



Original Article

Characterization of Oceanographic Fields in the Gulf of Tonkin and Their Relationship with ENSO Using Cluster Analysis

Giap Ngoc Anh, Nguyen Kim Cuong*, Bui Minh Tuan

VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

Received 25th August 2025

Revised 15th October 2025; Accepted 30th October 2025

Abstract: In this study, the Self-Organizing Maps (SOM) method, combined with the K-means algorithm, was employed to analyze the seasonal variability of oceanic physical fields in the Gulf of Tonkin by grouping similar patterns of surface currents, temperature, and salinity into clusters. Analyzing the transitions between these patterns over time enables the identification and assessment of climate characteristics in these physical fields, utilizing a substantially larger amount of information compared to traditional climatological averages. SOM serves to simplify complex nonlinear features into easily observable two-dimensional relationships, while K-means assists in recognizing and highlighting significant spatial characteristics of the clusters. The results indicate that this approach provides a clearer and more intuitive depiction of the seasonal differences and variability of the physical fields, particularly improving the identification of the timing of current reversals and sea surface temperature changes. Furthermore, the study reveals that some representative patterns occur more frequently during El Niño years, whereas others are more prominent in La Niña years. These findings offer a novel approach, contributing to a better understanding of the relationship between oceanic physical fields in the Gulf of Tonkin and ENSO.

Keywords: SOM, K-Means, ENSO, Gulf of Tonkin.

* Corresponding author.

E-mail address: cuongnk@hus.edu.vn

<https://doi.org/10.25073/2588-1094/vnuees.5403>

Nghiên cứu đặc trưng các trường vật lý biển khu vực Vịnh Bắc Bộ và mối liên hệ với ENSO bằng phương pháp phân cụm

Giáp Ngọc Ánh, Nguyễn Kim Cương*, Bùi Minh Tuân

Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội

334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam

Nhận ngày 25 tháng 8 năm 2025

Chỉnh sửa ngày 15 tháng 10 năm 2025; Chấp nhận đăng ngày 30 tháng 10 năm 2025

Tóm tắt: Trong nghiên cứu này, phương pháp bản đồ tự tổ chức (Self-Organizing Maps, SOM) kết hợp với thuật toán K-means được sử dụng để phân tích sự biến đổi theo mùa của các trường vật lý biển tại Vịnh Bắc Bộ, bằng cách nhóm các hình thể tương đồng của dòng chảy, nhiệt độ và độ muối bề mặt biển thành các cụm. Việc phân tích sự chuyển đổi giữa các hình thể trong các cụm theo thời gian cho phép nhận diện và đánh giá các đặc trưng khí hậu của các trường vật lý biển này dựa trên lượng thông tin lớn hơn rất nhiều so với phương pháp trung bình khí hậu truyền thống. Trong đó, SOM có vai trò đơn giản hóa các đặc trưng phi tuyến tính phức tạp thành những mối quan hệ hai chiều dễ quan sát, còn K-means hỗ trợ nhận diện và làm nổi bật các đặc trưng không gian quan trọng của các nhóm. Kết quả cho thấy phương pháp này giúp mô tả rõ ràng và trực quan hơn sự khác biệt cũng như biến đổi theo mùa của các trường vật lý biển, điển hình là khả năng xác định chính xác hơn thời điểm đảo ngược của dòng chảy và nhiệt độ bề mặt biển. Ngoài ra, nghiên cứu cũng chỉ ra rằng một số hình thể điển hình xuất hiện thường xuyên hơn trong những năm El Niño, trong khi các hình thể khác gia tăng trong những năm La Niña. Những phát hiện này mở ra một hướng tiếp cận mới, góp phần làm sáng tỏ mối quan hệ giữa các trường vật lý biển ở Vịnh Bắc Bộ và ENSO.

Từ khóa: SOM, K-Means, ENSO, Gulf of Tonkin.

1. Mở đầu

Vịnh Bắc Bộ là vịnh biển nửa kín tại phía tây bắc Biển Đông đóng vai trò quan trọng trong kinh tế biển, địa chính trị và khoa học của Việt Nam. Đây là một trong những vịnh lớn nhất khu vực Đông Nam Á với diện tích xấp xỉ 126.250 km². Chiều ngang của vịnh dao động từ 220 km tại khu vực cửa vịnh hẹp nhất đến khoảng 310 km tại vị trí rộng nhất. Vịnh Bắc Bộ kết nối với Biển Đông thông qua eo biển Quỳnh Châu, rộng khoảng 25 km, nằm giữa bán đảo Lô Lô Châu và

đảo Hải Nam và cửa chính kéo dài từ đảo Côn Cỏ (Việt Nam) đến mũi Oanh Ca trên đảo Hải Nam (Trung Quốc) với chiều rộng khoảng 200 km. Nhờ vị trí địa lý đặc biệt, Vịnh Bắc Bộ giữ vai trò chiến lược quan trọng về kinh tế, quốc phòng và an ninh, đồng thời là cửa ngõ giao thương quan trọng của Việt Nam với khu vực và thế giới [1].

Mặc dù đã có nhiều nghiên cứu về các trường vật lý biển tại khu vực Vịnh Bắc Bộ nhưng vẫn còn tồn tại một số hạn chế nhất định. Wu và cộng sự (2008) sử dụng dữ liệu nhiệt độ, độ muối thực

* Tác giả liên hệ.

Địa chỉ email: cuongnk@hus.edu.vn

<https://doi.org/10.25073/2588-1094/vnuees.5403>

tế và dữ liệu ứng suất gió vệ tinh (NASA) theo trung bình mùa, kết hợp với các mô hình số để nghiên cứu cơ chế hoàn lưu xoáy thuận ở Vịnh Bắc Bộ theo mùa. Mô hình số được áp dụng vì cho phép mô phỏng có kiểm soát các yếu tố động lực như gió, thủy triều và dòng qua eo biển, đồng thời tách riêng ảnh hưởng của từng cơ chế lên hoàn lưu mùa, kể cả trong điều kiện thiếu dữ liệu quan trắc đầy đủ. Kết quả nghiên cứu chỉ ra rằng dòng chảy qua eo biển Quỳnh Châu đóng vai trò then chốt trong việc duy trì hoàn lưu xoáy thuận quanh năm ở Vịnh Bắc Bộ, ngay cả khi gió mùa thay đổi theo mùa [2].

Yang và cộng sự (2013) sử dụng dữ liệu gió và dòng chảy thực đo để nghiên cứu hoàn lưu theo mùa ở Vịnh Bắc Bộ. Phương pháp nghiên cứu so sánh vận tốc nước quan trắc với mô hình Global-FVCOM, đồng thời kết hợp các thí nghiệm thành phần để đánh giá vai trò của các cơ chế vật lý trong tuần hoàn mùa hè. Kết quả cho thấy hoàn lưu xoáy thuận mùa hè hình thành chủ yếu nhờ sự kết hợp giữa gió mùa trên nền nước phân tầng và dòng chảy triều từ eo biển Quỳnh Châu, trong khi gió mùa và dòng chảy sông Hồng chỉ làm tăng cường độ xoáy thuận ở tầng mặt [3].

Hà Thanh Hương (2016) ứng dụng phương pháp bình phương tối thiểu có trọng số xây dựng mặt cong trơn thể hiện sự phân hóa cấu trúc nhiệt độ theo trường bề mặt và theo độ sâu áp dụng cho khu vực Vịnh Bắc Bộ. Mục tiêu là xác định cấu trúc nhiệt độ gần đúng nhất làm trường ban đầu cho các mô hình nghiên cứu các trường thủy văn biển trên khu vực nghiên cứu theo các tháng đại diện các mùa trung bình nhiều năm [4]. Rogowski và cộng sự (2019) sử dụng mô hình ROMS kết hợp ảnh vệ tinh, radar HF, khảo sát thủy văn và quan trắc phao để đánh giá xu thế dòng chảy theo mùa ở phía tây Vịnh Bắc Bộ, đặc biệt là lưu lượng nước từ sông Hồng. Kết quả trung bình mùa cho thấy sự hình thành và lan truyền của dòng ven biển theo mùa là đặc điểm động lực chính, cùng với ảnh hưởng của bão và dòng sông đến động lực khu vực này [5].

Gao và cộng sự [6-9] đã sử dụng mô hình POM để mô phỏng hoàn lưu ở Vịnh Bắc Bộ theo trung bình tháng và theo mùa. Các nghiên cứu

này cho thấy vào mùa hè, vùng phía bắc xuất hiện xoáy thuận, còn phía nam lại hình thành xoáy nghịch, sang mùa đông thì chiều dòng chảy đảo ngược. Dòng chảy qua eo biển Quỳnh Châu cũng thay đổi theo mùa, trong khi gió mùa và dòng chảy từ Biển Đông là hai yếu tố chi phối chính. Nhiệt độ và thủy triều tuy không làm thay đổi cấu trúc chung của dòng chảy nhưng lại ảnh hưởng đến cường độ, phạm vi và sự hình thành của khối nước lạnh. Ngoài ra, cường độ và phạm vi dòng chảy còn dao động theo từng năm và từng tháng, chịu tác động đồng thời của gió mùa, dòng chảy Biển Đông, nhiệt độ, thủy triều và địa hình. Khối nước lạnh này thường bắt đầu hình thành vào mùa xuân, mạnh nhất vào mùa hè và dần biến mất khi sang mùa thu.

Với sự phát triển của các cơ sở dữ liệu tái phân tích quy mô lớn như HYCOM, ECMWF, ... việc nghiên cứu sự biến động theo thời gian của các yếu tố động lực biển ngày càng khả thi và chi tiết hơn. Những nguồn dữ liệu này cung cấp hàng chục nghìn trường số liệu liên tục với độ phân giải cao cả về không gian lẫn thời gian giúp phân tích rõ ràng hơn các quá trình hải dương học. Tuy nhiên, khối lượng dữ liệu lớn như vậy đòi hỏi phải có phương pháp phân tích phù hợp để xử lý và khai thác hiệu quả. Trong số đó, phương pháp phân cụm được đánh giá là công cụ hữu ích, cho phép phân loại dữ liệu thành các nhóm có đặc điểm tương đồng và theo dõi sự biến động của các trường vật lý theo thời gian. Nhờ vậy, các xu thế, quy luật biến đổi hay những bất thường trong các yếu tố động lực biển có thể được nhận diện rõ ràng, hỗ trợ tốt cho nghiên cứu, dự báo và quản lý tài nguyên biển.

Phương pháp phân cụm Bản đồ tự tổ chức (Self-Organizing Maps: SOM) và K-Means là những công cụ phân tích dữ liệu hiện đại, đặc biệt hữu ích trong việc khai thác thông tin ẩn và nhận diện các mẫu đặc trưng từ những tập dữ liệu phức tạp. Hai thuật toán này cho phép đơn giản hóa các bộ dữ liệu quy mô lớn giúp giảm chiều dữ liệu nhưng vẫn giữ được các đặc trưng quan trọng, từ đó hỗ trợ việc phân tích hiệu quả hơn. Trong hải dương học, SOM và K-Means được ứng dụng rộng rãi để trích xuất và phân loại các đặc trưng động lực và thủy văn của biển như

dòng chảy, nhiệt độ, độ muối hay mực nước biển. Nhờ khả năng phát hiện các cấu trúc ẩn trong dữ liệu, các phương pháp này giúp xác định các chế độ thủy văn khác nhau, nhận diện sự biến động cũng như mối liên hệ với các hiện tượng khí hậu lớn.

Trong lĩnh vực hải dương học trên thế giới đã có một số nghiên cứu sử dụng thuật toán phân cụm cho nghiên cứu. Liu và cộng sự (2006) đã đánh giá hiệu suất của SOM trong trích xuất đặc trưng dòng chảy từ dữ liệu biển so với các phương pháp truyền thống. Kết quả cho thấy thuật toán có thể nhận dạng các mẫu từ dữ liệu nhiễu và tách biệt các mẫu dòng chảy mạnh liên quan đến thời tiết khắc nghiệt. Phương pháp này cũng được so sánh với phương pháp EOF và chứng minh khả năng chọn lọc mẫu vượt trội cùng cách áp dụng thực tiễn với nhiều ví dụ [10]. Vilibic và cộng sự (2016) phát triển hệ thống dự báo dòng chảy bề mặt biển dựa trên SOM huấn luyện từ dữ liệu dòng chảy radar HF và dự báo khí tượng cho khu vực ven biển Adriatic phía bắc và so sánh với hệ thống ROMS. Kết quả chỉ ra hệ thống SOM có kỹ năng dự báo tốt hơn, đặc biệt trong điều kiện gió mạnh và có thể cải thiện thêm với dữ liệu chất lượng cao và dài hơn [11]. Dey và cộng sự (2023) áp dụng thuật toán SOM để phân tích các kiểu dòng chảy mặt tại Vịnh Bengal giai đoạn 1993–2015, xác định bốn nhóm chính liên quan đến các hệ thống dòng chảy. Nghiên cứu cho thấy sự biến đổi của các kiểu này liên quan đến ENSO, IOD và cho thấy tần suất xuất hiện của các đặc trưng dòng chảy là khác nhau [12].

Ngoài các nghiên cứu sử dụng SOM trong hải dương học thì thuật toán K-means cũng được sử dụng rộng rãi. Lemenkova và cộng sự (2019) sử dụng thuật toán K-means để phân tích dữ liệu địa chất của rãnh Mariana dựa trên các tham số như độ dày trầm tích, vị trí kiến tạo, núi lửa, độ sâu và độ dốc địa hình. Kết quả cho thấy số cụm tối ưu là năm cụm và cung cấp các chương trình tính toán trong nghiên cứu tương tự [13]. Sun và cộng sự (2021) áp dụng K-means trên dữ liệu nhiệt độ và độ muối từ CESM2 và WOA để xác định các chế độ thủy văn ở các vùng biển Amundsen, Bellingshausen và Ross. Nghiên cứu chỉ ra CESM2 có độ lệch so với WOA nhưng vẫn

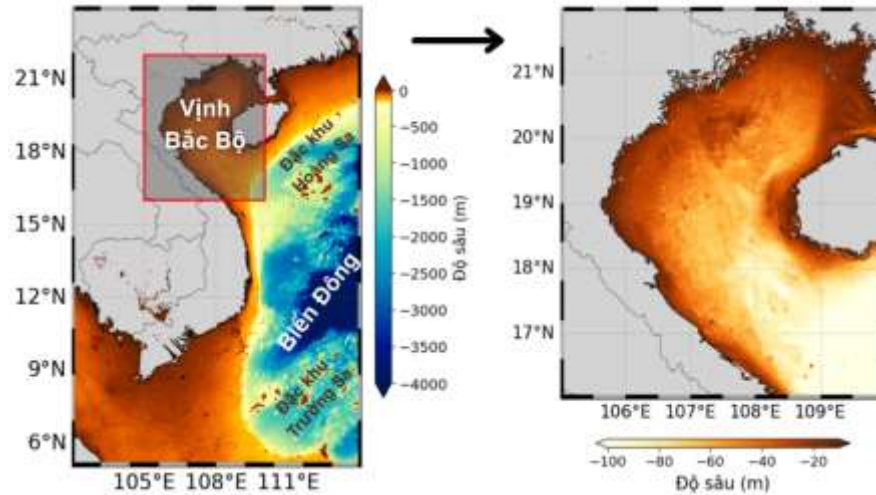
tái hiện tốt sự khác biệt tương đối giữa các chế độ và phạm vi phân bố của chúng [14]. Saastamoinen và cộng sự (2021) thảo luận việc phân loại đáy biển Baltic thành khu vực mềm và cứng bằng phương pháp học không giám sát kết hợp với các kỹ thuật mô phỏng thị giác con người. Kết quả sử dụng bộ lọc Gabor trước khi phân cụm K-means giúp cải thiện đáng kể so với chỉ dùng K-means hoặc màu Lab [15]. Romero và cộng sự (2024) áp dụng kỹ thuật phân cụm K-means trên dữ liệu nhiệt độ và độ muối của toàn bộ đại dương để xác định các khối nước có cấu trúc nhiệt - muối tương đồng ở nhiều quy mô không gian khác nhau. Kết quả chỉ ra các cấu trúc động lực chính của đại dương ảnh hưởng mạnh đến cấu trúc nhiệt - muối theo chiều thẳng đứng và cho phép xác định các ranh giới khu vực dựa trên đặc trưng thủy văn thay vì ranh giới địa lý cố định [16].

Việc kết hợp SOM và K-Means trong nghiên cứu hải dương học mở ra hướng tiếp cận mới, dù hiện vẫn còn ít nghiên cứu thực hiện. Một trong số ít các nghiên cứu tiêu biểu trong lĩnh vực này là của Bùi Minh Tuấn và cộng sự (2024), kết hợp hai thuật toán này nhằm phát hiện các mô hình synop đặc trưng cho sự khởi đầu gió mùa Nam Á và Đông Nam Á trong các năm ENSO, qua đó khắc phục những hạn chế của phương pháp truyền thống. Kết quả nghiên cứu cho thấy trong pha âm của ENSO, thời điểm khởi đầu gió mùa có xu hướng trung bình muộn hơn, áp cao cận nhiệt đới trên Biển Đông duy trì lâu hơn, trong khi các điều kiện khô hạn gia tăng ở khu vực Ấn Độ Dương gần xích đạo. Bên cạnh đó, gió tây và đối lưu sâu trong pha âm cũng suy yếu rõ rệt, phản ánh sự chi phối mạnh mẽ của ENSO đối với các đặc trưng biển – khí quyển trong thời kỳ khởi phát gió mùa [17, 18]. Những phát hiện này cho thấy tiềm năng của phương pháp kết hợp trong việc nâng cao độ tin cậy của các nghiên cứu về hiện tượng khí hậu – hải dương [17].

Bài báo này ứng dụng kết hợp cả hai phương pháp SOM và K-means để phân cụm đồng thời và đánh giá các biến động của các trường dòng chảy mặt, nhiệt độ mặt biển và độ muối mặt biển để làm sáng tỏ những biến động mùa cũng như đánh giá mối liên hệ với ENSO. So sánh sự xuất

hiện các dạng dòng chảy trên Vịnh Bắc Bộ theo thời gian trong các năm ENSO có thể làm sáng

tỏ hơn về biến động trường dòng chảy trên Vịnh Bắc Bộ (Hình 1).



Hình 1. Khu vực nghiên cứu.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Cơ sở dữ liệu

Dữ liệu sử dụng trong nghiên cứu này được thu thập trong giai đoạn từ năm 1994 đến hết năm 2015 bao gồm ba nguồn số liệu chính. Thứ nhất, dữ liệu vận tốc dòng chảy, nhiệt độ và độ muối mặt biển sử dụng từ GOFS 3.1 (41-layer HYCOM + NCODA Global 1/12° Reanalysis, GLBv0.08 grid, 1994–2015), do Naval Research Laboratory cung cấp. Dữ liệu có độ phân giải cao (1/12°) trong vùng nghiên cứu, tần suất 3 giờ [18, 19]. Thứ hai, các pha ENSO (El Niño, La Niña, trung tính) được xác định dựa trên Southern Oscillation Index (SOI), tính từ chênh lệch áp suất mực nước biển giữa Tahiti và Darwin (1981–2010) và dữ liệu SOI được thu thập từ Cơ quan Quản lý Khí quyển và Đại dương Quốc gia Hoa Kỳ (NOAA - National Oceanic and Atmospheric Administration). Giá trị SOI âm kéo dài tương ứng với El Niño, giá trị SOI dương kéo dài tương ứng với La Niña và giá trị SOI gần 0 tương ứng với pha trung tính [20].

2.2. Phương pháp phân cụm

Bản đồ tự tổ chức (Self-Organizing Map – SOM) do Kohonen (1982) đề xuất là một thuật

toán phân cụm mạnh mẽ, bao gồm hai thành phần chính: lớp đầu vào (input layer) và lớp cạnh tranh (competition layer) [21]. Lớp đầu vào trong nghiên cứu này được xây dựng từ các véc-tơ chứa giá trị trung bình ngày của dòng chảy, nhiệt độ mặt biển và độ muối tại các điểm lưới trên khu vực Vịnh Bắc Bộ. Trong giai đoạn 1994-2015 có tổng cộng khoảng 8.035 véc-tơ đầu vào. Lớp cạnh tranh cũng được biểu diễn dưới dạng véc-tơ có cấu trúc tương tự lớp đầu vào nhưng với số lượng ít hơn và được sắp xếp trên một lưới hai chiều, gọi là các véc-tơ trọng số (weight vectors). Trong nghiên cứu này, lưới cạnh tranh có kích thước 15×15, tương ứng với 225 véc-tơ trọng số. Các véc-tơ đầu vào có đặc trưng tương đồng sẽ được ánh xạ tới các véc-tơ trọng số thông qua quá trình huấn luyện của SOM.

Trong nghiên cứu của Vesanto và cộng sự (2000), nghiên cứu nhấn mạnh việc lựa chọn kích thước lưới SOM cần đảm bảo đủ lớn để tạo ra một tập các véc-tơ nguyên mẫu đa dạng, nhằm đại diện chính xác cho cấu trúc phức tạp và đa dạng của dữ liệu đầu vào [22]. Do đó, bộ dữ liệu gồm 8035 véc-tơ, kích thước lưới 15 x 15 (tương đương 225 nút) là một lựa chọn hợp lý bởi kích thước này nằm trong khoảng phù hợp để cân bằng giữa khả năng biểu diễn chính xác đặc trưng dữ liệu và tối ưu hóa tính toán cho các bước

phân cụm kế tiếp. Việc chọn kích thước quá nhỏ sẽ làm giảm chi tiết biểu diễn dữ liệu, dẫn đến thiếu sót trong đại diện, trong khi kích thước quá lớn có thể làm tăng tính toán không cần thiết và gây ra hiện tượng quá khớp.

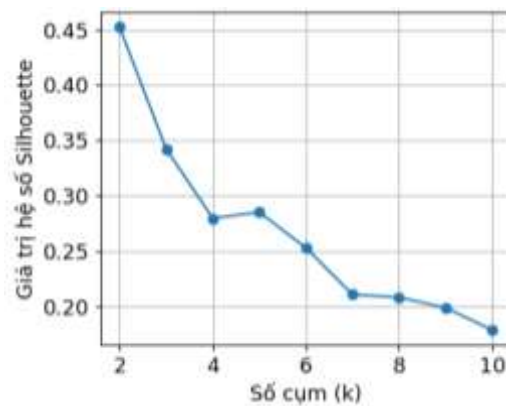
Trước khi huấn luyện, các biến (dòng chảy, nhiệt độ mặt, độ muối) được chuẩn hóa nhằm loại bỏ sự khác biệt về đơn vị và quy mô. Các véc-tơ trọng số sau đó được khởi tạo ngẫu nhiên. Trong quá trình huấn luyện, SOM xác định véc-tơ trọng số gần nhất với véc-tơ đầu vào dựa trên khoảng cách Euclid, véc-tơ này được gọi là Best Matching Unit (BMU). BMU sau đó được hiệu chỉnh để tiến gần hơn tới véc-tơ đầu vào, đồng thời các véc-tơ trọng số lân cận (trong bán kính ảnh hưởng) cũng được điều chỉnh theo hướng tương tự. Những véc-tơ ở xa BMU sẽ ít hoặc không bị điều chỉnh. Quá trình này được lặp lại nhiều lần cho đến khi tập véc-tơ trọng số phản ánh được cấu trúc của dữ liệu đầu vào (chi tiết trong Kohonen, 1982).

Sau giai đoạn huấn luyện, các véc-tơ trọng số tiếp tục được phân cụm bằng thuật toán K-means. Bước này cho phép nhóm các véc-tơ trọng số tương đồng vào những cụm lớn hơn, từ đó hỗ trợ phân tích và nhận diện đặc trưng của từng cụm. Để xác định số cụm tối ưu sau khi phân cụm trọng số SOM bằng K-means, nghiên cứu sử dụng hệ số Silhouette – một chỉ số định lượng phổ biến trong học không giám sát, phản

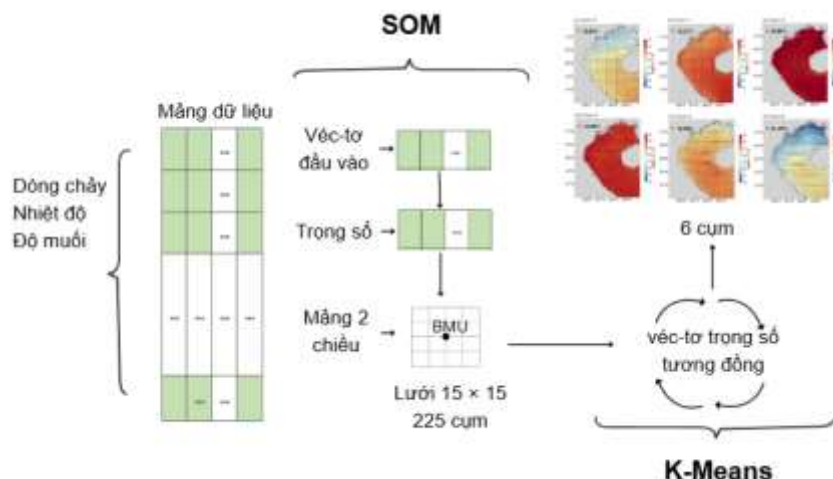
ánh mức độ gắn kết nội cụm và phân tách giữa các cụm. Hệ số Silhouette cho mỗi điểm được tính theo công thức:

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (1)$$

trong đó $a(i)$ là khoảng cách trung bình giữa điểm i và các điểm trong cùng cụm, còn $b(i)$ là khoảng cách trung bình nhỏ nhất giữa điểm i và các cụm khác. Giá trị trung bình của $s(i)$ trên toàn bộ dữ liệu biểu thị chất lượng tổng thể của phân cụm; giá trị càng cao, các cụm càng tách biệt và đồng nhất hơn [23, 24]. Hình 2 dưới đây biểu diễn hệ số Silhouette khi xác định số cụm K-means nghiên cứu này.



Hình 2. Hệ số Silhouette.



Hình 3. Sơ đồ phương pháp thực hiện.

Từ Hình 2 cho thấy hệ số Silhouette đạt giá trị lớn nhất tại $k = 2$, sau đó giảm dần khi số cụm tăng lên. Tuy nhiên, việc lựa chọn số cụm tối ưu còn cần xem xét ý nghĩa vật lý và cấu trúc không gian của dữ liệu, do đó nghiên cứu chọn $k = 6$ để đảm bảo sự cân bằng giữa tính phân tách, ổn định và khả năng diễn giải khoa học của các cụm. Cuối cùng, sơ đồ phương pháp thực hiện được mô tả rõ ràng trong Hình 3.

Thêm vào đó, để xác định tần suất xuất hiện của từng cụm trong toàn bộ giai đoạn nghiên cứu, nghiên cứu sử dụng công thức tính tần suất xuất hiện dựa trên tỷ lệ thời gian cụm đó xuất hiện so với tổng thời gian dữ liệu. Cách tiếp cận này cho phép so sánh trực quan mức độ phổ biến của từng cụm trong toàn bộ giai đoạn. Cụ thể, tần suất xuất hiện của cụm i được xác định theo công thức:

$$F_i = \frac{T_i}{T} \times 100 \quad (2)$$

Trong đó, T_i là tổng thời gian mà cụm i xuất hiện, còn T là tổng thời gian của toàn bộ chuỗi dữ liệu nghiên cứu.

Để phân tích tần suất xuất hiện của từng cụm theo các pha ENSO, trước hết cần xác định rõ các đại lượng liên quan. Ký hiệu $n_{i,p}$ là tổng số ngày mà cụm i xuất hiện trong pha ENSO p (trong đó p có thể là El Niño, La Niña hoặc pha trung tính). Đồng thời, N_i được định nghĩa là tổng số ngày xuất hiện của cụm i trong toàn bộ chuỗi dữ liệu nghiên cứu, không phân biệt pha ENSO. Khi đó, tần suất xuất hiện của cụm i trong pha p được tính bằng tỷ lệ giữa $n_{i,p}$ và N_i , thể hiện mức độ đóng góp của pha ENSO đó vào sự hình thành cụm i trong toàn bộ giai đoạn nghiên cứu. Công thức được biểu diễn như sau:

$$F_{i,p} = \frac{n_{i,p}}{N_i} \quad (3)$$

3. Kết quả và thảo luận

3.1. Biến động của các trường vật lý biển

3.1.1. Biến động trường dòng chảy trung bình mặt biển

Hình 4 thể hiện trường dòng chảy trung bình mặt biển được phân thành 6 cụm đại diện cho cả

giai đoạn từ 1994-2015, kèm theo đó là giá trị tần suất xuất hiện. Cụm 0 có tần suất xuất hiện 12,89% với chế độ dòng chảy yếu và xuất hiện chủ yếu vào giai đoạn mùa hè và thường xuất hiện trong giai đoạn chuyển tiếp vào tháng 3-4 hàng năm (Hình 7a). Trong cụm này, vận tốc dòng chảy dao động từ 0,1-0,3 m/s trên toàn vịnh. Cụm 1 có tần suất xuất hiện 12,21% và dòng chảy có vận tốc 0,1-0,3 m/s. Tại kinh độ 109°E xuất hiện dải dòng chảy có vận tốc 0,3 m/s di chuyển về phía bắc đến vĩ độ 19°N. Xu thế dòng chảy này thường xuất hiện vào tháng 5-6 (Hình 7a).

Cụm 2 đánh dấu sự tăng lên rõ rệt về vận tốc dòng chảy và mức độ phức tạp trong hướng dòng chảy với tần suất xuất hiện 27,6%. Ở phía cửa Vịnh, dòng chảy di chuyển theo chiều kim đồng hồ có dải vận tốc khoảng 0,3-0,4 m/s thể hiện rõ sự xâm nhập của dòng chảy nền ngoài cửa vịnh. Phần còn lại của vịnh có dòng chảy yếu dao động 0,1-0,3 m/s với hướng Đông Bắc thịnh hành. Trường dòng chảy trong cụm 2 thể hiện trường dòng chảy mùa hè và hoạt động kéo dài từ tháng 6 đến tháng 9 hàng năm (Hình 7a). Trong khi đó, cụm 3 xuất hiện xu hướng dòng chảy có sự đối lập hoàn toàn so với cụm 2 và tần suất xuất hiện giảm còn 13,98%. Dòng chảy từ eo Quỳnh Châu đi vào Vịnh theo hướng Tây – Tây Nam, sau đó di chuyển theo chiều dọc bờ xâm nhập sâu xuống phía nam. Tại phía nam đảo Hải Nam xuất hiện 1 dải dòng chảy có vận tốc lớn hơn, trong đó một phần dòng chảy di chuyển vào Vịnh với hướng Tây Bắc, sau đó đổi hướng về phía đông nam và hợp lưu với dòng chảy ven bờ. Vận tốc cực đại trong cụm này đạt khoảng 0,4 m/s tại vùng ven bờ và dòng chảy phía nam đảo Hải Nam. Cụm này xuất hiện vào giai đoạn chuyển tiếp từ gió mùa mùa hè sang gió mùa mùa đông khi gió mùa Đông Bắc bắt đầu hoạt động thường xuất hiện vào cuối tháng 9 đến đầu tháng 11 (Hình 7a).

Cụm 4 đặc trưng bởi vận tốc dòng chảy tăng mạnh hơn so với các giai đoạn trước nhưng tần suất xuất hiện chỉ chiếm 12,05%. Hình thể dòng chảy trong cụm 4 tương tự cụm 3 nhưng vận tốc dòng chảy tại phía nam đảo Hải Nam về phía nam tăng cường, đạt giá trị cực đại khoảng 0,5 m/s. Đồng thời, dòng chảy từ eo Quỳnh Châu

đi vào Vịnh cũng ghi nhận vận tốc tối đa khoảng 0,4 m/s. Xu thế dòng chảy mặt này chủ yếu xuất hiện khi gió mùa Đông Bắc trong tháng 11-12 (Hình 7a). Mặt khác, cụm 5 có tần suất xuất hiện 21,27% và thể hiện xu thế suy yếu của dòng chảy trên phần lớn vịnh. Chỉ khu vực phía đông nam còn duy trì vận tốc cao, đạt cực đại khoảng 0,4-0,6 m/s, trong khi các vùng trung tâm và ven bờ phía bắc giảm đáng kể so với cụm 4. Xu thế dòng chảy này xuất hiện chủ yếu trong mùa đông và bắt đầu từ cuối tháng 12 đến cuối tháng 3 (Hình 7a).

Qua phân tích xu thế và thời gian xuất hiện trong năm có thể thấy rõ biến động của trường dòng chảy trong mùa đông và mùa hè cũng như trong giai đoạn chuyển tiếp. Trong mùa đông có sự xen kẽ giữa các hình thể dòng chảy tùy theo sự hoạt động của gió mùa trong khi trong mùa hè trường dòng chảy hoạt động tương đối ổn định từ cuối tháng 5 đến cuối tháng 9 (Hình 7a).

3.1.2. Biến động trường độ muối trung bình mặt biển

Hình 5 mô tả phân bố trường độ muối và trường dòng chảy trung bình mặt biển, thể hiện sự thay đổi rõ rệt của hai yếu tố này theo không gian trong Vịnh Bắc Bộ. Sự phân bố độ muối trong các cụm 0, 1 và 2 thể hiện xu thế giảm dần giá trị độ muối cùng với sự thay đổi hướng lan truyền của khối nước ven bờ trong vịnh. Cụm 0 đặc trưng bởi vùng có độ muối thấp phân bố dọc bờ phía tây vịnh và lan truyền theo hướng dọc bờ. Tại khu vực cửa sông Hồng có độ muối dưới 31 PSU và di chuyển xuống phía nam theo hướng dòng chảy. Ngược lại, khu vực đông nam vịnh ghi nhận độ muối cao nhất đạt khoảng 34 PSU, cao hơn rõ rệt so với khu vực ven bờ phía tây và phía bắc vịnh. Trong khi đó, cụm 1 vẫn duy trì vùng nước có độ muối thấp dọc bờ, nhưng hướng lan truyền thay đổi khi dòng chảy ven bờ vận chuyển khối nước này từ phía nam lên phía bắc Vịnh. Ở phía đông nam vịnh, độ muối tại khu vực này giảm nhẹ còn 33–33,5 PSU so với ghi nhận trong cụm 0. Cụm 2 cho thấy sự suy giảm toàn bộ giá trị độ muối trong khu vực nghiên cứu so với các cụm trước. Hướng lan truyền của dòng chảy ven bờ trong cụm 2 tương tự cụm 1 nhưng các giá trị độ muối đều thấp hơn. Cụ thể, dọc bờ Vịnh khu vực 19°N có độ muối giảm xuống dưới

32 PSU; tại cửa sông Hồng có độ muối duy trì mức 30 PSU; trong khi đó, vùng phía nam Vịnh ghi nhận độ muối khoảng 32,5 PSU, thấp hơn rõ rệt so với cụm 1.

Cụm 3 đặc trưng bởi sự phân hóa độ muối giữa khu vực ven bờ và khu vực trong lòng Vịnh. Độ muối tại khu vực từ cửa sông Hồng khoảng 30 PSU theo dòng chảy lan truyền về phía nam của Vịnh và độ muối tăng dần lên 31-32 PSU khi tiến gần ra Biển Đông. Ở khu vực đông bắc Vịnh, ven bờ ghi nhận độ muối giảm xuống còn khoảng 31 PSU, sau đó lan rộng dần ra khu vực trung tâm. Nhìn chung, cụm 3 có giá trị độ muối cao nhất chỉ đạt khoảng 32,5 PSU, thấp hơn đáng kể so với khu vực đông nam Vịnh trong cụm 0.

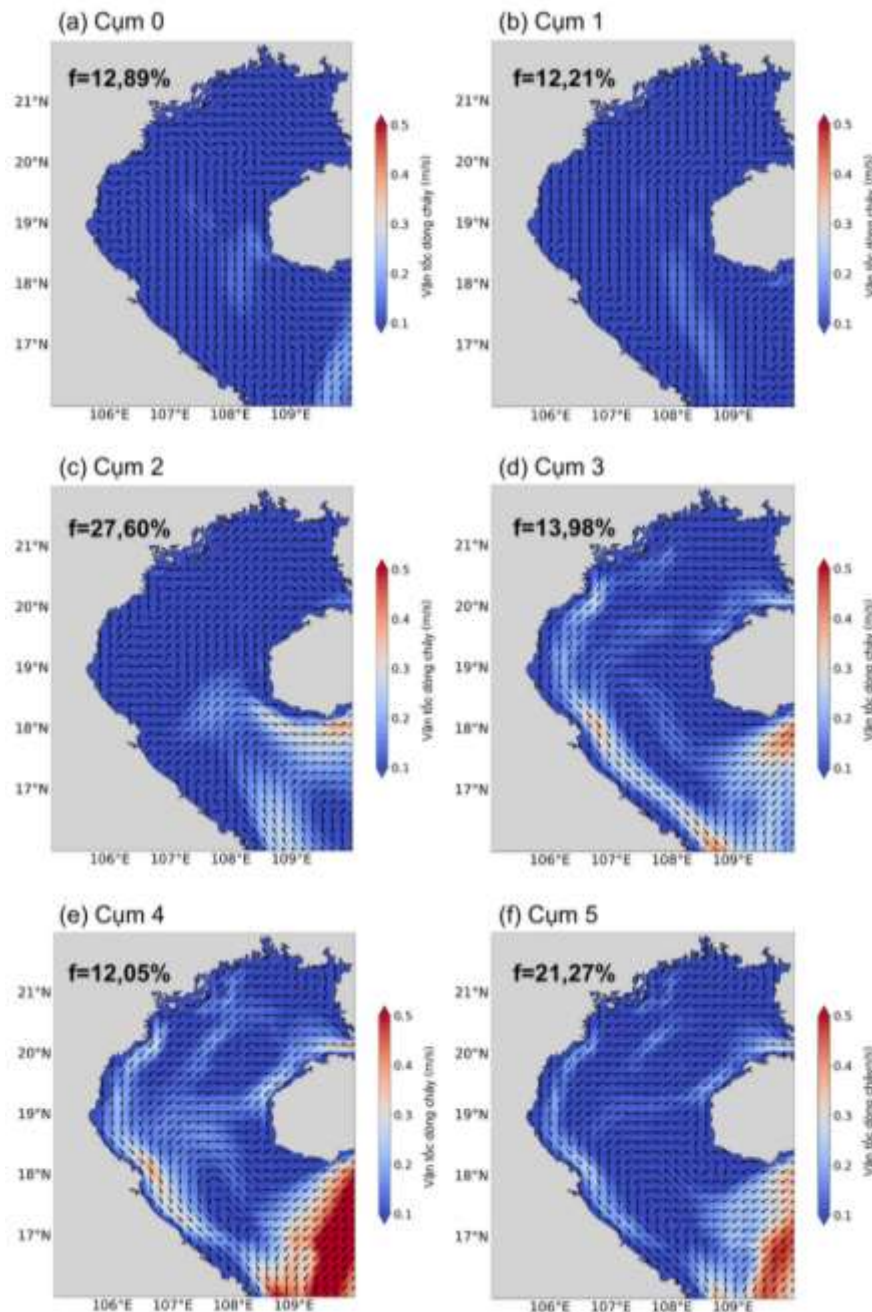
Cụm 4 và cụm 5 tiếp tục cho thấy những đặc điểm đáng chú ý trong phân bố dòng chảy kết hợp với độ muối tại Vịnh Bắc Bộ. So với các cụm trước, hai cụm này vẫn giữ được cấu trúc tổng thể tương đồng với cụm 3, nhưng giá trị độ muối có sự thay đổi rõ rệt, đặc biệt khi so sánh với cụm 2. Ở cụm 4, vùng ven bờ từ khu vực cửa sông Hồng duy trì dải độ muối 30 PSU, song phạm vi lan truyền về phía nam Vịnh thu hẹp hơn so với cụm 3. Trong khi đó, cụm 5 cho thấy khu vực phía ngoài Vịnh đạt giá trị độ muối cao nhất đến 34 PSU. So sánh giữa các cụm cho thấy khu vực đông nam Vịnh, độ muối có xu hướng tăng lên từ mức 31-32 PSU trong các cụm trước lên đến 34 PSU ở cụm 5. Như vậy, cả hai cụm 4 và 5 phản ánh sự chuyển dịch độ muối từ vùng ven bờ có giá trị thấp sang vùng xa bờ và đông nam Vịnh có giá trị cao hơn, đặc biệt rõ trong cụm 5.

3.1.3. Biến động trường nhiệt độ trung bình mặt biển

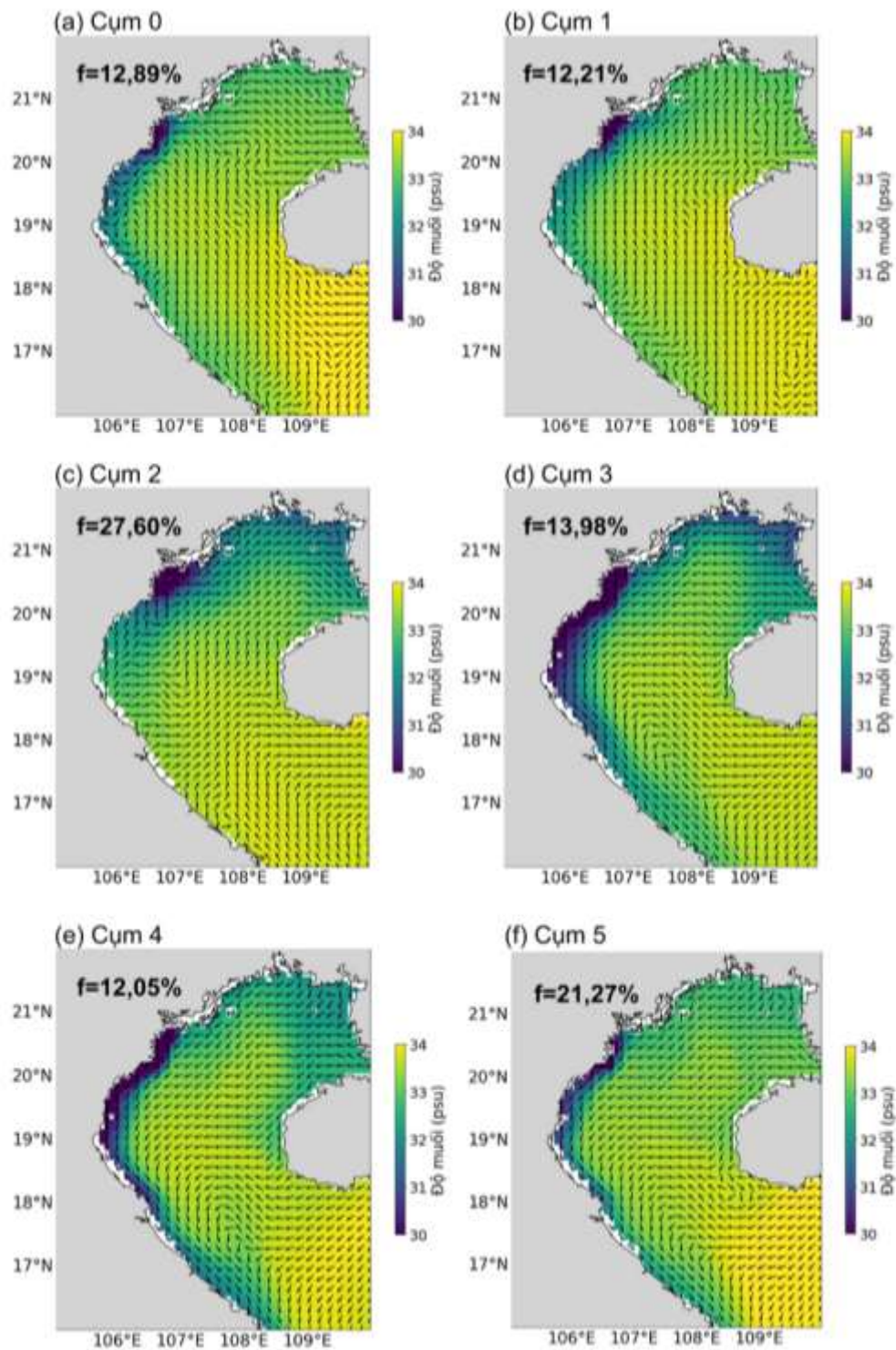
Hình 6 cho thấy nhiệt độ trung bình mặt biển tại Vịnh Bắc Bộ có sự phân hóa rõ rệt về không gian qua sáu cụm đặc trưng. Cụm 0 thể hiện nền nhiệt thấp với sự chênh lệch giữa các khu vực trong Vịnh. Ở phía đông nam Vịnh, nhiệt độ trung bình khoảng 24 °C, sau đó giảm dần vào trung tâm Vịnh còn 22 °C, trong khi khu vực phía bắc chỉ đạt 20-21 °C. Sang cụm 1, nền nhiệt mặt biển tăng rõ rệt so với cụm 0 nhưng tần suất xuất hiện 12,21% không khác biệt đáng kể so với cụm 0. Vùng ven bờ phía bắc Vịnh ghi nhận nhiệt độ 23-24 °C, khu vực trung tâm đạt 24-25

°C, còn phía nam Vịnh có nhiệt độ cao nhất, khoảng 27-28 °C. Trên toàn Vịnh, nhiệt độ giảm dần từ nam lên bắc theo hướng dòng chảy, đồng thời ranh giới các dải nhiệt độ không còn thể hiện rõ rệt như cụm 0. Cụm 2 đặc trưng bởi nền nhiệt

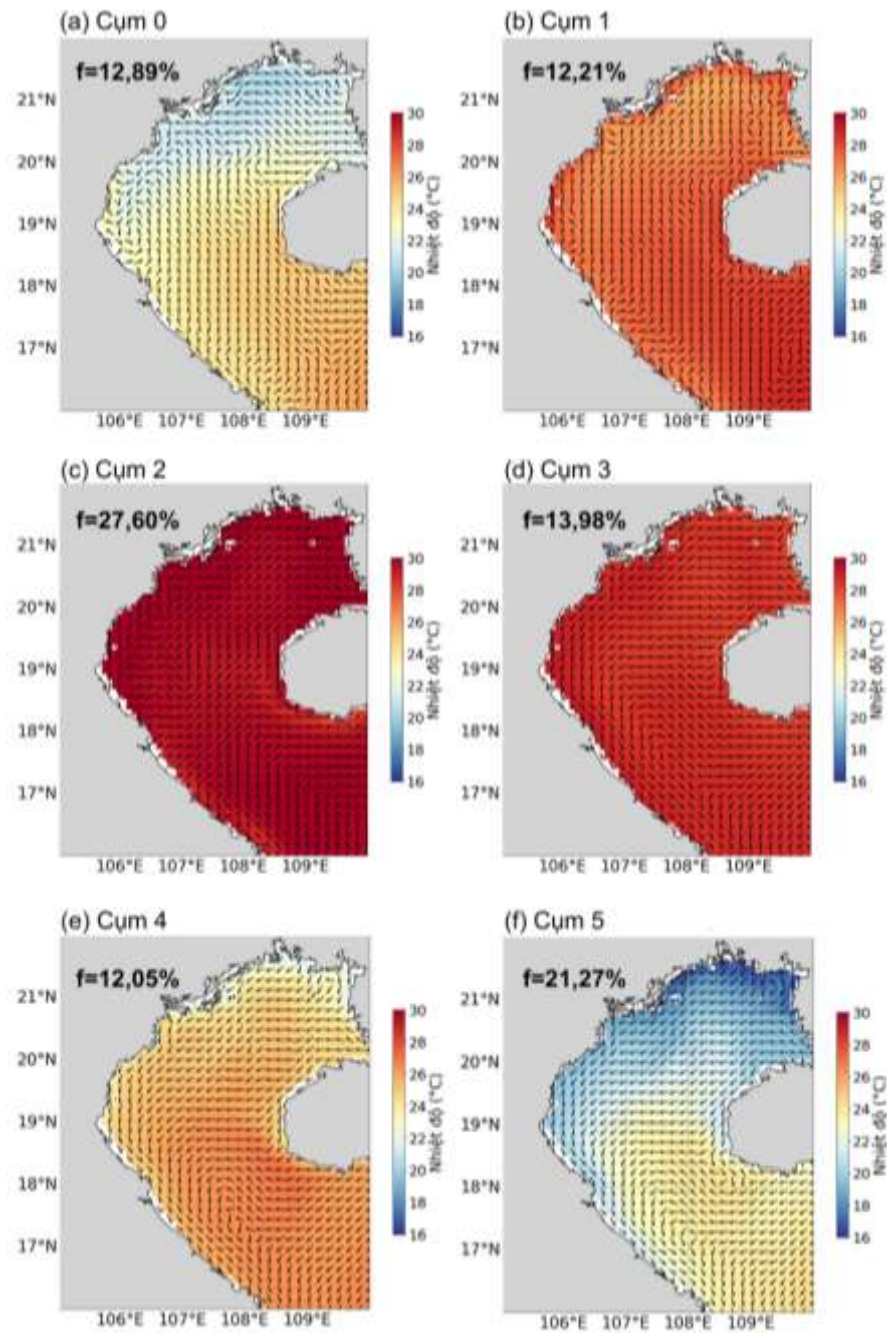
tăng cao và phân bố ổn định trên toàn Vịnh có tần suất xuất hiện chiếm đến 27,6%. Nhiệt độ trung bình mặt biển đạt khoảng 30 °C, bao phủ diện rộng và đi kèm với hướng dòng chảy Đông Bắc thịnh hành, thể hiện đặc trưng mùa hè.



Hình 4. Trường dòng chảy trung bình mặt biển gồm 6 cụm: (a)-Cúm 0, (b)-Cúm 1, (c)-Cúm 2, (d)-Cúm 3, (e)-Cúm 4, (f)-Cúm 5.



Hình 5. Trường độ muối kết hợp trường dòng chảy trung bình mặt biển gồm 6 cụm: (a)-Cù 0, (b)-Cù 1, (c)-Cù 2, (d)-Cù 3, (e)-Cù 4, (f)-Cù 5.



Hình 6. Trường nhiệt độ kết hợp trường dòng chảy trung bình mặt biển gồm 6 cụm: (a)-Cúm 0, (b)-Cúm 1, (c)-Cúm 2, (d)-Cúm 3, (e)-Cúm 4, (f)-Cúm 5.

Từ cụm 3 đến cụm 4 phản ánh giai đoạn chuyển tiếp của nền nhiệt mặt biển tại Vịnh Bắc Bộ. Ở cụm 3, nền nhiệt bắt đầu suy giảm nhẹ so với cụm 2, với nhiệt độ trung bình mặt biển duy

trì khoảng 29 °C và phân bố tương đối đồng đều trên toàn khu vực. Sang cụm 4, sự suy giảm về nền nhiệt độ trung bình mặt biển rõ rệt hơn được ghi nhận với tần suất xuất hiện chỉ 12,05%. Nhiệt

độ tại khu vực phía bắc vịnh giảm xuống 23–24 °C, trong khi phía nam vịnh duy trì mức cao nhất nhưng chỉ còn khoảng 25 °C. Sự thay đổi này thể hiện quá trình chuyển từ giai đoạn nhiệt độ cao sang giai đoạn giảm dần, với sự chênh lệch rõ giữa các khu vực bắc và phía nam vịnh.

Cụm 5 ghi nhận giai đoạn nhiệt độ mặt biển giảm xuống mức thấp nhất trong toàn bộ các cụm đặc trưng với tần suất xuất hiện 21,27% trên tổng thời gian dữ liệu. Khu vực ven bờ phía bắc vịnh có nhiệt độ trung bình khoảng 16–20 °C, trong khi phía còn lại của Vịnh duy trì mức 21–22 °C. Nhìn tổng thể, chuỗi cụm từ 0 đến 5 phản ánh sự thay đổi liên tục của nhiệt độ mặt biển tại Vịnh Bắc Bộ theo từng giai đoạn, với các giá trị cực trị và ranh giới nhiệt độ thể hiện rõ không gian. Sự phân bố này cho phép nhận diện xu thế tăng hoặc giảm nhiệt độ ở từng khu vực cụ thể, gắn liền với sự biến đổi tổng thể theo thời gian và không gian của toàn bộ khu vực nghiên cứu.

3.2. Mối quan hệ giữa các hình thái đặc trưng và ENSO

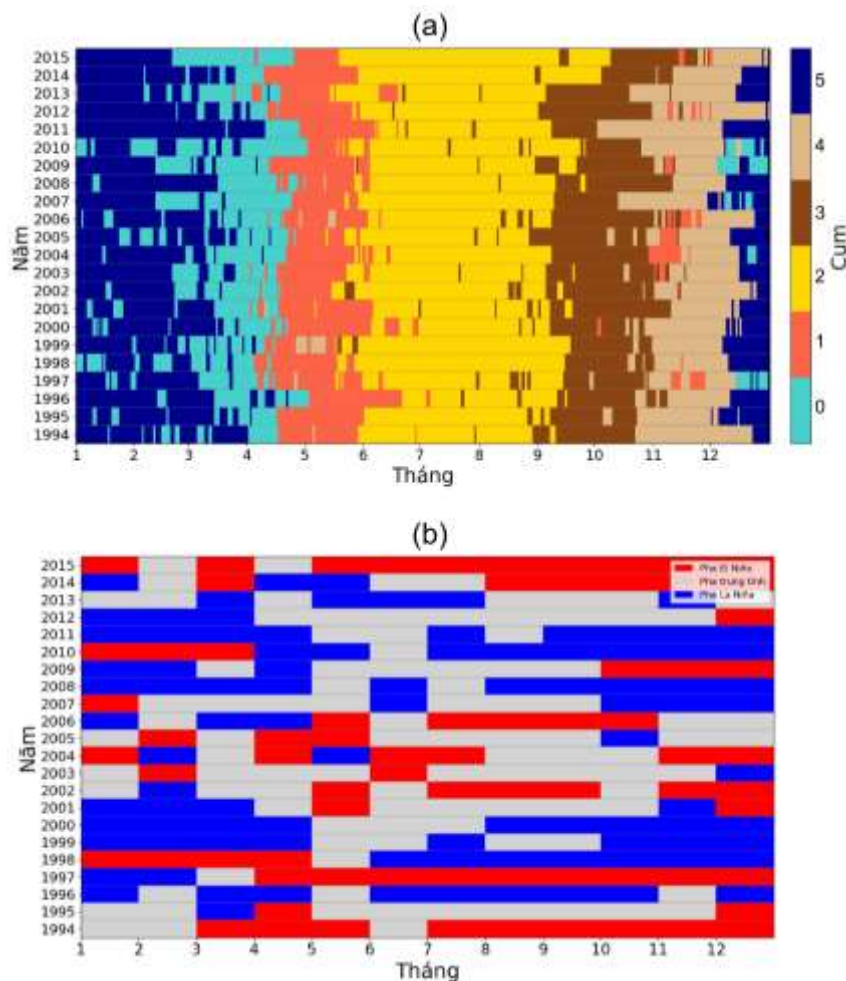
3.2.1. Phân bố các hình thái đặc trưng theo thời gian

Hình 7a cho thấy sự xuất hiện của các cụm có thời gian diễn ra khác nhau khác nhau tạo nên một chu kỳ rõ rệt trong năm tại khu vực Vịnh Bắc Bộ. Cụ thể, cụm 5 xuất hiện đầu tiên, kéo dài gần 4 tháng từ cuối mùa đông đến đầu mùa xuân (từ cuối tháng 12 đến giữa tháng 3 năm sau), trở thành cụm có thời gian tồn tại dài nhất trong toàn bộ chuỗi. Ngay sau đó, cụm 0 và cụm 1 lần lượt chiếm ưu thế trong giai đoạn mùa xuân, với thời gian tồn tại xấp xỉ nhau, chỉ khoảng hơn một tháng mỗi cụm (từ giữa tháng 3 đến cuối tháng 5). Bước sang mùa hè, cụm 2 thống trị khu vực Vịnh Bắc Bộ từ tháng 6 đến giữa tháng 9 với khoảng thời gian trung bình gần 3,5 tháng. Kế tiếp là cụm 3 xuất hiện vào mùa thu, kéo dài từ giữa tháng 9 đến cuối tháng 10. Cuối cùng, cụm 4 xuất hiện từ cuối tháng 10 đến đầu tháng 12, đánh dấu giai đoạn chuyển tiếp từ mùa thu sang đầu đông.

Kết hợp Hình 7b về chỉ số SOI xác định các pha ENSO với Hình 7a cho thấy phân bố cụm

trong năm cho thấy ENSO có những bất thường rõ rệt trong thời gian xuất hiện của các cụm khí hậu tại Vịnh Bắc Bộ. Trong các năm El Niño mạnh như 1997–1998 và 2009–2010, cụm 0 xuất hiện sớm vào cuối năm làm giảm đáng kể hoặc loại bỏ sự hiện diện của cụm 5 trong những tháng đầu năm sau. Ngược lại, La Niña 2010–2011 cho thấy cụm 5 kéo dài và ổn định hơn. Một số năm như 1996 và 2011 ghi nhận cụm 0 kéo dài đến tận đầu tháng 5 do giai đoạn chuyển từ La Niña sang pha trung tính. Bên cạnh đó, cụm 1 cũng có sự biến động mạnh khi các pha El Niño hoặc khi pha trung tính chuyển pha nhanh (1997–1998, 2001 và 2013) dẫn đến cụm 1 đến sớm hơn bình thường. Ngược lại, La Niña mạnh các năm 1996, 2000 và 2013 khiến cụm 5 kéo dài, lấn vào thời gian cụm 1 xuất hiện. Cụm 2 với thời gian tồn tại lâu nhất, đôi khi duy trì đến đầu tháng 10 như El Niño 2014–2015. Thêm vào đó, pha La Niña (2007, 2011) làm cụm 4 đến sớm, lấn át cụm 3, trong khi El Niño (1994, 2002, 2012, 2015) khiến cụm 4 kéo dài đến nửa cuối tháng 12. Đáng chú ý, cụm 1 đôi khi xuất hiện xen kẽ trong giai đoạn cụm 4 thịnh hành, như 1997, 2004, 2009 (El Niño) hoặc 2003, 2005, 2006, 2012 (trung tính). Những bất thường này cho thấy ENSO liên kết chặt chẽ thời gian xuất hiện và duy trì của các cụm đặc trưng vật lý biển.

Thời điểm xuất hiện của các cụm theo mùa phản ánh trực tiếp biến đổi của trường dòng chảy, độ muối và nhiệt độ tại Vịnh Bắc Bộ. Cụm 5 xuất hiện vào cuối đông – đầu xuân khi dòng chảy suy yếu trên hầu hết khu vực, độ muối ngoài khơi tăng cao đến 34 PSU, còn nhiệt độ giảm xuống mức thấp nhất 16–20 °C ở phía bắc Vịnh. Sang các cụm 0 và 1 trong giai đoạn mùa xuân, dòng chảy duy trì vận tốc thấp 0,1–0,3 m/s, độ muối ven bờ giảm dưới 31 PSU, trong khi nhiệt độ tăng dần từ 23 °C lên đến 27–28 °C. Đến mùa hè, cụm 2 thống trị với dòng chảy mạnh hơn (0,3–0,4 m/s), độ muối giảm trên toàn Vịnh và nhiệt độ đạt cực đại khoảng 30 °C. Cụm 3 và 4 trong giai đoạn mùa thu thể hiện dòng chảy đổi hướng Tây – Tây Nam với vận tốc 0,4–0,5 m/s, độ muối và nhiệt độ giảm dần so với mùa hè. Chu kỳ khép lại với cụm 5 khi nhiệt độ giảm sâu và dòng chảy suy yếu, báo hiệu sự chuyển sang mùa đông mới.



Hình 7. (a): Phân bố các hình thể cụm theo thời gian, (b): Chỉ số SOI xác định các pha ENSO.

3.2.2. Tần suất các cụm xuất hiện theo pha ENSO

Hình 8 thể hiện thời điểm và tần suất xuất hiện của các cụm đặc trưng theo từng pha ENSO. Trong các năm trung tính, cụm 0 có suất hiện chủ yếu từ tháng 1–5, trong khi ở pha El Niño, cụm này xuất hiện sớm hơn, từ tháng 12 năm trước và kết thúc vào tháng 4 năm sau, với tần suất đạt 0,6–0,8 vào tháng 12. Trong pha La Niña, tần suất tương tự nhưng chỉ khoảng 0,4–0,6 vào các tháng 1, 4 và 5.

Trong năm trung tính, cụm 1 xuất hiện từ tháng 4–6 và 10–11 với tần suất đạt cao nhất 0,6–0,8 vào tháng 10. Trong pha El Niño, cụm này đến muộn hơn bắt đầu vào tháng 11 và đặc

biệt có tần suất cao nhất 0,8–1,0 vào tháng 12. Tuy nhiên, cụm này có số lần xuất hiện rất thấp trong giai đoạn tháng 11 và 12 (Hình 7a), do đó, giá trị tần suất cao này chỉ phản ánh sự xuất hiện dị thường của cụm 1 trong giai đoạn này. Trong pha La Niña, sự xuất hiện của cụm 1 sớm hơn, kéo dài từ tháng 3–7, tần suất đạt cực đại 0,8–1,0 vào các tháng 3 và 7, rồi giảm trong tháng 10–11. Một cách tương tự, số lần xuất hiện của cụm 1 trong tháng 3 và tháng 7 rất thấp, do đó, tần suất cao cũng được xem là sự xuất hiện dị thường của cụm 1 trong giai đoạn này.

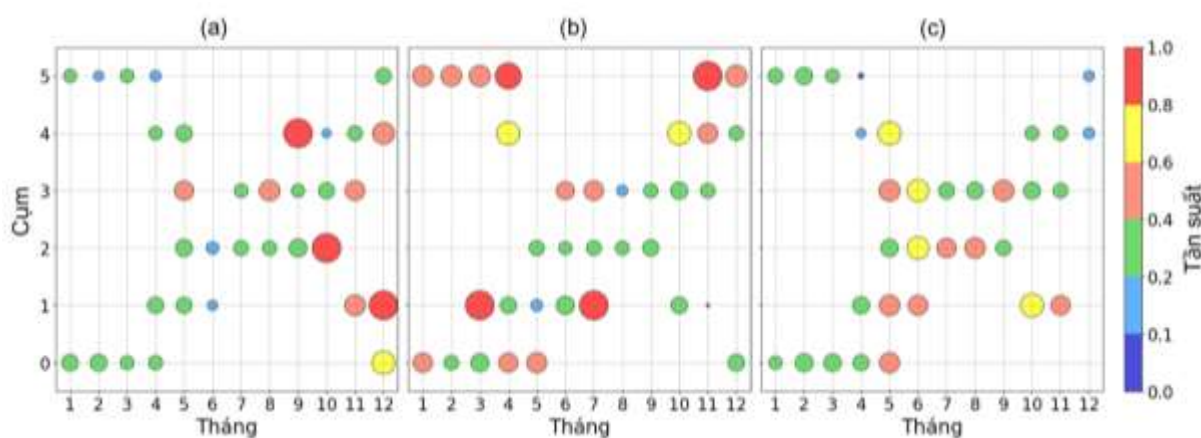
Pha trung tính, cụm 2 có tần suất từ tháng 5–9, đạt cao nhất 0,6–0,8 vào tháng 6. Trong pha El Niño, tần suất có giá trị tương tự nhưng thấp hơn, dao động trong khoảng 0,2–0,4. Trong pha

La Niña, cụm 2 kéo dài từ tháng 5–10, đạt cực đại 0,8–1,0 vào tháng 10.

Cụm 3, tần suất không có sự thay đổi đáng kể giữa các pha ENSO. Pha trung tính, cụm 3 có tần suất từ tháng 5–11, đạt cao nhất 0,6–0,8 vào tháng 6. Trong pha El Niño, tần suất xuất hiện vào các tháng 5 và 7–11, đạt 0,4–0,6 vào các tháng 5, 8 và 11. Trong pha La Niña, tần suất bắt đầu muộn hơn, từ tháng 6–11.

Cụm 4 có tần suất từ tháng 4–5 và 10–12, đạt cao nhất 0,6–0,8 vào tháng 5 trong năm trung tính. Trong pha El Niño, tần suất vào các tháng

4–5 và 9–12, đạt cực đại 0,8–1,0 vào tháng 9 và 0,4–0,6 vào tháng 12. Trong pha La Niña, cụm 4 có tần suất vào tháng 4, 10–12, đạt 0,6–0,8 trong tháng 4 và 10, và 0,4–0,6 vào tháng 11. Như vậy, tần suất của cụm 4 sớm hơn trong pha El Niño, khi dòng chảy hướng tây nam và nhiệt độ trung bình mặt biển bắt đầu giảm. Pha trung tính, cụm 5 có tần suất thấp khoảng 0,1–0,4 từ tháng 1–5 và tháng 12, tương tự như vậy trong pha El Niño. Trong pha La Niña, tần suất tăng rõ rệt, đạt 0,4–0,6 vào các tháng 12, 1–3, và cực đại 0,8–1,0 vào tháng 4 và 11.



Hình 8. Tần suất các cụm đặc trưng theo pha ENSO (a- Pha El Niño, b- Pha La Niña, c- Pha Trung tính). Diện tích bong bóng thể hiện tần suất của mỗi nút.

4. Kết luận

Trong nghiên cứu này, các thuật toán SOM và K-means đã được áp dụng để phân tích đặc trưng của các trường vật lý biển (dòng chảy, nhiệt độ, độ muối) tại Vịnh Bắc Bộ trong giai đoạn 1994–2015 và mối liên hệ của chúng với các pha ENSO. Các kết quả chính có thể tóm tắt như sau:

Thời điểm xuất hiện các cụm có sự khác biệt rõ rệt giữa pha El Niño và La Niña, trong đó El Niño thường tập trung mạnh vào cuối năm (tháng 9–12), trong khi La Niña chiếm ưu thế vào đầu và cuối năm. Sự thay đổi này phản ánh tác động điều tiết khí hậu mạnh mẽ của ENSO đối với khu vực. Về trường dòng chảy, nghiên cứu cho thấy sự biến đổi hướng và cường độ chịu ảnh hưởng đồng thời của chuyển mùa, lưu lượng sông và hải lưu từ Biển Đông. Các phân tích về

độ muối và nhiệt độ cũng làm rõ bức tranh đối lập: một bên là tác động bởi nước ngọt từ hệ thống sông Hồng – Thái Bình, bên còn lại chịu ảnh hưởng của khối nước mặn và ấm từ Biển Đông.

Phân tích phân cụm mang lại lợi thế so với các phương pháp truyền thống khi nghiên cứu biến động các trường vật lý biển. Thứ nhất, phương pháp này mô tả rõ hơn sự chuyển tiếp giữa các trạng thái khí hậu do ENSO chi phối. Thứ hai, nó cho phép giải thích đặc điểm không gian và thời gian của hệ thống khí hậu qua tần suất các cụm thay vì chỉ dựa vào các chỉ số tổng hợp. Việc áp dụng thuật toán SOM và K-Means cho phép phân tích dữ liệu theo ngày thay vì dựa trên giá trị trung bình tháng hay trung bình mùa như các nghiên cứu trước, nhờ đó không bỏ sót các dấu hiệu chuyển tiếp đặc trưng của dữ liệu.

Do đó, nhờ ưu thế trong việc nhận diện các đặc trưng không gian và thời gian, phương pháp phân cụm mở ra một hướng tiếp cận mới, giúp bổ sung thêm góc nhìn so sánh với các phương pháp truyền thống khi nghiên cứu đặc trưng các trường vật lý biển trong khu vực Vịnh Bắc Bộ.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Quỹ Đổi mới sáng tạo Vingroup (VINIF) trong Dự án mã số VinIF.2023.DA151.

Tài liệu tham khảo

- [1] D. V. Bat, Seafloor Topography and Geomorphological Characteristics of the Gulf of Tonkin, Vietnam, Marine Geology Division, 2024 (in Vietnamese).
- [2] D. Wu, Y. Wang, X. Lin, J. Yang, On the Mechanism of the Cyclonic Circulation in the Gulf of Tonkin in the Summer, *J. Geophys. Res. Oceans*, Vol. 113, No. C9, 2008, pp. 1-15, <https://doi.org/10.1029/2007JC004208>.
- [3] Y. Ding, C. Chen, R.C. Beardsley, X. Bao, M. Shi, Y. Zhang, Z. Lai, R. Li, H. Lin, N.T. Viet, Observational and Model Studies of the Circulation in the Gulf of Tonkin, South China Sea, *J. Geophys. Res. Oceans*, Vol. 118, No. 12, 2013, pp. 6495-6510, <https://doi.org/10.1002/2013JC009455>.
- [4] H. T. Huong, Determine the Temperature Structure of the Tonkin Gulf, *VNU Journal of Science: Earth and Environmental Sciences*, Vol. 32, No. 3S, 2016, pp. 109-115 (in Vietnamese).
- [5] P. Rogowski, J. Z. Garay, K. Shearman, E. Terrill, J. Wilkin, T. H. Lam, Air-Sea-Land Forcing in the Gulf of Tonkin, *Oceanogr*, Vol. 32, No. 2, 2019, pp. 150-161, <https://doi.org/10.5670/oceanogr.2019.223>.
- [6] J. Gao, H. Xue, F. Chai, M. Shi, Modeling the Circulation in the Gulf of Tonkin, South China Sea, *Ocean Dyn*, Vol. 63, No. 8, 2013, pp. 979-993, <https://doi.org/10.1007/s10236-013-0636-y>.
- [7] J. Gao, M. Shi, B. Chen, P. Guo, D. Zhao, Responses of the Circulation and Water Mass in the Beibu Gulf to the Seasonal Forcing Regimes, *Acta Ocean. Sin*, Vol. 33, No. 7, 2014, pp. 1-11, <https://doi.org/10.1007/s13131-014-0506-6>.
- [8] J. Gao, B. Chen, M. Shi, Summer Circulation Structure and Formation Mechanism in the Beibu Gulf, *Science China Earth Sciences*, Vol. 58, No. 2, 2015, pp. 286-299, <https://doi.org/10.1007/s11430-014-4916-2>.
- [9] J. Gao, G. Wu, H. Ya, Review of the Circulation in the Beibu Gulf, South China Sea, *Cont. Shelf Res*, Vol. 138, 2017, pp. 106-119, <http://dx.doi.org/10.1016/j.csr.2017.02.009>.
- [10] Y. Liu, R. H. Weisberg, C. N. Mooers, Performance Evaluation of the Self-organizing Map for Feature Extraction, *J. Geophys. Res. Oceans*, Vol. 111, No. C5, 2006, <https://doi.org/10.1029/2005JC003117>.
- [11] I. Vilibić, J. Sepic, H. Mihanović et al., Self-Organizing Maps-based Ocean Currents Forecasting System, *Sci. Rep.*, Vol. 6, No. 1, 2016, pp. 22924, <https://doi.org/10.1038/srep22924>.
- [12] S. Dey, R. Sikhakolli, D. P. Dogra, S. Sil, On the Variability of Ocean Surface Current in the Bay of Bengal using Self-organizing Map (SOM), *Deep Sea Res. Part I: Oceanographic Research Papers*, Vol. 199, 2023, pp. 104103, <https://doi.org/10.1016/j.dsr.2023.104103>.
- [13] P. Lemenkova, K-means Clustering in R Libraries and for Grouping Oceanographic Data, *International Journal of Informatics and Applied Mathematics*, Vol. 2, No. 1, 2019, pp. 1-26, <https://doi.org/10.6084/m9.figshare.9891203>.
- [14] Q. Sun, C. M. Little, A. M. Barthel, L. Padman, A Clustering-based Approach to Ocean Model-data Comparison Around Antarctica, *Ocean Sci.*, Vol. 17, No. 1, 2021, pp. 131-145, <https://doi.org/10.5194/os-2020-51>.
- [15] K. Saastamoinen, S. Penttinen, Visual Seabed Classification Using K-means Clustering, *CIELAB Colors and Gabor-filters*, *Procedia Computer Science*, 2021, <https://doi.org/10.1016/j.procs.2021.09.016>.
- [16] E. Romero, E. Portela, L. T. Fernandez, L. S. Velasco, Detection of Coherent Thermohaline Structures Over the Global Ocean Using Clustering, *Deep Sea Res. Part I: Oceanographic Research Papers*, Vol. 209, 2024, pp. 104344, <https://doi.org/10.1016/j.dsr.2024.104344>.
- [17] T. B. Minh, K. C. Nguyen, D. Q. Van, T. P. Van, T. Cong, N. T. Minh, An Application of Cluster Analysis in Investigating Characteristics of the South and Southeast Asian Monsoon Onset in ENSO Years, *SOLA*, Vol. 20, 2024, pp. 386-391, <https://doi.org/10.2151/sola.2024-051>.

- [18] A. Wallcraft, S. N. Carroll, K. A. Kelly, K. V. Rushing, Hybrid Coordinate Ocean Model (HYCOM) Version 2.1. User's Guide, 2003.
- [19] https://tds.hycom.org/thredds/catalogs/GLBv0.08/exprt_53.X.html (accessed on: September 10th, 2025).
- [20] A. Barnston, Why are there so Many ENSO Indexes, Instead of Just One, NOAA Climate, Gov, 2015.
- [21] T. Kohonen, Self-organized Formation of Topologically Correct Feature Maps, Biol. Cybern, Vol. 43, No. 1, 1982, pp. 59-69, <https://doi.org/10.1007/BF00337288>.
- [22] J. Vesanto, E. Alhoniemi, Clustering of the Self-Organizing Map, IEEE Transactions on Neural Networks, Vol. 11, No. 3, pp. 586-600, <https://doi.org/10.1109/72.846731>.
- [23] Y. Januzaj, E. Beqiri, A. Luma, Determining the Optimal Number of Clusters using Silhouette Score as a Data Mining Technique, International Journal of Online & Biomedical Engineering, Vol. 19, No. 4, 2023, <https://doi.org/10.3991/ijoe.v19i04.37059>.
- [24] G. Vardakas, L. Papakostas, A. Likas, Deep Clustering Using the Soft Silhouette Score: Towards Compact and Well-Separated Clusters, Towards Compact and Well-separated Clusters, arXiv preprint arXiv:2402.00608, 2024, <https://doi.org/10.48550/arXiv.2402.00608>.