

Nghiên cứu ứng dụng mạng nơ-ron nhân tạo dự đoán độ cao sóng gần bờ tại khu vực biển Cửa Đại, tỉnh Quảng Ngãi

Vũ Văn Ngọc^{1*}, Tạ Đức Hải²

¹Phòng Thí nghiệm trọng điểm Quốc gia về Động lực học Sông Biển;

²Khoa Hàng không vũ trụ, Học viện Kỹ thuật Quân sự.

*Email: vuvannoc85@gmail.com

Nhận bài: 01/4/2023; Hoàn thiện: 15/5/2023; Chấp nhận đăng: 10/6/2023; Xuất bản: 25/6/2023.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.88.2023.147-153>

TÓM TẮT

Mạng nơ-ron nhân tạo (ANN) được coi là một công cụ mạnh để giải quyết các bài toán có tính phi tuyến, phức tạp, mối quan hệ giữa các thông số không dễ thiết lập một cách tường minh. Mạng nơ-ron nhân tạo được ứng dụng trong các lĩnh vực khoa học kỹ thuật khác nhưng chưa được ứng dụng nhiều đối với nghiên cứu thủy động lực vùng ven biển. Sử dụng ANN có ưu điểm đáng kể trong xử lý dữ liệu lớn, tốc độ giải nhanh với độ chính xác đảm bảo. Trong nghiên cứu này, nhóm tác giả sử dụng mạng nơ-ron nhân tạo để tính toán mô phỏng và dự đoán chiều cao sóng gần bờ thuộc khu vực biển Cửa Đại tỉnh Quảng Ngãi trong giai đoạn tháng 6/2016 và đầu tháng 7/2016. Tập dữ liệu huấn luyện mạng ANN bao gồm các dữ liệu đầu vào được lấy từ số liệu sóng ngoài khơi của NOAA, dữ liệu đầu ra được lấy từ kết quả tính của mô hình Mike 21 SW (Spectral Wave) tại biển Cửa Đại. Ngoài ra, nhằm tăng độ tin cậy khi xây dựng mô hình Mike 21 SW, kết quả đầu ra của mô hình Mike 21 SW sẽ được so sánh với kết quả đo thực tế tại biển Cửa Đại trong vòng 7 ngày đầu tháng 7/2016. Kết quả tính toán mô phỏng và dự đoán độ cao sóng gần bờ của mạng ANN được so sánh với kết quả của mô hình Mike 21 SW và kết quả đo thực tế.

Từ khóa: Mạng nơ-ron nhân tạo; Độ cao sóng có nghĩa; Dự đoán sóng; Mike 21SW.

1. MỞ ĐẦU

Trên thế giới hiện nay, việc áp dụng các giải thuật thông minh hay mô hình ANN trong các bài toán tính toán mô phỏng cũng như dự đoán đang dần trở nên phổ biến hơn. Một số nghiên cứu công bố gần đây trên thế giới và trong nước đã cho thấy sự phát triển mạnh mẽ của các ứng dụng ANN trong nghiên cứu sóng biển. Tác giả Scott C. James và đồng nghiệp đã mô phỏng sóng trên miền tính 2D tại khu vực biển hồ, vịnh Monterey bằng mô hình SWAN và mô hình học máy (ML), so sánh tương quan kết quả đạt được độ chính xác trên 90% [1]. Nhóm nghiên cứu của Jadran Berbic dự báo sóng trong vịnh kín Adriatic tại các điểm có dữ liệu quan trắc dài hạn theo chuỗi thời gian và kết quả cho độ chính xác đều trên 90% [2]. D.J. Peres và các cộng sự sử dụng ANN với dữ liệu đầu vào là tốc độ gió U, V tái phân tích của NCEP/NCAR ứng dụng cho vùng biển RON ở Italia, kết quả phân tích bằng ANN được so sánh với kết quả từ mô hình tiên tiến như WaveWatch III/CFSR và chỉ ra mô hình hoạt động tốt, có tính thực tiễn cao [3]. Bên cạnh đó, có các nghiên cứu khác như dự báo sóng phản xạ do các công trình trong bể cảng, dự báo độ đục của nước biển, dự báo sự lan truyền sóng gần bờ từ dữ liệu sóng toàn cầu [4-6]. Tại Việt Nam, một số nghiên cứu có sử dụng ANN trong lĩnh vực thủy văn và hải văn cũng đã được công bố [7-8]. Trong lĩnh vực động lực học sóng biển phải kể đến tác giả Đặng Văn Tô và Trần Hồng Thái. Tác giả Đặng Văn Tô và các đồng nghiệp phục hồi dữ liệu sóng biển bằng mạng ANN với chương trình OceanANN, tập số liệu sóng có sẵn được cố ý làm thất thoát và sau đó khôi phục lại, kết quả đạt được cho thấy số liệu khôi phục có độ chính xác đạt trên 98% [9]. Nhóm nghiên cứu của tác giả Trần Hồng Thái dự báo độ cao sóng tại trạm Cồn Cỏ (Quảng Trị, Việt Nam) được thực hiện bằng mô hình ANN hồi quy, theo các hạn dự báo 6, 12, 18 và 24 giờ. Trong đó, tác giả sử dụng mô hình đơn biến và mô hình hai biến, kết quả cho thấy độ tin cậy chấp nhận được của mô hình hai biến, hệ số tương quan đạt trên 58% [10]. Có thể thấy rằng, các

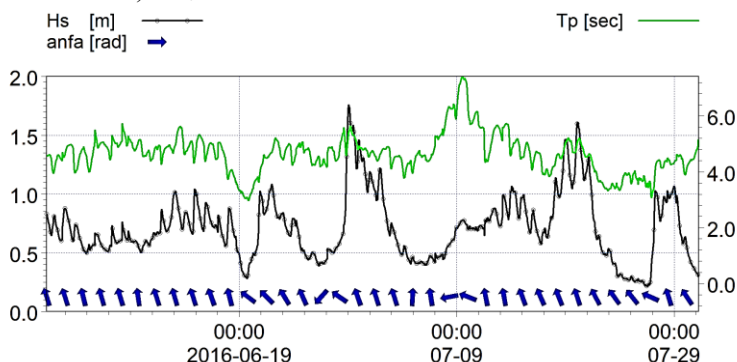
nghiên cứu sử dụng ANN đối với lĩnh vực nghiên cứu thủy động lực biển đang rất mới và thu hút được sự chú ý của các nhà khoa học trong và ngoài nước.

Sóng gần bờ đóng vai trò quan trọng đối với nghiên cứu diễn biến tương tác thủy động lực và công trình biển, nó cũng là đầu vào cho thiết kế các công trình bảo vệ ven biển như đê, kè, đập chắn sóng, cảng biển, v.v. Bên cạnh đó, sóng gần bờ rất có ý nghĩa đối với các hoạt động kinh tế vùng ven biển tại Việt Nam (khai thác và nuôi trồng thủy hải sản, vận tải biển, du lịch biển). Trên thực tế, sóng lớn sẽ gây ra xói lở bãi biển, gây hư hỏng và sạt trượt công trình, cản trở và gây mất an toàn cho tàu thuyền khai thác tại vùng biển. Tính toán lan truyền sóng có thể thực hiện bằng nhiều mô hình toán khác nhau và đạt độ chính xác ở mức chấp nhận được (mô hình Mike 21 SW, DELFT3D, TELEMAC, SWAN). Vận hành khai thác các mô hình này đòi hỏi nhiều bộ tham số, kinh nghiệm và kiến thức chuyên sâu để đạt được hiệu quả cao. Để thực hiện được một mô phỏng cần thời gian dài từ lúc xử lý số liệu đầu vào đến tính toán và thiết lập các thông số vật lý đặc trưng, chạy kịch bản mô phỏng. Đặc biệt, việc mô phỏng dài hạn rất khó khăn do tốc độ tính toán của máy tính có hạn dẫn đến thời gian mô phỏng mất hàng tuần hoặc hơn nữa. Việc ứng dụng ANN để giải quyết bài toán mô phỏng, dự đoán các thông số thủy động lực vùng ven biển như đã chỉ ra ở trên đều cho kết quả tốt, có độ chính xác đảm bảo, đáp ứng số liệu đầu vào dài hạn. Tuy nhiên, để thu được mô hình ANN phù hợp, yêu cầu dữ liệu huấn luyện mạng phải đủ lớn, dài hạn và có độ tin cậy cao. Việc đo đạc khảo sát sóng ngoài khơi và sóng gần bờ đòi hỏi kinh phí lớn, đôi khi không thực hiện được do yếu tố thời tiết bất lợi. Trong bài báo này, nhóm tác giả nghiên cứu ứng dụng mạng ANN tính toán mô phỏng độ cao sóng gần bờ tại khu vực biển Cửa Đại tỉnh Quảng Ngãi. Mô hình mạng ANN được huấn luyện là cơ sở cho tính toán dự đoán độ cao sóng trong khoảng thời gian ngắn hạn 7 ngày tiếp theo. Quá trình xây dựng dữ liệu huấn luyện mạng sử dụng nguồn dữ liệu mô phỏng được trình bày trong phần 2. Dữ liệu mô phỏng sóng ngoài khơi được lấy từ mô hình WAVEWATCH III, dữ liệu sóng gần bờ lấy từ mô hình Mike 21 SW, tập dữ liệu này thay cho dữ liệu thực đo để huấn luyện mạng ANN. Kết quả tính toán mô phỏng và dự đoán dựa trên mạng ANN được thực hiện trong thời gian tháng 6/2016 và đầu tháng 7/2016 được trình bày trong phần tiếp theo. Cuối cùng là kết luận và các kết quả đạt được của bài báo.

2. XÂY DỰNG DỮ LIỆU HUẤN LUYỆN MẠNG ANN

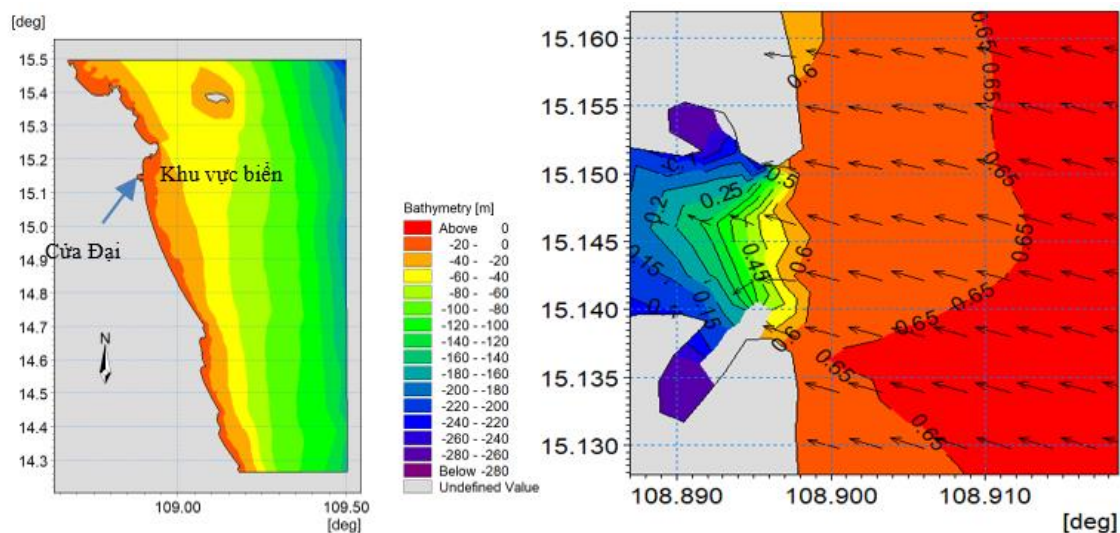
Dữ liệu huấn luyện mạng nơ-ron

Việc tiến hành đo đạc sóng biển dài ngày tạo dữ liệu huấn luyện mạng ANN rất khó khăn cả về mặt kỹ thuật lẫn chi phí tài chính. Để giải quyết bài toán xây dựng dữ liệu huấn luyện mạng ANN nhóm nghiên cứu sử dụng mô hình WAVEWATCH III và mô hình Mike 21 SW tạo kết quả mô phỏng số tại khu vực ngoài khơi và khu vực gần bờ biển Cửa Đại. Hình 1 dưới đây mô tả dữ liệu đầu vào của mạng ANN là số liệu sóng ngoài khơi của NOAA (National Oceanic and Atmospheric Administration) được tính toán bởi mô hình WAVEWATCH III.

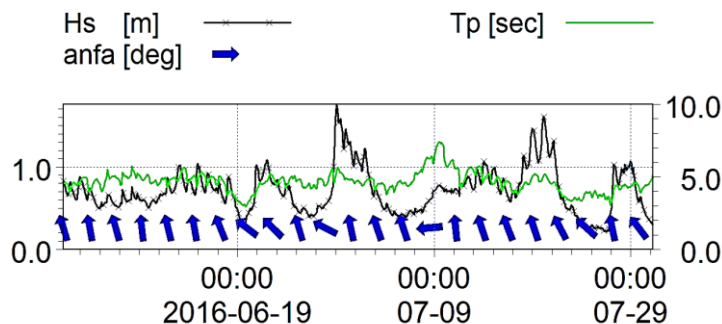


Hình 1. Biểu đồ số liệu sóng ngoài khơi biển Cửa Đại từ 01/6/2016 đến 30/7/2016.

Mô hình Mike 21 SW được thiết lập với phạm vi địa hình bao trùm toàn bộ khu vực biển Cửa Đại tỉnh Quảng Ngãi (hình 2.a), giới hạn ở phía Nam và phía Bắc ở vĩ tuyến $14,3^{\circ}$ và $15,5^{\circ}$, giới hạn phía Đông nằm ở vùng ngoài khơi tại kinh tuyến $109,5^{\circ}$. Điều kiện biên phía ngoài khơi là các thông số sóng ngoài khơi (độ cao sóng: H_s , chu kỳ sóng: T_p , hướng sóng: α) được lấy từ nguồn dữ liệu sóng toàn cầu ERA5 (hình 2.b). Số liệu sóng để kiểm định mô hình là chiều cao sóng được đo khảo sát ở gần bờ thuộc khu vực Cửa Đại tỉnh Quảng Ngãi, đã đo đạc từ 30/6/2016 đến 7/7/2016 (hình 3).

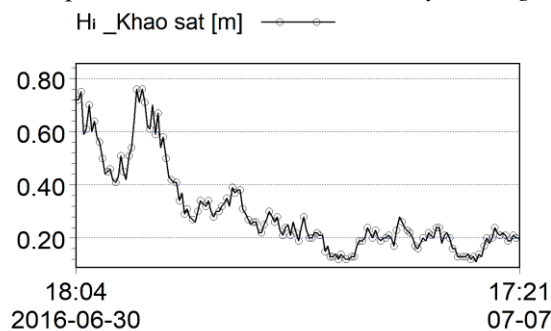


a. Địa hình khu vực biển và trường sóng gần bờ biển Cửa Đại.



b. Điều kiện biên phía ngoài khơi.

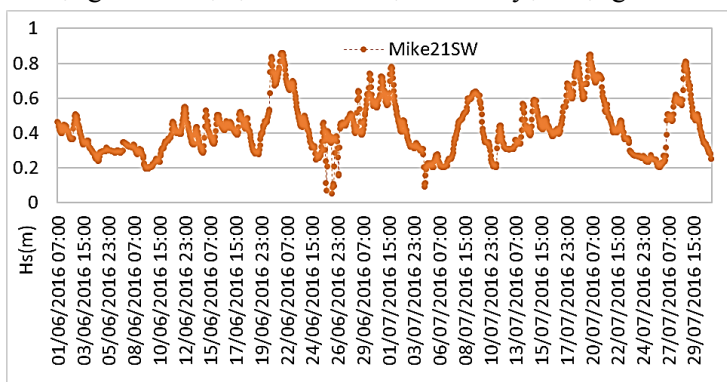
Hình 2. Thiết lập và kiểm định mô hình lan truyền sóng Mike 21 SW.



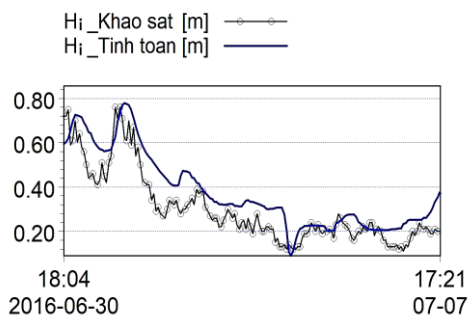
Hình 3. Biểu đồ độ cao sóng gần bờ H_i (m) được khảo sát đo đạc tại biển Cửa Đại tỉnh Quảng Ngãi từ 30/6/2016 đến 7/7/2016.

Quá trình mô phỏng sóng gần bờ tại khu vực biển Cửa Đại được thiết lập từ 01/6/2016 đến 30/7/2016, thời gian tiến hành tính toán mô phỏng theo mô hình Mike 21 SW là 48 h. Kết quả độ cao sóng H_i (m) được lấy theo ba khoảng thời gian từ 01/6/2016 đến 29/6/2016, khoảng thời gian từ 30/6/2016 đến 7/7/2016 và từ 7/7/2016 đến 30/7/2016. Đoạn kết quả thứ hai được sử dụng để kiểm định mô hình dự báo của mạng ANN (hình 4).

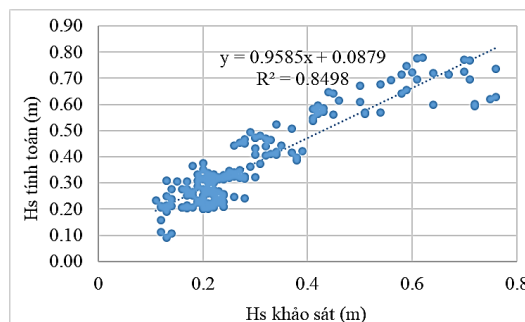
Hình 5a dưới đây là kết quả kiểm định độ cao sóng trong thời gian từ 30/6/2016 đến 7/7/2016, độ cao sóng tính toán theo mô hình Mike 21 SW tương đối phù hợp với kết quả khảo sát độ cao sóng tại khu vực gần bờ biển Cửa Đại. Hình 5b thể hiện tương quan giữa độ cao sóng đo được và độ cao sóng tính toán, hệ số tương quan đạt xấp xỉ 85%. Kết quả này cho thấy mô hình Mike 21 SW đủ tin cậy để sử dụng như là bộ tạo nhân dữ liệu huấn luyện mạng ANN.



Hình 4. Số liệu độ cao sóng gần bờ H_i (m) tính toán bằng Mike 21SW từ 1/6/2016 đến 30/7/2016.



a. Đồ thị độ cao sóng gần bờ theo mô hình Mike 21 SW và khảo sát từ ngày 30/6/2016 đến 7/7/2016.



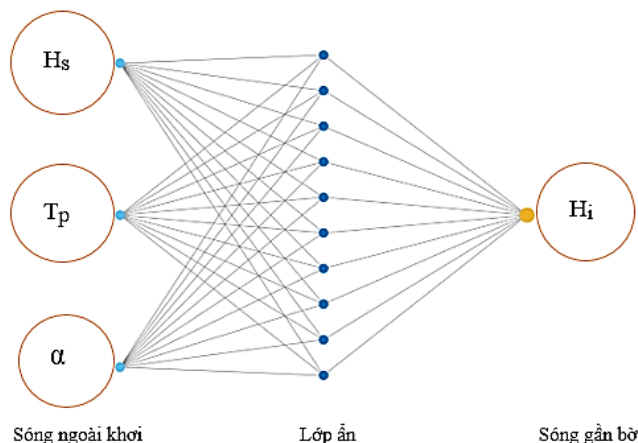
b. Tương quan kết quả mô phỏng và kết quả khảo sát.

Hình 5. Chiều cao sóng gần bờ được khảo sát và tính toán bằng Mike 21SW.

3. XÂY DỰNG MẠNG ANN MÔ PHỎNG VÀ DỰ ĐOÁN ĐỘ CAO SÓNG GẦN BỜ

3.1. Lựa chọn mô hình mạng ANN

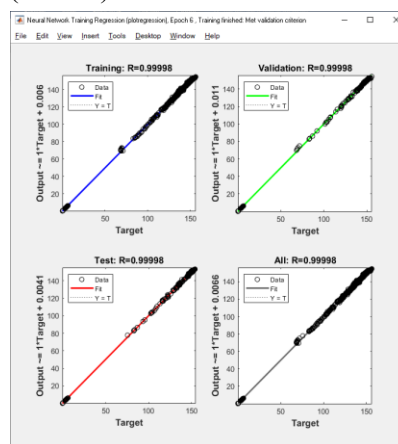
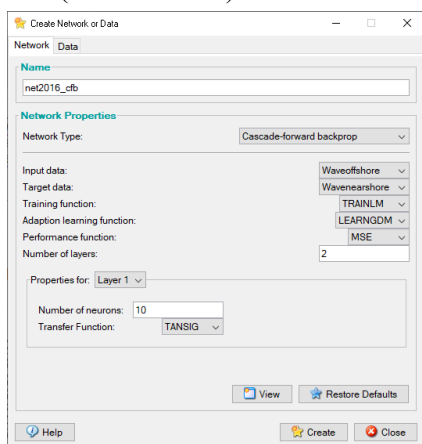
Trong nghiên cứu này, việc dự báo sóng gần bờ tại khu vực biển Cửa Đại tỉnh Quảng Ngãi được thực hiện bằng mạng nơ-ron nhân tạo đa lớp (Multi-layer Perceptron - MLP) và giải thuật lan truyền ngược (Backpropagation algorithm). Để huấn luyện mạng nơ-ron cần chuẩn bị một tập dữ liệu huấn luyện (bao gồm dữ liệu đầu vào và nhãn dữ liệu). Quá trình huấn luyện mạng nơ-ron sử dụng giải thuật lan truyền ngược là quá trình tối ưu hóa trọng số để đưa sai số về giá trị nhỏ nhất. Hình 6 mô tả cấu trúc mạng MLP với đầu vào là thông số sóng ngoài khơi: chiều cao sóng H_s , chu kỳ sóng T_p , hướng sóng α ; kế tiếp là lớp ẩn, số lượng nơ-ron ở lớp ẩn sẽ được xác định cụ thể trong phần tiếp theo, cuối cùng là đầu ra của mạng ANN, chiều cao sóng gần bờ H_i .



Hình 6. Cấu trúc mạng ANN MLP 3 đầu vào và 1 đầu ra.

3.2. Xây dựng mô hình mạng ANN trên phần mềm Matlab

Mạng ANN được khởi tạo với giải thuật lan truyền ngược (back propagation algorithm) trên phần mềm Matlab, hàm kích hoạt được lựa chọn là hàm *tansig()*, số lớp là ẩn là 1, số nơ-ron ở lớp ẩn 10 nơ-ron. Dữ liệu huấn luyện mạng chia theo tỷ lệ: tập huấn luyện (training set) chiếm 70%, tập tối ưu (validation set) chiếm 15% và tập thử nghiệm (test set) chiếm 15%.

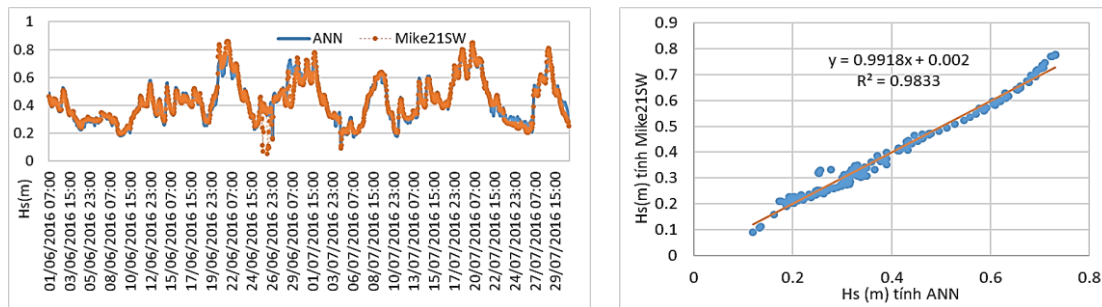


a. Khởi tạo mạng ANN trên phần mềm Matlab.

b. Kết quả huấn luyện mạng.

Hình 7. Quá trình tạo mạng ANN và thực hiện mô phỏng sóng gần bờ Cửa Đại từ 1/6/2016 đến 29/6/2016.

Hình 7b cho thấy quá trình huấn luyện, kiểm định, tính thử nghiệm đạt độ chính xác cao với hệ số tương quan trên 99%. Độ cao sóng gần bờ H_s tính bằng ANN và mô hình Mike 21SW trong gian đoạn từ 01/6/2016 đến 29/6/2016 cho kết quả hoàn toàn trùng nhau. Để thấy rõ hơn về hiệu quả của việc sử dụng ANN trong bài toán dự đoán chiều cao sóng, nhóm tác giả tiếp tục sử dụng mạng ANN đã được huấn luyện kể trên, dự đoán sóng gần bờ Cửa Đại trong thời gian 7 ngày đầu tháng 7/2016 và các ngày còn lại của tháng 7/2016 với dữ liệu đầu vào là sóng ngoài khơi từ nguồn WaveWatch III. Kết quả thể hiện như hình 8a cho thấy rõ rệt sự tương đồng giữa kết quả tính toán bằng hai công cụ khác nhau, tương quan kết quả giữa hai mô hình đạt 98,33%, trong khi đó sử dụng ANN dự đoán cho kết quả trong khoảng 10 s, nhanh hơn rất nhiều so với tính toán bằng Mike21SW mất đến gần 48 h mô phỏng tính toán. Do độ tương quan giữa kết quả dự đoán bằng ANN và Mike 21SW lớn, chúng ta hoàn toàn có thể kết luận về mức độ tương quan về kết quả dự đoán bằng ANN và dữ liệu đo thực tế trong thời gian đầu tháng 7/2016, giá trị này có thể đạt trên mức 84%, tương tự như hình 5b.



a. Kết quả Hs tính bằng ANN và Mike 21SW.

b. Tương quan Hs từ ngày 30/6/2016 đến 7/7/2016 tính bằng ANN và Mike 21SW.

Hình 8. Kết quả quá trình mô phỏng từ 30/6/2016 đến 7/7/2016 bằng ANN.

4. KẾT LUẬN

Nghiên cứu đã cho thấy rõ khả năng và tính ứng dụng ANN hay mạng nơ-ron nhân tạo trong việc tính toán mô phỏng, dự đoán độ cao sóng gần bờ từ số liệu sóng ngoài khơi. Bằng cách sử dụng hai mô hình gồm Mike 21SW và ANN, nhóm tác giả đã dự đoán tham số độ cao sóng Hs (m) tại khu vực gần bờ biển Cửa Đại tỉnh Quảng Ngãi với xuất phát là sóng khởi điểm ngoài khơi. Mô phỏng được thực hiện trong các khoảng thời gian từ 1/6-29/6/2016. Đáng lưu ý, trong giai đoạn dự đoán từ 30/6/2016 đến 7/7/2016 có số liệu khảo sát sóng ngoài thực tế để kiểm tra các mô hình, kết quả cho thấy độ chính xác cao. Tương quan kết quả tính bằng ANN so với số liệu khảo sát đạt trên 80%, so với Mike 21SW đạt đến trên 98%, tốc độ giải của ANN cho một tháng dữ liệu chỉ diễn ra trong 10s và nhanh hơn rất nhiều so với mô hình Mike 21 SW.

Lời cảm ơn: Nghiên cứu này được thực hiện trong khuôn khổ đề tài KHCN tiềm năng cấp Bộ Nông Nghiệp và Phát triển nông thôn “Nghiên cứu mạng nơ-ron nhân tạo ANN dự đoán tham số sóng gần bờ khu vực Cửa Đại tỉnh Quảng Ngãi”.

TÀI LIỆU THAM KHẢO

- [1]. Scott C. James, Yushan Zhang, Fearghal O'Donncha, “A machine learning framework to forecast wave conditions, *Coastal Engineering*”, **Volume 137**, Pages 1-10, (2018).
- [2]. J. Berbić, E. Ocvirik, D. Carević, G. Lončar, “Application of neural networks and support vector machine for significant wave height prediction”, *Oceanologia*, **Volume 59**, Issue 3, pp. 331-349, (2017).
- [3]. D.J. Peres, C. Iuppa, L. Cavallaro, A. Cancelliere, E. Foti, “Significant wave height record extension by neural networks and reanalysis wind data, *Ocean Modelling*, **Volume 94**, Pages 128-140, (2015).
- [4]. Hyun-Doug Yoon, Daniel T. Cox, Munki Kim, “Prediction of time-dependent sediment suspension in the surf zone using artificial neural network”, *Coastal Engineering*, **Volume 71**, Pages 78-86, (2013).
- [5]. Yunwei Wang, Jun Chen, Hui Cai, Qian Yu, Zeng Zhou, “Predicting water turbidity in a macro-tidal coastal bay using machine learning approaches”, *Estuarine, Coastal and Shelf Science*, **Volume 252**, (2021), 107276, ISSN 0272-7714, <https://doi.org/10.1016/j.ecss.2021.107276>.
- [6]. Sooyoul Kim, Tracey H.A. Tom, Masahide Takeda, Hajime Mase, “A framework for transformation to nearshore wave from global wave data using machine learning techniques: Validation at the Port of Hitachinaka, Japan”, *Ocean Engineering*, **Volume 221**, 108516, (2021).
- [7]. Lê Văn Nghinh, Hoàng Thanh Tùng, Nguyễn Ngọc Hải, “Nghiên cứu ứng dụng mạng nơ-ron thần kinh vào dự báo lũ các sông ở tỉnh Bình Định và Quảng Trị”, *Tạp chí Khoa học Kỹ thuật Thủy lợi và Môi trường*, Số 14, (2006).
- [8]. Nguyễn Đăng Tinh, “Ứng dụng mạng nơ-ron nhân tạo để dự báo mưa và dòng chảy làm cơ sở cho công tác phòng tránh và giảm nhẹ thiên tai hạn hán trên một số lưu vực sông thuộc vùng Tây nguyên Việt Nam”, *Tạp chí Khoa học Kỹ thuật Thủy lợi và Môi trường*, Số 22, (2008).
- [9]. Đặng Văn Tô, “Phục hồi dữ liệu sóng biển bằng mạng nơ-ron nhân tạo”, *Tạp chí Khoa học và Công nghệ biển T10*, Số 1. Tr 17 – 25, (2010).

- [10]. Trần Hồng Thái, Mai Văn Khiêm, Nguyễn Bá Thùy, Bùi Mạnh Hà, Phạm Khánh Ngọc, “*Xây dựng mô hình mạng nơ ron hồi quy dự báo độ cao sóng có nghĩa tại trạm Cồn Cỏ, Quảng Trị, Việt Nam*”, Tạp chí Khí tượng Thủy văn, Tập. 4 Số. 736(1), Tr 73-84, (2022).

ABSTRACT

Research on application of Artificial neural network to predict high waves near the shore in Cua Dai sea area, Quang Ngai province

Artificial Neural Networks (ANNs) are considered as a powerful tool to solve complex and nonlinear problems, especially in cases where relationships between processes are not explicitly established. Currently, artificial neural networks are widely applied in various fields of science and technology and in hydrodynamic studies in coastal estuaries is no exception. Using ANN has significant advantages in large data processing, fast resolution with guaranteed accuracy. In this study, the authors used an artificial neural network to calculate, simulate and predict high waves near the coast in Cua Dai sea area in Quang Ngai province during the period of June 2016 and early July. 2016. The ANN training dataset includes starting data taken from NOAA offshore wave data, the output data is taken from the calculation results of Mike 21SW mathematical model at Cua Dai Sea. Ability to enhance reliability when building the Mike 21SW mathematical model, the output results of the Mike 21SW model will be compared with the actual measurement results at Cua Dai beach within the first 7 days of July 2016. The results of computational simulation and prediction of high near-shore broadcast rates of the ANN are compared with the results of the Mike 21SW mathematical model and the actual survey results.

Keywords: ANN; Significant wave height; Wave prediction; Mike 21SW.