

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/333995366>

Ứng dụng mạng Long Short-Term Memory (LSTM) để dự báo mực nước tại trạm Quang Phục và Cửa Cấm, Hải Phòng, Việt Nam

Article · September 2018

CITATION

1

READS

192

2 authors, including:



[Hung Ho Viet](#)

Thuy Loi University

20 PUBLICATIONS 7 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Hydraulics; Free surface flow; dam break problem [View project](#)

ỨNG DỤNG MẠNG LONG SHORT-TERM MEMORY (LSTM) ĐỂ DỰ BÁO MỨC NƯỚC TẠI TRẠM QUANG PHỤC VÀ CỬA CẤM, HẢI PHÒNG, VIỆT NAM

Lê Xuân Hiền^{1,2}; Hồ Việt Hùng¹

Tóm tắt: Trong bài báo này, mô hình Bộ nhớ gần xa (Long Short-Term Memory - LSTM) được sử dụng để dự báo mực nước sông mà không cần các số liệu địa hình và dự báo mưa. Dữ liệu cần thiết cho mô phỏng là mực nước theo giờ tại các trạm thủy văn ở thành phố Hải Phòng. Mô hình được thiết lập để dự báo mực nước cho trạm Quang Phục và trạm Cửa Cấm trước 5 giờ (dự báo từ 1 giờ đến 5 giờ). Mặc dù mô hình không yêu cầu các dữ liệu về khí hậu, địa hình nhưng kết quả dự báo có độ chính xác cao. Trong trường hợp dự báo mực nước trước 3 giờ, hệ số NSE (hệ số Nash) cho giá trị trên 97,8% và giá trị RMSE (sai số căn quân phương) nhỏ hơn 0,10 m cho cả 2 trạm. Kết quả này cho thấy rằng, mô hình LSTM mà các tác giả đề xuất dự báo chính xác mực nước theo thời gian thực, có thể áp dụng mô hình này để cảnh báo lũ trên các sông của Việt Nam.

Từ khóa: Hải Phòng, dự báo mực nước, LSTM, DNN, Quang Phục, Cửa Cấm.

1. MỞ ĐẦU

Diễn biến mực nước sông là một quá trình phức tạp, biến đổi theo không gian và thời gian. Việc dự báo chính xác mực nước là một trong những yêu cầu cấp bách nhằm giảm thiểu các rủi ro do lũ gây ra và có ý nghĩa quan trọng trong việc xây dựng phương án phòng, chống lũ. Các mô hình truyền thống được sử dụng để dự báo ở Việt Nam cũng như trên thế giới là các mô hình số về thủy lực và thủy văn. Các mô hình này yêu cầu một số lượng lớn các dữ liệu đầu vào như: đặc điểm lưu vực, địa hình, dự báo lượng mưa, quan hệ mưa – dòng chảy, quan hệ lưu lượng - mực nước theo thời gian tại một số vị trí.

Một trong những giải pháp hiệu quả cho việc dự báo là sử dụng mô hình Mạng thần kinh nhân tạo (Artificial Neural Network – ANN). Trên thế giới các mô hình ANN đã được sử dụng rộng rãi trong dự báo lũ từ những năm 1990 (Sung, J.Y. và các cộng sự, 2017). Cùng với đó, các nhà nghiên cứu đã áp dụng các thuật toán vào mô hình Mạng thần kinh để làm tăng độ chính xác

của kết quả dự báo. Trong những năm gần đây, phương pháp Học sâu (Deep learning) dựa trên nền tảng là các mạng thần kinh nhân tạo đang phát triển rất nhanh và thu hút sự quan tâm của nhiều nhà khoa học. Có thể kể đến nghiên cứu của Chen, J.F. và các cộng sự (2014), ở đây thuật toán Cuckoo Search đã được áp dụng để dự báo dòng chảy đến hồ Hòa Bình, Việt Nam; Nguyễn Thanh Tùng (2016) đã sử dụng phương pháp Random Forest cũng để dự báo dòng chảy đến hồ này; Trương Xuân Nam và các cộng sự (2016) đã sử dụng phương pháp Học sâu để dự báo lưu lượng nước đến hồ Hòa Bình. Các nghiên cứu này đều có một điểm chung, đó là sử dụng dữ liệu của mùa kiệt với bước thời gian quan trắc 10 ngày làm đầu vào và đưa ra dự báo dòng chảy cho 10 ngày sau đó. Các kết quả nghiên cứu đều khẳng định khả năng tiềm ẩn của mô hình mạng thần kinh. Mô hình ANN cũng được áp dụng để dự báo mực nước ở trạm Hirakata, Nhật Bản (Kim, S. và các cộng sự, 2017). Nghiên cứu này chỉ sử dụng dữ liệu mực nước thực đo theo giờ ở các trạm thượng lưu để dự báo mực nước ở hạ lưu trước 3, 6 và 9 giờ. Kết quả đạt được cho thấy mô hình mạng thần kinh có thể áp dụng cho việc cảnh báo lũ trên

¹ Khoa Kỹ thuật Tài nguyên nước, Trường Đại học Thủy lợi

² Dept. of Construction & Disaster Prevention Engineering, Kyungpook National University, Sangju, Korea

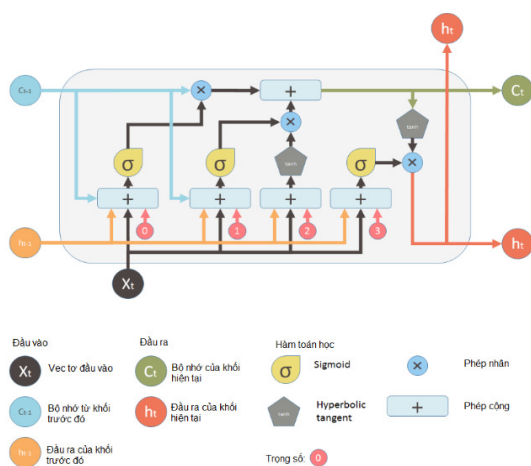
sông. Wang, Y. và các cộng sự (2017) đã sử dụng mạng thần kinh LSTM để dự báo chất lượng nước cho hồ Taihu, Trung Quốc. Nghiên cứu chỉ ra rằng mô hình LSTM đưa ra dự báo chính xác hơn so với các mô hình mạng thần kinh khác.

Trong bài báo này, các tác giả sử dụng mô hình LSTM để dự báo mực nước sông mà không cần các thông tin dự báo lượng mưa, cũng như số liệu địa hình và tình hình sử dụng đất. Mô hình này được áp dụng để dự báo mực nước trước 1h, 2h, 3h, 4h và 5h tại trạm Quang Phục trên sông Văn Úc và trạm Cửa Cấm trên sông Cấm.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1 Mô hình LSTM

Mô hình đề xuất dựa trên mô hình mạng thần kinh sâu LSTM, đây là một dạng đặc biệt của RNN (Recurrent Neural Network - Mạng thần kinh hồi quy). LSTM được giới thiệu bởi Hochreiter và Schmidhuber (1997) nhằm giải quyết các bài toán về phụ thuộc xa (long-term dependency).



Hình 1. Cấu trúc của mô hình LSTM
(Nguồn: Internet)

Theo Olah (2015), mọi mạng hồi quy đều có dạng là một chuỗi các mô đun lặp đi lặp lại của một mạng thần kinh, mỗi mô đun này thường có cấu trúc đơn giản được gọi là một tầng “tanh”. LSTM cũng có kiến trúc dạng chuỗi như vậy và

thay vì chỉ có 1 tầng mạng thần kinh như RNN chuẩn thì chúng có tới 4 tầng và tương tác với nhau một cách đặc biệt. Cấu trúc của mô hình mạng thần kinh LSTM được thể hiện ở Hình 1. Cốt lõi của LSTM bao gồm trạng thái tế bào (cell state) và cổng (gate). Trạng thái tế bào giống như băng chuyền, chạy xuyên suốt qua tất cả các nút mạng giúp thông tin được truyền đạt dễ dàng, còn cổng là nơi sàng lọc thông tin đi qua nó, chúng được kết hợp bởi một tầng mạng sigmoid. Một LSTM gồm có 3 cổng để duy trì hoạt động trạng thái của tế bào.

Bước đầu tiên của mô hình LSTM được gọi là tầng cổng quên (forget gate layer). Bước này sẽ quyết định xem thông tin nào cần bỏ đi từ trạng thái tế bào. Đầu vào cho bước này là h_{t-1} (giá trị đầu ra tại thời điểm $t-1$) và x_t (dữ liệu đầu vào); đầu ra f_t là một số trong khoảng từ 0 đến 1 cho mỗi số trong trạng thái tế bào C_{t-1} .

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Trong đó: σ là hàm sigmoid, W_f và b_f lần lượt là trọng số và tham số của tầng cổng quên.

Các bước tiếp theo sẽ quyết định thông tin lưu vào trạng thái tế bào và cập nhật giá trị cho trạng thái. Bao gồm một tầng sigmoid hay còn được gọi là cổng vào (input gate layer, i_t) và một véc tơ giá trị được tạo từ tầng tanh.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * C_t \quad (4)$$

Trong đó: C_{t-1} và C_t là trạng thái tế bào lần lượt ở thời điểm $t-1$ và t ; W_C và b_C lần lượt là trọng số và tham số của trạng thái tế bào.

Ở bước cuối cùng, giá trị đầu ra (h_t) sẽ được quyết định bởi trạng thái của tế bào muốn xuất ra (output gate, o_t).

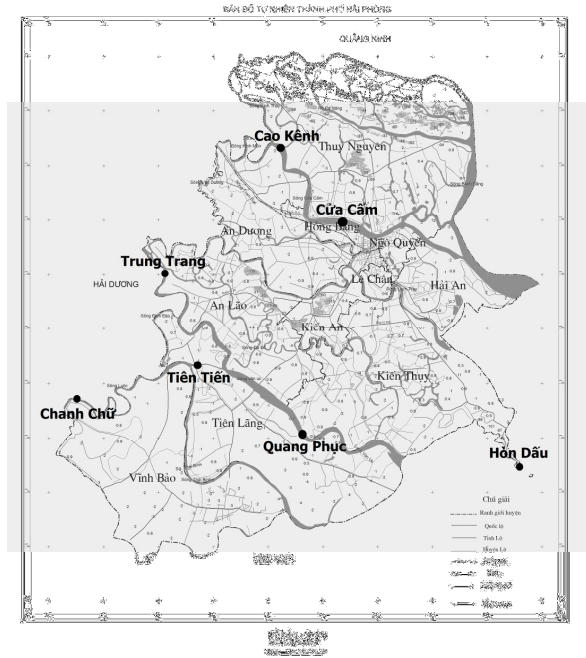
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

2.2 Thu thập dữ liệu về khu vực nghiên cứu

Dữ liệu được thu thập bao gồm: điều kiện tự nhiên, đặc điểm khí tượng, thủy văn, hải văn. Các số liệu lượng mưa và mực nước theo giờ tại các trạm thủy văn có trong 19 ngày, bắt đầu từ 0h ngày 14/7/2011. Khu vực nghiên cứu gồm

sông Văn Úc và sông Cấm thuộc địa phận thành phố Hải Phòng (Hình 2). Đây là khu vực chịu ảnh hưởng của thủy triều. Các dữ liệu phục vụ cho bài báo này được thể hiện trong Bảng 1.



Hình 2. Hệ thống sông khu vực TP. Hải Phòng
(Nguồn: Viện Kỹ thuật tài nguyên nước, 2011)

Bảng 1. Thống kê các số liệu đã thu thập

TT	Trạm	Sông	Yếu tố đo
1	Chanh Chữ	Luộc	H, X
2	Tiên Tiến	Mới	H
3	Trung Trang	Văn Úc	H
4	Quang Phục*	Văn Úc	H
5	Cao Kênh	Kinh Thầy	H, X
6	Cửa Cấm*	Cấm	H, X
7	Hòn Dấu	Biển Đông	H

(Nguồn: Viện Kỹ thuật tài nguyên nước, 2011)

Trong Bảng 1: H là mực nước; X là lượng mưa; * là các trạm cần dự báo mực nước.

Đoạn sông Văn Úc chảy qua Hải Phòng từ ngã ba Gù ra đến biển dài 45 km. Đây là sông sâu và rộng nhất trong số các sông ở hạ du sông Thái Bình, với chiều rộng trung bình từ 500 đến 800m. Dưới ngã ba Gù khoảng 1 km,

sông Văn Úc phân ra một nhánh chính là sông Lạch Tray đổ ra cửa Lạch Tray. Sông Cấm là ranh giới giữa hệ thống An Hải và Thủy Nguyên, toàn bộ sông Cấm thuộc địa phận Hải Phòng. Đây chính là nhánh của sông Kinh Môn, có chiều dài 23 km, bắt đầu từ ngã ba Hợp Thành đến nhập lưu vào sông Bạch Đằng để đổ ra biển qua cửa Nam Triệu. Sông Cấm có chiều rộng khoảng 200-700m.

2.3 Phương pháp đánh giá

Để đánh giá hiệu quả của mô hình dự báo, chúng tôi sử dụng hai trị số, đó là RMSE (Root Mean Squared Error – sai số căn quân phương) và NSE (Nash Sutcliffe Efficiency – hệ số Nash)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (8)$$

Trong đó: O_i , \bar{O} và P_i lần lượt là giá trị thực đo, giá trị thực đo trung bình và giá trị dự báo của mẫu thứ i tương ứng. Mô hình dự báo cho kết quả tốt nếu RMSE nhỏ và NSE lớn.

3. THIẾT LẬP THÔNG SỐ MÔ HÌNH

Mô hình LSTM được đề xuất để dự báo mực nước trong nhiều trường hợp, từ 1 giờ đến 5 giờ tại trạm Quang Phục (sông Văn Úc) và trạm Cửa Cấm (sông Cấm). Mỗi mô hình LSTM được hiệu chỉnh (training) và kiểm định (test) để dự báo mực nước lần lượt cho từng trạm. Các thông số của mô hình được hiệu chỉnh và kiểm định để đảm bảo mô hình cho kết quả tốt nhất, chi tiết về các thông số này được tóm tắt trong Bảng 2. Sau quá trình hiệu chỉnh và kiểm định, các thông số tốt nhất của mô hình cho mỗi trường hợp đã được lựa chọn để phục vụ việc dự báo. Dữ liệu đầu vào của mô hình LSTM dựa trên mực nước thực đo tại các trạm thủy văn trong 3 giờ gần nhất ($t-2$, $t-1$, $t-0$), hoặc trong 6 giờ gần nhất (từ $t-5$ đến $t-0$).

Bảng 2. Các thông số của mô hình LSTM

Đặc trưng	Chi tiết
Mục tiêu dự báo	Mức nước tại trạm Quang Phục và Cửa Cấm trước 1h, 2h, 3h, 4h, 5h
Dữ liệu đầu vào	Lượng mưa tại khu vực nghiên cứu. Mức nước thực đo tại các trạm thủy văn trong 3 giờ: t-2, t-1, t-0. Mức nước thực đo tại các trạm thủy văn trong 6 giờ: từ t-5 đến t-0
Cấu trúc mô hình	TensorFlow với BasicLSTMCell
Thông số hiệu chỉnh mô hình	Số lượng lớp ẩn: 10, 20; 50 Hệ số học: 0,1; 0,5; 0,01; 0,05; 0,001; 0,005 Số lượng Epoch: 10.000; 20.000; 50.000

Để dự báo mực nước cho trạm **Quang Phục**, mực nước thực đo tại các trạm Trung Trang, Chanh Chũ, Tiên Tiến, Hồn Dầu và **Quang Phục** đã được sử dụng. Tương tự như vậy, mực nước thực đo tại các trạm Cao Kênh, Hồn Dầu và **Cửa Cấm** được dùng để dự báo mực nước cho trạm **Cửa Cấm**. Các dữ liệu về lượng mưa tại 2 khu vực này đã được đưa vào để kiểm định, tuy nhiên việc đưa thêm các số liệu này không làm các kết quả dự báo tốt hơn. Việc này có thể giải thích như sau: diện tích mặt sông nhỏ và lượng mưa không lớn, ngoài ra còn có tác động của bốc hơi và thấm nên ảnh hưởng của mưa là không đáng kể; mặt khác, lượng mưa ở khu giữa trên mỗi đoạn sông còn được thể hiện trong chính mực nước thực đo của trạm cần dự báo. Vì vậy trong bài báo này, kết quả dự báo chỉ phụ thuộc vào dữ liệu mực nước thực đo tại các trạm. Trong mô hình thủy lực sông, các mực nước này chính là các biên của mô hình.

Để hiệu chỉnh và kiểm định mô hình, dữ liệu được sử dụng bao gồm 456 bản ghi là số liệu mực nước theo giờ, từ 0h ngày 14/7/2011 đến 23h ngày 01/8/2011, thời điểm này đang là mùa lũ ở Hải Phòng. Bộ dữ liệu này được chia thành 2 tập: tập dữ liệu hiệu chỉnh gồm 408 bản ghi để hiệu chỉnh mô hình nhằm chọn các thông số tốt nhất; tập dữ liệu kiểm định gồm 48 bản ghi để đánh giá hiệu quả của mô hình với các thông số đã chọn. Các dữ liệu đã thu thập cho thấy, mực

nước lớn nhất ở cả 2 trạm Quang Phục (2,26m) và Cửa Cấm (2,09m) đều rơi vào ngày 30/7/2011. Các giá trị này nằm trong tập dữ liệu hiệu chỉnh nhằm đảm bảo đưa ra kết quả chính xác hơn cho quá trình dự báo đỉnh lũ.

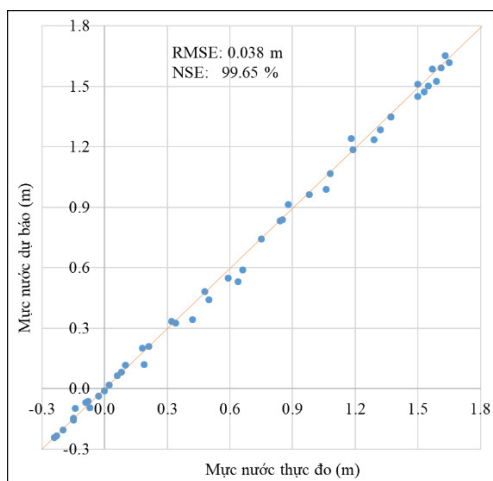
Một điểm lưu ý nữa là, không có quy tắc nào trong việc lựa chọn cấu trúc mô hình cũng như các thông số mô hình (Kim, S., và các cộng sự, 2017). Việc lựa chọn cấu trúc mô hình và thông số mô hình dựa trên việc đánh giá về kích cỡ dữ liệu và quá trình thử để đảm bảo chọn được các thông số phù hợp nhất cho nghiên cứu này.

4. KẾT QUẢ NGHIÊN CỨU

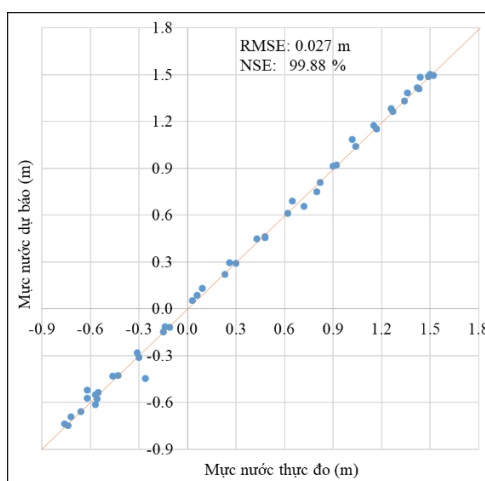
4.1 Dự báo mực nước trạm Quang Phục

Kết quả dự báo mực nước cho trạm Quang Phục trong các trường hợp từ 1 giờ đến 5 giờ được thể hiện tương ứng trong các Hình 3(a), Hình 4 và Hình 5.

Có thể thấy rằng mô hình dự báo cho kết quả rất ấn tượng, đặc biệt cho các trường hợp dự báo từ 1-3 giờ với sai số trung bình nhỏ hơn 0,095m và hệ số NSE trên 97,8%. Hình 3 (a) mô tả sự tương quan chặt chẽ giữa kết quả dự báo và giá trị thực đo trong trường hợp dự báo mực nước (MN) trạm Quang Phục trước 1 giờ, hệ số Nash lên tới 99,7% và RMSE chỉ 0,038m. Trong trường hợp dự báo trước 4 giờ và 5 giờ (Hình 5), mặc dù hệ số NSE tương đối tốt (lần lượt là 94% và 92%) nhưng giữa kết quả dự báo và thực đo có chênh lệch nhỏ (1 giờ) về thời gian xuất hiện đỉnh lũ.

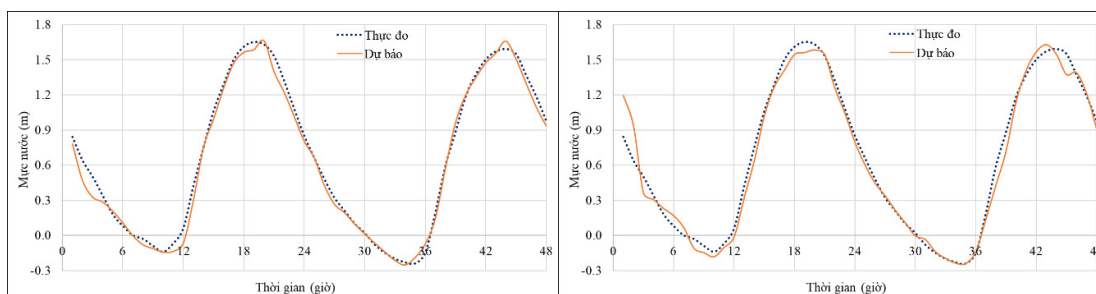


(a)

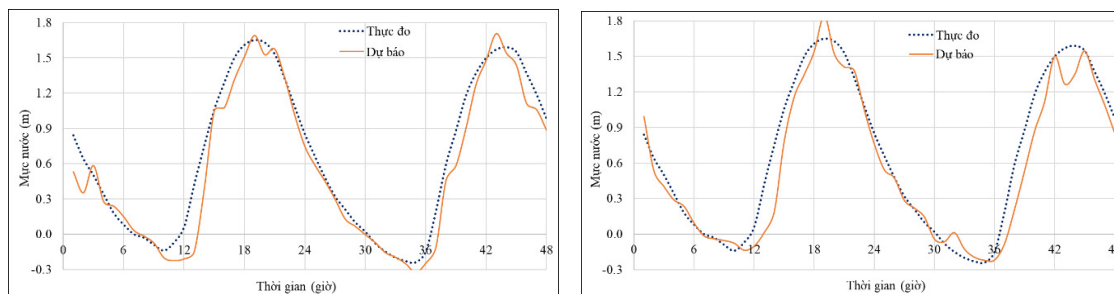


(b)

Hình 3. So sánh MN thực đo với dự báo 1 giờ tại Quang Phục (a) và Cửa Cấm (b)



Hình 4. Dự báo MN 2 giờ (trái) và 3 giờ (phải) tại Quang Phục



Hình 5. Dự báo MN 4 giờ (trái) và 5 giờ (phải) tại Quang Phục

Bảng 3 tổng hợp các kết quả kiểm định tốt để dự báo mực nước trạm Quang Phục. nhất và các thông số mô hình đã được lựa chọn

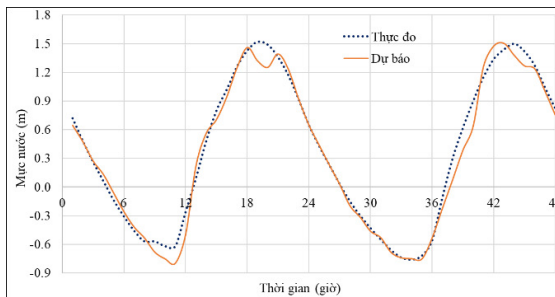
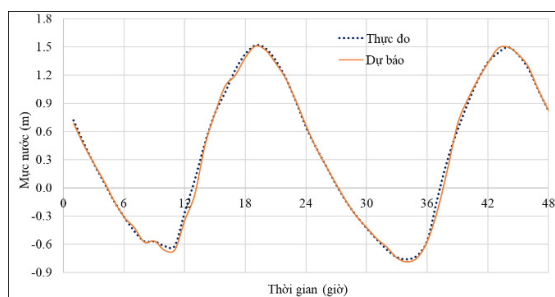
Bảng 3. Kết quả kiểm định dự báo mực nước cho trạm Quang Phục

Thời gian dự báo(giờ)	Chiều dài dữ liệu	Số trạm đầu vào	Số lượng lớp ẩn	Hệ số học	Số lượng Epoch	RMSE (m)	NSE (%)
1	6	5	50	0.01	50000	0.038	99.7%
2	3	5	50	0.001	20000	0.064	99.0%
3	6	5	50	0.1	20000	0.095	97.8%
4	6	5	50	0.1	20000	0.155	94.1%
5	6	5	50	0.1	20000	0.177	92.4%

4.2 Dự báo mực nước trạm Cửa Cấm

Đối với trạm Cửa Cấm, kết quả kiểm định tốt nhất của các trường hợp được thể hiện trong Bảng 4 và được mô tả chi tiết ở Hình 3(b), Hình

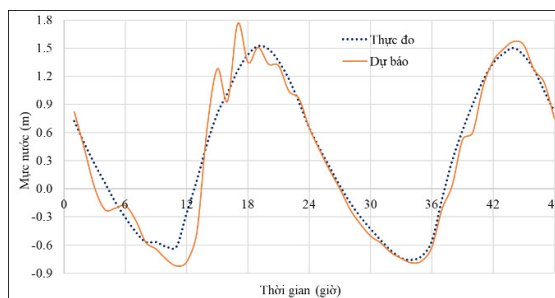
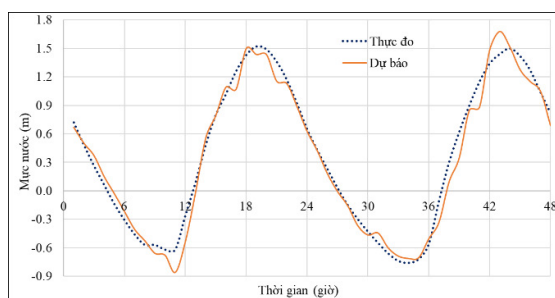
6 và Hình 7. Trong cả 5 trường hợp, kết quả dự báo mực nước trạm Cửa Cấm tốt hơn một chút so với kết quả của trạm Quang Phục.



Hình 6. Dự báo MN 2 giờ (trái) và 3 giờ (phải) tại Cửa Cấm

Hệ số NSE của trạm Cửa Cấm rất cao, giá trị NSE nhỏ nhất là 94,1% khi dự báo cho 5 giờ. Đặc biệt, với trường hợp dự báo từ 1 đến 3 giờ giá trị này là trên 98%, sai số RMSE tương ứng nhỏ hơn 0,107m. Hình 3(b) so sánh giữa kết quả

tính toán mực nước với thực đo khi dự báo 1 giờ cho thấy rằng, hầu như không có sự sai lệch đáng kể nào giữa giá trị dự báo và thực đo. Hệ số NSE lên tới 99,88% và sai số trung bình chỉ xấp xỉ 2,7cm.



Hình 7. Dự báo MN 4 giờ (trái) và 5 giờ (phải) tại Cửa Cấm

Trong trường hợp dự báo trước 4 giờ và 5 giờ, mặc dù kết quả dự báo có hệ số NSE rất cao (97,5% và 94,1%), nhưng Hình 7 cho thấy có sự sai lệch giữa giá trị dự báo và thực đo. Tuy nhiên, các giá trị dự báo có xu hướng cao hơn so

với giá trị thực đo tại đỉnh lũ, điều này cho thấy kết quả mô phỏng có xu hướng an toàn hơn. Khi dự báo thời gian dài, kết quả dự báo sẽ có độ chính xác cao nếu chuỗi số liệu thực đo đủ dài.

Bảng 4. Kết quả kiểm định dự báo mực nước cho trạm Cửa Cấm

Thời gian dự báo (giờ)	Chiều dài dữ liệu	Số trạm đầu vào	Số lượng lớp ẩn	Hệ số học	Số lượng Epoch	RMSE (m)	NSE (%)
1	3	3	20	0.001	20000	0.027	99.9%
2	3	3	50	0.1	50000	0.043	99.7%
3	3	3	20	0.05	20000	0.107	98.0%
4	3	3	20	0.1	20000	0.121	97.5%
5	3	3	20	0.005	20000	0.185	94.1%

Việc thay đổi các thông số mô hình sẽ ảnh hưởng tới kết quả dự báo, tuy nhiên ảnh hưởng này là không đáng kể. Việc thay đổi chiều dài dữ liệu đầu vào (tăng kích thước dữ liệu từ t-2 lên t-5) không ảnh hưởng nhiều tới độ chính xác của kết quả dự báo, thậm chí trong một số trường hợp còn giảm độ chính xác (trạm Cửa Cấm). Điều này có thể giải thích được, vì ảnh hưởng của sự thay đổi mực nước trong 3 bước thời gian gần nhất là rõ rệt nhất. Ngoài ra, do kích thước của tập dữ liệu chưa đủ lớn, nên khi tăng các giá trị như số lượng lớp ẩn, số lượng epoch hoặc chiều dài dữ liệu thì độ chính xác của kết quả dự báo không thay đổi nhiều.

5. KẾT LUẬN

Bài báo này đã mô tả chi tiết quá trình xây dựng mô hình mạng thần kinh LSTM để dự báo mực nước sông trước 1 giờ, 2 giờ, 3 giờ, 4 giờ và 5 giờ tại trạm Quang Phục và Cửa Cấm ở thành phố Hải Phòng. Mô hình mà các tác giả đề xuất không sử dụng các dữ liệu dự báo mưa, chỉ sử dụng thông tin về mực nước thực đo đã có tại các trạm thủy văn ở thượng lưu và hạ lưu để dự báo mực nước cho 1 trạm thủy văn ở trung lưu. Các số liệu mực nước có ý nghĩa như các biên của mô hình thủy lực. Đối với vùng chịu ảnh hưởng của thủy triều, dòng chảy

êm, nên cần cả biên trên và biên dưới khi tính toán. Kết quả kiểm định mô hình cho thấy sự ổn định và độ chính xác cao trong dự báo. Kết quả dự báo mực nước từ 1 đến 3 giờ có hệ số NSE trên 97,8% đối với trạm Quang Phục và trên 98% với trạm Cửa Cấm. Điều này thể hiện sự tương quan rất lớn giữa giá trị dự báo và giá trị thực đo.

Có thể thấy rằng, không có quy tắc cụ thể nào cho việc lựa chọn các thông số của mô hình như hệ số học, số lượng chuỗi, số lượng lớp ẩn, số lượng Epoch cũng như cấu trúc của mô hình. Việc lựa chọn các thông số này dựa vào quá trình thử và đánh giá sai số. Các thông số tốt nhất đã được chọn và đánh giá qua quá trình hiệu chỉnh và kiểm định mô hình.

Mạng thần kinh sâu có ưu điểm là đơn giản hơn so với các mô hình thủy văn, thủy lực, đặc biệt nó có thể dự báo cho vùng bị ảnh hưởng của thủy triều trong trường hợp thiếu số liệu địa hình, khi mà không thể sử dụng các mô hình thủy lực, thủy văn. Lúc đó, mô hình mạng thần kinh LSTM là giải pháp tốt cho việc dự báo mực nước theo thời gian thực, thay thế các mô hình đã có. Vì vậy, có thể ứng dụng mô hình này để cảnh báo lũ trên các sông suối của Việt Nam.

TÀI LIỆU THAM KHẢO

- Trương Xuân Nam và Nguyễn Thanh Tùng (2016). *Deep learning: Ứng dụng cho dự báo lưu lượng nước đến hồ chứa hòa bình*. Hội nghị khoa học thường niên, Đại học Thủy lợi.
- Viện Kỹ thuật tài nguyên nước (2011). *Quy hoạch thủy lợi chống ngập úng thành phố Hải Phòng*. Trường Đại học Thủy lợi.
- Chen, J.F., Hsieh, H.N., and Do, Q.H. (2014). *Forecasting Hoabinh Reservoir's Incoming Flow: An Application of Neural Networks with the Cuckoo Search Algorithm*. Information. 5, 570-586.
- Kim, S. and Tachikawa, Y. (2017). *Real-time river-stage prediction with artificial neural network based on only upstream observation data*. Annual Journal of Hydraulic Engineering, JSCE, Vol. 61.
- Nguyen, T.T. (2015). *An II-regression Random Forests Method For Forecasting of Hoa Binh Reservoir's Incoming Flow*. International Conference on Intelligent Systems and Knowledge Engineering.
- Olah, C. (2015). *Understanding LSTM Networks*. GITHUB blog, Retrieved from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Posted on August 27, 2015.

- Sung, J.Y., Lee, J., Chung, I.M., and Heo, J.H. (2017). *Hourly Water Level Forecasting at Tributary Affected by Main River Condition*. KSCE Journal of Civil Engineering. 9, 644.
- Wang, Y., Zhou, J., Chen. K., Wang, Y., and Liu. L. (2017). *Water Quality Prediction Method Based on LSTM Neural Network*. International Conference on Intelligent Systems and Knowledge Engineering.

Abstract:

**USING LONG SHORT-TERM MEMORY NEURAL NETWORK TO FORECAST
WATER LEVEL AT THE QUANG PHUC AND THE CUA CAM STATIONS
IN HAI PHONG, VIETNAM**

In this article, the LSTM (Long Short-Term Memory) model is applied to predict the river water level without utilization of rainfall – forecast information and terrain data. The data required for simulation are hourly water levels at hydrological stations in Hai Phong city. The model was formulated to predict water level at the Quang Phuc station and the Cua Cam station, in Hai Phong city for many cases from 1 to 5 hours of lead time. Although the model does not require many input data such as climate, geography, land-use for rainfall-runoff simulation, the prediction results are very stable and reliable: the Nash – Sutcliffe efficiency (NSE) is higher than 97.8% and the root mean square error (RMSE) is lower than 0.10m for 3 hours of lead time prediction. The result illustrated that the LSTM model is able to produce the river water level time series and useful for the practical flood forecasting.

Keywords: Hai Phong, water level prediction, LSTM, DNN, Quang Phuc, Cua Cam.

Ngày nhận bài: 21/3/2018

Ngày chấp nhận đăng: 29/7/2018