

TẠP CHÍ

KHÍ TƯỢNG THỦY VĂN

Journal of Hydro - Meteorology

ISSN 2525 - 2208



TỔNG CỤC KHÍ TƯỢNG THỦY VĂN
Viet Nam Meteorological and Hydrological Administration

Số 756
12-2023



Q. TÔNG BIÊN TẬP
PGS. TS. Đoàn Quang Trí

- | | |
|------------------------------|-----------------------------|
| 1. GS. TS. Trần Hồng Thái | 14. PGS. TS. Mai Văn Khiêm |
| 2. GS. TS. Trần Thực | 15. PGS. TS. Nguyễn Bá Thủy |
| 3. GS. TS. Mai Trọng Nhuận | 16. TS. Tống Ngọc Thanh |
| 4. GS. TS. Phan Văn Tân | 17. TS. Đinh Thái Hưng |
| 5. GS. TS. Nguyễn Kỳ Phùng | 18. TS. Võ Văn Hòa |
| 6. GS. TS. Phan Đình Tuấn | 19. TS. Nguyễn Đắc Đồng |
| 7. GS. TS. Nguyễn Kim Lợi | 20. GS. TS. Kazuo Saito |
| 8. PGS. TS. Nguyễn Văn Thắng | 21. GS. TS. Jun Matsumoto |
| 9. PGS. TS. Dương Văn Khảm | 22. GS. TS. Jaecheol Nam |
| 10. PGS. TS. Dương Hồng Sơn | 23. TS. Keunyong Song |
| 11. TS. Hoàng Đức Cường | 24. TS. Lars Robert Hole |
| 12. TS. Bạch Quang Dũng | 25. TS. Sooyoul Kim |
| 13. PGS. TS. Đoàn Quang Trí | |

Giấy phép xuất bản

Số: 225/GP-BTTTT - Bộ Thông tin Truyền thông cấp ngày 08/6/2015

Tòa soạn

Số 8 Pháo Đài Láng, Đống Đa, Hà Nội
Điện thoại: 024.39364963
Email: Tapchikttv@gmail.com

Chế bản và In tại:

Công ty TNHH Đầu Tư Nông Nghiệp Việt Nam
ĐT: 0243.5624399

Ảnh bìa: Trạm Quan trắc Khí tượng bề mặt Phú Quốc

Giá bán: 40.000 đồng

TẠP CHÍ KHÍ TƯỢNG THỦY VĂN
SỐ 756 - 12/2023

MỤC LỤC

Bài báo khoa học

- 1 Trà Nguyễn Quỳnh Nga, Trần Thị Kim, Ôn Bảo Hạng, Nguyễn Thị Bẩy:** Đánh giá nguy cơ xói lở dưới ảnh hưởng của dòng chảy trước và sau khi có kè mô hàn trên sông Tiền đoạn chảy qua thành phố Sa Đéc, Đồng Tháp
- 14 Nguyễn Thị Bẩy, Trần Thị Kim, Trần Thị Thúy An, Trà Nguyễn Quỳnh Nga:** Đánh giá ảnh hưởng của khai thác cát đến diễn biến đáy đoạn sông Tiền chảy qua tỉnh Vĩnh Long
- 29 Phạm Thị Thanh Hòa, Vũ Ngọc Quang, Lê Thanh Nghị, Đoàn Thị Nam Phương, Nguyễn Minh Hải:** Nghiên cứu khả năng ứng dụng thuật toán Random Forest và ảnh vệ tinh Sentinel-2 trong phân loại lớp phủ mặt đất tỉnh Quảng Bình trên nền tảng Google Colab
- 42 Ngô Thị Xuân, Nguyễn Châu Mỹ Duyên, Bùi Tá Long:** Mô phỏng ô nhiễm bụi PM_{2.5} và phân tích các yếu tố liên quan – Trường hợp tỉnh Cà Mau, Việt Nam
- 59 Nguyễn Quốc Tuấn, Nguyễn Hồng Quân, Nguyễn Thị Hải:** So sánh mức độ ảnh hưởng của một số vật liệu đối với quá trình cố định Asen trong đất bị ô nhiễm
- 70 Nguyễn Việt Hằng, Phạm Duy Huy Bình, Phạm Quang Nam, Nguyễn Bách Tùng, Trịnh Tuấn Long:** Đánh giá tiềm năng năng lượng gió cho phát triển năng lượng tái tạo trên khu vực Biển Đông, Việt Nam
- 83 Bùi Mạnh Hà, Nguyễn Bá Thủy, Phạm Khánh Ngọc, Phạm Văn Tiến:** Kết quả bước đầu thử nghiệm thuật toán XGBoost dự báo nước dâng do bão tại trạm Hòn Dấu
- 95 Vũ Văn Sáng:** Nghiên cứu lựa chọn công nghệ truyền tin không dây năng lượng thấp (LORA/LPWAN) ứng dụng trong truyền dữ liệu trạm khí tượng tự động thuộc Đài Khí tượng Thủy văn khu vực Tây Nguyên

Bài báo khoa học

Đánh giá nguy cơ xói lở dưới ảnh hưởng của dòng chảy trước và sau khi có kè mỏ hàn trên sông Tiền đoạn chảy qua thành phố Sa Đéc, Đồng Tháp

Trà Nguyễn Quỳnh Nga^{1,2*}, Trần Thị Kim³, Ôn Bảo Hạng^{1,2}, Nguyễn Thị Bầy^{1,2}

¹ Trường Đại học Bách khoa Tp.HCM; tnqnga@hcmut.edu.vn;
hangbaoon1609@gmail.com; ntbay@hcmut.edu.vn

² Đại học Quốc gia Thành phố Hồ Chí Minh; tnqnga@hcmut.edu.vn;
hangbaoon1609@gmail.com; ntbay@hcmut.edu.vn

³ Trường Đại học Tài nguyên và môi trường Tp.HCM; ttkim@hcmunre.edu.vn

*Tác giả liên hệ: Email: tnqnga@hcmut.edu.vn; Tel.: +84-347292972

Ban Biên tập nhận bài: 5/9/2023; Ngày phản biện xong: 16/10/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Đồng Tháp ở phía thượng nguồn sông Cửu Long, là nơi thường xuyên xảy ra sạt lở bờ. Bài báo này tập trung nghiên cứu ảnh hưởng của chế độ thủy động lực trước và sau khi có các kè mỏ hàn trên đoạn sông Tiền chảy qua khu vực thành phố Sa Đéc, tỉnh Đồng Tháp. Phương pháp nghiên cứu chính là ứng dụng mô hình thủy lực MIKE21 HD kết hợp với module bùn cát MT, được hiệu chỉnh và kiểm định với số liệu thực đo tại trạm Mỹ Thuận và trạm Cao Lãnh cho kết quả có độ tương thích cao. Mô hình được tính toán theo 2 kịch bản trước và sau khi 7 kè mỏ hàn được xây dựng. Kết quả cho thấy các kè mỏ hàn làm thay đổi luồng dòng chảy. Nhìn chung là mức độ xói giảm ở sát bờ lồm nhưng mức độ xói ở giữa dòng khá đáng kể (0,8 m/6 tháng). Diện tích xói tăng so với khi chưa có kè. Mặc dù các kè mỏ hàn có khả năng giảm xói hiệu quả ở sát bờ lồm (quanh các kè sát bờ lồm có hiện tượng bồi), nhưng ở đầu các kè mỏ hàn cũng bị xói nhiều do dòng chảy mạnh dẫn đến sạt lở kè. Kết quả cũng góp phần giúp các nhà quản lý có thêm góc nhìn về hiệu quả của kè mỏ hàn ở đoạn sông cong này.

Từ khóa: MIKE21; Thủy động lực học dòng chảy; Xói lở; Kè mỏ hàn; Sông Tiền; Sa Đéc.

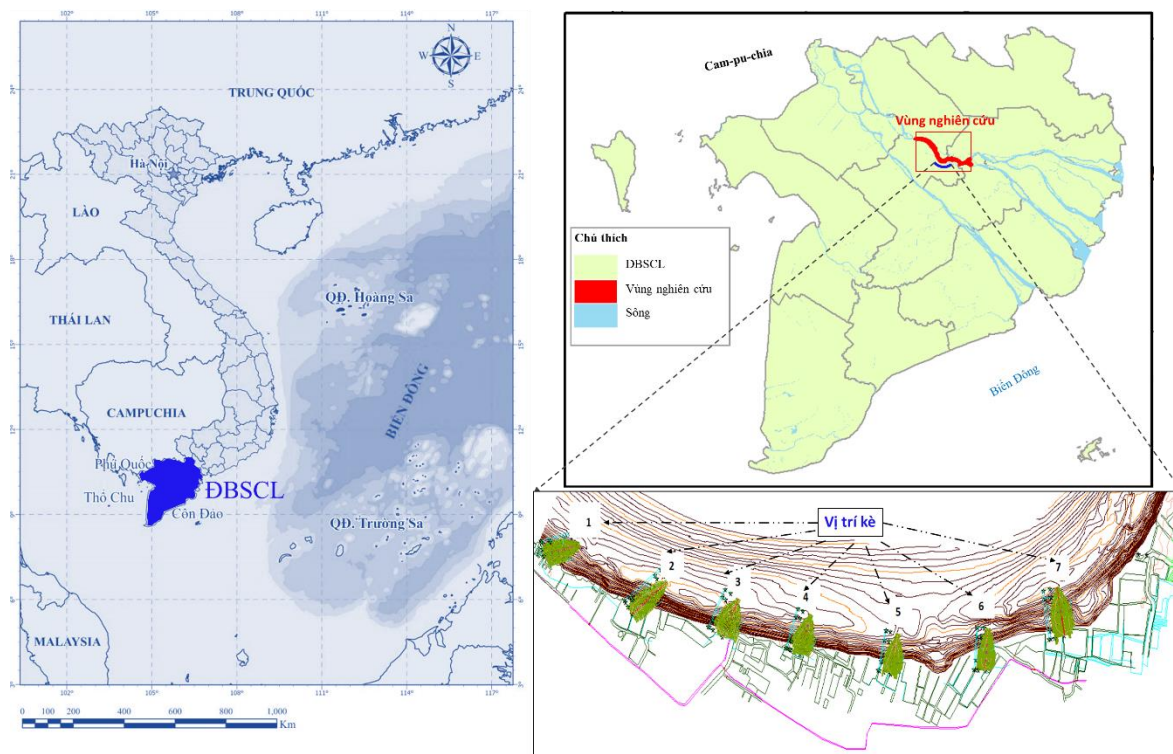
1. Mở đầu

Hiện tượng bồi, xói là kết quả của quá trình tác động qua lại giữa dòng chảy và lòng sông thông qua chuyển động của bùn cát. Diễn biến lòng sông rất phức tạp, chịu ảnh hưởng của nhiều yếu tố như tác dụng xâm thực của dòng chảy; các vận động kiến tạo của địa chất, các điều kiện khí hậu, thời tiết, điều kiện địa hình, địa mạo, lớp phủ thực vật và con người. Xói lòng sông là một trong những nguyên nhân dẫn đến sạt lở bờ. Các nghiên cứu về diễn biến bồi, xói đáy và sạt lở bờ sông đã thu hút được rất nhiều sự quan tâm của các nhà khoa học trên thế giới. Nghiên cứu ứng dụng mô hình 1D kết hợp điều tra thực địa [1] hoặc kết hợp mô hình 1D-2D [2] để tính toán tác động của lũ lụt đến sự nứt vỡ đê hoặc sạt lở kè, từ đó thiết kế được các kè hướng dòng làm giảm tác động của dòng chảy đến sạt lở. Tác giả [3] nghiên cứu phát triển mô hình 2D có sự kết hợp với mô hình vật lý để tính toán và so sánh hiện tượng xói lở bờ sông Old Rhine. Tuy nhiên nghiên cứu này cũng có một số kết quả chưa phù hợp với số liệu đo đạc do có nhiều yếu tố chưa được xem xét trong mô hình 2D. Nghiên cứu [4] kết hợp lý thuyết với thực nghiệm từ mô hình thí nghiệm đã đề xuất chỉnh sửa một vài thông số lý thuyết trong tính toán xói mòn sông để phù hợp với thực nghiệm hơn. Tác giả

[5] nghiên cứu ảnh hưởng của tỷ trọng bùn đến vận tốc khởi động bùn ven biển bằng các thí nghiệm trong phòng thí nghiệm với kênh dài 22 m và mẫu bùn từ cửa sông Huangmaohai. Dựa trên kết quả thực nghiệm, một công thức thực nghiệm về vận tốc khởi động của bùn ven biển với tỷ trọng khác nhau được đưa ra. Nghiên cứu [6] kết hợp tính toán CFD và thí nghiệm đo vận tốc hạt (PIV) để đề xuất được phương trình tính toán xói mòn. Nghiên cứu này đã chứng minh được phương trình rất phù hợp để dự báo xói sâu cục bộ. Nhìn chung, với mong muốn hiểu được bản chất và quy luật của quá trình bồi lắng, xói lở, cho đến nay đã có rất nhiều cách tiếp cận và nhiều phương pháp khác nhau như nghiên cứu lý thuyết và công thức thực nghiệm từ thí nghiệm, phân tích tài liệu, đo đạc hiện trường, mô hình vật lý, mô hình toán.

Trong nước cũng đã có nhiều cách tiếp cận để tìm hiểu các quá trình thủy động lực học, vận chuyển bùn cát, xói lở và bồi tụ, đặc biệt là ở đồng bằng sông Cửu Long (ĐBSCL) như khảo sát thực địa [7–9], mô hình thủy động lực [8, 10–14] và viễn thám [14–17]. Trong số đó, mô hình thủy động lực có khả năng thực hiện được theo nhiều kịch bản, kết hợp được nhiều yếu tố tác động ở quá khứ, hiện tại và tương lai cho kết quả tốt [10, 11, 13]. Có rất nhiều mô hình tính toán được bồi tụ, xói lở được sử dụng rộng rãi hiện nay như: CCHE2D, EFDC ECOMSED, HEC-RAS, TELEMAC, Delft3D, MIKE 21,... trong đó MIKE21 thường được sử dụng rất nhiều vì đã có nhiều tính toán ứng dụng cho các sông ngòi ở nước ta đạt hiệu quả tốt [10, 12, 18–20]. Vì vậy, trong nghiên cứu này, mô hình thủy động lực MIKE 21 được lựa chọn do tính hiệu quả và phổ biến trong nghiên cứu chế độ dòng chảy, vận chuyển trầm tích và diễn biến hình thái ở Việt Nam nói chung và ĐBSCL nói riêng.

Ở ĐBSCL, tình hình sạt lở diễn biến phức tạp trong nhiều năm qua, có trên 665 điểm sạt lở nghiêm trọng [7]. Trong đó, tỉnh Đồng Tháp ở thượng nguồn của ĐBSCL có tỷ lệ xói lở bờ sông cao. Sông Tiền và sông Hậu chảy qua địa phận tỉnh Đồng Tháp với tổng chiều dài khoảng 140 km (110 km sông Tiền và 40 km sông Hậu) mang lại cho tỉnh nhiều lợi ích về kinh tế và xã hội. Tuy nhiên, tình trạng xói lở bờ sông thường xuyên xảy ra làm ảnh hưởng nhiều đến đời sống dân sinh và sự phát triển kinh tế - xã hội của địa phương. Theo thống kê của Chi cục Thủy lợi Đồng Tháp, tổng diện tích đất sạt lở chiếm khoảng 330 ha trên toàn tỉnh, bình quân 22 ha/năm (tính từ năm 2005-2020) [21]. Theo báo cáo thường niên của tỉnh Đồng Tháp, nguyên nhân khiến lòng sông bị thay đổi mạnh và sạt lở bờ sông có thể là do



Hình 1. Vùng nghiên cứu và vị trí kênh mô hàn.

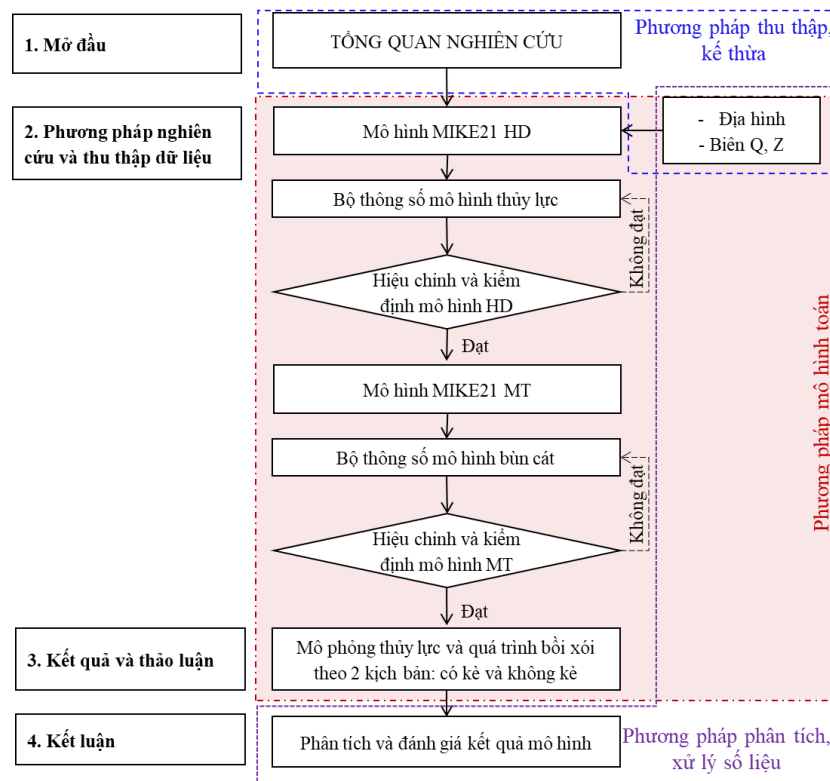
dòng thủy động lực mạnh tương tác với nền đất yếu, hình thái sông cong và có nhiều cồn cát làm phân dòng, ở đây là vùng giáp nước chịu ảnh hưởng của cả dòng chảy lũ ở phía thượng nguồn chảy xuống và dòng triều chảy lên. Với tình hình sạt lở bờ liên tục và nghiêm trọng nơi đây, nhiều giải pháp công trình bảo vệ bờ được xây dựng nhằm giảm thiểu thiệt hại và mất mát [22]. Tuy nhiên, vẫn có nhiều sự cố sạt lở bờ xảy ra ở những đoạn công trình đã được bảo vệ [23–25]. Do đó, mục đích chính của nghiên cứu này là đánh giá tác động của dòng chảy đến xói lở lòng sông trong hai trường hợp trước và sau khi có kè mở hàn cho đoạn sông Tiền qua thành phố Sa Đéc. Kết quả thu được từ nghiên cứu này sẽ góp phần giúp các nhà quản lý đánh giá được hiệu quả của kè mở hàn, đồng thời lựa chọn được phương án bảo vệ bờ phù hợp với khu vực.

Vì vậy, khu vực nghiên cứu được chọn là đoạn sông Tiền qua công trình kè chống sạt lở ở khu vực xã An Hiệp (TP. Sa Đéc), giới hạn từ cầu Cao Lãnh đến cầu Mỹ Thuận được thể hiện bằng vùng màu đỏ như trong Hình 1.

2. Phương pháp nghiên cứu và thu thập dữ liệu

2.1. Phương pháp nghiên cứu

Nghiên cứu này được tóm tắt bằng sơ đồ cấu trúc nghiên cứu như trong Hình 2.



Hình 2. Sơ đồ cấu trúc nghiên cứu.

Trong đó, mô hình thủy động lực MIKE 21 (Viện Thủy lực Đan Mạch - DHI) với module thủy lực (*Hydrodynamic - HD*) và bùn cát (*Mud transport - MT*) được sử dụng kết hợp để tính toán sự biến đổi địa hình đáy khu vực nghiên cứu là phương pháp chính. Các module này giải phương trình nước nông 2 chiều, phương trình chuyển tải phù sa và phương trình liên tục bùn cát đáy [26].

2.2. Các chỉ tiêu đánh giá hiệu chỉnh và kiểm định mô hình

Hiệu chỉnh và kiểm tra các thông số của mô hình đối với khu vực sao cho kết quả tính toán phù hợp với số liệu thực đo tại các trạm. Phương pháp biểu đồ, đồ thị và phương pháp số được

sử dụng trong quá trình hiệu chỉnh và kiểm định. Phương pháp số có sử dụng một số hệ số để đánh giá sai số giữa số liệu mô phỏng và thực đo:

Hệ số Nash-Sutcliffe (NSE), được xác định theo công thức:

$$NSE = 1 - \frac{\sum_{i=1}^n (Y_i^{do} - Y_i^{MIKE})^2}{\sum_{i=1}^n (Y_i^{do} - Y^{TB})^2} \quad (1)$$

Tỉ số độ lệch quan trắc tiêu chuẩn (RSR), được xác định theo công thức:

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_{i=1}^n (Y_i^{do} - Y_i^{MIKE})^2}}{\sqrt{\sum_{i=1}^n (Y_i^{do} - \bar{Y}^{MIKE})^2}} \quad (2)$$

Phần trăm sai số (PBIAS):

$$PBIAS = \frac{\sum_{i=1}^n (Y_i^{do} - Y_i^{MIKE})}{\sum_{i=1}^n (Y^{TB})} \times 100 \quad (3)$$

Trong đó: Y_i^{do} là giá trị quan trắc thứ i ; Y_i^{MIKE} là giá trị mô phỏng thứ i ; Y^{TB} là giá trị quan trắc trung bình; \bar{Y}^{MIKE} là giá trị trung bình mô phỏng.

Bảng 1 là các tiêu chuẩn dùng để đánh giá hiệu quả của mô hình.

Bảng 1. Các tiêu chuẩn đánh giá cho chỉ số NSE, RSR và PBIAS [27].

Hiệu quả mô phỏng	NSE	RSR	PBIAS	
			Dòng chảy	Chất lượng nước
Rất tốt	0,75 - 1,00	0,00 - 0,50	<± 10%	<± 25%
Tốt	0,65 - 0,75	0,50 - 0,60	± 10% - ± 15%	± 25% - ± 40%
Thỏa mãn	0,50 - 0,65	0,60 - 0,70	± 15% - ± 25%	± 40% - ± 70%
Không thỏa mãn	< 0,50	> 0,70	> ± 25%	> ± 70%

2.3. Các kịch bản tính toán

Theo thông tin của UBND tỉnh Đồng Tháp, công trình kè chống sạt lở ở khu vực xã An Hiệp đã khởi công thực hiện năm 2012, công trình có tổng chiều dài 4.500 m gồm 7 mỏ hàn với hơn 1.400 m kè bờ (vị trí như mô tả trong Hình 1). Để đánh giá ảnh hưởng của dòng chảy và mức độ xói lở lòng dẫn quanh các kè mỏ hàn này, 2 kịch bản mô phỏng như sau:

- Kịch bản trước khi có kè - KB1: thời gian tính toán từ 07/2008 đến 12/2008;
- Kịch bản sau khi có kè - KB2: thời gian tính toán từ 07/2017 đến 12/2017.

2.4. Dữ liệu đầu vào mô hình

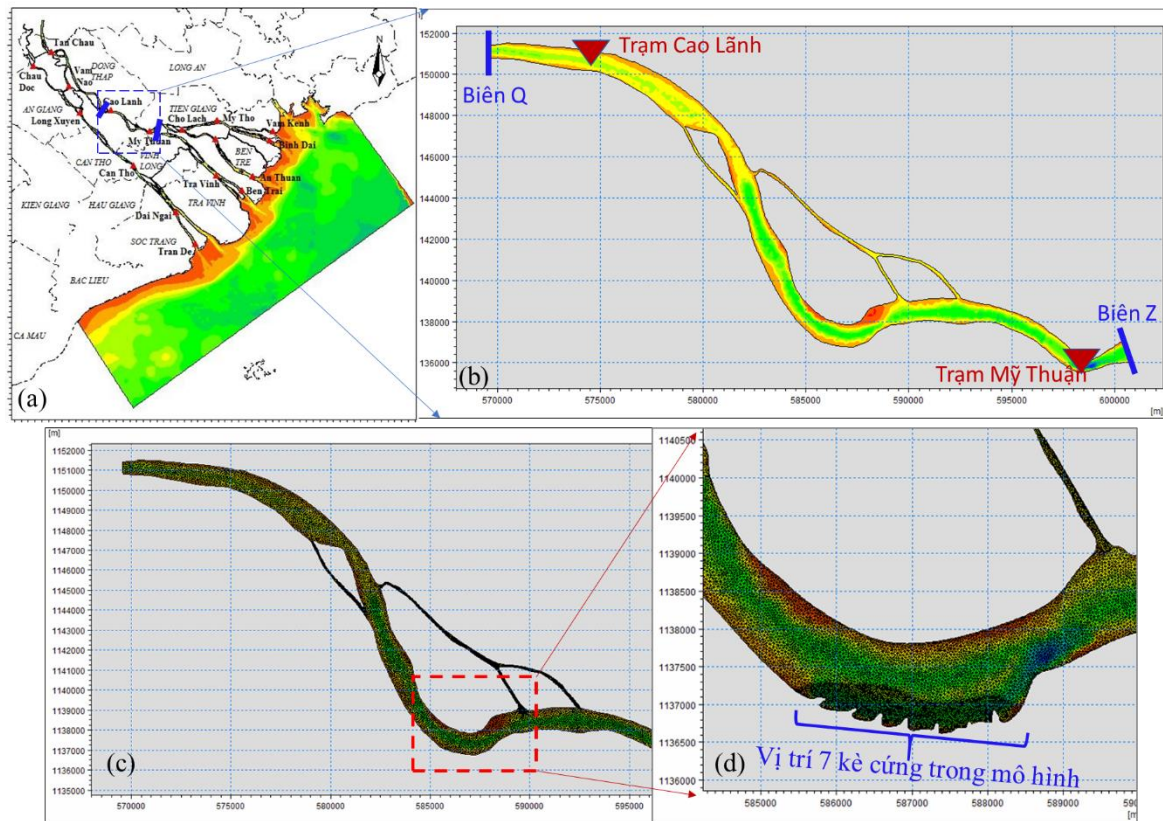
Vùng tính có 2 biên lỏng là biên Q và biên Z. Số liệu từ các biên này được kế thừa từ [10] (mô hình 2D toàn vùng - Hình 3a). Để đảm bảo độ tin cậy thì mô hình được hiệu chỉnh và kiểm định lại với số liệu thực đo từ trạm thủy văn Cao Lãnh và trạm Mỹ Thuận (vị trí các trạm thủy văn trong Hình 3b). Trong đó:

- Biên cho mô hình thủy lực HD: Dữ liệu biên lưu lượng (tại biên Q) - mực nước (tại biên Z) là dữ liệu theo giờ trong thời gian từ 01/07 đến 31/12/2008 (tính toán cho trường hợp không kè - KB1) và từ 01/07 đến 31/12/2017 (tính toán cho trường hợp có kè - KB2).

- Biên cho mô hình chuyển tải bùn cát MT: Dữ liệu biên phù sa lơ lửng TSS theo giờ tại 2 biên Q, biên Z.

Dữ liệu địa hình năm 2013 (Hình 3b) của khu vực nghiên cứu được kế thừa từ đề tài [20].

Lưới tính được xây dựng là lưới tam giác phi cấu trúc gồm 22.541 nút, 40.375 phần tử, góc nhỏ nhất của các phần tử tam giác là 30°, khoảng cách giữa các nút trong vùng tính từ 10 m đến 50m. Lưới tính toán cho trường hợp có kè (Hình 3d) được chia mịn hơn lưới không kè (Hình 3c).



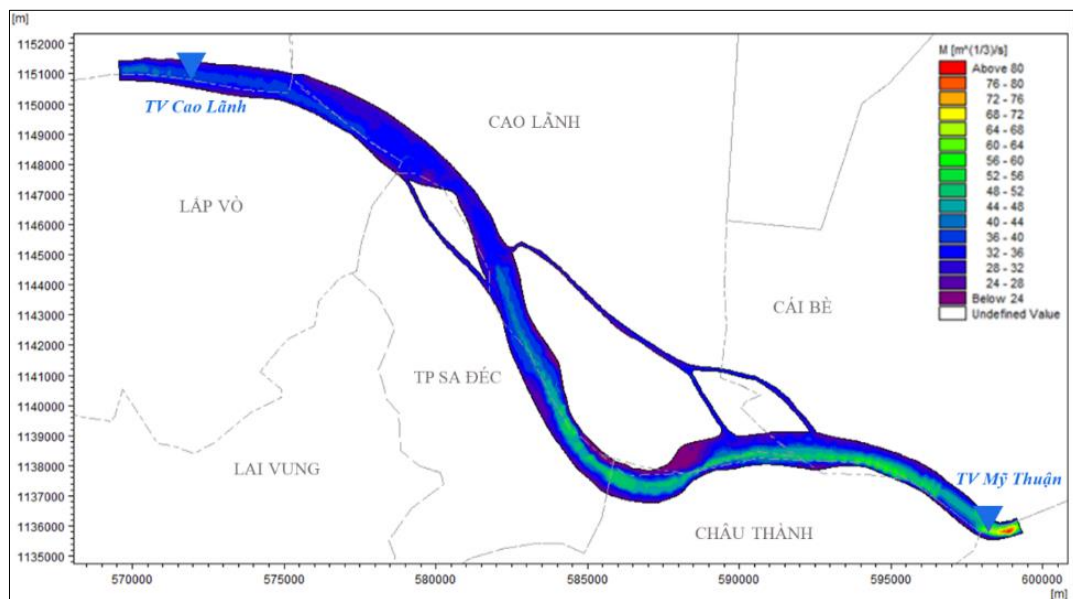
Hình 3. Vùng tính lớn để lấy biên (a) [10]; vị trí biên và địa hình khu vực nghiên cứu (b); lưới tính trường hợp không kè (c) và có kè (d).

2.5. Thông số thiết lập mô hình

Bước thời gian tính toán: 30s.

Điều kiện ban đầu mô hình thủy lực: Mực nước tĩnh, vận tốc bằng 0 m/s.

Hệ số Manning's M dùng để hiệu chỉnh mô hình thủy lực HD. Hệ số M được thay đổi tuyến tính theo độ sâu địa hình đáy sông với M từ 20 đến 85 m^{1/3}/s (Hình 4).



Hình 4. Hệ số Manning M.

Khu vực nghiên cứu có đặc điểm cấu trúc địa chất được mô tả trong Bảng 2.

Bảng 2. Đặc điểm địa chất vùng sông Tiền qua Sa Đéc [20].

Lớp	Độ dày lớp (m)	Phân bố các cấp phối hạt trong từng lớp (%)		
		Hạt cát	Hạt bụi	Hạt sét
Lớp 1	2,4	38,5	47	14,5
Lớp 2	13,6	39,3	44,2	16,5
Lớp 3	14,1	88,7	7	4,3
Lớp 4	3,9	30,2	36,6	33,2

Với đặc điểm như trên, bộ thông số mô hình chuyển tải bùn cát MT như sau:
 - Vận tốc lắng đọng của các cấp hạt: $w_{s1} = 0,003$ m/s; $w_{s2} = 0,04$ m/s; $w_{s3} = 0,091$ m/s.
 - Thông số dùng để hiệu chỉnh mô hình MT là: ứng suất tới hạn bồi τ_{cd} (N/m²); ứng suất tới hạn xói τ_{ce} (N/m²); tốc độ xói E (kg/m²/s); độ nhám đáy k_n (m). Bộ thông số mô hình MT như trong Bảng 3.

Bảng 3. Bộ thông số mô hình MT.

Thông số	Ký hiệu	Đơn vị	Giá trị
Ứng suất tới hạn xói	τ_{ce}	Lớp 1	0,3
		Lớp 2	0,35
		Lớp 3	0,4
		Lớp 4	0,73
Ứng suất tới hạn bồi	τ_{cd}		0,05
			2×10^{-6}
Tốc độ xói	E_0	Lớp 1	5×10^{-6}
		Lớp 2	5×10^{-5}
		Lớp 3	5×10^{-5}
		Lớp 4	$2,5 \times 10^{-5}$
Độ nhám đáy	k_n	m	0,0005

3. Kết quả và thảo luận

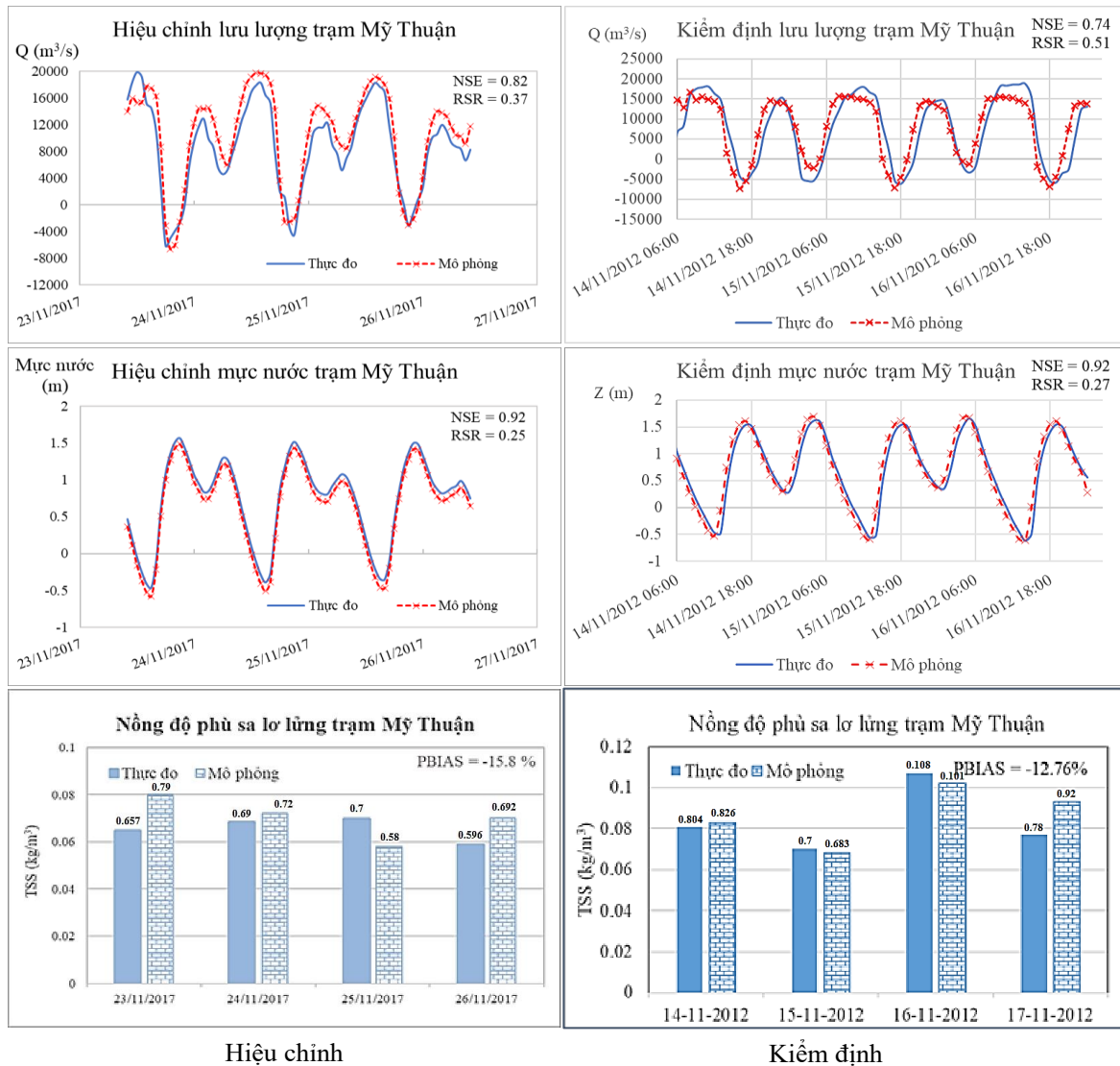
3.1. Kết quả hiệu chỉnh và kiểm định mô hình

Mô hình được hiệu chỉnh trong 3 ngày từ 10h 23/11/2017 đến 10h ngày 26/11/2017 và kiểm định trong 3 ngày sử dụng dữ liệu từ 6h ngày 14/11/2012 đến 0h ngày 17/11/2012. Kết quả hiệu chỉnh và kiểm định mô hình được thể hiện trong Hình 5, Hình 6 và Bảng 4.

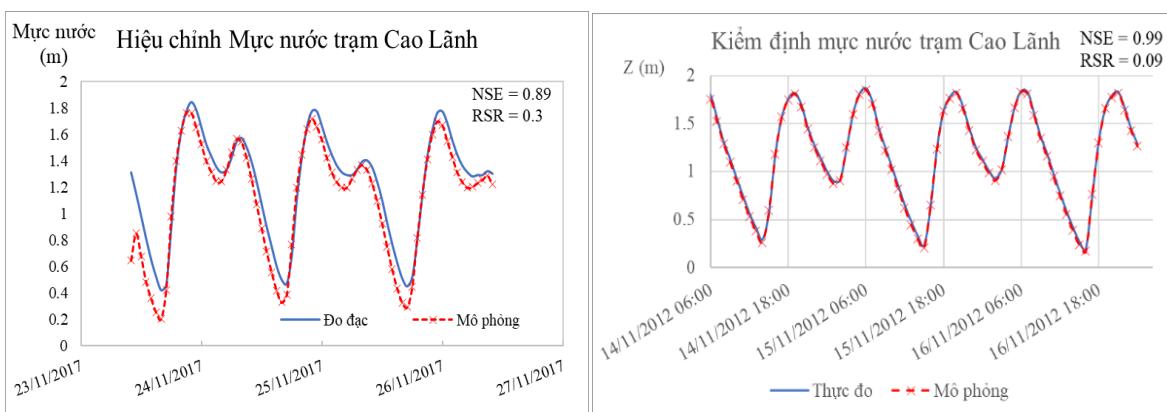
Bảng 4. Kết quả đánh giá mô hình.

STT	Trạm	Hiệu chỉnh			Kiểm định		
		NSE	RSR	PBIAS	NSE	RSR	PBIAS
Mức nước	Cao Lãnh	0,89	0,3	-	0,99	0,09	-
		0,92	0,25	-	0,92	0,27	-
Lưu lượng	Mỹ Thuận	0,82	0,37	-	0,74	0,51	-
TSS		-	-	-15,8	-	-	-12,76

Kết quả hiệu chỉnh và kiểm định mô hình thủy lực tại 2 trạm Cao Lãnh và Mỹ Thuận tương đối tốt. Số liệu mô phỏng so với số liệu thực đo ở các trạm này có tương quan rất tốt với mực nước (NSE > 0,9; RSR < 0,3) và lưu lượng (NSE > 0,75; RSR < 0,5). Riêng với mô hình chuyển tải bùn cát MT chỉ hiệu chỉnh và kiểm định tại trạm Mỹ Thuận do trong vùng tính chỉ có số liệu TSS tại trạm này. Hệ số PBIAS được áp dụng để đánh giá hiệu quả mô hình vì dữ liệu hàm lượng phù sa thực đo không liên tục (TSS đo 8 lần mỗi ngày). Kết quả tính toán và thực đo phù sa trung bình ngày được so sánh với nhau cho sai số phần trăm ở mức tốt (PBIAS < ±15%). Như vậy, mô hình MIKE21 với những thông số đã thiết lập đảm bảo được độ tin cậy để mô phỏng cho đoạn sông này trong thời gian dài.



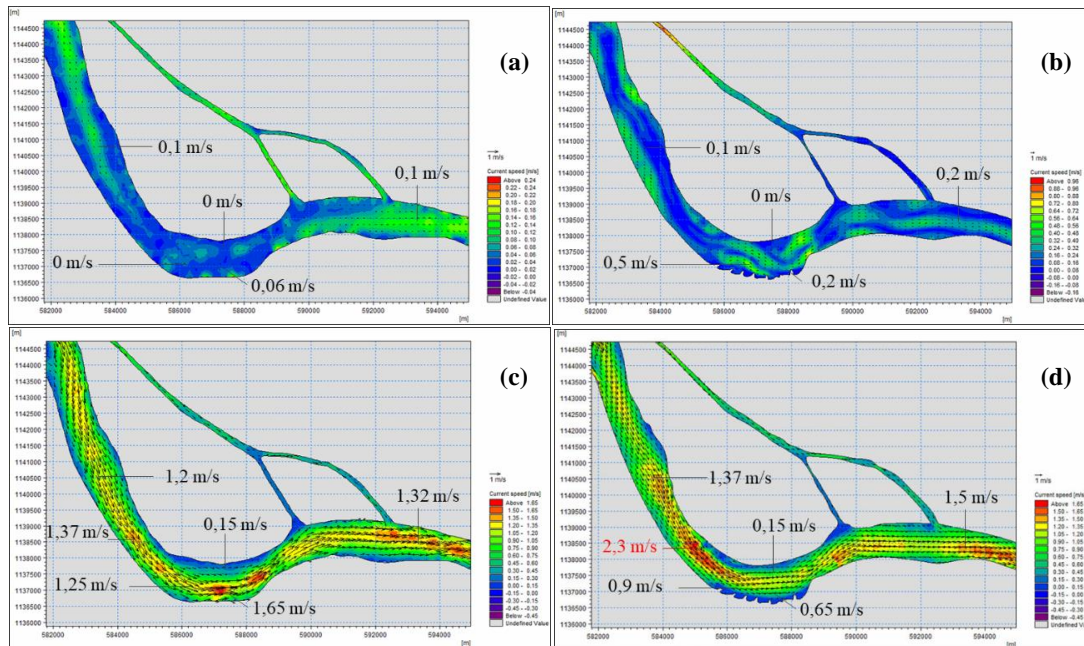
Hình 5. Kết quả hiệu chỉnh và kiểm định tại trạm Mỹ Thuận.



Hình 6. Kết quả hiệu chỉnh và kiểm định tại trạm Cao Lãnh.

3.2. Kết quả mô hình thủy lực

Các kết quả mô phỏng trường vận tốc của 2 kịch bản được biểu diễn tại 2 thời điểm: dòng chảy xuống mạnh nhất và dòng triều lên mạnh nhất (Hình 7). Trong đó, với kết quả dòng chảy xuống mạnh nhất sẽ được biểu diễn cùng 1 thang màu, còn với dòng triều lên ở 2 kịch bản sẽ biểu diễn bằng 2 thang màu khác nhau để thấy rõ kết quả cần phân tích.



Hình 7. Phân bố vận tốc dòng chảy khi không kè và có kè: (a) KB - Triều lên; (b) KB2 - Triều lên; (c) KB1 - Triều xuống; (d) KB2 - Triều xuống.

Ở pha triều lên, khu vực này là vùng giáp nước, với dòng chảy tự nhiên khi chưa có kè (KB1) thì vận tốc dòng chảy gần như bằng không (Hình 7a). Khi có kè (KB2), ngay tại đoạn cong xuất hiện các nhiễu động, vận tốc dòng chảy khoảng 0,2 m/s đến 0,5 m/s (Hình 7b). Tuy nhiên ở pha triều xuống, vận tốc dòng chảy ngay tại đoạn cong tương đối lớn và áp sát vào phía bờ lõm, vận tốc dòng chảy tối đa đạt tới 1,65 m/s và có xu hướng đâm thẳng vào bờ lõm (Hình 7c). Đây có khả năng là một trong những nguyên nhân làm cho xói lở bờ dẫn đến sạt bờ. Ở đoạn sông cong thường xảy ra dòng chảy thứ cấp, là một trong những yếu tố gây xói lở bờ lõm, nhưng trong mô hình 2D không tính được tác động của dòng chảy thứ cấp này. Phía bờ lồi của đoạn cong này, vận tốc dòng chảy thấp (khoảng 0,15 m/s), những khu vực này có khả năng sẽ có gây bồi.

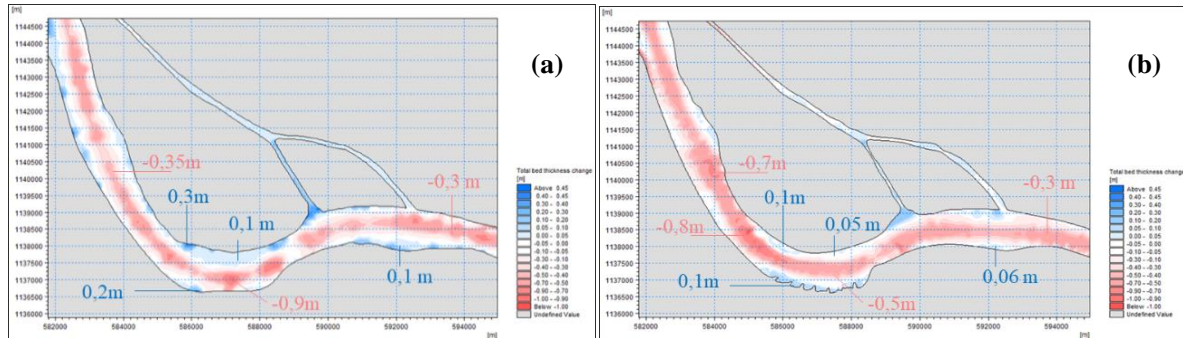
Khi có kè bảo vệ bờ ở đoạn sông cong (Hình 7d), phân bố vận tốc có sự thay đổi so với khi chưa có kè. Cụ thể, ở trước đoạn cong, vận tốc dòng chảy tăng mạnh, vận tốc tối đa đạt tới trên 2,3 m/s. Khu vực có kè làm thay đổi luồng phân bố vận tốc dòng chảy, dòng chảy lúc này không còn áp sát bờ lõm nữa mà đã phân bố ra giữa dòng với tốc độ dòng chảy cũng giảm hơn, đạt khoảng 0,65 m/s. Với vận tốc dòng chảy ở pha triều lên và pha triều xuống thay đổi liên tục và dòng chảy ở pha triều xuống chiếm ưu thế hơn trong mùa lũ, đây có khả năng sẽ tác động vào các kè ngay đoạn cong gây xói dần các chân kè, lâu ngày có thể làm ảnh hưởng đến chất lượng của các kè.

3.3. Mức độ bồi, xói lòng dẫn

Kết quả mô phỏng diễn biến bồi, xói lòng dẫn theo kịch bản sau 6 tháng mùa lũ, với màu xanh đại diện cho vùng bồi, màu đỏ đại diện cho vùng xói (màu đậm thể hiện mức độ bồi/xói nhiều, màu nhạt thể hiện mức độ bồi/xói ít) (Hình 8).

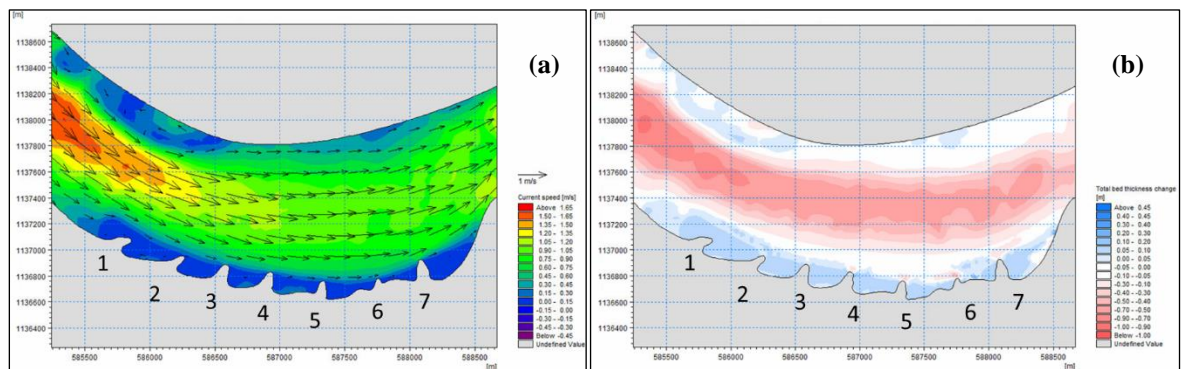
Hình 8a với kịch bản là dòng chảy tự nhiên khi chưa có kè, kết quả chỉ ra rằng, tại các vị trí phía bờ lõm, tốc độ xói cao đạt khoảng 0,9 m, phía bờ lồi, bồi tụ là chủ yếu với tốc độ 0,3 m sau 6 tháng. Đây là đoạn sông cong mạnh (cong 90 độ), địa hình lòng dẫn lệch về phía bờ lõm nên phân bố vận tốc dòng chảy ép sát phía bờ lõm hơn làm cho bờ lõm ngày càng bị đào xói thêm, bên bờ lồi chịu tác động dòng chảy yếu hơn, lắng đọng nhiều hơn nên ngày càng bồi thêm. Hình 8b là kết quả mức độ bồi xói khi có kè bảo vệ bờ. Với tác dụng của kè bờ, luồng dòng chảy không còn ép sát bờ mà bị đẩy ra giữa hơn, xói đáy cũng phân bố ở giữa

dòng với tốc độ 0,5-0,8 m/6 tháng. Quanh các kè mỏ hàn gần như dòng chảy yếu (vận tốc gần bằng không) do đó xảy ra bồi lắng nhẹ với tốc độ khoảng 0,1 m/6 tháng. Sát phía bờ lồi cũng là vùng bồi lắng nhẹ với mức độ khoảng 0,05 đến 0,1 m sau 6 tháng.



Hình 8. Mức độ bồi, xói lòng dẫn trong trường hợp không kè và có kè: (a) KB1; (b) KB2.

Để xem kỹ hơn tác dụng của các kè đến chế độ thủy động lực học dòng chảy và bồi, xói đáy thì kết quả được trích rõ hơn thể hiện trong Hình 9. Dòng chảy giữa các kè gần như bằng 0, do đó, giữa các kè có hiện tượng bồi nhẹ (0,1 m sau 6 tháng màu lữ). Ở vị trí đầu các kè có vận tốc từ 0,4-0,7 m/s, vận tốc này lớn hơn vận tốc khởi động bùn cát của vùng (vận tốc khởi động bùn cát khoảng 0,35 m/s [28]), do đó đây có thể sẽ là nguyên nhân gây ra xói lở ở các vị trí đầu các kè. Cụ thể, vị trí kè số 3 và số 4 đến kè số 7, vận tốc ở đầu các mũi kè tương đối lớn (trên 0,4 m/s), tương ứng với kết quả bồi xói (Hình 9b), tốc độ xói ở đầu các mũi kè từ 0,1 m đến 0,3 m trong 6 tháng mùa lũ. Thực tế cũng cho thấy đỉnh kè 4 và đỉnh kè 7 bị sạt, cụ thể, tính từ cuối năm 2013 đến nay, đã xuất hiện tình trạng sạt lở bờ giữa các kè mỏ hàn số 2, số 3, giữa kè số 4 và 5 và năm 2017 sạt kè số 7 [29]. Các kè bị sự số cũng đã và đang được khắc phục, tuy nhiên với tình hình diễn biến bồi, xói phức tạp ngay đoạn sông cong, trong tương lai sẽ rất dễ tiếp tục bị hư hại nếu không có những giải pháp phù hợp.



Hình 9. Vận tốc dòng chảy (a) và mức độ bồi, xói (b) tại vị trí kè: (a) Dòng chảy; (b) Bồi, xói lòng dẫn.

3.4. Thảo luận

Đoạn Sa Đéc được biết đến với tình hình sạt lở bờ nghiêm trọng [20]. Theo thống kê từ Viện kỹ thuật biển, từ những số liệu khảo sát và thu thập trong giai đoạn từ 2003 đến 2015, có thể thấy quá trình xói bồi sông Tiền khu vực An Hiệp - cồn Linh trong giai đoạn này như Hình 10. Ngay đoạn cong xã An Hiệp, phía bờ trái (cồn Linh) bồi tụ nhiều với chiều rộng bồi khoảng 1100 m sau 10 năm, tuy nhiên ở bờ phải khu vực An Hiệp bị sạt lở nghiêm trọng với bề rộng sạt lở ăn sâu 500 m.

Trên đoạn sông cong đi qua xã An Hiệp, từ sau năm 2012 đã được nhà nước đầu tư xây dựng 7 kè mỏ hàn để chống sạt lở (Hình 11), tuy nhiên, do vận tốc dòng chảy ở đây rất lớn áp sát vào bờ lõm, địa hình đáy đoạn này có các hố xói sâu trên 40 m, do đó tình hình sạt lở vẫn diễn biến phức tạp. Cụ thể, tính từ cuối năm 2013 đến nay, đã xuất hiện tình trạng sạt lở bờ giữa các kè mỏ

hàn số 2, số 3, giữa kè số 4 và 5 và năm 2017 sát kè số 7 [20]. Mặc dù đã có kè nhưng do dòng chảy tự nhiên ở đây khá mạnh, lại chảy qua đoạn sông cong và co hẹp nên tình hình sạt lở vẫn xảy ra mạnh.



Hình 10. Diễn biến đường bờ sông Tiền khu vực An Hiệp - cồn Linh (2003-2015) [30].



Hình 11. Một kè mở hàn khu vực xã An Hiệp, huyện Châu Thành [20].

Nguyên nhân gây bồi bên phía bờ lồi và xói bờ lõm: khu vực này là đoạn sông cong tự nhiên, địa hình có lạch sâu ép sát bờ lõm, hố xói sâu trên 40 m, chế độ dòng chảy phức tạp do chịu ảnh hưởng của chế độ dòng chảy lũ ở thượng lưu và chịu ảnh hưởng của triều ở hạ lưu, do đó, có thể gây ra dòng xoáy và dòng nhiễu động trong nội bộ dòng chảy.

4. Kết luận

Nghiên cứu ứng dụng mô hình MIKE21 đã tính toán được chế độ thủy lực và chuyên tải bùn cát, biến đổi đáy tại khu vực sông Tiền đoạn qua thành phố Sa Đéc. Với mục đích của nghiên cứu là đánh giá nguy cơ xói lở dưới tác động của dòng chảy trước và sau khi xây dựng kè bờ, kết quả đạt được như sau:

- Dòng chảy trước khi có kè có vận tốc dòng chảy tối đa đạt 1,65m/s áp sát bờ lõm ở đoạn sông cong, dẫn đến xói sâu lòng dẫn (0,9 m/6 tháng) và bồi nhẹ ở bờ lồi (0,1-0,3 m/6 tháng).

- Sau khi xây dựng kè, vận tốc dòng chảy tăng ở phía trước đoạn có kè (vận tốc đạt tối đa 2,3 m/s) và giảm dần vận tốc trong khu vực có kè, đồng thời các kè này làm thay đổi luồng dòng chảy. Nhìn chung là mức độ xói giảm ở sát bờ lõm nhưng mức độ xói ở giữa dòng cũng khá đáng kể (0,8 m/6 tháng). Diện tích xói tăng so với khi chưa có kè. Bồi lắng phân bố ở phía bờ lồi, diện tích bồi giảm hơn so với khi không có kè.

- Mặc dù các kè mở hàn có khả năng giảm xói hiệu quả ở sát bờ lõm (quanh các kè sát bờ lõm có hiện tượng bồi), nhưng ở đầu các kè mở hàn cũng bị xói dẫn đến sạt lở kè.

Bên cạnh những kết quả đạt được như trên, trong bài báo này còn hạn chế là chưa tính toán được quá trình bồi xói trong thời gian dài hơn (trên 1 năm), do đó chưa thấy được tác động của dòng chảy mùa lũ và dòng chảy mùa kiệt.

Khu vực này là đoạn sông cong tự nhiên có chế độ dòng chảy mạnh và phức tạp. Địa hình lòng sông có lạch sâu ép sát bờ lõm, có hố xói sâu. Dòng chảy chịu ảnh hưởng của cả dòng chảy lũ ở thượng lưu và triều ở hạ lưu, do đó, dòng chảy phức tạp với dòng xoáy và dòng nhiễu động, dễ dẫn đến xói lở khó khắc phục. Nghiên cứu góp phần giúp các nhà quản lý có thêm góc nhìn về hiệu quả của kè mở hàn ở đoạn sông cong này.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: T.N.Q.N.; Lựa chọn phương pháp nghiên cứu: N.T.B., T.T.K., T.N.Q.N.; Xử lý số liệu: O.B.H.; Tính toán mô hình: O.B.H., T.N.Q.N.; Phân tích kết quả: T.T.K., T.N.Q.N.; Viết bản thảo bài báo: T.N.Q.N.; Chỉnh sửa bài báo: N.T.B., T.T.K., T.N.Q.N.

Lời cảm ơn: Chúng tôi xin cảm ơn Trường Đại học Bách khoa, ĐHQG-HCM đã hỗ trợ cho nghiên cứu này.

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. Abdella, K.; Mekuanent, F. Application of hydrodynamic models for designing structural measures for river flood mitigation: the case of Kulfo River in southern Ethiopia. *Model. Earth Syst. Environ.* **2021**, 7, 2779–2791.
2. Viero, D.P.; D’Alpaos, A.; Carniello, L.; Defina, A. Mathematical modeling of flooding due to river bank failure. *Adv. Water Resour.* **2013**, 59, 82–94.
3. Abderrezzak, K.E.K.; Moran, A.D.; Tassi, P.; Ata, R.; Hervouet, J.M. Modelling river bank erosion using a 2D depth-averaged numerical model of flow and non-cohesive, non-uniform sediment transport. *Adv. Water Resour.* **2016**, 93, 75–88.
4. Govers, G.; Giménez, R.; Van, O.K. Rill erosion: exploring the relationship between experiments, modelling and field observations. *Earth Sci. Rev.* **2007**, 84(3-4), 87–102.
5. Xu, D.; Bai, Y.; Ji, C.; Williams, J. Experimental study of the density influence on the incipient motion and erosion modes of muds in unidirectional flows: the case of Huangmaohai Estuary. *Ocean Dyn.* **2015**, 65, 187–201.
6. Mansouri, A.; Arabnejad, H.; Shirazi, S.A.; McLaury, B.S. A combined CFD/experimental methodology for erosion prediction. *J. Frict. Wear.* **2015**, 332, 1090–1097.
7. Hoài, H.C.; Bầy, N.T.; Khôi, Đ.N.; Nga, T.N.Q. Phân tích nguyên nhân gây gia tăng xói lở bờ sông ở đồng bằng sông Cửu Long. *Tạp chí Khí tượng Thủy văn* **2019**, 703, 42–50.
8. Tân, N.P.; Trí, V.P.Đ.; Thành, V.Q. Ứng dụng mô hình thủy lực hai chiều mô phỏng đặc tính thủy lực và tính toán bồi xói ở vùng cửa sông Định An. *Tạp chí Khoa học Đại học Cần Thơ* **2014**, 31, 8–17.
9. Hùng, N.N.; Hưng, D.Q. Đánh giá hiện trạng hình thái lòng dẫn một số đoạn sông Tiền và sông Vàm Nao bằng tài liệu địa chấn nông phân giải cao. *Tạp chí Khoa học và Công nghệ Việt Nam* **2020**, 62(1), 17–23.
10. Thuy, N.T.D.; Khoi, D.N.; Nhan, D.N.; Nga, T.N.Q.; Bay, N.T.; Phung, N.K. Modelling accresion and erosion processes in the bassac and mekong rivers of the vietnamese mekong delta. in APAC 2019: Proceedings of the 10th International Conference on Asian and Pacific Coasts, 2019, Hanoi, Vietnam. 2020.
11. Binh, D.; Kantoush, S.A.; Saber, M.; Mai, N.P.; Maskey, S.; Phong, D.T.; Sumi, T. Long-term alterations of flow regimes of the Mekong River and adaptation strategies for the Vietnamese Mekong Delta. *J. Hydrol.: Reg. Stud.* **2020**, 32, 100742.
12. Van, C.T.; Tuan, L.A.; Tuan, N.C. Application of two-dimensional hydrodynamic model (MIKE 21FM) to simulate the sediment regime on Hau river, piloted in Long Xuyen city-An Giang province. *VNU HCM J. Earth Sci. Environ.* **2021**, 5(SI2), SI20–SI33.
13. Hoàng, T.B.; Dương, N.B.; Phong, N.C. Chế độ vận chuyển bùn cát vùng đồng bằng sông Cửu Long trong kịch bản phát triển thượng nguồn. *Tạp chí Khoa học và công nghệ thủy lợi* **2019**, 57, 47–57.
14. Tien, N.N.; Thuc, D.N.; Luan, N.T.; Vinh, V.D. Study and assessment of the situations and causes of erosion along the Hau riverbank in An Giang province during the period 2009–2019. *VN J. Marine Sci. Technol.* **2021**, 21(4), 493–506.
15. Khoi, D.N.; Duc, D.T.; Lien, P.T.H.; Loi, P.T.; Thuy, N.T.D.; Phung, N.K.; Bay, N.T. Morphological change assessment from intertidal to river-dominated zones using multiple-satellite imagery: A case study of the Vietnamese Mekong Delta. *J. Reg. Stud. Mar. Sci.* **2020**, 34, 101087.

16. Anthony, E.J.; Brunier, G.; Besset, M.; Goichot, M.; Dussouillez, P.; Lap, N.V. Liên hệ giữa xói lở nhanh ở đồng bằng sông Cửu Long và các hoạt động của con người. *Sci. Rep.* **2015**, *5*, 14745.
17. Điệp, N.T.H.; Minh, V.Q.; Trường, P.N.; Thành, L.K.; Vinh, T.L.Q. Diễn tiến tình hình sạt lở ven bờ sông Tiền và sông Hậu, vùng Đồng bằng sông Cửu Long. *Tạp chí Khoa học Trường Đại học Cần Thơ* **2019**, *55*, 125–133.
18. Đoàn, V.P. Ứng dụng mô hình Mike 21 đánh giá tác động của công trình kè cửa sông đến trường thủy động lực khu vực Cửa Tùng, tỉnh Quảng Trị. Đại học Quốc gia Hà Nội, 2020.
19. Bình, H.T.; Châm, Đ.Đ.; Thảo, Đ.T.; Hạnh, L.Đ.; Sơn, N.T.; Huấn, N.M.; Trinh, N.Q. Nghiên cứu các quá trình thủy động lực tích hợp (sóng, dòng chảy và mực nước) bằng MIKE 21/3 coupled model FM vùng biển Đà Nẵng. *Tạp chí Khí tượng Thủy văn* **2022**, *735*, 1–11.
20. Bầy, N.T. Nghiên cứu xác định nguyên nhân, cơ chế và đề xuất các giải pháp khả thi về kỹ thuật, hiệu quả về kinh tế nhằm hạn chế xói lở, bồi lắng cho hệ thống sông Đồng Bằng Sông Cửu Long. Báo cáo tổng kết đề tài thuộc Chương trình Tây Nam Bộ, 2021.
21. Chi cục Thủy lợi. Sở Nông nghiệp & phát triển nông thôn Đồng Tháp. Đồng Tháp thiệt hại hơn 8,2 tỉ đồng do sạt lở 2020. Trục tuyến: <http://vmha.gov.vn/public/kttv-voi-san-xuat-va-doi-song-106/dong-thap-thiet-hai-hon-8-2-ti-dong-do-sat-lo-7052.html>.
22. Dương, H. Cần 400 tỉ để xây kè chống sạt lở ở Đồng Tháp. Báo điện tử Pháp Luật thành phố Hồ Chí Minh 2023. Trục tuyến: <https://plo.vn/can-400-ti-de-xay-ke-chong-sat-lo-o-dong-thap-post720943.html>
23. Tài, N. Kè bảo vệ bờ sông hơn trăm tỷ đồng liên tục sạt lở. VnExpress 2023. Trục tuyến: <https://vnexpress.net/ke-105-ty-o-dong-thap-sat-lo-lan-thu-4-4607070.html>.
24. Ánh, N. Đồng Tháp đề nghị hỗ trợ vốn để xử lý cấp bách sạt lở. Cơ quan phát ngôn chính thống của UBND tỉnh Đồng Tháp 2021. Trục tuyến: https://dongthap.gov.vn/chi-tiet-bai-viet/-/asset_publisher/1mOzUrGkrdAE/content/id/8694611
25. Có, L. Đồng Tháp: Kè chống sạt lở chưa xong đã sạt lở, sụt lún. Trang Tin Điện Tử Truyền Hình Quốc Hội Việt Nam 2023. Trục tuyến: <https://www.quochoitv.vn/dong-thap-ke-chong-sat-lo-chua-xong-da-sat-lo-sut-lun>
26. DHI. MIKE 21 & MIKE 3 Flow Model FM - Hydrodynamic and Transport Module, in Scientific documentation. 2012.
27. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and water quality models: Performance measures and evaluation criteria. *J. Trans. ASABE* **2015**, *58(6)*, 1763–1785.
28. Huệ, V.H. Giải pháp công trình khắc phục sạt lở cồn Thanh Long. *Tạp chí Khí tượng Thủy văn*, **2023**, *754*, 26–43.
29. Thuận, P. Dân ở xã An Hiệp bao giờ mới hết lo sạt lở. Báo Đồng Tháp, 2015. 04.2015. Trục tuyến: <https://www.baodongthap.vn/xa-hoi/dan-o-xa-an-hiep-bao-gio-moi-het-lo-sat-lo-48274.aspx>.
30. Hùng, L.M. Nghiên cứu dự báo xói bồi lòng dẫn và đề xuất giải pháp phòng chống cho hệ thống sông ở ĐBSCL. Báo cáo kết quả nghiên cứu đề tài nhà nước KC08-15 Viện Khoa học Thủy lợi miền Nam, 2004.

The assessment of bank erosion susceptibility under the influence of flow before and after the construction of groyne embankments on the Tien River section flowing through Sa Dec City, Dong Thap province

Tra Nguyen Quynh Nga^{1,2*}, Tran Thi Kim³, On Bao Hang^{1,2}, Nguyen Thi Bay^{1,2}

¹ Ho Chi Minh City University of Technology; tnqnga@hcmut.edu.vn;
hangbaon1609@gmail.com; ntbay@hcmut.edu.vn

² Vietnam National University Ho Chi Minh City; tnqnga@hcmut.edu.vn;
hangbaon1609@gmail.com; ntbay@hcmut.edu.vn

³ Ho Chi Minh City University of Natural Resources and Environment;
ttkim@hcmunre.edu.vn

Abstract: Dong Thap province is located upstream of the Mekong River, where there is a dense river network and frequent bank erosion. This paper focuses on studying the influence of the hydrodynamic regime before and after groyne embankments on the section of the Tien River that flows through Sa Dec city, Dong Thap province. The main research method is to utilize the MIKE21 model with a hydrodynamic module (HD) combined with a mud transport module (MT). The model has been calibrated and validated to ensure high compatibility. The model is calculated based on two scenarios: one before the construction of the seven groyne embankments, and one after. The results indicate that the groyne embankments alter the distribution of flow. As a result, the level of erosion decreases near the concave bank, but there is a significant level of erosion in the middle of the stream (0,8 m/6 months). The erosion area has increased compared to when there are no embankments. While groyne embankments can effectively reduce erosion near concave banks (resulting in a lightly accreted area), it is important to note that the tops of groyne embankments are still susceptible to significant erosion caused by strong currents. This erosion can ultimately lead to embankment failure. The study also contributes to assisting managers in gaining a broader perspective on the effectiveness of groyne embankments in this meandering river section.

Keywords: MIKE21; Hydrodynamics; Erosion; Groyne embankments; Tien River; Sa Dec City.

Bài báo khoa học

Đánh giá ảnh hưởng của khai thác cát đến diễn biến đáy đoạn sông Tiền chảy qua tỉnh Vĩnh Long

Nguyễn Thị Bầy^{1,2*}, Trần Thị Kim³, Trần Thị Thúy An³, Trà Nguyễn Quỳnh Nga^{1,2}

¹ Trường Đại học Bách Khoa; ntbay@hcmut.edu.vn; tnqnga@hcmut.edu.vn

² Đại học Quốc Gia Tp.HCM; ntbay@hcmut.edu.vn; tnqnga@hcmut.edu.vn

³ Trường Đại học Tài nguyên và Môi trường Tp.HCM; ttkim@hcmunre.edu.vn; antran.nrec@gmail.com

*Tác giả liên hệ: ntbay@hcmut.edu.vn; Tel.: +84-902698585

Ban Biên tập nhận bài: 5/9/2022; Ngày phản biện xong: 17/10/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Khai thác cát sông quá mức làm thay đổi kết cấu địa chất, gây ra sạt lở nghiêm trọng là một trong những vấn đề được xã hội quan tâm hiện nay. Nghiên cứu này tập trung đánh giá diễn biến lòng dẫn dưới ảnh hưởng của suy giảm phù sa và hoạt động khai thác cát tại khu vực sông Tiền đoạn chảy qua tỉnh Vĩnh Long. Kết quả mô phỏng bồi, xói đáy cho thấy ở đoạn sông này có diện tích bồi lắng phân bố tương đối nhiều ở các đoạn sông thẳng, tốc độ bồi lắng năm 2017 thấp hơn so với năm 2008. Trong khi đó, xói đáy xảy ra ở các đoạn sông co hẹp đột ngột, có vận tốc dòng chảy lớn, địa hình đáy sâu. Kết quả mô phỏng cho thấy hoạt động khai thác cát có những tác động tích cực và tiêu cực khác nhau. Nếu khai thác không đúng quy mô, vị trí có thể gây ra những vấn đề biến đổi thủy động lực khó lường dẫn đến nguy cơ sạt lở bờ. Tuy nhiên, khai thác cát đúng quy định lại đem đến kết quả tốt, giúp làm khơi thông dòng chảy và làm cho địa hình đáy trở về hình dạng ban đầu. Các kết quả đạt được là cơ sở trong việc đề xuất các biện pháp giảm thiểu xói lở phục vụ công tác quản lý rủi ro thiên tai do sạt lở.

Từ khóa: Sông Tiền; Khai thác cát; Chuyên tải phù sa; Vận chuyển bùn cát; Suy giảm phù sa.

1. Mở đầu

Sự biến đổi của lòng dẫn trong các dòng sông diễn ra liên tục theo không gian và thời gian, dưới tác động của cả yếu tố tự nhiên và nhân tạo. Những sự thay đổi của một dòng sông/kênh có thể là về: kích thước, hình dạng hay hình thái trên mặt bằng... Mặc dù vậy, sự quan tâm hàng đầu vẫn là: cơ chế sạt lở và phát triển các phương pháp để dự đoán tương tác của dòng sông. Trong đó, việc ứng dụng mô hình toán số để mô phỏng bồi xói lòng dẫn, từ đó phân tích nguyên nhân và cơ chế đang được thực hiện rộng rãi và có độ tin cậy cao [1-7].

Đồng bằng sông Cửu Long (ĐBSCL) được hình thành vài ngàn năm với sự bồi lắng phù sa từ sông Mê Kông [8]. Hiện nay, lượng trầm tích ở ĐBSCL đã giảm đáng kể bị ảnh hưởng từ các quá trình tự nhiên và các hoạt động của nhân tạo như khai thác cát, xây đập, thay đổi sử dụng đất [9]. Trong giai đoạn 2012-2013, Kondolf đã cho thấy tải lượng trầm tích đổ ra biển của sông Mekong đã giảm 75% từ 160 triệu tấn/năm trước khi có đập [10] xuống còn 40 triệu tấn/năm [11]. Nghiên cứu [12] cũng cho thấy nếu tất cả các đập dự kiến được xây dựng, dòng trầm tích lơ lửng sẽ suy giảm 50% so với mức hiện tại (đo đạc tại Tân Châu) và gây ra hậu quả cho sinh kế và hệ sinh thái địa phương. Theo số liệu đo đạc từ Đài khí tượng thủy văn khu vực Nam bộ, trong các năm 2008 và 2017 lượng nước về trạm Tân Châu và

Tạp chí Khí tượng Thủy văn **2023**, 756, 14-28; doi:10.36335/VNJHM.2023(756), 14-28 <http://tapchikttv.vn>

Châu Đốc không thay đổi nhưng lượng bùn cát giảm đi 1/3. Nếu thống kê số điểm sạt lở theo thời gian và thể tích nước cùng lượng bùn cát từ dòng chính Mekong đổ vào ĐBSCL, nhận thấy có một sự trùng hợp về thời điểm mà các điểm sạt lở gia tăng (sau 2010). Điều này chứng tỏ sự thiếu hụt bùn cát, gây xói lở nghiêm trọng cho hệ thống sông ĐBSCL.

Ngoài tác động xây dựng đập ở thượng nguồn, hoạt động khai thác cát cũng là một nguyên nhân làm giảm lượng bùn cát, gia tăng xói lở ở khu vực ĐBSCL [13–16]. Lượng cát hàng năm bị khai thác khoảng 57 triệu tấn, trong đó 86% là cát chủ yếu diễn ra mạnh trong giai đoạn từ 2008 đến 2012 [17]. Khai thác cát có cả hậu quả ngắn hạn và dài hạn [18]. Tác động ngắn hạn của khai thác cát bao gồm sạt bờ, xói đáy và biến đổi mực nước ngầm [19]. Trong khi đó, về lâu dài sẽ gây nên xâm nhập mặn và tạo nên các vùng ngập lũ do hạ thấp mực nước ngầm hay nên hạ thấp cốt nền [20]. Các tác động của việc khai thác cát đã được nghiên cứu bởi nhiều chuyên gia trên thế giới, điển hình như nhóm nghiên cứu [21] đã nghiên cứu ảnh hưởng của việc đào cát trong kênh đến chế độ thủy văn của đồng bằng sông Châu Giang, Trung Quốc. Nghiên cứu cho thấy, bên cạnh các tác động tích cực như giảm nguy cơ thiệt hại do lũ lụt, cải thiện điều kiện giao thông đường thủy và cung cấp nước nhiều hơn cho các khu vực đang phát triển kinh tế thì vấn đề khai thác cát cũng đem lại những tiêu cực nhất định bao gồm làm tăng độ dốc và sự mất ổn định của bờ sông, gây gián đoạn giao thông thủy ở các hồ nạo vét thượng nguồn trong mùa khô và xâm nhập nước lợ [21]. Bên cạnh đó, nghiên cứu [22] cũng đã cho thấy tác động từ việc khai thác cát đối với khu vực Bestari Jaya, Selangor là lượng khai thác quá lớn so với khả năng tái tạo của tự nhiên, việc giảm tải trọng như vậy có thể gây ra xói lở tại hạ lưu và bờ sông, thay đổi hướng dòng chảy; ngoài ra việc khai thác cát cũng có thể tạo ra các vực sâu, làm mất đi các rãnh, tăng độ đục [22]. [23] đã nghiên cứu tác động của việc khai thác cát ở bốn khía cạnh môi trường là vật lý, sinh học, hóa học và môi trường nhân tạo. Các tác động của việc khai thác cát sông lên môi trường vật lý là mở rộng và hạ thấp lòng sông; trong môi trường sinh học tác động bao trùm là làm giảm đa dạng sinh học và trải dài từ hệ động thực vật thủy sinh và ven bờ đến toàn bộ khu vực đồng bằng ngập lũ; môi trường hóa học là làm giảm chất lượng nước, không khí và đất do ô nhiễm và môi trường nhân tạo là cơ sở hạ tầng bị hư hỏng, điều kiện làm việc tồi tệ cho người lao động, khả năng tiếp cận nguồn nước hạn chế và thiệt hại về nông nghiệp [23].

Hiện nay, khai thác cát trên dòng chính Mekong ngoài lãnh thổ Việt Nam gây thiếu hụt bùn cát nghiêm trọng trên dòng chảy về ĐBSCL. Trong lãnh thổ Việt Nam, khai thác cát quá mức đã tạo ra những hố cục bộ làm thay đổi dòng chảy vốn đã ổn định [24–27]. Lượng bùn cát đổ về từ thượng lưu không đủ bùn cát để bù đắp vào các hố, làm các hố cát thay đổi không kiểm soát được [28–30]. Bên cạnh đó, việc khai thác cát gần bờ, sau quá trình khai thác để lại hố sâu sát bờ gây mất ổn định bờ [16].

Hiện tượng xói lở, biến hình lòng dẫn tại Vĩnh Long hiện nay đang diễn biến rất phức tạp, rộng khắp trên phạm vi toàn tỉnh [31]. Vấn đề sạt lở tại tỉnh Vĩnh Long cũng đã thu hút được sự quan tâm của rất nhiều chuyên gia, cụ thể như nghiên cứu [32] đã thể hiện xã Đồng Phú thuộc thành phố Vĩnh Long tỉnh Vĩnh Long là một trong những khu vực điểm nóng rất thường xuyên xảy ra sạt lở. Các đặc điểm như đoạn cong ven sông, cồn giữa sông là nơi thường xảy ra sạt lở do ảnh hưởng của dòng chảy. Cũng tại khu vực tỉnh Vĩnh Long, nghiên cứu của tác giả [33] cho thấy, dòng chảy hướng vào đầu cồn là một trong những nguyên nhân chính gây sạt lở, để khắc phục tình trạng sạt lở nêu trên nghiên cứu đã đề xuất giải pháp mở hàn và kè. Cũng với hướng nghiên cứu trên, và áp dụng cho khu vực bờ sông Cổ Chiên tác giả này đã xác định được những nguyên nhân gây mất ổn định bờ như bùn cát từ thượng nguồn suy giảm, lòng dẫn bị hạ thấp, tỷ lệ phân lưu dòng chảy và dòng chủ lưu áp sát bờ [34].

Trên địa bàn tỉnh đã xảy ra khoảng 200 tuyến/điểm sạt lở, làm mất 5-6 km bờ sông, kênh, rạch. Năm 2022, đoạn từ vàm Mương Lộ đến rạch Bà Bông, bờ sông Cổ Chiên đã xảy ra vụ sạt lở nghiêm trọng với tổng diện tích sạt lở lên đến 41.516 m² [35]. Theo [18], dựa trên bản

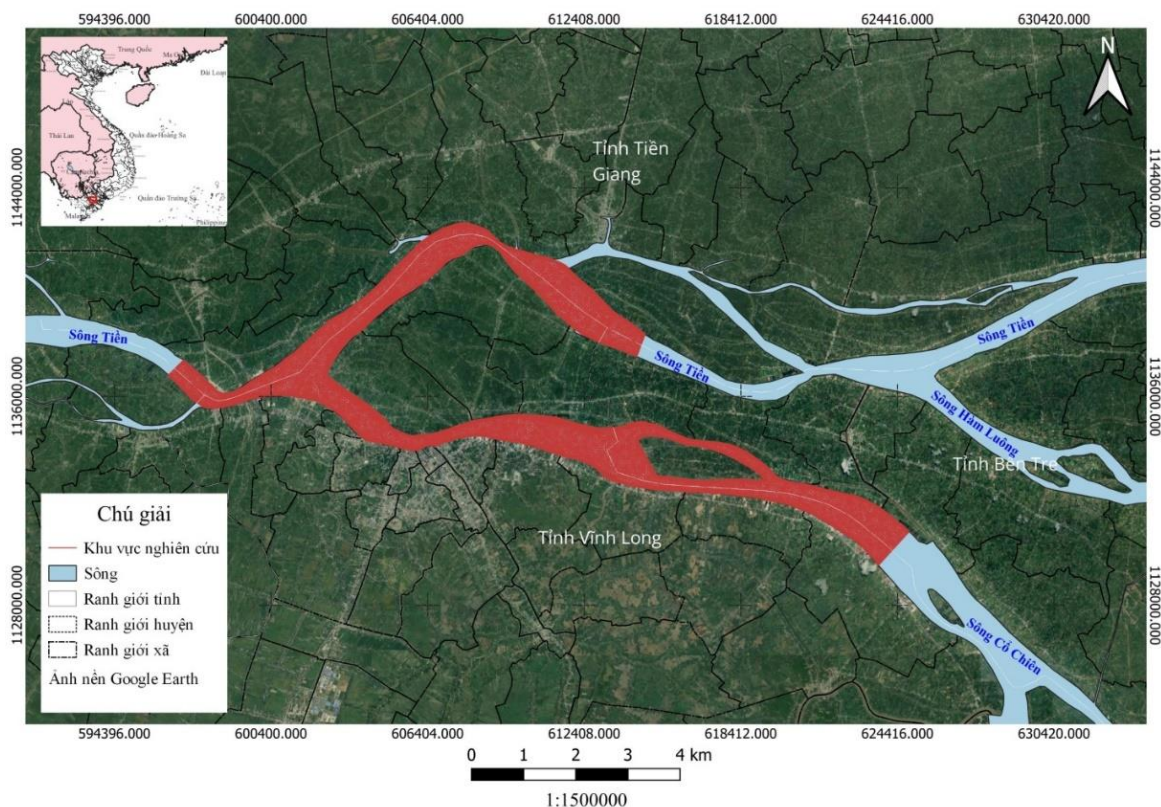
đồ mật độ thuyền khai thác cát, nghiên cứu này đã xác định được một số điểm nóng khai thác cát lớn dọc biên giới Campuchia trên sông Tiền. Các vị trí khai thác cũng ở trên sông Tiền là cù lao Long Khánh, luồng hướng ra Vàm Nao, cù lao Thuận Đông và đoạn Sa Đéc. Ở hạ lưu Mỹ Thuận, điểm nóng lớn được phát hiện quanh ngã ba Mỹ Tho, sông Cổ Chiên, gần thành phố Vĩnh Long và xung quanh An Phước. Điểm nóng nhất ở hạ lưu là hạ lưu cầu Cổ Chiên [18]. Do đó, nghiên cứu này tập trung vào mô phỏng diễn biến đáy cho đoạn sông Tiền chảy qua Tp. Vĩnh Long với hai nhánh chính là nhánh sông Tiền và nhánh sông Cổ Chiên khi có ảnh hưởng của khai thác cát và thiếu hụt phù sa.

Từ các nhận định trên, nghiên cứu này tập trung đánh giá diễn biến lòng dẫn khi suy giảm phù sa và ảnh hưởng từ hoạt động khai thác cát tại khu vực sông Tiền đoạn chảy qua tỉnh Vĩnh Long. Mục tiêu chính của nghiên cứu là tập trung đánh giá diễn biến đáy với 03 kịch bản tương ứng với trường hợp: năm 2017, khi có suy giảm phù sa (giảm 1/3 so với năm 2008), năm 2008, khi chưa có suy giảm phù sa và năm 2017, khi có thêm yếu tố khai thác cát. Kết quả của nghiên cứu là cơ sở, hỗ trợ trong việc đề xuất các biện pháp giảm thiểu xói lở bờ sông phục vụ công tác quản lý rủi ro thiên tai do sạt lở.

2. Phương pháp nghiên cứu và dữ liệu thu thập

2.1. Phạm vi nghiên cứu

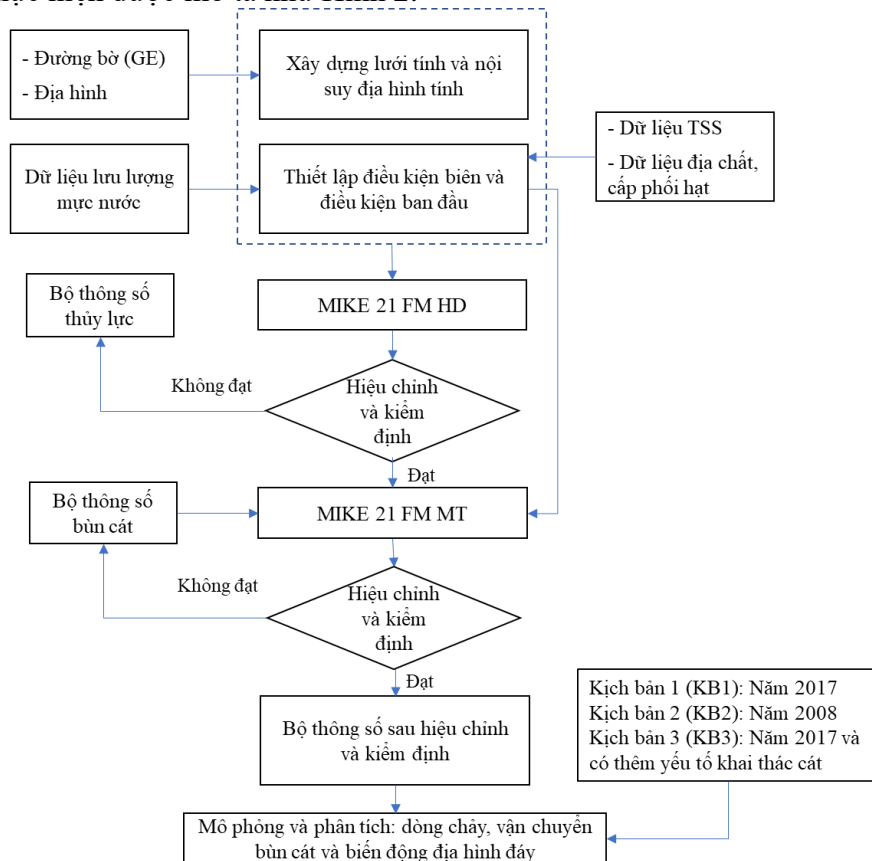
Phạm vi nghiên cứu là đoạn sông Tiền đoạn chảy qua tỉnh Vĩnh Long (Hình 1). Theo khảo sát của Chi cục Thủy lợi Vĩnh Long, hiện Vĩnh Long có 4 khu vực sạt lở mạnh và nguy hiểm (trên 10m/năm) dài 4.690 m, 7 khu vực sạt lở trung bình (từ 5 đến dưới 10 m/năm) và còn lại là các điểm sạt lở nhỏ (dưới 5 m/năm). Các khu vực sạt lở mạnh là: Trên sông Hậu có khu vực từ vàm kinh Hai Quý đến phà Cần Thơ (Bình Minh) dài 690 m; trên sông Tiền hiện khu vực đầu cù lao Minh (phần địa phận xã An Bình, Long Hồ) đang nguy hiểm; sông Cổ Chiên có khu vực ấp Phước Định 1 và 2, xã Bình Hoà Phước (Long Hồ) là khu vực sạt lở lớn.



Hình 1. Khu vực nghiên cứu.

2.2. Mô hình MIKE 21

Nghiên cứu này tập trung tính toán chế độ thủy động lực và vận chuyển bùn cát khu vực đoạn sông Tiền chảy qua tỉnh Vĩnh Long bằng mô hình MIKE 21 FM với module HD và MT của (DHI). Mô hình MIKE 21 HD là gói phần mềm mô phỏng chế độ thủy động lực học 2 chiều. Lý thuyết mô hình dựa trên hệ phương trình Saint - Venant với lưới phi cấu trúc linh hoạt. Mô hình MIKE 21 HD cho phép mô phỏng các đặc trưng thủy lực, môi trường và hình thái hai chiều...[36]. MIKE 21 MT là mô hình vận chuyển bùn, mô hình kết hợp giữa mô hình nhiều lớp, được sử dụng để mô tả vận chuyển bùn cát kết dính và không kết dính [37]. Các bước thực hiện được mô tả như Hình 2.



Hình 2. Khung thực hiện.

2.3. Dữ liệu thu thập

Số liệu địa hình năm 2008 được thu thập từ Sở Tài Nguyên và Môi Trường Tỉnh Vĩnh Long. Bên cạnh đó, dữ liệu địa hình đáy năm 2017 kế thừa từ [38].

Số liệu đo thủy văn trong 3 ngày (từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013) tại 03 vị trí trong vùng nghiên cứu được dùng để hiệu chỉnh mô hình và 0h ngày 7/7/2013 đến 0h ngày 10/7/2013 để kiểm định mô hình, được thu thập từ đề tài [31].

Bảng 1. Vị trí tọa độ của 03 trạm đo.

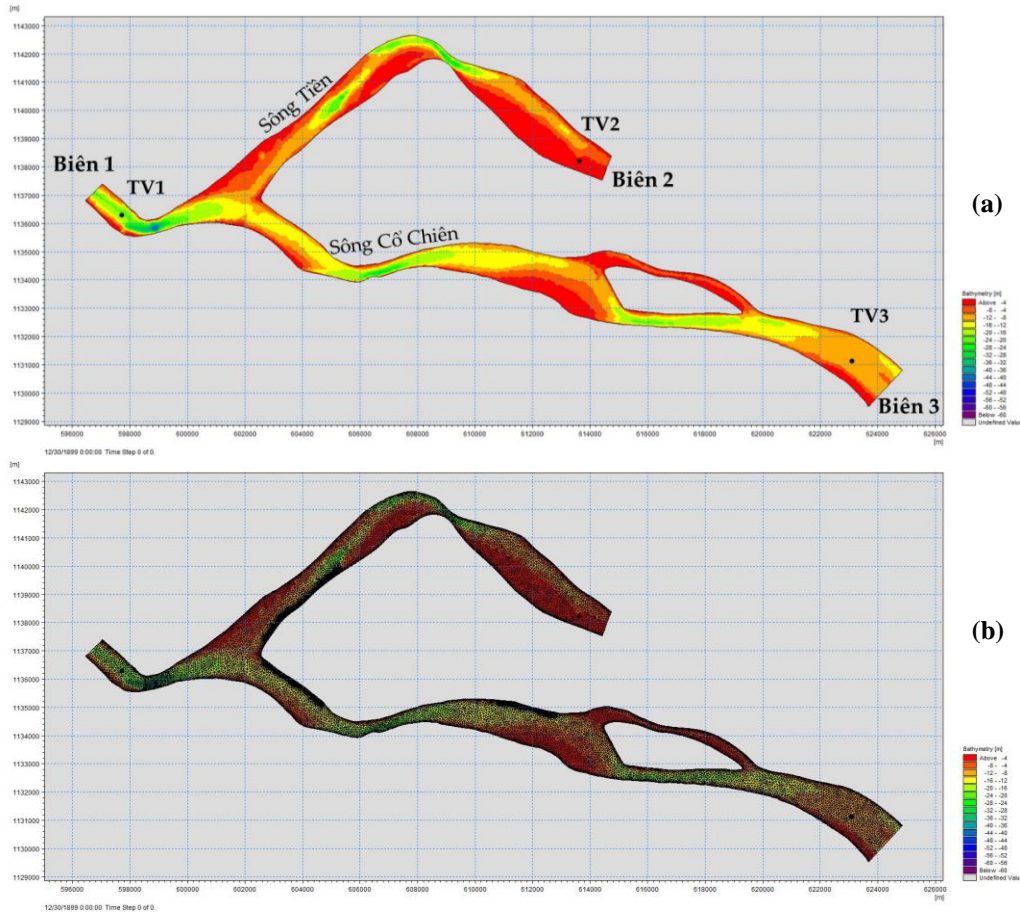
Trạm	Kinh độ	Vĩ độ
TV1	105°53'32,32"E	10°16'41,24"N
TV2	106°2'15,82"E	10°17'41,92"N
TV3	106°7'26,19"E	10°13'50,23"N

Đặc điểm cấu trúc địa chất được thu thập từ Liên Đoàn Bản đồ địa chất miền Nam, 2017.

2.4. Thiết lập mô hình

2.4.1. Lưới tính

Phạm vi không gian của vùng tính được thể hiện như Hình 3. Lưới tính được tạo thành từ 28.926 nút và 53.023 phần tử (Hình 1).



Hình 3. Vùng tính, biên và vị trí các trạm hiệu chỉnh và kiểm định (a) và lưới tính (b).

2.4.2. Điều kiện biên

Điều kiện biên thủy lực và phù sa được trích xuất từ mô hình MIKE 21 từ đề tài [31]. Vùng lớn đã được hiệu chỉnh và kiểm định với độ tin cậy cao [38].

Vùng tính có 03 biên lỏng (Hình 3a) gồm: Biên 1 nằm ở đầu sông Tiền (trước cầu Mỹ Thuận), biên 2 nằm ở nhánh sông Tiền, đoạn giáp với tỉnh Tiền Giang và biên 3 nằm ở nhánh sông Cổ Chiên, đoạn qua huyện Măng Thít.

2.4.3. Đặc điểm cấu trúc địa chất và cấp phối hạt

Đặc điểm cấu trúc địa chất tại khu vực nghiên cứu được mô tả trong Bảng 2, cụ thể như sau:

Khu vực thị xã Vĩnh Long (kí hiệu mẫu VL1): Bờ phải sông Cổ Chiên, từ trên mặt đến độ sâu 46m được cấu tạo gồm 5 lớp chủ yếu: Lớp cát lấp dày 1,2m; Lớp bùn sét màu xám đen dày 1m; Lớp bùn sét màu xám xanh, xám đen đến màu đen dày 8,3m; Lớp cát hạt mịn màu xám đen, chứa ít sét, màu xám đen dày 10,5m; Lớp bùn sét, sét pha màu xám nâu đến màu đen, dày 24m.

Khu vực thị trấn Cái Vồn (kí hiệu mẫu VL2): Bờ trái sông Hậu, từ trên mặt đến độ sâu 36m được cấu tạo gồm 4 lớp chủ yếu: Lớp sét xám đen, dẻo chảy dày 2,5m; Lớp bùn sét màu

xám nâu đến xám đen dày 13,5m; Lớp sét pha màu xám nâu, đen trạng thái chảy lẫn ít cát dày 7,3m; Lớp sét màu xám xanh đến màu đen trạng thái chảy dày 12m.

Khu vực thị trấn Trà Ôn (kí hiệu mẫu VL3): Bờ trái sông Hậu từ trên mặt đến độ sâu 20m có cấu tạo 3 lớp chủ yếu: Lớp cát thô dày 2m; Lớp bùn sét màu đen, nâu đen chứa lớp kẹp cát hạt mịn màu xám xanh dày 11,5m; Lớp sét pha màu đen, nâu có lẫn ít vỏ sò, hến trạng thái từ chảy đến mềm dẻo dày 6,5m.

Các cấp hạt tính toán cho mô hình gồm 6 cấp hạt với đường kính trung bình các cấp hạt là 0,005mm, 0,0075mm, 0,03mm, 0,075mm, 0,175mm và 0,375mm. Việc chọn những thông số này dựa vào số liệu phân tích mẫu tại một số vị trí trong vùng nghiên cứu vào tháng 3/2013. Thành phần cấp hạt trong các lớp được trình bày trong Bảng 2.

Bảng 2. Bảng thống kê cấp hạt phân bố theo các lớp.

Số hiệu mẫu	Dày (m)	Thành phần cỡ hạt								Hạt sét	Tổng số
		Hạt cát				Hạt bụi					
		To	Vừa	Nhỏ	Thật nhỏ	Bụi	To	Nhỏ			
		Đường kính các hạt (mm) tính theo tỷ lệ %									
	2-1	1-0,5	0,5-0,25	0,25-0,1	0,1-0,05	0,05-0,01	0,01-0,005	<0,005	%		
VL 1 - 1	1,2				1	7	9	13	70	100	
VL 1 - 2	1					9	34	16	41	100	
VL 1 - 3	8,3			8	87	2	3	-	-	100	
VL 1 - 4	10,5				2	12	30	12	44	100	
VL 1 - 5	24				1	16	43	12	28	100	
VL 2 - 1	2,5				1	8	21	16	54	100	
VL 2 - 2	13,5				1,5	13	26	15	44,5	100	
VL 2 - 3	7,3				31	20	16	8	25	100	
VL 2 - 4	12				1	11	27	15,5	45,5	100	
VL 3 - 1	2					11	25	18	46	100	
VL 3 - 2	11,5				1	19	37	9	34	100	
VL 3 - 3	6,5			5,5	13	30	25	5,5	21	100	

2.4.4. Kích bản tính toán

Mô hình được thiết lập và tính toán với ba kích bản sau:

Kịch bản 1 (KB1): Năm 2017, lấy địa hình và thủy văn 2017 để tính toán. Kích bản này đưa ra nhằm đánh giá địa hình đáy khi đã có suy giảm phù sa (giảm 1/3 so với năm 2008).

Kịch bản 2 (KB2): Năm 2008, lấy địa hình và thủy văn 2008 để tính toán. Kích bản này đưa ra nhằm đánh giá địa hình đáy khi chưa có suy giảm phù sa.

Kịch bản 3 (KB3): Năm 2017, lấy địa hình và thủy văn 2017 để tính toán, có thêm yếu tố khai thác cát.

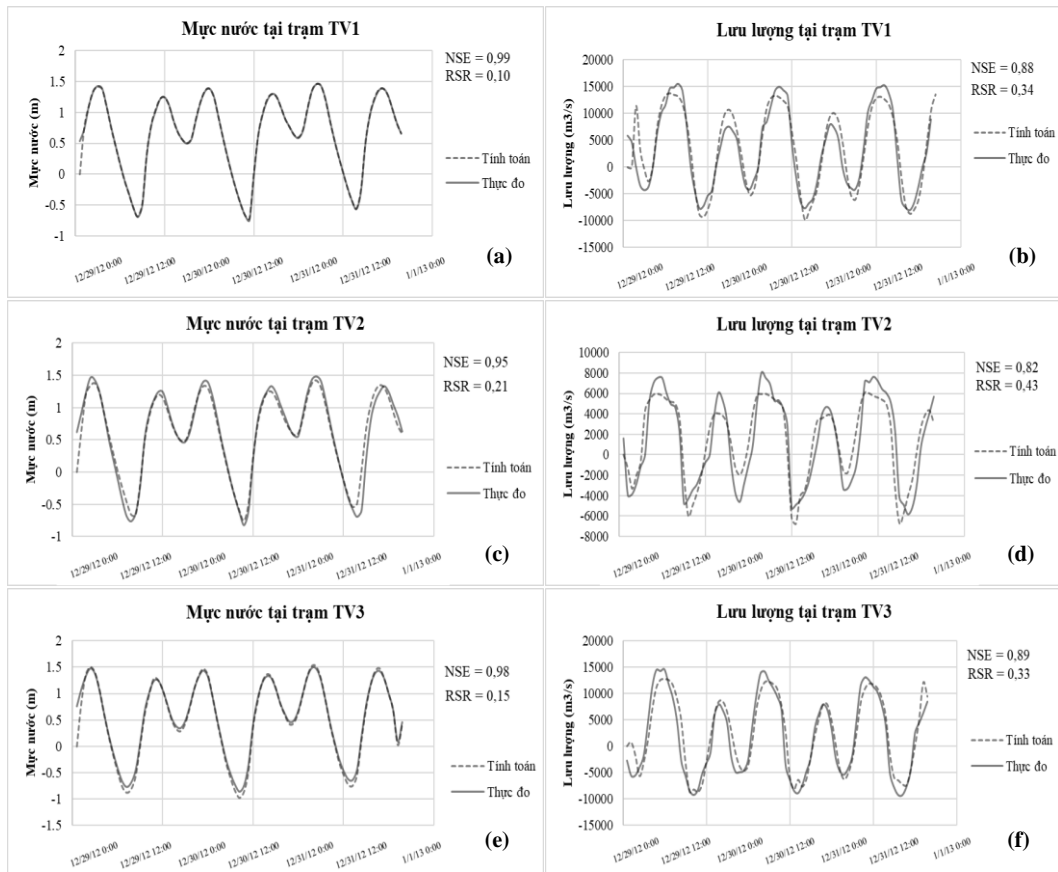
3. Kết quả và thảo luận

3.1. Hiệu chỉnh mô hình

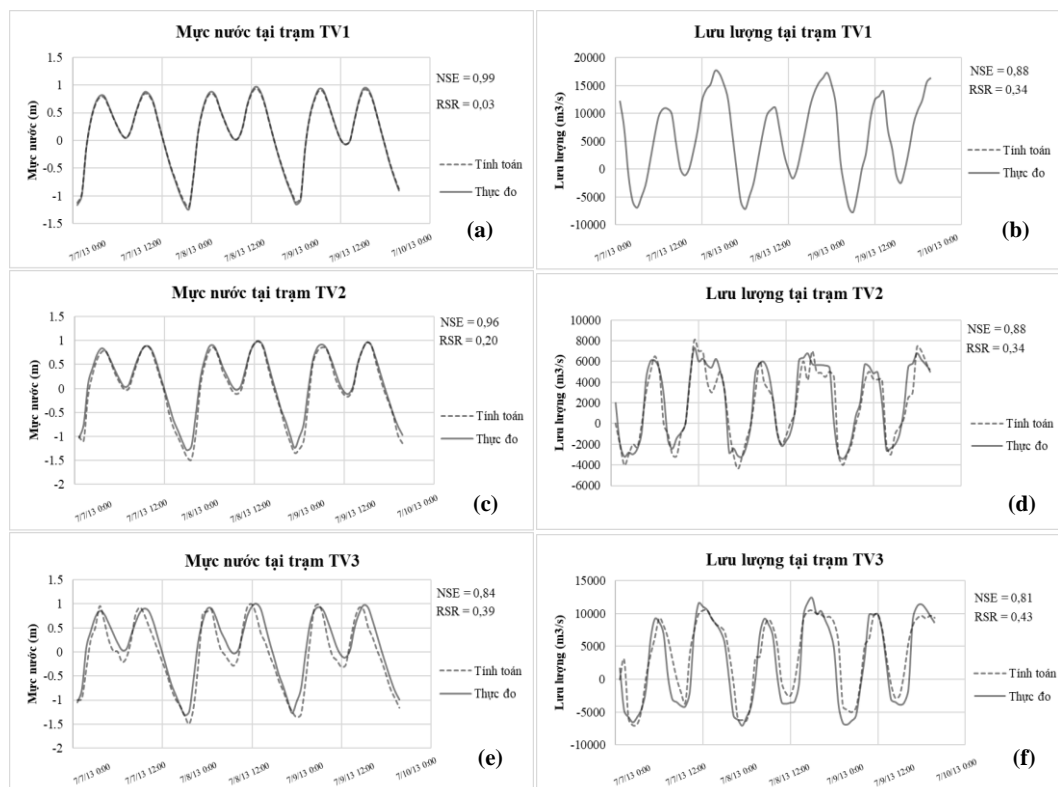
3.1.1. Mô hình thủy lực

Hiệu chỉnh được thực hiện tính toán trong 3 ngày từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013 với bước thời gian tính là 720s (Hình 4). Thông qua việc đánh giá bằng chỉ số Nash-Sutcliffe, ta nhận thấy rằng sai số mực nước giữa tính toán và thực đo ở các trạm đều hầu như không đáng kể (với NSE > 0,8) (Hình 4). Tương tự như mực nước, kết quả so sánh lưu lượng tính toán và thực đo ở các trạm cho thấy chỉ số NSE cũng đều > 0,8.

Sau khi hiệu chỉnh mô hình, giá trị lưu lượng và mực nước tại trạm TV1, TV2 và TV3 tiếp tục được trích xuất từ 0h ngày 7/7/2013 đến 0h ngày 10/7/2013 để phục vụ kiểm định mô hình (Hình 5).



Hình 4. Kết quả so sánh giữa mực nước tính toán và thực đo tại trạm TV1 (a), TV2 (c), TV3 (e) và lưu lượng tính toán và thực đo tại trạm TV1 (b), TV2 (d), TV3 (f) từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013.

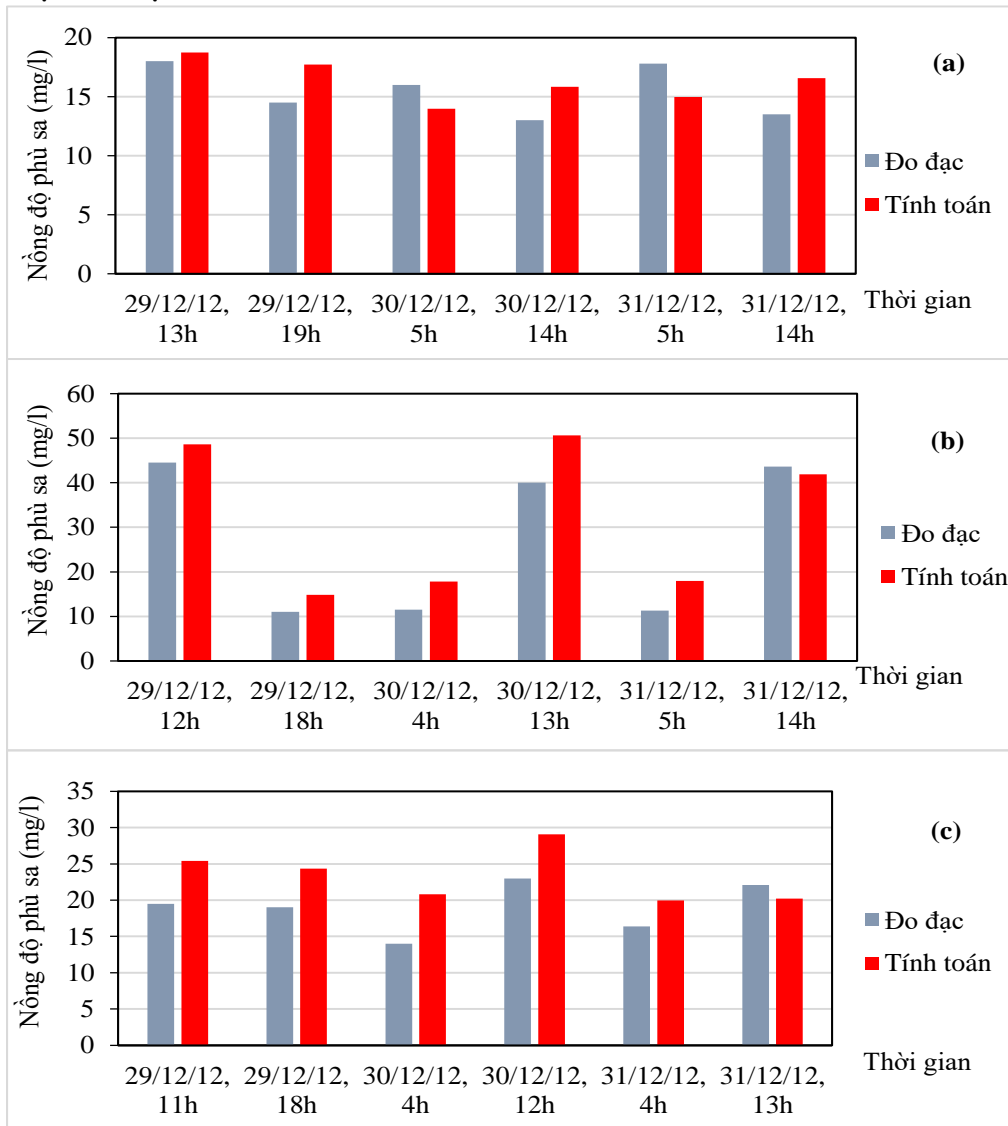


Hình 5. Kết quả kiểm định giữa mực nước tính toán và thực đo tại trạm TV1 (a), TV2 (c), TV3 (e) và lưu lượng tính toán và thực đo tại trạm TV1 (b), TV2 (d), TV3 (f) từ 0h ngày 7/7/2013 đến 0h ngày 10/7/2013.

Như vậy, từ các kết quả hiệu chỉnh và kiểm định mô hình đạt được như trên, ta nhận thấy các yếu tố thủy động lực học do mô hình MIKE 21 FM mô phỏng phù hợp với thực tế. Nên ta có thể dùng các thông số thủy động lực học về độ nhám của mô hình là hợp lý. Bộ thông số mô hình gồm: Hệ số nhám n phân bố tuyến tính theo độ sâu trong khoảng từ 0,015 đến 0,05 và điều kiện ban đầu với mực nước và vận tốc bằng 0.

3.1.2. Mô hình vận chuyển bùn cát

Kết quả tính toán về nồng độ phù sa được xuất ra tại các vị trí từ TV1 tới TV3 (được mô tả trong Hình 3a) trong 3 ngày từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013 để so sánh với số liệu thực đo. Kết quả so sánh nồng độ phù sa giữa tính toán và đo đạc sau khi đã hiệu chỉnh được thể hiện như Hình 6.



Hình 6. Kết quả so sánh nồng độ phù sa giữa tính toán và thực đo tại trạm TV1 (a), TV2 (b), TV3 (c) từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013.

Kết quả so sánh (Hình 6) cho thấy, không có sự chênh lệch quá lớn giữa số liệu tính toán và thực đo tại 03 trạm TV1, TV2 và TV3 từ 0h ngày 29/12/2012 đến 0h ngày 1/1/2013; % sai số vẫn nằm trong giới hạn cho phép. Cụ thể, % sai số trung bình đối với trạm TV1, TV2, TV3 lần lượt là 7,90%, 12,90% và 11,81%.

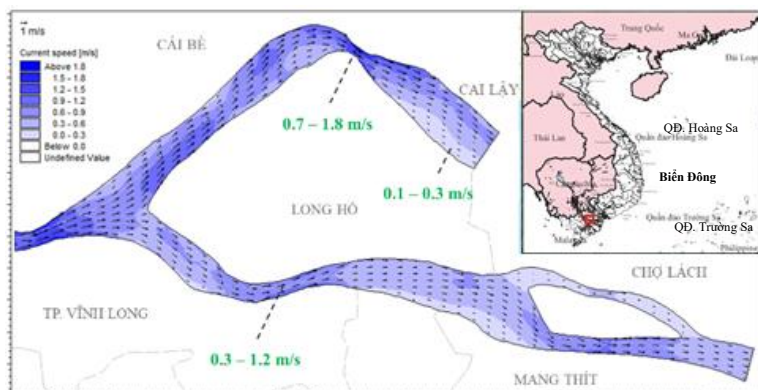
Sau khi hiệu chỉnh và kiểm định mô hình, được bộ thông số tính toán cho mô hình vận chuyển bùn cát của vùng nghiên cứu như Bảng 3.

Bảng 3. Bộ thông số tính toán vận chuyển bùn cát.

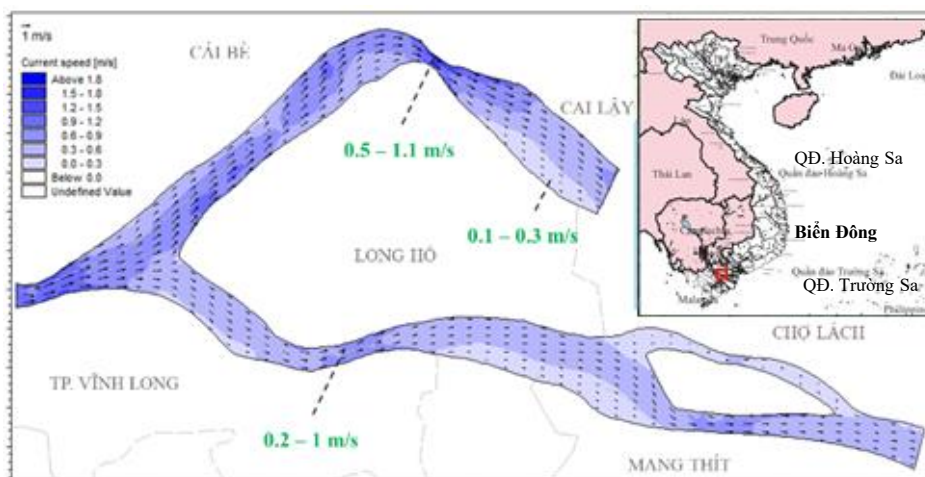
Thông số	Ký hiệu	Đơn vị	Giá trị
Vận tốc lắng đọng	w_s	m/s	$w_{s1} = 0,000022$
			$w_{s2} = 0,00005$
			$w_{s3} = 0,0008$
			$w_{s4} = 0,005$
			$w_{s5} = 0,021$
			$w_{s6} = 0,055$
Ứng suất tới hạn bồi	τ_{cd}	N/m ²	0,05
Ứng suất tới hạn xói	τ_{ce}	N/m ²	Lớp 1: 0,25
			Lớp 2: 0,3
			Lớp 3: 0,45
			Lớp 4: 0,5
			Lớp 5: 0,8
Tốc độ xói	E_0	kg/m ² /s	Lớp 1: 0,0001
			Lớp 2: 0,00008
			Lớp 3: 0,00005
			Lớp 4: 0,00001
			Lớp 5: 0,000005
Độ nhám đáy	k_n	m	0,0001

3.2. Tính toán diễn biến đáy theo KB1 và KB2

Đoạn sông Tiền qua khu vực Mỹ Thuận được xem như một nút thắt khống chế của hình thái sông với phía thượng lưu là đoạn sông phân lạch, cong và phía hạ lưu là đoạn sông phân dòng rẽ nhánh. Quá trình diễn biến lòng sông và đặc trưng hình thái sông của khu vực này chịu ảnh hưởng của điều kiện dòng chảy, dòng bùn cát, diễn biến lòng sông của các đoạn sông phía thượng lưu, đặc biệt là đoạn Sa Đéc. Sông Tiền khu vực này bắt đầu phân lưu, bên bờ phải với sự hình thành sông Cỏ Chiên nên chế độ thủy lực thủy văn của vùng này rất phức tạp.



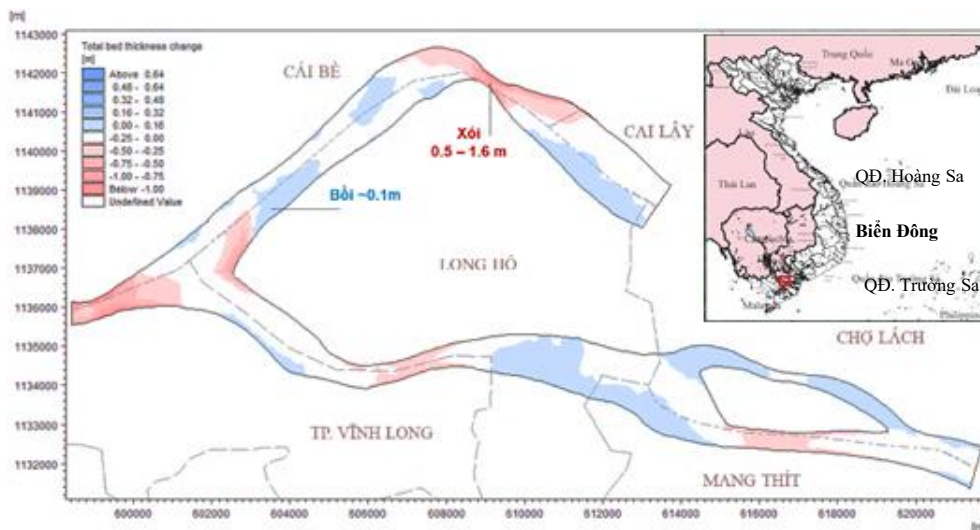
Hình 7. Vận tốc dòng chảy theo KB1.



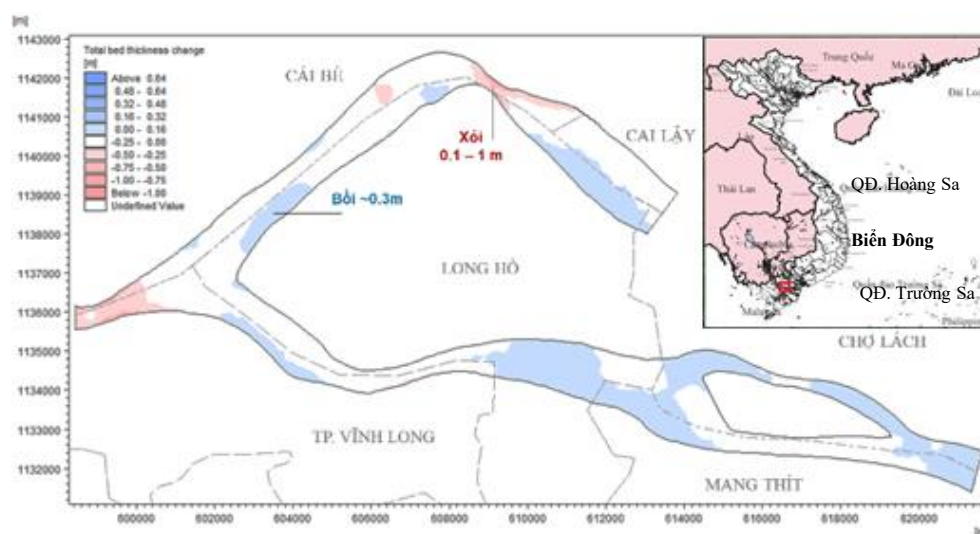
Hình 8. Vận tốc dòng chảy theo KB2.

Theo kết quả từ mô hình thủy lực cho thấy, dòng chủ lưu, luồng sâu của lòng sông sau khi ép sát bờ tả của Mỹ Thuận và khi ra khỏi khu vực co hẹp thì chuyển về phía bờ hữu. Sau khi tách dòng, đi vào sông Cỏ Chiên thì vận tốc lại tăng đột ngột khi qua đoạn sông co hẹp, dòng chủ lưu ép sát bờ tại khu vực TP. Vĩnh Long (Hình 7 và Hình 8). Vận tốc dòng chảy tại vị trí co hẹp đoạn sông Cỏ Chiên này đạt 1,2 m/s theo KB1 và 1 m/s theo KB2. Trong khi đó, vận tốc dòng chảy bên phía sông Tiền đoạn qua huyện Long Hồ cũng đáng kể, đạt tới 1,8 m/s theo KB1 và 1,1 m/s theo KB2. Kết quả tính toán trường vận tốc cho thấy kết quả tính toán thủy lực năm 2017 theo KB1 lớn hơn nhiều so với vận tốc theo KB2.

Về diễn biến lòng dẫn theo 2 KB: Kết quả mô phỏng bồi, xói đáy (Hình 9 và Hình 10) cho thấy ở đoạn này có diện tích bồi lắng phân bố tương đối nhiều ở các đoạn sông thẳng, tốc độ bồi lắng đạt khoảng 0,1m (KB1 năm 2017) và 0,3 m (KB2 năm 2008). Xói đáy xảy ra ở các đoạn sông co hẹp đột ngột, có vận tốc dòng chảy lớn, địa hình đáy sâu. Tốc độ xói đáy mạnh khoảng 0,5-1,6 m sau theo KB1, lớn hơn nhiều so với KB2 (đạt khoảng 0,1-1 m). Như kết quả phân tích thủy lực, tại vị trí co hẹp này, vận tốc dòng chảy ở KB1 lớn hơn KB2, bên cạnh đó, lượng phù sa trong KB1 suy giảm so với KB2 (chỉ còn 1/3 so với KB2). Đó là lý do mà tốc độ xói theo KB1 tăng đáng kể so với KB2. Hồ xói sâu trên sông Cỏ Chiên hình thành tại đoạn bờ thành phố Vĩnh Long đạt đến -40 m, gây ra sập, sụt bờ kè phía bờ phải sông Cỏ Chiên.



Hình 9. Bồi, xói đáy theo KB1.

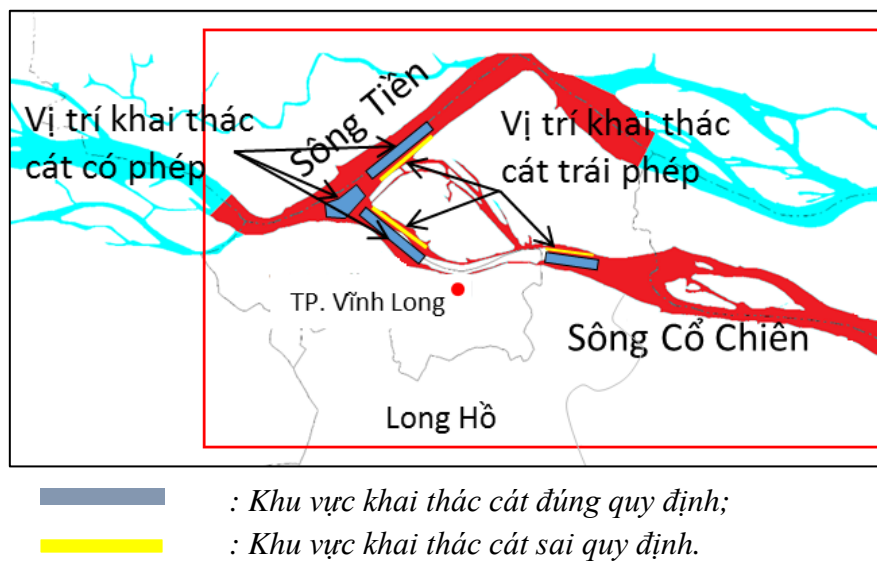


Hình 10. Bồi, xói đáy theo KB2.

So với nghiên cứu [34], tính toán vận tốc tại đoạn Cỏ Chiên chảy qua Vĩnh Long cũng cho thấy dòng chảy sau khi phân lưu (đoạn sông Cỏ Chiên), kết hợp với địa hình sông cong, lồi về phía Vĩnh Long nên dòng chủ lưu áp sát bờ sông phía Vĩnh Long. Kết quả cho thấy dòng chủ lưu ảnh hưởng đến bờ kè sông Cỏ Chiên, đoạn Phường 1, Tp. Vĩnh Long gây nên mất ổn định bờ. Như vậy, kết quả tính toán cũng cho thấy kết quả mô phỏng tương đồng về trường vận tốc.

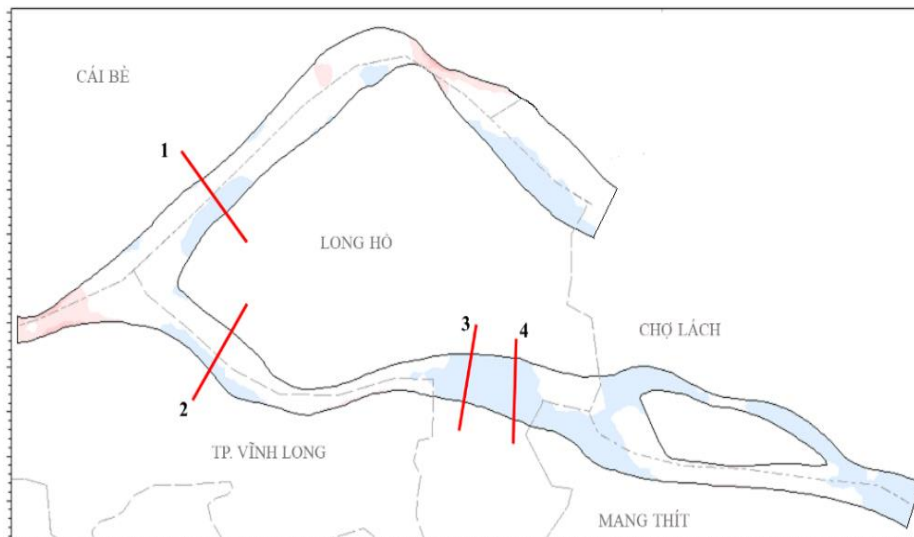
3.3. Ảnh hưởng của khai thác cát đến diễn biến lòng dẫn (KB3)

Toàn tỉnh có 31 khu vực mỏ cát (theo Sở Tài nguyên và Môi trường Vĩnh Long) được phép hoạt động khai thác giai đoạn 2016-2020. Ở đoạn sông này (Hình 11), do có diện tích bồi lắng tương đối nhiều, vì vậy nơi đây được quy hoạch cho phép khai thác cát theo quy định. Vị trí khai thác cát trong vùng này gồm 4 khu vực và được thể hiện trên **Error! Reference source not found.** Công suất khai thác cát là 200.000 m³/năm đối với mỗi khai trường.

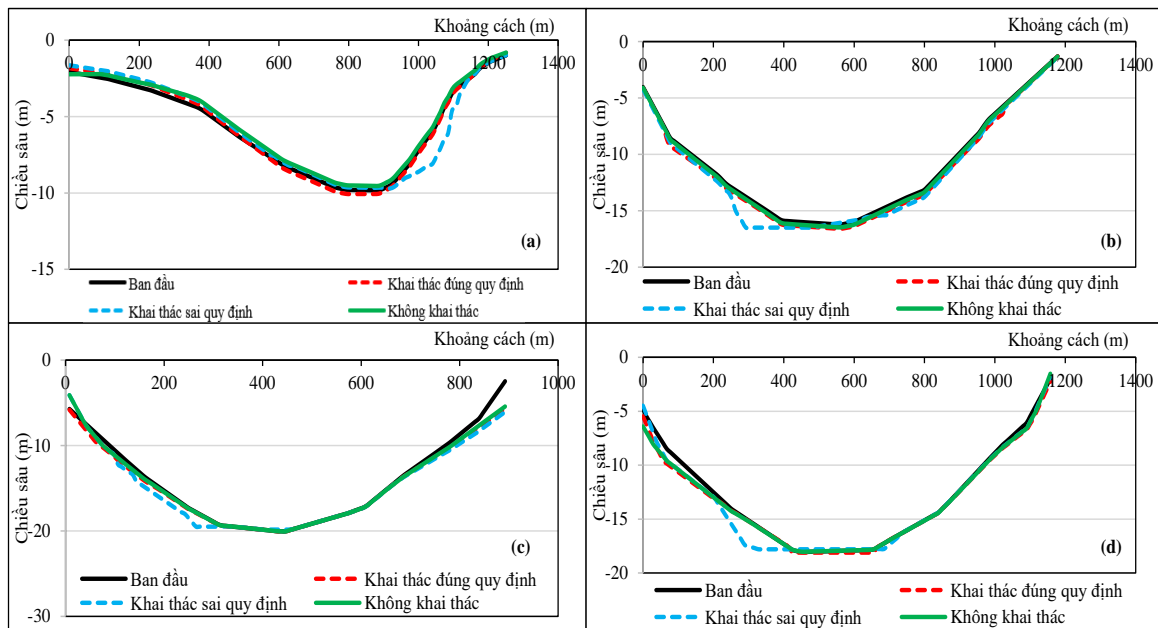


Hình 3. Vị trí khai thác cát theo KB3.

Nhóm nghiên cứu tiến hành thử nghiệm tính toán các nguồn khai thác cát đúng và sai quy định để đánh giá tác động của việc khai thác cát đến vấn đề sạt lở bờ sông. Kết quả mô phỏng tác động khai thác cát đến diễn biến lòng dẫn được xem xét tại 4 mặt cắt (MC) 1, 2, 3, 4 (Hình 12). Diễn biến tại mặt cắt 1, 2, 3, 4 được thể hiện như Hình 13.



Hình 4. Vị trí 4 mặt cắt theo KB3.



Hình 5. Diễn biến lòng dẫn tại MC 1 (a), MC 2 (b), MC 3 (c), MC 4(d).

Theo kết quả tính toán bồi, xói (Hình 13), địa hình đáy trong trường hợp nếu không có khai thác cát (đường liền nét màu xanh lá) so với địa hình ban đầu (đường liền nét màu đen) hầu hết các khu vực tại mặt cắt 2, 3, 4 có hiện tượng bồi ở giữa dòng và xói gần bờ. Qua kết quả tính toán diễn biến lòng dẫn có tính đến kịch bản khai thác cát, nhận thấy rằng nếu khai thác đúng theo quy định (đường nét đứt màu đỏ) thì sẽ không làm ảnh hưởng nhiều đến xói lở bồi lắng lòng sông khu vực này.

Kết quả cho thấy rằng hoạt động khai thác cát có những tác động tích cực và tiêu cực khác nhau. Nếu khai thác không đúng quy mô, vị trí có thể gây ra những vấn đề biến đổi thủy động lực khó lường dẫn đến nguy cơ sạt lở bờ. Cụ thể đối với xã An Bình, khu vực hiện tại hoạt động khai thác cát đang diễn ra, khai thác cát ở đây chủ yếu diễn ra ở sông Tiền thuộc địa phận Vĩnh Long, bên bờ Tiền Giang hiện chưa khai thác. Dòng chủ lưu đang bị lệch về phía bờ phải làm lưu tốc dòng chảy gần bờ tăng lên. Điều này ngược với kết quả tính toán thủy lực. Từ đó cho thấy rằng, vị trí khai thác cát là hết sức quan trọng, thậm chí cho kết quả ngược lại đối với biến đổi thủy lực khu vực khai thác.

Tóm lại, theo kết quả tính toán tác động đến xói lở bờ sông ở khu vực TP. Vĩnh Long, thì khai thác cát trên sông Cổ Chiên là không có lợi cho việc giảm thiểu sạt lở khu vực này, do vậy nên hạn chế khai thác ở đây. Tuy nhiên, khai thác cát đúng quy định lại đem đến kết quả tốt, giúp làm khơi thông dòng chảy và làm cho hình dạng địa hình đáy trở về ban đầu.

Thực tế cho thấy việc khai thác cát trái phép đang diễn ra tràn lan không chỉ ở Vĩnh Long. Nghiên cứu [15] báo cáo rằng việc khai thác cát (ít nhất là trong năm 2018) đã bị cấm ở hai tỉnh Tiền Giang và Bến Tre thông qua nguồn dữ liệu cũng cấp khai thác cát thực tế ở các tỉnh. Tuy nhiên, một thống kê khác cho thấy trữ lượng cát bị khai thác ở Tiền Giang vẫn là 0,51 triệu m³ cát. Những nghiên cứu như vậy ngụ ý rằng việc khai thác trái phép là một vấn đề đang tiếp diễn đối với khu vực này. Điều này cũng chứng minh các thống kê về khai thác cát vẫn còn bị hạn chế, đó cũng chính là hạn chế của nghiên cứu khi chưa mô phỏng diễn biến đáy dưới ảnh hưởng của khai thác cát với đầy đủ dữ liệu về các khai trường. Nghiên cứu [18] đã xác định vấn đề khai thác cát trái phép hiện đang tiếp tục diễn ra từ 2015 đến 2020, và không dừng hẳn. Việc khai thác cát trái phép được ghi nhận nhiều nhất là vào lúc

ban đêm. Tuy nhiên, trong nghiên cứu này, việc đánh giá diễn biến đáy dưới ảnh hưởng của khai thác cát theo thời gian vẫn chưa được thực hiện.

Nghiên cứu [3] cũng xác định hố xói gần bờ tại khu vực sạt lở Tp. Vĩnh Long (Hòa Ninh) được hình thành từ hoạt động khai thác cát. Hố xói này ngày càng mở rộng và ảnh hưởng đến độ ổn định bờ. Đoạn khai thác cát như mô phỏng của [3] có xu thế xói bồi xen kẽ với tốc độ trung bình từ 0,1÷0,2 m/năm. Kết quả này khá tương đồng với KB2 về tốc độ bồi ở những đoạn thẳng (đoạn khai thác cát) của nhóm nghiên cứu.

4. Kết luận

Nghiên cứu phân tích diễn biến hình thái đáy có xét đến yếu tố suy giảm phù sa và khai thác cát tại khu vực sông Tiền đoạn chảy qua tỉnh Vĩnh Long thông qua tính toán chế độ thủy động lực và chuyển tải phù sa bằng mô hình MIKE 21 FM. Kết quả mô phỏng bồi, xói đáy cho thấy, ở đoạn sông Tiền qua khu vực Mỹ Thuận có diện tích bồi lắng phân bố tương đối nhiều ở các đoạn sông thẳng, tốc độ bồi lắng đạt khoảng 0,1m (KB1 năm 2017) và 0,3 m (KB2 năm 2008). Xói đáy xảy ra ở các đoạn sông co hẹp đột ngột, có vận tốc dòng chảy lớn, địa hình đáy sâu. Hố xói sâu trên sông Cổ Chiên hình thành tại đoạn bờ thành phố Vĩnh Long đạt đến -40 m, gây ra sập, sụt bờ kè phía bờ phải sông Cổ Chiên. Đối với ảnh hưởng của hoạt động khai thác cát, kết quả cho thấy rằng hoạt động khai thác cát có những tác động tích cực và tiêu cực khác nhau. Nếu khai thác không đúng quy mô, vị trí có thể gây ra những vấn đề biến đổi thủy động lực khó lường dẫn đến nguy cơ sạt lở bờ. Tuy nhiên, khai thác cát đúng quy định cũng đem đến kết quả tốt, giúp làm khơi thông dòng chảy và làm cho địa hình đáy trở về hình dạng ban đầu.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: T.N.Q.N., T.T.K., N.T.B.; Lựa chọn phương pháp nghiên cứu: T.N.Q.N., T.T.K., N.T.B.; Xử lý số liệu: T.N.Q.N.; Mô hình hóa: T.N.Q.N., T.T.K.; Phân tích kết quả: T.N.Q.N., T.T.K., N.T.B.; Chính sửa bài báo: T.T.K., N.T.B., T.T.T.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 sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Dakheel, A.; Al-Aboodi, A.; Abbas, S. Assessment of Annual Sediment Load Using Mike 21 Model in Khour Al-Zubair Port, South of Iraq. *Basrah J. Eng. Sci.* **2022**, *22*, 108–114.
2. Hồng, N.V.; An, Đ.T. Đánh giá diễn biến hình thái sông dựa trên sự kết hợp mô hình hóa và phân tích ảnh viễn thám. *Tạp chí Khí tượng Thủy văn* **2021**, *726*, 36–46.
3. Huệ, V.H. Đánh giá ổn định bờ sông Cổ Chiên, tỉnh Vĩnh Long. *Tạp chí Khí tượng Thủy văn* **2023**, *752*, 12–28.
4. Kim, T.T.; Minh, N.T.T.; Nga, T.N.Q.; Bầy, N.T.; Phùng, N.K. Nghiên cứu diễn biến đáy khu vực cửa sông Hậu thuộc tỉnh Sóc Trăng. *Tạp chí Khí tượng Thủy văn* **2020**, *717*, 44–55.
5. Novico, F.; Priohandono, Y.A. Analysis of erosion and sedimentation patterns using software of Mike 21 HDFM-MT in The Kapuas Murung River Mouth Central Kalimantan Province. *Bull. Mar. Geol.* **2012**, *27(1)*, 35–53.
6. Wibowo, M.; Khoirunnisa, H.; Wardhani, K.S.; Wijayanti, R. Pemodelan Pola Sedimentasi di Muara Cisadane untuk Mendukung Pengembangan Terpadu Pesisir Ibukota Negara. *Jurnal Kelautan Tropis* **2022**, *25(2)*, 179–190.
7. Zhang, K.; Li, Q.; Zhang, J.; Shi, H.; Yu, J.; Guo, X.; Du, Y. Simulation and analysis of Back siltation in a navigation channel using MIKE 21. *J. Ocean Univ. China* **2022**, *21(4)*, 893–902.

8. Hanebuth, T.J.; Proske, U.; Saito, Y.; Nguyen, V.L.; Ta, T.K.O. Early growth stage of a large delta—Transformation from estuarine-platform to deltaic-progradational conditions (the northeastern Mekong River Delta, Vietnam). *Sediment. Geol.* **2012**, *261*, 108–119.
9. Binh, D.V.; Kantoush, S.; Sumi, T. Changes to long-term discharge and sediment loads in the Vietnamese Mekong Delta caused by upstream dams. *Geomorphology* **2020**, *353*, 107011.
10. Kondolf, G.M.; Rubin, Z.K.; Minear, J. Dams on the Mekong: Cumulative sediment starvation. *Water Resour. Res.* **2014**, *50*(6), 5158–5169.
11. Nowacki, D.J.; Ogston, A.S.; Nittrouer, C.A.; Fricke, A.T.; Van, P.D.T. Sediment dynamics in the lower Mekong River: Transition from tidal river to estuary. *J. Geophys. Res.: Oceans.* **2015**, *120*(9), 6363–6383.
12. Bussi, G.; Darby, S.E.; Whitehead, P.G.; Jin, L.; Dadson, S.J.; Voepel, H.E.; Vasilopoulos, G.; Hackney, C.R.; Hutton, C.; Berchoux, T.; Parsons, D.R.; Nicholas, A. Impact of dams and climate change on suspended sediment flux to the Mekong delta. *Sci. Total Environ.* **2021**, *755*, 142468.
13. Manh, N.V. et al. Future sediment dynamics in the Mekong Delta floodplains: Impacts of hydropower development, climate change and sea level rise. *Global Planet. Change*, **2015**, *127*, 22–33.
14. Hackney, C.R. et al. River bank instability from unsustainable sand mining in the lower Mekong River. *Nat. Sustainability* **2020**, *3*(3), 217–225.
15. Jordan, C. et al. Sand mining in the Mekong Delta revisited-current scales of local sediment deficits. *Sci. Rep.* **2019**, *9*(1), 17823.
16. Kim, T.T. et al. Assessment of the impact of sand mining on bottom morphology in the Mekong River in an Giang Province, Vietnam, using a hydro-morphological model with GPU computing. *Water* **2020**, *12*(10), 2912.
17. Bravard, J.P.; Goichot, M.; Gaillot, S. Geography of sand and gravel mining in the Lower Mekong River. First survey and impact assessment. *EchoGéo* **2013**, *26*, 1–20.
18. Gruel, C.R. et al. New systematically measured sand mining budget for the Mekong Delta reveals rising trends and significant volume underestimations. *Int. J. Appl. Earth Obs.* **2022**, *108*, 102736.
19. Park, E. et al. Dramatic decrease of flood frequency in the Mekong Delta due to river-bed mining and dyke construction. *Sci. Total Environ.* **2020**, *723*, 138066.
20. Loc, H.H. et al. Intensifying saline water intrusion and drought in the Mekong Delta: From physical evidence to policy outlooks. *Sci. Total Environ.* **2021**, *757*, 143919.
21. Luo, X.L. et al. Effects of in-channel sand excavation on the hydrology of the Pearl River Delta. *China. J. Hydrol.* **2007**, *343*(3-4), 230–239.
22. Ashraf, M.A. et al. Sand mining effects, causes and concerns: A case study from Bestari Jaya, Selangor, Peninsular Malaysia. *Sci. Res. Essays.* **2011**, *6*(6), 1216–1231.
23. Rentier, E.; Cammeraat, L. The environmental impacts of river sand mining. *Sci. Total Environ.* **2022**, *838*, 155877.
24. Anthony, E.J. et al. Linking rapid erosion of the Mekong River delta to human activities. *Sci. Rep.* **2015**, *5*(1), 14745.
25. Kim, T.T. et al. Assessment of the impact of sand mining on bottom morphology in the Mekong River in an Giang Province, Vietnam, using a hydro-morphological model with GPU computing. *Water* **2020**, *12*(10), 2912.
26. Brunier, G. et al. Recent morphological changes in the Mekong and Bassac river channels, Mekong delta: The marked impact of river-bed mining and implications for delta destabilisation. *Geomorphology* **2014**, *224*, 177–191.

27. Barman, B.; Kumar, B.; Sarma, A.K. Turbulent flow structures and geomorphic characteristics of a mining affected alluvial channel. *Earth Surface Processes Landforms*, **2018**, 43(9), 1811–1824.
28. Kondolf, G.M. PROFILE: hungry water: effects of dams and gravel mining on river channels. *Environ. Manage.* **1997**, 21(4), 533–551.
29. Barman, B.; Kumar, B.; Sarma, A.K. Dynamic characterization of the migration of a mining pit in an alluvial channel. *Int. J. Sediment Res.* **2019**, 34(2), 155–165.
30. Padmalal, D.; Maya, K. Sand mining: environmental impacts and selected case studies. 2014: Springer.
31. Bầy, N.T. Nghiên cứu hiện tượng bồi lắng, sạt lở bờ sông, xác định nguyên nhân, đề xuất các giải pháp phòng chống, khắc phục ở tỉnh Vĩnh Long. 2014 - 2015.
32. Điệp, N.T.H. et al. Diễn tiến tình hình sạt lở ven bờ sông Tiền và sông Hậu, vùng Đồng bằng sông Cửu Long. *Tạp chí Khoa học Trường Đại học Cần Thơ* **2019**, 55, 125–133.
33. Huệ, V.H. Giải pháp công trình khắc phục sạt lở cồn Thanh Long. *Tạp chí Khí tượng Thủy văn* **2023**, 754, 26–43.
34. Huệ, V.H. Giải pháp công trình ứng phó với dòng chủ lưu áp sạt bờ sông Cổ Chiên, khu vực TP. Vĩnh Long. *Tạp chí Khí tượng Thủy văn* **2023**, 753, 23–36.
35. Tú, L.X. Dự án Đánh giá ổn định bờ sông Cổ Chiên (Khu vực từ rạch Cái Cá đến rạch Bà Bông) tỉnh Vĩnh Long. Viện Khoa học Thủy lợi miền Nam, 2023.
36. Hydrodynamic Module. 2012: Scientific documentation.
37. Hydrodynamic and Transport Module. 2012: Scientific documentation.
38. Thuy, N.T.D. et al. Modelling accretion and erosion processes in the bassac and mekong rivers of the vietnamese mekong delta. in APAC 2019: Proceedings of the 10th International Conference on Asian and Pacific Coasts, 2019, Hanoi, Vietnam. 2020. Springer.

Assessing the impact of sand mining on the riverbed of the Tien River flowing through Vinh Long Province

Nguyen Thi Bay^{1,2*}, Tran Thi Kim³, Tran Thi Thuy An³, Tra Nguyen Quynh Nga^{1,2}

¹ University of Technology; ntbay@hcmut.edu.vn; tnqnga@hcmut.edu.vn

² Vietnam National University Ho Chi Minh City; ntbay@hcmut.edu.vn; tnqnga@hcmut.edu.vn

³ Ho Chi Minh City University of Natural Resources and Environment; ttkim@hcmunre.edu.vn; antran.nrec@gmail.com

Abstract: The excessive extraction of river sand, which causes significant alterations to the geomorphology, is a current social concern. This study focuses on assessing the changes in the riverbed under the influence of decreasing sediment supply and sand extraction activities in the area of the Tien River, flowing through Vinh Long province. The results of bed change simulations show that accretion is relatively high in straight river sections, and the accretion rate in 2017 was lower than in 2008. On the other hand, erosion occurs in narrow, fast-flowing river sections with deep riverbeds. The simulation results indicate that sand extraction has both positive and negative effects. If sand extraction is carried out incorrectly in terms of scale and location, it can lead to unpredictable alterations in hydraulic dynamics and the risk of bank erosion. However, properly regulated sand extraction can yield positive results by improving flow conditions and restoring the original riverbed shape. These findings provide a basis for proposing measures to minimize erosion and assist in disaster risk management due to bank erosion.

Key words: Tien River; Riverbed; Sandmining; Sediment Transport; Sediment deficit.

Bài báo khoa học

Nghiên cứu khả năng ứng dụng thuật toán Random Forest và ảnh vệ tinh Sentinel-2 trong phân loại lớp phủ mặt đất tỉnh Quảng Bình trên nền tảng Google Colab

Phạm Thị Thanh Hòa^{1,2*}, Vũ Ngọc Quang³, Lê Thanh Nghị^{1,2}, Đoàn Thị Nam Phương^{1,2}, Nguyễn Minh Hải¹

¹ Khoa Trắc địa - Bản đồ và Quản lý đất đai, Trường Đại học Mở - Địa chất;

phamthithanhhoa@humg.edu.vn; lethanhngghi@humg.edu.vn;

doanthinamphuong@humg.edu.vn; nguyenminhhai@humg.edu.vn

² Nhóm nghiên cứu Công nghệ Địa tin học trong Khoa học Trái đất (GES), Trường Đại học Mở - Địa chất; phamthithanhhoa@humg.edu.vn; lethanhngghi@humg.edu.vn;

doanthinamphuong@humg.edu.vn

³ Khoa Công trình, Trường Đại học Công nghệ giao thông vận tải; quangvn@utt.edu.vn

*Tác giả liên hệ: phamthithanhhoa@humg.edu.vn; Tel.: +84-977732505

Ban Biên tập nhận bài: 8/9/2023; Ngày phản biện xong: 18/10/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Trong kỷ nguyên công nghệ mới, phương pháp học máy (*Machine learning*) dần thay thế các phương pháp truyền thống trong lĩnh vực viễn thám. Một trong những thuật toán có độ chính xác cao trong phân loại là *Random Forest* (Rừng ngẫu nhiên - RF). Cùng với đó, thay vì phân loại ảnh trên các phần mềm thương mại, nền tảng đám mây Google Colab giúp tối ưu hóa thời gian xử lý với nguồn thư viện phong phú và đặc biệt phù hợp với phương pháp học máy. Vì vậy, nghiên cứu đã tiến hành phân loại lớp phủ mặt đất sử dụng thuật toán Random Forest trên nền tảng Google Colab, thực nghiệm tại tỉnh Quảng Bình với thời gian là tháng 8 năm 2022. Ảnh vệ tinh Sentinel-2 được lựa chọn do độ phân giải không gian cao hơn so với các ảnh miễn phí khác. Đồng thời, nghiên cứu cũng so sánh kết quả phân loại RF trong hai trường hợp: (1) sử dụng bốn kênh ảnh có độ phân giải 10m của ảnh Sentinel-2, (2) kết hợp 4 kênh ảnh trên và các ảnh chỉ số NDVI, NDWI, NDBI. Cả hai trường hợp đều đạt độ chính xác tổng thể trên 90% và Kappa trên 0,9, cho thấy tính khả thi của thuật toán RF. Trong đó, trường hợp (2) đạt độ chính xác cao hơn, khẳng định rằng việc sử dụng các chỉ số quang phổ giúp làm tăng thông tin và cải thiện kết quả phân loại.

Từ khóa: Random Forest; Sentinel-2; Lớp phủ bề mặt; Google Colab.

1. Mở đầu

Trong điều kiện tự nhiên, lớp phủ mặt đất (*land cover*) tích hợp và phản ánh khí hậu, địa chất, đất đai và hệ sinh vật sẵn có của một khu vực tại một thời điểm, theo tháng hoặc năm, có thể hàng thập kỷ hoặc lâu hơn. Nó được xem là nguồn thông tin đầu vào quan trọng trong các nghiên cứu lũ lụt, hạn hán, xói mòn, cũng như cần thiết trong quản lý, giám sát đối tượng lớp phủ mặt đất [1-2]. Nhiều nghiên cứu và nhiều phương pháp được lựa chọn để theo dõi lớp đối tượng này. Cho đến nay, việc chiết tách lớp phủ sử dụng phương pháp chủ yếu là viễn thám (*Remote sensing*) [3]. Việc sử dụng tư liệu viễn thám trong thành lập bản đồ lớp phủ tương đối đơn giản và khá nhanh chóng, được đánh giá là mang lại hiệu quả tốt, vừa có thể tiết kiệm được chi phí và công sức. Công nghệ viễn thám đang ở giai

đoạn phát triển vượt trội, với số lượng ảnh viễn thám và các ứng dụng không ngừng phát triển qua các năm. Số lượng lớn ảnh miễn phí với nhiều độ phân giải khác nhau được dùng trong thành lập bản đồ lớp phủ [4–7].

Trong kỷ nguyên công nghệ mới, chuyển đổi số mang lại nhiều sự đột phá trong nhiều lĩnh vực với việc xuất hiện của Trí tuệ nhân tạo (*Artificial Intelligence - AI*) và Học máy (*Machine Learning*). Khi đó, khoa học công nghệ gắn liền với việc xử lý nguồn dữ liệu lớn (*Big data*) và phương tiện hiện đại. Hiện nay, việc kết hợp mô phỏng chủ đề vào các thuật toán *Machine learning* trở thành một hướng nghiên cứu mới mà nhiều nhà khoa học quan tâm [8–9], trong đó đặc biệt nhấn mạnh sự kết hợp của học máy và lĩnh vực viễn thám. Các kỹ thuật phân loại dựa trên học máy xuất hiện và trở thành hướng tiếp cận mới trong nghiên cứu lớp phủ mặt đất [10]. Một trong những thuật toán học máy có giám sát mang tính khả thi là rừng ngẫu nhiên (*Random Forest - RF*). Các nhà khoa học đã sử dụng *Random Forest* trong thành lập bản đồ lớp phủ mặt đất với độ chính xác cao. Nghiên cứu [11] cho thấy thuật toán *RF* mang lại sự phân loại lớp phủ mặt đất ở phía nam Tây Ban Nha với độ chính xác tổng thể là 92% và chỉ số *Kappa* là 0,92. Trong khi kết quả này ở nghiên cứu của [12] tương ứng là 84,6% và *Kappa* 0,808. Các nghiên cứu [13–14] đánh giá *RF* có độ chính xác cao hơn một số phương pháp phân loại khác như *Maximum Likelihood*, khoảng cách tối thiểu, cây quyết định, mạng Nơ ron nhân tạo và Máy vectơ hỗ trợ (*Support Vector Machine*). Nhìn chung, các nghiên cứu đã chứng minh tính hiệu quả của thuật toán rừng ngẫu nhiên trong nghiên cứu lớp phủ [11–14].

Một vấn đề khác cần chú ý là việc phân loại lớp phủ mặt đất truyền thống thường yêu cầu khối lượng tính toán khổng lồ, đôi khi gây ra áp lực trong quá trình phân tích và xử lý ảnh viễn thám. Do đó, cần lựa chọn một nền tảng xử lý cho phép giảm bớt sự phụ thuộc vào tài nguyên cơ sở hạ tầng máy tính, cũng như giảm bớt gánh nặng về dung lượng ổ cứng máy tính. Một trong những nền tảng cho phép thực hiện trong lĩnh vực viễn thám là *Google Colab*. Sự xuất hiện của *Google Colab* giúp các nhà nghiên cứu thực thi mã xử lý ảnh thông qua kết nối Internet và đặc biệt phù hợp với phương pháp học sâu [15] và học máy [16].

Hầu hết các nghiên cứu trước đây về lớp phủ mặt đất chưa tiếp cận sử dụng nền tảng *Google Colab*. Đồng thời, muốn khẳng định tính hiệu quả về độ chính xác của kỹ thuật học máy, nhóm nghiên cứu đã tích hợp thuật toán *Random Forest* trên nền tảng *Google Colab* để phân loại lớp phủ mặt đất ở tỉnh Quảng Bình, trên cơ sở sử dụng ảnh vệ tinh *Sentinel-2*. Như vậy, thay vì phân loại ảnh trên các phần mềm thương mại, nghiên cứu tiến hành lập trình Python trong môi trường *Google Colab*, giúp tối ưu hóa thời gian xử lý ảnh, tận dụng ưu điểm đơn giản và nguồn thư viện phong phú của ngôn ngữ Python. Ảnh *Sentinel-2* được lựa chọn do được tích hợp sẵn trên nền tảng điện toán đám mây, miễn phí và có độ phân giải không gian cao hơn so với ảnh khác như *Landsat*, *Modis*. Đây chính là hướng tiếp cận mới cho khu vực tỉnh Quảng Bình khi mà trước đây có rất ít các nghiên cứu về lớp phủ.

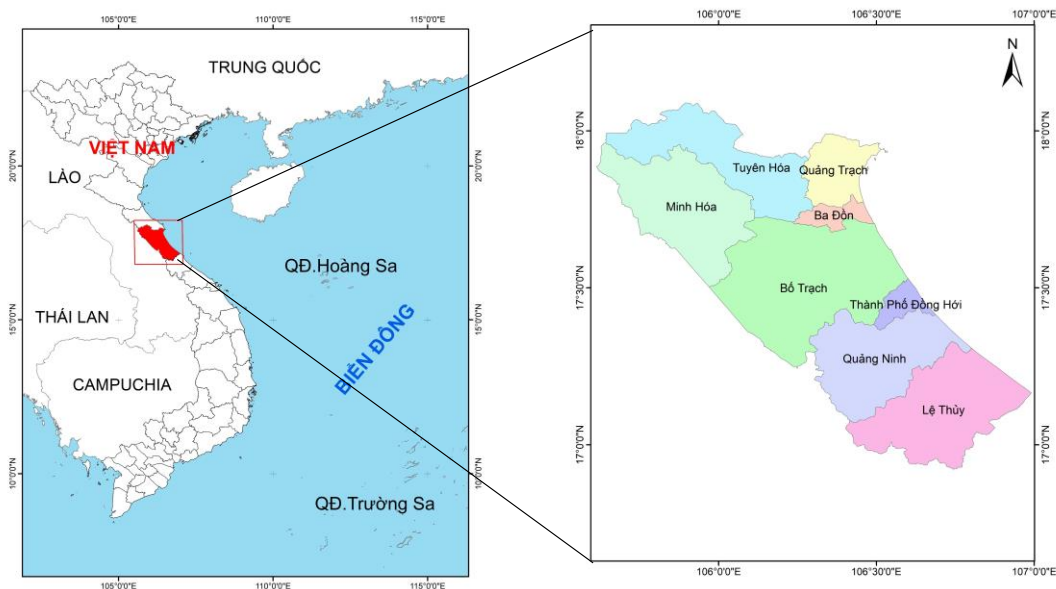
Nghiên cứu sử dụng thuật toán *Random Forest* được tiến hành theo hai hướng: (1) sử dụng bốn kênh ảnh có độ phân giải 10 m (kênh 2, kênh 3, kênh 4, kênh 8) của *Sentinel-2*, (2) sử dụng bốn kênh ảnh có độ phân giải 10m trên và bổ sung thêm các ảnh chỉ số phổ *NDVI* (*Normalized Difference Vegetation Index* - chỉ số thực vật khác biệt chuẩn hóa), *NDWI* (*Normalized Difference Water Index* - chỉ số nước khác biệt chuẩn hóa), *NDBI* (*Normalized Difference Built-up Index* - chỉ số xây dựng khác biệt chuẩn hóa). Mục tiêu cụ thể của nghiên cứu là: (1) Phân loại đối tượng lớp phủ mặt đất ở tỉnh Quảng Bình; (2) Đánh giá tiềm năng của thuật toán *Random Forest* thông qua kết quả đánh giá độ chính xác; (3) So sánh hai hướng tiếp cận; từ đó lựa chọn hướng tối ưu trong phân loại lớp phủ ở tỉnh Quảng Bình khi sử dụng thêm các chỉ số phổ làm tăng lượng thông tin và khả năng nhận biết từng đối tượng đặc trưng; (4) Đánh giá tiềm năng của nền tảng *Google Colab*.

2. Tài liệu thu thập và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Quảng Bình là tỉnh thuộc vùng duyên hải Bắc Trung Bộ, Việt Nam với giới hạn tọa độ địa lý từ 16°55' đến 18°05' vĩ độ Bắc và từ 105°37' đến 107°00' kinh độ Đông. Tỉnh tiếp giáp với tỉnh Hà Tĩnh ở phía Bắc, phía Nam giáp tỉnh Quảng Trị, phía Đông giáp biển với chiều dài trên 116,04 km và đường biên giới phía Tây giáp Lào có tổng chiều dài 222,118 km.

Tỉnh Quảng Bình nằm ở sườn Đông dãy Trường Sơn, với địa hình đồi núi cao hiểm trở, bề ngang hẹp và dốc, nghiêng từ Tây sang Đông và chia cắt ở các khu vực phía Tây của tỉnh. Địa hình phân chia thành các tiểu vùng: Vùng núi cao tập trung ở phía Tây của Tỉnh và nằm dọc theo sườn Đông dãy Trường Sơn; Vùng gò đồi và trung du; Vùng đồng bằng và vùng cát ven biển. Phần lớn diện tích tỉnh là đồi núi (chiếm trên 85% diện tích tự nhiên), các dạng địa hình khác là trung du, đồng bằng duyên hải và bãi cát ven biển chỉ chiếm gần 15%. Do vậy, diện tích đất canh tác lúa tương đối hạn chế và thường xuyên gánh chịu lũ lụt bất thường vào mùa mưa do hệ thống sông suối ngắn, dốc, chảy từ Tây sang Đông (theo Cổng thông tin điện tử tỉnh Quảng Bình).



Hình 1. Vị trí khu vực nghiên cứu.

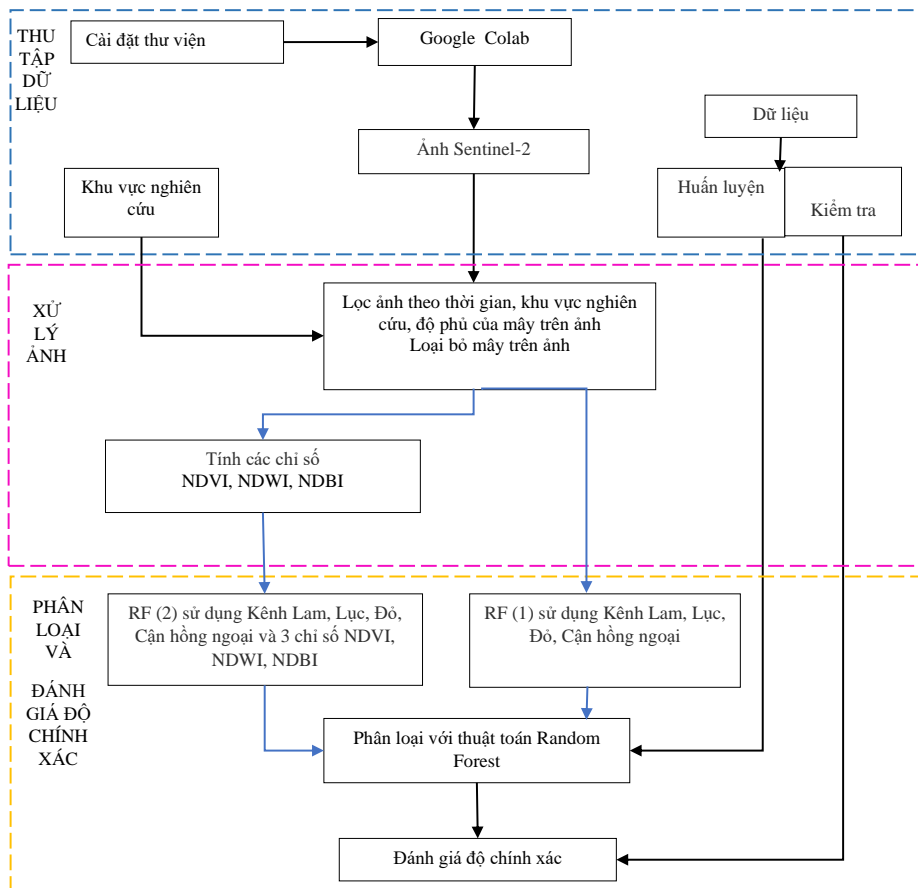
2.2. Dữ liệu ảnh vệ tinh

Nghiên cứu sử dụng ảnh Sentinel-2 với độ che phủ mây dưới 30%, phủ trùm khu vực nghiên cứu vào tháng 8 năm 2022. Sentinel-2 là vệ tinh được nghiên cứu và phát triển bởi ESA - cơ quan hàng không vũ trụ châu Âu, có nhiệm vụ quan sát Trái đất, hỗ trợ các nghiên cứu giám sát thảm thực vật, lớp phủ mặt đất cũng như các tai biến thiên nhiên.

Hệ thống này bao gồm hai vệ tinh quay quanh cực trên cùng một quỹ đạo nhưng lệch pha nhau 180°, được thiết kế với chu kỳ lặp 5 ngày (kết hợp cả hai vệ tinh), độ rộng dải chụp 290 km. Ảnh Sentinel-2 đa phổ (*MultiSpectral Instrument*) gồm 13 kênh phổ: bốn kênh ở độ phân giải không gian 10 m, sáu kênh ở 20 m và ba kênh ở độ phân giải 60 m (theo ESA).

2.3. Phương pháp nghiên cứu

Hình 2 thể hiện sơ đồ quy trình ứng dụng thuật toán Random Forest và ảnh Sentinel-2 trên nền tảng Google Colab để phân loại lớp phủ mặt đất.



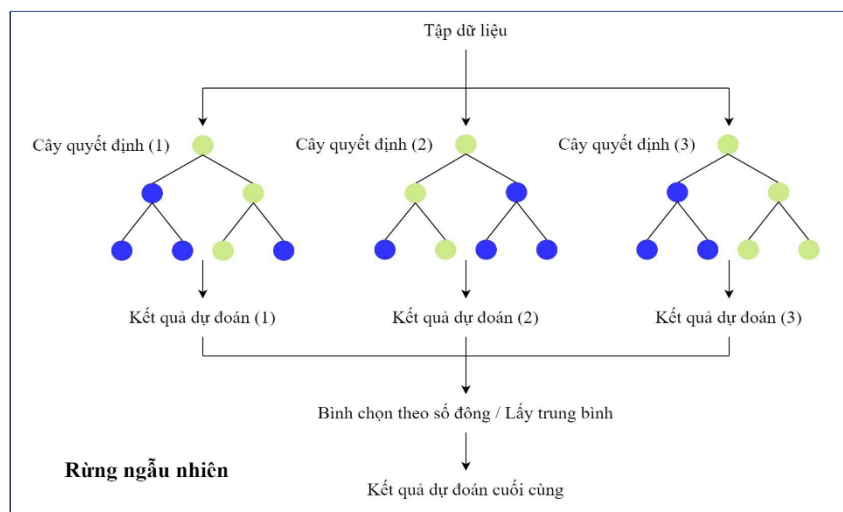
Hình 2. Sơ đồ quy trình.

2.3.1. Thuật toán Random Forest (Rừng ngẫu nhiên - RF)

Random Forest được đề xuất bởi Breiman [17] vào năm 2001. Đây là một thuật toán học máy có giám sát dễ sử dụng và linh hoạt, đồng thời cũng được sử dụng phổ biến để giải quyết nhiệm vụ phân loại và hồi quy.

Random Forest được hiểu là “Rừng ngẫu nhiên”, bắt nguồn từ thuật toán Decision tree (Cây quyết định), nó phát triển nhiều cây quyết định và kết hợp chúng với nhau. Với phân loại Random Forest, cây quyết định được tạo bằng cách sử dụng các tập hợp con ngẫu nhiên khác nhau của dữ liệu và tính năng nhất định. Mỗi cây sẽ cung cấp dự đoán của nó để phân loại. Random Forest dựa trên đa số kết quả dự đoán và lấy kết quả phổ biến nhất làm đầu ra cuối cùng.

Sử dụng thuật toán Random Forest, cần đặc biệt chú ý tham số “number of trees”- số lượng cây quyết định



Hình 3. Thuật toán Random Forest (<https://interactivechaos.com/en/wiki/random-forest>)

bởi nó ảnh hưởng lớn đến kết quả và độ chính xác dự báo [17]. Ngoài ra, còn có các tham số khác liên quan đến số lượng thuộc tính dùng để xây dựng cây.

Random Forest hoạt động như sau:

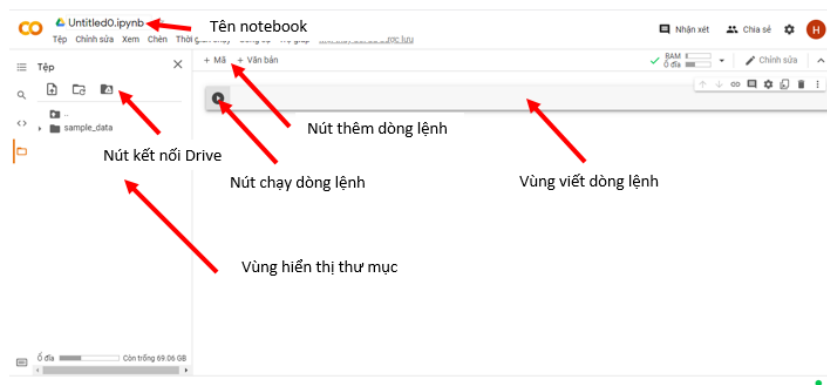
Từ tập dữ liệu D ban đầu, tạo dữ liệu ngẫu nhiên (mẫu bootstrap) Dj có kích thước N sử dụng phương pháp “bagging”. Những dữ liệu còn lại không được lựa chọn tham gia vào quá trình huấn luyện gọi là dữ liệu “out-of-bag”.

Tiếp đó, sử dụng mẫu Dj làm dữ liệu huấn luyện: Các tập con dữ liệu lấy mẫu ngẫu nhiên D1, D2,..., Dk để thiết lập nên các cây quyết định T1, T2,...,Tk. Khi đó, sử dụng các cây đã tạo để nhận kết quả dự đoán. Cuối cùng, tổng hợp kết quả từ các cây quyết định theo hình thức: với bài toán phân loại, lựa chọn theo đa số (tức là số phiếu bầu cao nhất), với bài toán hồi quy, lấy trung bình các giá trị dự đoán.

2.3.2. Nền tảng Google Colab

Bên cạnh sự xuất hiện của Google Earth Engine (GEE) - một nền tảng lưu trữ và xử lý dữ liệu lớn (Big data), Google Research tạo ra Google Colaboratory (còn gọi là Google Colab) cho phép thực thi trên nền tảng đám mây. Google Colab thay thế cho Google Earth Engine khi dễ dàng xử lý dữ liệu trong tài nguyên máy tính [18]. Google Colab phù hợp để thực thi các bài toán đòi hỏi khối lượng tính toán khổng lồ, đi kèm với các thư viện hỗ trợ cho dự án deep learning/machine learning và khoa học dữ liệu.

Google Colab có ưu điểm là được thiết kế chạy mã Python thông qua trình duyệt, cho phép thay thế các phần mềm xử lý ảnh, cũng như không cần nâng cấp phần cứng máy tính. Khi sử dụng Google Colab, các cơ sở hạ tầng như bộ nhớ, khả năng xử lý, đơn vị xử lý đồ họa (GPU) và đơn vị xử lý tensor (TPU) được cung cấp quyền truy cập miễn phí. Và đặc biệt Google Colab giải quyết được các bài toán trong nhiều lĩnh vực với chi phí thấp và thời gian ngắn.



Hình 4. Giao diện Google Colab (<https://interactivechaos.com/en/wiki/random-forest>).

2.3.3. Công thức xác định các chỉ số phổ trong nghiên cứu

Phân loại lớp phủ mặt đất trong nghiên cứu sử dụng thuật toán Random Forest được tiến hành theo hai hướng: (1) sử dụng bốn kênh ảnh có độ phân giải 10m (kênh 2, kênh 3, kênh 4, kênh 8 - tương ứng kênh lam, lục, đỏ và cận hồng ngoại) của Sentinel-2, (2) sử dụng bốn kênh ảnh có độ phân giải 10m và bổ sung thêm các ảnh chỉ số phổ. Nhiều nghiên cứu đã sử dụng các chỉ số phổ để cải thiện kết quả phân loại: nghiên cứu chỉ sử dụng chỉ số NDVI [14, 19], trong khi nghiên cứu sử dụng kết hợp nhiều loại chỉ số phổ khác nhau [20].

Trong nghiên cứu này, các chỉ số phổ được sử dụng bao gồm:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \tag{2}$$

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \tag{3}$$

Trong đó NDVI (*Normalized Difference Vegetation Index*): chỉ số thực vật khác biệt chuẩn hóa [21]. Giá trị NDVI nằm trong khoảng từ -1 đến 1, giúp định lượng sức khỏe và mật độ của thảm thực vật. NDVI cao đại diện cho khu vực có độ che phủ thảm thực vật cao, giá trị âm thể hiện khu vực mặt nước; NDWI (*Normalized Difference Water Index*): chỉ số nước khác biệt chuẩn hóa. Giá trị NDWI giúp phân biệt giữa vùng nước so với các đối tượng thực vật [22]; NDBI (*Normalized Difference Built-up Index*): chỉ số xây dựng khác biệt chuẩn hóa. Giá trị NDBI dùng để nhấn mạnh các khu vực xây dựng [23]; NIR là giá trị bức xạ của sóng hồng ngoại gần; RED là giá trị bức xạ của bước sóng đỏ; GREEN là giá trị bức xạ của bước sóng màu lục; SWIR là giá trị bức xạ của hồng ngoại sóng ngắn.

3. Kết quả và thảo luận

3.1. Thực hiện phân loại

Python đã tích hợp sẵn một số thư viện trong Google Colab như Geopandas/panda - Matplotlib... Bên cạnh đó, Google Colab kết nối với Google Earth Engine để lấy dữ liệu ảnh Sentinel-2 cho khu vực nghiên cứu, với thời gian là tháng 8 năm 2022, độ phủ mây dưới 30%. Kết quả hiển thị trên geemap - gói Python để phân tích và trực quan hóa không gian địa lý. Từ tập dữ liệu ảnh thu thập được, tiến hành loại bỏ mây dựa trên kênh ảnh “probability” (COPERNICUS/S2_CLOUD_PROBABILITY). Sau đó sử dụng hàm “median” và “clip” để ghép các ảnh trong bộ sưu tập và cắt thành một ảnh theo ranh giới khu vực nghiên cứu.

Trong nghiên cứu này, gần 1800 điểm được lựa chọn trong quá trình phân loại, trong đó sử dụng 70% làm dữ liệu huấn luyện (training data) và 30% được sử dụng làm dữ liệu đánh giá (validation data). Tỷ lệ này cũng thường được sử dụng trong các bài toán phân loại bằng Random Forest [20]. Số lượng cây được lựa chọn là 300 trên cơ sở các thử nghiệm của nhóm nghiên cứu. Như vậy theo hướng (1), phân loại RF với việc lựa chọn 4 kênh ảnh dùng để huấn luyện được thể hiện như hình 5 (trên nền tảng Google Colab).

```
[ ] # Select bands for training
bands4 = ['B2', 'B3', 'B4', 'B8']
# Get training
trainingsample4 = S2_StudyArea.select(bands4).sampleRegions(**{
    'collection': trainingfc,
    'properties': ['Class'],
    'scale': 10
})
validationSample4 = S2_StudyArea.select(bands4).sampleRegions(**{
    'collection': testingfc,
    'properties': ['Class'],
    'scale': 10
})
##### RF Classifier Model Building
# ee.Classifier.smileRandomForest(numberOfTrees, variablesPerSplit, minLeafPopulation, bagFraction, maxNodes, seed)
# numberOfTrees: 300, variablesPerSplit: null, minLeafPopulation: 1, maxNodes: null
RFclassifier4 = ee.Classifier.smileRandomForest(300).train(**{
    'features': trainingsample4,
    'classProperty': 'Class',
    'inputProperties': bands4,
})
# Classify the image
RFClassified4 = S2_StudyArea.select(bands4).classify(RFclassifier4)
```

Hình 5. Code trên Google Colab phân loại theo hướng (1) sử dụng 4 kênh ảnh 10m của Sentinel-2.

Theo hướng (2), tính toán bổ sung các chỉ số đã nêu trong mục 2.3.3 với code thể hiện trong hình 6. Hình 7 thể hiện code phân loại theo hướng (2).

```
##### Image Classification with indices: NDVI, NDWI, BSI
# NDVI=(NIR-R)/(NIR+R)
NDVI=S2_StudyArea.normalizedDifference(['B8', 'B4']).rename('NDVI')
#NDWI = (G-NIR)/(G+NIR)
NDWI=S2_StudyArea.normalizedDifference(['B3', 'B8']).rename('NDWI')
#NDBI = (NIR-SWIR1)/(NIR+SWIR1)
NDBI=S2_StudyArea.normalizedDifference(['B11', 'B8']).rename('NDBI')
```

Hình 6. Code tính toán các chỉ số NDVI, NDWI, NDBI.

```

1 # Select 4 bands and 3 spectral indice for training
bands7 = ['B2', 'B3', 'B4', 'B8', 'NDVI', 'NDWI', 'NDBI']
# Get training
trainingsample7 = S2_StudyArea3.select(bands7).sampleRegions(**{
    'collection':trainingfc,
    'properties': ['Class'],
    'scale': 10
})
validationsample7 = S2_StudyArea3.select(bands7).sampleRegions(**{
    'collection':testingfc,
    'properties': ['Class'],
    'scale': 10
})

##### RF Classifier Model Building
# ee.Classifier.smileRandomForest(numberOfTrees, variablesPerSplit, minLeafPopulation, bagFraction, maxNodes, seed)
#numberOfTrees: 300,variablesPerSplit:null,minLeafPopulation: 1,maxNodes: null})
RFclassifier7 = ee.Classifier.smileRandomForest(300).train(**{
    'features': trainingsample7,
    'classProperty': 'Class',
    'inputProperties': bands7,
})
# Classify the image
RFclassified7 = S2_StudyArea3.select(bands7).classify(RFclassifier7)
    
```

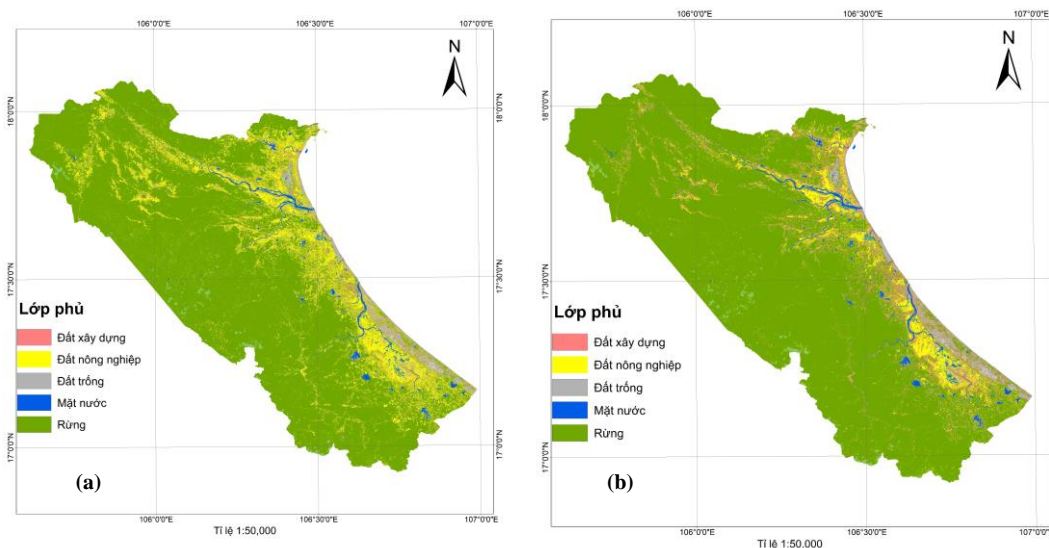
Hình 7. Code phân loại theo hướng (2) sử dụng 4 kênh ảnh 10m của Sentinel-2 và 3 chỉ số phổ.

3.2. Kết quả phân loại

Hình 8 và hình 9 tương ứng là ảnh Sentinel-2 cắt theo khu vực tỉnh Quảng Bình trên giao diện Geemap trong Google Colab và bản đồ kết quả phân loại lớp phủ mặt đất theo hai hướng tiếp cận.



Hình 8. Ảnh Sentinel-2 khu vực tỉnh Quảng Bình vào tháng 8 năm 2022 (Tổ hợp màu thực) trên nền Geemap.



Hình 9. Kết quả phân loại lớp phủ mặt đất tỉnh Quảng Bình với Random Forest: (a) Trường hợp 1: sử dụng bốn kênh ảnh B2, B3, B4, B8; (b) Trường hợp 2: sử dụng bốn kênh ảnh B2, B3, B4, B8 và 3 ảnh chỉ số NDVI, NDWI, NDBI.

Hình 9 thể hiện kết quả phân loại của mô hình Random Forest của khu vực nghiên cứu trong hai trường hợp đã trình bày ở phần 2.3.3. Các lớp phủ được phân loại bao gồm: Rừng, Đất nông nghiệp (gồm đất trồng lúa, hoa màu, cây ăn quả...), Đất trống (gồm đất trống, cát, đá sỏi, đồi trọc ...), Mặt nước (sông, hồ, kênh mương) và Đất xây dựng (gồm dân cư và cơ sở hạ tầng). Nhìn chung kết quả phân loại cho thấy diện tích rừng tỉnh Quảng Bình chiếm tỉ lệ lớn, tập trung chủ yếu ở vùng núi phía Tây tỉnh Quảng Bình. Vùng đất trống (cát) nằm ven biển phía Đông của tỉnh. Các lớp khác như đất nông nghiệp, đất xây dựng phân bố rải rác; vùng mặt nước như sông, hồ, kênh mương được phân loại chi tiết trên ảnh. Tuy nhiên với kết quả phân loại trong trường hợp (1), phần diện tích đất nông nghiệp lớn hơn so với trường hợp (2). Để kiểm tra kết quả phân loại nào tốt hơn, nhóm đã tiến hành đánh giá độ chính xác và đánh giá trực quan trên hai ảnh phân loại.

3.3. Đánh giá độ chính xác

Độ chính xác của hai hướng tiếp cận trên được đánh giá bằng ma trận nhầm lẫn (*Confusion Matrix*), độ chính xác tổng thể (*OA-Overall Accuracy*), Kappa, độ chính xác của nhà sản xuất (PA) và độ chính xác của người dùng (*UA- user's accuracy*).

Ma trận nhầm lẫn là bảng tóm tắt hiệu suất của thuật toán phân loại, trong đó bao gồm giá trị phân loại và giá trị kiểm tra/tham chiếu. Mỗi hàng và mỗi cột sẽ tương ứng với từng lớp phủ được phân loại và lớp kiểm tra/tham chiếu. Đường chéo chính của ma trận liệt kê số lượng pixel được phân loại chính xác. Các giá trị còn lại cho biết mỗi lớp phủ được phân loại thế nào khi so với lớp kiểm tra, lớp nào được phân loại đúng nhiều nhất và lớp nào bị phân loại nhầm vào lớp khác.

Hệ số Kappa [24] là hệ số trong thống kê, dùng để đo đặc độ đồng thuận giữa các thành phần định tính (phân loại), là thước đo mức độ phù hợp tổng thể của ma trận.

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (4)$$

Trong đó P_o là giá trị đồng thuận quan sát được giữa các biến đánh giá; P_e là xác suất của sự đồng thuận ngẫu nhiên P_o và P_e được tính toán dựa trên dữ liệu trong bảng ma trận để tính toán xác suất ngẫu nhiên cho mỗi nhóm.

Một thước đo cơ bản là độ chính xác tổng thể, được tính bằng cách chia các pixel được phân loại chính xác (tổng các giá trị trong đường chéo chính) cho tổng số pixel được kiểm tra.

$$OA = \frac{N}{T} \quad (5)$$

Trong đó N là Số lượng các pixel được phân loại chính xác; T là Tổng số lượng các pixel được kiểm tra.

Bên cạnh độ chính xác tổng thể, độ chính xác phân loại của từng lớp riêng lẻ có thể được tính theo cách tương tự: độ chính xác của người dùng (còn gọi là Độ chính xác sai sót/sử dụng) và độ chính xác của nhà sản xuất PA - Producer's accuracy (còn gọi là Độ chính xác thực hiện). Độ chính xác của nhà sản xuất được tính bằng cách chia số pixel chính xác trong một lớp chia cho tổng số pixel được lấy từ dữ liệu tham chiếu. Độ chính xác của người dùng UA-User's accuracy, về cơ bản cho biết mức độ thường xuyên mà lớp trên bản đồ sẽ thực sự hiện diện trên mặt đất, được tính là các pixel được phân loại chính xác trong một lớp chia cho tổng số pixel được phân loại trong lớp đó [2]. Độ chính xác của kết quả phân loại theo hai hướng tiếp cận trên được thể hiện trong bảng 1 đến bảng 4.

Bảng 1. Ma trận nhầm lẫn của phương pháp phân loại RF 4 kênh ảnh (đơn vị: pixel).

Lớp phủ	Đất xây dựng	Đất nông nghiệp	Đất trống	Mặt nước	Rừng
Đất xây dựng	167	0	3	0	1
Đất nông nghiệp	5	97	0	0	6
Đất trống	3	5	64	1	1
Mặt nước	0	2	0	77	0
Rừng	2	5	0	0	98

Bảng 2. Ma trận nhầm lẫn của phương pháp phân loại RF sử dụng 4 kênh ảnh và 3 chỉ số phổ (đơn vị: pixel).

Lớp phủ	Đất xây dựng	Đất nông nghiệp	Đất trống	Mặt nước	Rừng
Đất xây dựng	166	1	3	0	1
Đất nông nghiệp	5	98	0	1	4
Đất trống	0	5	67	1	1
Mặt nước	0	1	0	78	0
Rừng	1	1	0	0	103

Bảng 3. Độ chính xác tổng thể (OA) và hệ số kappa trong 2 trường hợp.

Trường hợp phân loại RF	Độ chính xác tổng thể (OA) %	Kappa
(1) 4 kênh ảnh	93,7	0,92
(2) 4 kênh ảnh và 3 kênh ảnh chỉ số	95,3	0,94

Bảng 4. Độ chính xác thực hiện PA và độ chính xác sử dụng UA trong 2 trường hợp.

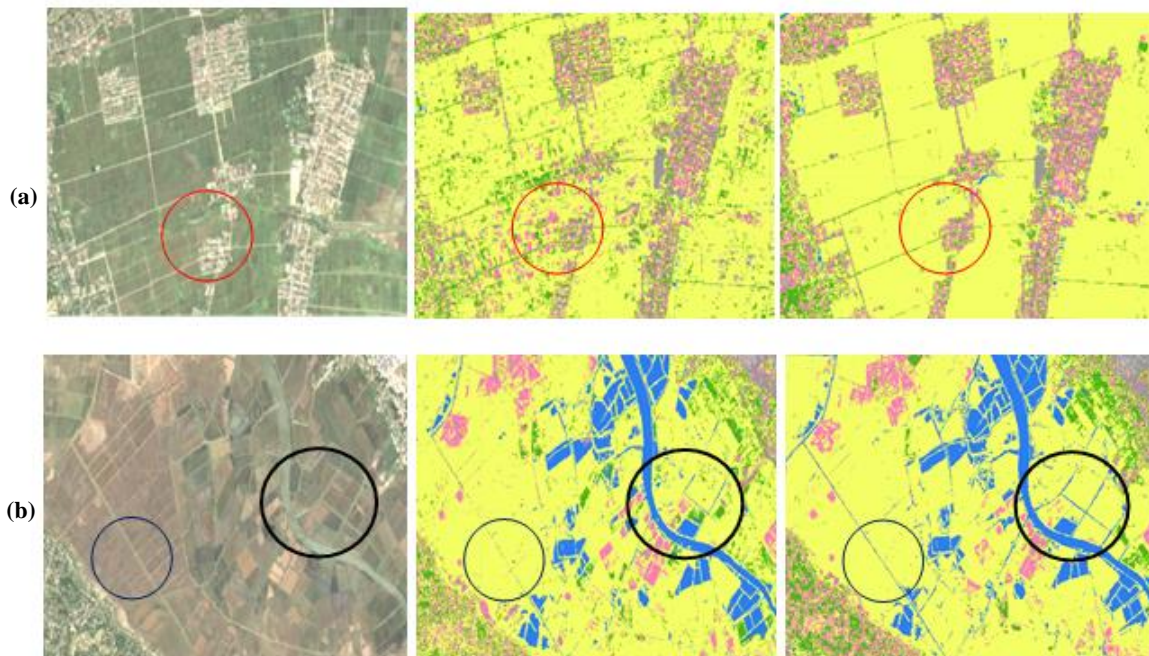
Trường hợp phân loại RF \ Lớp phủ	RF 4 kênh ảnh		RF 4 kênh ảnh và 3 kênh chỉ số	
	UA (%)	PA (%)	UA (%)	PA (%)
Đất xây dựng	94,4	97,7	96,5	97,1
Đất nông nghiệp	89,0	89,8	92,5	90,7
Đất trống	95,5	86,5	95,7	90,5
Mặt nước	98,7	97,5	97,5	98,7
Rừng	92,5	93,3	94,5	98,1

3.4. Thảo luận

Qua số liệu từ bảng 1 đến bảng 4, dễ dàng nhận thấy kết quả phân loại đều đạt mức cao (trên 90% với độ chính xác tổng thể OA và trên 0,9 với Kappa) khi sử dụng thuật toán Random Forest. Các lớp đối tượng về cơ bản có sự phân tách với các lớp còn lại. Trong đó, cả hai trường hợp đều cho thấy kết quả lớp mặt nước được đánh giá có mức độ phân loại đạt cao nhất, số pixel lẫn rất ít so với lớp đất nông nghiệp và không lẫn bất kỳ pixel nào với các lớp khác (Bảng 1 và Bảng 2), giá trị PA và UA gần như đạt mức tuyệt đối (Bảng 4). Điều này thể hiện rằng mẫu của lớp mặt nước thể hiện đúng đặc trưng riêng.

Các thứ tự mức chính xác giảm dần từ rừng, đất xây dựng đến đất nông nghiệp và đất trống. Về đất nông nghiệp, pixel lớp này bị nhận nhầm với pixel rừng do phản xạ phổ nên màu sắc đối tượng khá giống nhau; còn một phần bị nhầm với đất xây dựng vì đất này bao gồm cả đất dân cư, có thể xảy ra trường hợp nhận lớp thực vật trong khu dân cư thành đất nông nghiệp. Đối với vùng đất trống, giá trị PA ở hai trường hợp đều thấp hơn, pixel bị lẫn chủ yếu với đất nông nghiệp (Bảng 1, Bảng 2) do ảnh thu nhận trong thời kỳ mùa khô, một số ruộng khô trên ảnh có màu gần tương đồng với màu đất trống (cụ thể là các vùng đồi trọc), tuy nhiên mức độ lẫn không lớn, và các chỉ số đánh giá độ chính xác đều cao (trên 85%). Một số pixel đất trống lẫn với đất xây dựng trong trường hợp (1), nhưng trong trường hợp (2) thì không lẫn do có sử dụng thêm chỉ số NDBI. Tương tự với lớp Rừng, trong trường hợp (2) pixel rừng lẫn với đất nông nghiệp ít hơn trường hợp (1) nhờ vào việc kết hợp chỉ số NDVI.

Với hai hướng tiếp cận, kết quả tốt hơn khi phân loại với thuật toán Random Forest trong trường hợp có sử dụng thêm các chỉ số phổ ngoài các kênh ảnh có sẵn trên ảnh Sentinel-2. Cụ thể trong bảng 3, kết quả mô hình (1) RF 4 kênh ảnh có độ chính xác tổng thể OA là 93,7% và hệ số Kappa là 0,92, thấp hơn so với mô hình (2) (OA = 95,3%, Kappa = 0,94). Bảng 4 cho thấy các giá trị UA và PA trong trường hợp (2) cơ bản cao hơn trường hợp (1). Nghiên cứu tiến hành so sánh hai ảnh kết quả phân loại như trong hình 10.



Hình 10. (a, b) So sánh kết quả phân loại lớp phủ mặt đất: Cột 1 là ảnh chụp một phần ảnh Sentinel-2, cột 2 là kết quả phân loại RF 4 kênh ảnh (RF trường hợp 1), cột 3 là kết quả phân loại RF 4 kênh ảnh và 3 chỉ số phổ NDVI, NDWI, NDBI (RF trường hợp 2). Các vòng tròn màu đỏ làm nổi bật các vùng có sự khác biệt giữa kết quả phân loại RF trường hợp (1) và (2) liên quan đến dân cư-đất nông nghiệp. Các vòng tròn màu đen thể hiện sự khác biệt nhận biết lớp mặt nước giữa RF trường hợp (1) và (2).

Hình 10a mô tả một phần khu vực với đất nông nghiệp và đất xây dựng. Có thể nhận thấy rằng, mô hình RF với trường hợp 1 không phân loại tốt: đất nông nghiệp lẫn với màu xanh của rừng, pixel dân cư rải rác, lẫn vùng dân cư và đất nông nghiệp; trong khi trường hợp có bổ sung các chỉ số phổ cho kết quả phân loại rõ ràng (khu vực được khoanh hình tròn đỏ). Hình 10b thể hiện đối tượng mặt nước được phân loại khá rõ trong cả hai trường hợp. Tuy nhiên, tại khu vực vòng tròn đen, kết quả phân loại trong trường hợp sử dụng 4 kênh ảnh cho thấy một số vùng mặt nước, cụ thể là các kênh mương, bị nhầm lẫn với lớp đất nông nghiệp. Mặt khác, mô hình RF (2) sử dụng 4 kênh ảnh Sentinel-2 và 3 chỉ số phổ đã phân loại chính xác các lớp phủ này do được bổ sung chỉ số về mặt nước NDWI. Việc đánh giá trực quan cho thấy, sự nhầm lẫn phổ giữa các lớp phủ mặt đất có mức độ cao hơn trong mô hình RF 4 kênh phổ so với kết quả của mô hình RF trong trường hợp còn lại.

So sánh kết quả đạt được với nghiên cứu [25], mặc dù sử dụng dữ liệu ảnh khác nhau, nhưng kết quả phân loại lớp phủ tỉnh Quảng Bình đạt được độ chính xác cao khi sử dụng thuật toán Random Forest. Nghiên cứu [25] cho độ chính xác tổng thể 88,8% và hệ số kappa 0.85 trong khi RF (1) là 93,7% và 0,92, RF (2) là 95,3% và 0,94. Các kết quả của [25] của cũng khá tương đồng với bài báo này ở độ chính xác PA và UA: phân loại RF cho kết quả tốt nhất ở đối tượng mặt nước, rừng, còn đất xây dựng và đất nông nghiệp có chỉ số PA và UA thấp hơn.

Nhìn chung, trên cơ sở so sánh độ chính xác và so sánh trực quan hai hướng tiếp cận của thuật toán Random Forest trong bài toán phân loại (Hình 10), khẳng định rằng việc sử

dụng các chỉ số quang phổ giúp làm tăng thông tin và cải thiện kết quả phân loại. Trong khi đó, công cụ Google Colab đã hoàn thành nhiệm vụ chạy các tập lệnh với thời gian là 308 giây. Con số này cũng thể hiện lợi thế về thời gian tính toán của Google Colab [18], đồng thời cho thấy tiềm năng của nền tảng này trong phân loại lớp phủ mặt đất.

4. Kết luận

Nghiên cứu đã phân loại 5 nhóm lớp phủ mặt đất tỉnh Quảng Bình vào tháng 8 năm 2022 sử dụng ảnh Sentinel-2 và thuật toán Random Forest, đồng thời cũng đánh giá hiệu quả của việc kết hợp các băng tần khác nhau của ảnh Sentinel-2 với các chỉ số phổ đặc trưng. Với các điểm đào tạo (*training data*) trong nghiên cứu, kết quả thử nghiệm cho thấy rằng mô hình RF (2) với 4 kênh phổ (các kênh lam, xanh lục, đỏ và cận hồng ngoại) kết hợp với các chỉ số phổ NDVI, NDWI, NDBI đã đạt được độ chính xác phân loại tổng thể là 95,3% ($\kappa = 0,94$), tốt hơn so với mô hình RF 4 kênh phổ. Nhìn chung, với hướng đi mới này trong phân loại lớp phủ mặt đất, cần có thêm các nghiên cứu toàn diện hơn như sử dụng thêm các ảnh độ phân giải cao, các chỉ số phổ, hoặc kết hợp các phương pháp, mô hình khác nhau.

Bên cạnh việc khẳng định tính khả thi của thuật toán Random Forest trong bài toán phân loại, nghiên cứu cũng đánh giá hiệu quả xử lý của nền tảng Google Colab với thời gian nhanh chóng. Đây là một công cụ thể hiện sự vượt trội trong lĩnh vực viễn thám, khi người dùng có thể bắt đầu mã hóa các mô hình khoa học sử dụng ngôn ngữ lập trình Python thông qua các trình duyệt.

Nghiên cứu đã thực hiện đầy đủ theo mục tiêu đề ra, tuy nhiên còn có một số hạn chế như sau: Google Colab miễn phí nhưng bị giới hạn về thời gian, bị ngắt kết nối khi không có tương tác của người dùng, phụ thuộc vào Internet và Google Drive. Vì vậy, việc sử dụng một phiên bản khác như Google Colab Pro sẽ mang lại hiệu quả hơn trong tương lai.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: P.T.T.H.; Thu thập dữ liệu: L.T.N., N.M.H.; Xử lý số liệu: V.N.Q., Đ.T.N.P.; Thực hiện lập trình: P.T.T.H.; Viết bản thảo bài báo: P.T.T.H., V.N.Q.; Chỉnh sửa bài báo: P.T.T.H., N.M.H.

Lời cảm ơn: Bài báo hoàn thành nhờ vào kết quả của đề tài nghiên cứu khoa học cấp cơ sở mã số T23 - 41 được hỗ trợ kinh phí từ Trường Đại học Mở Địa chất, Hà Nội: “Nghiên cứu ứng dụng phương pháp Machine Learning với thuật toán Random Forest trong thành lập bản đồ lớp phủ bề mặt”.

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. Comber, A.; Fisher, P.; Wadsworth, R. What is Land Cover? *Environ. Plann. B: Plann. Des.* **2005**, *32*, 199–209.
2. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62(1)*, 77–89.
3. Aplin, P. Remote sensing: Land cover. *Prog. Phys. Geogr.* **2004**, *28*, 283–293.
4. Liu, J.; Zhuang, D.; Luo, D.; Xiao, X. Land-cover classification of China: Integrated analysis of AVHRR imagery and geophysical data. *Int. J. Remote Sens.* **2003**, *24*, 2485–2500.
5. Peng, X.; Wang, J.; Raed, M.; Gari, J. Land cover mapping from RADARSAT stereo images in a mountainous area of southern Argentina. *Can. J. Remote Sens.* **2003**, *29(1)*, 75–87.

6. Makinde, E.O.; Oyelade, E.O. Land cover mapping using Sentinel-1 SAR and Landsat 8 imageries of Lagos State for 2017. *Environ. Sci. Pollut. Res.* **2020**, *27(1)*, 66–74.
7. Song, K.; Wang, Z.; Liu, Q.; Liu, D.; Ermoshin, V.V.; Ganzei, S.S.; Zhang, B.; Ren, C.; Zeng, L.; Du, J. Land use/land cover (LULC) classification with MODIS time series data and validation in the Amur River Basin. *Environ. Sci. Pollut. Res.* **2011**, *32(1)*, 9–15.
8. Sen, J.; Mehtab, S.; Sen, R.; Dutta, A.; Kherwa, P.; Ahmed, S.; Berry, P.; Khurana, S.; Singh, S.; Cadotte, D.; Anderson, D.; Ost, K.; Akinbo, R.; Daramola, O.; Lainjo, B. *Machine Learning: Algorithms, Models, and Applications*. IntechOpen Series Artificial Intelligence, 2022, 7, pp. 133.
9. Bansal, R.; Singh, J.; Kaur, R. Machine learning and its applications: A Review. *J. Appl. Sci. Comput.* **2019**, *6(6)*, 1392–1398.
10. Yuh, Y.G.; Tracz, W.; Matthews, H.D.; Turner, S.E. Application of machine learning approaches for land cover monitoring in northern Cameroon. *Ecol. Inf.* **2023**, *74*, 101955.
11. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.
12. Tokar, O.; Olena, V.; Lubov, K.; Havryliuk, S.; Korol, M. Using the Random Forest Classification for Land Cover Interpretation of Landsat Images in the Prykarpattia Region of Ukraine. *Proceeding of the 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT) 2018*, pp. 241–244.
13. Kulkarni, A.D.; Lowe, B. Random Forest Algorithm for Land Cover Classification. *Int. J. Recent Innovation Trends Comput. Commun.* **2016**, *4(3)*, 58–63.
14. Tran, V.A.; Le, M.H.; Tran, H.H.; Le, T.N.; Tran, T.A.; Nguyen, C.C.; Ha, T.K. Land cover mapping in Camau province by machine learning algorithms using Sentinel-2 imagery. *Proceeding of the 43th Asian Conference on Remote Sensing (ACRS 2022)*, 2022.
15. Pessoa, T.; Medeiros, R.; Nepomuceno, T.; Bian, G.B.; Albuquerque, V.H.C.; Filho, P.P. Performance analysis of google colab as a tool for accelerating deep learning applications. *IEEE Access.* **2018**, *6*, 61677–61685.
16. Bisong, E. *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. Apress Publishers, 2019, pp. 709.
17. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45(1)*, 5–32.
18. Hải, P.M.; Quang, N.N. Nghiên cứu thử nghiệm kết hợp môi trường làm việc Google Colaboratory và phương pháp học máy (Machine learning) trong phân loại ảnh viễn thám. *Tap chí Khoa học Đo đạc và Bản đồ* **2020**, *43*, 13–17.
19. Jin, Y.; Liu, X.; Chen, Y.; Liang, X. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. *Int. J. Remote Sens.* **2018**, *39(23)*, 8703–8723.
20. Tassi, A.; Gigante, D.; Modica, G.; Di Martino, L.; Vizzari, M. Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sens.* **2021**, *13(12)*, 2299. Doi: 10.3390/rs13122299.
21. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8(2)*, 127–150.
22. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Remote Sens. Environ.* **1996**, *17(7)*, 1425–1432.

23. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*(3), 583–594.
24. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46.
25. Phong, D.H.; Huệ, N. Giám sát và kiểm kê phát thải khí nhà kính (CO₂ tương đương) trên cơ sở phân loại lớp phủ bằng ảnh Sentinel 1 tỉnh Quảng Bình. *Tap chí Khí tượng Thủy văn* **2022**, *735*, 63–73.

Research on the potential of applying Random Forest algorithm and Sentinel-2 to classify land cover in Quang Binh province on the Google Colab platform

Pham Thi Thanh Hoa^{1,2*}, Vu Ngoc Quang³, Le Thanh Nghi^{1,2}, Doan Thi Nam Phuong^{1,2}, Nguyen Minh Hai¹

¹ Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology; phamthithanhhoa@humg.edu.vn; lethanhngghi@humg.edu.vn; doanthinamphuong@humg.edu.vn, nguyenminhhai@humg.edu.vn

² Geomatics in Earth Sciences Research Group, Hanoi University of Mining and Geology; phamthithanhhoa@humg.edu.vn; lethanhngghi@humg.edu.vn; doanthinamphuong@humg.edu.vn

³ Engineering Faculty, University of Transport Technology; quangvn@utt.edu.vn

Abstract: In the era of technology, Machine learning has gradually replaced traditional methods in the field of remote sensing. One of the supervised machine learning algorithms which has high accuracy in land cover classification is Random Forest. Besides, instead of classifying images by commercial software, the cloud platform - Google Colab helps optimize image processing with a massive variety of libraries and is especially appropriate for machine learning methods. Thus, the article classified land cover using the Random Forest algorithm on the Google Colab platform, a case study in Quang Binh province in August 2022. Sentinel-2 image was chosen since it has a spatial resolution higher than other free images. At the same time, a comparison was performed in two cases: (1) using four image bands with 10m spatial resolution, (2) using four image bands with 10m spatial resolution and NDVI, NDWI, NDBI indices. The results with two approaches have an overall accuracy higher than 90% and Kappa above 0,9, showing the feasibility of Random Forest algorithm. Regarding accuracy, the second case has better results, confirming that the use of spectral indices helps increase information and improve classification results.

Keywords: Random Forest; Sentinel-2; Landcover; Google Colab.

Bài báo khoa học

Mô phỏng ô nhiễm bụi PM_{2.5} và phân tích các yếu tố liên quan – Trường hợp tỉnh Cà Mau, Việt Nam

Ngô Thị Xuân^{1,2}, Nguyễn Châu Mỹ Duyên^{1,2}, Bùi Tá Long^{1,2*}

¹ Khoa Môi trường và Tài nguyên, Trường Đại học Bách Khoa Tp. Hồ Chí Minh; ngothixuana3qt@gmail.com; nguyenduyen91@hcmut.edu.vn; longbt62@hcmut.edu.vn

² Đại học Quốc gia Thành phố Hồ Chí Minh; ngothixuana3qt@gmail.com; nguyenduyen91@hcmut.edu.vn; longbt62@hcmut.edu.vn

*Tác giả liên hệ: longbt62@hcmut.edu.vn; Tel.: +84–918017376

Ban Biên tập nhận bài: 5/9/2023; Ngày phản biện xong: 23/10/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Ô nhiễm không khí PM_{2.5} là tác nhân tác động nghiêm trọng tới sức khỏe con người không chỉ tại các khu đô thị lớn, mà còn ảnh hưởng tới các tỉnh không gần khu đô thị, trong trường hợp nghiên cứu này là tỉnh Cà Mau. Hiện tại ủy ban nhân dân tỉnh đang trong quá trình xây dựng kế hoạch quản lý chất lượng môi trường không khí cho toàn tỉnh. Trong đó nhiệm vụ đánh giá ô nhiễm bụi mịn được đặt ra. Để giải quyết mục tiêu này, cặp các mô hình WRF/CMAQ được sử dụng một mặt hình thành bản đồ phân bố ô nhiễm, mặt khác tìm ra sự phụ thuộc giữa mức độ ô nhiễm PM_{2.5} với cả yếu tố phát thải, lẫn khí tượng. Bộ dữ liệu kiểm kê phát thải nhân tạo và sinh học (tự nhiên) từ nguồn số liệu kiểm kê phát thải toàn cầu ECCAD (Emissions of atmospheric compounds and compilation of ancillary data) được sử dụng. Kết quả mô phỏng được thực hiện cho 2 tháng đặc trưng của mùa khô 2020 cho thấy nồng độ PM_{2.5} trung bình 24 giờ của tháng 03/2020 chưa vượt ngưỡng giới hạn cho phép của QCVN 05:13/BTNMT, giá trị nồng độ dao động trong khoảng 7,82-51,72 µg/m³. Giá trị nồng độ PM_{2.5} trung bình 24 giờ giai đoạn tháng 04/2020 là 7,05-114,42 µg/m³. Kết quả phân tích sự phụ thuộc nồng độ vào phát thải và khí tượng cũng đã được phân tích, làm rõ.

Từ khóa: Mô hình WRF/CMAQ; PM_{2.5}; Khí tượng; Phát thải; Cà Mau.

1. Giới thiệu

Chất lượng không khí có tác động lớn đến đời sống, sức khỏe của nhân dân [1–4]. Việt Nam là quốc gia đang phát triển với lượng phát thải khổng lồ phát sinh từ số lượng phương tiện giao thông, các khu công nghiệp, nhà máy nhiệt điện [5–8]; dẫn tới Việt Nam đang phải đối diện với những hậu quả của ô nhiễm không khí, nhất là ô nhiễm bụi PM (*Particulate Matter*). Gần đây, vấn đề ô nhiễm bụi mịn (PM_{2.5}) trở thành mối quan tâm do những tác động tiềm ẩn đến sức khỏe cộng đồng [9–12]; gây ra giảm thiểu tầm nhìn [13, 14] và biến đổi khí hậu [15, 16]. Tuy nhiên các nghiên cứu tại phía Nam Việt Nam chủ yếu tập trung vào Tp.HCM, cùng một số tỉnh thành phụ cận như Bình Dương, Đồng Nai thuộc vùng phát triển kinh tế trọng điểm [17–19], trong khi đó các tỉnh thành khác chưa được quan tâm đúng mức. Theo quy định của Luật Bảo vệ môi trường năm 2022, Ủy ban nhân dân cấp tỉnh có trách nhiệm ban hành và tổ chức thực hiện kế hoạch quản lý chất lượng môi trường không khí cấp tỉnh. Từ đó nhiệm vụ đánh giá chất lượng môi trường không khí của các tỉnh thành, trong đó có Cà Mau được đặt ra. Bụi PM_{2.5} dùng để chỉ các hạt bụi có đường kính động học bé hơn hoặc bằng 2.5 µm [10, 20]. Đây là một thông số quan trọng đánh giá mức độ ô nhiễm môi

trường không khí [21]. Nồng độ bụi PM_{2.5} chịu ảnh hưởng từ các nguồn phát thải nhân tạo như từ các phương tiện giao thông, đốt sinh khối, và đốt nhiên liệu hóa thạch. Các chất lưu huỳnh đioxit (SO₂), nitơ oxit (NO_x), ammonia (NH₃), carbon đen (BC), carbon hữu cơ (OC), và các hợp chất hữu cơ bay hơi không metan (NMVOCs) là những chất điển hình được tạo ra. Một mặt, bụi PM_{2.5} xuất phát từ các nguồn đốt (combustion sources), được xem là dạng PM sơ cấp [20]; mặt khác, bụi PM_{2.5} cũng được hình thành trong khí quyển, được gọi là PM thứ cấp từ các phản ứng hóa học liên quan đến các tiền chất (precursors) dạng khí như SO₂, NO_x, NH₃, và các hợp chất hữu cơ dễ bay hơi (VOCs) [22–24]. Sự hình thành PM thứ cấp phụ thuộc đáng kể vào số lượng tiền chất có trong khí quyển và các điều kiện khí tượng [25], [23]. Sự nhạy cảm đáng kể của nồng độ bụi PM_{2.5} đối với các yếu tố khí tượng thông qua sự biến đổi trong các quá trình vật lý và hóa học [26–29]. Với các thành phần hạt có kích thước nhỏ, bụi PM_{2.5} có thời gian lưu từ vài ngày đến vài tuần và dễ dàng bị phân tán trong không khí [24]. Mặt khác, khi chiều cao lớp biên khí quyển (*Planetary Boundary Layer Height - PBLH*) giảm dẫn đến sự hạn chế quá trình hòa trộn theo chiều dọc và gây ra sự tích lũy của các hạt PM [32, 33]. Cùng với đó, bụi PM_{2.5} có thể được loại bỏ bằng sự rửa trôi qua mưa [21, 24, 34, 35].

Gần đây, các mô hình chất lượng không khí đã được sử dụng rộng rãi để làm rõ vai trò các quá trình vật lý và hóa học liên quan đến ô nhiễm bụi PM_{2.5} [36]. Hiện nay các mô hình có thể thực hiện việc mô phỏng chất lượng không khí từ quy mô toàn cầu (*global scale*) đến quy mô địa phương (*local scale*) với các độ phân giải khác nhau tùy thuộc vào các mục đích nghiên cứu khác nhau [36]. Mô hình toàn cầu GEOS-Chem với độ phân giải từ 5 đến 4° [37], mô hình khu vực như *Weather Research and Forecasting Model/Community Multiscale Air Quality Model (WRF/CMAQ)* với các mức độ phân giải từ 36 đến 4 km [38]. Trong đó, hệ thống mô hình CMAQ đã được sử dụng nhiều nhất bởi khả năng ứng dụng mô phỏng ô nhiễm không khí đa quy mô (đô thị và cấp khu vực), cũng như đa chất ô nhiễm (chất oxy hóa, lắng đọng axit và bụi PM).

Để góp phần vào việc hoạch định chính sách, phục vụ phát triển kinh tế - xã hội và sự nghiệp phát triển bền vững của tỉnh Cà Mau, nghiên cứu này được thực hiện với mục tiêu ứng dụng WRF/CMAQ kết hợp với mô hình thống kê đa biến nhằm: (1) Mô phỏng phân bố PM_{2.5} phạm vi toàn thành phố và (2) Thực hiện đánh giá mức độ ảnh hưởng của 2 yếu tố phát thải, khí tượng lên nồng độ. Kết quả nghiên cứu cho phép đề xuất các giải pháp giảm thiểu tối ưu phát thải, đề xuất lộ trình từ trung hạn tới dài hạn để đảm bảo chi phí và lợi ích kinh tế, hiệu quả về môi trường.

2. Số liệu sử dụng và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Cà Mau nằm ở ven biển ở cực Nam Việt Nam, có dạng hình chữ V, là một trong bốn tỉnh thuộc vùng kinh tế trọng điểm Đồng Bằng Sông Cửu Long. Phần đất liền có tọa độ điểm cực Nam 8°30' vĩ độ Bắc, điểm cực Bắc 9°33' vĩ Bắc, điểm cực Đông 105°24' kinh Đông và điểm cực Tây 104°43' kinh Đông. Phía Bắc giáp tỉnh Kiên Giang, phía Đông Bắc giáp tỉnh Bạc Liêu, phía Đông và Đông Nam giáp biển Đông và phía Tây giáp Vịnh Thái Lan [39]. Diện tích phần đất liền của tỉnh là 5.329 km², bằng 13,13% diện tích vùng đồng bằng sông Cửu Long và bằng 1,58% diện tích cả nước [40]. Địa hình của tỉnh Cà Mau tương đối bằng phẳng và thấp, cao trình phổ biến từ 0,5-1 m so với mực nước biển. Phía Bắc địa hình thấp (cao trung bình từ 0,2-0,5 m), phía Nam có địa hình cao hơn (trung bình 0,2-0,8 m). Bên cạnh đó, địa hình của tỉnh còn bị chia cắt nhiều bởi hệ thống sông rạch chằng chịt. Trong những năm qua tình hình kinh tế - xã hội tỉnh Cà Mau không ngừng phát triển và đạt được những thành tựu đáng kể. Tổng sản phẩm trong tỉnh (GRDP) năm 2009 đạt 13.021 tỷ đồng, GRDP năm 2020 (theo giá so sánh năm 2010) đạt 41.318 tỷ đồng, gấp 1,58 lần so với GRDP năm

2010 (26.037 tỷ đồng). GRDP bình quân đầu người năm 2020 đạt 52,6 triệu đồng (tương đương 2.267 USD), gấp 2,44 lần so với năm 2010 (21,6 triệu đồng/người) [41].

Theo báo cáo hiện trạng môi trường tỉnh Cà Mau năm 2015-2020, nhìn chung các giá trị hàm lượng tổng bụi giai đoạn 2015-2020 tại các vị trí quan trắc qua các năm đều nằm trong giới hạn quy định của QCVN 05:2013/BTNMT (trung bình 1 giờ) (Quy chuẩn kỹ thuật quốc gia về môi trường không khí xung quanh) [39]. Dù hầu hết các giá trị quan trắc đều nằm trong giới hạn cho phép so với QCVN 05:2013/BTNMT tuy nhiên ô nhiễm bụi có xu hướng duy trì ở ngưỡng cao tại khu vực đô thị, khu vực sản xuất công nghiệp đặc biệt là các khu vực gần các trục giao thông. Diễn biến hàm lượng bụi đang có xu hướng tăng dần và gây tác động đến sinh hoạt cũng như sức khỏe người dân trong khu vực.

2.2. Dữ liệu

2.2.1. Dữ liệu khí tượng thực đo

Để kiểm định mô hình khí tượng WRF, bộ số liệu khí tượng thực đo tại các trạm nằm ở huyện U Minh, xã Hòa Thành, huyện Cái Nước, huyện Đầm Dơi, xã Tân Thuận, huyện Năm Căn, xã Tam Giang, huyện Ngọc Hiển, huyện Trần Văn Thời thể hiện ở Bảng 1 và Hình 1 theo chuỗi thời gian trung bình giờ của tháng 03 và tháng 04 năm 2020 được lấy từ trang web <https://www.wunderground.com/> [42]. Tại mỗi trạm đo, nhóm các số liệu về nhiệt độ, vận tốc gió, hướng gió, độ ẩm và áp suất được thu thập, trên cơ sở đó thực hiện so sánh các số liệu này với số liệu chiết xuất từ kết quả chạy mô hình WRF để đánh giá tính chính xác và mức độ sai số của kết quả mô hình khí tượng cho ra

Bảng 1. Vị trí các trạm khí tượng.

STT	Ký hiệu	X	Y	Vị trí
1	KT1	104,9000	9,3000	Huyện U Minh
2	KT2	105,2000	9,1000	Xã Hòa Thành
3	KT3	105,0557	8,9868	Huyện Cái Nước
4	KT4	105,1941	8,9944	Huyện Đầm Dơi
5	KT5	105,3000	9,0000	Xã Tân Thuận
6	KT6	105,0000	8,8000	Huyện Năm Căn
7	KT7	105,2000	8,8000	Xã Tam Giang
8	KT8	105,1000	8,7000	Huyện Ngọc Hiển
9	KT9	105,0000	9,1000	Huyện Trần Văn Thời

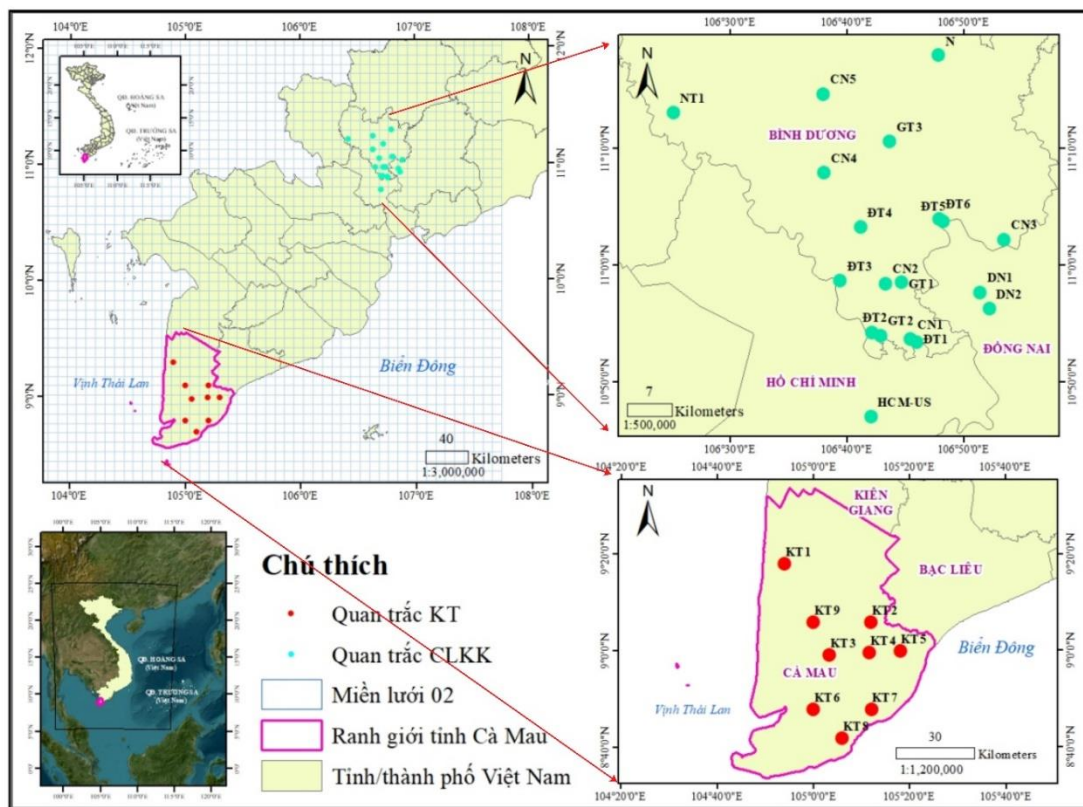
2.2.2. Dữ liệu chất lượng không khí quan trắc

Bộ dữ liệu quan trắc chất lượng không khí năm 2020 để đánh giá hiệu quả của mô phỏng WRF/CMAQ được thu thập từ Trung tâm quan trắc môi trường tỉnh Đồng Nai [43], Trung tâm quan trắc môi trường tỉnh Bình Dương [44] và Trung tâm quan trắc môi trường tp. Hồ Chí Minh [45]. Giai đoạn tháng 03 và 04 nồng độ bụi PM_{2.5} được chuyển đổi từ nồng độ TSP thu thập từ (1) 16 điểm đo đặc tại tỉnh Bình Dương (gồm N, NT1, DT1, DT2, DT3, DT4, DT5, DT6, GT1, GT2, GT3, CN1, CN2, CN3, CN4, và CN5) cho 4 thời điểm quan sát lúc 07.00, 09.00, 11.00, 13.00, 15.00 và 17.00; (2) 2 trạm quan trắc tự động ở Đồng Nai (DN1, DN2); (3) 1 trạm tự động ở Thành phố Hồ Chí Minh (HCM-US). Các vị trí quan trắc được mô tả trong Bảng 2 và Hình 1.

Bảng 2. Vị trí các trạm quan trắc.

STT	Ký hiệu	X	Y	Mô tả khu vực quan trắc
1	N	106,7969	11,2997	Nghĩa trang Liệt sĩ huyện Phú Giáo
2	NT1	106,4178	11,2170	Nông trường cao su Thanh An
3	ĐT1	106,7653	10,8889	Trung tâm hành chính thị xã Dĩ An
4	ĐT2	106,7014	10,9036	UBND thị xã Thuận An

STT	Ký hiệu	X	Y	Mô tả khu vực quan trắc
5	ĐT3	106,6564	10,9772	Trụ sở TT QT TNMT - số 26 Huỳnh Văn Nghệ
6	GT1	106,7442	10,9750	Ngã tư Miếu Ông Cu
7	GT2	106,7144	10,8989	Ngã tư cầu Ông Bó
8	CN1	106,7569	10,8936	Công ty Wimbledon - Đường số 22, KCN Sóng Thần II
9	CN2	106,7214	10,9733	Công ty Victory cách CCN Thuận Giao 300m (điểm GS cũ)
10	CN3	106,8908	11,0361	Khu vực mỏ đá xã Thường Tân, huyện Tân Uyên
11	CN4	106,6331	11,1314	Công ty Orion, đường NA3, KCN Mỹ Phước II
12	ĐT4	106,6858	11,0536	Trung tâm hành chính, thành phố mới Bình Dương
13	ĐT5	106,8036	11,0625	Khu trung tâm thị trấn Mỹ Phước
14	ĐT6	106,7975	11,0656	Thị trấn Uyên Hưng, huyện Tân Uyên
15	CN5	106,6322	11,2442	Khu đô thị - khu công nghiệp Bàu Bàng
16	GT3	106,7267	11,1758	Gần ngã 3 công xanh
17	DN1	106,8558	10,9605	Khu dân cư
18	DN2	106,8693	10,9370	Nút giao thông
19	HCM-US	106,7007	10,7831	Tổng lãnh sự quán Hoa Kỳ



Hình 1. Vị trí các trạm quan trắc chất lượng không khí và khí tượng.

2.2.3. Dữ liệu phát thải

Bộ dữ liệu kiểm kê phát thải được sử dụng trong nghiên cứu này được lấy từ nguồn số liệu kiểm kê phát thải toàn cầu ECCAD (*Emissions of atmospheric compounds and compilation of ancillary data*) (<https://eccad3.sedoo.fr/>) cho tháng 03 và 04 năm 2020. Có ba bộ dữ liệu phát thải được sử dụng, bao gồm: CAMS-GLOB-AIR dữ liệu lượng khí thải toàn cầu từ aircraft, CAMS-GLOB-ANT dữ liệu phát thải nhân tạo toàn cầu, CAMS-GLOB-BIO Lượng phát thải toàn cầu của 29 VOC sinh học từ thực vật. Các bộ dữ liệu được thu thập, sử dụng ở dạng thông lượng trung bình hàng hán (kg.m⁻².s⁻¹) với các mức độ phân giải khác nhau trên phạm vi toàn cầu (0,5° × 0,5°, 0,25° × 0,25° và 0,1° × 0,1°) (*Atmosphere and Service*

2018) được tính toán phân bố không gian về các ô lưới của miền tính Domain D02 (có bao hàm khu vực nghiên cứu là tỉnh Cà Mau) với độ phân giải tăng lên ~ 9 km × 9 km.

2.3. Phương pháp nghiên cứu

Cặp mô hình WRF/CMAQ được ứng dụng để mô phỏng ô nhiễm PM_{2.5} tại tỉnh Cà Mau, lấy năm 2020 nghiên cứu. Mô hình Chất lượng Không khí Đa quy mô Cộng đồng (CMAQ) là mô hình chất lượng không khí bằng số dựa trên các nguyên tắc khoa học đầu tiên để dự đoán nồng độ của các loại khí và hạt trong không khí cũng như sự lắng đọng của các chất ô nhiễm này trở lại bề mặt Trái đất. Để hỗ trợ cộng đồng người dùng CMAQ, EPA và Đại học Bắc Carolina tại Đồi Chapel đã tổ chức trao đổi email cho người dùng và cung cấp đào tạo cho người dùng mới về hệ thống mô hình hóa CMAQ bằng cách tổ chức Trung tâm Hệ thống Phân tích và Mô hình hóa Cộng đồng (CMAS), nơi phân phối phần mềm CMAQ [46].

2.3.1. Thiết lập mô hình WRF/CMAQ kết hợp

Miền tính toán PM_{2.5} mô phỏng lồng nhau hai cấp D01 và D02 được thiết lập trong mô hình WRF/CMAQ, trong đó miền tính D01 có độ phân giải không gian xấp xỉ 27 km × 27 km là miền tính lớn nhất, với 61 cột và 81 hàng, tổng số ô lưới tính của miền D01 là 4661 ô; miền tính D02 có độ phân giải không gian xấp xỉ 9 km × 9 km là miền tính thứ 2 lồng trong miền tính D01, với 50 cột và 56 hàng, tổng số ô lưới của miền D02 là 2592 ô. Các thông số kỹ thuật của các miền tính toán D01 và D02 được trình bày bên dưới. Miền tính D01 bao phủ hết toàn bộ Việt Nam; miền tính D02 bao gồm hầu hết các tỉnh phía Nam, đây là khu vực có thể có ảnh hưởng đáng kể nhất đối với khu vực nghiên cứu. Các điều kiện biên của D02 được xác định theo D01. Các thông số kỹ thuật của các miền tính toán D01, D02 được trình bày trong Bảng 3.

Bảng 3. Miền không gian tính toán mô phỏng PM_{2.5} cho hệ thống mô hình kết hợp WRF/CMAQ.

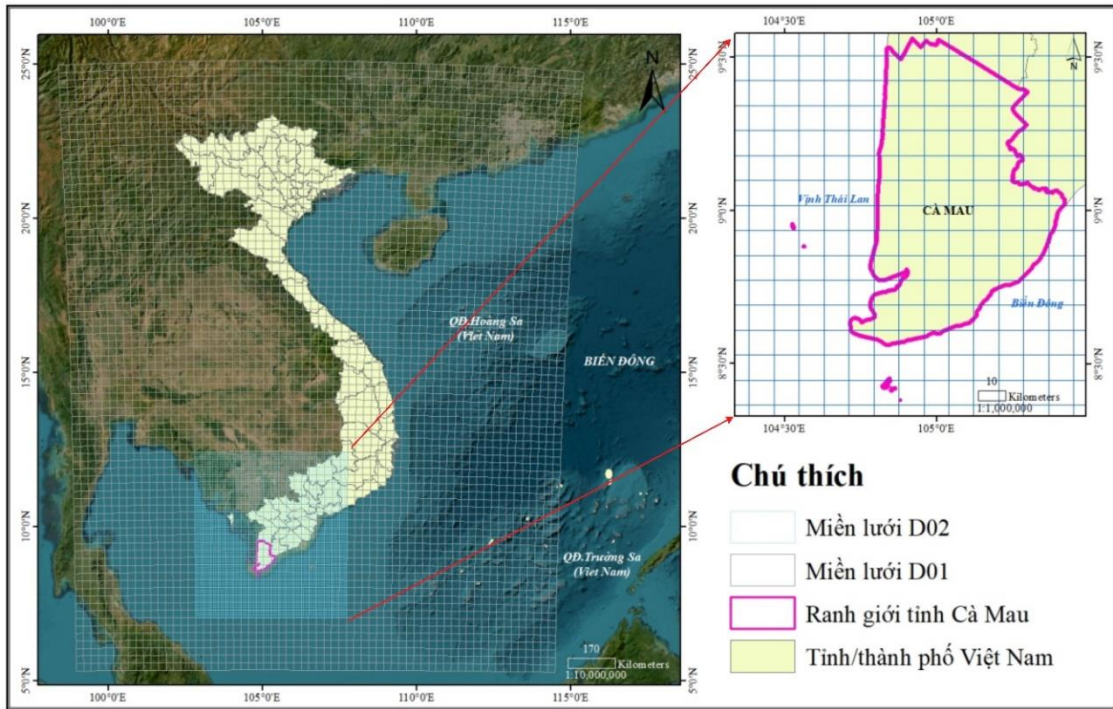
Thông số	Miền tính D01	Miền tính D02
Phạm vi miền tính	Toàn bộ Việt Nam	Các tỉnh phía Nam Việt Nam (bao gồm tỉnh Cà Mau)
Kích thước không gian X [km] × Y [km]	1593 x 2133	432 x 486
Số nút lưới tính N _x × N _y	60 × 80	49 × 55
Tổng số ô lưới tính	4661	2592
Kích thước ô lưới (km)	27	9
Tọa độ tâm miền	(15,03°; 105,394°)	(9,733119°; 105,297835°)
Hệ tọa độ thiết lập	Asia Lambert Conformal Conic	Asia Lambert Conformal Conic
Loại miền tính thiết lập	The nested modelling domain	The nested modelling domain

Trong nghiên cứu này, WRF phiên bản 3.8 được sử dụng với bộ dữ liệu khí tượng toàn cầu GFS được cung cấp bởi NCEP [47]. Bước hiệu chỉnh WRF sử dụng chuỗi số liệu khí tượng thời gian như nhiệt độ, tốc độ gió và hướng gió đo được tại 9 vị trí nằm ở các huyện xã, bao gồm KT1 (huyện U Minh), KT2 (xã Hòa Thành), KT3 (huyện Cái Nước), KT4 (huyện Đầm Dơi), KT5 (xã Tân Thuận), KT6 (huyện Năm Căn), KT7 (xã Tam Giang), KT8 (huyện Ngọc Hiển) và KT9 (huyện Trần Văn Thời) trong khoảng thời gian các tháng 03 và 04 năm 2020 để kiểm định mô hình WRF. Trong nghiên cứu, việc sử dụng đầu ra WRF trung bình 1h để phục vụ chạy mô hình CMAQ tính toán lan truyền PM_{2.5}.

Phương pháp đồng nhất dữ liệu (fusion data) giữa các kết quả quan trắc và các kết quả mô phỏng [17] được áp dụng. Công thức (1) ước lượng hai tham số hồi quy α và β phản ánh phương trình tương quan được sử dụng để hiệu chỉnh kết quả mô phỏng từ CMAQ như sau:

$$CMAQ_{x,corrected} = \alpha CMAQ_x^\beta \tag{1}$$

Trong đó, $CMAQ_x$ là nồng độ $PM_{2.5}$ ban đầu từ kết quả của mô hình CMAQ tại vị trí trạm quan trắc x vào thời điểm cụ thể t (giờ); $CMAQ_{x, Corrected}$ là nồng độ $PM_{2.5}$ đã được hiệu chỉnh tại vị trí trạm quan trắc x vào thời điểm cụ thể t (giờ) và α, β là các hệ số của phương trình tương quan sử dụng để hiệu chỉnh.



Hình 2. Lưới D01 và D02 được sử dụng chạy WRF/CMAQ.

Từ đó, xác định được các phương trình tương quan để hiệu chỉnh nồng độ bụi $PM_{2.5}$ mô phỏng lần lượt có các tháng như sau: phương trình $y = 1,3002x^{0,9794}$ cho tháng 03/2020; phương trình $y = 1,2607x^{0,9905}$ cho tháng 04/2020.

2.3.2. Mô hình hồi quy tuyến tính đa biến (Multiple Linear Regression - MLR)

Mô hình MLR (Multiple Linear Regression – MLR) được xây dựng để làm rõ mối tương quan giữa nồng độ bụi $PM_{2.5}$ lần lượt đối với các biến khí tượng (i), các biến tải lượng phát thải tiền chất (ii) và cả O2 loại biến (iii). Tất cả các dữ liệu nồng độ $PM_{2.5}$ được thiết lập như (1) dưới dạng trung bình hàng giờ và hàng ngày của các tháng 03 và 04 tương ứng với dữ liệu của các biến khí tượng và phát thải tiền chất. Mô hình được thiết lập chi tiết có dạng như ở công thức bên dưới [16].

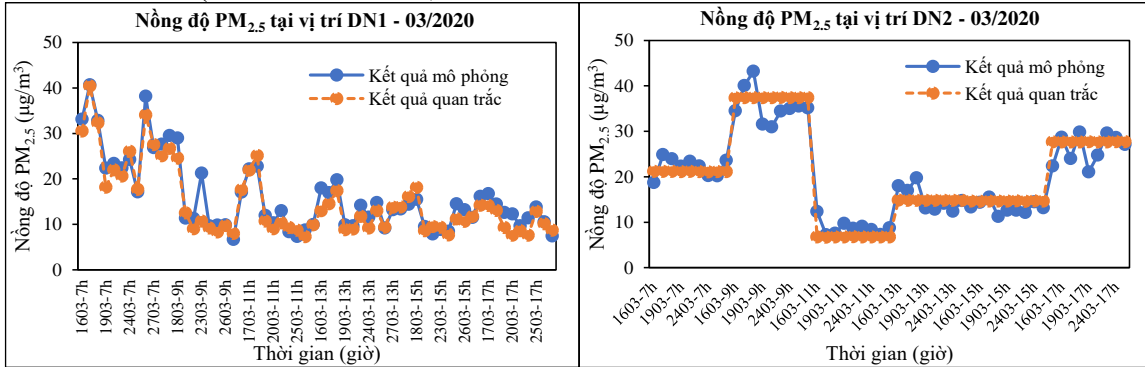
$$y_{i,j} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{k-1} x_{k-1} + \beta_k x_k + \varepsilon \quad (2)$$

Trong đó α là điểm cắt trên trục tung; $\beta, \beta_1, \beta_2, \dots, \beta_{k-1}$ và β_k là độ dốc hay còn gọi là các hệ số hồi quy (các tham số) với các trường hợp (i), (ii), (iii) giá trị k lần lượt là 3, 7 và 10; j là sự tương ứng cho từng thời điểm tháng thực hiện; i là sự tương ứng với các trường hợp (i), (ii), (iii) (i từ 1 đến 3); và ε là biến ngẫu nhiên (số hạng sai số). Đối với trường hợp (i) mối quan hệ giữa nồng độ $PM_{2.5}$ ($C_{PM_{2.5}}$) và 03 yếu tố khí tượng gồm vận tốc gió x_1 (W_s), hướng gió x_2 (W_d) và nhiệt độ x_3 (T) hay $C_{PM_{2.5}} = F(W_s, W_d \text{ và } T)$. Đối với trường hợp (ii) mối quan hệ giữa nồng độ $PM_{2.5}$ ($C_{PM_{2.5}}$) và 07 loại phát thải tiền chất SO_2 là E_{SO_2}, x_1 ; NO_x là E_{NO_x}, x_2 ; NH_3 là E_{NH_3}, x_3 ; $NMVOCs$ là E_{NMVOCs}, x_4 ; OC là E_{OC}, x_5 ; CO là E_{CO}, x_6 và BC là E_{BC}, x_7 hay $C_{PM_{2.5}} = F(E_{SO_2}, E_{NO_x}, E_{NH_3}, E_{NMVOCs}, E_{OC}, \text{ và } E_{CO} \text{ và } E_{BC})$ cho tỉnh Cà Mau. Đối với trường hợp (iii) mối quan hệ $C_{PM_{2.5}} = F(W_s, W_d, T, E_{SO_2}, E_{NO_x}, E_{NH_3}, E_{CO}, E_{NMVOCs}, E_{OC}, \text{ và } E_{BC})$, kết quả các phương trình hồi quy được xây dựng thể hiện mối tương quan đa biến giữa 03 yếu tố khí tượng và 07 loại phát thải tiền chất đối với giá trị nồng độ bụi $PM_{2.5}$. Các mô hình MLR được xây dựng cho toàn bộ khu vực nghiên cứu.

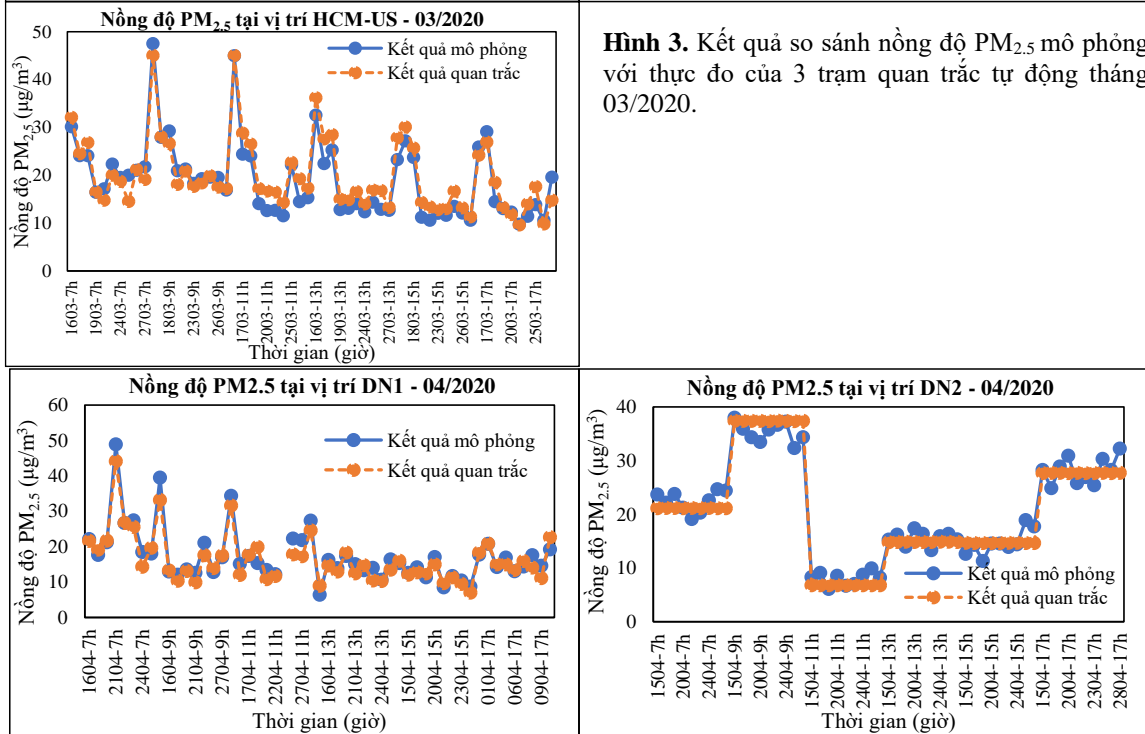
3. Kết quả và thảo luận

3.1. Đánh giá hiệu quả mô hình WRF/CMAQ

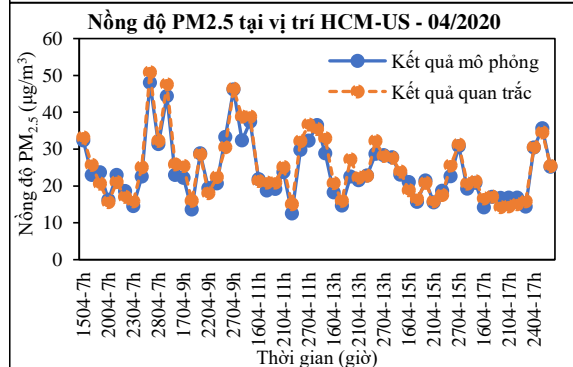
Kết quả mô phỏng nồng độ $PM_{2.5}$ cho tháng 03/2020 và tháng 04/2020 cho khu vực Cà Mau được xuất ra tại 3 điểm quan trắc tự động DN1, DN2 và HCM-US và so sánh với các giá trị quan trắc. Kết quả kiểm định dựa trên các chỉ số thống kê để đánh giá khả năng mô phỏng của mô hình WRF/CMAQ và mối tương quan giữa kết quả mô phỏng, đo đạc cũng được thể hiện ở (Hình 3 và Hình 4).



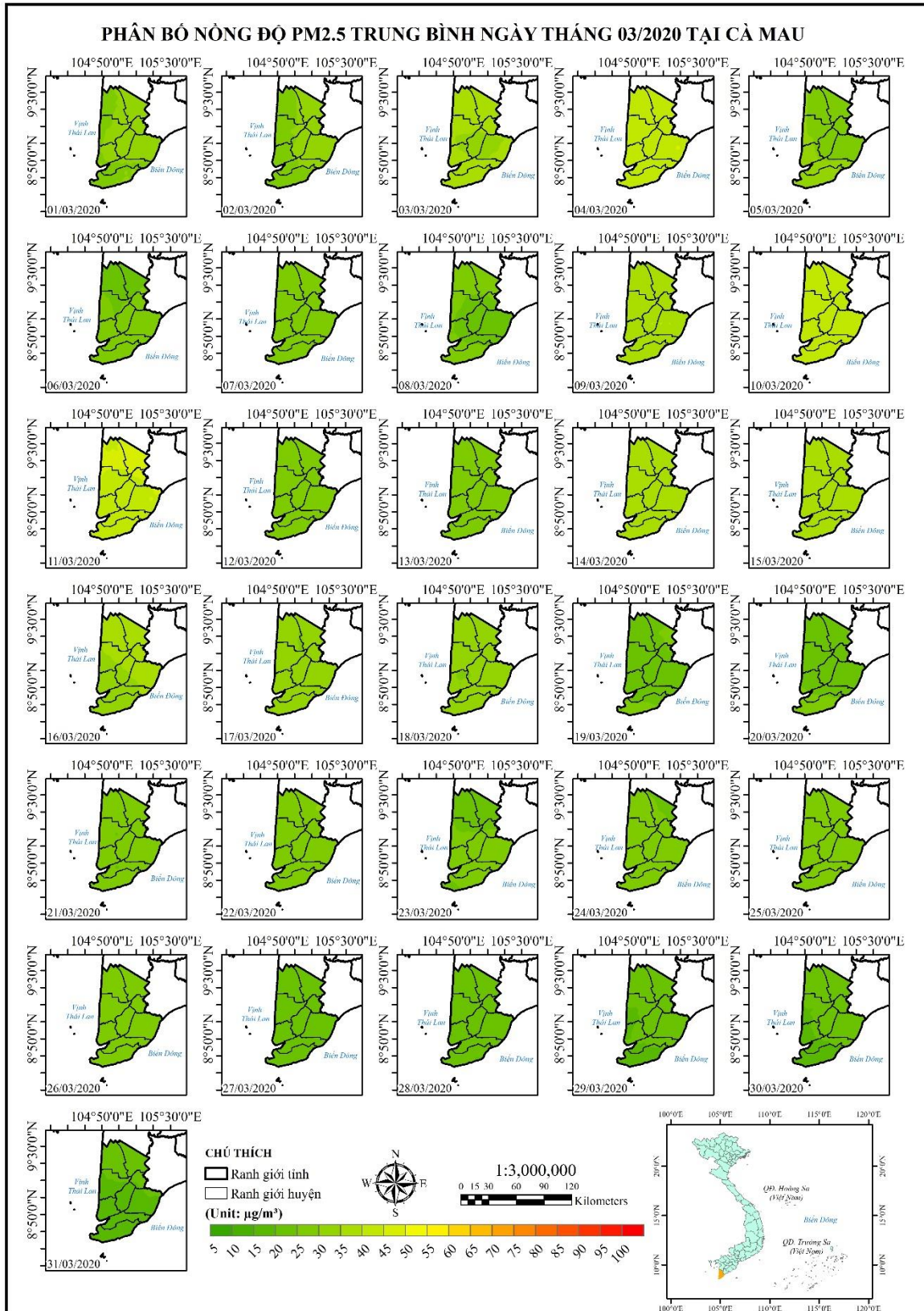
Hình 3. Kết quả so sánh nồng độ $PM_{2.5}$ mô phỏng với thực đo của 3 trạm quan trắc tự động tháng 03/2020.



Hình 4. Kết quả so sánh nồng độ $PM_{2.5}$ mô phỏng với thực đo của 3 trạm quan trắc tự động tháng 04/2020.



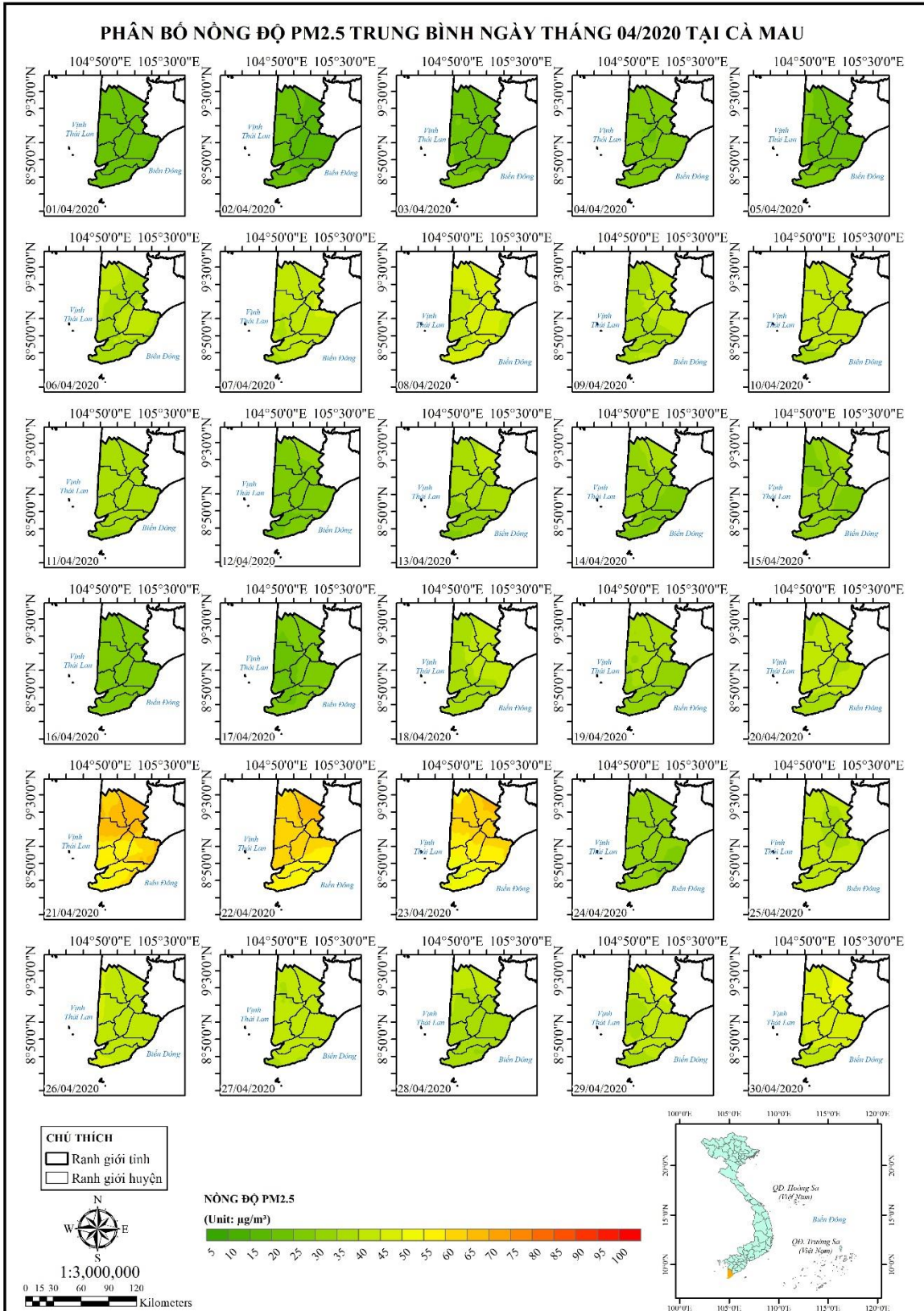
Vào tháng 03/2020, giá trị chỉ số Nash - Sutcliffe (NSE) tại các trạm đều đạt mức tốt với sự tương quan cao. Tại trạm DN1 giá trị $NSE_{DN1} = 0,905$ ($NSE > 0,9$); trạm DN2 với $NSE_{DN2} = 0,918$ ($NSE > 0,9$); trạm HCM-US giá trị $NSE_{HCM-US} = 0,880$ ($NSE > 0,8$). Vào tháng 03/2020, giá trị chỉ số Nash - Sutcliffe (NSE) tại các trạm đều hầu như đạt mức tốt với sự tương quan cao. Tại trạm DN1 giá trị $NSE_{DN1} = 0,574$ ($NSE > 0,5$); trạm DN2 giá trị $NSE_{DN2} = 0,952$ ($NSE > 0,9$). trạm HCM-US giá trị $NSE_{HCM-US} = 0,936$ ($NSE > 0,9$).



Hình 5. Phân bố không gian nồng độ PM_{2.5} trung bình ngày tại Cà Mau tháng 03/2020.

3.2. Sự phân bố không gian nồng độ PM_{2.5}

Nồng độ bụi PM_{2.5} trung bình ngày tháng 03/2020 có sự phân bố về mặt không gian thay đổi nhẹ theo giai đoạn đầu tháng 01/03/2020 - 08/03/2020, giữa tháng 09/03/2020 - 20/03/2020 và các ngày cuối tháng từ 21/03/2020 - 31/03/2020. Vào giai đoạn đầu tháng, nồng độ PM_{2.5} phân bố đều các huyện trên địa bàn tỉnh và không vượt quá giới hạn cho phép của QCVN 05:13/BTNMT (mức quy định 50 µg/m³). Đến giai đoạn giữa tháng, nồng độ có



Hình 6. Phân bố không gian nồng độ PM_{2.5} trung bình ngày tại Cà Mau tháng 04/2020.

xu hướng gia tăng. Vào các ngày 09, 10 và 11 nồng độ PM_{2.5} cao hơn các ngày còn lại, ngày có nồng độ cao nhất là ngày 11/03/2020. Cuối tháng nồng độ có xu hướng giảm dần từ ngày 21/03/2020 đến ngày 31/03/2020. Với dao động từ 7,82-51,72 µg/m³, nồng độ PM_{2.5} trung bình ngày cao nhất là ngày 11/03/2020 với mức dao động từ 23,82-51,72 µg/m³, và thấp nhất vào ngày 24/03/2020 với mức dao động từ 7,82-25,58 µg/m³. Tại thành phố Cà Mau nồng độ PM_{2.5} thường cao hơn các huyện còn lại. Về phân bố nồng độ PM_{2.5} trung bình hàng giờ trong ngày tháng 03/2020 (Hình 5), nồng độ trung bình hàng giờ ở các ngày có xu hướng dao động khá tương đồng, với nồng độ cao thường xuất hiện từ 9 đến 15 giờ. Nồng độ giảm dần vào buổi chiều tối và sáng sớm. Nhìn chung nồng độ PM_{2.5} trên toàn địa bàn tỉnh Cà Mau vào tháng 3 đều không vượt ngưỡng cho phép của QCVN 05:13/BTNMT.

Nồng độ PM_{2.5} trung bình ngày của tháng 04/2020 có sự phân bố không gian thay đổi nhẹ theo giai đoạn đầu giữa tháng 01/04/2020 - 20/04/2020 và các ngày cuối tháng từ 21/04/2020 - 30/04/2020. Đầu tháng, nồng độ PM_{2.5} phân bố đều các huyện trên địa bàn tỉnh và không vượt quá giới hạn cho phép của QCVN 05:13/BTNMT (mức quy định 50 µg/m³). Cuối tháng, nồng độ có xu hướng gia tăng. Vào các ngày 21, 22 và 23 nồng độ PM_{2.5} cao vượt mức cho phép tại hầu hết các huyện và thành phố như thành phố Cà Mau, huyện Ngọc Hiển, huyện Cái Nước, huyện Đầm Dơi, huyện Phú Tân, huyện Thới Bình, huyện Trần Văn Thời và huyện U Minh. Đặc biệt ngày 22/04 nồng độ PM_{2.5} vượt giới hạn trên toàn địa bàn tỉnh Cà Mau. Với dao động từ 7,05-114,42 µg/m³, 3/30 ngày trong tháng đã vượt quá giới hạn cho phép của QCVN 05:13/BTNMT. Trong đó, nồng độ PM_{2.5} trung bình ngày cao nhất là ngày 21/04/2020 với mức dao động từ 14,94-91,05 µg/m³, và thấp nhất vào ngày 02/04/2020 với mức dao động từ 7,05-18,83 µg/m³. Về phân bố nồng độ PM_{2.5} trung bình hàng giờ trong ngày tháng 04/2020 (Hình 6), nồng độ trung bình hàng giờ ở các ngày có xu hướng dao động khá tương đồng, với nồng độ cao thường xuất hiện từ 6 giờ đêm đến 9 giờ sáng và từ 18 đến 23 giờ.

3.3. Sự phụ thuộc nồng độ PM_{2.5} vào các yếu tố khí tượng

Mối quan hệ giữa nồng độ PM_{2.5} với 03 biến khí tượng (x_1 - vận tốc gió W_s , x_2 - hướng gió W_d và x_3 - nhiệt độ T) xem xét (*meteorological variables*) được xác định bằng các hệ số β_k trong các phương trình hồi quy tuyến tính đa biến (*Multiple Linear Regression*) tương ứng gắn với các biến khí tượng x_1, x_2, x_3 .

Bảng 4. Phương trình tương quan giữa giá trị nồng độ PM_{2.5} với vận tốc gió, hướng gió và nhiệt độ, hay $C_{PM_{2.5}} = F(W_s, W_d \text{ và } T)$ tại Cà Mau tại thời điểm tháng 03 và tháng 04.

Ký hiệu	Phương trình tương quan xây dựng	
KT1	$-115,8590+0,8609x_1-0,0450x_2-0,0265x_3$	$R^2 = 0,0472$
KT2	$46,1239+1,0330x_1-0,0177x_2-0,0599x_3$	$R^2 = 0,0096$
KT3	$19,9078+0,7899x_1-0,0181x_2+0,0272x_3$	$R^2 = 0,0090$
KT4	$25,1164+0,5927x_1-0,0200x_2+0,0128x_3$	$R^2 = 0,0065$
KT5	$37,6099+0,8148x_1-0,0299x_2-0,0264x_3$	$R^2 = 0,0083$
KT6	$-34,9415+0,0783x_1-0,0269x_2+0,2171x_3$	$R^2 = 0,0093$
KT7	$23,4061+0,9358x_1-0,0305x_2+0,0159x_3$	$R^2 = 0,0107$
KT8	$7,9582+0,0555x_1-0,0319x_2+0,0770x_3$	$R^2 = 0,0054$
KT9	$16,6495+1,6425x_1-0,0406x_2+0,0419x_3$	$R^2 = 0,0322$

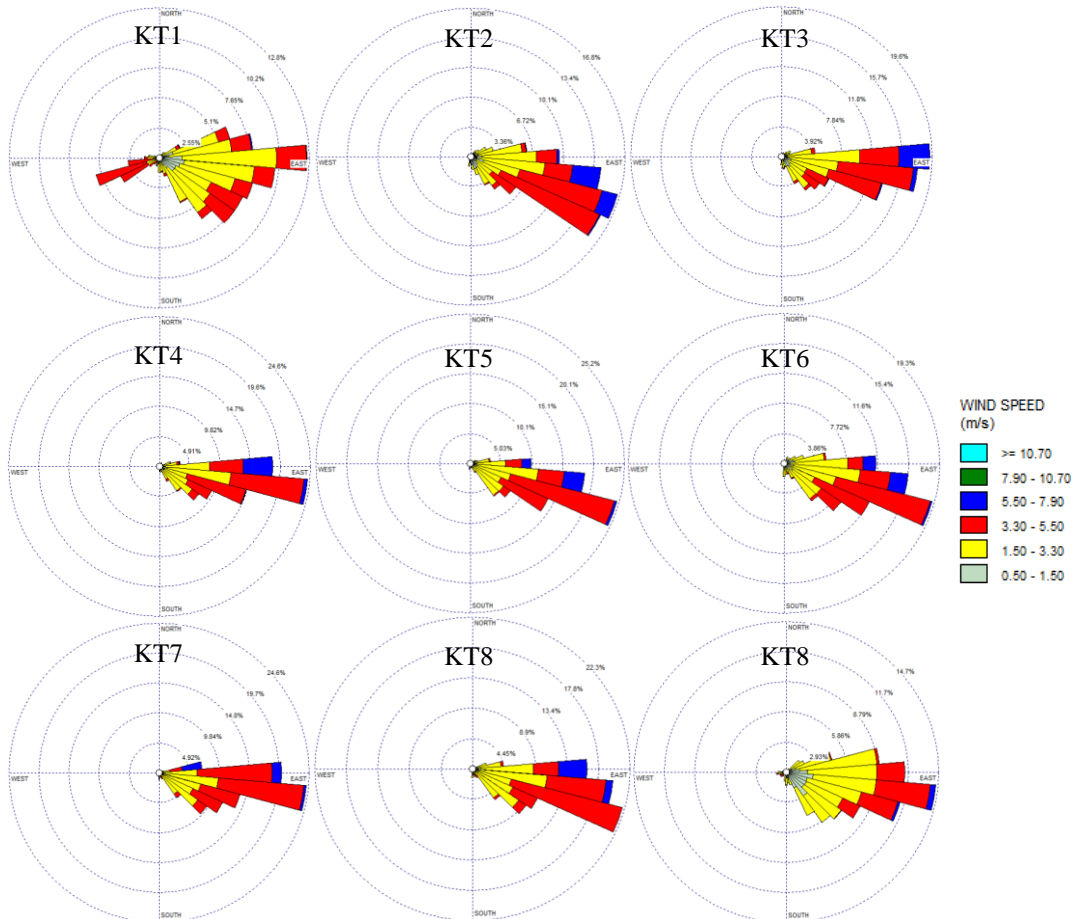
Từ các hàm số tương quan có thể nhận thấy tại toàn bộ vị trí từ KT1 đến KT9 giá trị nồng độ bụi PM_{2.5} có một mối tương quan thuận với xu hướng tăng lần lượt 0,8609 µg/m³, 1,0330 µg/m³, 0,7899 µg/m³, 0,5927 µg/m³, 0,8148 µg/m³, 0,0783 µg/m³, 0,9358 µg/m³, 0,0555 µg/m³, 1,6425 µg/m³ đối với mỗi 1 đơn vị (1 m/s) vận tốc gió W_s tăng lên (khi hướng gió W_d và nhiệt độ T không thay đổi). Ngược lại yếu tố hướng gió có sự tác động nghịch khi đối với mỗi 1 đơn vị (1°K) hướng gió W_d thay đổi theo chiều hướng tăng thì nồng độ bụi

PM_{2.5} lại có xu hướng giảm lần lượt là 0,0450 μg/m³, 0,0177 μg/m³, 0,0181 μg/m³, 0,0200 μg/m³, 0,0299 μg/m³, 0,0269 μg/m³, 0,0305 μg/m³, 0,0319 μg/m³, 0,0406 μg/m³ (khi vận tốc gió W_s và hướng gió W_d không thay đổi). Tại các vị trí KT3, KT4, KT6, KT7, KT8 và KT9 giá trị nồng độ bụi PM_{2.5} có mối tương quan thuận khi 1 đơn vị (1°K) nhiệt độ T tăng lên lần lượt là 0,0272 μg/m³, 0,0128 μg/m³, 0,2171 μg/m³, 0,0159 μg/m³, 0,0770 μg/m³, 0,0419 μg/m³. Ngược lại giá trị nồng độ bụi PM_{2.5} có mối tương quan nghịch khi 1 đơn vị (1°K) nhiệt độ T tăng lên tại các vị trí KT1, KT2 và KT5 lần lượt là 0,0265 μg/m³, 0,0599 μg/m³ và 0,0264 μg/m³. Như vậy, trong cả 03 yếu tố khí tượng phân tích thì vận tốc gió có đóng góp cho việc gia tăng mức nồng độ PM_{2.5} tại tỉnh Cà Mau giai đoạn tháng 03/2020 và 04/2020 (tỷ lệ thuận), hướng gió tỷ lệ nghịch đến việc gia tăng nồng độ PM_{2.5}, trong khi đó nhiệt độ có mức độ tác động phức tạp hơn tùy thuộc vào từng khu vực khác nhau.

Bảng 5. Mô tả khí tượng khu vực nghiên cứu tại tháng 03 và 04/2020.

STT	Ký hiệu	Vận tốc gió trung bình (m/s)	Hướng gió	Loại gió chủ yếu
1	KT1	2,57	Đông - Đông Nam	Gió yếu
2	KT2	2,71	Đông - Đông Nam	Gió yếu
3	KT3	2,64	Đông - Đông Nam	Gió yếu
4	KT4	3,01	Đông - Đông Nam	Gió yếu
5	KT5	3,05	Đông - Đông Nam	Gió yếu
6	KT6	2,73	Đông - Đông Nam	Gió yếu
7	KT7	3,31	Đông - Đông Nam	Gió yếu
8	KT8	3,17	Đông - Đông Nam	Gió yếu
9	KT9	2,07	Đông - Đông Nam	Gió yếu

Giai đoạn tháng 03, 04/2020, hướng gió chủ yếu tại khu vực nghiên cứu là Đông -Đông Nam, vận tốc gió tại các vị trí khá thấp, vận tốc gió trung bình 2,07-3,31 m/s, chủ yếu là gió yếu (39-54,4%) và gió nhẹ (26-44,3%) (Hình 7). Tại vị trí KT1, gió yếu chiếm 54,4% và gió



Hình 7. Bản đồ hoa gió tại tỉnh Cà Mau thời điểm tháng 03 và 04/2020.

nhẹ chiếm 26%; Tại vị trí KT2 gió yếu chiếm 39% và gió nhẹ chiếm 30,5%; Tại vị trí KT3 gió yếu chiếm 40,4% và gió nhẹ chiếm 28,6%; Tại vị trí KT4 gió yếu chiếm 45,1% và gió nhẹ chiếm 33,7%; Tại vị trí KT5 gió yếu chiếm 47,1% và gió nhẹ chiếm 34,9%; Tại vị trí KT6 gió yếu chiếm 45,3% và gió nhẹ chiếm 29,3%; Tại vị trí KT7 gió yếu chiếm 45,1% và gió nhẹ chiếm 44,3%; Tại vị trí KT8 gió yếu chiếm 51% và gió nhẹ chiếm 34,5%; Tại vị trí KT9 gió yếu chiếm 51% và gió rất nhẹ chiếm 29,3%.

3.4. Sự phụ thuộc nồng độ PM_{2.5} vào phát thải

Trong mục này thực hiện đánh giá mức độ đóng góp phát thải theo từng khu vực và theo từng loại tiền chất phát thải. Tỷ lệ đóng góp gây ô nhiễm đối với từng khu vực và từng loại tiền chất từ các nhóm ngành/hoạt động nhân tạo (sectors) cũng được phân tích. Kết quả phân tích, đánh giá mức đóng góp phát thải theo từng tiền chất trên toàn địa bàn tỉnh Cà Mau và tỉ lệ đóng góp của từng nhóm ngành/hoạt động theo từng tiền chất.

Đối với mối quan hệ phụ thuộc nồng độ PM_{2.5} với các phát thải SO₂, NO_x, NH₃, NMVOCs, OC, CO, và BC được thực hiện và được biểu diễn dưới dạng: $C_{PM2.5} = F(E_{SO2}, E_{NOx}, E_{NH3}, E_{NMVOCs}, E_{OC}, E_{CO} \text{ và } E_{BC})$. Kết quả các phương trình hồi quy được xây dựng thể hiện mối tương quan đa biến giữa 07 loại phát thải tiền chất ($x_1, x_2, x_3, x_4, x_5, x_6$ và x_7) đối với giá trị nồng độ bụi PM_{2.5} tại tỉnh Cà Mau ($y_{2,j}$) được thể hiện trong Bảng 4.

Bảng 6. Phương trình tương quan đa biến giữa giá trị nồng độ PM_{2.5} với 07 loại phát thải tiền chất, hay $C_{PM2.5} = F(E_{SO2}, E_{NOx}, E_{NH3}, E_{NMVOCs}, E_{OC}, E_{CO} \text{ và } E_{BC})$.

Thời gian	Phương trình tương quan	
03/2020	$y_{2,1} = 505,9025 + 3,4778x_2 - 27,3894x_3 + 4,6143x_4 + 22,7662x_5 - 0,0353x_6$	$R_1^2 = 0,1681$
04/2020	$y_{2,2} = -422174 + 4,3551x_1 - 99857,237x_2 + 2,8993x_4 + 0,6892x_5 - 0,0005x_6$	$R_2^2 = 0,1716$

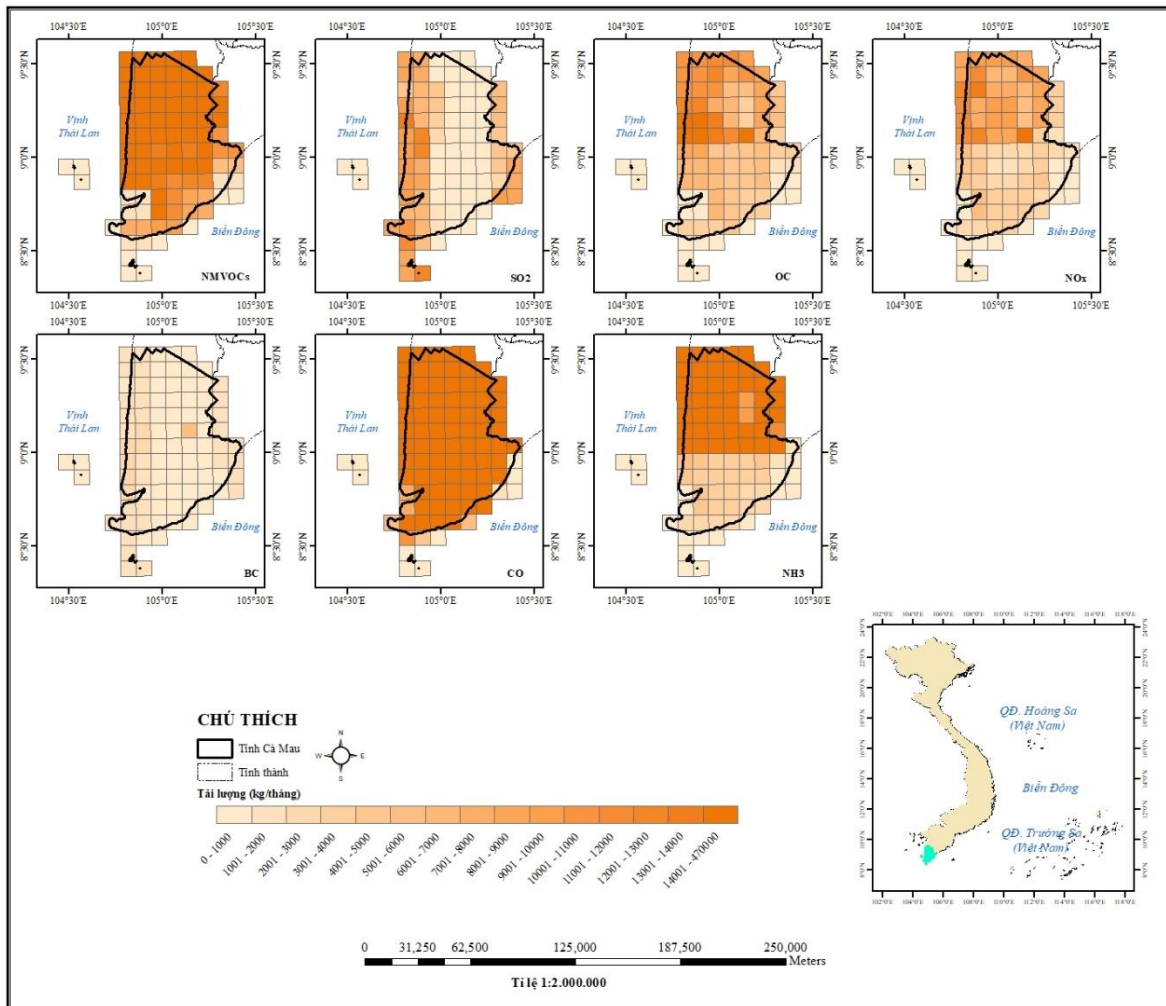
Hệ số tương quan tại các tháng 03/2020 và 04/2020 lần lượt có giá trị là $R_1 = 0,4101$, $R_2 = 0,4143$. Sự tương quan đa biến của 07 phát thải tiền chất đặc trưng như trên đều đã giải thích được những ảnh hưởng trong sự thay đổi ở mức giá trị nồng độ bụi PM_{2.5} tại toàn địa bàn tỉnh Cà Mau. Giá trị nồng độ bụi PM_{2.5} tại tháng 03/2020 có xu hướng tăng đối với mỗi 1 đơn vị (1 kg/ngày) phát thải của chất NO_x hoặc NMVOCs hoặc OC tăng lên (khi 06 tiền chất còn lại không thay đổi), cụ thể đối với E_{NOx} nồng độ PM_{2.5} tăng 3,4778 (µg/m³); đối với E_{NMVOCs} nồng độ PM_{2.5} tăng 4,6143 (µg/m³); đối với E_{OC} nồng độ PM_{2.5} tăng 22,7662 (µg/m³). Ngược lại, giá trị nồng độ bụi PM_{2.5} có xu hướng giảm đối với mỗi 1 đơn vị (1 kg/ngày) phát thải của chất NH₃ hoặc CO tăng lên (khi 06 tiền chất còn lại không thay đổi), cụ thể đối với E_{NH3} nồng độ PM_{2.5} giảm 27,3894 (µg/m³); đối với E_{CO} nồng độ PM_{2.5} giảm 0,0353 (µg/m³). Từ đó, có thể thấy rằng trong cả 07 phát thải tiền chất phân tích thì NO_x, NMVOCs và OC có đóng góp cho việc gia tăng mức nồng độ PM_{2.5} tại Cà Mau vào tháng 03/2020.

Giá trị nồng độ bụi PM_{2.5} tại tháng 04/2020 có xu hướng tăng đối với mỗi 1 đơn vị (1 kg/ngày) phát thải của chất SO₂ hoặc NMVOCs hoặc OC tăng lên (khi 06 tiền chất còn lại không thay đổi), cụ thể đối với E_{SO2} nồng độ PM_{2.5} tăng 4,3551 (µg/m³); đối với E_{NMVOCs} nồng độ PM_{2.5} tăng 2,8993 (µg/m³); đối với E_{OC} nồng độ PM_{2.5} tăng 0,6892 (µg/m³). Từ đó, có thể thấy rằng trong cả 07 phát thải tiền chất phân tích thì SO₂, NMVOCs và OC có đóng góp cho việc gia tăng mức nồng độ PM_{2.5} tại Cà Mau vào tháng 04/2020.

Trong đó, vào tháng 03/2020 mức độ đóng góp của các tiền chất đến nhóm ngành/hoạt động chính sau: (1) Sản xuất điện với phát thải NO_x nhiều nhất (chiếm 52%), kế đến là CO (46%); (2) Đốt chất thải nông nghiệp phát thải CO nhiều nhất (chiếm 87%), kế đến là NMVOCs (5%); (3) Quy trình công nghiệp phát thải CO nhiều nhất (chiếm 53%), kế đến là NMVOCs (31%); (4) Đất nông nghiệp phát thải NH₃ nhiều nhất (chiếm 84%), kế đến là CO (14%); (5) Vận tải đường bộ phát thải CO nhiều nhất (chiếm 88%), kế đến là NMVOCs (8%); (6) Chất thải rắn và nước thải phát thải NMVOCs nhiều nhất (chiếm 42%), kế đến là NH₃ (37%); (7) Khu dân cư và các khu vực khác phát thải CO nhiều nhất (chiếm 76%), kế đến là

NMVOCs (11%); (8) Tàu phát thải SO₂ nhiều nhất (chiếm 90%), kể đến là NO_x (5%); (9) chăn nuôi từ nông nghiệp phát thải NH₃ nhiều nhất (chiếm 98%); (10) sử dụng nhiên liệu hóa thạch phát thải NMVOCs nhiều nhất (chiếm 61%).

Giai đoạn tháng 04/2020 mức độ đóng góp của các tiền chất đến nhóm ngành/hoạt động chính như sau: (1) Sản xuất điện với phát thải NO_x nhiều nhất (chiếm 51%), kể đến là CO (44%) và SO₂ (3%); (2) Đốt chất thải nông nghiệp phát thải CO nhiều nhất (chiếm 88%), kể đến là NMVOCs (5%); (3) Quy trình công nghiệp phát thải CO nhiều nhất (chiếm 45%), kể đến là NMVOCs (26%) và SO₂ (14%); (4) Đất nông nghiệp phát thải NH₃ nhiều nhất (chiếm 74%), kể đến là CO (16%) và NO_x (10%); (5) Vận tải đường bộ phát thải CO nhiều nhất (chiếm 86%), kể đến là NMVOCs (9%) và NO_x (4%); (6) Chất thải rắn và nước thải phát thải NMVOCs nhiều nhất (chiếm 49%), kể đến là NH₃ (42%); (7) Khu dân cư và các khu vực khác phát thải CO nhiều nhất (chiếm 78%), kể đến là NMVOCs (11%); (8) Tàu phát thải NO_x nhiều nhất (chiếm 48%), kể đến là SO₂ (43%); (9) chăn nuôi từ nông nghiệp phát thải NH₃ nhiều nhất (chiếm 100%); (10) sử dụng nhiên liệu hóa thạch phát thải NMVOCs nhiều nhất (chiếm 72%).



Hình 8. Phân bố không gian tải lượng các phát thải tiền chất tại Cà Mau thời điểm tháng 03/2020.

4. Kết luận

Nghiên cứu đã đạt được các kết quả chính sau đây:

Thứ nhất, đã mô phỏng được nồng độ PM_{2.5} cho hai tháng mùa khô là tháng 03 và 04 năm 2020 tại tỉnh Cà Mau. Kết quả mô phỏng cho thấy nồng độ PM_{2.5} trung bình 24 giờ của tháng 03/2020 chưa vượt ngưỡng giới hạn cho phép của QCVN 05:13/BTNMT, giá trị nồng độ dao động trong khoảng 7,82-51,72 μg/m³. Giá trị nồng độ PM_{2.5} trung bình 24 giờ giai

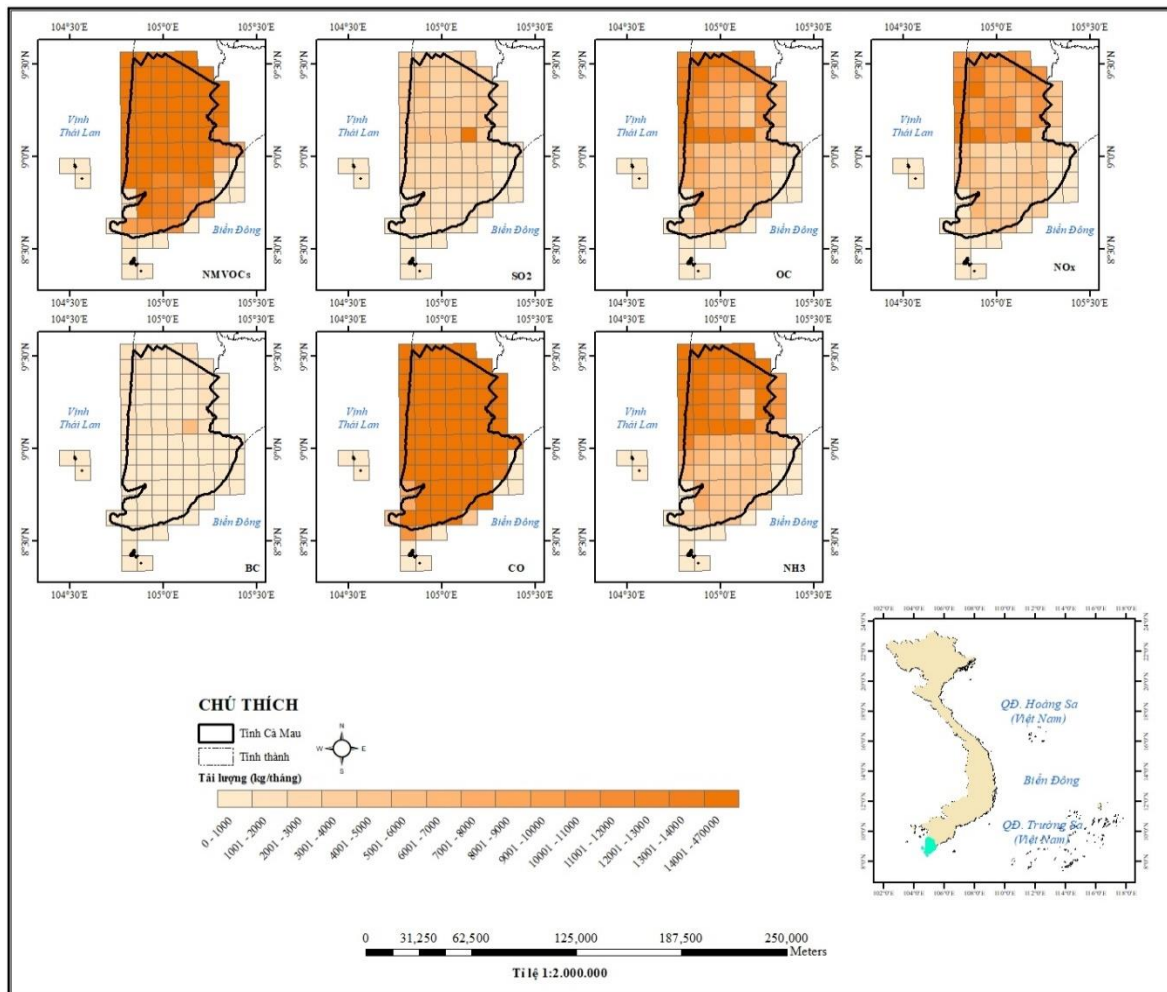
đoạn tháng 04/2020 là 7,05-114,42 $\mu\text{g}/\text{m}^3$. Vào tháng 04 đã có 3/30 ngày vượt ngưỡng giá trị cho phép từ 1,21-1,27 lần tùy vào thời điểm và vị trí. Có thể thấy được nồng độ $\text{PM}_{2.5}$ của tháng 04 cao hơn tháng 03 và nồng độ $\text{PM}_{2.5}$ khu vực thành phố Cà Mau cao hơn các vị trí còn lại.

Thứ hai, đã mô phỏng được các yếu tố khí tượng cho cả hai tháng 03 và 04 năm 2020. Vận tốc gió tại khu vực nghiên cứu khá thấp, chủ yếu là gió yếu và gió nhẹ, hướng gió chủ yếu là Đông - Đông Nam. Vận tốc gió trung bình của khu vực trong hai tháng là 2,07-3,31 m/s, gió yếu dao động từ 39-54,4% và gió nhẹ dao động từ 26-44,3%.

Thứ ba, đã đánh giá đóng góp của các phát thải tiền chất vào ô nhiễm $\text{PM}_{2.5}$. Dựa trên kết quả phân tích trong nghiên cứu, 07 tiền chất SO_2 , NO_x , NH_3 , CO , BC , OC , NMVOCs là các dạng tiền chất có đóng góp chính đến ô nhiễm bụi $\text{PM}_{2.5}$ tại tỉnh Cà Mau.

Thứ tư, đã tìm ra mối tương quan giữa giá trị $\text{PM}_{2.5}$ và các yếu tố khí tượng. Cụ thể trong ba yếu tố khí tượng phân tích thì vận tốc gió có đóng góp cho việc gia tăng mức nồng độ $\text{PM}_{2.5}$ tại tỉnh Cà Mau giai đoạn tháng 03/2020 và 04/2020 (tỷ lệ thuận), hướng gió tỷ lệ nghịch đến việc gia tăng nồng độ $\text{PM}_{2.5}$, trong khi đó nhiệt độ có mức độ tác động phức tạp hơn tùy thuộc vào từng khu vực khác nhau.

Trong nghiên cứu này vẫn tồn tại một số hạn chế có thể dẫn đến sự không chắc chắn trong kết quả đạt được và tạo ra sai số cho các phân tích, đánh giá như sự hạn chế về dữ liệu nồng độ $\text{PM}_{2.5}$ thực đo. Cụ thể, bộ số liệu thực đo được sử dụng để kiểm định mô hình là từ các trạm quan trắc tại Tp. HCM, Bình Dương và Đồng Nai. Dù có những hạn chế nhất định trong kết quả nghiên cứu được trình bày nhưng kết quả của đề tài đã cung cấp được những thông tin có giá trị liên quan đến đóng góp của phát thải tiền chất cũng như khí tượng cho sự hình thành $\text{PM}_{2.5}$, tạo cơ sở cho các nghiên cứu tiếp theo tại tỉnh Cà Mau.



Hình 9. Phân bố không gian tải lượng các phát thải tiền chất tại Cà Mau thời điểm tháng 04/2020.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu và đặt ra bài toán: B.T.L.; Viết bản thảo, chỉnh sửa bài báo: B.T.L.; Xử lý số liệu: N.T.X.; Viết báo cáo kỹ thuật; chỉnh sửa bài báo: N.T.X.; GIS, biện luận phần khí tượng, xây dựng phương trình tương quan: N.C.M.D.

Lời cảm ơn: Chúng tôi xin cảm ơn Trường Đại học Bách Khoa, ĐHQG-HCM đã hỗ trợ cho nghiên cứu này.

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. Jerrett, M. Atmospheric science: The death toll from air-pollution sources. *Nature* **2015**, 525, 330–331. Doi:10.1038/525330a.
2. GreenID. Report of Air Quality in 2017. Cau Giay, Ha Noi, 2017.
3. Landrigan, P.J. et al. The lancet commission on pollution and health. *Lancet* **2018**, 391, 462–512. Doi:10.1016/S0140-6736(17)32345-0.
4. Chen, L.; Zhu, J.; Liao, H.; Yang, Y.; Yue, X. Meteorological influences on PM_{2.5} and O₃ trends and associated health burden since China's clean air actions. *Sci. Total Environ.* **2020**, 744, 140837. Doi: 10.1016/j.scitotenv.2020.140837.
5. Hung, T.D. et al. Applied GIS technologies and satellite in monitoring content of PM_{2.5} dust variation in the North of Vietnam in 2000-2005-2010. National Scientific Conference on Meteorology, Hydrology, Environment and Climate Change. 2017, pp. 476–482.
6. Nguyen. T.T.N. et al. Current status of PM_{2.5} pollution and its mitigation in Vietnam. *Glob. Environ. Res.* **2018**, 22, 073–083.
7. MONRE. Vietnam National Environmental Report 2017. Ha Noi Capital. 2018.
8. Do, T.N.N. et al. Application of WRF-Chem to simulate air quality over Northern Vietnam. *Environ. Sci. Pollut. Res.* **2021**, 28, 12067–12081. Doi: 10.1007/s11356-020-08913-y.
9. Li, J. et al. Estimation of PM_{2.5} mortality burden in China with new exposure estimation and local concentration-response function. *Environ. Pollut.* **2018**, 243, 1710–1718. doi: 10.1016/j.envpol.2018.09.089.
10. Toledo, T. et al. WRF-SMOKE-CMAQ modeling system for air quality evaluation in São Paulo megacity with a 2008 experimental campaign data. *Environ. Sci. Pollut. Res.* **2018**, 25, 36555–36569.
11. Pope III, C.A. et al. Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. *JAMA.* **2002**, 287, 1132–1141. Doi: 10.1001/jama.287.9.1132.
12. Cohen, A.J. et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* **2017**, 389, 1907–1918. Doi:10.1016/S0140-6736(17)30505-6.
13. Zhao, P. et al. Long-term visibility trends and characteristics in the region of Beijing, Tianjin, and Hebei, China. *Atmos. Res.* **2011**, 101, 711–718. Doi: 10.1016/j.atmosres.2011.04.019.
14. Zhao, H. et al. Characteristics of visibility and particulate matter (PM) in an urban area of Northeast China. *Atmos. Pollut. Res.* **2013**, 4, 427–434. Doi: 10.5094/APR.2013.049.
15. Bond, T.C. et al. Bounding the role of black carbon in the climate system: A scientific assessment. *J. Geophys. Res. Atmos.* **2013**, 118, 5380–5552. Doi: 10.1002/jgrd.50171.
16. Tai, A.P.K. et al. Meteorological modes of variability for fine particulate matter (PM_{2.5}) air quality in the United States: Implications for PM_{2.5} sensitivity to climate change. *Atmos. Chem. Phys.* **2012**, 12, 3131–3145. Doi: 10.5194/acp-12-3131-2012.
17. Long, B.T.; Phong, N.H.; Duyen, N.C.M. Spatio-temporal modelling and related factors analysis of PM 2.5 pollution : A case study of Ho Chi Minh city, Vietnam. *Phys. Chem.*

- Earth*. **2023**, *131*, 103417. Doi: 10.1016/j.pce.2023.103417.
18. Long, B.T.; Nhi, N.H.T.; Phong, N.H. Chronic and acute health effects of PM_{2.5} exposure and the basis of pollution control targets. *Environ. Sci. Pollut. Res.* **2023**. Doi: 10.1007/s11356-023-27936-9.
 19. Long, B.T.; Phong, N.H.; Duyen, N.C.M. Linking air quality, health, and economic effect models for use in air pollution epidemiology studies with uncertain factors. *Atmos. Pollut. Res.* **2021**, *12*(7), 1011118. Doi:10.1016/j.apr.2021.101118.
 20. Chi, N.N.H.; Oanh, N.T.K. Photochemical smog modeling of PM_{2.5} for assessment of associated health impacts in crowded urban area of Southeast Asia. *Environ. Technol. Innov.* **2021**, *21*, 101241. Doi: 10.1016/j.eti.2020.101241.
 21. Huy, D.H. et al. Fine particulate matter (PM_{2.5}) in Ho Chi Minh City: Analysis of the status and the temporal variation based on the continuous data from 2013-2017. *Sci. Technol. Dev. J. Nat. Sci.* **2018**, *2*(5), 30–137.
 22. Squizzato, S. et al. Factors determining the formation of secondary inorganic aerosol: A case study in the Po Valley (Italy). *Atmos. Chem. Phys.* **2013**, *13*, 1927–1939. Doi: 10.5194/acp-13-1927-2013.
 23. Hallquist, M. et al. The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmos. Chem. Phys.* **2009**, *9*, 5155–5236. Doi: 10.5194/acp-9-5155-2009.
 24. Hien, T.T.; Chi, N.D.T.; Nguyen, N.T.; Vinh, L.X.; Takenaka, N.; Huy, D.H. Current status of fine particulate matter (PM_{2.5}) in vietnam's most populous city, Ho Chi Minh City. *Aerosol Air Qual. Res.* **2019**, *19*, 2239–2251. Doi:10.4209/aaqr.2018.12.0471.
 25. Zhang, B.N.; Oanh, N.T.K. Photochemical smog pollution in the Bangkok Metropolitan Region of Thailand in relation to O₃ precursor concentrations and meteorological conditions. *Atmos. Environ.* **2002**, *36*, 4211–4222.
 26. Zhang, L.; Liao, H.; and Li, J. Impacts of Asian summer monsoon on seasonal and interannual variations of aerosols over eastern China. *J. Geophys. Res. Atmos.* **2010**, *115*, 1–20. Doi:10.1029/2009JD012299.
 27. Han, X.; Zhang, M.; Gao, J.; Wang, S.; and Chai, F. Modeling analysis of the seasonal characteristics of haze formation in Beijing. *Atmos. Chem. Phys.* **2014**, *14*, 10231–10248. Doi:10.5194/acp-14-10231-2014.
 28. Huang, X.; Wang, Z. and Ding, A. Impact of Aerosol-PBL Interaction on Haze Pollution: Multiyear Observational Evidences in North China. *Geophys. Res. Lett.* **2018**, *45*, 8596–8603. Doi:10.1029/2018GL079239.
 29. Zhang, Y. et al. Long-term trends in the ambient PM_{2.5}- and O₃-related mortality burdens in the United States under emission reductions from 1990 to 2010. *Atmos. Chem. Phys.* **2018**, *18*, 15003–15016. Doi:10.5194/acp-18-15003-2018.
 30. Yang, Y. et al. Dust-wind interactions can intensify aerosol pollution over eastern China. *Nat. Commun.* **2017**, *8*, 1–8. Doi:10.1038/ncomms15333.
 31. Zhang, Z.; Wang, K. Stilling and recovery of the surface wind speed based on observation, reanalysis, and geostrophic wind theory over China from 1960 to 2017. *J. Clim.* **2020**, *33*, 3989–4008. Doi:10.1175/JCLI-D-19-0281.1.
 32. Gao, Y. et al. Modeling the feedback between aerosol and meteorological variables in the atmospheric boundary layer during a severe fog-haze event over the North China Plain. *Atmos. Chem. Phys.* **2015**, *15*, 4279–4295. Doi:10.5194/acp-15-4279-2015.
 33. Chen, Z. et al. The control of anthropogenic emissions contributed to 80% of the decrease in PM_{2.5} concentrations in Beijing from 2013 to 2017. *Atmos. Chem. Phys.* **2019**, *19*, 13519–13533. Doi:10.5194/acp-19-13519-2019.
 34. Racherla, P.N.; Adams, P.J. Sensitivity of global tropospheric ozone and fine particulate matter concentrations to climate change. *J. Geophys. Res. Atmos.* **2006**, *111*, 1–11. Doi:10.1029/2005JD006939.
 35. Wang, X. et al. PM_{2.5} pollution in China and how it has been exacerbated by terrain and meteorological conditions. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 105–120. Doi:10.1175/BAMS-D-16-0301.1.

36. Tao, H. et al. Impacts of improved modeling resolution on the simulation of meteorology, air quality, and human exposure to PM_{2.5}, O₃ in Beijing, China. *J. Clean. Prod.* **2020**, 243, 118574. Doi:10.1016/j.jclepro.2019.118574.
37. Henze, D.K.; Seinfeld, J.H.; Shindell, D.T. Inverse modeling and mapping US air quality influences of inorganic PM 2.5 precursor emissions using the adjoint of GEOS-Chem. *Atmos. Chem. Phys.* **2009**, 9, 5877–5903. Doi:10.5194/acp-9-5877-2009.
38. Wong, D.C. et al. WRF-CMAQ two-way coupled system with aerosol feedback: Software development and preliminary results. *Geosci. Model Dev.* **2012**, 5, 299–312. Doi:10.5194/gmd-5-299-2012.
39. Ca Mau CEM. Report on the current state of the environment in Ca Mau province in 2015-2020, Ca Mau, 2020.
40. Ca Mau CEM. Environmental planning of Ca Mau province until 2010 and orientation to 2020, Ca Mau, 2021.
41. The economy of Ca Mau province has achieved a good growth rate, the scale of the economy has continuously increased over 17 years of implementing Resolution No. 21-NQ/TW of the Politburo. Ca Mau Provincial Party Committee's electronic information portal, Ca Mau, 2021.
42. Weather underground. Data. <https://www.wunderground.com/> (01/08/2023)
43. Dong Nai CEM. Air quality monitoring report of Dong Nai province. Dong Nai, 2021.
44. Binh Duong CEM, Report on air quality monitoring results in Binh Duong province in 2020. Thu Dau Mot. 2021.
45. AirNow Department of State, World air quality report 2020. 2020.
46. United States Environmental Protection Agency. Community Multiscale Air Quality Modeling System (CMAQ). 2021.
47. NCEP. NCEP FNL Operational Model Global Tropospheric Analyses, Continuing from July 1999. 2018.

Simulating PM_{2.5} dust pollution and analyzing related factors – The case of Ca Mau province, Vietnam

Ngô Thị Xuan^{1,2}, Nguyễn Châu My Duyên^{1,2}, Bùi Tạ Long^{1,2*}

¹ Faculty of Environment and Natural Resources, Ho Chi Minh City University of Technology; ngothixuana3qt@gmail.com; nguyenduyen91@hcmut.edu.vn; longbt62@hcmut.edu.vn

² Vietnam National University Ho Chi Minh City; ngothixuana3qt@gmail.com; nguyenduyen91@hcmut.edu.vn; longbt62@hcmut.edu.vn

Abstract: PM_{2.5} air pollution is a factor that seriously impacts human health not only in large urban areas, but also affects rural provinces, in this case study, Ca Mau province. Currently, the Provincial People's Committee is in the process of developing an air environment quality management plan for the entire province. In which the task of assessing fine dust pollution is set. To solve this goal, the pair of WRF/CMAQ models is used on the one hand to form a pollution distribution map, and on the other hand to find the dependence between the level of PM_{2.5} pollution and emissions factors and meteorology. The anthropogenic and biological (natural) emissions inventory data set from the global emission inventory source ECCAD (Emissions of atmospheric compounds and compilation of ancillary data) is used. Simulation results performed for 2 typical months of the 2020 dry season show that the average 24-hour PM_{2.5} concentration in March 2020 has not exceeded the allowable limit of QCVN 05:13/BTNMT, the level concentration fluctuates between 7.82-51.72 µg/m³. The average 24-hour PM_{2.5} concentration value in April 2020 is 7.05-114.42 µg/m³. The results of analysis of concentration dependence on emissions and meteorology were also analyzed and were clarified.

Keywords: WRF/CMAQ models; PM_{2.5}; Meteorology; Emission; Ca Mau.

Bài báo khoa học

So sánh mức độ ảnh hưởng của một số vật liệu đối với quá trình cố định Asen trong đất bị ô nhiễm

Nguyễn Quốc Tuấn^{1*}, Nguyễn Hồng Quân¹, Nguyễn Thị Hải^{1,2}

¹ Khoa Địa chất, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, 334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam; nqt18@vnu.edu.vn; nguyenhongquan160198@gmail.com

² Phòng thí nghiệm Trọng điểm Địa môi trường và Ứng phó Biến đổi Khí hậu, trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, 334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam; nguyenhai128@vnu.edu.vn

*Tác giả liên hệ: nqt18@vnu.edu.vn; Tel.: +84-345311281

Ban Biên tập nhận bài: 12/9/2023; Ngày phản biện xong: 30/10/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Nghiên cứu so sánh sự ảnh hưởng của ba vật liệu chính là bentonite (B), than hoạt tính (AC), và vật liệu tổng hợp của chúng (BAC) với tỉ lệ áp dụng 2% đối với khả năng cố định Asen (As) trong đất bị ô nhiễm. Khả năng cố định As trong đất được đánh giá bằng: i) quy trình lọc kết tủa tổng hợp (SPLP), quy trình chiết độc tính (TCLP), và dung dịch 0,1M HCl để đánh giá khả năng rửa trôi; ii) tách chiết tuần tự (Sequential Extraction) để đánh giá các dạng liên kết của As trong đất. Kết quả chỉ ra sự có mặt của các vật liệu đã giúp tăng độ pH của đất và thay đổi tính linh động của As trong đất. Hiệu suất cố định As tăng tới 52% trong đất được xử lý bằng vật liệu BAC. Kết quả cũng chỉ ra As trong dạng tồn tại dễ trao đổi đã được chuyển sang dạng tồn tại bền hơn và ít linh động hơn, đặc biệt đối với đất được xử lý bằng BAC. Nghiên cứu chỉ ra rằng việc kết hợp các vật liệu đem lại hiệu quả cao trong xử lý đất bị ô nhiễm, do đó cần phải được tiếp tục nghiên cứu để đánh giá tính hiệu quả không những đối với As mà còn với các kim loại nặng khác, và xem xét tác động của vật liệu đối với sức khỏe của đất.

Từ khóa: Asen; Bentonite; Than hoạt tính; Cố định Asen trong đất.

1. Mở đầu

Asen (As) là một nguyên tố vô cơ phổ biến, dễ dàng được tìm thấy ở ngoài tự nhiên, cũng như trong sinh hoạt hàng ngày. Ngày nay, với việc phát triển của quá trình công nghiệp hóa, As xuất hiện nhiều trong sản phẩm công nghiệp, ảnh hưởng ô nhiễm đến môi trường sống, và trở thành mối quan tâm lớn ở nhiều khu vực trên thế giới và Việt Nam [1]. Ô nhiễm As trong đất có thể ảnh hưởng tới hệ sinh thái, nước ngầm, môi trường vi sinh vật, thực vật, và theo chuỗi thức ăn có tác hại trực tiếp lên con người. Theo cơ quan Bảo vệ Môi trường Mỹ (USEPA), As là một nguyên tố độc hại nguy hiểm và đã đặt ra nồng độ tối đa cho phép của As trong nước uống là 10 µg/L. Việc cơ thể con người tiếp xúc lâu dài hoặc liều lượng cao với As có thể gây ra hàng loạt bệnh liên quan đến hệ sinh sản và tuần hoàn cũng như nguy cơ ung thư ở nhiều cơ quan [2].

Trên thế giới, các báo cáo về mức độ ô nhiễm As cho thấy thách thức đang ngày càng lớn trong vấn đề xử lý ô nhiễm As. Theo báo cáo khảo sát về mức độ ô nhiễm đất ở Trung Quốc năm 2014, ô nhiễm As đã vượt quá giới hạn cho phép 2,7% [3, 4]. Ở châu Âu, theo khảo sát của dự án Baltic Soil Survey (BSS), hàm lượng As trong đất nông nghiệp ở khoảng 5,7 mg/kg; trong đó, nồng độ As ở phía Nam châu Âu cao gấp 3 lần so với phía Bắc [5]. Tại Hoa Kỳ, theo số liệu báo cáo có khoảng từ 100.000 - 1.000.000 ha đất nông nghiệp có hàm

lượng As trong đất từ 200 ppm trở lên và hàng chục triệu ha có dư lượng As trong đất từ 20 - 30 ppm [6]. Ở Việt Nam, theo nghiên cứu [7], tổng hàm lượng As của khoảng 80% đất nông nghiệp khu vực Đồng Bằng sông Hồng trong phạm vi khảo sát của nghiên cứu vượt quá giới hạn tối đa cho phép đối với đất nông nghiệp Việt Nam (12 mg/kg) gây nguy cơ an toàn lương thực.

Ngày nay, quá trình đô thị hóa, công nghiệp hóa diễn ra rất nhanh và mạnh, là nguyên nhân dẫn đến sự thu hẹp của đất nông nghiệp. Việc tận dụng các khu đất xung quanh các khu khai thác mỏ đã bỏ hoang làm đất nông nghiệp đang là một xu thế của một số quốc gia trên thế giới như Hàn Quốc. Tuy nhiên, đã có nhiều báo cáo liên quan đến ô nhiễm As và kim loại nặng trong đất canh tác xung quanh khu mỏ bỏ hoang, và vấn đề xử lý ô nhiễm As và kim loại nặng cần được quan tâm chú ý đối với các khu vực như này [1].

Chính vì vậy, việc xử lý đất canh tác tại các khu mỏ bỏ hoang bị ô nhiễm As đang ngày càng nhận được sự quan tâm chú ý của các nhà nghiên cứu. Có thể kể đến một số phương pháp để loại bỏ As trong đất như dùng thực vật hấp thụ, vi sinh vật, điện phân [8–10]. Tuy nhiên, các phương pháp này hoặc đòi hỏi thời gian xử lý lâu dài hoặc chi phí xử lý tốn kém. Phương pháp rửa đất có thể mang lại hiệu quả bằng cách sử dụng những dung môi như tartaric acid, citric acid, oxalic acid... nhưng việc xử lý dung dịch sau quá trình rửa đất cũng là một vấn đề hết sức khó khăn [11, 12].

Gần đây, phương pháp cố định sử dụng các vật liệu bổ sung vào đất đang ngày càng được tập trung nghiên cứu vì phương pháp này được đánh giá là một phương pháp đầy hứa hẹn và tiềm năng trong việc xử lý ô nhiễm As và kim loại nặng trong đất [13]. So sánh với các phương pháp xử lý ô nhiễm As trong đất khác như xử lý bằng các biện pháp sinh học (hấp thụ thực vật...), xử lý nhiệt, hay rửa đất, phương pháp cố định As bằng các vật liệu đem lại hiệu quả cao, ổn định, và tiết kiệm chi phí cũng như thời gian xử lý [14]. Tuy nhiên, việc áp dụng vật liệu nào trong việc cố định As cần xem xét kỹ lưỡng để đem lại hiệu quả cao nhất, cũng như dễ tiếp cận, không những có giá thành giá rẻ, mà còn ít tác động tiêu cực đến môi trường.

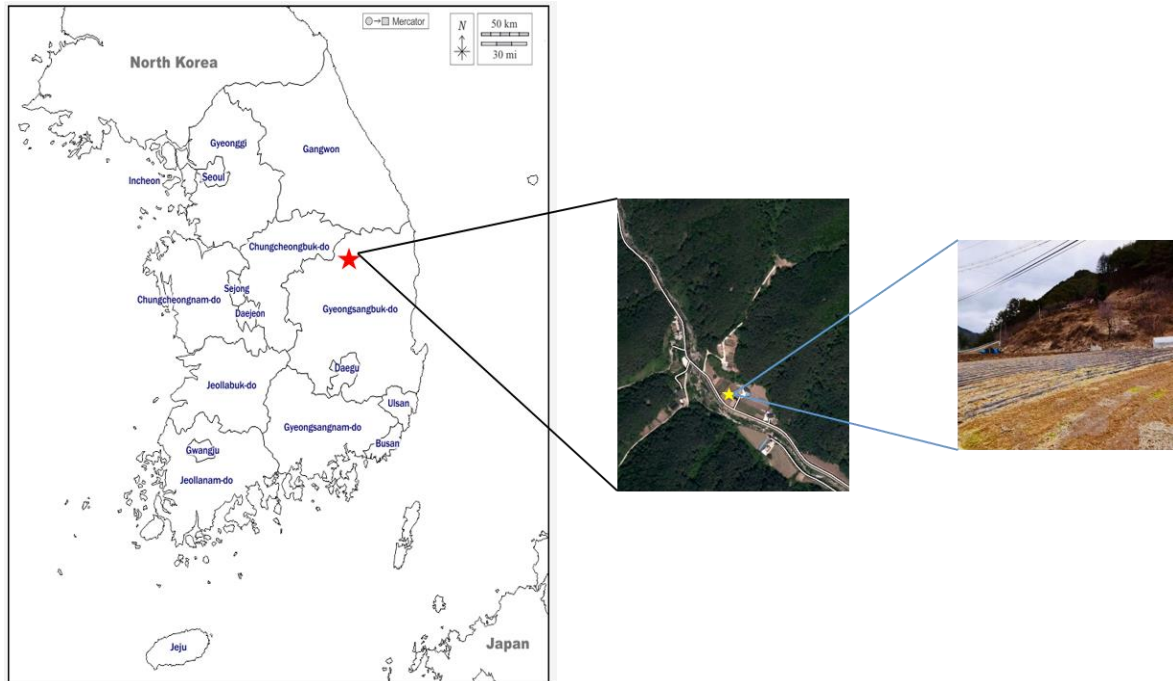
Các hợp chất khoáng sét là một trong những vật liệu hiện đang thu hút sự chú ý do giá thành rẻ, số lượng lớn và diện tích bề mặt, cũng như tính chất hấp phụ tốt [15]. Bentonite là một loại khoáng sét bao gồm các lớp tứ diện Si và bát diện Al liên kết với nhau của các nhóm hydroxyl. Bentonite mang điện tích âm và có thể được trung hòa bởi các cation trao đổi [16]. Do đó, bentonite được dùng nhiều trong xử lý ô nhiễm nước bởi các hợp chất hữu cơ như methylene blue [17]. Bên cạnh đó, than hoạt tính cũng là một vật liệu hấp phụ được dùng nhiều trong quá trình xử lý ô nhiễm vô cơ hay hữu cơ môi trường nước do diện tích bề mặt hấp phụ của chúng lớn, và tính chất thân thiện với môi trường [18]. Tuy nhiên, việc áp dụng bentonite và than hoạt tính trong việc xử lý ô nhiễm As trong đất vẫn chưa có. Đặc biệt, việc kết hợp hai vật liệu này nhằm tăng cường khả năng cố định As trong đất hiện vẫn chưa được nghiên cứu.

Chính vì vậy, nghiên cứu này được thực hiện nhằm đánh giá tính hiệu quả và khả năng của việc áp dụng các vật liệu như bentonite, than hoạt tính, và vật liệu tổng hợp giữa chúng đối với quá trình cố định As trong đất.

2. Phương pháp nghiên cứu

2.1. Chuẩn bị đất và vật liệu

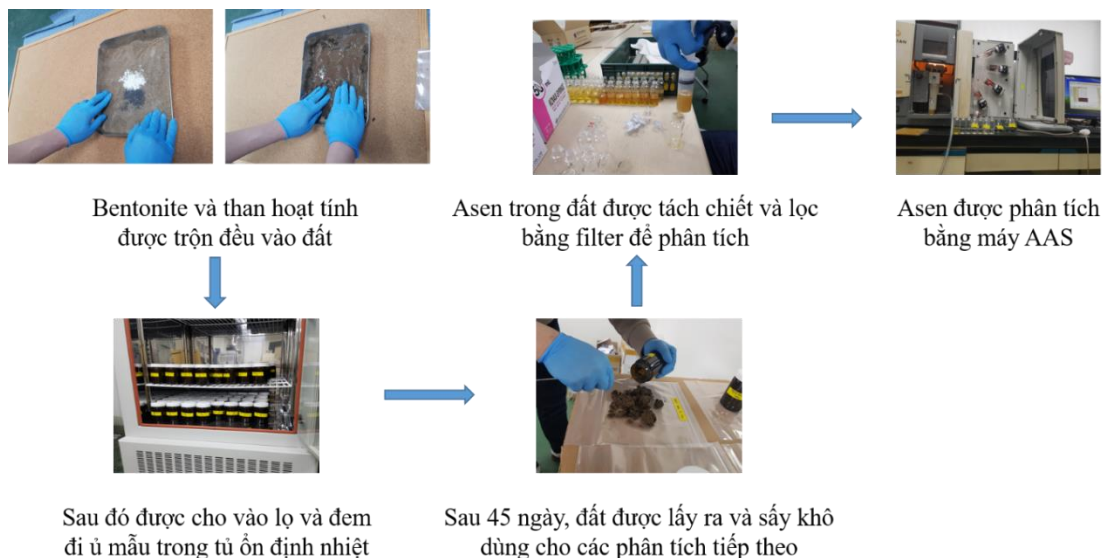
Đất bị ô nhiễm bởi As được lấy từ lớp bề mặt của một khu ruộng gần khu mỏ bỏ hoang ở tỉnh miền Trung Hàn Quốc (Hình 1). Sau khi lấy mẫu, đất được sấy khô ở nhiệt độ phòng trong 7 ngày và được sàng ở cỡ hạt 2 mm để dùng cho các thí nghiệm sau này. Vật liệu dùng trong nghiên cứu này là bentonite (B) và than hoạt tính (AC) được mua thương mại từ công ty Hóa chất Daejung. Vật liệu tổng hợp giữa bentonite và than hoạt tính (BAC) được tạo ra bằng cách trộn đều 2 chất trên theo tỉ lệ 50:50.



Hình 1. Sơ đồ địa điểm lấy mẫu.

2.2. Thiết kế thí nghiệm

Đất bị ô nhiễm bởi As được trộn với các vật liệu ở cùng một tỉ lệ là 2% giữa vật liệu/đất theo khối lượng. Sau khi đất được trộn với vật liệu, đất được cho vào các lọ lần lượt được dán nhãn CTR (đất không trộn vật liệu), B (đất được trộn bentonite), AC (đất được trộn than hoạt tính), BAC (đất được trộn cả bentonite và than hoạt tính). Sau đó, các lọ được đặt vào tủ ổn định nhiệt (incubator) trong 45 ngày và luôn được giữ ẩm ở 70%. Sau 45 ngày, các mẫu đất lần lượt được lấy ra để tiến hành các thí nghiệm về đo độ rửa trôi, và dạng tồn tại của As trong đất. Thí nghiệm được mô tả theo Hình 2.



Hình 2. Quy trình thực hiện nghiên cứu.

2.3. Phân tích mẫu

Để đánh giá khả năng rửa trôi của As trong các mẫu chứa và không chứa các vật liệu, ba phương pháp tách chiết được áp dụng bao gồm phương pháp SPLP (US EPA 1312), TCLP (US EPA 1311), và 0,1M HCl.

Để đánh giá khả năng cố định As của các vật liệu, hiệu số cố định được sử dụng theo công thức sau:

Hiệu suất cố định (%) = $(C_{CTR} - C_i)/C_{CTR}$. Trong đó, C_{CTR} là nồng độ As trong đất không chứa vật liệu, C_i là nồng độ As trong đất đã xử lý bằng các vật liệu B, AC, and BAC.

Để đánh giá các dạng tồn tại của As trong đất, nghiên cứu thực hiện tách chiết tuần tự dựa theo nghiên cứu [19]. Trong đó, As sẽ bám vào và tồn tại ở 5 dạng: dễ trao đổi, bám vào các hợp chất carbonate, bám vào các hợp chất oxit Fe và Mn, bám vào các hợp chất hữu cơ, và bám vào các phần còn lại trong đất. Trong 5 dạng tồn tại đó, dạng dễ trao đổi và các hợp chất carbonate là nơi mà As dễ linh động nhất, và có khả năng bị hấp thụ nhiều nhất bởi thực vật. Do đó, việc xem xét sự dịch chuyển của As trong các dạng tồn tại cũng là một cách hiệu quả trong việc đánh giá sự ảnh hưởng của các vật liệu đối với quá trình cố định As trong đất. Hàm lượng As trong đất được đo bằng máy quang phổ hấp thụ nguyên tử (AAS, AA240, Varian, Australia).

2.3. Phân tích dữ liệu

Nghiên cứu được lặp lại ba lần để lấy dữ liệu phân tích sự khác biệt của các vật liệu. Dữ liệu được phân tích bằng phương pháp ANOVA với kiểm định Tukey ($p < 0,05$) bằng phần mềm thống kê SPSS 20,0. Sau đó, nghiên cứu dùng phần mềm Origin 9,1 mô tả dữ liệu.

3. Kết quả và thảo luận

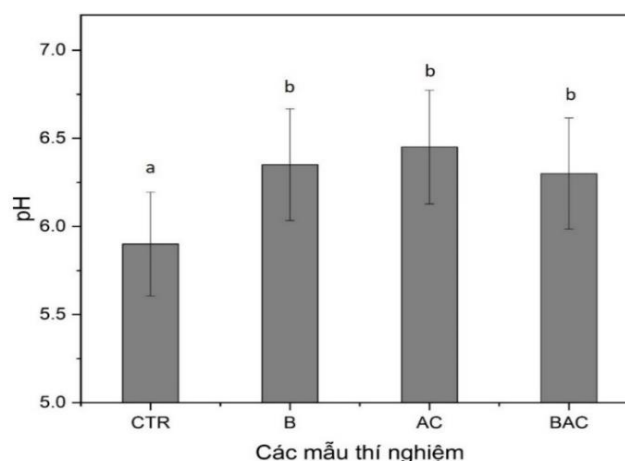
3.1. Ảnh hưởng của vật liệu đến độ pH của đất

Đất dùng trong nghiên cứu là đất bị ô nhiễm As với nồng độ tương đối cao với hàm lượng $58,5 \pm 2,78$ mg/kg (Bảng 1). Đất có độ chua nhẹ ở pH 5,9, hàm lượng hữu cơ thấp 2,2% và có thành phần vật chất chủ yếu ở dạng cát bùn.

Bảng 1. Một số đặc tính lý hóa của đất dùng trong nghiên cứu.

	pH	As (mg/kg)	Hàm lượng hữu cơ (%)	Kết cấu
Đất	$5,9 \pm 0,29$	$58,5 \pm 2,78$	$2,2 \pm 0,28$	Cát bùn

Ảnh hưởng của vật liệu đến các đặc tính lý hóa của đất được đánh giá thông qua độ pH đất (Hình 3). Độ pH của đất là một yếu tố quan trọng ảnh hưởng đến hoạt động của kim loại trong đất và trực tiếp hoặc gián tiếp làm thay đổi tính di động và độc tính của chúng đối với thực vật và vi sinh vật [20]. Nghiên cứu đã chỉ ra rằng các vật liệu như bentonite, than hoạt tính góp phần cải tạo chất lượng đất, làm tăng đáng kể pH trong đất từ $5,9 \pm 0,29$ trong mẫu đất đối sánh (không chứa vật liệu) đến $6,45 \pm 0,32$ trong mẫu đất có chứa than hoạt tính sau 45 ngày trong tủ ổn nhiệt. Kết quả cũng cho thấy pH trong tất cả các mẫu đều có sự khác biệt đáng kể ($p < 0,05$) so với mẫu đất đối sánh. Kết quả này phù hợp với một số nghiên cứu trước đó, phải kể đến [21] báo cáo rằng, bằng việc sử dụng bentonite trong xử lý đất nhiễm chì, pH của đất tăng từ 5,18 lên tới 6,83 sau 2 năm xử

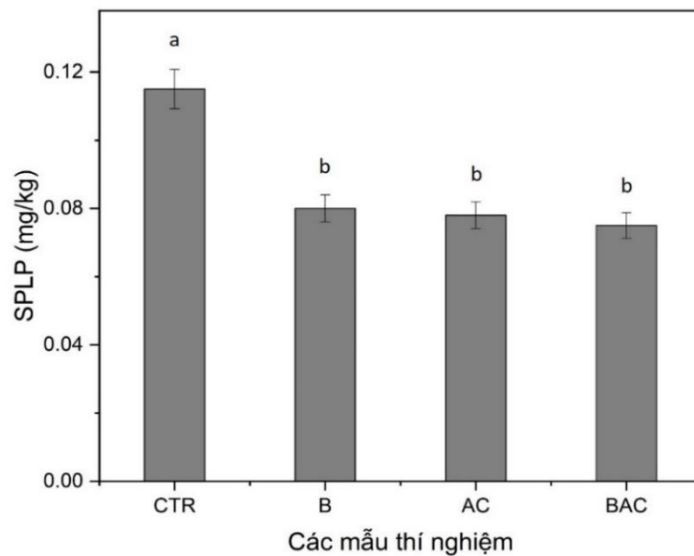


Hình 3. Độ pH của đất trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính. BAC: Đất có chứa cả bentonite và than hoạt tính).

lý. Nghiên cứu [18] cũng chỉ ra rằng AC được bổ sung vào đất ô nhiễm đồng, giúp cải thiện pH của đất từ 2,85 lên 4,2 sau 28 ngày.

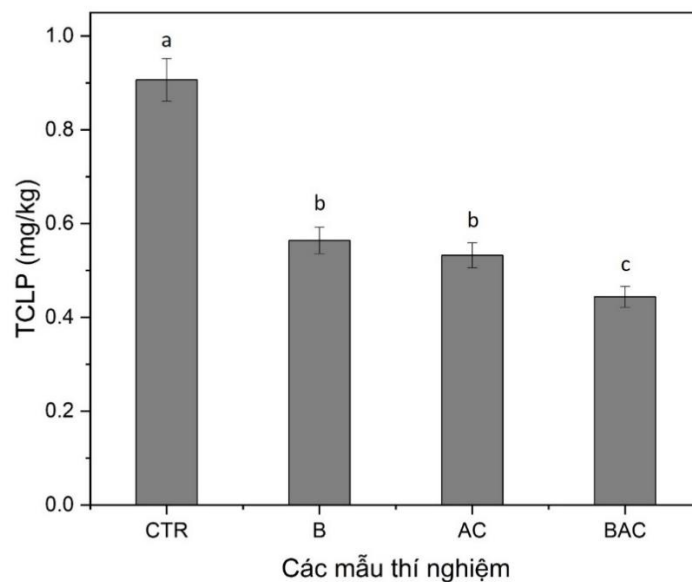
3.2. Ảnh hưởng của vật liệu đến tính rửa trôi của As trong đất

Ảnh hưởng của B, AC, và BAC đến khả năng rửa trôi của As được đánh giá bằng 3 quy trình tách chiết, bao gồm: SPLP, TCLP và HCl 0,1M. Kết quả cho thấy As trong cả 3 quy trình tách chiết đều có sự khác biệt đáng kể ($p < 0,05$) so với mẫu đất đối sánh. Đối với quy trình tách chiết SPLP, đất được xử lý bằng B, AC và BAC có hàm lượng As trung bình là $0,08 \pm 0,001$ mg/kg so với đất đối sánh (CTR) là $0,12 \pm 0,005$ mg/kg (Hình 4).



Hình 4. Nồng độ As tách chiết bằng quy trình lọc kết tủa tổng hợp (SPLP) trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính, BAC: Đất có chứa cả bentonite và than hoạt tính).

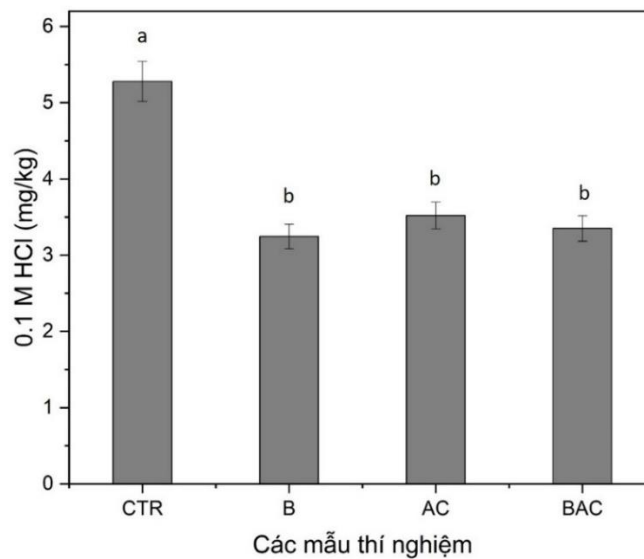
Nồng độ As trong dung dịch tách chiết tăng dần lần lượt đối với quy trình tách chiết TCLP và dung dịch chiết HCl 0,1M. Đối với quy trình tách chiết TCLP, đất được cải tạo bằng BAC cho hiệu quả cao nhất trong việc cố định As trong đất (Hình 5). Nồng độ As trong dung dịch chiết bằng quy trình TCLP trong mẫu đất có chứa BAC cho kết quả thấp nhất với nồng độ ở mức $0,44 \pm 0,02$ mg/kg, so với nồng độ As trong dung dịch chiết từ mẫu đất đối sánh là $0,9 \pm 0,04$ mg/kg. Trong khi đó, đất được xử lý bằng B và AC cho cùng hàm lượng As được chiết bằng quy trình TCLP với hàm lượng lần lượt là $0,56 \pm 0,03$ và $0,53 \pm 0,03$ mg/kg.



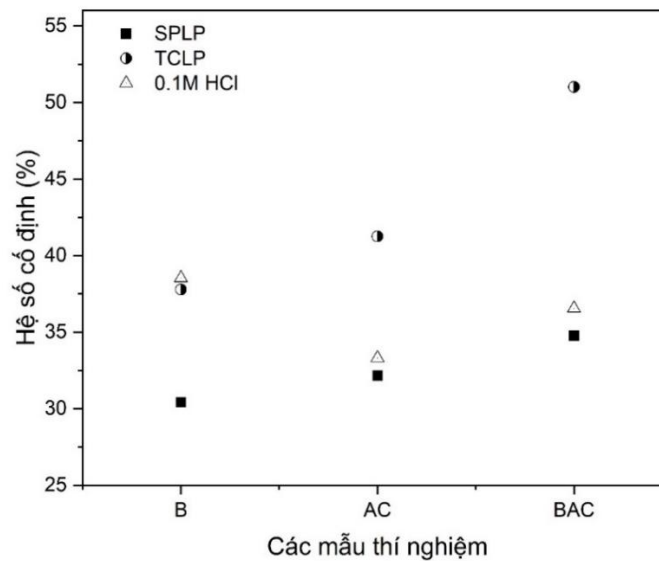
Hình 5. Nồng độ As tách chiết bằng quy trình chiết độc tính (TCLP) trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính, BAC: Đất có chứa cả bentonite và than hoạt tính).

Tuy nhiên, bằng quy trình tách chiết với dung dịch HCl 0,1M, đất được xử lý bằng B cho thấy hiệu quả cố định As trong đất cao nhất (Hình 6). Hàm lượng As được chiết bằng

dung dịch HCl 0,1M trong mẫu đất có chứa B thấp nhất với nồng độ $3,2 \pm 0,12$ mg/kg, so với đất đối sánh là $5,28 \pm 0,26$ mg/kg. Từ kết quả này, nghiên cứu chỉ ra rằng bằng cách tăng cường độ ion của dung dịch chiết, nồng độ As trong tất cả các thí nghiệm xử lý đều tăng theo thứ tự các quy trình tách chiết lần lượt từ SPLP < TCLP < 0,1M HCl.



Hình 6. Nồng độ As tách chiết bằng dung dịch 0,1M HCl trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính, BAC: Đất có chứa cả bentonite và than hoạt tính).

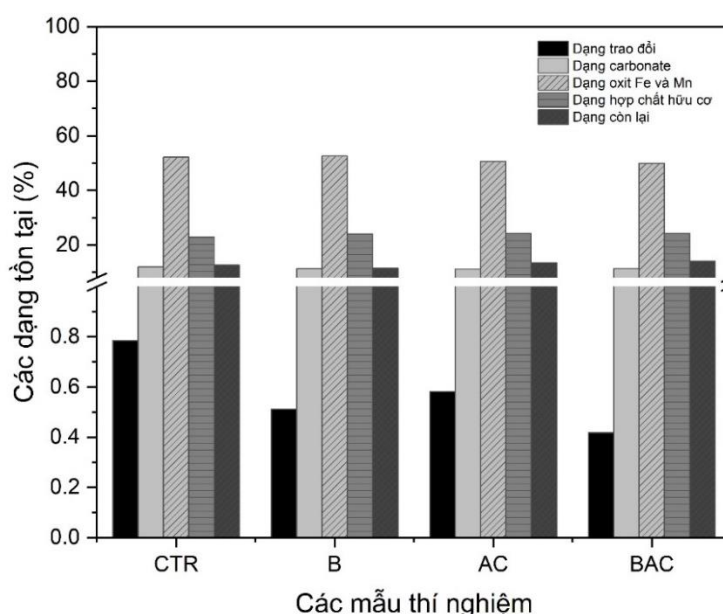


Hình 7. Hệ số cố định của As bằng các quy trình tách chiết trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính, BAC: Đất có chứa cả bentonite và than hoạt tính).

Hình 7 cho ta thấy hệ số cố định As khác nhau với mỗi quy trình tách chiết và mỗi mẫu đất được xử lý bằng vật liệu khác nhau. Nhìn chung, các vật liệu đều cho thấy khả năng cố định As tương đối tốt trong các quy trình tách chiết khác nhau, trung bình từ 30 - 40%. Đặc biệt, đất được xử lý bằng BAC cho hệ số cố định As cao nhất đối trong quy trình tách chiết bằng TCLP, lên tới 52%. Có thể thấy rằng, việc áp dụng các vật liệu như B hay AC đã góp phần làm thay đổi về tính chất vật lý và hóa học của đất, như pH. Từ đó ảnh hưởng trực tiếp đến tính linh động của As trong đất [22]. Hơn thế nữa, các nhóm chức của B và AC được chỉ ra trong các nghiên cứu trước đây như -OH group, Si-O, H-O-H, C=O đóng vai trò quan trọng trong việc cố định As thông qua trao đổi cation, hấp phụ hoặc tạo phức nhóm chức [23, 24].

3.3. Ảnh hưởng của vật liệu đến các dạng tồn tại của As trong đất

Kết quả của quá trình tách chiết tuần tự cho thấy As chủ yếu tồn tại trong các dạng vật chất hữu cơ và oxit Fe và Mn (Hình 8). Trong tất cả các mẫu có chứa vật liệu và mẫu đối sánh, As tồn tại trong dạng oxit Fe và Mn chiếm trung bình 50% về khối lượng. Phần trăm As tồn tại trong dạng oxit Fe và Mn trong mẫu đối sánh và mẫu đất có chứa B lần lượt là 52,1 và 52,5%; trong khi đó As trong mẫu chứa BAC cho tỉ lệ thấp nhất với 49%. Asen tồn tại trong các dạng hợp chất hữu cơ và dạng carbonate trong đất ở tất cả các mẫu có chứa vật liệu và không chứa vật liệu đều không có sự khác biệt đáng kể, chiếm tỉ lệ trung bình lần lượt là 11% và 24%. Tuy nhiên, sự khác biệt thể hiện rõ nhất ở dạng trao đổi và các dạng bền trong đất. Đối với dạng trao đổi, As chiếm tỉ lệ thấp nhất ở 0,4% trong mẫu đất chứa BAC, và cao nhất là 0,8% trong mẫu đối sánh. Ngược lại, As tồn tại trong các dạng bền trong đất chiếm tỉ lệ cao nhất trong mẫu chứa BAC là 14%, so với mẫu không chứa vật liệu (mẫu đối sánh) là 12%.



Hình 8. Các dạng tồn tại của As trong đất trong mỗi mẫu thí nghiệm. (CTR: Đất đối sánh (không có vật liệu), B: Đất có chứa bentonite, AC: Đất có chứa than hoạt tính, BAC: Đất có chứa cả bentonite và than hoạt tính).

Mặc dù tổng hàm lượng As trong đất biểu thị mức độ chung của ô nhiễm As nhưng nó không cung cấp thông tin về các dạng hóa học hoặc khả năng di chuyển, cũng như tính sinh khả dụng của nó. Trong khi đó, tính sinh khả dụng của As liên quan đến các phân tử bị hấp thụ (ví dụ như các dạng dễ trao đổi và cacbonat) [13]. Kết quả của tách chiết tuần tự As chứng tỏ rằng, các vật liệu đã góp phần chuyển hóa các dạng tồn tại của As trong đất từ dạng dễ trao đổi, sang các dạng tồn tại bền hơn, không còn linh động và dễ dàng thấm vào nước ngầm hay hấp thụ bởi thực vật.

Bảng 2 cho thấy sự so sánh về mức độ cố định As trong đất của một số vật liệu như biochar, carbon foam, hay nanosilica với vật liệu được dùng trong nghiên cứu này. Có thể thấy rằng đây là những vật liệu hoặc là có nguồn gốc hữu cơ tự nhiên như: *hydroxyapatite - biochar from rice straw* và *ferrihydrate - supported animal derived biochar* hoặc là có nguồn gốc nhân tạo như: carbon foam và nano-silica với đặc điểm chung là cho hiệu suất cố định As tương đối cao. Tuy nhiên, đây là các vật liệu có giá thành cao, không có sẵn, quá trình tổng hợp vật liệu phức tạp. Các vật liệu được dùng trong nghiên cứu này tương đối rẻ, dễ dàng tìm thấy trên thị trường. Qua nghiên cứu cũng cho thấy hiệu suất cố định As ở mức tương đối, đặc biệt là đối với vật liệu BAC. Đây là kết quả đầy hứa hẹn cho việc cố định As

trong đất thông qua việc giảm tính linh động của As bằng việc sử dụng các vật liệu giá rẻ, có sẵn, trong đó việc kết hợp hai hay nhiều vật liệu có thể sẽ đem lại hiệu quả cao.

Bảng 2. So sánh khả năng cố định As trong đất của một số vật liệu.

Vật liệu	Tỉ lệ áp dụng	Hiệu suất cố định	Loại đất	Tài liệu tham khảo
Hydroxyapatite - biochar from rice straw	1%	28%	Đất gần khu công nghiệp	[25]
Ferrihydrite - supported animal derived biochar	3%	40%	Khu vực luyện kim	[26]
Goethite-based carbon foam	20%	80%	Khu vực luyện kim	[27]
Mercapto propyltrimethoxysilane- and ferrous sulfate-modified nano-silica	3%	80%	Đất nông nghiệp	[28]
Bentonite	2%	38%	Đất nông nghiệp gần khu mỏ bỏ hoang	Nghiên cứu này
Than hoạt tính	2%	42%	Đất nông nghiệp gần khu mỏ bỏ hoang	Nghiên cứu này
Bentonite + Than hoạt tính	2%	52%	Đất nông nghiệp gần khu mỏ bỏ hoang	Nghiên cứu này

4. Kết luận

Nghiên cứu đã chỉ ra tiềm năng của B, AC, cũng như vật liệu tổng hợp của chúng (BAC) trong việc cố định As trong đất bị ô nhiễm. Sự kết hợp giữa bentonite và than hoạt tính (BAC) như một phương pháp xử lý đầy hứa hẹn, tăng cường đáng kể khả năng cố định As trong đất và chuyển các dạng tồn tại của As trong đất sang các dạng tồn tại khác có trạng thái ít linh hoạt hơn.

Những phát hiện này góp phần có giá trị trong việc tìm phương án hữu hiệu cho quá trình xử lý đất bị ô nhiễm As cũng như các kim loại nặng khác một cách bền vững, nhằm giảm thiểu ô nhiễm, bảo vệ môi trường, và bảo vệ sức khỏe con người.

Tuy nhiên, bên cạnh tính hiệu quả của việc áp dụng các vật liệu đối với quá trình cố định As trong đất thì cũng cần xem xét đến khía cạnh ảnh hưởng của nó đối với chất lượng đất, hệ sinh thái đất như vi sinh vật, thực vật. Đây cũng là hạn chế của nghiên cứu, cần được bổ sung trong các nghiên cứu sau.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: N.Q.T.; Lựa chọn phương pháp nghiên cứu: N.Q.T.; Thực hiện thí nghiệm: N.H.Q.; Phân tích mẫu: N.H.Q.; Xử lý số liệu: N.Q.T., N.T.H.; Viết bản thảo bài báo: N.Q.T.; Chỉnh sửa bài báo: N.T.H., N.Q.T.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội trong đề tài mã số TN.23.14.

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. Nejad, D.Z.; Kim, J.W.; Jung, M.C. Reclamation of arsenic contaminated soils around mining site using solidification/stabilization combined with revegetation. *Geosci. J.* **2017**. *21*(3), 385–396. Doi:10.1007/s12303-016-0059-0.
2. Agency, U.S.E.P. Arsenic compounds. Available online: epa.gov/sites/default/files/2021-04/documents/arsenic_april_2021.pdf (accessed December 12, 2021). 2021.
3. MEP, M. National soil pollution survey bulletin. Ministry of Environmental Protection and Ministry of Land and Resources Beijing, 2014.

4. Zhang, L.; Hu, J.; Li, C.; Chen, Y.; Zheng, L.; Ding, D.; Shan, S. Synergistic mechanism of iron manganese supported biochar for arsenic remediation and enzyme activity in contaminated soil. *J. Environ. Manage.* **2023**, *347*, 119127. <https://doi.org/10.1016/j.jenvman.2023.119127>.
5. Tarvainen, T.; Albanese, S.; Birke, M.; Poňavič, M.; Reimann, C.; The GEMAS Project Team. Arsenic in agricultural and grazing land soils of Europe. *Appl. Geochem.* **2013**, *28*, 2–10. <https://doi.org/10.1016/j.apgeochem.2012.10.005>.
6. Freeman, G.B.; Schoof, R.A.; Ruby, M.V.; Davis, A.O.; Dill, J.A.; Liao, S.C.; Lapin, C.A.; Bergstrom, P.D. Bioavailability of arsenic in soil and house dust impacted by smelter activities following oral administration in cynomolgus monkeys. *Fundam. Appl. Toxicol.* **1995**, *28*(2), 215–22. Doi:10.1006/faat.1995.1162.
7. Phuong, N.M.; Kang, Y.; Sakurai, K.; Iwasaki, K.; Kien, C.N.; Noi, N.V.; Son, L.T. Arsenic contents and physicochemical properties of agricultural soils from the Red River Delta, Vietnam. *Soil Sci. Plant Nutr.* **2008**, *54*(6), 846–855. Doi:10.1111/j.1747-0765.2008.00312.x.
8. Yang, J. et al. Phytoaccumulation of As by *Pteris vittata* supplied with phosphorus fertilizers under different soil moisture regimes—a field case. *Ecol. Eng.* **2019**, *138*, 274–280.
9. Liang, J. et al. Tolerance and Bioaccumulation of arsenate by *Aspergillus oryzae* TLWK-09 isolated from arsenic-contaminated soils. *Water Air Soil Pollut.* **2018**, *229*, 1–11.
10. Jeon, E.K. et al. In situ electrokinetic remediation of As-, Cu-, and Pb-contaminated paddy soil using hexagonal electrode configuration: a full scale study. *Environ. Sci. Pollut. Res.* **2015**, *22*, 711–720.
11. He, J. et al. Removal of arsenic from contaminated soils by combining tartaric acid with dithionite: An efficient composite washing agent. *J. Environ. Chem. Eng.* **2023**, *11*(3), 109877. <https://doi.org/10.1016/j.jece.2023.109877>.
12. Fazle Bari, A.S.M. et al. Soil washing of arsenic from mixed contaminated abandoned mine soils and fate of arsenic after washing. *Chemosphere* **2022**, *296*, 134053. <https://doi.org/10.1016/j.chemosphere.2022.134053>.
13. Palansooriya, K.N. et al. Soil amendments for immobilization of potentially toxic elements in contaminated soils: A critical review. *Environ. Int.* **2020**, *134*, 105046. <https://doi.org/10.1016/j.envint.2019.105046>.
14. Derakhshan Nejad, Z. et al. Effects of fine fractions of soil organic, semi-organic, and inorganic amendments on the mitigation of heavy metal(loid)s leaching and bioavailability in a post-mining area. *Chemosphere* **2021**, *271*, 129538. <https://doi.org/10.1016/j.chemosphere.2021.129538>.
15. Nabilah, B. et al. Methylene Blue biodecolorization and biodegradation by immobilized mixed cultures of *Trichoderma viride* and *Ralstonia pickettii* into SA-PVA-Bentonite matrix. *Arabian J. Chem.* **2023**, *16*(8), 104940. <https://doi.org/10.1016/j.arabjc.2023.104940>.
16. Hebbar, R.S.; Isloor, A.M.; Ismail, A. Preparation and evaluation of heavy metal rejection properties of polyetherimide/porous activated bentonite clay nanocomposite membrane. *RSC Adv.* **2014**, *4*(88), 47240–47248.
17. Oussalah, A.; Boukerroui, A. Alginate-bentonite beads for efficient adsorption of methylene blue dye. *Euro-Mediterr. J. Environ. Integr.* **2020**, *5*, 1–10.
18. Quoc, T.N.; Jung, M.C. Sequential Application of Column Leaching and Plant Uptake Tests to Assess the Effect of Various Commercial Amendments on Cu Immobilization in Ultra-High Cu-Contaminated Soil. *Toxics* **2022**, *10*(4), 185. Doi: 10.3390/toxics10040185.

19. Wenzel, W.W. et al. Arsenic fractionation in soils using an improved sequential extraction procedure. *Anal. Chim. Acta* **2001**, 436(2), 309–323. [https://doi.org/10.1016/S0003-2670\(01\)00924-2](https://doi.org/10.1016/S0003-2670(01)00924-2).
20. Zhong, X. et al. Factors influencing heavy metal availability and risk assessment of soils at typical metal mines in Eastern China. *J. Hazard. Mater.* **2020**, 400, 123289. <https://doi.org/10.1016/j.jhazmat.2020.123289>.
21. Vrînceanu, N.O. et al. Assessment of using bentonite, dolomite, natural zeolite and manure for the immobilization of heavy metals in a contaminated soil: The Copșa Mică case study (Romania). *Catena* **2019**, 176, 336–342. <https://doi.org/10.1016/j.catena.2019.01.015>.
22. Xing, W. et al. Effect of soluble phosphate and bentonite amendments on lead and cadmium bioavailability and bioaccessibility in a contaminated soil. *Sci. Total Environ.* **2023**, 900, 166370. <https://doi.org/10.1016/j.scitotenv.2023.166370>.
23. Que, W. et al. Appraising the effect of in-situ remediation of heavy metal contaminated sediment by biochar and activated carbon on Cu immobilization and microbial community. *Ecol. Eng.* **2019**, 127, 519–526. <https://doi.org/10.1016/j.ecoleng.2018.10.005>.
24. Zhirong, L.; Uddin, M.A.; Zhanxue, S. FT-IR and XRD analysis of natural Na-bentonite and Cu(II)-loaded Na-bentonite. *Spectrochim. Acta, Part A* **2011**, 79(5), 1013–1016. <https://doi.org/10.1016/j.saa.2011.04.013>.
25. Ahmed, W. et al. Simultaneous immobilization of lead and arsenic and improved phosphorus availability in contaminated soil using biochar composite modified with hydroxyapatite and oxidation: Findings from a pot experiment. *Environ. Res.* **2023**, 235, 116640. <https://doi.org/10.1016/j.envres.2023.116640>.
26. Hong, C. et al. Simultaneous and long-term effective immobilization of lead, cadmium and arsenic in multi-contaminated soil by ferrihydrite-supported animal-derived biochar. *J. Environ. Chem. Eng.* **2023**, 11(3), 109989. <https://doi.org/10.1016/j.jece.2023.109989>.
27. Janeiro-Tato, I. et al. Goethite-based carbon foam nanocomposites for concurrently immobilizing arsenic and metals in polluted soils. *Chemosphere* **2022**, 301, 134645. <https://doi.org/10.1016/j.chemosphere.2022.134645>.
28. Cao, P. et al. Mercapto propyltrimethoxysilane- and ferrous sulfate-modified nano-silica for immobilization of lead and cadmium as well as arsenic in heavy metal-contaminated soil. *Environ. Pollut.* **2020**, 266, 115152. <https://doi.org/10.1016/j.envpol.2020.115152>.

Comparing the effect of various amendments for Arsenic immobilization in contaminated soil

Nguyen Quoc Tuan^{1*}, Nguyen Hong Quan¹, Nguyen Thi Hai^{1,2}

¹ Faculty of Geology, Hanoi University of Science, Vietnam National University; nqt18@vnu.edu.vn; nguyenhongquan160198@gmail.com

² VNU Key Laboratory of Geo-environment and Climate change Response, University of Science, Vietnam National University, Hanoi; nguyenhai128@vnu.edu.vn

Abstract: The study compared the effects of three materials: bentonite, activated carbon, and their combination with an application rate of 2% on Arsenic (As) immobilization in contaminated soil. The As immobilization efficiency in soil was evaluated by the following methods: i) Synthetic Precipitation Leaching Procedure (SPLP), Toxicity Characteristic Leaching Procedure (TCLP), and extracted with 0.1M HCl solution to evaluate As leaching behaviour; and ii) Sequential Extraction to evaluate the fractions of As in soil. The results showed that the presence of materials help increase soil pH, and change the mobility of As

in soil. The As immobilization efficiency increased up to 52% in soil treated with combination of bentonite and activated carbon (BAC). The results also showed that As in exchangeable form were transformed into the immobile fraction in soil, especially for BAC treated soil. The study indicates that the combination of materials is highly effective in remediating contaminated soil, so further research needed to evaluate its effectiveness not only against As but also against other heavy metals and consider its impact on soil health.

Keywords: Arsenic; Bentonite; Activated carbon; Heavy metal immobilization in soil.

Bài báo khoa học

Đánh giá tiềm năng năng lượng gió cho phát triển năng lượng tái tạo trên khu vực Biển Đông, Việt Nam

Nguyễn Việt Hằng¹, Phạm Duy Huy Bình¹, Phạm Quang Nam¹, Nguyễn Bách Tùng¹,
Trịnh Tuấn Long^{1*}

¹ Trung tâm Động lực học Thủy khí Môi trường, trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội; viethang1485@gmail.com; phambinh@hus.edu.vn; nampq@hus.edu.vn; bachtung_cefd@hus.edu.vn; trinhtuanlong@hus.edu.vn

*Tác giả liên hệ: trinhtuanlong@hus.edu.vn; Tel: +84-969622708

Ban Biên tập nhận bài: 5/10/2022; Ngày phản biện xong: 1/11/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Trong nghiên cứu này, mô hình khí hậu khu vực RegCM4 (RCM) được sử dụng để mô phỏng trường gió từ đó xác định năng lượng gió tiềm năng cho khu vực Biển Đông trong giai đoạn 20 năm (2000-2019). Trường vận tốc gió được mô phỏng từ mô hình RCM khi so sánh với số liệu tái phân tích ERA5 sử dụng dữ liệu quan trắc gió từ 13 trạm khí tượng hải văn ven biển và đảo của Việt Nam cũng như 10 trạm đo gió mực cao 80 m. Kết quả khi so sánh với dữ liệu quan trắc gió cho thấy RCM mô phỏng tốt và phù hợp hơn so với sử dụng số liệu từ ERA5. Phân bố không gian của kết quả tính toán năng lượng gió từ mô hình RCM cũng có sự tương đồng so với dữ liệu được công bố trong “Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam” của Bộ Tài nguyên và Môi trường và báo cáo của Ngân hàng thế giới năm 2009. Mặc dù, nghiên cứu vẫn còn có một số hạn chế tuy nhiên, kết quả của nghiên cứu cũng là nguồn thông tin sử dụng tham khảo tốt trong việc quy hoạch và xây dựng các dự án điện gió trên biển.

Từ khóa: Năng lượng gió; Biển Đông; Việt Nam; RegCM4.

1. Mở đầu

Năng lượng tái tạo (NLTT) đang được xem là giải pháp thay thế tiềm năng cho những nguồn năng lượng hóa thạch đang dần cạn kiệt do tác động của biến đổi khí hậu trên toàn cầu [1]. Các giải pháp sử dụng năng lượng tái tạo sẽ giúp giảm thiểu khí thải gây ô nhiễm môi trường, giảm sự phụ thuộc vào năng lượng hóa thạch và góp phần quan trọng trong việc chuyển đổi sang một hệ thống năng lượng sạch, giảm thiểu tác động tiêu cực đến môi trường và sức khỏe con người. Trong bối cảnh xã hội càng phát triển thì vai trò của năng lượng tái tạo lại càng được thể hiện rõ, việc đảm bảo nguồn cung cấp năng lượng tái tạo bền vững và ổn định đang là mục tiêu quan trọng của toàn nhân loại [2].

Các nguồn năng lượng tái tạo bao gồm năng lượng sinh học, thủy điện, nhiệt địa cầu, năng lượng mặt trời, năng lượng gió, năng lượng sóng và năng lượng thủy triều. Đứng trước bài toán lớn về năng lượng của thế giới và khủng hoảng điện năng, năng lượng gió được xem là một trong những giải pháp tiềm năng đáp ứng nhu cầu một cách bền vững [3]. Năng lượng gió đang là một trong những hình thức năng lượng tái tạo phổ biến nhất với một trong những điểm mạnh là tính sạch sẽ và không gây khí thải. Một trong những ưu điểm khác của năng lượng gió là khả năng tái tạo và tiềm năng còn rất lớn. Điều này giúp đảm bảo rằng năng lượng gió có thể sử dụng trong suốt thời gian dài mà không cần lo ngại về cạn kiệt nguồn tài nguyên. Tuy nhiên, việc xác định tiềm năng gió cũng rất quan trọng, bởi đó là nền tảng giúp

các nhà hoạch định chính sách và hỗ trợ nhà đầu tư đưa ra các quyết định đầu tư và phát triển các dự án năng lượng gió.

Hiện nay, trên thế giới xác định tiềm năng năng lượng gió nói riêng và năng lượng tái tạo nói chung vẫn luôn là chủ đề lớn được các nhà khoa học và các nhà quản lý quan tâm. Theo báo cáo của Ủy ban Liên chính phủ về Biến đổi Khí hậu (IPCC) [4], cho đến năm 2050, 80% nguồn cung cấp năng lượng cả thế giới có thể đến từ các nguồn năng lượng tái tạo và năng lượng gió sẽ đóng vai trò tạo điện vào năm 2050. Cùng với sự phát triển nhanh chóng của công nghệ, các dự án năng lượng gió ngoài khơi đã trở thành một xu hướng nghiên cứu tiềm năng tại nhiều quốc gia [5]. Tại Tây Ban Nha cũng như một số nước Châu Âu như Đan Mạch, Hà Lan, Thụy Điển..., việc sản xuất năng lượng gió ngoài khơi đã trải qua sự tăng trưởng đáng kể trong 5 năm với công suất lắp đặt lên tới gần 8000 MW, tuy nhiên sự phát triển của công nghệ này đang bị hạn chế với một số đặc điểm về địa hình xã hội cũng như thiếu sự ổn định về môi trường nhằm đảm bảo đầu tư [6]. Tại Nhật Bản, tiềm năng năng lượng gió tại bờ biển khu vực Kanto được ước tính dựa trên các chỉ số kinh tế và xã hội thông qua Hệ thống thông tin địa lý (GIS) [7]. Nếu không xem xét bất kỳ yếu tố kinh tế hoặc xã hội nào, tổng tiềm năng năng lượng gió dọc bờ biển khu vực này sẽ đạt 287 TWh/năm.

Việt Nam là một trong những quốc gia được biết đến có tiềm năng lớn trong việc phát triển năng lượng gió. Không chỉ dọc đường bờ biển trải dài từ Bắc xuống Nam mà ngay cả các vùng đồi núi, thung lũng, và khu vực ngoài khơi đều là những nơi có khả năng xem xét việc khai thác năng lượng gió hiệu quả. Việt Nam cũng đã và đang có rất nhiều những nghiên cứu về mặt kỹ thuật để phát triển năng lượng gió. Trong báo cáo “Tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam” [8] của bộ Tài nguyên và Môi trường công bố tháng 3 năm 2022 nhận định rằng các khu vực có khả năng khai thác năng lượng gió tốt nhất ở nước ta là khu vực Bình Định, Ninh Thuận, Bình Thuận đến Cà Mau và một phần vùng biển trung tâm Vịnh Bắc Bộ. Báo cáo trên đã sử dụng số liệu tái phân tích tại độ cao 100 m để tính toán, kết quả cho thấy tiềm năng năng lượng gió ở mức tốt đến rất tốt tại một số khu vực biển Ninh Thuận đến Bà Rịa - Vũng Tàu với tốc độ trung bình từ 8-10 m/s và mật độ năng lượng trung bình phổ biến từ 600 đến 700 W/m². Ngoài ra, phê duyệt Quy hoạch phát triển điện lực quốc gia thời kỳ 2021-2030, tầm nhìn đến 2050, gọi tắt là Quy hoạch điện VIII (PDP8) [9] được Thủ tướng Chính phủ phê duyệt theo Quyết định 500/QĐ-TTg ngày 15/5/2023 cũng đã cung cấp nhiều thông tin về khả năng phát triển điện gió tại Việt Nam, bao gồm điện gió khu vực đất liền và điện gió ngoài khơi theo từng giai đoạn cho đến năm 2045. Trong đó tại khu vực điện gió trên bờ và gần bờ, tổng quy mô tiềm năng điện gió khá lớn khoảng 217 GW và chủ yếu là tiềm năng gió thấp (4,5-5,5 m/s) khoảng 163 GW. Chi phí đầu tư điện gió trong tương lai có xu hướng giảm dần, đặt trong bối cảnh giai đoạn 2045, mặc dù nhiều nhận định cho rằng chỉ có khu vực tiềm năng gió cao (trên 6 m/s) và gió tiềm năng trung bình (5,5-6 m/s) mới có hiệu quả về mặt kinh tế tuy nhiên tổng năng lượng gió tiềm năng của hai khu vực này tại vùng trên bờ và gần bờ Việt Nam tương đối thấp (lần lượt là 24 GW và 30 GW) và tập chung chủ yếu tại Tây Nam Bộ, Tây Nguyên và Nam Trung Bộ. Trong khi đó điện gió ngoài khơi tại Việt Nam có tổng tiềm năng kỹ thuật khoảng 160 GW và khu vực có tiềm năng gió cao và tiềm lực kinh tế tốt nằm ở phía Nam Trung Bộ (Bình Thuận, Ninh Thuận, Khánh Hòa) với tổng tiềm năng khoảng 80 GW với tốc độ gió trên 7-9 m/s, ngoài ra, khu vực còn lại ở Trà Vinh, Hà Tĩnh và Quảng Ninh có tốc độ gió thấp hơn (6-7 m/s). Bên cạnh đó, dự thảo này cũng đưa ra một số giải pháp nhằm tăng cường phát triển điện gió tại Việt Nam, như chú trọng đầu tư vào các cơ sở hạ tầng, phát triển các dự án điện gió quy mô lớn, tăng cường năng lực kỹ thuật và tăng cường quản lý, giám sát các hoạt động điện gió.

Nhìn chung, các nghiên cứu về tiềm năng điện gió đã cung cấp những thông tin có giá trị về hiện trạng năng lượng gió tại Việt Nam cũng như đưa ra nhận định về các khu vực tiềm năng để phát triển. Bên cạnh đó, nhằm phục vụ cho việc tiếp cận dễ hơn các thông tin về tiềm năng năng lượng gió đối với các bên liên quan, tập bản đồ (Atlas) gió và năng lượng gió được

coi là một công cụ hỗ trợ hữu ích và đem lại hiệu quả. Sáng kiến về bộ atlas năng lượng gió đầu tiên trên thế là của Cộng đồng Châu Âu được xây dựng từ những năm giữa thế kỷ XX. Ngoài ra, tại một số nước phát triển như Mỹ, Châu Âu, Trung Quốc... các bộ atlas về năng lượng gió đã được xây dựng từ những năm cuối của thế kỷ XX. Ở khu vực Châu Á, atlas năng lượng gió đã được nhiều nước trong khu vực xây dựng và phát triển như Nhật Bản, Hàn Quốc, Trung Quốc. Đặc biệt, trong dự án đánh giá tài nguyên năng lượng gió do Ngân hàng thế giới (WB) tài trợ, công ty TrueWind của Hoa Kỳ (2021) [10] đã xây dựng tập atlas năng lượng gió cho khu vực Đông Nam Á, trong đó có Việt Nam. Tuy vậy bộ atlas này mới chỉ thể hiện những tiềm năng chung về năng lượng gió mà chưa xem xét tới các yếu tố về địa hình, phân loại sử dụng đất... để xác định cụ thể khu vực có khả năng xây dựng các nhà máy điện gió. Và đứng trước những bài toán đặt ra về năng lượng gió đối với sự phát triển năng lượng tại Việt Nam, cần thiết có cách xây dựng một bộ dữ liệu chuẩn và đánh giá nhằm khai thác và phát triển hiệu quả tiềm năng năng lượng gió trong tương lai, nhất là khu vực ven biển và ngoài khơi.

2. Khu vực, nguồn số liệu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Nghiên cứu này tập trung chủ yếu vào khu vực Biển Đông, nhất là khu vực ven bờ và ngoài khơi nhằm mục tiêu xây dựng và đánh giá một bộ cơ sở tiềm năng năng lượng gió ven bờ. Biển Đông là vùng biển lớn thứ tư trên thế giới, nơi có tuyến hàng hải nhộn nhịp đứng thứ hai thế giới, nằm giữa lục địa Đông Nam Á và chuỗi các đảo từ Đài Loan ở phía bắc, Philippine ở phía Đông đến Borneo ở phía nam, là phần rìa phía tây của vành đai Thái Bình Dương. Theo phương Bắc - Nam, Biển Đông trải dài từ vĩ độ 25° xuống đến vĩ độ 5° Bắc và từ kinh độ 100° đến 120° kinh độ Đông. Biển Đông nằm trọn trong khu vực nhiệt đới bắc bán cầu, là vành đai nhận được lượng bức xạ mặt trời trực tiếp nhiều nhất so với các vành đai khác. Biển Đông có chế độ khí hậu mang tính chất gió mùa nhiệt đới hải dương và được chia thành hai khu vực khí hậu phía bắc và phía nam cùng với nhiều đặc điểm riêng không giống với đặc điểm khí hậu trên đất liền.

Bên cạnh đó, biển Đông cũng là khu vực có tiềm năng năng lượng gió biển lớn. Tại khu vực từ Bình Thuận đến Cà Mau, khoảng cách từ bờ ra đến 300 km là nơi có tốc độ gió đạt từ 7-11 m/s, cũng là nơi tiềm năng năng lượng gió rất lớn. Ngoài ra trên bản đồ phân bố tiềm năng gió trung bình ở độ cao 100 m cho thấy trên Biển Đông, vùng kéo dài dọc hướng Đông Bắc - Tây Nam từ eo biển Đài Loan tới vùng biển khu vực Đông Nam Bộ nước ta có tiềm năng năng lượng khá cao, hàng năm đạt 600-800 W/m². Còn phân tích theo độ sâu, địa hình và tốc độ gió trung bình năm dựa vào chuỗi số liệu 10 năm gần đây, khu vực biển ven bờ Việt Nam được phân chia thành 5 khu vực như sau: Quảng Ninh - Quảng Trị (biển thoải, nông, mật độ năng lượng gió vừa); Quảng Bình - Quảng Ngãi (Biển thoải, hẹp, mật độ năng lượng gió thấp); Bình Định - Ninh Thuận (biển nông hẹp, mật độ năng lượng gió thấp); Bình Thuận - Mũi Cà Mau (biển thoải, nông, mật độ năng lượng gió cao); Mũi Cà Mau - Kiên Giang (biển nông, mật độ năng lượng gió vừa). Dựa trên một số đặc điểm cơ bản về khu vực Biển Đông tại Việt Nam, có thể thấy đây là khu vực có tiềm năng phát triển năng lượng gió rất lớn cho đất nước [11].

2.2 Số liệu nghiên cứu

Trong nghiên cứu này, năng lượng gió được tính toán sử dụng kết quả mô phỏng trường gió từ mô hình khí hậu khu vực RegCM4 (RCM) [12], sử dụng các điều kiện biên và điều kiện ban đầu từ dữ liệu tái phân tích toàn cầu ERA5 nhằm mô phỏng trường gió cho giai đoạn 20 năm từ năm 2000 đến hết năm 2019. Do mục tiêu xây dựng bản đồ phân bố gió ngoài khơi khu vực Biển Đông, các biến gió mực bề mặt 10 m và mực cao 80 m từ kết quả mô hình RCM cùng với số liệu tái phân tích ERA5 sẽ được đánh giá với số liệu quan trắc tương ứng.

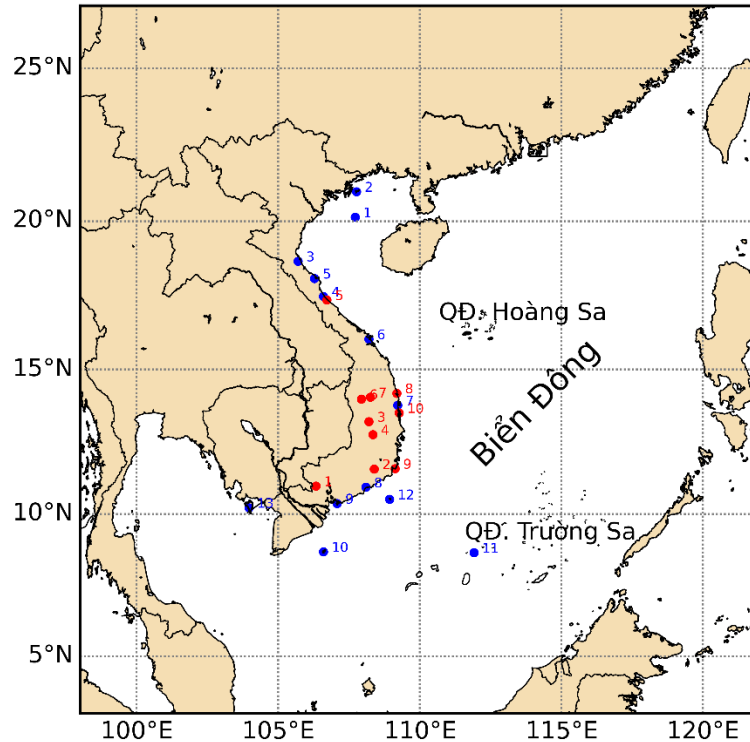
Sau đó, dựa trên các kết quả đánh giá, tốc độ gió tại mực 100 m sẽ được sử dụng làm đầu vào xây dựng nên bản đồ năng lượng gió tiềm năng cho khu vực biển Đông.

2.2.1. Số liệu quan trắc gió

Trong nghiên cứu này sử dụng hai bộ dữ liệu gió quan trắc gồm: (1) Số liệu quan trắc gió cực đại ngày (V_{max}) tại 13 trạm khí tượng ven biển được sử dụng để so sánh, đánh giá khả năng mô phỏng của mô hình ở mực thấp. Đây đều là các trạm khí tượng ven biển hoặc ở đảo, đo gió ở độ cao tiêu chuẩn mực 10 m, do vậy ảnh hưởng của địa hình hoặc các công trình xây dựng, vật cản là không đáng kể so với các trạm khí tượng nằm sâu trong đất liền. Số liệu V_{max} hàng ngày thu thập được có từ năm 1961 đến nay, tuy nhiên nghiên cứu này chỉ sử dụng dữ liệu trong 20 năm, giai đoạn 2000-2019. Vị trí của 13 trạm được thể hiện trên bảng 1 và được đánh dấu bằng chấm tròn màu xanh trên hình 1; (2) Số liệu từ 10 trạm quan trắc gió tại các cột đo gió tầng cao tại mực độ cao 80 m thu thập từ gói thầu “Đo gió phục vụ xây dựng quy hoạch điện gió và các dự án điện gió” thuộc khuôn khổ dự án “Xây dựng khung pháp lý và hỗ trợ kỹ thuật cho kết nối lưới điện và phát triển điện gió ở Việt Nam”. Chương trình được tài trợ bởi Bộ Môi trường, Bảo tồn Thiên nhiên và An toàn Hạt nhân Liên bang (BMU), Cộng hòa Liên bang Đức do *Deutsche Gesellschaft International Zusammenarbeit* (GIZ) [13] và Bộ Công Thương (MoIT) Việt Nam phối hợp thực hiện. Mục tiêu của chương trình là hỗ trợ chính quyền địa phương triển khai quy hoạch điện gió tại các tỉnh, đồng thời hỗ trợ các nhà phát triển tư nhân trong các dự án điện gió. Số liệu gió từ dự án GIZ có trong giai đoạn năm 2012-2017, tùy thuộc vào từng điểm đo với tần suất mỗi 10 phút. Thông tin về các trạm đo được thể hiện trên bảng 1, vị trí các điểm đo được đánh dấu bằng đường chấm đỏ trên hình 1.

Bảng 1. Danh sách và thông tin các trạm đo gió mực 10 m (13 trạm) và 80 m (10 trạm).

STT	Tên Trạm (Mực 10 m)	Kinh Độ	Vĩ Độ	STT	Tên Trạm (Mực 80 m)	Kinh Độ	Vĩ Độ
1	Bạch Long Vĩ	107,717	20,133	1	An Ninh Đông	10,973	106,332
2	Cô Tô	107,767	20,983	2	Đà Loan	11,562	108,391
3	Vinh	105,700	18,667	3	Ea Drăng	13,206	108,205
4	Đồng Hới	106,600	17,483	4	Ea Phê	12,758	108,344
5	Kỳ Anh	106,283	18,083	5	Hải Ninh	17,358	106,721
6	Đà Nẵng	108,200	16,033	6	Ia Dêr	13,984	108,256
7	Quy Nhơn	109,217	13,767	7	Kon Dong	14,045	108,256
8	Phan Thiết	108,100	10,933	8	Mỹ Thanh	14,177	109,190
9	Vũng Tàu	107,083	10,367	9	Thanh Hải	11,580	109,128
10	Côn Đảo	106,600	8,683	10	Xuân Hòa	13,508	109,273
11	Trường Sa	111,917	8,650				
12	Phú Quý	108,933	10,517				
13	Phú Quốc	103,967	10,217				



Hình 1. Vị trí đo gió mực 10 m tại các trạm khí tượng (màu xanh) và các trạm đo gió mực cao 80 m (màu đỏ).

2.2.2. Số liệu tái phân tích ERA5

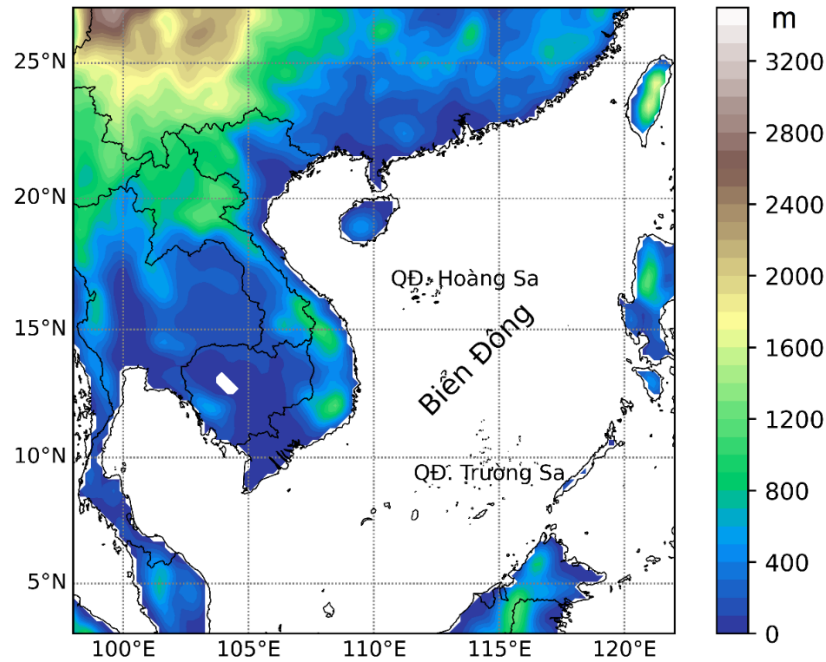
ERA5 là số liệu tái phân tích thế hệ thứ 5 của Trung tâm Dự báo hạn vừa Châu Âu (ECMWF). Đây là bộ số liệu toàn cầu có độ phân giải cao và được sử dụng phổ biến nhất trong cộng đồng hiện nay, với độ phân giải ngang là $0,25 \times 0,25$ độ kinh vĩ, có từ năm 1979 cho đến nay. ERA5 có số liệu với độ phân giải thời gian cao nhất là từng giờ được tính toán và nội suy từ kết quả tổ hợp mô hình độ phân giải thời gian 3 giờ và phương ngang $0,5$ độ kinh vĩ [14]. Kỹ thuật đồng hóa 4D-Var được sử dụng với hệ thống IFS (*Integrated Forecasting System*) [15] để tạo ra ERA5.

2.3 Phương pháp nghiên cứu

2.3.1. Mô hình RCM

Trong nghiên cứu này phiên bản thứ tư của mô hình khí hậu khu vực RegCM được sử dụng. RCM [16] sử dụng dữ liệu địa hình như độ cao, sử dụng đất, bề mặt biển,... cũng như các điều kiện ban đầu và điều kiện biên. RCM có thể chạy với điều kiện biên từ các mô hình khí hậu toàn cầu GCM (*Global climate model*) [17]. Trong nghiên cứu này, số liệu tái phân tích ERA5 được sử dụng làm điều kiện biên và điều kiện ban đầu cho mô hình RCM [18]. So với các phiên bản trước, phiên bản RCM thế hệ thứ tư được cải tiến hơn, bao gồm: một số sơ đồ tham số hóa mới như sơ đồ quá trình đất bề mặt *CML*, sơ đồ lớp biên hành tinh *UW*, lớp biên *Holtslag*, sơ đồ chuyển đổi phát xạ và một số cấu hình mới linh hoạt và dễ áp dụng hơn với các trình biên dịch khác nhau [19].

Trong nghiên cứu này, RCM được chạy với độ phân giải là 10 km, sử dụng sơ đồ tham số hóa đối lưu *Kain-Fritsch*, sơ đồ lớp biên hành tinh *UW*, sơ đồ âm *SUBEX*. Đây là bộ cấu hình phù hợp nhất cho mô phỏng hoàn lưu khí quyển cho khu vực Đông Nam Á cũng như Biển Đông Việt Nam [12]. Sản phẩm trường gió kinh hướng và vĩ hướng đầu ra từng giờ được sử dụng để so sánh trực tiếp với quan trắc cũng như làm đầu vào tính toán năng lượng gió tiềm năng.



Hình 2. Miền tính và bản đồ địa hình cho mô hình RCM mô phỏng trường gió.

2.3.2. Phương pháp tính toán năng lượng từ trường gió

Để đánh giá tiềm năng năng lượng gió tại một độ cao nào đó của khu vực, cần phải biết giá trị tốc độ gió ở độ cao đó. Tuy nhiên, các trạm khí tượng tiêu chuẩn hiện nay đo gió ở mực 10 m, các trạm đo gió ở độ cao lớn như các trạm quan trắc cao không nhiều, nên để xác định đặc trưng gió ở độ cao lớn phải tính toán một cách gián tiếp dựa vào tốc độ gió quan trắc được từ các trạm khí tượng bề mặt bằng một hàm phân bố gió theo độ cao.

Phân bố gió theo độ cao ở từng khu vực, từng thời điểm cụ thể phụ thuộc không chỉ vào độ gồ ghề của mặt đệm mà cả tầng kết nhiệt của khí quyển và một số yếu tố khác. Hiện nay, phần lớn các nghiên cứu trên thế giới cũng như ở Việt Nam đều sử dụng quy luật loga để ước lượng gián tiếp phân bố gió theo độ cao. Các nghiên cứu cho thấy sử dụng hàm phân bố loga mặc dù đơn giản nhưng phù hợp với tính toán tốc độ gió trong khí quyển từ mặt đến độ cao khoảng 150 m. Vì vậy, nghiên cứu này lựa chọn hàm loga để tính toán thông qua một công cụ chuyển đổi trong mô hình RCM: Về mặt lý thuyết, khi biết tốc độ gió V_1 ở độ cao Z_1 có thể tính được tốc độ gió V_z ở độ cao Z theo công thức sau [20]:

$$\frac{V_z}{V_1} = \frac{\ln(z/z_0)}{\ln(z_1/z_0)} \quad (1)$$

Suy ra: $V_z = V_1 \frac{\ln(z/z_0)}{\ln(z_1/z_0)}$

Trong đó V_z là tốc độ gió ở độ cao cần tính Z (ví dụ tại 100 m, 150 m), V_1 là tốc độ gió quan trắc bề mặt, Z_0 là độ gồ ghề của mặt đệm, mức Z_1 là độ cao của thiết bị đo gió ($Z_1 = 10$ m).

Do độ cao cần tính thường lớn hơn độ cao đo gió mặt đất ($Z > Z_1$) nên $V_z > V_1$ hay tốc độ gió tăng theo độ cao. Ngoài ra, mức độ tăng lên của tốc độ gió theo độ cao phụ thuộc vào độ gồ ghề của mặt đệm (Z_0). Khi độ gồ ghề của mặt đệm càng lớn thì tốc độ gió ở độ cao cần tính (V_z) càng tăng nhanh.

2.3.3. Phương pháp ước lượng năng lượng gió qua mật độ không khí và tốc độ gió

Mật độ năng lượng gió trung bình E (W/m^2) tại một nơi trong thời gian T nào đó (năm, mùa, tháng ...) được tính theo công thức:

$$E = \frac{1}{2} \rho \frac{1}{N} \sum_{i=1}^N v_i^3 \quad (2)$$

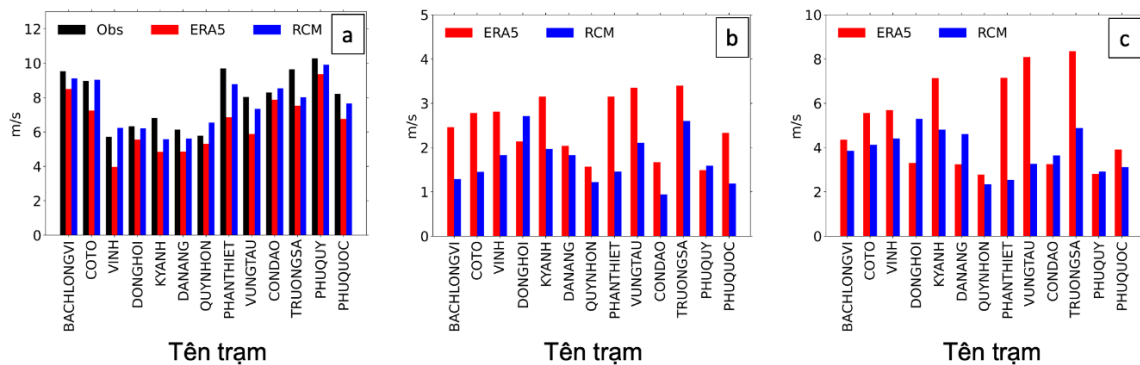
Trong đó ρ là mật độ không khí được giả định là hằng số $1,225 \text{ kg/m}^3$, v_i là tốc độ gió tức thời (m/s), N là dung lượng mẫu.

Như vậy, năng lượng gió tiềm năng là đại lượng được tính toán từ tốc độ gió tuyệt đối mà không xem xét đến hướng gió, do vậy, đặc trưng về tốc độ gió là cơ sở để đánh giá tiềm năng của gió trên lãnh thổ.

3. Kết quả và thảo luận

3.1. Đánh giá khả năng mô phỏng tốc độ gió cực đại ngày

Kết quả so sánh giá trị mô phỏng tốc độ gió cực đại ngày (V_{\max}) trung bình của RCM và ERA5 so với trạm quan trắc tại 13 trạm khí tượng ven biển được thể hiện trên biểu đồ hình 3a. Một điểm cần chú ý khi các trạm được lựa chọn đều là các trạm nằm tại khu vực ven biển, tốc độ gió ít chịu ảnh hưởng bởi địa hình hơn so với các trạm khác. Biểu đồ trên so sánh độ giá trị trung bình của 20 năm (2000-2019) của bộ dữ liệu tái phân tích ERA5 và dữ liệu mô hình khí hậu khu vực RCM trích xuất và so với dữ liệu quan trắc tại điểm trạm (Obs). Kết quả cho thấy, tốc độ gió cực đại ngày trung bình ở các trạm phía Nam lớn hơn phía Bắc, ngoại trừ tại hai trạm đảo Bạch Long Vĩ và Cô Tô với trung bình dao động từ 9,1 đến 9,7 m/s. Trong khi đó các trạm dọc khu vực biển miền Trung từ Vinh đến Quy Nhơn có tốc độ gió cực đại ngày trung bình không quá lớn, phổ biến trong khoảng từ 5,9 đến 7,4 m/s. Ở khu vực Phan Thiết trở vào khu vực biển phía Nam, tốc độ gió cực đại ngày trung bình khá lớn, phổ biến trong khoảng 8,2 đến 10,4 m/s. Phần lớn đây đều là những trạm đảo, ít bị ảnh hưởng bởi địa hình và có mặt thoáng của biển, do vậy cường độ gió ổn định và thường lớn hơn so với các trạm ven biển ở đất liền. RCM mô phỏng tốc độ gió cực đại ngày trung bình tốt, và nắm bắt được sự biến thiên giữa các trạm, bao gồm cả các trạm đảo và các trạm ven biển. Trong khi đó, ERA5 có xu hướng thiên thấp hơn so với quan trắc, đặc biệt tại một số trạm như Vinh, Phan Thiết, Vũng Tàu. Như vậy có thể thấy RCM có thể mô phỏng giá trị gió cực đại trung bình tốt hơn so với ERA5, đặc biệt là ở các trạm ven biển.



Hình 3. So sánh các chỉ số thống kê vận tốc gió V_{\max} trung bình (a) và sai số MAE (b) và RMSE (c) của RCM và ERA5 so với giá trị quan trắc tại 13 trạm khí tượng.

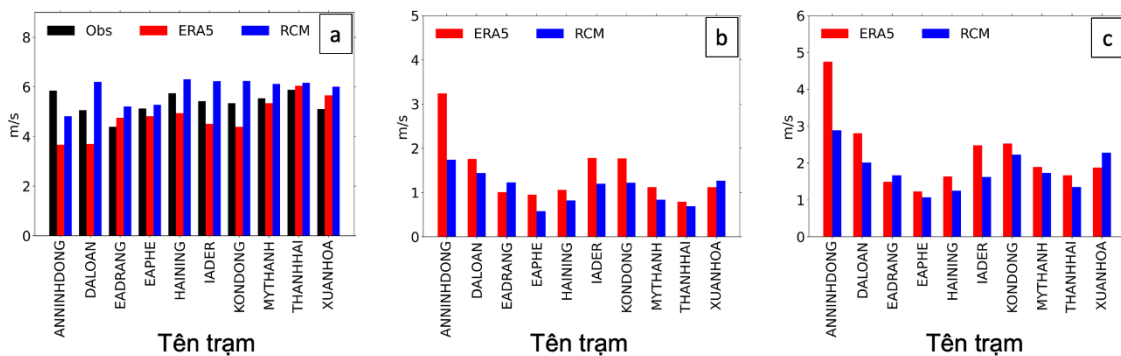
Biểu đồ so sánh sai số trung bình tuyệt đối (MAE) và sai số quân phương (RMSE) của vận tốc gió cực đại trung bình ngày từ hai nguồn dữ liệu tái phân tích ERA5 và dữ liệu xây dựng từ mô hình RCM so với giá trị quan trắc tại 13 trạm khí tượng ven biển được thể hiện trên hình 3b và 3c. Kết quả cho thấy hầu hết các giá trị sai số của RCM thấp hơn so với sai số từ dữ liệu tái phân tích ERA5. Sự khác biệt của sai số trung bình tuyệt đối của dữ liệu tái phân tích ERA5 lớn nhất so với dữ liệu từ mô hình RCM tại trạm Phan Thiết khoảng 2 m/s. Tại một số ít trạm như trạm Phú Quý, dữ liệu RCM có chênh lệch sai số trung bình tuyệt đối

cao hơn kết quả dữ liệu ERA5 khoảng 0,1 m/s hay trạm Đồng Hới cao hơn xấp xỉ 0,5 m/s. Ngoài ra, biểu đồ cũng thống kê sai số trung bình tuyệt đối lớn nhất MAE của dữ liệu tái phân tích ERA5 nằm tại trạm Trường Sa với xấp xỉ 3,5 m/s, trong khi đó sai số nhỏ nhất của dữ liệu mô hình RCM nằm tại trạm Côn Đảo với xấp xỉ 1 m/s.

Phần lớn kết quả sai số quân phương ghi nhận các giá trị V_{max} từ dữ liệu ERA5 cao hơn so với các giá trị được xây dựng từ mô hình RCM, đặc biệt chênh lệch lớn nhất tại các trạm Phan Thiết, Vũng Tàu, Trường Sa từ 4-4,5 m/s. Tuy nhiên cũng có một số trạm có sai số quân phương V_{max} từ dữ liệu mô hình RCM lớn hơn dữ liệu tái phân tích ERA5 như các trạm Đồng Hới, Đà Nẵng chênh lệch xấp xỉ 3 m/s, trạm Phú Quý, Côn Đảo chênh lệch từ 0,1-0,5 m/s. Ngoài ra, có thể thấy trong biểu đồ, sai số V_{max} lớn nhất từ dữ liệu tái phân tích ERA5 nằm tại trạm Trường Sa là xấp xỉ 8,5 m/s và sai số nhỏ nhất từ dữ liệu mô hình RCM ở tại trạm Quy Nhơn xấp xỉ 2,5 m/s.

3.2. Đánh giá khả năng mô phỏng tốc độ gió trung bình mực 80 m

Biểu đồ so sánh độ chính xác của vận tốc gió trung bình (V_{tb}) trong hai bộ dữ liệu tái phân tích ERA5 và dữ liệu xây dựng từ mô hình RCM với dữ liệu quan trắc Obs của 10 trạm đo mực cao thuộc dự án GIZ được thể hiện trên hình 4. Có thể thấy, có sự khác biệt giữa kết quả tốc độ gió trung bình từ dữ liệu tái phân tích ERA5 và mô hình RCM với dữ liệu quan trắc. Một số trạm có kết quả dữ liệu vận tốc gió trung bình ERA5 thấp hơn đáng kể so với dữ liệu quan trắc như trạm An Ninh Đông, thấp hơn xấp xỉ 2 m/s; trạm Đà Loan thấp hơn xấp xỉ 1,5 m/s; trạm IaDer, Kon Dong thấp hơn xấp xỉ 1 m/s,... Bên cạnh đó, cũng có một số trạm dữ liệu từ RCM lớn hơn dữ liệu quan trắc như các trạm: Đà Loan cao hơn xấp xỉ 3 m/s; trạm IaDer, Kon Dong cao hơn xấp xỉ 2 m/s; trạm Hải Ninh cao hơn xấp xỉ 1,5 m/s, trạm My Thanh cao hơn xấp xỉ 1 m/s. Ngoài ra, tại một số trạm như EaDrang, trạm Xuân Hòa, trạm Thanh Hải, kết quả vận tốc gió trung bình từ cả hai nguồn dữ liệu ERA5 và RCM đều ghi nhận có sai số nhỏ khi so với kết quả quan trắc. Có thể thấy, so với số liệu quan trắc, hầu hết các kết quả vận tốc gió trung bình từ mô hình RCM đều có độ chênh lệch ít hơn kết quả của dữ liệu ERA5.

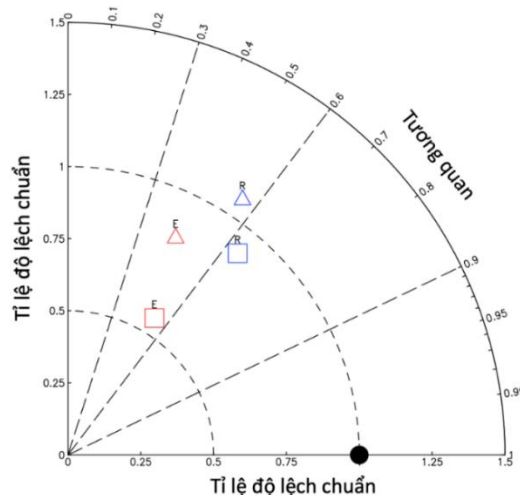


Hình 4. So sánh các chỉ số thống kê vận tốc gió trung bình (V_{tb}) (a), sai số MAE (b) và sai số RMSE (c) của RCM và ERA5 với quan trắc tại 10 trạm đo gió mực cao 80 m.

Sai số V_{tb} của RCM đều thấp hơn ERA5 từ 0,5-1 m/s, đặc biệt tại trạm An Ninh Đông, chênh lệch lên tới xấp xỉ 2m/s. Bên cạnh đó cũng có một số trạm có sai số V_{tb} từ dữ liệu mô hình RCM cao hơn so với dữ liệu ERA5 như trạm EaDrang cao hơn xấp xỉ 0,3 m/s, trạm Xuân Hòa cao hơn 0,1 m/s. Ngoài ra, trong biểu đồ cũng thấy, sai số trung bình tuyệt đối MAE lớn nhất của dữ liệu tái phân tích ERA5 ở trạm An Ninh Đông với xấp xỉ 3,5 m/s, trong khi đó sai số nhỏ nhất của dữ liệu mô hình RCM tại trạm EaPhe với xấp xỉ 0,75 m/s. Sự chênh lệch đáng kể giữa sai số quân phương RMSE của hai bộ dữ liệu tái phân tích ERA5 và dữ liệu mô hình RCM được thể hiện trên biểu đồ hình 4c. Với sai số quân phương, tại trạm An Ninh Đông, dữ liệu ERA5 có sai số lớn nhất, xấp xỉ 5 m/s và có sự chênh lệch lớn nhất

so với dữ liệu mô hình RCM (xấp xỉ 2 m/s). Đa số tại các trạm, kết quả sai số từ dữ liệu tái phân tích ERA5 đều lớn hơn dữ liệu mô hình RCM. Tuy nhiên bên cạnh đó, tại một số trạm như EaDrang, Xuân Hòa, dữ liệu từ mô hình RCM có sai số lớn hơn dữ liệu tái phân tích ERA5 với chênh lệch nhỏ từ 0,1-0,5 m/s. Ngoài ra, sai số của mô hình RCM nhỏ nhất ở tại trạm EaPhe với xấp xỉ 1,5 m/s.

Như vậy, tùy thuộc trạm kết quả mô phỏng RCM và ERA5 lại cho sai số với quan trắc là khác nhau, do đó, để đánh giá được toàn diện khả năng sử dụng kết quả mô phỏng gió từ mô hình RCM so với sử dụng trực tiếp dữ liệu tái phân tích ERA5, biểu đồ Taylor được sử dụng với tập dữ liệu được nội suy về 13 trạm đo gió 10 m (hình tam giác) và 10 trạm đo gió mực cao 80 m (hình vuông). Các điểm thông tin về trạm và điểm đo đã được cung cấp chi tiết trong bảng 1 và hình 1. Kết quả so sánh từng cặp giữa RCM và ERA5 cho thấy trong cả 2 trường hợp RCM đều cho kết quả mô phỏng gió tốt hơn đáng kể so với ERA5. Độ lệch chuẩn của ERA5 trong cả 2 trường hợp so sánh gió tại mực 10 m và 80 m đều nhỏ hơn so với giá trị quan trắc. Trong khi đó sử dụng RCM với độ phân giải cao hơn đã nắm bắt được tính biến động của số liệu gió tốt hơn trong cả 2 trường hợp với hệ số tương quan cao hơn và độ lệch chuẩn xấp xỉ so với giá trị quan trắc. Ngoài ra, kết quả đánh giá cũng cho thấy ERA5 và RCM cho mô phỏng gió trung bình tốt hơn so với gió V_{max} (hình vuông tốt hơn hình tam giác tại Hình 5). Điều này có thể lý giải do cách mô phỏng liên tục hóa trường gió từ mô hình chưa thể hiện hết được các nhiễu động ngẫu nhiên trong khí quyển. Do vậy, kết quả mô phỏng chưa thể nắm bắt chính xác được các giá trị gió cực đại.



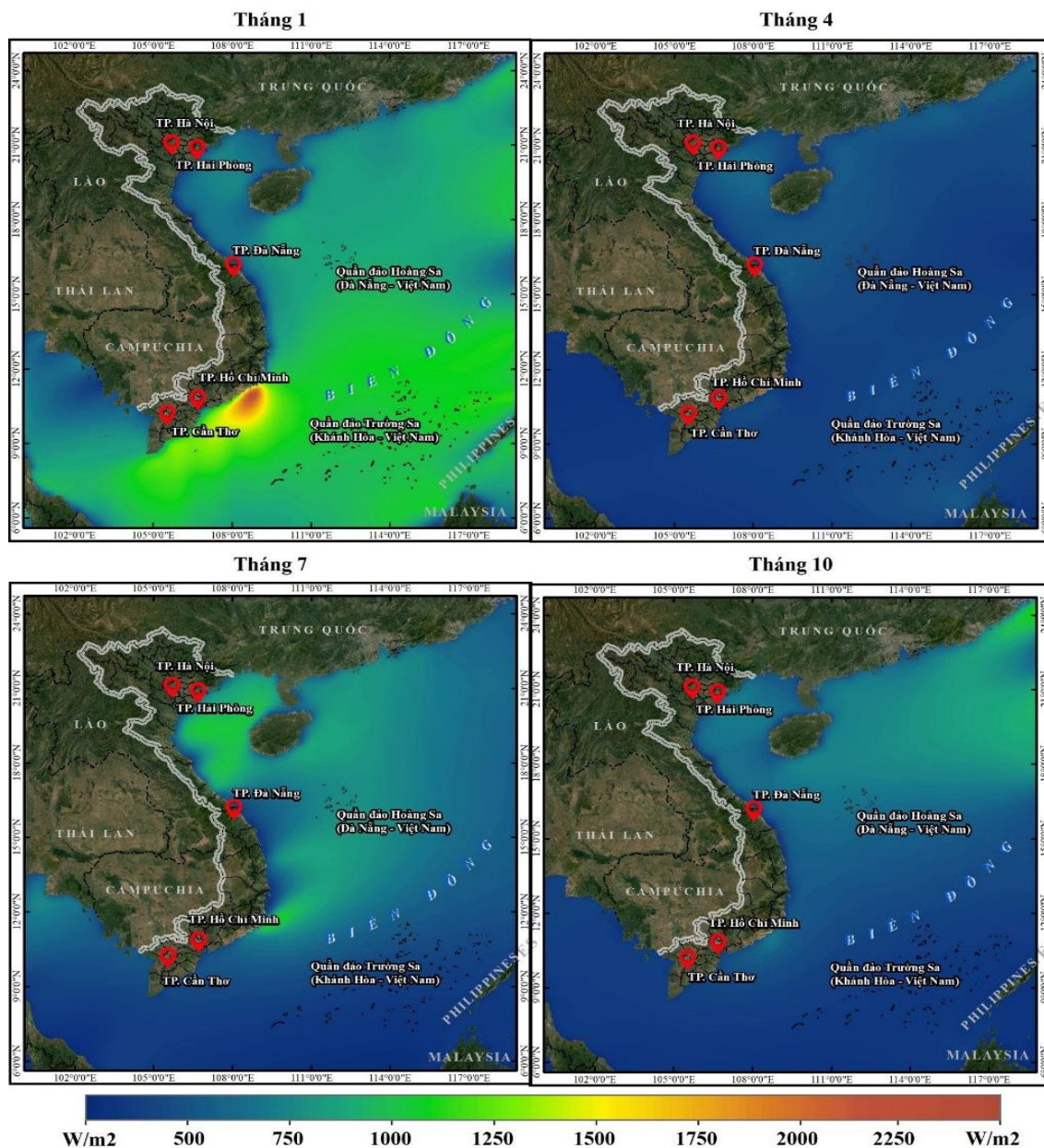
Hình 5. Biểu đồ Taylor so sánh số liệu gió mô phỏng từ mô hình RCM (màu xanh) và ERA5 (màu đỏ) với số liệu quan trắc, Hình tam giác tính cho tập giá trị gió V_{max} mực 10 m, còn hình vuông tính toán cho tập giá trị gió trung bình mực 80 m.

Từ phân tích các chỉ số thống kê cũng như phân tích từ biểu đồ Taylor cho thấy mặc dù chưa mô phỏng được hoàn toàn chính xác tốc độ gió trung bình và tốc độ gió cực đại ngày tại tất cả các điểm đo, tuy nhiên RCM cho thấy độ chính xác cao hơn trong cả hai trường hợp tại mực 10 m và mực độ cao 80 m so với số liệu ERA5. Do đó với mục tiêu sử dụng một bộ số liệu tốc độ gió xây dựng bản đồ tiềm năng năng lượng gió có độ phân giải cao, nghiên cứu này đề xuất sử dụng kết quả mô phỏng từ mô hình RCM.

3.3. Phân tích đánh giá tiềm năng năng lượng gió

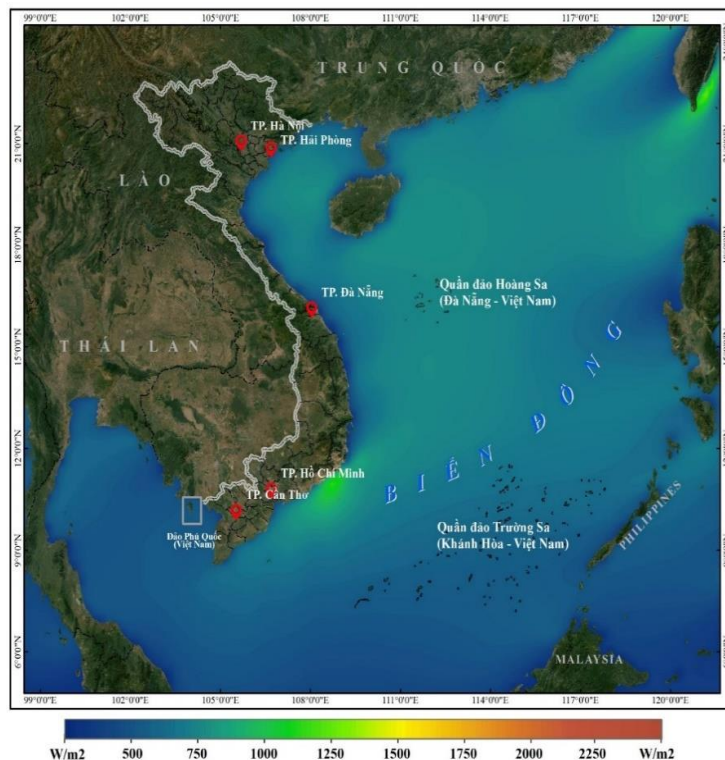
Tốc độ gió mô phỏng từ mô hình RCM được sử dụng để tính toán tiềm năng năng lượng gió sử dụng phương pháp như đã được trình bày trong mục 3.2.3. Về quy mô mùa, tiềm năng năng lượng gió trên toàn Biển Đông ở độ cao 100 m lớn nhất vào tháng 1 và thấp nhất vào tháng 4 (Hình 6), đây cũng là kết quả có sự tương đồng khi so sánh với kết quả đánh giá tiềm

năng lượng gió được công bố trong “*Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam*” [8] của Bộ TM&MT. Tuy nhiên so với kết quả được công bố của Bộ TM&MT, các giá trị được mô phỏng bởi mô hình RCM có xu hướng thiên cao hơn. Cụ thể vào tháng 1, năng lượng gió trên toàn Biển Đông được mô phỏng bởi RCM có mật độ năng lượng gió trung bình từ 600-1800 W/m², trong khi kết quả được công bố trong “*Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam*” là 300-1300 W/m², bên cạnh đó, cũng trong báo cáo này, mật độ năng lượng gió trung bình vào tháng 4 được công bố từ dưới 100-500 W/m² và thiên thấp hơn so với kết quả mô phỏng trong tháng 4 của mô hình RCM là dưới 200-600 W/m². Ở các vùng biển phía Bắc (Vịnh Bắc Bộ, Quảng Trị - Quảng Ngãi và Bắc Biển Đông), mật độ năng lượng gió lớn nhất trong các tháng 1, 7 với các giá trị lần lượt là 300-600 W/m² và 800-1200 W/m², thấp nhất trong các tháng 4, 10 với giá trị lần lượt dưới 650 W/m² và dưới 750 W/m². Tương tự tại các vùng biển ven bờ phía Nam (Bình Định trở vào, Giữa và Nam Biển Đông), mật độ năng lượng gió lớn nhất trong các tháng chính đông và thấp hơn trong các tháng chính hè, mật độ năng lượng gió trung bình các tháng 1, 4, 7, 10 lần lượt là: từ 1200-1800 W/m²; 300-600 W/m²; 400-1200 W/m² và 400-600 W/m².



Hình 6. Bản đồ năng lượng gió tại độ cao 100 m từ mô hình RCM các tháng chính mùa (tháng 1, 4, 7, 10).

Kết quả tính toán năng lượng gió trung bình năm tại độ cao 100 m của mô hình RCM với độ phân giải 10 km được thể hiện trên hình 7, so sánh kết quả tính này với tiềm năng năng lượng gió của WB cho thấy: Phân bố không gian tiềm năng năng lượng gió sử dụng mô phỏng RCM khá tương đồng với kết quả của WB, về cơ bản các vùng ven biển có tiềm năng năng lượng gió lớn nhất là Bắc vịnh Bắc Bộ, Ninh Thuận đến Bà Rịa - Vũng Tàu, trong đó đáng chú ý nhất là vùng biển Ninh Thuận - Bình Thuận. Tuy nhiên về mặt trị số, kết quả mô phỏng bằng RCM thiên cao hơn tại một số vị trí so với WB. Trong kết quả công bố mới nhất của WB (2021), khu vực Bình Thuận có mật độ năng lượng gió trung bình từ 400-500W/m² thì kết quả tính toán sử dụng RCM lại có mật độ năng lượng gió trung bình trên 1000 W/m². Bên cạnh đó, tại phía Bắc vịnh Bắc Bộ, mật độ năng lượng gió tính toán từ mô hình RCM trung bình từ 500-750 W/m², trong khi theo kết quả công bố của WB, tại khu vực này chỉ từ 300-500 W/m².



Hình 7. Bản đồ năng lượng gió trung bình năm tại độ cao 80 m từ mô hình RCM.

4. Kết luận và kiến nghị

Trong nghiên cứu này đã sử dụng mô hình khí hậu khu vực RCM để mô phỏng trường gió trên khu vực Biển Đông cho giai đoạn 20 năm, từ 2000 đến 2019. Kết quả mô phỏng được so sánh với số liệu tái phân tích ERA5 và dữ liệu quan trắc tại 13 trạm khí tượng mực tiêu chuẩn 10 m cũng như dữ liệu gió tại độ cao 80 m từ dự án đo gió của GIZ. Trường gió mô phỏng bởi RCM sau đó được sử dụng để tính toán năng lượng gió tiềm năng cho khu vực Biển Đông tại nhiều mực độ cao, trong đó kết quả tại mực 100 m được phân tích và so sánh với một số dữ liệu tương tự. Một số kết quả chính của nghiên cứu được tóm tắt như sau: (1) Tốc độ gió mô phỏng bởi RCM có độ chính xác cao hơn với sai số MAE và RMSE đều nhỏ hơn so với sai số từ ERA5 ở phần lớn các trạm, cả ở mực 10 m và mực 80 m. Khi so sánh tương quan và độ lệch chuẩn, mô phỏng từ RCM cũng thể hiện tốt hơn đáng kể so với ERA5. (2) Có sự tương đồng về phân bố không gian năng lượng gió trên Biển Đông giữa kết quả sử dụng trường gió từ mô hình RCM với dữ liệu từ “*Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam*” [8] của Bộ TM&MT cũng như dữ liệu từ “*Global Wind Atlas*” tài trợ bởi WB. Về mặt giá trị, năng lượng gió tính toán từ mô phỏng trường gió

bởi RCM nhìn chung thiên cao hơn so với kết quả trong “*Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam*” của Bộ TM&MT.

Mặc dù vậy, nghiên cứu này vẫn tồn tại một số hạn chế như: (1) Các tính toán từ tốc độ gió sang năng lượng gió mới chỉ sử dụng công thức lý thuyết mà chưa xét đến những giả thiết như ảnh hưởng chi tiết mật độ và tương tác với các công trình...; (2) Số liệu để so sánh, đánh giá còn thiếu và chủ yếu ở khu vực đất liền, ven biển. Tuy vẫn còn một số tồn tại như đã đề cập ở trên, song kết quả của nghiên cứu này cũng là một nguồn tham khảo đối với phát triển quy hoạch và xây dựng dự án điện gió trên biển trong tương lai.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: T.T.L., N.V.H.; Lựa chọn phương pháp nghiên cứu: P.D.H.B., N.V.H., P.Q.N.; Xử lý số liệu: N.V.H., P.Q.N., N.B.T.; Viết bản thảo bài báo: N.V.H., T.T.L.; Chỉnh sửa bài báo: T.T.L.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội trong đề tài mã số TN.22.17.

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. Dincer, I. Renewable energy and sustainable development: a crucial review. *Renew. Sustain. Energy Rev.* **2000**, 2, 157–175. [https://doi.org/10.1016/S1364-0321\(99\)00011-8](https://doi.org/10.1016/S1364-0321(99)00011-8).
2. Tâm, T.M. Tìm hiểu về năng lượng tái tạo. *Tạp chí khoa học Đại học Văn Lang* **2021**, 61–66.
3. Archer, C.L.; Jacobson, M.Z. Evaluation of global wind power. *J. Geophys. Res.* **2005**, 110, 1–20. <https://doi.org/10.1029/2004JD005462>.
4. IPCC. Climate Change 2022_ Impacts, Adaptation and Vulnerability_ Working Group II contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 2022. doi: 10.1017/9781009325844.
5. Nguyễn, Đ.V.; Nhiên, N.T.T. Vì sao nên phát triển điện gió ngoài khơi ở Việt Nam. *Tạp chí Khoa học Công nghệ Việt Nam* **2020**, 11, 20–22.
6. De Palacio, C.R.; Santos, A.C.; Perez, J.; Diez, D. Offshore wind energy: A review of the current status, challenges and future development in Spain. *Renew. Sustain. Energy Rev.* **2016**, 64, 1–18. <https://doi.org/10.1016/j.rser.2016.05.087>.
7. Yamaguchi, A.; Ishihara, T. Assessment of offshore wind energy potential using mesoscale model and geographic information system. *Renew. Energy* **2014**, 69, 506–515. <https://doi.org/10.1016/j.renene.2014.02.024>.
8. Tổng cục Khí tượng Thủy văn, Bộ Tài nguyên và Môi trường. Báo cáo tiềm năng năng lượng gió, sóng ngoài khơi tại các vùng biển Việt Nam, Hà Nội, 2022, tr. 105.
9. Viện Năng Lượng. Quy hoạch phát triển Điện Lực Quốc gia thời kỳ 2021–2030 tầm nhìn đến 2045. Hà Nội, 2021, tr. 851.
10. AWS. Truepower, Wind Resource Atlas of Vietnam. 2011, pp. 17.
11. Tiềm năng phát triển năng lượng gió biển ở Việt Nam. 2017. Trực tuyến: <https://tietkiemnangluong.com.vn/tin-tuc/kinh-nghiem/t26058/tiem-nang-phat-trien-nang-luong-gio-bien-o-viet-nam.html>.
12. Duc, T.N.; Tangang, F.T.; Santisirisomboon, J.; Cruz, F.; Tuan, L.T.; Nguyen-Xuan, T., Aldrian, E. Performance evaluation of RegCM4 in simulating extreme rainfall and temperature indices over the CORDEX-Southeast Asia region. *Int. J. Climatol.* **2017**, 37(3), 1634–1647.
13. Bộ Công Thương, Bộ Môi trường, Bảo tồn Thiên nhiên và An toàn Hạt nhân Liên bang, Cộng Hòa Liên bang Đức. Chương trình Hỗ trợ Năng lượng GIZ. 2021, tr. 24.

14. Ruiz, S.G.; Eduardo, J.; Barriga, C. Wind power assessment in the Caribbean region of Colombia, using ten-minute wind observations and ERA5 data. *Energy Rev.* **2021**, *172*, 158–176. <https://doi.org/10.1016/j.renene.2021.03.033>.
15. Wedi, N.P.; Bauer, P.; Deconinck, W.; Diamantakis, M.; Hamrud, M.; Kühnlein, C.; Malardel, S.; Mogensen, K.; Mozdzyński, G.; Smolarkiewicz, P.K. The modelling infrastructure of the Integrated Forecasting System: Recent advances and future challenges. Tech. Memorandum 2015, pp. 50.
16. Wu, J.; Han, Z.Y.; Yan, Y.P.; Sun, C.Y.; Xu, Y.; Shi, Y. Future changes in wind energy potential over China using RegCM4 under RCP emission scenarios. *Adv. Clim. Chang. Res.* **2021**, *12*, 596–610. <https://doi.org/10.1016/j.accre.2021.06.005>.
17. Koné, B.; Diedhiou, A.; Touré, N.D.E.; Sylla, M.B.; Giorgi, F.; Anquetin, S.; Bamba, A. Diawara. Sensitivity study of the regional climate model RegCM4 to different convective schemes over West Africa. *ESD* **2018**, *9*, 1261–1278. <http://dx.doi.org/10.5194/esd-2018-44>.
18. Gu, H.; Wang, G. Assessing future climate changes and extreme indicators in east and south Asia using the RegCM4 regional climate model. *Clim. Chang.* **2012**, *3317*, 1–114. <https://doi.org/10.1007/s10584-012-0411-y>
19. Bộ Tài nguyên và Môi trường. Kịch bản Biến đổi Khí hậu, nước biển dâng cho Việt Nam. Nhà xuất bản Tài nguyên - Môi trường và Bản đồ Việt Nam, Hà Nội, 2020.
20. Hagspiel, S.; Antonis, P.; Schmid, M.; Andersson, G. Copula - based modeling of stochastic wind power in Europe and implications for the Swiss power grid. *Appl. Energy* **2012**, *96*, 33–44. <https://doi.org/10.1016/j.apenergy.2011.10.039>.

Assessment of wind energy potential for renewable energy development in the East Sea, Vietnam

Nguyen Viet Hang¹, Pham Duy Huy Binh¹, Pham Quang Nam¹, Nguyen Bach Tung¹, Trinh Tuan Long^{1*}

¹ Center for Environmental Fluid Dynamics, University of Science, Vietnam National University, Hanoi; viethang1485@gmail.com; phambinh@hus.edu.vn; nampq@hus.edu.vn; bachtung_cefd@hus.edu.vn; trinhtuanlong@hus.edu.vn

Abstract: In this study, the Regional Climate Model version 4 (RCM) is used to simulate the wind field, allowing for the determination of potential wind energy in the Vietnam East Sea region during the 20-year period (2000-2019). The wind speed is simulated using the RCM, then compared with the reanalysis data (ERA5) using wind observations at 13 coastal and island meteorological stations, at 10 meters and 10 wind measurement stations at 80 meters. The comparison between RCM simulation and ERA5 to wind speed observation shows that the RCM model performs well and is more suitable than using ERA5 data. The spatial distribution of calculated wind energy results from the RCM also exhibits similarities to the data published in the “Potential Wind Energy and Offshore Wave Reports over Vietnam Coastal Areas” by the Ministry of Natural Resources and Environment and the World Bank’s 2009 report. However, the study still has limitations, such as the use of theoretical formulas and not considering boundary layer effects. Nevertheless, the study's results provide valuable reference information for offshore wind energy project planning and development.

Keywords: Wind Energy; Vietnam East Sea; RegCM4.

Bài báo khoa học

Kết quả bước đầu thử nghiệm thuật toán XGBoost dự báo nước dâng do bão tại trạm Hòn Dấu

Bùi Mạnh Hà¹, Nguyễn Bá Thủy^{1*}, Phạm Khánh Ngọc¹, Phạm Văn Tiến²

¹ Trung tâm Dự báo khí tượng thủy văn quốc gia; manhamhc@gmail.com;
thuybanguyen@gmail.com; ngocpkchibo@gmail.com

² Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu; phamvantien@gmail.com

*Tác giả liên hệ: thuybanguyen@gmail.com; Tel.: +84-975853471

Ban Biên tập nhận bài: 8/10/2022; Ngày phản biện xong: 1/11/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Trong nghiên cứu này, thuật toán tăng cường độ dốc cấp cao XGBoost (Extreme Gradient Boosting, sau đây gọi là mô hình XGBoost) được ứng dụng để xây dựng công cụ dự báo nước dâng do bão tại Hòn Dấu. Mô hình XGBoost được xây dựng với 4 phương án sử dụng dữ liệu khác nhau (04 mô hình): mô hình XGBoost đơn biến, mô hình XGBoost đa biến I, mô hình XGBoost đa biến II và mô hình XGBoost sử dụng dữ liệu chéo. Bộ dữ liệu trong 28 cơn bão ảnh hưởng tới trạm Hòn Dấu giai đoạn 2002-2021 được thu thập để xây dựng các mô hình và kiểm định kết quả dự báo. Kết quả thử nghiệm mô hình XGBoost dự báo nước dâng do bão cho thấy, mô hình XGBoos đơn biến cho độ tin cậy thấp ở tất cả các thời hạn dự báo. Trong khi đó, hai mô hình XGBoos đa biến và mô hình sử dụng dữ liệu chéo đều cho kết quả tin cậy cao, với phần lớn hệ số tương quan giữa dự báo và quan trắc đều trên 80%. Kết quả của nghiên cứu làm cơ sở lựa chọn công cụ dự báo nước dâng do bão tại Hòn Dấu tùy thuộc vào hiện trạng số liệu quan trắc khí tượng, hải văn.

Từ khóa: Dự báo nước dâng do bão; XGBoost; Machine Learning; AI.

1. Mở đầu

Thông thường, các mô hình số như POM [1–2], ROMS [3–4], ADCIRC [5–6], SuWAT [7–10], Delft3D [11–12] được sử dụng để dự báo nước dâng do bão có độ tin cậy cao do sử dụng các phương trình vật lý chính xác, nhưng thường đi kèm với đó là cần tài nguyên tính toán lớn, thời gian vận hành khá đáng kể. Với ưu điểm là linh hoạt và mạnh mẽ, có khả năng xác định các mối quan hệ phức tạp trong dung lượng lớn dữ liệu đầu vào nên trí tuệ nhân tạo (AI: Artificial Intelligence) sẽ phù hợp khi áp dụng để triển khai học máy chuyên sâu, phân tích dữ liệu đa yếu tố nhanh và chính xác. Dự báo nước dâng do bão theo hướng sử dụng phương pháp học máy đã được nhiều nhà khoa học trên thế giới nghiên cứu và phát triển mạnh mẽ trong thời gian gần đây. Mạng nơ-ron nhân tạo (ANN: Artificial Neural Networks) đã được sử dụng phổ biến trong dự báo độ cao nước dâng do bão [13–19]. Với nguyên lý kết hợp các mô hình học tập có sai số cao thành một cây học tập mạnh hơn theo kiểu tuân tự nhằm mục đích xử lý bài toán học máy có giám sát với độ tin cậy cao mà mô hình XGBoost được ứng dụng nhiều trong dự báo liên quan đến lĩnh vực khí tượng thủy văn, quản lý rủi ro thiên tai, trong đó có dự báo nước dâng do bão.

Trong nghiên cứu [20–21] đã ứng dụng mô hình học máy XGBoost để dự báo độ cao nước dâng lớn nhất cho một số khu vực ven biển Hoa Kỳ, kết quả nghiên cứu cho thấy mô hình XGBoost dự báo nước dâng bão có độ tin cậy tương đương mô hình ADCIRC. Nghiên cứu [22] sử dụng mô hình XGBoost dự báo nước dâng do bão khu vực Phúc Kiến và Quảng

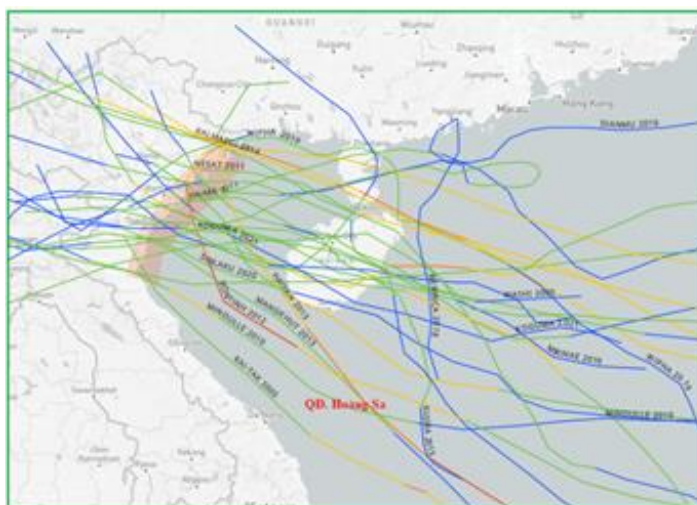
Đông, Trung Quốc cho thấy mô hình XGBoost dự báo tốt hơn so với mô hình BPNN (*Backpropagation Neural Network*) và SVR (*Support Vector Regression*). Nghiên cứu [23] sử dụng bốn mô hình XGBoost, MLR (*Multiple Linear Regression*), SVR, RF (*Random Forest*) để dự tính sóng leo trên bãi biển dốc cho thấy mô hình XGBoost có hiệu suất dự báo vượt trội so với ba mô hình còn lại là MLR, SVR và RF. Nhóm nghiên cứu [24] đã sử dụng mô hình GN-VSIDM là sự kết hợp của mô hình GN (*Gaussian Noise*) với mô hình VSIDM (*Diffusion Model based on the Vibrating String*) và mô hình XGBoost để ước tính thiệt hại do nước dâng bão cho thấy mô hình tổ hợp GN-VSIDM-XGBoost là mô hình ước tính thiệt hại do nước dâng bão tối ưu. Chỉ số RMSE và R^2 đánh giá mô hình GN -VSIDM- XGBoost lần lượt là 0,1089 và 0,8292. Nghiên cứu [25] đã sử dụng mô hình XGBoost để dự báo mực nước ngầm tại Selangor, Malaysia cho thấy rằng mô hình XGBoost có kết quả dự báo tốt hơn so với mô hình ANN và SVR.

Việt Nam là khu vực thường xuyên chịu ảnh hưởng của ATNĐ, bão và tác động kèm theo là nước dâng do bão. Do vậy, dự báo chính xác độ cao và thời điểm xuất hiện nước dâng do bão là rất quan trọng nhằm giảm thiểu rủi ro do nước dâng bão đến cơ sở hạ tầng và thiệt hại về người. Công nghệ dự báo nước dâng do bão bằng mô hình số hiện tại đang phát triển rất mạnh, có thể nói chung là gần đến giới hạn về thuật toán mô hình. Ưu điểm của mô hình số dự báo nước dâng do bão đã được khẳng định nhưng để dự báo chi tiết thì mô hình số đòi hỏi nguồn tài nguyên tính toán lớn, điều này làm cho việc nghiên cứu áp dụng mô hình số bị giới hạn khi cần dự báo chi tiết trên phạm vi không gian rộng. Để khắc phục các vấn đề về tài nguyên tính toán khi dự báo bằng mô hình số cũng như sự hạn chế về số liệu quan trắc mực nước tại các trạm khí tượng hải văn ven biển hiện nay tại Việt Nam thì nghiên cứu và xây dựng mô hình học máy dự báo nước dâng do bão là hướng tiếp cận hợp lý. Trong nghiên cứu này, nghiên cứu sử dụng mô hình XGBoost với các dữ liệu quan trắc khác nhau để dự báo độ cao nước dâng do bão tại Hòn Dấu, đây là trạm có số liệu quan trắc các yếu tố khí tượng, hải văn đầy đủ và dài nhất so với các trạm khí tượng hải văn trên khác cả nước. Các giá trị dự báo của mô hình sau đó sẽ được so sánh với các giá trị thực đo để đánh giá kỹ năng của mô hình XGBoost trong nghiệp vụ dự báo nước dâng do bão. Tính mới của nghiên cứu này

2. Số liệu và phương pháp nghiên cứu

2.1. Số liệu phục vụ xây dựng mô hình XGBoost

Bộ dữ liệu đầu vào để xây dựng mô hình XGBoost dự báo nước dâng tại Hòn Dấu gồm gió, khí áp, mực nước quan trắc tại trạm Hòn Dấu, Hòn Ngư, Sơn Trà, các tham số bão trong các cơn bão ảnh hưởng trực tiếp đến trạm Hòn Dấu gồm vị trí tâm bão, khí áp tại tâm bão, vận tốc gió cực đại, và hướng di chuyển tại các bước thời gian 06 giờ. Bộ dữ liệu quan trắc và các tham số bão trong 20 năm gần đây (2002-2021) được thu thập, bao gồm 28 cơn bão có khả năng gây nước dâng tại Hòn Dấu. Quỹ đạo các cơn bão được thu thập thể hiện trên hình 1. Trong đó, 80% dữ liệu dành cho huấn luyện mô hình (training) và 20% dữ liệu dùng để kiểm định mô hình (testing). Dữ liệu trong 02 cơn bão



Hình 1. Quỹ đạo các cơn bão trong 20 năm (2002-2021) với số liệu được sử dụng xây dựng mô hình dự báo nước dâng tại trạm Hòn Dấu.

Dữ liệu trong 02 cơn bão

Wutip (năm 2013) và Doksuri (2017) được sử dụng để kiểm định mô hình, đây là 2 cơn bão mạnh, gây nước dâng lớn đáng kể (lớn hơn 0,5m) tại Hòn Dấu. Bộ dữ liệu trong thời gian 02 cơn bão này ảnh hưởng sẽ không tham gia vào huấn luyện mô hình để đảm bảo tính khách quan.

Mô hình XGBoost được thiết kế dự báo nước dâng do bão tại Hòn Dấu với 04 phương án như trên bảng 1. Các siêu tham số chính để thực hiện hiệu chỉnh mô hình XGBoost trong quá trình huấn luyện gồm: learning_rate, max_depth, l1_reg, l2_reg, subsample, gamma, min_child_weight, n_estimators và early_stopping. Phạm vi điều chỉnh bộ siêu tham số này được thể hiện trong bảng 2. Để điều chỉnh bộ siêu tham số này, nghiên cứu lựa chọn thư viện Optunn nhằm hỗ trợ tự động điều chỉnh tham số mô hình để mô hình có thể đạt được hiệu năng tốt nhất, kết quả thể hiện trong bảng 3.

Bảng 1. Thiết lập mô hình XGBoost cho trạm Hòn Dấu.

TT	Mô hình	Tham số đầu vào	Tham số dự báo
1	Mô hình XGBoost đơn biến	SL tại Hòn Dấu	SS
2	Mô hình XGBoost đa biến I (06 biến)	LG, LT, CAP, MWS, MV, HWS tại Hòn Dấu	SS
3	Mô hình XGBoost đa biến II (09 biến)	SL, OWS, OP, LG, LT, CAP, MWS, MV, HWS tại Hòn Dấu	SS
4	Mô hình XGBoost dữ liệu chéo (13 biến)	SL (Hòn Dấu) OWS, OP (Hòn Ngự và Sơn Trà), LG, LT, CAP, MWS, MV, HWS	SS

Trong đó SS là độ cao nước dâng dự báo tại trạm; SL là mực nước quan trắc tại trạm; OWS là vận tốc gió quan trắc tại trạm; OP là áp suất quan trắc tại trạm; CAP là khí áp tại tâm bão; MWS là vận tốc gió cực đại trong bão; LG là kinh độ (tâm bão); LT là vĩ độ (tâm bão); MV là tốc độ di chuyển của bão; HWS là hướng di chuyển của bão.

Bảng 2. Phạm vi điều chỉnh các siêu tham số của mô hình XGBoost.

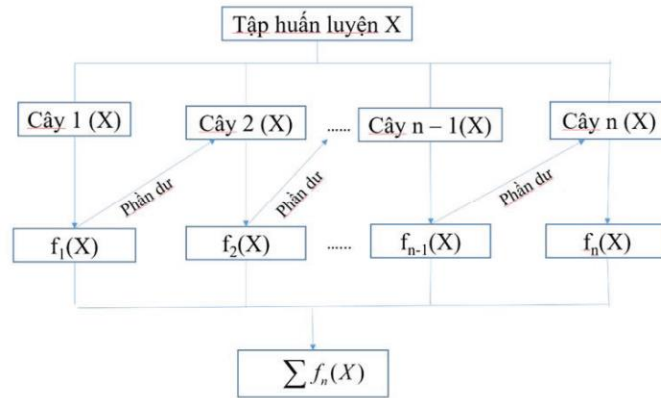
Siêu tham số	Phạm vi điều chỉnh	Siêu tham số	Phạm vi điều chỉnh
learning_rate	[0.00001 ÷ 0.50]	gamma	[1e-8 ÷ 1.0]
max_depth	[1 ÷ 20]	min_child_weight	[1e-8 ÷ 1.0]
l1_reg	[0.00001 ÷ 10.0]	n_estimators	[10 ÷ 1000]
l2_reg	[0.00001 ÷ 10.0]	early_stopping	20
subsample	[0.0 ÷ 1.0]		

Bảng 3. Bộ các siêu tham số tối ưu của mô hình XGBoost tại Hòn Dấu.

Siêu tham số	Mô hình đơn biến	Mô hình đa biến I	Mô hình đa biến II	Sử dụng chéo dữ liệu
learning_rate	0,5	0,4	0,43	0,49
max_depth	18	19	14	16
l1_reg	2,2e-05	4,6e-05	0,4	0,53
l2_reg	4,15	0,32	0,37	2,2e-03
subsample	0,94	0,88	0,91	0,98
gamma	0,00015	0,05	3,99e-05	3,2e-04
min_child_weight	0,00048	0,3e-06	2,36e-06	2,3e-04
n_estimators	644	756	666	558
early_stopping	20	20	20	20

2.2. Mô hình XGBoost

Mô hình XGBoost sử dụng thuật toán tăng cường độ dốc cấp cao, là sự mở rộng của thuật toán Gradient Tree Boosting (GTB) được đề xuất bởi Friedman [26]. Nguyên lý cơ bản được sử dụng trong mô hình XGBoost là việc kết hợp các cây mô hình học tập có độ sai số cao thành một cây mô hình học tập mạnh hơn theo kiểu tuần tự, nghĩa là đào tạo các mô hình mới tốt hơn từ việc kết hợp các mô hình yếu trước đó để bù đắp các thiếu sót trong các mô hình trước, với sơ đồ trên hình 2.



Hình 2. Sơ đồ thuật toán của mô hình XGBoost.

Mô hình XGBoost nhằm mục đích xử lý bài toán học máy có giám sát cho độ tin cậy cao. Với phương pháp học máy chuyên sâu thông thường chỉ nhận nguồn vào là dạng thô, khi đó phải quy đổi sang n-vector trong khoảng trống số thực thì XGBoost nhận nguồn dữ liệu đầu vào là dạng bảng với mọi kích cỡ dữ liệu và dạng tài liệu gồm có cả phân loại. XGBoost có tốc độ huấn luyện nhanh do có thể tính toán song song khi sử dụng tất cả các lõi CPU trong quá trình đào tạo. XGBoost là sự mở rộng của thuật toán GTB tuy nhiên kèm theo đó là những cải tiến để tối ưu thuật toán, bộ nhớ đệm của cấu trúc dữ liệu, sự kết hợp tối ưu giữa phần mềm và phần cứng nên có khả năng ứng dụng với bộ dữ liệu lớn. Thư viện của XGBoost triển khai thuật toán cây quyết định tăng cường độ dốc. Tăng cường độ dốc là một cách tiếp cận trong đó các mô hình mới được tạo ra để dự đoán phần dư hoặc sai sót của các mô hình trước đó rồi cộng lại với nhau để đưa ra dự đoán cuối cùng. XGBoost được gọi là độ dốc tăng cường vì nó sử dụng thuật toán giảm độ dốc để giảm thiểu tổn thất khi thêm mới các mô hình. Cách tiếp cận này hỗ trợ cả các vấn đề về mô hình dự đoán hồi quy và phân loại.

Lý thuyết của thuật toán như sau, giả sử rằng chúng ta có một tập huấn luyện có N mẫu:

$$X = \{X_1, X_2, X_3, \dots, X_n\} \tag{1}$$

Với thông số đầu ra xác định là:

$$Y = \{Y_1, Y_2, Y_3, \dots, Y_n\} \tag{2}$$

Như vậy trong thuật toán XGBoost, tại vòng lặp đầu tiên một cây học tập bất kỳ được tạo ra và ước lượng các giá trị đầu ra $f_1(X)$. Các thông số ước lượng này sẽ có sai khác với giá trị chính xác y một lượng giá trị được gọi là phần dư. Phần dư có thể hiểu là biểu thị cho sự sai số của mô hình. Muốn mô hình học tập tốt thì chúng ta phải giảm giá trị phần dư này đi. Để thực hiện điều này, cây học tập thứ 2 sẽ được thiết lập để ước lượng các giá trị của phần dư đó (không phải là giá trị y). Tương tự như trên, khi ước lượng phần dư $G_1(X)$, cây học tập thứ 2 sẽ ước lượng được giá trị $f_2(X)$ tạo ra phần dư $G_2(X)$. Để ước lượng phần dư $G_2(X)$, ta lại tiếp tục tạo ra cây học tập thứ 3. Quá trình lặp cứ tiếp tục như vậy. Cuối cùng thì giá trị ước lượng sẽ là $\sum f_n(X)$. Để nâng cao hiệu suất làm việc của mô hình XGBoost, hàm mất mát sẽ được thêm vào và có dạng sau:

$$J(\mathcal{F})=L(\mathcal{F})+Q(\mathcal{F}) \tag{3}$$

Trong đó các tham số của mô hình được huấn luyện ký hiệu là \mathcal{L} , L là hàm mất mát; Q là thành phần được thêm vào thường được gọi là regularization nhằm đánh giá mức độ phức tạp của mô hình. Việc thêm vào thành phần regularization giúp làm hài hòa các tham số thu được của mô hình học máy và tránh hiện tượng mô hình trở nên quá khớp (overfitting). Theo kinh nghiệm, việc sử dụng hàm mục tiêu được chuẩn hóa như trong công thức (3) sẽ giúp mô hình được lựa chọn có xu hướng sử dụng các hàm đơn giản và có thể dự đoán. Mô hình càng đơn giản sẽ cho phép tránh hiện tượng quá khớp càng tốt. Do dựa vào mô hình học tập dạng cây, giá trị ước đoán cuối cùng sẽ là:

$$y_i^n = \sum_{t=1}^n f_t(X_i) \quad (4)$$

Hàm mất mát ở vòng lặp thứ t có dạng:

$$J = \sum_{i=1}^n L(y_i, y_i^t) + \sum_{i=1}^n Q(f_n) \quad (5)$$

Giá trị ước lượng đầu ra y_i ở vòng lặp thứ t , y_i^t được tính như sau:

$$y_i^0 = \sum_{t=1}^n f_t(X_i) = y_i^{t-1} + f_t(X_i) \quad (6)$$

Giá trị regularization $Q(f_n)$ có thể sử dụng công thức sau để xác định:

$$Q_n = \gamma^T + \frac{1}{2} \mu \sum_{i=1}^T W_i^2 \quad (7)$$

Trong đó γ là độ phức tạp của các lá trong cây quyết định; T là số lá trong cây quyết định; μ là hệ số phóng đại hàm phạt; W là véc tơ điểm số cho các lá. Khi đó phân tích bậc 2 Taylor sẽ được sử dụng trong hàm mất mát ở thuật toán XGBoost thay thế cho phân tích bậc nhất được sử dụng trong thuật toán GTB.

Giả thiết rằng hàm tối ưu cho quá trình học tập là hàm MSE, khi đó hàm mục tiêu sẽ được viết thành:

$$J_i^0 = \sum_{i=1}^n \left[g_i W_{qi} + \frac{1}{2} (h_i W_{q0}^2) \right] + \gamma^T + \frac{1}{2} \gamma \sum_{i=1}^T W_i^2 \quad (8)$$

Trong công thức (8), các hằng số đã được loại bỏ $q()$ là hàm số dùng để gán dữ liệu cho là tương ứng; g_i và h_i là đạo hàm bậc nhất và bậc hai của hàm mất mát MSE. Hàm mất mát được xác định bằng tổng của các giá trị mất mát cho từng mẫu do mỗi mẫu chỉ tương ứng với 01 lá cho nên hàm mất mát có thể được xác định bằng tổng các giá trị mất mát của từng lá. Do đó, công thức (8) được viết lại như sau:

$$J^{()} \approx \sum_{i=1}^T \left[\left(\sum_{i_{dj}} g_i \right) W_j + \frac{1}{2} \left(\sum_{i_{dj}} h_i + \gamma \right) W_j^2 \right] = \sum_{i=1}^T \left[G_j W_j + \frac{1}{2} (h_j + \gamma) W_j^2 \right] + \gamma^T \quad (9)$$

Khi đó, bài toán tối ưu của hàm mất mát có thể chuyển thành bài toán tìm giá trị nhỏ nhất của hàm bậc 2. Nói cách khác, sau khi phân chia một nút nhất định trong cây ra quyết định, sự thay đổi hiệu suất của mô hình có thể đánh giá dựa trên hàm mất mát. Nếu hiệu suất của mô hình được cải thiện sau khi thực hiện sự phân chia nút này, sự phân chia đó sẽ được chấp nhận, ngược lại việc tách nút sẽ dừng lại.

2.3. Phương pháp đánh giá dự báo

- Sai số quân phương trung bình (Root Mean Square Error - RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2} \quad (10)$$

- Hệ số tương quan Pearson (r):

$$r = \frac{\sum_{i=1}^n (O_i - \bar{O})(F_i - \bar{F})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 (F_i - \bar{F})^2}} \quad (11)$$

Trong đó N là độ dài chuỗi số liệu, F_i và O_i là các biến trong tập dữ liệu ứng với giá trị dự báo và quan trắc ở thời điểm i , \bar{F} là giá trị dự báo trung bình của chuỗi F , \bar{O} là giá trị quan trắc trung bình của chuỗi O .

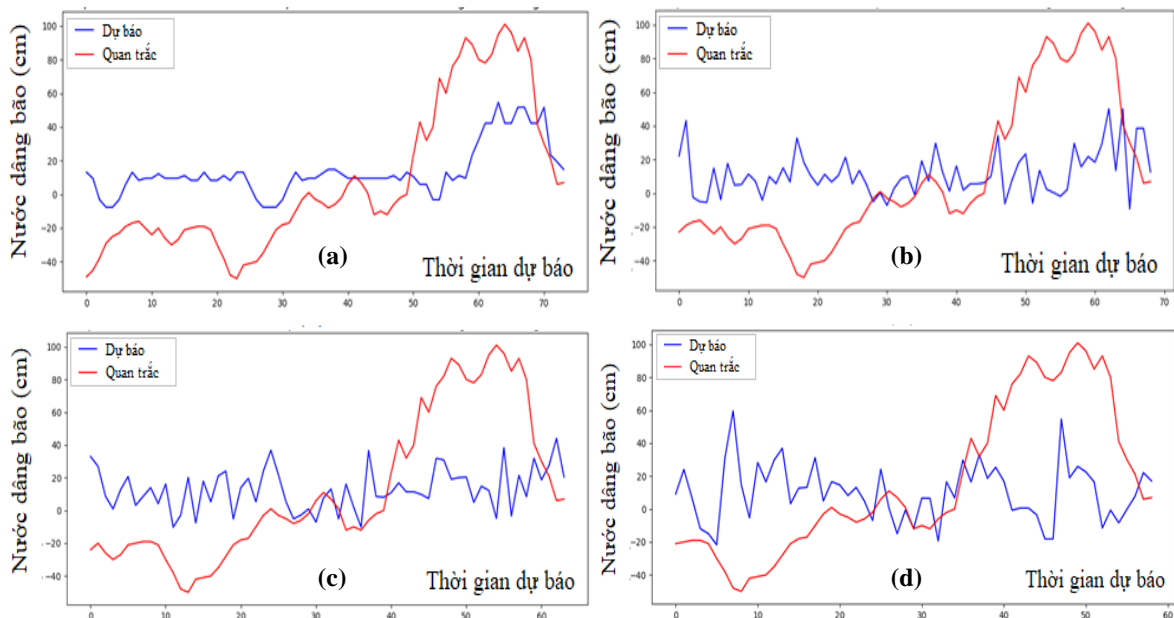
Mô hình XGBoost dự được xây dựng dựa trên dữ liệu dạng chuỗi thời gian tuần tự và liên tục, do đó cần phải được xử lý các giá trị mực nước bị khuyết thiếu. Để lấp đầy các giá

trị khuyết thiếu về mực nước tại trạm, trong nghiên cứu này sử dụng phương pháp nội suy tuyến tính.

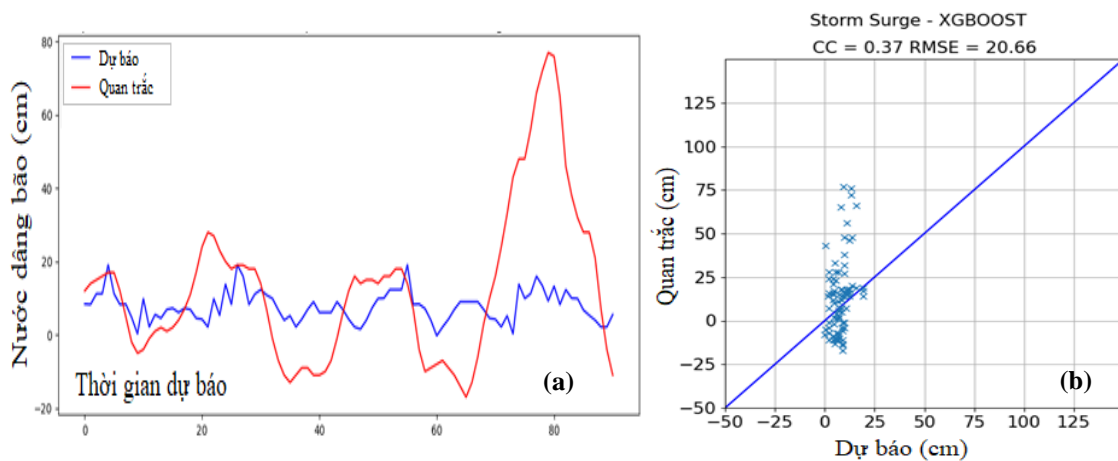
3. Kết quả và thảo luận

3.1. Mô hình XGBoost đơn biến

Đối với mô hình đơn biến (chỉ sử dụng số liệu quan trắc mực nước tại trạm), diễn biến độ cao nước dâng tính toán và thực đo tại Hòn Dấu trong bão Duksuri (9/2014) thể hiện trên hình 3 cho thấy sai số dự báo lớn ở tất cả các thời hạn dự báo, 06, 12, 18 và 24 giờ, với hệ số tương quan R và sai số bình phương trung bình RMSE, tương ứng là 0,67 (39,3 cm), 0,19 (43,1 cm), 0,19 (43,8 cm) và -0,09 (49,8 cm). Kết quả dự báo đối với bão Wutip (9/2013) cũng có sai số lớn ở tất cả các thời hạn dự báo (Hình 4). Như vậy, có thể thấy trường hợp mô hình XGBoost đơn biến chỉ sử dụng độ cao mực nước quan trắc trạm để dự báo nước dâng cho thời hạn 06, 12, 18 và 24 giờ sẽ không cho kết quả tin cậy, do bởi với chỉ độ cao mực nước quan trắc không phản ánh được xu thế nước dâng mà còn cần những yếu tố tương quan khác.



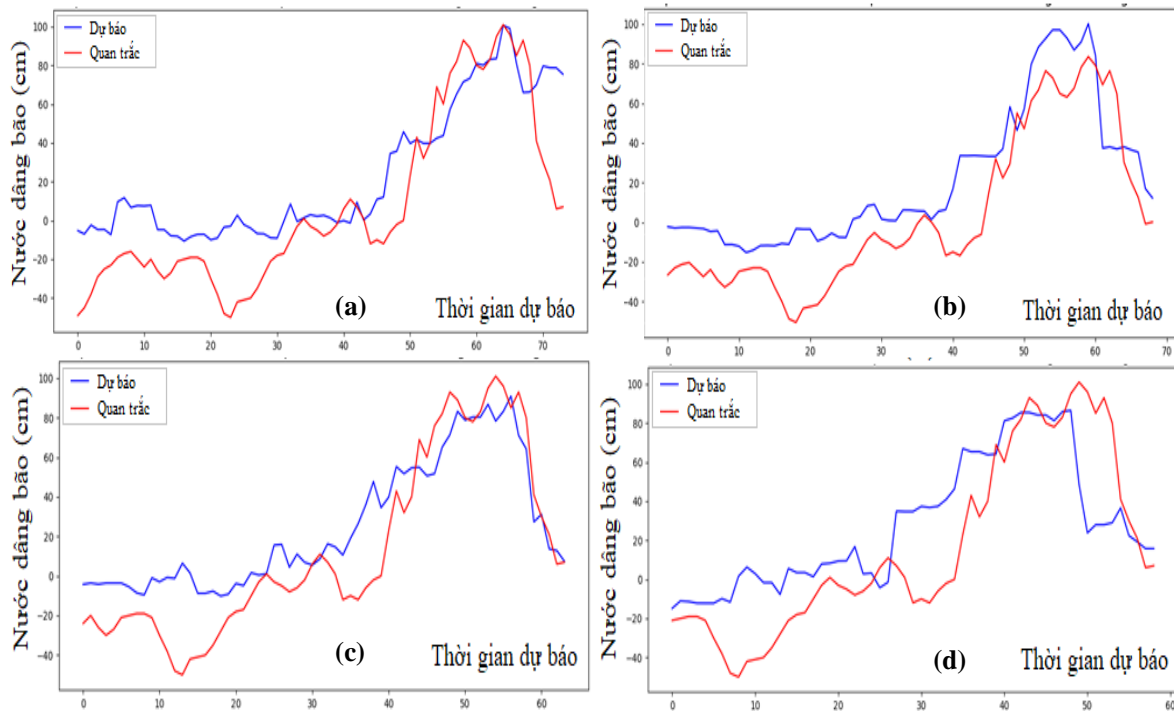
Hình 3. So sánh độ cao nước dâng dự báo bằng mô hình XGBoost đơn biến và quan trắc trong bão Duksuri (9/2017) tại Hòn Dấu tại các thời hạn dự báo: (a) 6 giờ, (b) 12 giờ, (c) 18 giờ, (d) 24 giờ.



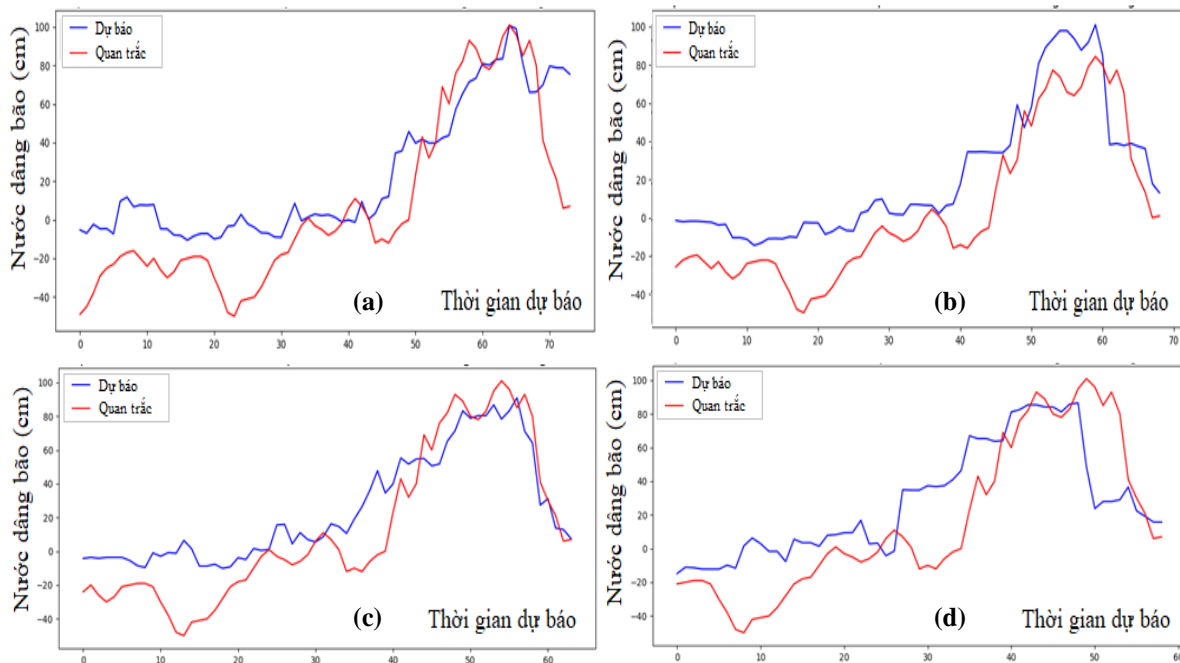
Hình 4. So sánh độ lớn nước dâng dự báo bằng mô hình XGBoost đơn biến và quan trắc tại Hòn Dấu trong bão Wutip (9/2013) thời hạn 6 giờ: (a) Diễn biến độ cao nước dâng; (b) Biểu đồ phân tán nước dâng do bão.

3.2. Mô hình XGBoost đa biến I

Đối với mô hình XGBoost đa biến I, trong trường hợp ngoài số liệu quan trắc mực nước tại trạm Hòn Dấu, các tham số bão (kinh độ và vĩ độ tâm bão; khí áp tại tâm bão; vận tốc gió cực đại; tốc độ di chuyển; hướng di chuyển). Biểu đồ so sánh độ lớn nước dâng dự báo và quan trắc tại Hòn Dấu trong bão Duksuri (9/2014) với các hạn dự báo 06, 12, 18 và 24 giờ trên hình 5 cho thấy kết quả dự báo độ lớn nước dâng đã được cải thiện. Hệ số tương quan R (và chỉ số RMSE) trong trường hợp này với thời hạn dự báo 06, 12, 18 và 24 giờ, tương ứng là 0,88 (27,0 cm), 0,88 (26,7 cm), 0,93 (24,5 cm) và 0,86 (28,6 cm).



Hình 5. So sánh độ cao nước dâng dự báo bằng mô hình XGBoost đa biến I và quan trắc trong bão Duksuri (9/2017) tại Hòn Dấu theo các thời hạn dự báo: (a) 6 giờ, (b) 12 giờ, (c) 18 giờ, (d) 24 giờ.

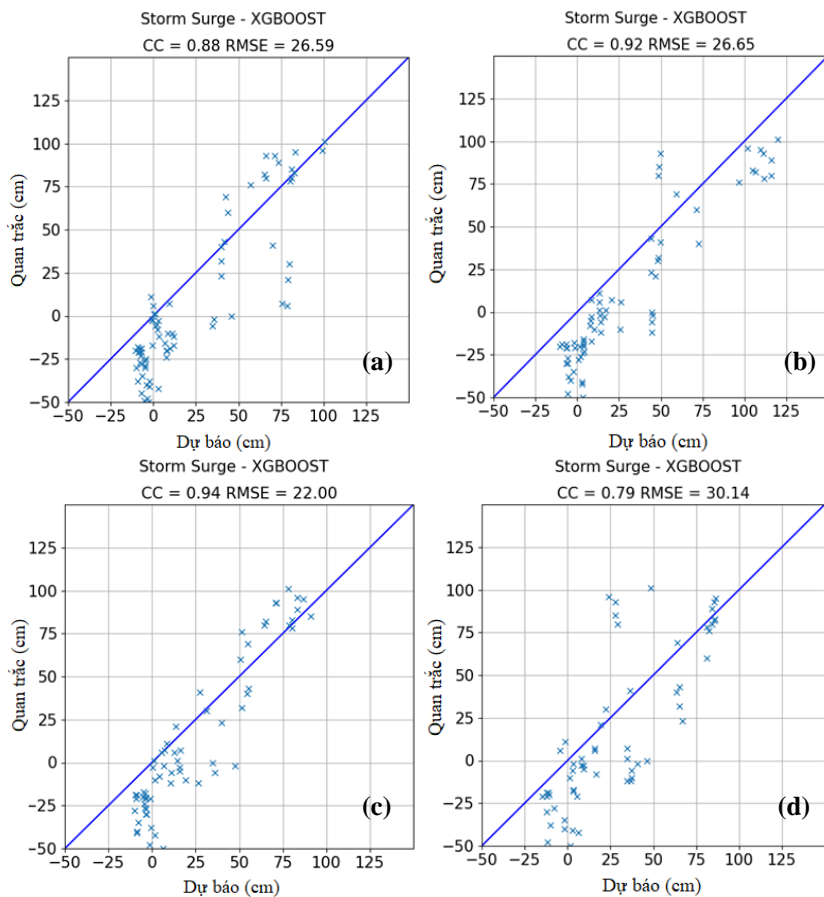


Hình 6. So sánh độ cao nước dâng dự báo bằng mô hình XGBoost đa biến II và quan trắc trong bão Duksuri (9/2017) tại Hòn Dấu theo các thời hạn dự báo, (a) 6 giờ, (b) 12 giờ, (c) 18 giờ, (d) 24 giờ.

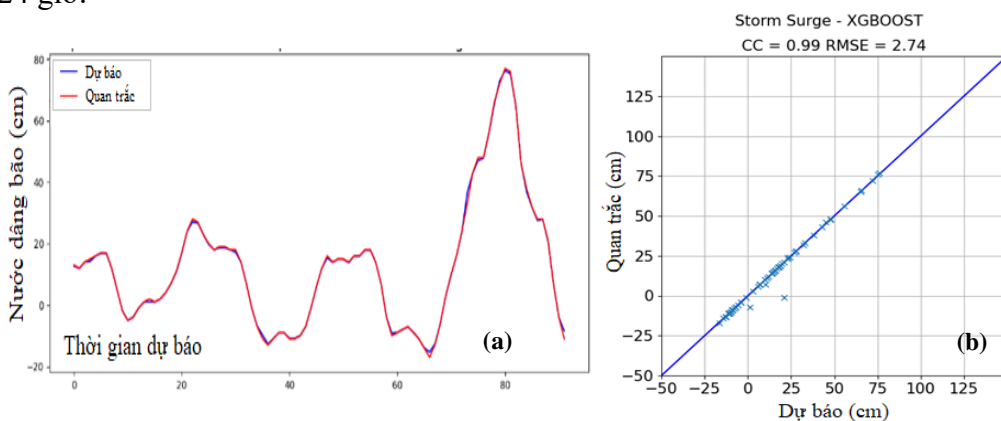
3.3. Mô hình XGBoost đa biến II

Đối với mô hình XGBoost đa biến II, ngoài các tham số được sử dụng ở mô hình XGBoost đa biến I, đã sử dụng thêm số liệu quan trắc tại trạm, gồm mực nước quan trắc, vận tốc gió và khí áp. Kết quả dự báo và so sánh với quan trắc cho các hạn dự báo 06, 12, 18 và 24 giờ thể hiện trên hình 6, biểu đồ phân tán thể hiện trên hình 7. So sánh kết quả dự báo với mô hình XGBoost đa biến I, sai số dự báo tại thời hạn 06 giờ là tương đương, thời hạn 12 và 18 giờ có sai số nhỏ hơn, với hệ số tương quan và RMSE tương ứng là 0,92 (26, 7 cm) và 0,94 (22 cm). Tuy nhiên, sai số dự báo thời hạn 24 giờ lớn hơn kết quả của mô hình XGBoost đa biến I với $R = 0,79$ và $RMSE = 30,14$.

Trong khi đó, với trường hợp bão Wutip (9/2013), kết quả so sánh giữa dự báo và quan trắc độ lớn nước dâng tại Hòn Dấu trên hình 8 cho thấy mô hình có độ tin cậy cao cả ở thời hạn 06 và 24 giờ.



Hình 7. Biểu đồ phân tán giữa độ lớn nước dâng dự báo bằng mô hình XGBoost đa biến II và quan trắc tại Hòn Dấu trong bão Doksuri (9/2017) thời hạn: 6 giờ (a), 12 giờ (b), 18 giờ (c) và 24 giờ (d).

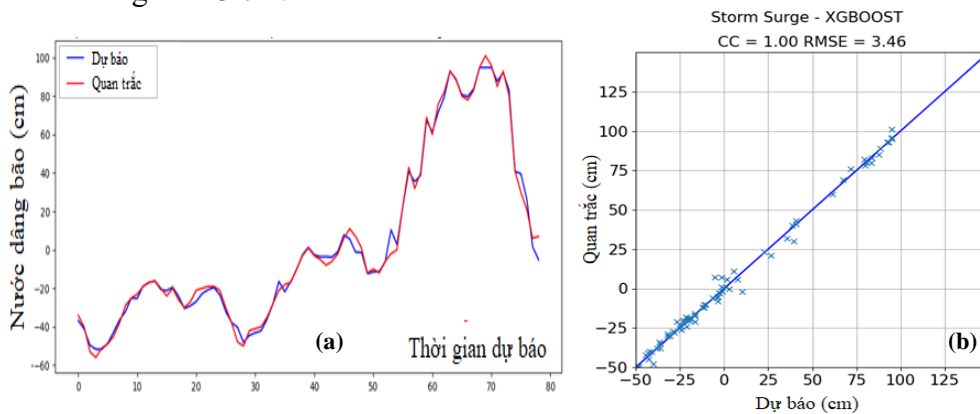


Hình 8. So sánh độ lớn nước dâng dự báo bằng mô hình XGBoost đa biến II và quan trắc tại Hòn Dấu trong bão Wutip (9/2013) thời hạn 24 giờ: (a) Diễn biến độ cao nước dâng, (b) Biểu đồ phân tán nước dâng do bão.

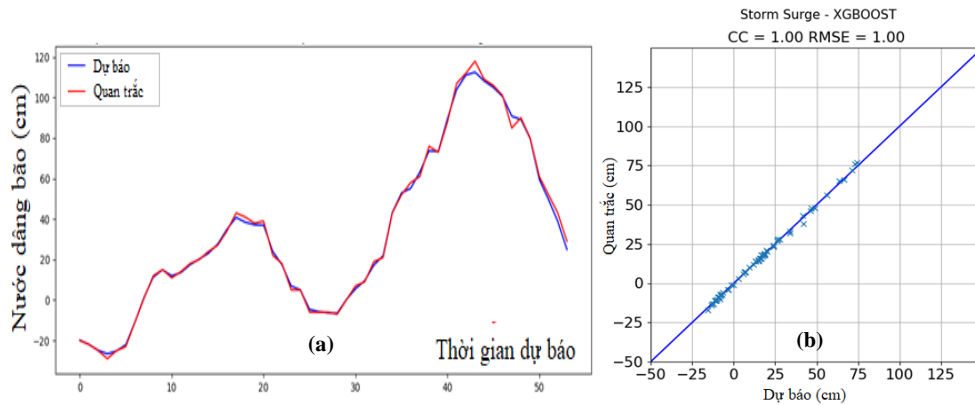
3.4. Mô hình XGBoost sử dụng số liệu chéo

Mô hình XGBoost sử dụng số liệu chéo, bao gồm mực nước quan trắc tại Hòn Dấu, vận tốc gió, khí áp quan trắc tại trạm Hòn Ngự và Sơn Trà và các tham số bão như trường hợp

Mô hình XGBoost đơn biến I và II. Kết quả dự báo độ cao nước dâng thời hạn 24 giờ với bão Duksiri (9/2017) và Wutip (9/2013) thể hiện trên hình 9 và 10 tương ứng. Có thể thấy rằng với mô hình sử dụng số liệu chéo có độ tin cậy với thời hạn dự báo 24 giờ cao, sai số RMSE chỉ khoảng từ 1-3 cm.



Hình 9. So sánh độ lớn nước dâng dự báo bằng Mô hình XGBoost dữ liệu chéo và quan trắc tại Hòn Dấu trong bão Duksiri (9/2017) thời hạn 24 giờ: (a) diễn biến độ cao nước dâng, (b) Biểu đồ phân tán nước dâng do bão.



Hình 10. So sánh độ lớn nước dâng dự báo bằng Mô hình XGBoost dữ liệu chéo và quan trắc tại Hòn Dấu trong bão Wutip (9/2013) thời hạn 24 giờ: (a) diễn biến độ cao nước dâng, (b) Biểu đồ phân tán nước dâng do bão.

Qua kết quả phân tích ở trên cho thấy, ngoài mô hình XGBoost đơn biến, các mô hình XGBoost đa biến I, mô hình XGBoost đa biến II và mô hình XGBoost sử dụng dữ liệu chéo đều có thể ứng dụng để dự báo nước dâng do bão cho thời hạn 06, 12, 18 và 24 giờ, độ tin cậy quả dự báo phần lớn đạt trên 80%, tương đương và cao hơn các mô hình dự báo số trị đang dự báo nghiệp vụ tại Việt Nam. Do vậy, tùy thuộc vào hiện trạng số liệu quan trắc ở thời điểm dự báo, dự báo viên có thể lựa chọn mô hình phù hợp để dự báo nước dâng do bão tại trạm Hòn Dấu.

4. Kết luận

Trong nghiên cứu này, mô hình XGBoost được ứng dụng để xây dựng công cụ dự báo nước dâng do bão tại trạm Hòn Dấu cho các hạn dự báo 06, 12, 18 và 24 giờ. Mô hình XGBoost được thiết lập với 04 phương án sử dụng dữ liệu khác nhau (04 mô hình) với tên gọi: Mô hình XGBoost đơn biến, mô hình XGBoost đa biến I, mô hình XGBoost đa biến II và mô hình XGBoost sử dụng dữ liệu chéo (sử dụng số liệu quan trắc gió, khí áp tại Hòn Ngự và Sơn Trà). Bộ dữ liệu độ cao mực nước quan trắc tại trạm Hòn Dấu, vận tốc gió, khí áp quan trắc tại trạm Hòn Dấu, Hòn Ngự và Sơn Trà, các tham số bão (vị trí tâm bão, khí áp tại tâm bão, vận tốc gió cực đại, tốc độ di chuyển và hướng di chuyển) trong 28 cơn bão ảnh hưởng tới khu vực trạm Hòn Dấu giai đoạn 2002-2021 được thu thập để xây dựng và kiểm

định mô hình XGBoost. Trong đó, khoảng 80% dữ liệu được đưa vào huấn luyện và 20% dữ liệu, tương ứng với số liệu trong 02 cơn bão (Duksuri, tháng 9/2017 và Wutip, tháng 9/2013) dùng để kiểm định các mô hình. Kết quả cho thấy, mô hình XGBoost đơn biến cho độ tin cậy thấp ở tất cả các thời hạn dự báo. Trong khi đó, hai mô hình XGBoost đa biến và mô hình sử dụng dữ liệu chéo đều cho kết quả tin cậy cao, với phần lớn hệ số tương quan giữa dự báo và quan trắc đều trên 80%. Trong đó, mô hình XGBoost sử dụng dữ liệu chéo cho độ tin cậy cao nhất với thời hạn dự báo 24 giờ. Tuy nhiên, các mô hình được xây dựng cần tiếp tục được thử nghiệm thêm với nhiều cơn bão trước khi triển khai ứng dụng trong dự báo nghiệp vụ. Ngoài ra, ứng dụng mô hình XGBoost vào xây dựng công cụ dự báo nước dâng do bão ở các trạm khí tượng hải văn khác cũng cần triển khai thực hiện để nâng cao năng lực dự báo hải văn tại Việt Nam.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: B.M.H., N.B.T.; Thu thập, xử lý số liệu: P.K.N., B.M.H., P.V.T.; Phân tích kết quả: N.B.T., P.K.N., B.M.H.; Viết bản thảo bài báo: B.M.H., N.B.T.; Chỉnh sửa bài báo: B.M.H., N.B.T., P.V.T.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi đề tài nghiên cứu khoa học cấp Bộ Tài nguyên và Môi trường mã số: TNMT.2022.06.04. Tập thể tác giả xin chân thành cảm ơn.

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. Ca, V.T.; Hiều, P.Đ.; Hiền, N.X.; Đạo, N.X. Mô hình dự báo nước dâng do bão có tính đến thủy triều. *Tạp chí Khí tượng Thủy văn* **2008**, 568, 25–33.
2. Ou, S.H.; Liao, J.M.; Tsai, C.Y.; Hsu, T.W. Numerical studies of typhoon-induced storm surge using POM and finite element depth-averaged model in Taiwan. Proceedings 4th Chinese-German Joint Symposium on Hydraulic and Ocean Engineering, Darmstadt, Germany, 2008.
3. Thủy, N.B.; Ngọc, P.K.; Tiến, D.Đ.; Tiến, T.Q.; Hole, L.R.; Kristensen, N.M.; Röhrs, J. Mô hình Roms 2D dự báo nước dâng do bão và gió mùa tại Việt Nam. *Tạp chí Khí tượng Thủy văn* **2016**, 665, 36–40.
4. Qin, G.; Fang, Z.; Zhao, S.; Meng, Y.; Sun, W.; Yang, G.; Wang, L.; Feng, T. Storm Surge Inundation Modulated by Typhoon Intensities and Tracks: Simulations Using the Regional Ocean Modeling System (ROMS). *J. Mar. Sci. Eng.* **2023**, 11, 1112. <https://doi.org/10.3390>.
5. Li, Z.; Li, S.; Hu, P.; Mo, D.; Li, J.; Du, M.; Yan, J.; Hou, Y.; Yin, B. Numerical study of storm surge-induced coastal inundation in Laizhou Bay, China. *Phys. Oceanogr.* **2022**, 9, 1–14. <https://doi.org/10.3389/fmars.2022.952406>
6. Taflanidis, A.A.; Kennedy, A.B.; Westerink, J.J.; Smith, J.; Cheung, K.F.; Hope, M.; Tanaka, S. Rapid assessment of wave and surge risk during landfalling hurricanes: Probabilistic approach. *J. Waterway Port Coastal Ocean Eng.* **2012**, 139(3), 171–182. [https://doi.org/10.1061/\(ASCE\)WW.1943-5460.0000178](https://doi.org/10.1061/(ASCE)WW.1943-5460.0000178).
7. Chiến, Đ.Đ.; Thủy, N.B.; Sáo, N.T.; Thái, T.H.; Kim, S. Nghiên cứu tương tác sóng và nước dâng do bão bằng mô hình số trị. *Tạp chí Khí tượng Thủy văn* **2014**, 647, 19–24.
8. Cường, H.Đ.; Thủy, N.B.; Hưởng, N.V.; Tiến, D.Đ. Đánh giá nguy cơ bão và nước dâng do bão tại ven biển Việt Nam. *Tạp chí Khí tượng Thủy văn* **2018**, 684, 29–36.
9. Thái, T.H.; Trí, Đ.Q.; Hoàng, Đ.V. Nghiên cứu mô phỏng tác động của sóng và nước dâng bão khu vực ven biển miền Trung. *Tạp chí Khí tượng Thủy văn* **2018**, 687, 1–14.

10. Thủy, N.B. Nghiên cứu lựa chọn mô hình dự báo nghiệp vụ nước dâng do bão vào dự báo nghiệp vụ tại Việt Nam. Báo cáo tổng kết đề tài nghiên cứu khoa học cấp Bộ. Hà Nội, 2016.
11. Cát, V.M.; Lân, V.V. Mô phỏng nước dâng do bão và xây dựng bản đồ ngập lụt đảo Phú Quốc. *Tap chí Khoa học Kỹ thuật thủy lợi và môi trường* **2017**, 56, 16–23.
12. Hiền, N.X. Nghiên cứu nước dâng do bão có tính đến ảnh hưởng của sóng và áp dụng cho vùng ven biển Hải Phòng. Luận án tiến sĩ. Hà Nội, 2013.
13. Sztobryn, M. Forecast of storm surge by means of artificial neural network. *J. Sea Res.* **2003**, 49(4), 317–322.
14. You, S.; Seo, J.W. Storm surge prediction using an artificial neural network model and cluster analysis. *Geology Nat. Hazards* **2009**, 53996115.
15. Lee, T.L. 2006. Neural network prediction of a storm surge. *Ocean Eng.* **2006**, 33(3), 483–494.
16. Lee, T.L. Predictions of typhoon storm surge in Taiwan using artificial neural networks. *Ocean Eng.* **2009**, 40(11), 1200–1206.
17. Kim, S.; Matsumi, Y.; Pan, S.; Mase, H. 2016. A real-time forecast model using artificial neural network for after- runner storm surges on the Tottori coast, Japan. *Ocean Eng.* **2016**, 122(6), 44–53.
18. Kim, S.; Pan, S.; Mase, H. Artificial neural network-based storm surge forecast model: Practical application to Sakai Minato, Japan. *J. Ocean Res.* **2019**, 199134715.
19. Chao, W.T.; Young, C.C.; Hsu, T.W.; Liu, W.C. Long-lead-time prediction of storm surge using artificial neural networks and effective typhoon parameters: revisit and deeper insight. *J. Ocean Res.* **2020**, 12(9), 2394.
20. Pacheva, B.; Arorab, P.; del-Castillo-Negrete, C.; Valseth, E.; Dawson, C. A framework for flexible peak storm surge prediction. *Coastal Eng.* **2023**, 1–31.
21. Pacheva, B.; Valseth, E.; Dawson, C. Learning storm surge with gradient boosting. *Ocean Modell.* **2022**, 1–14.
22. Sun, H.; Wang, J.; Ye, W. A Data Augmentation-Based Evaluation System for Regional Direct Economic Losses of Storm Surge Disasters. *Int. J. Environ. Res. Public Health* **2021**, 18, 2918.
23. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, Aug. 2016, pp. 785–794.
24. Du, X.; Li, X.; Zhang, S.; Zhao, T.; Hou, Q.; Jin, X.; Zhan, J. High-accuracy estimation method of typhoon storm surge disaster loss under small sample conditions by information diffusion model coupled with machine learning models. *Int. J. Disaster Risk Reduct.* **2022**, 82, 103307.
25. Osman, A.I.A.; Ahmed, A.N.; Chow, M.F.; Huang, Y.F.; El-Shafie, A. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Eng. J.* **2021**, 12(2), 1545–1556.
26. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Statist.* **2001**, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>.

Initial results of testing the XGBoost algorithm to predict storm surge at Hon Dau station

Bui Manh Ha¹, Nguyen Ba Thuy^{1*}, Pham Khanh Ngoc¹, Pham Van Tien²

¹ National Center for Hydro-Meteorological Forecasting; manhamhc@gmail.com; thuybanguyen@gmail.com; ngocpkchibo@gmail.com

² Institute of Meteorology, Hydrology and Climate Change; phamvantien@gmail.com

Abstract: In this study, the high-level gradient boosting algorithm XGBoost (Extreme Gradient Boosting, hereinafter referred to as the XGBoost model) is applied to build a storm surge forecasting tool at Hon Dau. The XGBoost model is built with 4 different data usage options (04 models): univariate XGBoost model, multivariate XGBoost model I, multivariate XGBoost model II and XGBoost model using cross-sectional data. A data set of 28 storms affecting Hon Dau station in the period 2002-2021 was collected to build models and test forecast results. Test results of the XGBoost model predicting storm surges show that the univariate XGBoos model has low reliability at all forecast horizons. Meanwhile, the two multivariate XGBoos models and the model using cross-sectional data both give highly reliable results, with most correlation coefficients between forecasts and observations above 80%. The results of the study serve as the basis for selecting storm surge forecasting tools at Hon Dau depending on the current status of meteorological and oceanographic monitoring data.

Keywords: Storm surge forecasting; XGBoost; Machine Learning; AI.

Bài báo khoa học

Nghiên cứu lựa chọn công nghệ truyền tin không dây năng lượng thấp (LORA/LPWAN) ứng dụng trong truyền dữ liệu trạm khí tượng tự động thuộc Đài Khí tượng Thủy văn khu vực Tây Nguyên

Vũ Văn Sáng^{1*}

¹ Trung tâm Mạng lưới khí tượng thủy văn quốc gia; sangvu2210@gmail.com

*Tác giả liên hệ: sangvu2210@gmail.com; Tel.: +84-796188866

Ban Biên tập nhận bài: 8/10/2023; Ngày phản biện xong: 6/11/2023; Ngày đăng bài: 25/12/2023

Tóm tắt: Hiện nay, công nghệ truyền tin không dây năng lượng thấp (LORA/LPWAN) đang trở nên phổ biến và ngày càng được ứng dụng nhiều trong các lĩnh vực như viễn thông, truyền tin, trao đổi thông tin số liệu, trong đó có số liệu khí tượng thủy văn. Mạng lưới trạm tự động thuộc Đài Khí tượng Thủy văn khu vực Tây nguyên phần lớn được lắp đặt tại những khu vực vùng sâu, vùng xa và giáp danh gần biên giới, mạng thông tin viễn thông sóng yếu, điều kiện thông tin, truyền tin gặp nhiều khó khăn, nên gây ra nhiều thách thức cho công tác điều tra cơ bản, giám sát, cảnh báo và dự báo, phòng chống thiên tai. Chất lượng truyền nhận thông tin nhiều khi có gián đoạn dẫn đến việc xác định trạm hoạt động hay không hoạt động khó chính xác; gây khó khăn trong việc tìm nguyên nhân, đánh giá hoạt động của các trạm. Nghiên cứu tập trung vào việc ứng dụng những ưu điểm và sự phù hợp của công nghệ truyền tin không dây năng lượng thấp (LORA/LPWAN) nhằm nâng cao hiệu quả truyền phát số liệu của các trạm khí tượng thủy văn tự động phục vụ công tác điều tra cơ bản khí tượng thủy văn tại khu vực Tây Nguyên.

Từ khóa: Công nghệ truyền tin không dây năng lượng thấp LPWAN; Công nghệ truyền tin bằng công nghệ LORA.

1. Mở đầu

Trên thế giới, việc lựa chọn và ứng dụng công nghệ truyền tin không dây năng lượng thấp theo thời gian thực rất phát triển, một số nghiên cứu nổi bật như: Với nghiên cứu nhóm tác giả [1]: Hệ thống giám sát chất lượng không khí thời gian thực công suất thấp sử dụng LPWAN sử dụng công nghệ LoRa. Nhóm nghiên cứu một hệ thống giám sát chất lượng không khí thời gian thực công suất thấp dựa trên công nghệ truyền thông không dây LoRa. Hệ thống tích hợp truyền được số liệu về môi trường không khí (NO₂, SO₂, O₃, CO, PM₁, PM₁₀ và PM_{2.5}). Tại Ý, nhóm nghiên cứu [2] đã nghiên cứu mô đun thử nghiệm truyền số liệu không dây giữa các trạm thời tiết tự động bằng công nghệ LoRaWAN. Với ưu điểm về chi phí thấp, điện năng thấp rất phù hợp ứng dụng trong các lĩnh vực phòng chống và giảm nhẹ thiên tai bằng cách sử dụng một mạng lưới nhiều trạm quan trắc thời tiết tự động. Trong những năm gần đây, các nhà khoa học trong nước cũng đã nghiên cứu về công nghệ truyền tin không dây. Đối với nghiên cứu của tác giả Trung tâm Công nghệ vi điện tử và tin học, Viện ứng dụng công nghệ: Xây dựng mạng cảm biến không dây sử dụng công nghệ truyền thông LORA cho bài toán giám sát và điều khiển trong nông nghiệp công nghệ cao: kết quả đạt được nhằm mục đích giám sát các thông số liên quan đến sinh trưởng, phát triển của cây

trồng như: Ánh sáng, nhiệt độ, độ ẩm, toàn bộ số liệu này được truyền về trung tâm xử lý, từ các thông số thu thập, có thể điều chỉnh lượng nước tưới, độ ẩm không khí hay cường độ ánh sáng phù hợp theo nhu cầu phát triển của cây trồng [3–5]. Nghiên cứu tiếp cận công nghệ sử dụng khinh khí cầu thả ở tầng bình lưu tích hợp công nghệ trạm thu phát thông tin để giám sát, dẫn đường, tìm kiếm cứu hộ và đo đạc thông số vật lý môi trường tầng khí quyển [6, 7], kết quả đạt được: Chế tạo thành công bóng Pilot có gắn thiết bị truyền tin; thử nghiệm trạm thu phát bằng công nghệ LPWAN cho khoảng cách > 250 km, dữ liệu gửi về theo thời gian thực, 10 lần/1s; đo gió từ mặt đất đến 25-30 km trong mọi điều kiện thời tiết. Với nghiên cứu Hệ thống giám sát thành phố thông minh Bến Tre do Công ty Cổ phần Phát triển Hạ tầng và Đầu tư Công nghệ Châu Long thực hiện: Hệ thống có mạng cảm biến IoT thu thập dữ liệu chất lượng môi trường khí tại các điểm quan trọng trong thành phố để phục vụ quản lý môi trường sinh thái đô thị; đo đạc độ mặn, mực nước ở sông Bến Tre để thiết lập các dự báo điều khiển phân luồng giao thông, ứng phó với thiên tai. Phủ sóng LPWAN hỗ trợ kết nối nền tảng đô thị IoT phục vụ cảm biến môi trường, điều khiển đèn chiếu sáng thông minh, cảm biến và điều hành xử lý ô nhiễm, rác thải và quản lý phương tiện di chuyển dùng chung trong tương lai.

Nghiên cứu về quá trình thiết lập thông số LoRa ứng với các khoảng cách khác nhau trong mô hình IoT sử dụng mạng không dây LoRa, qua đó đánh giá được khả năng hoạt động thực tế cho mạng LoRa cho ứng dụng IoT tầm xa, đặc biệt là thiết lập thông số để đạt được hiệu quả cao khi truyền tin không dây LoRa bị tác động bởi môi trường xung quanh bao gồm các tòa nhà cao tầng, cây cối [8].

Trong bối cảnh các loại hình thiên tai, đặc biệt là các hiện tượng thời tiết nguy hiểm như lũ, mưa lớn, bão,... có xu hướng ra tăng tần suất lẫn cường độ, uy hiếp trực tiếp đến tính mạng và tài sản của nhân dân, việc cung cấp thông tin và lựa chọn phương án truyền tin cảnh báo theo thời gian thực là cần thiết. Do vậy, việc lựa chọn công nghệ LoRa là một giải pháp phù hợp truyền số liệu tự động để khắc phục những hạn chế tại khu vực Tây Nguyên góp phần tăng cường cung cấp thông tin phục vụ dự báo nhằm giảm thiểu thiệt hại tài sản do thiên tai gây ra.

2. Phương pháp nghiên cứu lựa chọn công nghệ và số liệu sử dụng

2.1. Giới thiệu khu vực nghiên cứu

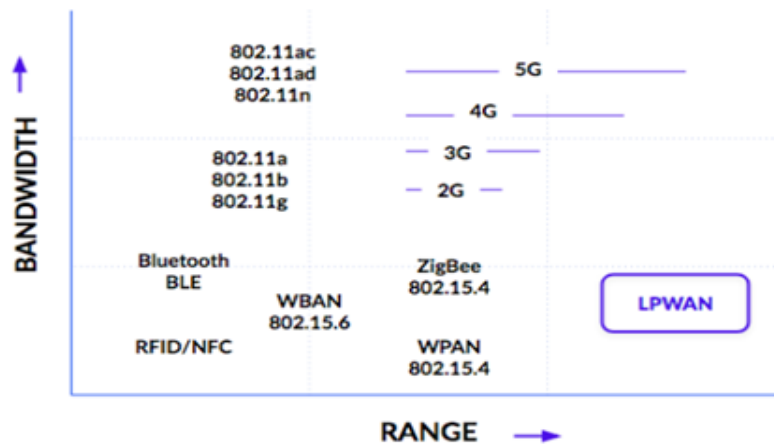
Đặc điểm địa hình ở khu vực Tây Nguyên, khí hậu khu vực Tây Nguyên cũng chia thành ba tiểu vùng khí hậu, gồm Bắc Tây Nguyên (tương ứng với các tỉnh Kon Tum và Gia Lai), Trung Tây Nguyên (tương ứng với các tỉnh Đắk Lắk và Đắk Nông), Nam Tây Nguyên (tương ứng với tỉnh Lâm Đồng). Trung Tây Nguyên có độ cao thấp hơn và nền nhiệt độ cao hơn hai tiểu vùng phía Bắc và Nam. Mạng lưới trạm khí tượng thủy văn thường tập trung ở những vùng núi khó khăn, thượng nguồn các sông suối để phục vụ dự báo, cảnh báo tốt cũng như các trạm đo mưa rất cần lắp đặt trên các vùng có nguy cơ lũ quét, sạt lở đất, vì vậy các khu vực đặt trạm quan trắc khí tượng thủy văn thường không có hạ tầng viễn thông tốt dẫn đến chất lượng đường truyền gặp nhiều khó khăn, đặc biệt khi có hiện tượng thời tiết nguy hiểm đường truyền hay xảy ra mất tín hiệu. Với đặc thù địa hình phức tạp và một số khu vực có hạ tầng viễn thông rất hạn chế nói trên, việc truyền số liệu của nhiều trạm quan trắc trong khu vực Tây Nguyên lại càng gặp nhiều khó khăn hơn nữa.

2.2. Đặc điểm một số công nghệ LPWAN hiện nay

Mạng diện rộng năng lượng thấp LPWAN (*Low Power Wide Area Network*) là các công nghệ không dây, có nhiều công nghệ nhưng có bốn công nghệ chính, gồm: LoRa, Sigfox, NB-IoT và LTE-M.

Internet vạn vật (*Internet of Things- IoT*) gồm nhiều thiết bị kết nối với nhau để thu thập và trao đổi dữ liệu. Các giao thức mạng không dây hiện tại cho thiết bị IoT hiện có như

Bluetooth LE, Wi-Fi, 802.15.4-based Zigbee [9, 10]. Tuy nhiên, các giao thức mạng không dây này bị giới hạn ở việc cung cấp truyền thông khoảng cách rộng và hiệu quả năng lượng. Mạng diện rộng năng lượng thấp LPWAN được sử dụng nhiều hơn do khắc phục được những hạn chế giao thức mạng không dây khác [11].



Hình 1. So sánh LPWAN với các công nghệ khác.

Với công nghệ Sigfox: Cung cấp giải pháp kết nối không dây toàn cầu, nhằm triển khai giải pháp công nghệ LPWAN trong IoT. Cách tiếp cận của Sigfox tương tự như các nhà khai thác mạng di động, nhưng khác biệt là cung cấp các thiết bị tiêu thụ năng lượng thấp nên chi phí thấp. Sigfox sử dụng cấu trúc mạng hình sao và mỗi trạm gốc giao tiếp với đám mây thông qua các liên kết điểm - điểm. Kết nối giữa thiết bị với trạm của Sigfox là vô tuyến. Tuy nhiên, nó cũng có một số nhược điểm là khi gửi dữ liệu trở lại các cảm biến/thiết bị thường bị nhiễu tín hiệu [12].

Với công nghệ NB-IoT (*Narrow Band - IoT*) là một công nghệ IoT băng thông hẹp được chuẩn hóa bởi 3GPP. Được ứng dụng nhiều ở vùng phủ sóng trong nhà; giảm chi phí và tiêu thụ năng lượng thấp hơn, hỗ trợ các tính năng ứng dụng có độ trễ. Công nghệ NB-IoT chủ yếu được sử dụng vào các ứng dụng IoT thông lượng thấp. Kết nối của công nghệ này được đặc trưng bởi thông lượng thấp, truyền dữ liệu không thường xuyên. NB-IoT có thể hoạt động cùng với mạng GSM và LTE dưới dải tần số cần phải đăng ký cấp phép khi sử dụng [13].

Với công nghệ LTE-M (*Long Term Evolution for Machines*). LTE-M sử dụng dải tần số phải được đăng ký cấp phép sử dụng. Các thiết bị của LTE-M được kết nối không dây trực tiếp với mạng 4G-LTE hiện có. Do đó, không cần triển khai các trạm mới, tiết kiệm chi phí thuận tiện cho các nhà mạng. LTE-M tương thích với mạng LTE hiện có; có thể hỗ trợ thực hiện roaming thích hợp các ứng dụng di động. Đặc điểm công nghệ: LTE-M được thiết kế sao cho phù hợp với công nghệ LTE hiện có và có thể hoạt động ở bất kỳ băng tần nào của LTE [14].

Với Công nghệ LoRa: (*Long Range Radio - truyền không dây tầm xa*) là công nghệ lớp vật lý tức là chip được phát triển bởi Semtech. Công nghệ LoRa là một giao thức không dây mới được thiết kế để truyền thông tầm xa, năng lượng thấp. Giao thức phù hợp với các thiết bị thông minh hiện có và liên minh LoRa đang hoạt động đảm bảo khả năng tương tác giữa nhiều mạng trên toàn quốc. Dải tần LoRa sử dụng ít bị nhiễu điện từ, do đó tín hiệu có thể kéo dài một khoảng cách xa, thậm chí đi qua các tòa nhà, với rất ít năng lượng, phù hợp với các thiết bị IoT với dung lượng pin thấp. Do vậy, LoRa thành các thiết bị rẻ hơn rất khả quan. Mỗi trạm gốc (*gateway*) LoRa có thể xử lý hàng triệu node, cùng với thực tế là các tín hiệu có thể kéo dài khoảng cách, nên là cần ít cơ sở hạ tầng mạng hơn, do đó việc xây dựng mạng LoRa rẻ hơn. Các mạng LoRa có thể được đặt cùng với các thiết bị liên lạc khác, như các tháp điện thoại di động, tiết kiệm được chi phí đầu tư xây dựng; đồng thời với các tính năng khác của LoRa sử dụng thuật toán tốc độ dữ liệu thích ứng để giúp tối đa hóa tuổi thọ pin và

tiết kiệm tối đa năng lượng mạng của thiết bị; các giao thức của nó bao gồm nhiều lớp mã hóa, ở cấp độ mạng, ứng dụng và thiết bị, cho phép liên lạc an toàn; tính hai chiều của giao thức hỗ trợ các thông điệp quảng bá, cho phép chức năng cập nhật phần mềm. Do vậy, LoRa là một giải pháp lý tưởng cho IoT [15].

+ Nguyên lý hoạt động của LoRa: LoRa sử dụng kỹ thuật điều chế gọi là Chirp Spread Spectrum (CSS), nguyên lý này là dữ liệu sẽ được chia nhỏ bằng các xung cao tần để tạo ra tín hiệu có dải tần số cao hơn tần số của dữ liệu gốc; sau đó tín hiệu cao tần này tiếp tục được mã hoá theo các chuỗi trước khi truyền ra ăng ten để gửi dữ liệu; nguyên lý này giúp hạn chế độ phức tạp và có độ chính xác cần thiết của mạch thiết bị nhận để có thể giải mã và điều chế lại dữ liệu; hơn nữa LoRa không cần công suất phát lớn mà vẫn có thể truyền xa vì tín hiệu LoRa có thể được nhận ở khoảng cách xa ngay cả độ mạnh tín hiệu thấp hơn cả nhiều môi trường xung quanh; có phạm vi rất rộng tương tự như được cung cấp bởi mạng băng tần hẹp, độ bền cao và chống lại nhiễu và suy hao tín hiệu trong quá trình truyền tín thấp, toàn bộ khu vực hoặc thành phố đó được bảo chi bởi một trạm gốc (*gateway*).

2.3. Lựa chọn giải pháp công nghệ LoRa

a) Sự phù hợp của công nghệ LPWAN

Trước tiên về mặt dữ liệu, theo thực tế vận hành hiện nay ở các trạm thu thập dữ liệu đo đạc tự động ta có các thông số về băng thông yêu cầu truyền dẫn như sau:

+ Trạm khí tượng: Tần suất gửi 10 phút 1 lần.

+ Trạm thủy văn: tần suất gửi 10 phút 1 lần - mực nước, lượng mưa.

+ Đo mưa: tần suất gửi 10 phút 1 lần - lượng mưa.

Dựa trên khảo sát và tần suất gửi tin, có thể thấy rằng yêu cầu băng thông/tốc độ gửi dữ liệu của các trạm khoảng dưới 1 KB trên 1 phút.

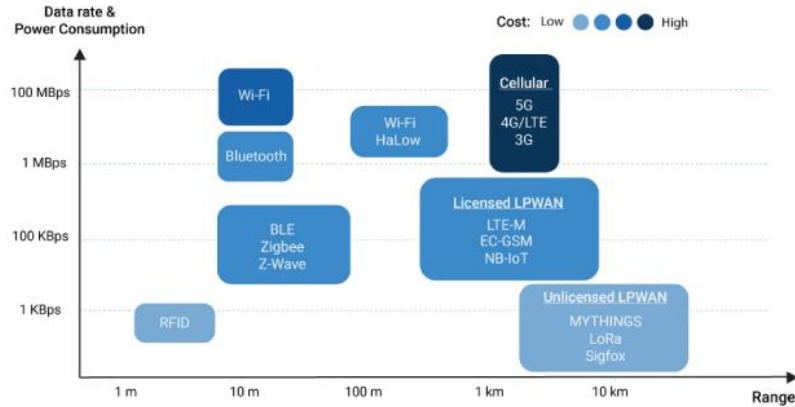
Như đã nêu tại phần trên, hiện nay có nhiều phương án công nghệ để kết nối thông tin cho các mạng không dây cảm biến. Hình 2 bên dưới đây là biểu so sánh băng thông/khoảng cách giữa cảm biến và điểm thu, giá thành của các công nghệ này:

+ Công nghệ LPWAN: rẻ về giá thành, xa về khoảng cách truyền dẫn và tốc độ/băng thông so sánh các công nghệ nối kết WSN dữ liệu có thể đạt đến gấp hơn 10 lần yêu cầu các trạm đo hiện nay, phù hợp để lựa chọn cho công nghệ truyền số liệu trạm tự động của ngành khí tượng thủy văn, đặc biệt là những nơi sóng di động yếu, hệ thống viễn thông hạn chế.

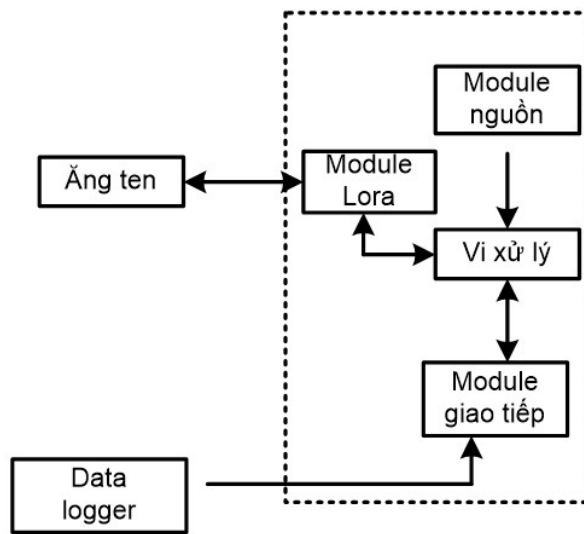
- Việc đầu tư mạng di động hoặc vệ tinh sẽ rất tốn kém và sử dụng quá nhiều năng lượng cho các thiết bị cảm biến tại hiện trường. Tương tự, mạng WiFi và Bluetooth cũng không phải lúc nào cũng là giải pháp tốt hoặc tiết kiệm chi phí, đặc biệt là hạn chế về khoảng cách truy.

- LPWANs (*Low Power Wide Area Networks* - mạng diện rộng năng lượng thấp) là một hiện tượng mới trong IoT, cung cấp giao tiếp tầm xa trên các loại pin nhỏ năng lượng thấp, rẻ tiền có tuổi thọ được duy trì trong nhiều năm, công nghệ này được xây dựng nhằm mục đích hỗ trợ các mạng IoT quy mô lớn trải dài, rộng khắp đặc biệt thích hợp với cảm biến nông/lâm/ngư nghiệp, môi trường và thời tiết [16].

Lựa chọn công nghệ hoạt động trên dải tần không cần xin cấp phép của Cục Tần số (không phải mất phí thuê tần số), tập trung dữ liệu của các trạm quan trắc vùng sâu vùng xa về một điểm chuyển tiếp khu vực lên mạng 3G/4G nối kết trung tâm. Với những ưu điểm của công nghệ LoRa được chỉ ra và thực tế đã có nhiều kinh nghiệm triển khai mạng LPWAN tại Việt Nam, trong đó đặc biệt nghiên cứu có liên quan về lĩnh vực thời tiết, khoảng cách truyền dẫn ở điều kiện nhìn thẳng không che khuất giữa hai ăng ten thu phát LPWAN có thể đạt đến vài chục km với lựa chọn thiết bị thu phát và ăng ten thích hợp. Do vậy, bằng sự so sánh các công nghệ hiện nay về sự phù hợp, có thể thấy LoRa là một giải pháp tiết kiệm chi phí, điện năng và tăng phạm vi phủ sóng ở các khu vực mà công nghệ di động thông thường chưa phổ biến, giải quyết được những bất cập đối với bài toán kết nối, truyền tin quan trắc khí tượng thủy văn.

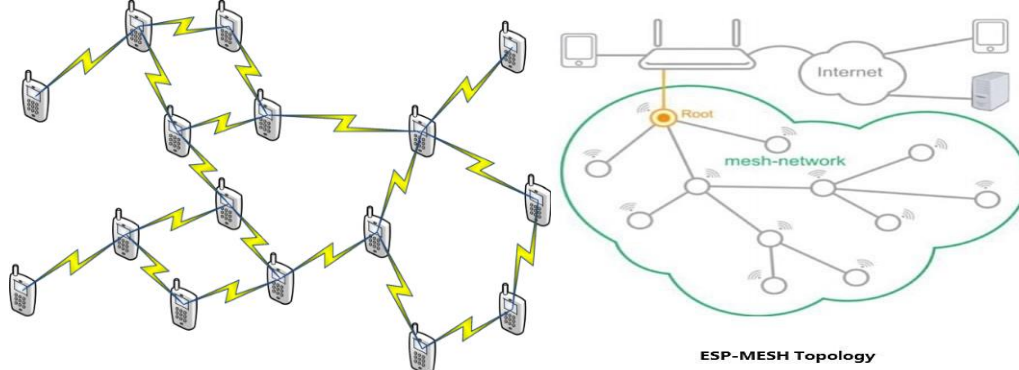


Hình 2. Biểu đồ so sánh các công nghệ nối kết WSN.



Hình 3. Sơ đồ cấu tạo Modem LPWAN.

b) Lựa chọn giải pháp mô hình truyền dữ liệu tính toán trạm thu phát truyền tin
 Mô hình lựa chọn truyền dữ liệu sử dụng công nghệ LORA/LPWAN trong mạng lưới khí tượng thủy văn tại khu vực Tây Nguyên được sử dụng theo mô hình Mesh Network, bởi một số những ưu điểm cụ thể được thể hiện trên hình 4.



Hình 4. Hình minh họa mô hình truyền truyền tin Mesh Network.

Mạng mesh là một cấu trúc liên kết mạng cục bộ trong đó các nút như cầu nối, thiết bị chuyển mạch và các thiết bị cơ sở hạ tầng khác kết nối trực tiếp, động và không phân cấp với càng nhiều nút khác càng tốt và kết hợp với nhau để định tuyến dữ liệu đến nơi sử dụng

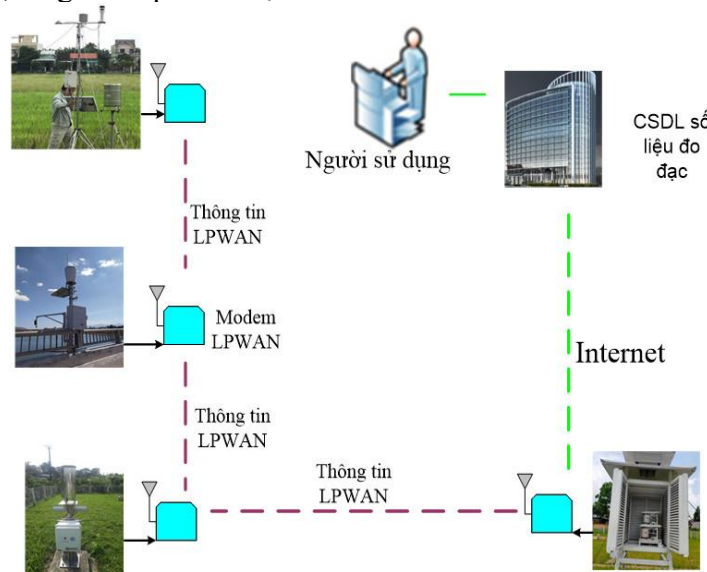
(khách hàng) một cách hiệu quả. Do không phụ thuộc vào nút mạng cụ thể nào, nên sử dụng mô hình mạng này cho phép tất cả các nút mạng có thể tham gia và quá trình truyền hoặc chuyển tiếp dữ liệu [17].

Mạng lưới truyền tin theo mô hình Mesh network có thể chuyển tiếp các dữ liệu, thông tin bằng cách sử dụng kỹ thuật flooding hoặc kỹ thuật routing. Đối với routing, dữ liệu được truyền dọc theo một đường dẫn bằng cách đi từ nút này sang nút khác cho đến khi nó đến điểm cuối. Để đảm bảo tất cả các đường dẫn đều khả dụng, mạng phải cho phép kết nối liên tục và có thể tự cấu hình lại xung quanh các nút, đường dẫn bị hỏng, trong đó sử dụng các thuật toán tự phục hồi như cầu nối đường dẫn ngắn nhất (*Shortest Path Bridging*). Tính năng tự phục hồi cho phép mạng dựa trên định tuyến hoạt động khi một nút bị hỏng hoặc khi kết nối không đáng tin cậy. Do đó, sử dụng mô hình truyền tin này thường đáng tin cậy, vì trong mô hình truyền tin thường có nhiều hơn một đường dẫn giữa nguồn và đích trong mạng [18].

3. Kết quả và thảo luận

3.1. Phương án kết nối thiết bị trạm

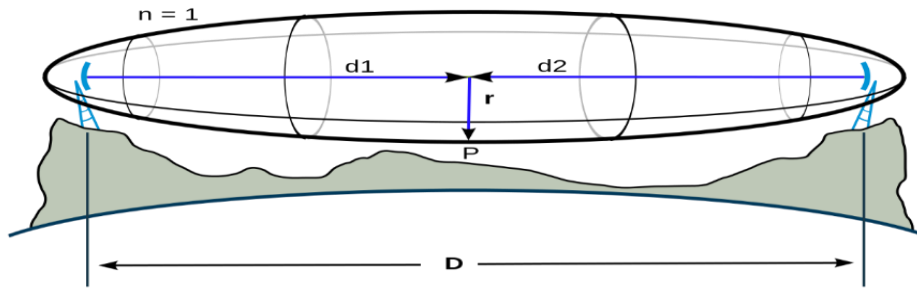
Các thiết bị trạm đo khí tượng, thủy văn được cài đặt đưa thông số dữ liệu qua cổng COM vào các mô-đem LPWAN để truyền dữ liệu lên các trạm truyền dẫn đường trục. Các trạm đo mưa bộ truyền phát dữ liệu có thể nối trực tiếp bộ đếm xung gàu nếu không có giao tiếp COM từ data logger. Từ các điểm đường trục sẽ có giao tiếp LTE (từ 2G đến 4G) để chuyển tải dữ liệu qua mạng dữ liệu di động gửi về một CSDL tại Tổng cục Khí tượng Thủy văn (hoặc hosting ở Hà Nội), có giao tiếp hiển thị trên bản đồ Internet - GIS.



Hình 5. Mô hình kết nối hệ thống.

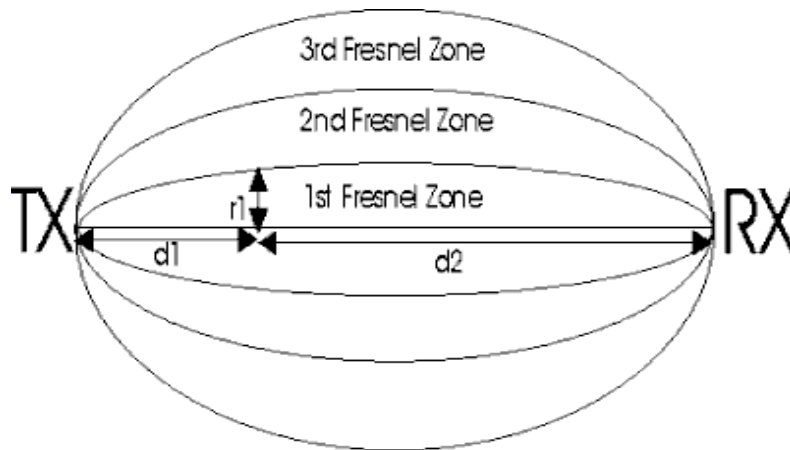
3.2. Kết quả nghiên cứu tính toán trạm thu phát truyền tin

Để có cơ sở tính toán vị trí đặt trạm thu phát truyền tin trong sử dụng công nghệ LoRa chúng ta cần quan tâm đến một thông số kỹ thuật quan trọng là tính toán Fresnel Zone [19, 20]. Vùng Fresnel là một công cụ hiệu quả giúp có thể tính toán được lượng suy hao của tín hiệu do các hiện tượng phản xạ, tán xạ, nhiễu xạ, mà tín hiệu gặp phải trên đường truyền, đặc biệt là nơi có nhiều đồi núi, nhiều cây cối và nhà cao tầng. Thực chất đây là một trong số rất nhiều các khối elip (3 chiều) đồng tâm, với 2 đỉnh elip là đầu thu và đầu phát tín hiệu. Ngoài đường truyền thẳng trong điều kiện môi trường truyền tin lý tưởng (không bị che chắn bởi vật cản), thì tín hiệu truyền từ đầu phát tới đầu thu có thể phản xạ, tán xạ theo nhiều đường khác nhau, trong vùng Fresnel này. Đó là nguyên nhân gây ra sự lệch pha của tín hiệu tại đầu thu, tín hiệu bị nhiễu, truyền về chậm và dung lượng có thể không chính xác.



Hình 6. Không gian Fresnel Zone được hình thành giữa trạm phát và trạm thu.

Các vùng này thường được ký hiệu là: F1, F2, F3 ... Nhưng thực tế thì chỉ có 3 vùng Fresnel đầu tiên là có ảnh hưởng nhiều đến sự lan truyền sóng vô tuyến.



Hình 7. Hình thành các vùng Fresnel.

Tín hiệu trong vùng Fresnel có thể lệch pha với tín hiệu gốc từ 0 đến 90 độ. Ở vùng Fresnel 2 là từ 90 độ đến 270 độ. Vùng Fresnel 3 là từ 270 độ đến 450 độ.

Thực tế cho thấy, không nên quá 40% vùng Fresnel bị che khuất bởi các vật cản. Tuy nhiên để đảm bảo truyền số liệu tốt nhất có thể và hiệu suất tối ưu thì không nên quá 20% vùng Fresnel bị che khuất bởi các vật cản.

Để có thể tính toán được vùng Fresnel có thể dùng công thức:

$$F_n = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}} \quad (1)$$

Trong đó F_n là bán kính vùng Fresnel thứ n ; d_1 và d_2 là khoảng cách từ điểm P (điểm tính bán kính) tới đầu phát và đầu thu. Nếu P nằm chính giữa đầu phát và đầu thu thì bán kính đó đạt giá trị cực đại. Khi đó $d_1 = d_2$; $D = d_1 + d_2$ (tổng khoảng cách từ điểm P (điểm tính bán kính) tới đầu phát và đầu thu); $\lambda = c/f$ (bước sóng); f là tần số, c là vận tốc truyền.

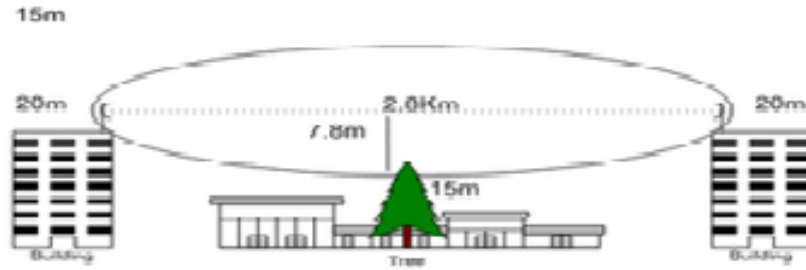
Ta sẽ có công thức đơn giản hơn:

$$r = 8,657 \sqrt{\frac{D}{f}} \quad (2)$$

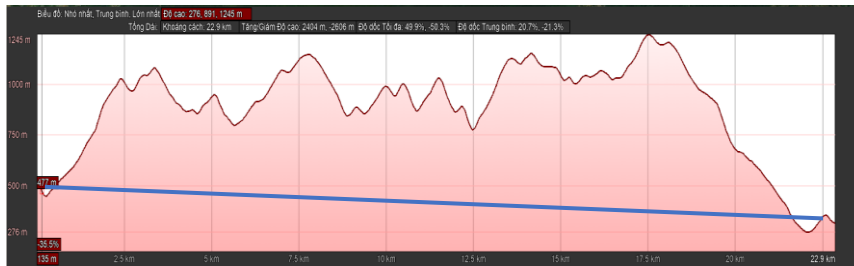
Nhờ việc tính được bán kính vùng Fresnel, có thể dễ dàng tính toán được chiều cao cần thiết của các cột ăng ten để có thể đảm bảo được truyền tín hiệu không bị cản trở quá nhiều (< 40%).

Khoảng không gian Fresnel Zone cần được rõ ràng, nếu có xuất hiện vật cản có thể ảnh hưởng đến tín hiệu truyền sóng, vật cản, đặc biệt là khu vực Tây Nguyên do địa hình cũng có nhiều đồi núi, nên vật cản chủ yếu là cây cối và núi tự nhiên.

Dựa vào phân tích trên, ta có thể tính toán được quá trình truyền sóng giữa các trạm thu phát và vị trí các trạm lặp trong mô hình truyền tin tại Đài KTTV khu vực Tây Nguyên.



Hình 8. Minh họa vật cản trong vùng Fresnel Zone.



Hình 9. Mô hình mô phỏng vật cản từ trạm Đăk Rìng - Đăk Lô tỉnh Kon Tum: 22.9 km.

4. Kết luận

Ngày nay, việc triển khai nghiên cứu ứng dụng công nghệ IoT (internet vạn vật) vào các lĩnh vực khoa học công nghệ và đời sống xã hội không còn mới, mà là xu thế tất yếu của các nước trên thế giới. Công nghệ truyền tin không dây năng lượng thấp (LORA/LPWAN) với ưu thế phủ sóng rộng, băng thông thấp, sử dụng ít năng lượng, hỗ trợ đa kết nối và thiết bị hoạt động ở dải tần không cần đăng ký xin cấp phép tần số đang trở nên phổ biến và ứng dụng nhiều trong các lĩnh vực viễn thông, truyền tin, trao đổi thông tin số liệu. Việc nghiên cứu công nghệ Lora để xây dựng mạng truyền dẫn diện rộng công suất thấp (LPWAN) mạng lưới quan trắc khí tượng thủy văn cho khu vực Tây nguyên là vấn đề cần thiết và có tính khả thi cao dựa trên các phân tích về ưu điểm, sự phù hợp của công nghệ nghiên cứu ứng dụng, nhằm tiến tới phát triển nghiên cứu với mục tiêu cụ thể như xây dựng giải pháp công nghệ LORA cho mạng lưới quan trắc khí tượng thủy văn phục vụ tác nghiệp dự báo khí tượng thủy văn và đặc biệt nghiên cứu thử nghiệm thành công hệ thống công nghệ LORA xây dựng mạng diện rộng công suất thấp (LPWAN) mạng lưới quan trắc KTTV cho khu vực Tây nguyên là phương án dự phòng truyền tin mới có tính bảo mật cao, giảm thiểu khả năng mất dữ liệu quan trắc khí tượng thủy văn góp phần tăng cường dự báo giảm thiểu thiệt hại do thiên tai gây ra.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: V.V.S.; Thu thập, xử lý số liệu: V.V.S.; Phân tích kết quả: V.V.S.; Viết bản thảo bài báo: V.V.S.; Chỉnh sửa bài báo: V.V.S.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi đề tài nghiên cứu khoa học cấp Bộ Tài nguyên và Môi trường “Nghiên cứu công nghệ LORA để thiết kế xây dựng mạng diện rộng công suất thấp (LPWAN) mạng lưới quan trắc khí tượng thủy văn cho khu vực Tây nguyên” mã số: TNMT.2022.04.12, thuộc Chương trình khoa học và công nghệ trọng điểm cấp Bộ “Nghiên cứu ứng dụng công nghệ trong xây dựng, phát triển Chính phủ số và Chuyển đổi số ngành tài nguyên và môi trường giai đoạn 2021 - 2025”, mã chương trình TNMT.04/21-25. Tập thể tác giả xin chân thành cảm ơn.

Lời cam đoan: Tác giả cam đoan bài báo này là công trình nghiên cứu của tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây.

Tài liệu tham khảo

1. Liu, S.; Xia, C.; Zhao, Z. Hội nghị quốc tế IEEE lần thứ 13 năm 2016 về công nghệ vi mạch tích hợp và trạng thái rắn (ICSICT), 2016.

2. Pietrosevoli, E.; Rainone, M.; Zennaro, M. On extending the wireless communications range of weather stations using LoRaWAN. Proceedings of EAI International Conference on Smart Objects and Technologies for Social Good, Valencia, Spain, September 2019, 25–27.
3. Đưa, N.V.; Long, N.H.; Kha, P.Đ.; Thương, T.T.; Cảnh, N.V.; Tấn, Đ.T. Xây dựng mạng cảm biến không dây sử dụng công nghệ truyền thông LORA cho bài toán giám sát và điều khiển trong nông nghiệp công nghệ cao. *Tap chí khoa học công nghệ* **2019**, 50, 27–31.
4. Thanh, N.Đ.; Ferrero, F.; Khải, L.C.V.; Thương, N.H.N.; Tuấn, P.V.; Huy, L.Q. Mạng công nghiệp LORAWAN cho Thành phố Đà Nẵng: Giải pháp cho ứng dụng IoT tầm xa năng lượng thấp. *Tap chí khoa học và công nghệ, Đại học Đà Nẵng* **2020**, 18(1), 7–10.
5. Anh, N.B.Q.; Diễm, T.P.; Quốc, C.A. Tìm hiểu về một nghiên cứu liên quan đến LoRa của hệ thống nông nghiệp thông minh Next Farm. Khoa Kỹ thuật & Công nghệ, Trường ĐH Quy Nhơn. 2019.
6. Báo cáo tổng kết đề tài VT-CN.04/17-20. Nghiên cứu tiếp cận công nghệ sử dụng kính khí cầu thả ở tầng bình lưu tích hợp công nghệ trạm thu phát thông tin để giám sát, dẫn đường tìm kiếm cứu hộ và đo đạc thông số vật lý môi trường tầng khí quyển. đề tài thuộc chương trình KHCN Vũ trụ giai đoạn 2016-2020, cơ quan chủ trì Trung tâm Tin học và Tính toán, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, 2020.
7. Báo cáo tổng kết đề tài VT-CN.02/18-20. Nghiên cứu, thiết kế và chế tạo mẫu tên lửa nghiên cứu (Sounding Rocket) đưa thiết bị khoa học để thử nghiệm thu thập dữ liệu khí quyển tầng cao” đề tài thuộc chương trình KHCN Vũ trụ giai đoạn 2016-2020. Học viện Kỹ thuật Quân sự, Bộ Quốc phòng.
8. Líc, T.V.; Nam, L.H. Mạng không dây LoRa cho ứng dụng IoT tầm xa. *Tap chí khoa học và Công nghệ Đại học Đà Nẵng* **2018**, 11(132), 50–53.
9. Murdyantoro, E.; Nugraha, A.W.W.; Wardhana, A.W.; Fadli, A.; Zulfa, M.I. A review of LoRa technology and its potential use for rural development in Indonesia. *AIP Conf. Proc.* **2019**, 2094, 020011.
10. Augustin, A.; Yi, J.; Clausen, T.; Townsley, W.M. A study of LoRa: Long range & low power networks for the internet of things. *Sensors* **2016**, 16, 1–18.
11. Murdyantoro, E.; Rosyadi, I.; Septian, H. Studi performansi Jarak Jangkauan LoRa OLG01 sebagai infrastruktur Konektivitas Nirkabel IoT Dinamika Rekayasa **2019**, 15(1), 47–56.
12. Trục tuyến: <https://www.daviteq.com/blog/vi/cong-nghe-sigfox-bien-moi-vat-tro-nen-song-dong/>
13. Trục tuyến: <https://giaiphapviettel.vn/tin-tuc/238/narrowband-iot-nbiot-la-gi.html>
14. Trục tuyến: <http://iottuonglai.com/lte-m-la-gi.html>
15. Trục tuyến: <https://viettuans.vn/cong-nghe-lora-la-gi>
16. Trục tuyến: <https://www.dienmayxanh.com/kinh-nghiem-hay/tim-hieu-ve-mang-lorawan-va-co-che-hoat-dong-nhu-t-1412680>
17. Trục tuyến: <http://iottuonglai.com/mesh-network-la-gi.html>
18. Trục tuyến: <https://vientin.com/blog/mo-hinh-ket-noi-mang-luoi-mesh-27-200>
19. Trục tuyến: <https://www.everythingrf.com/rf-calculators/fresnel-zone-calculator>.
20. LoRa/LoRaWAN tutorial 7: Fresnel Zone: Trục tuyến: <https://www.youtube.com/watch?v=HWOivbJjw7s>.

Research and select low-energy wireless communication technology (LORA/LPWAN) applied in automatic meteorological station data transmission under the Regional Hydro-Meteorological Central Highlands Observatory

Vu Van Sang^{1*}

¹ National Centre for Hydro-Meteorological Network (NCN); sangvu2210@gmail.com

Abstract: Currently, low-energy wireless communication technology (LORA/LPWAN) technology is becoming popular and is increasingly being applied in fields such as telecommunications, information transmission,...including hydrometeorology. The automatic station network of the Regional Hydro-Meteorological Central Highlands Station is mostly installed in remote areas and border, where the telecommunications network has weak signals and poor communication conditions. Therefore, it has caused many challenges for basic investigation, monitoring, warning, forecasting, and natural disaster prevention. The quality of information transmission and reception is often interrupted, it is difficult to accurately determine whether the station is active or inactive; causing difficulty in finding the cause and evaluating the operations of the stations. The study has focused on applying the advantages and suitability of low-power wireless communication technology (LORA/LPWAN) to improve the data transmission efficiency of automatic hydro-meteorological stations and effectively serve the basic investigation of hydrometeorology in the Central Highlands region.

Keywords: LPWAN Low-power wide-area network technology; Communication technology using LORA technology.

Table of content

- 1** Nga, T.N.Q.; Kim, T.T.; Hang, O.B.; Bay, N.T. The assessment of bank erosion susceptibility under the influence of flow before and after the construction of groyne embankments on the Tien River section flowing through Sa Dec City, Dong Thap Province. *J. Hydro-Meteorol.* **2023**, 756, 1–13.
- 14** Bay, N.T.; Kim, T.T.; An, T.T.T.; Nga, T.N.Q. Assessing the impact of sand mining on the riverbed of the Tien River flowing through Vinh Long Province. *J. Hydro-Meteorol.* **2023**, 756, 14–28.
- 29** Hoa, P.T.T.; Quang, V.N.; Nghi, L.T.; Phuong, D.T.N.; Hai, N.M. Research on the potential of applying Random Forest algorithm and Sentinel-2 to classify land cover in Quang Binh Province on the Google Colab platform. *J. Hydro-Meteorol.* **2023**, 756, 29–41.
- 42** Xuan, N.T.; Duyen, N.C.M.; Long, B.T. Simulating PM_{2.5} dust pollution and analyzing related factors – The case of Ca Mau Province, Vietnam. *J. Hydro-Meteorol.* **2023**, 756, 42–58.
- 59** Tuan, N.Q.; Quan, N.H.; Hai, N.T. Comparing the effect of various amendments for Arsenic immobilization in contaminated soil. *J. Hydro-Meteorol.* **2023**, 756, 59–69.
- 70** Hang, N.V.; Binh, P.D.H.; Nam, P.Q.; Tung, N.B.; Long, T.T. Assessment of wind energy potential for renewable energy development in the East Sea, Vietnam. *J. Hydro-Meteorol.* **2023**, 756, 70–82.
- 83** Ha, B.M.; Thuy, N.B.; Ngoc, P.K.; Tien, P.V. Initial results of testing the XGBoost algorithm to predict storm surge at Hon Dau station. *J. Hydro-Meteorol.* **2023**, 756, 83–94.
- 95** Sang, V.V. Research and select low-energy wireless communication technology (LORA/LPWAN) applied in automatic meteorological station data transmission under the regional hydro-meteorological Central Highlands observatory. *J. Hydro-Meteorol.* **2023**, 756, 95–104.