

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC NÔNG LÂM

HOÀNG QUÝ NHÂN

NGHIÊN CỨU ỨNG DỤNG TRÍ TUỆ NHÂN TẠO
VÀO DỰ BÁO LŨ SÔNG HỒNG TRONG
BỐI CẢNH BIẾN ĐỔI KHÍ HẬU

Ngành: Khoa học môi trường
Mã số: 9.44.03.01

TÓM TẮT LUẬN ÁN TIẾN SĨ KHOA HỌC MÔI TRƯỜNG

THÁI NGUYÊN - 2024

**Công trình được hoàn thành tại
Trường Đại học Nông Lâm – Đại học Thái Nguyên**

Người hướng dẫn khoa học:

- 1. PGS. TS. NGUYỄN XUÂN HOÀI**
- 2. PGS. TS. ĐỖ THỊ LAN**

Phản biện 1:

Phản biện 2:

Luận án sẽ được bảo vệ trước hội đồng chấm luận án cấp
Họp tại: **Trường Đại học Nông Lâm – Đại học Thái Nguyên**
Vào hồi giờ phút, ngày tháng năm

Có thể tìm hiểu luận án tại:

- Thư viện Quốc gia**
- Trung tâm Số - Đại học Thái Nguyên**
- Thư viện Trường Đại học Nông Lâm - ĐH Thái Nguyên**

DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CÓ LIÊN QUAN

1. Hoang Quy Nhan, Do Thi Lan, Khuat Thi Thanh Huyen, Nguyen Xuan Hoai, Ngo Van Manh (2019); Flood forecast by a deeplearning method at sontay station on the red river of Vietnam. The WORLD CONFERENCE of integrated land and water management and climate change in Vietnam and Japan, ISBN: 978-604-67-1316-4
2. Hoang Quy Nhan ^{1*}, Do Thi Lan ¹, Nguyen Xuan Hoai ²; Results of applying standardized methods of hydrographic data for stations 74162 - Son Tay - Viet Nam; The 42nd Asian Conference on Remote Sensing (ACRS2021). ISBN: 978-604-965-591-3.
3. Tuan-Dung Hoang, Hoang Quy Nhan, and Nguyen Vo Chau Ngan; (2022) *Chapter 2; Artificial Intelligence in Pollution Control and Management: Status and Future Prospects; Artificial Intelligence and Environmental Sustainability Challenges and Solutions in the Era of Industry 4.0*, P36-P56; Springer Nature Singapore Pte Ltd. 2022; https://doi.org/10.1007/978-981-19-1434-8_2
4. Hoang Q.N.¹, Do T.L.^{1*}, Nguyen N.S.H.¹ and Nguyen X.H.² (2023); *Application of remote sensing technology to monitor urban waste dumps and forecast waste collection points in waste management for smart cities*. 4thWORLD CONFERENCE ON WASTE MANAGEMENT 2023 (WCWM 2023).
5. Hoang Quy Nhan¹, Dang Thi Thai Ha^{2*}, Do Thi Lan¹, Nguyen Xuan Hoai³; (2023) *Assessing the impact of climate change on the operation of sonla hydropower reservoir on water quality in semisubmerged land and proposing measures for improvement*; TNU Journal of Science and Technology 228(S2): 162 – 167; ISSN 1859-2171, 2374-9098; e-ISSN 2615-9562.

PHẦN MỞ ĐẦU

1. Sự cần thiết

Dòng sông mang lại và cung cấp nhiều lợi ích đối với con người và môi trường, hệ sinh thái xung quanh dòng sông. Sông ngòi cung cấp nước cho sự sống hai bên bờ sông, giao thông và đi lại cho con người, sử dụng tưới tiêu cho nông nghiệp, phát triển công nghiệp năng lượng, là nguồn sống của sự phát triển. Dòng sông rất quan trọng đối với cuộc sống của con người, tuy nhiên nó cũng có thể gây ra những thiệt hại vô cùng lớn (Xuan-Hien Le và cs, 2019). Lũ lụt là nguyên nhân quan trọng gây ra những thiệt hại về kinh tế, xã hội cũng như thiệt hại về nhân mạng, cho các khu vực đông dân cư nằm cạnh và ở hạ lưu các con sông lớn. Lũ lụt cũng khác nhau ở các lưu vực sông trên thế giới và Việt Nam, nhưng hầu hết các trận lũ lớn tùy vào quy mô hoặc cường độ đều gây nên những thiệt hại đáng kể cho một khu vực rộng lớn (Trần Thanh Xuân và cs, 2018). Lũ ở miền bắc Việt Nam có tính chất chu kỳ, theo mùa, và có thể dự báo trước trong một khoảng thời gian nhất định, nhưng trong bối cảnh biến đổi khí hậu (BĐKH) trên toàn cầu hiện nay đang thay đổi nhanh chóng, với các kiểu thời tiết bất thường, không theo quy luật dẫn tới lũ lụt đang xảy ra với tần suất thường xuyên hơn, cường độ các trận lũ lụt cũng khó dự đoán hơn, lũ lụt xuất hiện mới vượt các trận lũ được ghi nhận trong lịch sử (Trình Thu Phương, 2021).

Trên thế giới và cả ở Việt Nam, dự báo chính xác dòng chảy lũ là một yêu cầu thiết yếu để tận dụng nguồn lợi từ lũ theo mùa và phòng tránh thiệt hại từ lũ lớn. Nhiệm vụ dự báo lũ, cảnh báo thiệt hại và giảm tác động xấu từ lũ lụt là rất quan trọng đối với việc lập kế hoạch và quản lý hệ thống tài nguyên nước. Tuy nhiên, dự báo chính xác lũ, mực nước, lưu lượng lũ sông là một vấn đề khó khăn vì phải phân tích các giai đoạn lũ sông là một quá trình động học phức tạp được đặc trưng bởi sự thay đổi theo không gian và thời gian. Ngoài ra, quá trình dòng chảy của sông là phi tuyến tính và chịu ảnh hưởng của nhiều yếu tố như lớp phủ bề mặt lưu vực sông, quá trình mưa cũng như địa hình lòng sông và đặc điểm khí hậu. Nhiều biện pháp dự báo đòi hỏi lượng dữ liệu khổng lồ để dự báo. Dự báo lũ có hai phương pháp dự báo lũ sông (Adikari, K.E. và cs., 2021), phương pháp đầu tiên bao gồm các mô hình toán học mô phỏng quá trình thủy động lực của dòng nước. Phương pháp mô phỏng hiện nay vẫn đang được sử dụng rộng rãi vì các mô hình toán học dựa trên các khái niệm về thủy lực và thủy văn có truyền thống lâu dài. Những mô hình này thường có xu hướng yêu cầu

các mối tương quan với lượng dữ liệu cần điều tra, cập nhật lớn cho dữ liệu đầu vào (như lượng mưa, vận hành hồ chứa, dữ liệu địa hình...) mà không phải lúc nào cũng có sẵn hoặc khó có thể điều tra được.

Việc áp dụng những công nghệ ứng dụng, xây dựng cơ sở cốt lõi mới của cuộc cách mạng công nghiệp 4.0 và sự phát triển khoa học của nhân loại trong thế kỉ 21. Các kỹ thuật về Trí tuệ nhân tạo (AI) đang được nghiên cứu ứng dụng rộng rãi trong nhiều ngành khoa học, lĩnh vực, trong đó lĩnh vực khoa học môi trường đang hết sức mới mẻ và có tính thời sự. Trên thực tế, AI đã bắt đầu được ứng dụng trong việc dự đoán các hiện tượng bất thường, nguy hiểm trong lĩnh vực Môi trường như dự báo mưa bão quy mô lớn, dự đoán cháy rừng, dự đoán và quản lý rác thải... (Veintimilla Reyes và cs., 2016) đã cải thiện việc sử dụng bảy phương pháp phổ biến truyền thống áp dụng AI để đánh giá dự báo lượng mưa của 42 thành phố dựa trên dữ liệu trong quá khứ. Thực nghiệm các mô hình AI cho thấy các phương pháp mới này cho kết quả vượt trội hơn so với phương pháp truyền thống. Các kỹ thuật Trí tuệ hiện đại áp dụng trong xử lý dữ liệu thời gian dựa trên Machine learning (ML), Deep learning (DL) như ANN, LSTM, RNN..., (Phan Thi Thu Hong và cs, 2018) đã được áp dụng và cho kết quả tốt trong dự báo lượng mưa ở Việt Nam.

Với sự phát triển của công nghệ thông tin và AI hiện đại đang tạo ra những cơ hội lớn để phát triển các giải pháp mới để giải quyết các vấn đề trong lĩnh vực dự báo lũ trong ngành khoa học môi trường, cuộc đua nghiên cứu trên thế giới về ứng dụng AI vào dự đoán các hiện tượng này trong bối cảnh BĐKH tạo nên sức mạnh mới trong các hướng nghiên cứu khoa học môi trường ứng dụng hiện nay. Xuất phát từ những vấn đề trên tôi lựa chọn đề tài: **“NGHIÊN CỨU ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO DỰ BÁO LŨ SÔNG HỒNG TRONG BỐI CẢNH BIẾN ĐỔI KHÍ HẬU”**. Luận án đặt ra mục tiêu nghiên cứu được lũ trên sông trên cơ sở cải thiện bộ dữ liệu thực tế, tính toán đảm bảo dự báo lũ trên sông trong khoảng thời gian nhất định, có độ tin cậy trong khoa học giúp dự báo rủi ro môi trường, đồng thời nâng cao hiệu quả của phòng chống lũ và quản lý tài nguyên nước, vận hành hồ chứa. Nhiệm vụ của luận án này sẽ hướng tới giải quyết một phần vấn đề còn tồn tại nêu trên, xử lý dữ liệu thủy văn trong khoa học môi trường (KHMT), ứng dụng AI cho các dự báo lũ, với đặc trưng lũ tại trạm thủy văn Sơn Tây, Sông Hồng của Việt Nam.

2. Mục tiêu của đề tài

2.1. Mục tiêu tổng quát

Góp phần nghiên cứu phương pháp xử lý cơ sở dữ liệu mực nước và ứng dụng mô hình AI vào dự báo lũ Sông Hồng trong bối cảnh BĐKH ở Việt Nam.

2.2. Mục tiêu cụ thể

1) Phân tích, đánh giá một số đặc trưng, cấu trúc, tính chất và các yếu tố ảnh hưởng đến bộ dữ liệu mực nước sông Hồng, đoạn chảy qua trạm thủy văn Sơn Tây, Hà Nội, giai đoạn 2011-2020. Xây dựng được bộ dữ liệu chuẩn hóa, huấn luyện và kiểm tra phục vụ cho dự báo mực nước lũ sông Hồng bằng AI.

2) Xây dựng phương pháp, áp dụng cơ sở lý thuyết và mô hình ứng dụng AI vào dự báo mực nước lũ sông Hồng tại trạm thủy văn Sơn Tây trong bối cảnh biến đổi khí hậu.

3) Dự báo mực nước lũ sông Hồng tại trạm thủy văn Sơn Tây thực nghiệm với 2 mô hình RNN-AI và LSTM-AI có kết quả, nhằm dự báo mực nước lũ sông, hỗ trợ cảnh báo, nhận diện lũ sông Hồng tại trạm thủy văn Sơn Tây, trong giai đoạn 2011-2020 có độ tin cậy.

4) Đánh giá độ tin cậy mô hình dự báo, đề xuất mô hình ra quyết định, điều tiết mực nước trong quản lý tài nguyên nước sông Hồng trong bối cảnh biến đổi khí hậu hiện nay.

3. Ý nghĩa của đề tài

3.1. Ý nghĩa khoa học

- Bổ sung cập nhật, xây dựng cơ sở dữ liệu về kết quả nghiên cứu chuẩn hóa bộ dữ liệu mực nước Sông Hồng tại trạm Sơn Tây, giai đoạn 2011-2020 góp phần làm sáng tỏ luận cứ, cơ sở khoa học và lý thuyết ứng dụng vào xây dựng bộ dữ liệu mực nước sông hoàn chỉnh vào dự báo lũ sông, sử dụng cho các mô hình AI trên nền dữ liệu sông Hồng quan trắc thực tế. Trên cơ sở dữ liệu chuẩn hóa có được thực nghiệm với 02 mô hình dự báo linh hoạt là mô hình RNN-AI và LSTM-AI theo hướng ứng dụng công nghệ mới, đưa ra các kết quả dự báo, xác định các yếu tố ảnh hưởng, thời gian dự báo và so sánh với các phương pháp dự báo khác. Kết quả đạt được của luận án đóng góp làm cơ sở khoa học, tài liệu tham khảo cho phương pháp dự báo lũ sông bằng AI trên cơ sở dữ liệu mực nước sông.

- Kết quả của luận án sẽ là tài liệu để tham khảo cho các nghiên cứu tiếp theo và tiếp tục thực hiện các nghiên cứu về dự báo lũ, xây dựng phần mềm, ra quyết định quản lý vận hành tài nguyên nước, trong bối cảnh công nghệ phát triển và BĐKH hiện nay nói riêng, và trong dự

báo trong ngành KHMT nói chung.

3.2. Ý nghĩa thực tiễn

- Kết quả nghiên cứu của luận án này dựa trên thực tiễn số liệu quan trắc tại trạm Sơn Tây, Sông Hồng. các kết quả chuẩn hóa xử lý cơ sở dữ liệu, mô hình dự báo lũ Sông Hồng bằng RNN và LSTM sẽ góp phần tạo tiền đề, đề xuất điều chỉnh các phương pháp dự báo lũ trong tương lai, phòng chống lũ trên lưu vực Sông Hồng và các con sông khác (khi có sự phối hợp đầy đủ trên hệ thống sông, mực nước sông và các quan hệ hồ chứa lớn trong lưu vực) trong bối cảnh biến đổi khí hậu hiện nay.

- Kết quả đạt được sẽ đóng góp phần nào hiệu quả nâng cao khả năng dự báo lũ cho trạm thủy văn Sơn Tây, nhận diện lũ trên sông, giúp vận hành quản lý hồ chứa, sử dụng tài nguyên nước hợp lý mà vẫn đảm bảo nhiệm vụ phòng, chống lũ cho Sông Hồng và tận dụng lợi ích của lũ sông theo mùa cho hạ du. Kết quả phục vụ trực tiếp trong công tác quản lý và sử dụng hợp lý tài nguyên nước hiện nay khi nguồn tài nguyên này đang ngày càng suy kiệt.

4. Những đóng góp mới của luận án

- Luận án sẽ cung cấp các dẫn liệu khoa học về phương pháp mới trong xử lý cấu trúc bộ dữ liệu Sông Hồng tại trạm Thủy văn Sơn Tây như: Đánh giá đặc điểm, tính chất của mực nước trong giai đoạn 2011-2020, đặc điểm tính chất mực nước lũ Sông Hồng giai đoạn 2011-2020, trong bối cảnh biến đổi khí hậu. Đề xuất phương thức tiên xử lý dữ liệu sử dụng các kỹ thuật toán học, thống kê và huấn luyện mô hình dự đoán mực sông Hồng bằng AI, ứng dụng dự báo lũ sông trong điều kiện quan trắc của Việt Nam.

- Luận án hỗ trợ xây dựng quy trình mới, cơ sở lý thuyết, phương pháp ứng dụng Thực nghiệm mô hình dự đoán mực nước sông Hồng sử dụng AI (RNN-AI và LSTM-AI) cho kết quả đáng tin cậy.

- Luận án thực hiện mô hình hóa, đưa ra một số phân tích về ảnh hưởng của mực nước sông Hồng đến sự phát triển của kinh tế xã hội và quản lý tài nguyên nước, đề xuất thử nghiệm mô hình cây quyết định (Decision Tree) cho quản lý tài nguyên này, tạo tài liệu trích dẫn khoa học cho hướng dự báo lũ trong tương lai và ra quyết định trong quản lý tài nguyên nước, khoa học môi trường trong thời kỳ công nghệ mới.

CHƯƠNG 1: TỔNG QUAN CÁC VẤN ĐỀ NGHIÊN CỨU

1.1. Các kết quả nghiên cứu tổng quan lũ sông trong bối cảnh biến đổi khí hậu

1.1.1. Lũ sông, cơ sở khoa học và nguyên nhân hình thành lũ

Lũ sông là hiện tượng nước dâng cao từ từ hoặc nhanh chóng trong một khoảng thời gian nhất định, sau đó giảm dần về mức bình thường. Khi có lũ, mức nước trong sông dâng cao và tăng nhanh, tốc độ dòng chảy của nước vượt quá mức bình thường. Lũ sông gây ra lụt khi nước sông dâng cao tràn qua đê ngăn, nước sông tràn vào các vùng bên ngoài dòng sông, làm ngập các khu vực có độ cao thấp hơn và các vùng lân cận, gây thiệt hại nhà cửa, đất đai, tài sản và con người. Lũ tên trong tiếng anh Flood, các khái niệm về đỉnh lũ là mực nước hay lưu lượng lũ là lượng nước được chảy qua một đơn vị mặt cắt ngang của sông trên một đơn vị thời gian như giây, thường được tính bằng m³/giây, đỉnh lũ H và lưu lượng nước lũ Q trong sông thường được đo đạc và quan tâm nhất trong một trận lũ sông (Nguyễn Thanh Sơn, 2003).

Lũ sông thường được phân cấp theo độ lớn, Căn cứ vào độ lớn đỉnh lũ trung bình nhiều năm, theo cục Quản lý tài nguyên nước, (2009) mùa lũ Sông Hồng có thể chia ra các cấp lũ như sau:

- Lũ nhỏ là lũ có mực nước đỉnh lũ thấp hơn đỉnh lũ trung bình nhiều năm;
- Lũ vừa là lũ có mực nước đỉnh lũ đạt mức đỉnh lũ trung bình nhiều năm;
- Lũ lớn là lũ có mực nước đỉnh lũ cao hơn mức đỉnh lũ trung bình nhiều năm;
- Lũ đặc biệt lớn là lũ có đỉnh cao hiếm thấy trong các thời kỳ quan trắc;

Các nguyên nhân chính hình thành lũ sông, Theo các nghiên cứu của Bùi Đình Lập, (2016) gồm có:

- Thứ nhất, Những trận mưa lớn kéo dài, hệ thống sông, thủy lợi không thể kịp thời tiêu dòng chảy tự nhiên, làm tăng mực nước trong sông gây ra lũ.
- Thứ hai, Lũ lụt còn có thể xảy ra khi điều tiết đập thủy điện, hồ chứa nước, đê điều, hoặc kè bị vỡ.
- Thứ ba, Các trận bão lớn, áp thấp nhiệt đới ven biển có thể làm nước biển dâng lên, nước chảy ngược tại các cửa sông.

1.1.2. Lợi ích và hậu quả của lũ sông

1.1.2.1. Lợi ích của lũ sông

Lũ sông có tính chất tuần hoàn theo mùa, cho phép nước sông chảy với lưu lượng lớn đến nhiều khu vực trên và dưới mặt đất. Nước lũ này có thể được lưu trữ và sử dụng cho các hồ đập, tạo nên lượng nước lưu

trữ phục vụ cho những mùa thiếu nước. Với dòng chảy lớn và nhanh, lũ sông cũng lọc các chất ô nhiễm ra khỏi sông và đất nuôi dưỡng để hỗ trợ hệ sinh thái và các khu vực màu mỡ cho nông nghiệp.

1.1.2.2. Hậu quả của lũ sông

Trong bối cảnh BĐKH, trái với lũ sông theo quy luật, những trận Lũ lụt rất lớn xảy ra ít thường xuyên hơn nhưng nó mang lại nhiều tác hại to lớn. Chúng khiến dòng sông lan rộng ra một khu vực rộng lớn hơn, đe dọa nhà cửa và cơ sở kinh doanh được xây dựng ở vùng ngập lũ.

Bảng 1.1. Thông kê thiệt hại về kinh tế do lũ lụt giai đoạn 2010-2019.

DVT: Ngàn tỷ đồng

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Tổng thiệt hại do lũ	2,3	5,3	5,1	3,9	2,6	2,5	6,2	27,3	11,5	14,2

Lũ lớn có thể gây ra rất nhiều thiệt hại nặng nề, như mất mạng con người, thiệt hại về tài sản và cơ sở hạ tầng, giao thông thủy, xói mòn và nguy cơ lở đất hai bên bờ sông, Lũ lụt tràn qua đê gây phá hủy mùa màng và thiệt hại về kinh tế lớn, Các mối đe dọa đối với cuộc sống của con người và các hệ sinh thái khác. Sau lũ là rủi ro sức khỏe do ô nhiễm nguồn nước, dịch chuyển nhà ở, tác động lớn đến kinh tế...

Lũ sông cho thấy nhiều mặt có lợi cũng như tác hại của lũ lớn cũng rất nghiêm trọng và khủng khiếp, ngày nay con người khai thác các nguồn nước để phục vụ nhiều mục đích khác nhau như cấp nước, giao thông, điều phối, sản xuất năng lượng, lương thực, kể cả về chính trị và quân sự..., Con người đã tận dụng triệt để nguồn tài nguyên này mà nhưng nghĩ rằng nó là vô hạn nhưng thực tế thì không phải như vậy. Các nguồn nước đang suy giảm trầm trọng, chất lượng nước đang đến bờ vực cảnh báo, các hiện tượng thiên tai liên quan đến nguồn nước trong các con sông xuất hiện với tần suất nhiều và phá vỡ nhiều kỷ lục được ghi nhận trong bối cảnh BĐKH hiện nay. Do đó việc nghiên cứu dự báo lũ sông là hết sức cần thiết và cần được nghiên cứu sâu rộng với nhiều phương pháp mới hơn nữa trong thời điểm bối cảnh BĐKH phức tạp hiện nay.

1.2. Ảnh hưởng của biến đổi khí hậu đến lũ sông

1.2.1. Biến đổi khí hậu

Biến đổi khí hậu là thuật ngữ được dùng để chỉ sự thay đổi của khí hậu do tác động chủ yếu của con người làm thay đổi các thành phần của khí quyển trái đất.

1.2.2. Ảnh hưởng của biến đổi khí hậu đến lũ sông

Biến đổi khí hậu có thể khiến lũ sông trở nên lớn hơn hoặc thường xuyên hơn ở một số nơi trước đây nhưng lại trở nên nhỏ hơn và ít thường xuyên hơn ở những nơi khác. Vì nhiệt độ ẩm hơn khiến nước bốc hơi nhiều hơn từ đất liền và đại dương, nên những thay đổi về quy mô và tần suất của các đợt mưa lớn có thể lần lượt ảnh hưởng đến quy mô và tần suất lũ lụt trên sông. Thay đổi về dòng chảy, thời gian cũng có thể ảnh hưởng đến các mô hình dự báo lũ lụt. Chúng ta có thể khẳng định rằng, BĐKH đang có diễn biến phức tạp, khó lường trước được những hậu quả do BĐKH gây ra, điều đó dẫn tới việc quan trọng là phải hiểu rõ các yếu tố hình thành và dự báo lũ trong thời kỳ công nghệ mới để dự báo lũ sông nhằm đảm bảo việc quản lý và sử dụng nguồn tài nguyên vô giá này. Vì vậy trong khuôn khổ luận án này tác giả sẽ đi sâu nghiên cứu các ứng dụng công nghệ AI vào dự báo lũ sông cho một khu vực cụ thể để từ đó có được cơ sở lý thuyết và kinh nghiệm thực nghiệm cho dự báo lũ trong ngành Khoa học môi trường với yêu cầu mới của phát triển khoa học kỹ thuật hiện nay.

1.3. Các nghiên cứu về dự báo lũ sông

1.3.1. Dự báo lũ sông trên thế giới

Trong nhiều hệ thống vận hành dự báo lũ và lũ sông trên thế giới, dự báo được đưa vào các mô hình định lượng mưa - dòng chảy - mực nước và để dự báo lưu lượng, tốc độ của dòng chảy và mực nước đỉnh lũ thể hiện ở chế độ Q-H, (Q là lưu lượng nước, H là chiều cao mực nước) trong khoảng thời gian từ vài giờ đến vài ngày, tùy thuộc vào quy mô của con sông hoặc lưu vực mà con sông ảnh hưởng tới.

Cùng với sự phát triển của các phương pháp dự báo lũ trên thế giới, phương thức kết xuất các kết quả dự báo được thể hiện dưới nhiều hình thức khác nhau.

1.3.2. Dự báo lũ sông ở Việt Nam

Khoa học dự báo thủy văn ở Việt Nam hiện nay dự báo vẫn là dựa trên mực nước sông phụ thuộc vào nhiều yếu tố như mưa, bay hơi, khí hậu, điều kiện địa hình, địa chất địa mạo, dòng chảy ngầm, thảm thực vật... Đó là một quá trình tự nhiên với đầy đủ các tính chất vật lý của nó, biểu hiện phạm trù giữa nguyên nhân và kết quả (Hoàng Thanh Tùng, 2021).

- +) Phương pháp tổng hợp phân tích hệ thống thông tin địa lý (GIS)
 - +) Phương pháp lưu vực tương tự:
 - +) Phương pháp thống kê xác suất và kết hợp với mô hình toán
- Các ứng dụng này, ở nước ta hiện nay các mô hình toán đã được

thiết lập và được sử dụng rộng rãi trong tính toán KTTV, dự báo tác nghiệp, phân tích tính toán diễn biến lòng sông, tính toán chiều cao lũ. Các mô hình toán thống kê thủy văn như mô hình TANK, đường lưu lượng đơn vị SCS, các mô hình tính toán lưu vực như SSARR, mô hình thủy lực... được ứng dụng vào Việt Nam từ những năm 1960. Các đặc trưng lũ của một trận lũ bao gồm đường quá trình lũ $Q \sim t$; đỉnh lũ Q_{max} , tổng lượng lũ W_{max} ; thời gian lũ T , thời gian lũ lên và lũ xuống TL, TX.

1.3.3. Một số bất cập trong dự báo lũ Sông Hồng ở Việt Nam

Hệ thống các trạm quan trắc có nhiệm vụ đảm bảo lưu trữ dữ liệu mực nước, lưu lượng nước chảy, lượng nước mưa Sông Hồng còn mỏng. Trong mùa lũ, các trạm quan trắc tự động không chủ động trong vận hành, mà vẫn cần đến yếu tố con người gây lãng phí nguồn lực, đặc biệt các năm có lũ nhỏ, lũ không đáng kể. Tuy nhiên, trong những năm lũ lớn, nhiều dự báo không thể chính xác vào các thời kỳ tiếp theo, nhiều hạn chế trong việc dự báo và kéo dài đã gây ảnh hưởng đến cấp nước, duy trì dòng chảy và ra quyết định kịp thời trên hệ thống cảnh báo lũ.

Trên cơ sở đó đã xác định cơ sở khoa học và thực tiễn dự báo lũ trên thế giới và Việt Nam hiện nay áp dụng các biện pháp khoa học kỹ thuật mới còn rất hạn chế, mới chỉ có bước đầu sử dụng các công nghệ mới về dữ liệu. Thường các biện pháp dự báo là dự đoán khả năng xuất hiện lũ, độ lớn của lũ sẽ dựa trên các cơ sở, phân tích nguyên nhân, mối quan hệ mưa lũ đặc thù của lưu vực, mối quan hệ giữa các nhân tố khí hậu và sự hình thành lũ thông qua các hàm toán học hoặc phân tích các mối tương quan tuyến tính và phi tuyến. Do tính chất phức tạp của dữ liệu đầu vào, các mối tương quan phi tuyến nhiều trong bối cảnh BĐKH và sự phát triển đô thị hóa nhanh dẫn tới các mối tương quan phi tuyến thay đổi liên tục, gây khó khăn cho việc thu thập và thiết lập mô hình dự báo dẫn tới khả năng dự báo còn kém, thời gian dự báo còn chậm và có sự sai lệch. Từ cơ sở dự báo lũ đề xuất điều chỉnh phân bổ lĩnh vực ra quyết định theo mô hình dự báo trong các thời kỳ lũ nhằm nâng cao hiệu quả ứng phó với bối cảnh BĐKH trên thế giới và Việt Nam nói chung và lưu vực Sông Hồng nói riêng.

1.4. Các nghiên cứu trí tuệ nhân tạo và lịch sử ứng dụng

1.4.1. Trí tuệ nhân tạo và dữ liệu lớn

Theo Hui Lin Ong và Cs., (2022) trí tuệ nhân tạo (Artificial Intelligence - AI) hay trí thông minh nhân tạo là trí tuệ được biểu diễn bởi bất cứ một hệ thống nhân tạo bằng ngôn ngữ lập trình được tạo ra.

Thuật ngữ này thường dùng để nói đến các máy tính có mục đích không nhất định và ngành khoa học nghiên cứu về các lý thuyết và ứng dụng của AI. Dữ liệu lớn - Big data là một thuật ngữ cho việc xử lý một tập hợp dữ liệu rất lớn và phức tạp mà các ứng dụng xử lý dữ liệu truyền thống không xử lý được. Big data bao gồm các thách thức như phân tích, thu thập, giám sát dữ liệu, tìm kiếm, chia sẻ, lưu trữ, truyền nhận, trực quan, truy vấn và tính riêng tư (Hamid Moeeni và cs., 2017). Thuật ngữ này thường chỉ đơn giản đề cập đến việc sử dụng các phân tích dự báo, phân tích hành vi người dùng, hoặc một số phương pháp phân tích dữ liệu tiên tiến khác trích xuất giá trị từ dữ liệu mà ít khi đề cập đến kích thước của bộ dữ liệu.

1.4.2. Lịch sử ứng dụng trí tuệ nhân tạo vào cuộc sống và khoa học

AI và dữ liệu lớn được phát triển từ những năm 1960 của thế kỉ 20. Hiện nay trong thế kỷ 21, thế giới đang chuyển mình nhanh chóng trong những các nghiên cứu về AI trong nhiều lĩnh vực khác nhau. AI chia thành hai trường phái tư duy: AI truyền thống và AI tính toán dự báo (Hui Lin Ong và Cs., 2022). Các bài toán điển hình được áp dụng các phương pháp AI gồm:

- Dự báo, Dự báo dựa trên dữ liệu;
- Xử lý ngôn ngữ tự nhiên, dịch tự động (dịch máy) và Chatterbot;
- Điều khiển phi tuyến và Robotics;
- Computer vision, Thực tại ảo và Xử lý ảnh;
- Lý thuyết trò chơi mô phỏng và Lập kế hoạch (Strategic planning);

Trò chơi AI và Computer game bot...

1.4.3. Ứng dụng trí tuệ nhân tạo vào khoa học môi trường và dự báo lũ lụt

Với những nhược điểm của mô hình dự báo cũ, hiện nay việc áp dụng mô hình mạng thần kinh nhân tạo Artificial Neural Network (ANN) mới bằng AI đã khắc phục các nhược điểm đó, thực hiện phân tích chuỗi thời gian và sự tương quan với nhiều yếu tố tác động tới dòng chảy, được sử dụng dự báo dài hạn dòng chảy rất phổ biến. Nhiều nghiên cứu cho thấy kết quả các mô hình ANN đã mang lại nhiều kết quả tích cực, có độ chính xác cao hơn các mô hình truyền thống khác (Vrijling, J.K., 2019).

Hướng nghiên cứu của luận án này là một trong những kỹ thuật hiện đại LSTM-AI và RNN-AI là sử dụng phương pháp dự báo của những quan sát mà sự mất mát không xảy ra. Phương pháp áp dụng trong các trường hợp khi mà số lượng quan sát có mất mát dữ liệu là nhỏ. Sau xử

lý dữ liệu thô, các thông tin dữ liệu được khôi phục sau mất mát là đúng với dữ liệu thực tế ban đầu thu được từ trạm đo thực tế từ đó tiến hành chạy mô hình các tập dữ liệu để dự báo kết quả và đưa ra các quyết định dựa trên các kinh nghiệm của AI và chuyên gia.

1.5. Bài toán mô hình trí tuệ nhân tạo trong dự báo lũ sông

1.5.1. Cơ sở dữ liệu cho mô hình dự báo lũ bằng trí tuệ nhân tạo

Dữ liệu quan trắc KTTV trên lưu vực Sông Hồng được hình thành từ khá sớm, độ dài chuỗi số liệu lại không đồng nhất phổ biến

1.5.2. Ngôn Ngữ lập trình Python

Ngôn ngữ lập trình Python là một trong những ngôn ngữ lập trình được ứng dụng phổ biến nhất hiện nay. Được ứng dụng trong ngành công nghệ thông tin, lập trình AI mang lại hiệu quả tốt.

1.5.3. Mô hình dự báo lũ sông bằng Trí tuệ nhân tạo và đề xuất mô hình cây ra quyết định

Mô hình AI/ML sẽ được học và nghiên cứu ứng dụng kết hợp với các phân tích về sự hình thành nên nước lũ tiềm năng, các quan hệ về quản lý hồ chứa, lượng mưa do các hình thể thời tiết cực đoạn như mưa lớn bất thường có thể gây lũ, từ đó dự báo mực nước lũ thông qua dữ liệu mực nước đã được thu thập thể hiện các mối tương quan này.

CHƯƠNG 2: ĐỐI TƯỢNG, NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1 ĐỐI TƯỢNG VÀ PHẠM VI LUẬN ÁN

2.1.1 Đối tượng nghiên cứu

Đối tượng nghiên cứu của luận án là Mực nước lũ trên Sông Hồng, đặc điểm mực nước tại trạm thủy văn Sơn Tây - Hà Nội.

Dựa trên cơ sở số liệu thu thập từ 29/11/2010 đến 13/02/2020. Trong Luận án này, đối tượng là các giá trị mực nước sông, mực nước lũ hay các cao trình mực nước tại trạm thủy văn Sơn Tây cao hơn mực nước Báo động 1 tại trạm Hà Nội, đo được gọi là đại diện cho những trận lũ sông.

2.1.2 Phạm vi nghiên cứu

- Luận án giới hạn nghiên cứu là: đoạn Sông Hồng đoạn sông từ trạm thủy văn Sơn Tây (74162) - tới trạm Hà Nội.

- Đối với đặc điểm mực nước: luận án đi sâu phân tích là các giá trị mực nước cao hơn mực nước trung bình năm, các mực nước sắp xi hoặc cao mực nước báo động 1 tại trạm Hà Nội, trong luận án được gọi

là đại diện cho những trận lũ.

2.2. Thời gian và địa điểm

2.2.1. Thời gian

Nghiên cứu được tiến hành từ năm 2019-2023 và số liệu tính toán có liên quan từ năm 2010 đến 2020.

2.2.2. Địa điểm nghiên cứu

Trạm thủy văn Sơn Tây - Hà Nội. Đại diện cho trạm phía sau điểm hợp lưu của 3 con Sông Đà, Sông Hồng, Sông Lô phía thượng lưu và hạ lưu là Sông Hồng chảy qua Hà Nội và đồng bằng châu thổ, trước khi đổ ra vịnh bắc bộ.

2.3. Nội dung nghiên cứu

- Xác định điều kiện tự nhiên, kinh tế xã hội từ đó xây dựng cơ sở khoa học đánh giá đặc điểm tính chất mực nước và lũ Sông Hồng đoạn chảy qua trạm thủy văn Sơn Tây và đưa ra giải pháp ứng dụng của AI để dự báo lũ Sông Hồng trong bối cảnh BĐKH.

- Xây dựng được mô hình hệ thống dự báo, hỗ trợ dự báo, ứng dụng được mô hình AI để dự báo lũ sông và nhận diện hiện tượng lũ lớn, nguy hiểm trong bối cảnh BĐKH.

- Đánh giá độ tin cậy của kết quả dự báo và so sánh được hệ thống hỗ trợ dự báo lũ sông qua mô hình AI để dự báo lũ với các phương pháp truyền thống dự báo hiện tượng lũ hiện nay. Đề xuất mô hình ra quyết định trong dự báo lũ phù hợp với điều kiện Việt Nam.

- Kiến nghị một số giải pháp hướng tới dự báo lũ sông và ra quyết định bằng AI trong bối cảnh BĐKH hiện nay.

2.4. Phương pháp nghiên cứu

2.4.1. Phương pháp nghiên cứu kế thừa và thống kê tài liệu thứ cấp

Thống kê, tổng hợp tài liệu, phân tích, đánh giá các mối liên hệ giữa các dữ liệu liên quan. Tổng hợp các kết quả sẽ phục vụ nghiên cứu dự báo lũ lớn tới đoạn sông chảy qua trạm thủy văn Sơn Tây – Hà Nội.

2.4.2. Phương pháp chuyên gia

Quá trình thiết lập bài toán và giải quyết vấn đề dự báo lũ, tham vấn các ý kiến nhận xét, đánh giá, tư vấn của các chuyên gia có kinh nghiệm trong lĩnh vực nghiên cứu đã được tham khảo, lựa chọn tiếp thu.

2.4.3. Phương pháp so sánh và đánh giá lựa chọn

Phương pháp này dùng để so sánh các nguồn dữ liệu theo dòng thời

gian hoặc so sánh nguồn thông tin giữa các nhóm đối tượng cung cấp để phân tích, nhận định vấn đề khi phân tích, đánh giá các kết quả nghiên cứu, kinh nghiệm trong lĩnh vực nghiên cứu đã được tham khảo, lựa chọn.

2.4.4. Phương pháp xử lý và giải đoán ảnh bằng hệ thống thông tin địa lý

Luận án đã vận dụng các định nghĩa và các kết quả thu thập về địa hình, khí hậu, lượng mưa, sử dụng phương pháp phân tích GIS để xác định các thay đổi địa hình và nhân tố khí hậu nhỏ hơn trong khu vực nghiên cứu.

2.4.5. Phương pháp mô hình hóa thực nghiệm trí tuệ nhân tạo

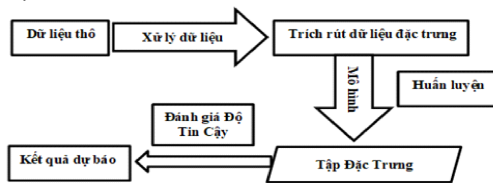
2.4.5.1. Thiết lập và đánh giá tính chất bộ Cơ sở dữ liệu về dự liệu mực nước Sông Hồng qua trạm thủy văn Sơn Tây.

Phương pháp xây dựng mô hình xử lý cơ sở dữ liệu, chuẩn hóa cơ sở dữ liệu mực nước

2.4.5.2. Phương pháp xây dựng mô hình AI dự báo lũ Sông Hồng

Phương pháp xây dựng mô hình LSTM-AI và LSTM-AI

2.4.5.3. Kiểm tra đặc trưng dữ liệu, trích rút cơ sở dữ liệu huấn luyện, Huấn luyện dự báo.



Hình 2.3. Sơ đồ quá trình huấn luyện

Độ tin cậy của các mô hình dự đoán dùng để đánh giá chất lượng dự báo (có thể là của một mô hình) (Murphy, 1993) đã đưa ra 3 khái niệm để chỉ mức độ “tốt” của một dự báo

2.4.5.4. Trích rút kết quả CSDL huấn luyện, dự báo và đánh giá độ tin cậy

Độ tin cậy của các mô hình dự đoán dùng để đánh giá chất lượng dự báo (có thể là của một hoặc nhiều mô hình) Murphy (1993) đã đưa ra 3 khái niệm để chỉ mức độ “tốt” của một dự báo

2.4.6. Phương pháp ra quyết định thống kê bằng mô hình hóa AI trong quá trình dự báo lũ sông

Các phương pháp ra quyết định thống kê có thể sử dụng trong mô

hình ML như cây quyết định theo lý thuyết quyết định Bayes; Mô hình ML có giám sát sử dụng thuật toán học hồi quy logic gồm: cây quyết định và phân lớp dựa trên tập huấn luyện. Thuật toán Cây quyết định (Rokach, L., 2015).

2.5. Các thành phần và cấu trúc mô hình

Phần cứng, phần mềm, cơ sở dữ liệu, bộ code mô hình và chuyên gia.

2.6. Phương pháp so sánh và đánh giá tổng hợp

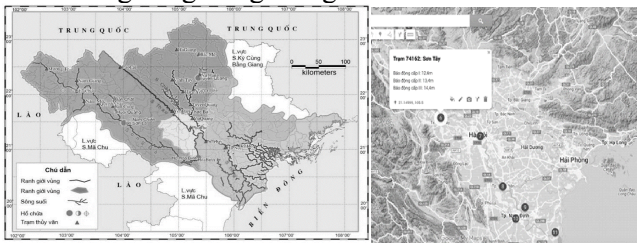
Phương pháp so sánh trong luận án được thể hiện qua các cách đối chiếu các kết quả, để thấy rõ tính vượt trội, các điểm giống và khác nhau của kết quả. Phương pháp nghiên cứu phân tích đánh giá tổng hợp là quá trình phân tích đánh giá kỹ lưỡng các vấn đề đưa ra và tổng hợp ngắn gọn lại với các nội dung chính của luận án.

CHƯƠNG 3: KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

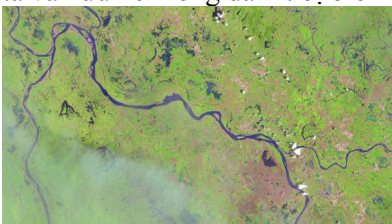
3.1. Đặc điểm mực nước và lũ Sông Hồng đoạn chảy qua thủ đô Hà Nội

3.1.1. Đặc điểm và vị trí địa lý Sông Hồng đoạn chảy qua thủ đô Hà Nội

Trạm thủy văn Sơn tây (74162) – Hà Nội là trạm thủy văn hết sức quan trọng, Trạm là đầu mối đoạn sông Hồng từ Sơn Tây đến Hà Nội - Thượng Cát nằm ở hạ lưu nơi hợp lưu của ba sông là Sông Đà, Sông Thao và Sông Lô, toàn bộ dòng chảy hình thành trên ba nhánh này đều đổ về đồng bằng Sông Hồng.



Qua phân tích bằng GIS - Ảnh Landsat cho thấy Khu vực lòng dẫn Sông Hồng đoạn từ trạm thủy văn Sơn Tây đến qua cổng Cẩm Đình trạm Long Biên, Hà Nội. Dòng chảy tương đối phức tạp, nằm ở giữa bờ tả và hữu nên lòng dẫn ít bị biến động hơn.



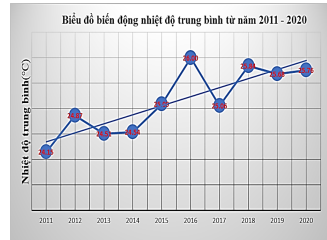
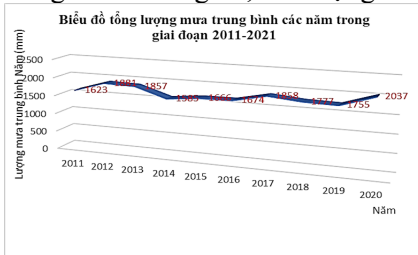
hơn, chỉ ở mức khoảng từ $0,25 \div 0,5$ công trình chỉnh trị ổn định luồng lạch

do ngành giao thông xây dựng, đặc biệt là các cụm kè chỉnh trị thuộc Dự án: Phát triển giao thông vùng đồng bằng Bắc Bộ, lòng dẫn Sông Hồng hiện nay vẫn đang tiếp tục biến đổi và chưa thể ổn định trong khoảng một vài năm tới.

3.1.2. Đặc điểm biến đổi khí hậu, khí tượng và lượng mưa tác động tại trạm thủy văn Sơn Tây

Nhiệt độ: Khu vực nghiên cứu nằm ở vùng cận nhiệt đới phía bắc Việt Nam. Khu vực này được đặc trưng bởi tính thời vụ riêng biệt, bao gồm "Mùa mưa" thường là từ tháng 4 đến tháng 9 và mùa khô thường là từ tháng 11 đến tháng 2 năm sau, với khoảng thời gian mùa xuân và mùa thu ít khác biệt.

Lượng mưa: Tổng lượng mưa hàng năm cho khu vực tại trạm thủy văn Sơn Tây trong năm 2022 là 2.037 mm, chủ yếu được diễn ra từ tháng 6 đến tháng 10, với lượng mưa tối đa là 600 mm vào tháng 10.



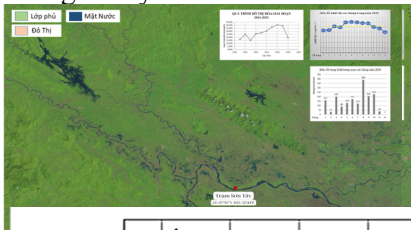
Các kết quả cho thấy rằng trong giai đoạn 10 năm, tăng trưởng kinh tế vào mức cao, hầu hết đều trên 5% một năm. Kết quả này cũng thể hiện môi trường khu vực nghiên cứu mà một môi trường đang phát triển, độ ổn định về cơ sở hạ tầng, hệ thống giao thông, khu dân cư chưa ổn định.

3.1.3. Đánh giá các yếu tố ảnh hưởng tới mực nước và lũ Sông Hồng đoạn chảy qua trạm thủy văn Sơn Tây

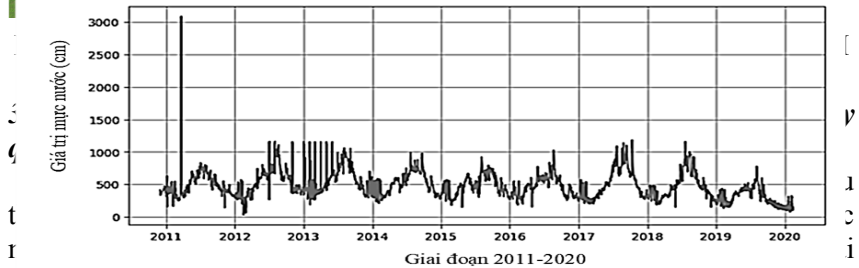
Dữ liệu lũ từ năm giai đoạn 2011 - 2020 được sử dụng để phân tích tương tác giữa sự gia tăng đóng góp lũ trên Sông Thao và mực nước hạ lưu Sông Hồng theo cấp mực nước khác nhau.

Kết quả phân tích cho thấy rằng trong giai đoạn 2011-2020 như sau: kết quả phân tích ảnh bằng công cụ LANDSAT trên GIS, cho thấy về lòng dẫn Sông Hồng không thay đổi nhiều, ít biến động trong những

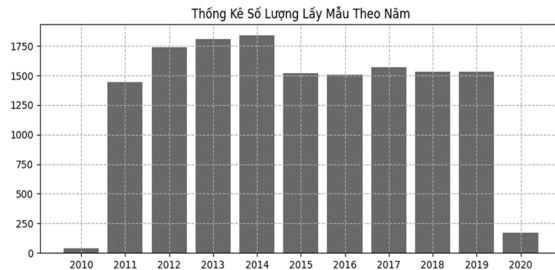
năm gần đây.



giá trị từ bộ dữ liệu thực đo. Để đánh giá thủy văn Sơn Tây, sử dụng công cụ trực quan dữ liệu thu được tại hình 3.2.1.1. Thủy văn Sơn Tây là phù hợp với yêu cầu.



được làm sạch để loại bỏ các điểm bất thường, các giá trị thiếu, chuẩn hóa về dạng chuỗi thời gian.



nước thu thập được tại năm từ 31/10/2010 đến 31/10/2019 là tương đối đầy đủ.

cần được xử lý theo:

- Chuẩn hóa về cơ sở dữ liệu dạng chuỗi thời gian

3.1.5. Phân tích tính chất lũ Sông Hồng đoạn chảy qua trạm thủy văn Sơn Tây.

Dữ liệu mực nước tại trạm thủy văn Sơn Tây cho thấy trên Sông Hồng giai đoạn 2010-2020 cho thấy hiện nay chủ yếu lũ trên Sông Hồng thường là lũ theo mùa, lũ do điều tiết hồ chứa và tần suất có xuất hiện lũ lớn bất thường là rất ít trong năm 2020.

3.2. Kết quả xây dựng Cơ sở dữ liệu chuẩn hóa cho dự báo lũ sông bằng trí tuệ nhân tạo.

3.2.1. Kết quả xây dựng cơ sở dữ liệu huấn luyện

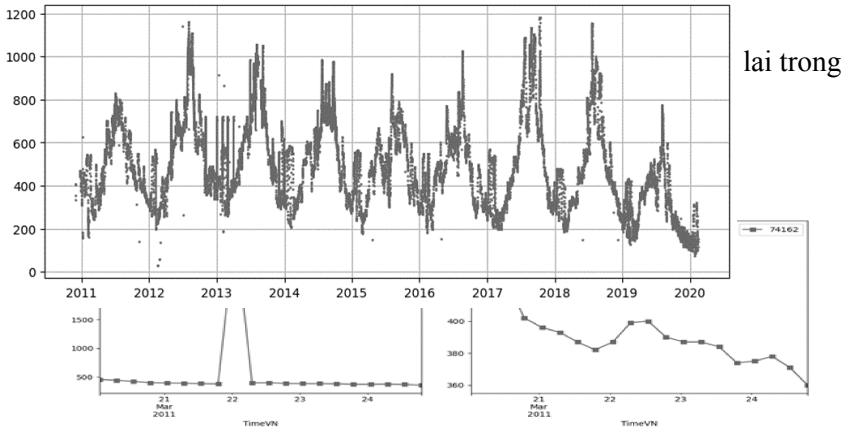
3.2.1.1. Kết quả xử lý dữ liệu Ngoại Lai và Sai số

Dữ liệu thực tế thu thập được xử lý các giá trị ngoại lai và sai số

theo hai phương pháp xử lý hiệu quả để phát hiện giá trị cực trị bao gồm Z-Scores và đồ thị Box-plot.

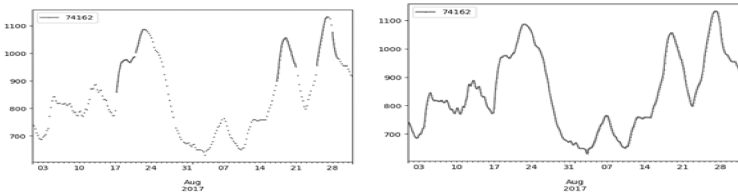
Do nhiều yếu tố chủ quan và khách quan, dẫn đến quá trình ghi nhận giá trị và gửi số liệu quan trắc về trung tâm bị sai sót, nhầm lẫn, mất mát so với giá trị thực tế. Đây là các dữ liệu được ghi nhận và lưu trữ theo thời gian, nhưng lại không phải là dữ liệu chuỗi thời gian.

Kết quả xử lý dữ liệu Ngoại Lai và Sai số



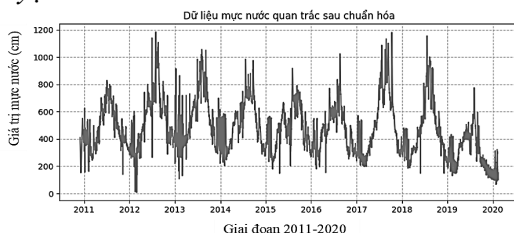
Dữ liệu thiếu, được bổ sung bằng cách thiết lập khoảng thời gian kiểm tra các dữ liệu thiếu trong khung 3h/1 lần quan trắc. Lọc các vị trí thiếu dữ liệu trong chuỗi thời gian của toàn bộ dữ liệu, gán và tạo một dataframe mới với index là các bước timestep thiếu và giá trị None.

3.2.1.2. Xử lý dữ liệu thiếu, mất mát



Kết quả thực hiện việc lọc các mốc thời gian lấy mẫu còn thiếu trong tập dữ liệu tương ứng với khoảng thời gian $t = 3h$ vào các thời điểm 1h, 4h, 7h, 10h, 13h, 16h, 19h, 22h trong ngày.

3.2.1.3. Chuẩn hóa về dữ liệu dạng chuỗi thời gian, tạo bộ dữ liệu huấn luyện



y đủ đã được chuẩn hóa

3.2.2. Đánh giá so sánh cơ sở dữ liệu mực nước thực tế và cơ sở dữ liệu mực nước sau xử lý trí tuệ nhân tạo

Dữ liệu mực nước thu thập được đều là các dữ liệu thô, cần phải được chuẩn hóa và làm sạch để loại bỏ được các điểm ngoại lai ra khỏi tập dữ liệu, các điểm ngoại lai có ảnh hưởng rất lớn tới độ chính xác của các mô hình dự đoán, dự báo.

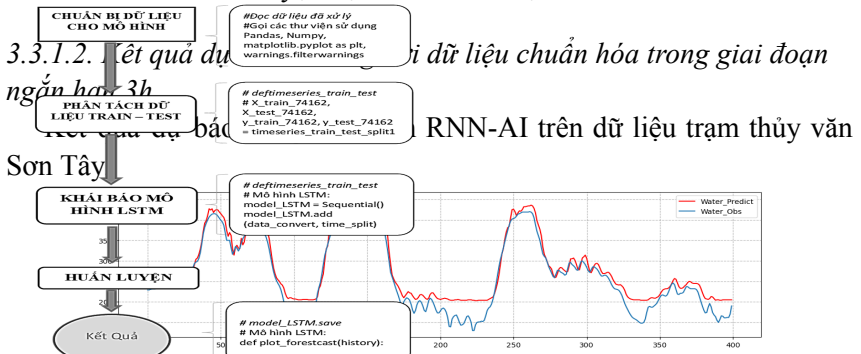
Chính vì vậy nếu không có phương pháp áp dụng AI này, việc phục hồi dữ liệu quan trắc thủy văn, và đảm bảo tính liên tục, có cơ sở thực tiễn là hết sức khó khăn, từ bộ dữ liệu được khắc phục sẽ cải thiện đáng kể khả năng dự báo và dự đoán cho các mô hình được áp dụng trong luận án và các mô hình khác về sau.

3.3. Kết quả huấn luyện mô hình dự báo lũ sông với cơ sở dữ liệu chuẩn hóa

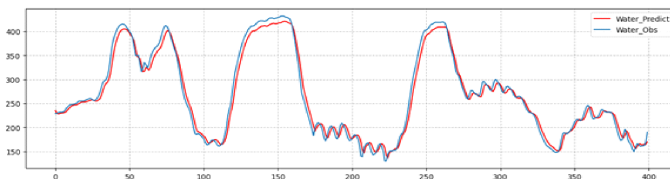
3.3.1. Kết quả huấn luyện mô hình dự báo với cơ sở dữ liệu chuẩn hóa trong giai đoạn ngắn hạn

3.3.1.1. Xây dựng mô hình dự báo với cơ sở dữ liệu chuẩn hóa trong giai đoạn ngắn hạn 3 giờ

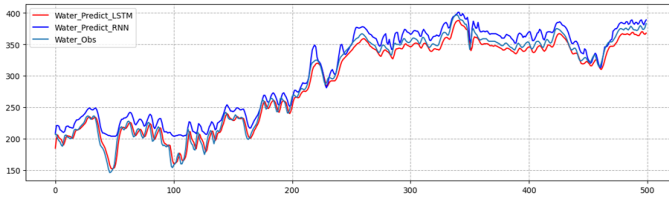
Trên cơ sở dữ liệu đầu vào đã được xử lý, trích rút đặc trưng, các mô hình ML, RNN-AI và LSTM-AI sẽ thực hiện các bước cấu hình, tham số và triển khai huấn luyện dự báo trên dữ liệu.



Kết quả dự báo bằng mô hình LSTM-AI trên dữ liệu trạm thủy văn Sơn Tây



Các kết quả này cho thấy độ chính xác của dự báo tương đối cao, khả năng ứng dụng của việc tối ưu hóa tham số của các thuật toán ML và DL để giải quyết các bài toán dự báo với dữ liệu thủy văn là rất khả quan trong thực tế.



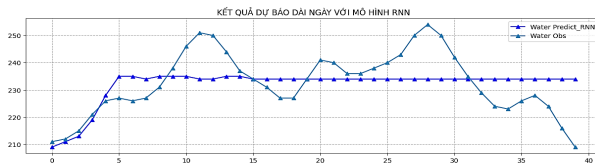
Kết quả dự báo mực nước (H) dựa trên yếu tố mực nước đầu vào của bộ cơ sở dữ liệu Test, phương pháp LSTM-AI này cho kết quả dự báo mực nước khoảng xung quanh giá trị quan trắc thực tế. Mực nước được biểu diễn dưới dạng các con số dự báo đã bao hàm độ không chắc chắn và các yếu tố tương quan tuyến tính, ngẫu nhiên có ảnh hưởng của các điều kiện đầu vào trong đó. Do đó, cần có một công cụ để đánh giá độ tin cậy, không chắc chắn của một mô hình để xử lý dữ liệu không chắc trong mô hình AI dự báo lũ.

3.3.2. Kết quả huấn luyện mô hình dự báo với dữ liệu chuẩn hóa trong giai đoạn trung hạn 5 ngày

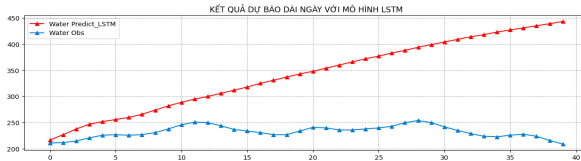
3.3.2.1. Xây dựng mô hình dự báo với dữ liệu chuẩn hóa trong giai đoạn trung hạn 5 ngày

Thay đổi một số thông số trong mô hình ngắn hạn, ta thu được mô hình dài hạn với tham số thời gian $d = 5$, của 2 mô hình dự báo RNN-AI và LSTM-AI trung hạn.

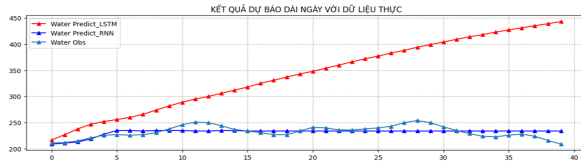
Lựa chọn phương pháp dự báo mực nước trung bình ngày lựa chọn tham số cho mô hình với số ngày tương ứng (day=5) hợp lý tùy vào từng lưu vực. Với thời gian trung hạn 5 ngày, địa hình hạ lưu như Sông Hồng thì việc thoát lũ trong 5 ngày là tương đối khả thi và có cảnh báo sớm. Kết quả dự báo lũ trung hạn với mô hình RNN-AI



Kết quả dự báo lũ trung hạn với mô hình LSTM-AI



So sánh kết quả dự báo lũ trạm thủy văn Sơn Tây bằng LSTM-AI và RNN-AI giai đoạn trung hạn 5 ngày



Lựa chọn phương pháp dự báo mực nước trung bình ngày lựa chọn tham số cho mô hình với số ngày tương ứng (day=5) cho kết quả có độ sai số khá cao. Điều này được giải thích như trong thời gian dự báo cơ sở dữ liệu trong giai đoạn có sự điều chỉnh của hồ chứa thượng lưu và hệ thống kênh thủy lợi, do vậy sai số của mô hình là tương đối lớn.

3.3.2.2. Tối ưu hóa cấu hình, tham số của hệ thống trí tuệ nhân tạo trong mô hình dự báo lũ sông

Các thông số trong mô hình trung và ngắn hạn đều đã được AI học qua dữ liệu mực nước, do đó thu được mô hình ngắn hạn từ 2 mô hình dự báo lũ 3 giờ là RNN-AI và LSTM-AI sau khi được tối ưu huấn luyện nhiều lần với nhiều cơ sở dữ liệu train và test khác nhau. Sử dụng ngôn ngữ Pythons, các thư viện Keras, Tensorflow trên thư viện AI để tối ưu các tham số phục vụ cho bài toán dự báo lũ tại trạm thủy văn Sơn Tây, Hà Nội. Với mô hình RNN-AI, sử dụng phương pháp kiểm chứng chéo 5 lần để tìm ra giá trị tham số tối ưu của mỗi mô hình. Với mô hình LSTM-AI, sử dụng hàm Auto.arima từ package của dự báo theo mô hình RNN.

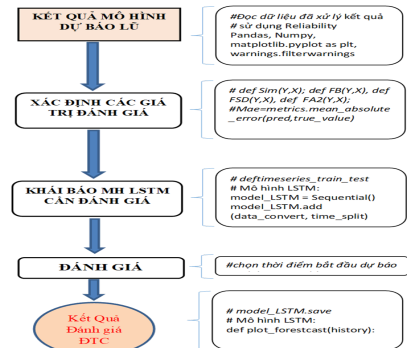
Kết quả tối ưu cấu hình, tham số cho mô hình sử dụng LSTM-AI, Để có thể áp dụng được các phương pháp ML cho bài toán dự báo lũ thì chỉ cần sử dụng dữ liệu mực nước do các yếu tố khác trong thời gian ngắn hạn và trung hạn thay đổi không nhiều (Nhiệt độ, diện tích thấm phủ, hệ số thấm...), thực hiện chuyển dữ liệu thành dữ liệu có giá trị = 1 (NULL), nhiều chiều thành hai chiều mực nước và thời gian, được cập nhật theo chuỗi. Điều này có nghĩa là để dự đoán cho thời điểm t , sẽ dựa vào dữ liệu của p thời điểm trước đó ($x(t-1)$, $x(t-2)$, ..., $x(t-p)$); trong đó, sử dụng giá trị $p=8$ (24h/3h) là phù hợp với mức yêu cầu của

cảnh báo lũ trên Sông Hồng, đảm bảo thời gian tiêu thoát lũ, chống ngập.

3.4. Đánh giá độ tin cậy của mô hình dự báo lũ sông bằng trí tuệ nhân tạo

3.4.1. Xây dựng các hàm và chỉ số đánh giá độ tin cậy

Xây dựng các hàm xác định độ tin cậy của của mô hình dự báo lũ sông bằng AI bằng các hàm và chỉ số độ tin cậy. Để đánh giá độ tin cậy của mô hình, thực hiện: Xây dựng mô hình DL (Deep learning), sử dụng giải thuật LSTM-AI để dự báo lũ. Mô hình ML được sử dụng với ngôn ngữ lập trình Python và sử dụng các thư viện cải thiện tốc độ tính toán của GoogleColab là PyTorch, TensorFlow, Keras...



3.4.2. Kết quả xác định độ tin cậy và dự đoán lũ tại trạm thủy văn Sơn Tây

Độ tin cậy dự báo của mô hình LSTM-AI cao hơn mô hình RNN-AI, khả năng ứng dụng của mô hình LSTM-AI lại một số mô hình khác để kiểm tra đối chứng dự báo lũ theo chuỗi thời gian là có kết quả rất tốt.

3.5. Đánh giá so sánh ưu nhược điểm của mô hình trí tuệ nhân tạo AI/ML với một số kết quả thực tiễn

3.5.1. Ưu điểm của trí tuệ nhân tạo vào ứng dụng trong dự báo lũ sông

AI có thể xử lý được khối lượng lớn công việc. Với nguồn dữ liệu lớn và đa dạng hơn, nhiều vấn đề cũng như công nghệ AI có thể đưa ra những dự đoán chính xác hơn con người và một số mô hình toán học khác. Đây là một lợi thế lớn của công nghệ AI hiện nay.

AI sử dụng các công nghệ ML và DL khác nhau, có thể tổng hợp và xử lý để có thể lấy dữ liệu này trong thời gian ngắn và nhanh chóng đưa ra các kết quả, có thể so sánh và biến đổi cập nhật nếu dữ liệu đó có sự thay đổi bất thường, các thông tin bất thường này được phân tích và ML chi tiết, được công cụ tính trong những lần dự báo tiếp theo.

Một trong những ưu điểm lớn của AI là khả năng dự báo trong thời gian ngắn, với các loại hình cần dự báo nhanh như thảm họa, có tác hại

rộng lớn, thì mô hình AI tỏ ra rất hiệu quả. Trong nghiên cứu thể hiện, thời gian chạy mô hình thử nghiệm kéo dài xấp xỉ 2-3 phút trong khi dự báo có thể trước thời gian từ 3 giờ đến 5 ngày cho kết quả tốt.

3.5.2. Hạn chế của trí tuệ nhân tạo vào ứng dụng trong dự báo lũ sông

Tuy vậy, AI vẫn có nhược điểm là nó không thể học cách suy nghĩ sáng tạo. AI có khả năng học hỏi theo thời gian với dữ liệu được cung cấp trước và kinh nghiệm trong quá khứ, nhưng không thể sáng tạo trong cách tiếp cận nào đó của nó. Do vậy, một số kịch bản có sự trùng lặp về các yếu tố gây ra biến động lớn sẽ làm khả năng giải thích và dự báo của AI kém đi. Việc sử dụng AI để tham khảo và so sánh trong một thời gian kiểm chứng có thể sẽ là khả năng cao cho các kết quả theo mô hình hộp đen (Black-box).

3.6. Mô hình Hệ thống cây ra quyết định cảnh báo lũ dựa trên mô hình trí tuệ nhân tạo

3.6.1. Kế hoạch, tổ chức và ứng dụng AI vào quản lý tài nguyên nước

Sau các kết quả huấn luyện mô hình, bộ dữ liệu kết quả dự báo thu được, tiến hành đọc, chuẩn hóa dữ liệu, phân chia dữ liệu thành các giá trị dự báo nguy hiểm, kiểm thử với các giá trị trong lịch sử và mức giá trị quy định để thực hiện việc ra quyết định cảnh báo nhận diện lũ bằng phương pháp cây quyết định.

Kết quả thực nghiệm cho thấy thuật toán ra quyết định được cập nhật liên tục từ bộ dữ liệu, cũng được cải tiến cho kết quả khả quan và có thể áp dụng trong thực tế. Kết quả sử dụng phương pháp cây quyết định (CQĐ) cho kết quả ra quyết định dự báo tốc độ lớn nhất trong dự báo khá chính xác, với hệ số $cc = 0.6$. Kết quả ra quyết định dự báo lũ xuất hiện hoặc lũ cảnh báo tại thời điểm huấn luyện chính xác hơn so với kết quả dự báo khác.



Hình 3.32. Mô hình cây ra quyết định bằng trí tuệ nhân tạo

3.6.2. Một số giải pháp ứng dụng AI trong khoa học môi trường và dự báo lũ

Đối với Việt Nam hiện nay, việc áp dụng AI vào khoa học môi

trường cũng cần phải có từng bước, từ lựa chọn mô hình AI phù hợp, đến lĩnh vực được thử nghiệm áp dụng, sau đó là nhân rộng, xây dựng hệ thống bài bản, quy mô. Một số loại hình có thể áp dụng mô hình AI trong khoa học môi trường ngay ở Việt Nam được đề xuất ở hình 3. 33.(Tuan-Dung Hoang và cs., 2022).



Hình 3.34. Mô hình AI thử nghiệm

Trên cơ sở nghiên cứu đặc điểm và mực nước lũ Sông Hồng, các mô hình AI đã được thực hiện việc phân tích các yếu tố ảnh hưởng, xử lý dữ liệu, chuẩn hóa cơ sở dữ liệu. Mô hình cũng đã xem xét các mối quan hệ giữa đặc trưng lũ, mực nước và các nhân tố khác theo thời gian. Đối với thời hạn từ ngắn hạn 3h đến trung hạn 5 ngày, Các mối tương quan cao nhất giữa các nhân tố được thống kê trong trong nhiều năm sẽ được lựa chọn để sử dụng dự báo đặc trưng lũ. Kết quả thực nghiệm dự báo tại trạm thủy văn Sơn Tây với hai mô hình RNN-AI và LSTM-AI trong giai đoạn 10 năm từ năm 2011 đến năm 2020 cho thấy kết quả đáng tin cậy và khả quan.

Mô hình AI với LSTM-AI để hỗ trợ dự báo lũ trên hệ thống Sông Hồng cho kết quả dự báo tốt với độ tin cậy cao. Giá trị Đáng tin cậy hơn các mô hình khác cũng như các phương pháp truyền thống. Mô hình RNN-AI và LSTM-AI là các mô hình cần phát triển thử nghiệm nhân rộng với nhiều con sông khác để dự báo mực nước trên các trạm thủy văn thêm đa dạng cơ sở dữ liệu huấn luyện và học tập của AI.

Với tiến bộ của khoa học kỹ thuật, các kết quả dự báo thu được, cần tiếp tục nghiên cứu tập huấn luyện, thử nghiệm và thực hiện trên nhiều bộ dữ liệu khác, mở ra một hướng nghiên cứu ứng dụng mới, các mô hình ra quyết định cảnh báo nhận diện, dự báo và cảnh báo lũ sớm bằng phương pháp sử dụng AI là phù hợp với xu thế hiện nay, thay thế con người như mô hình cây ra quyết định.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Ở nước ta và trên thế giới, hàng năm lũ lụt thường xuyên xảy ra gây thiệt hại không nhỏ đến dân sinh và kinh tế, những trận lũ lịch sử, lũ chưa từng có gây thiệt hại trầm trọng đến người và của cải vật chất ngày càng lớn hơn. Việc cảnh báo, dự báo và phòng chống lũ lụt là cấp thiết và phải được thực hiện bằng nhiều biện pháp

Tùy thuộc vào trình độ phát triển của khoa học kỹ thuật và đặc điểm hình thành lũ của mỗi nơi mà mỗi quốc gia cần có những biện pháp cụ thể trong cảnh báo, dự báo lũ và ứng phó với lũ khác nhau. Các phương pháp dự báo lũ hiện nay có thể chia làm hai loại: Phương pháp phân tích nguyên nhân hình thành và phương pháp thống kê xác suất. Phát triển và ứng dụng phần mềm, công nghệ thông tin, dự báo ứng phó với BĐKH trong thời kỳ phát triển công nghiệp 4.0 được cho là rất mới.

Luận án đã xác định được một số đặc điểm, tính chất lũ Sông Hồng và cấu trúc cơ sở dữ liệu mực nước Sông Hồng trong giai đoạn 2010-2020 của trạm thủy văn Sơn Tây, tạo cơ sở dữ liệu huấn luyện mới, dự trên các phương pháp mô hình AI mới hiện nay. Đã chỉnh sửa và nâng cấp cơ sở dữ liệu hoàn chỉnh dành cho huấn luyện phục vụ dự báo lũ Sông Hồng trong bối cảnh BĐKH. Các phương pháp này có thể áp dụng với nhiều con sông khác có điều kiện số liệu tương tự.

Các phân tích, đánh giá tổng quan ở trên khẳng định tầm quan trọng của công nghệ AI ngày càng được phát triển mạnh mẽ trên thế giới và trong nước. Luận án đã xác định được cơ sở khoa học và đưa ra giải pháp ứng dụng mô hình LSTM-AI và RNN-AI vào dự báo lũ trong bối cảnh BĐKH ở Sông Hồng, tại trạm thủy văn Sơn Tây, Hà Nội. Kết quả đã được đánh giá độ tin cậy và so sánh cho thấy Mô hình LSTM-AI có hiệu quả cao. Việc ứng dụng mô hình hệ thống AI nhằm dự báo, ứng dụng mô hình AI để nhận diện các hiện tượng lũ trong bối cảnh BĐKH đem lại kết quả dự báo nhanh chóng, tiết kiệm thời gian cho các nhà quản lý, cốt lõi của hệ thống vận hành quản lý tài nguyên nước đồng bộ, hướng tới tự động hóa hoàn toàn bằng mô hình cây ra quyết định.

Sử dụng AI trong khoa học môi trường đã đang và sẽ là đường hướng nghiên cứu trong tương lai thay thế việc cho các cảnh báo dạng mô hình toán thống kê truyền thống trước đây. Ở các nước phát triển, các nhà khoa học và quản lý nhận ra rằng họ cần phải chọn thiết kế và duy trì hệ thống hiện có, thay đổi để đáp ứng với việc cập nhật ứng

dụng AI vào cơ sở hạ tầng hiện có. Bằng việc áp dụng các công nghệ như AI vào các vấn đề của xã hội, các nhà khoa học và quản lý hiện đang xây dựng một nguồn lực khoa học ứng dụng bền vững mới.

2. Kiến nghị

- Hàng năm tiến hành thu thập dữ liệu về: Khí tượng thủy văn, mưa lớn, các quá trình đô thị hóa làm suy giảm diện tích thực vật và lớp phủ, mật độ và tầng diện tích mặt nước hồ chứa; để đánh giá, làm cơ sở dữ liệu cho các nghiên cứu ứng dụng tiếp theo.

- Tiếp tục nghiên cứu mô hình học máy (ML) và học sâu (DL) khác của AI đang ngày càng được cải tiến để ứng dụng vào lĩnh vực khoa học môi trường, tạo tiền đề để kiểm chứng, so sánh và tham khảo với các phương pháp khác, tiết kiệm chi phí hơn và sử dụng được nhiều lần hơn.