong đó y_i là giá trị thực tế biến mục tiêu, \hat{y}_i là giá trị dự đoán biến mục tiêu, \bar{y}_i là giá trị trung bình của tất cả các biến thực tế, y_i' giá trị tính toán tương ứng thứ i ; n là số lần phát báo. Mẫu số trong công thức là biến thiên của các giá trị thực tế.

2.5. Thiết kế mô hình dự báo lưu lượng dòng chảy

Nghiên cứu đã sử dụng 3 mô hình CNN, LSTM, CNN-LSTM huấn luyện với các bộ dữ liệu khí tượng thủy văn thu thập tại trạm thủy văn Sơn Tây trong hơn 60 năm. Tập dữ liệu phục vụ cho mô hình được chia thành 3 phần chính với các mục đích khác nhau: 80% cho huấn luyện mô hình (chuỗi dữ liệu này được đo đạc từ 01/01/1961 - 20/08/2010), 10% cho xác thực (từ ngày 21/08/2010- 30/10/2016) và 10% dùng để đánh giá kiểm tra hiệu suất mô hình sau khi huấn luyện xong (từ ngày 31/10/2016- 31/12/2022). Quá trình huấn luyện thử nghiệm nhiều lần đã giúp tạo ra mô hình tốt nhất dự báo lưu lượng dòng chảy. Các thông số của mô hình và cấu trúc mô hình được điều chỉnh để tối ưu hiệu suất, và kết quả tóm tắt các mô hình được trình bày ở Bảng 2.

Bảng 2. Cấu trúc và thông số các mô hình.

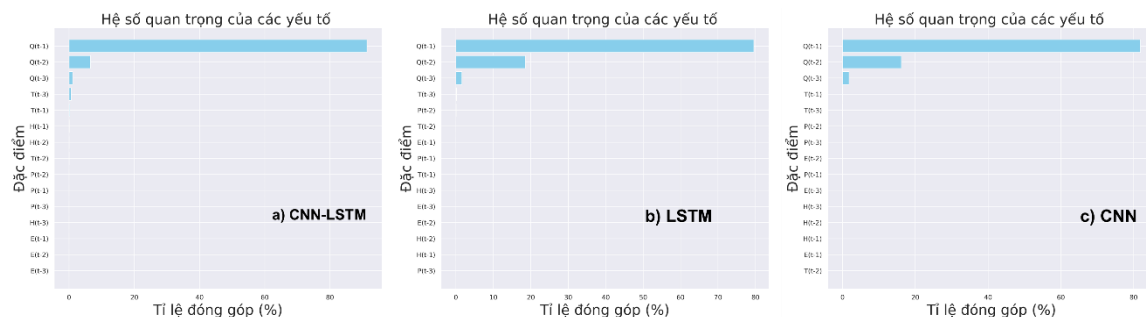
Thành phần	Mô hình CNN	Mô hình LSTM	Mô hình CNN-LSTM
Số lớp ẩn	4 lớp	2 lớp	8 lớp
Số lớp Conv1D	1	-	2
Số lớp Flatten	1	-	1
Số lớp Dense	2	1	2
MaxPooling1D	-	-	1
RepeatVector	-	-	1
Time Distributed	-	-	2
Kích hoạt lớp đầu ra	Sigmoid	Sigmoid	Sigmoid
Số tham số đã được đào tạo	3.649 (14,25KB)	20.545 (80,25KB)	54.921 (214,54KB)

Mỗi mô hình có cấu trúc khác nhau về số lớp ẩn, số lượng tham số đã được đào tạo. CNN có lớp tích chập Conv1D để trích xuất đặc trưng, trong khi LSTM có các lớp LSTM để xử lý chuỗi dữ liệu. CNN-LSTM kết hợp cả 2 loại nên số lớp ẩn nhiều hơn. Các tham số đã được đào tạo thể hiện khả năng học hỏi và độ phức tạp của mô hình.

3. Kết quả nghiên cứu và thảo luận

3.1. Kết quả hiệu chỉnh và kiểm định mô hình

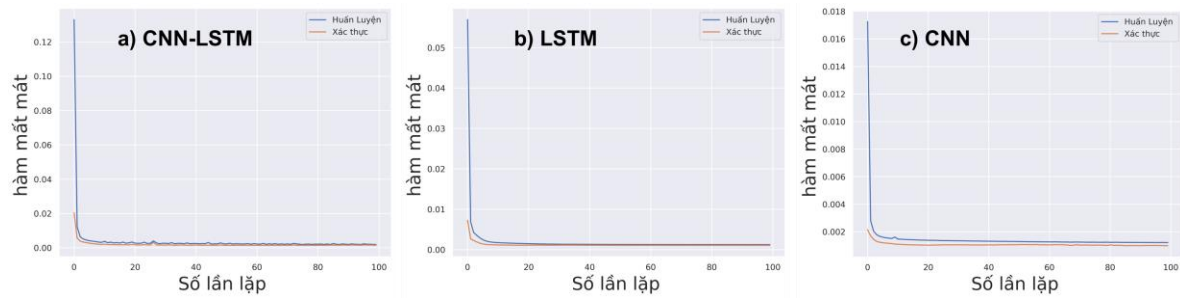
Kết quả hiệu chỉnh và kiểm định các mô hình CNN, LSTM và CNN-LSTM sử dụng số liệu thực đo tại trạm thủy văn Sơn Tây, Hà Nội. Hình 5 minh họa hệ số quan trọng của các đặc trưng đầu vào trong các mô hình dự báo dòng chảy. Các mô hình đều chỉ ra rằng đặc trưng lưu lượng tại thời điểm t-1 ($Q(t-1)$) là yếu tố quan trọng nhất, cho thấy tính nhất quán trong các kết quả. Tuy nhiên, sự khác biệt về tỉ lệ đóng góp giữa các mô hình cho thấy rằng cách tiếp cận mô hình hóa khác nhau có thể dẫn đến sự thay đổi trong mức độ quan trọng của



Hình 5. Hệ số quan trọng của các yếu tố khí tượng-thủy văn trong các mô hình dự báo.

các yếu tố khác. Các yếu tố được sắp xếp theo thứ tự giảm dần dự trên mức độ quan trọng của chúng. Đặc trưng lưu lượng tại thời điểm t-1 ($Q(t-1)$) có đóng góp lớn nhất, chiếm hơn 80% trong cả ba mô hình được thử nghiệm. Đặc trưng lưu lượng tại thời điểm t-2 ($Q(t-2)$) có

ảnh hưởng tương đối nhưng nhỏ hơn nhiều so với $Q(t-1)$. Các đặc trưng còn lại có đóng góp nhưng rất nhỏ không ảnh hưởng đáng kể đến hiệu suất của mô hình.

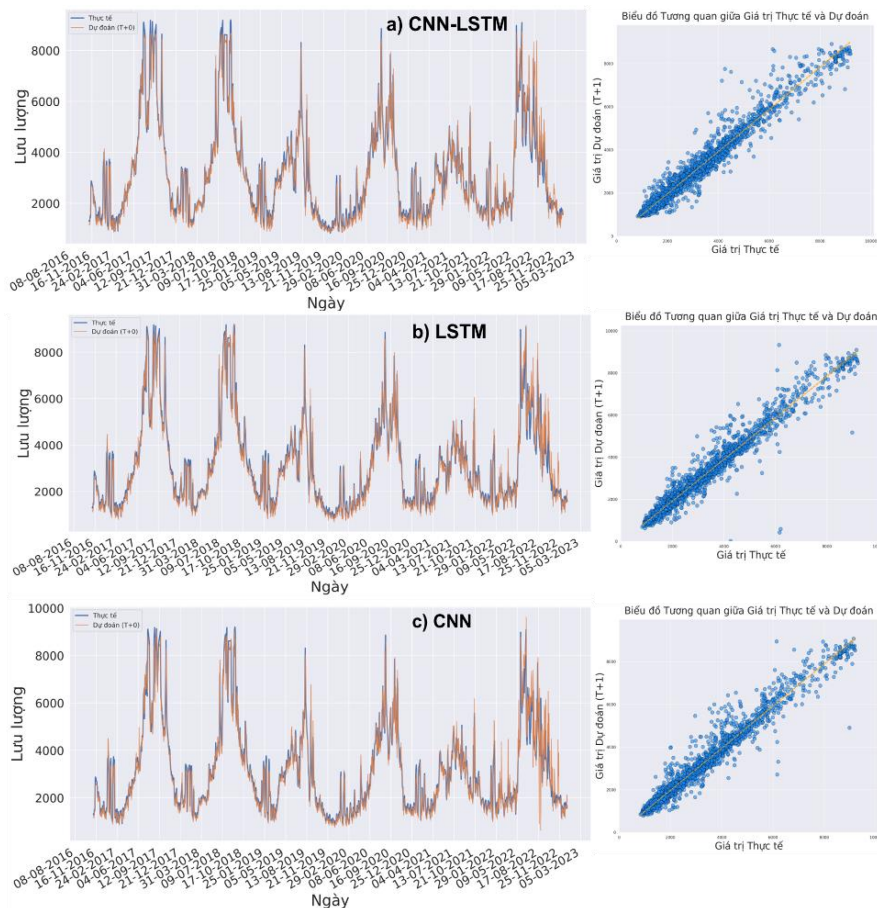


Hình 6. Đồ thị hàm tổn thất cho các mô hình.

Hình 6 hiển thị các đường cong của hàm mất mát các mô hình trong giai đoạn huấn luyện và xác thực thu được trong 100 lần lặp. Đường cong mất mát có xu hướng giảm dần theo số lần lặp. Hình 6a mô hình CNN-LSTM thể hiện hiệu suất tốt nhất trong dự đoán dòng chảy, với hàm tổn thất giảm nhanh và dao động ở mức thấp, cho thấy khả năng học và dự đoán chính xác trên cả tập dữ liệu huấn luyện và kiểm định. Trong khi đó, Hình 6b mô hình LSTM và Hình 6c mô hình CNN cũng đạt kết quả khá tốt, nhưng vẫn gặp khó khăn trong việc xử lý dữ liệu phức tạp.

Hình 7 hiển thị kết quả dự đoán lưu lượng dòng chảy được quan sát và mô phỏng từ mô hình CNN-LSTM (Hình 7a), LSTM (Hình 7b) và CNN (Hình 7c) trong khoảng thời gian từ 31/10/2016 - 31/12/2022 và các tiêu chí đánh giá hiệu suất các mô hình trong Bảng 3.

Các dòng mô phỏng Hình 7 cho thấy mô hình CNN-LSTM có hiệu suất tốt nhất trong số 3 mô hình được xem xét. Kết quả dự đoán dữ liệu dòng chảy mô hình CNN-LSTM (Hình 7a) với đường dự đoán gần như trùng khớp với dữ liệu thực tế, đặc biệt là ở các đỉnh lưu lượng. Quan sát trên biểu đồ tương quan Hình 7a các điểm trên biểu đồ tương quan của mô hình CNN-LSTM chủ yếu phân bố rất gần đường chéo (đường $y = x$), cho thấy sự dự báo của mô hình này gần với giá trị thực tế nhất. Độ phân tán của các điểm xung quanh đường $y = x$ là rất nhỏ thể hiện sự sai lệch của mô hình này là ít. Điều này đặc biệt rõ ràng ở các giá trị cao (các đỉnh lưu lượng lớn), khi mô hình dự báo khá chính xác. Điều này chứng minh được rằng CNN-LSTM có khả năng dự đoán lưu lượng dòng chảy với độ chính xác cao với



Hình 7. Biểu đồ thể hiện kết quả dự đoán lưu lượng dòng chảy của ba mô hình.

đỉnh lưu lượng. Quan sát trên biểu đồ tương quan Hình 7a các điểm trên biểu đồ tương quan của mô hình CNN-LSTM chủ yếu phân bố rất gần đường chéo (đường $y = x$), cho thấy sự dự báo của mô hình này gần với giá trị thực tế nhất. Độ phân tán của các điểm xung quanh đường $y = x$ là rất nhỏ thể hiện sự sai lệch của mô hình này là ít. Điều này đặc biệt rõ ràng ở các giá trị cao (các đỉnh lưu lượng lớn), khi mô hình dự báo khá chính xác. Điều này chứng minh được rằng CNN-LSTM có khả năng dự đoán lưu lượng dòng chảy với độ chính xác cao với

dữ liệu đầu vào phức tạp, có khả năng dự báo bất thường về lũ. CNN-LSTM thể hiện vượt trội trong việc dự báo các đỉnh dòng chảy, một yếu tố then chốt trong quản lý tài nguyên nước và cảnh báo lũ lụt. Mô hình LSTM với khả năng ghi nhớ các mối quan hệ dài hạn giúp ích rất nhiều cho mô hình dự đoán phi tuyến tính tuy nhiên LSTM vẫn có sai số trong dự đoán, hầu hết các dự đoán đỉnh chính lưu lượng dòng chảy đều có sự chênh lệch, nhất là khi dữ liệu có biến động lớn (Hình 7b) Các điểm trên biểu đồ tương quan của LSTM cũng phân bố khá gần với đường chéo, tuy nhiên có sự phân tán nhiều hơn so với mô hình CNN-LSTM, đặc biệt là ở các giá trị lưu lượng cực đoan (cao hoặc thấp). Độ phân tán của các điểm xung quanh đường $y = x$ lớn hơn so với CNN-LSTM. Mô hình CNN có thể dự đoán xu hướng chính lưu lượng nhưng có một số điểm giá trị dự đoán có sự chênh lệch lớn so với giá trị thực tế (Hình 7c). Biểu đồ tương quan của CNN cho thấy các điểm phân tán nhiều hơn so với cả hai mô hình trên, đặc biệt là ở các giá trị cực trị. Các điểm dữ liệu không tập trung sát đường $y = x$, độ phân tán của các điểm rất lớn, và điều này rõ ràng nhất ở các đỉnh và đáy của lưu lượng. Mô hình CNN dường như không thể học các chuỗi thời gian phức tạp mà chỉ dựa trên đặc trưng không gian, nên việc dự báo chuỗi thời gian của dòng chảy sẽ kém chính xác hơn. Mô hình CNN còn gặp khó khăn trong việc nắm bắt các mối quan hệ dài hạn trong dự đoán chuỗi thời gian.

3.2. Đánh giá kết quả dự báo

Nhằm kiểm tra hiệu suất của các mô hình CNN, LSTM và CNN-LSTM trong dự đoán lưu lượng dòng chảy. Sau khi huấn luyện các mô hình trên, dữ liệu thử nghiệm sẽ được đưa vào mô hình để dự đoán các biến mục tiêu. Các mô hình được đánh giá và so sánh bằng các chỉ số R^2 , NSE, RMSE. Kết quả được tóm tắt Bảng 3. Mô hình LSTM ($R^2 = 0,956$, $NSE = 0,953$, $RMSE = 402,139$) hoạt động tốt hơn mô hình CNN ($R^2 = 0,950$, $NSE = 0,948$, $RMSE = 422,375$). Mô hình LSTM hoạt động tốt hơn mô hình CNN trong dự đoán chuỗi thời gian nhưng mô hình kết hợp CNN-LSTM hoạt động tốt hơn các mô hình độc lập CNN và LSTM với giá trị chỉ số kiểm định đều cao hơn thể hiện: $R^2 = 0,960$, $NSE = 0,958$, $RMSE = 379,384$. Trong ba mô hình, giá trị RMSE của CNN-LSTM là thấp nhất cho thấy sự kết hợp CNN-LSTM đã cải thiện độ chính xác tổng thể, hoạt động tốt hơn các mô hình CNN và LSTM độc lập. Điều này là do LSTM có khả năng học sự phụ thuộc lâu dài và CNN có thể trích xuất các tính năng bất biến theo thời gian.

Bảng 3. Tiêu chí đánh giá mô hình.

Tiêu chí đánh giá	Mô hình CNN-LSTM	Mô hình LSTM	Mô hình CNN
R^2	0,960	0,956	0,950
NSE	0,958	0,953	0,948
RMSE	379,384	402,139	422,375

3.3. Thảo luận

Nghiên cứu này đã chứng minh rằng mô hình CNN-LSTM cho độ chính xác cao hơn các mô hình AI thông thường khi dự đoán lưu lượng dòng chảy. Với các chỉ số kiểm định vượt trội thể hiện khả năng mô hình dự đoán gần với thực tế. Có thể thấy được sự kết hợp giữa CNN và LSTM đã khai thác đối ưu điểm mạnh của từng mô hình riêng biệt. Kết quả này phù hợp với nghiên cứu [15] trong đó mô hình CNN-LSTM cũng được sử dụng để dự đoán dòng chảy và cho thấy hiệu suất vượt trội so với các mô hình AI thông thường như LSTM và các mạng nơ-ron sâu (DNN). Sự đồng nhất trong các kết quả này cho thấy rằng việc sử dụng mô hình CNN-LSTM là một công cụ mạnh mẽ và đáng tin cậy trong các ứng dụng dự báo chuỗi thời gian phức tạp.

4. Kết luận

Trong bài báo này, tác giả đề xuất mô hình kết hợp CNN-LSTM và ứng dụng trong dự báo lưu lượng dòng chảy với dữ liệu đầu vào là số liệu khí tượng thủy văn tại trạm thủy văn Sơn Tây, Hà Nội. Để chứng minh tính hiệu quả của mô hình CNN-LSTM, tác giả đã so sánh với các mô hình CNN và LSTM độc lập. Hiệu suất mô hình được đánh giá qua ba chỉ số R^2 , NSE, RMSE. Kết quả dự đoán các mô hình có thể được xếp hạng từ cao đến thấp như sau: CNN-LSTM > LSTM > CNN. Mô hình CNN-LSTM có khả năng dự đoán lưu lượng dòng chảy vượt trội về tính khái quát hóa và độ chính xác so với các mô hình khác, được thể hiện thông qua các tiêu chí đánh giá mô hình, CNN-LSTM có giá trị RMSE thấp nhất và giá trị R^2 , NSE cao nhất.

CNN-LSTM đã chứng minh sự vượt trội trong việc dự báo các đỉnh dòng chảy, một yếu tố then chốt trong quản lý tài nguyên nước và cảnh báo lũ lụt. Thông qua các việc đào tạo và kiểm định mô hình, các tác giả nhận thấy rằng CNN-LSTM có thể phân tích và kết hợp hiệu quả các dữ liệu đầu vào phức tạp bao gồm lưu lượng dòng chảy trong quá khứ và các yếu tố khí tượng như lượng mưa, nhiệt độ, độ ẩm. Nhờ vậy, mô hình CNN-LSTM không chỉ cải thiện độ chính xác quá trình dự báo mà vượt trội hơn hẳn các mô hình khác về mức độ phù hợp và sai số.

Tuy nhiên, để đảm bảo độ chính xác mô hình học sâu vào dự báo dòng chảy thì việc nghiên cứu bổ sung thêm các vấn đề như mô hình thiếu tính cập nhật trong các tình huống khẩn cấp; dữ liệu có thể bị ảnh hưởng bởi quá trình điều tiết dòng chảy hồ chứa và sông nhánh tại trạm thủy văn Sơn Tây là điều rất cần thiết.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: N.V.A.; Xử lý số liệu, chạy mô hình: N.V.A., N.T.T.L.; Viết bản thảo bài báo: N.T.T.L., N.V.A.

Lời cam đoan: Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Lima, A.R.; Cannon, A.J.; Hsieh, W.W. Forecasting daily streamflow using online sequential extreme learning machines. *J. Hydrol.* **2016**, *537*, 431–443.
2. Valipour, M.; Banihabib, M.E.; Behbahani, S.M.R. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J. Hydrol.* **2013**, *476*, 433–441.
3. Yu, X.; Zhang, X.; Qin, H. A data-driven model based on Fourier transform and support vector regression for monthly reservoir inflow forecasting. *J. Hydro-environ. Res.* **2018**, *18*, 12–24.
4. He, Z.; Wen, X.; Liu, H.; Du, J. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J. Hydrol.* **2014**, *509*, 379–386.
5. Humphrey, G.B.; Gibbs, M.S.; Dandy, G.C.; Maier, H.R. A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *J. Hydrol.* **2016**, *540*, 623–640.
6. Tan, Q.F.; Lei, X.H.; Wang, X.; Wang, H.; Wen, X.; Ji, Y.; Kang, A.Q. An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. *J. Hydrol.* **2018**, *567*, 767–780.
7. Bai, Y.; Chen, Z.; Xie, J.; Li, C. Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. *J. Hydrol.* **2016**, *532*, 193–206.
8. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*(8), 1735–1780.

9. Yang, J.; Kim, J. An accident diagnosis algorithm using long short-term memory. *Nucl. Eng. Technol.* **2018**, *50*(4), 582–588.
10. Peng, L.; Liu, S.; Liu, R.; Wang, L. Effective long short-term memory with differential evolution algorithm for electricity price prediction. *Energy* **2018**, *162*, 1301–1314.
11. Zhang, J.; Zhu, Y.; Zhang, X.; Ye, M.; Yang, J. Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929.
12. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE.* **1998**, *86*(11), 2278–2324.
13. Lu, W.; Li, J.; Li, Y.; Sun, A.; Wang, J. A CNN-LSTM-based model to forecast stock prices. *Complexity* **2020**, *2020*(1), 6622927.
14. Shen, X.; Ni, Z.; Liu, L.; Yang, J.; Ahmed, K. WiPass: 1D-CNN-based smartphone keystroke recognition using WiFi signals. *Pervasive Mob. Comput.* **2021**, *73*, 101393.
15. Ghimire, S.; Yaseen, Z.M.; Farooque, A.A.; Deo, R.C.; Zhang, J.; Tao, X. Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks. *Sci. Rep.* **2021**, *11*(1), 17497.
16. Sun, M.; Song, Z.; Jiang, X.; Pan, J.; Pang, Y. Learning pooling for convolutional neural network. *Neurocomputing* **2017**, *224*, 96–104.
17. Castangia, M.; et al. Transformer neural networks for interpretable flood forecasting. *Environ. Modell. Software.* **2023**, *160*, 105581.
18. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*(8), 5929–5955.
19. Kratzert, F.; et al. Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **2018**, *22*(11), 6005–6022.
20. Hu, C.; et al. Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water* **2018**, *10*(11), 1543.
21. Aldhyani, T.H.; Alzahrani, A. Framework for predicting and modeling stock market prices based on deep learning algorithms. *Electronics* **2022**, *11*(19), 3149.
22. Yao, R.; et al. Intrusion detection system in the advanced metering infrastructure: a cross-layer feature-fusion CNN-LSTM-based approach. *Sensors* **2021**, *21*(2), 626.
23. Guo, Q.; He, Z.; Wang, Z. Monthly climate prediction using deep convolutional neural network and long short-term memory. *Sci. Rep.* **2024**, *14*(1), 17748.
24. Ahmed, S.; et al. Forecasting the status of municipal waste in smart bins using deep learning. *Int. J. Environ. Res. Public Health.* **2022**, *19*(24), 16798.
25. Nohani, E.; et al. Monthly streamflow forecasting based on meteorological data from a nearby station. *Water Pract. Technol.* **2024**, *19*(5), 1659–1675.
26. Khosravi, K.; et al. Improving daily stochastic streamflow prediction: Comparison of novel hybrid data-mining algorithms. *Hydrol. Sci. J.* **2021**, *66*(9), 1457–1474.
27. Krause, P.; Boyle, D.; Bäse, F. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **2005**, *5*, 89–97.
28. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peer J. Comput. Sci.* **2021**, *7*, e623.
29. Althoff, D.; Rodrigues, L.N. Goodness-of-fit criteria for hydrological models: Model calibration and performance assessment. *J. Hydrol.* **2021**, *600*, 126674.
30. Xie, T.; et al. Hybrid forecasting model for non-stationary daily runoff series: a case study in the Han River Basin, China. *J. Hydrol.* **2019**, *577*, 123915.
31. Hodson, T.O. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev. Discuss.* **2022**, *2022*, 1–10.

Comparative evaluation of the prediction performance of three models CNN, LSTM and CNN-LSTM for predicting flow rate using machine learning algorithms

Nguyen Thi Thuy Linh^{1*}, Nguyen Van An²

¹ University of Education, University of Danang; linhpknguyen@gmail.com

² University of Education, University of Danang; nvan@ued.udn.vn

Abstract: Under the increasing impacts of climate change, streamflow prediction is an essential tool in water resource management and disaster response. Accurate streamflow forecasting is a highly complex problem that has garnered significant attention from researchers both domestically and internationally. In this study, a new model called CNN-LSTM is created by combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) networks for streamflow prediction. The CNN extracts temporal features, while the LSTM predicts the flow. The main objective of this paper is to compare the predictive performance of three models: CNN, LSTM, and CNN-LSTM, to determine which model can best predict streamflow. The model testing results show that CNN-LSTM, with the highest R^2 values ($R^2_{CNN} = 0.950$, $R^2_{LSTM} = 0.956$, $R^2_{CNN-LSTM} = 0.960$) and NSE values ($NSE_{CNN} = 0.948$, $NSE_{LSTM} = 0.953$, $NSE_{CNN-LSTM} = 0.958$), indicates higher predictive accuracy compared to the other two models. With the lowest RMSE ($RMSE_{CNN} = 422.375$, $RMSE_{LSTM} = 402.139$, $RMSE_{CNN-LSTM} = 379.384$), the CNN-LSTM model outperforms all conventional AI models. Therefore, CNN-LSTM has significant practical value in streamflow forecasting.

Keywords: Flow prediction; CNN, LSTM, CNN-LSTM.