



TUYỂN TẬP BÁO CÁO HỘI NGHỊ TOÀN QUỐC

KHOA HỌC TRÁI ĐẤT VÀ TÀI NGUYÊN VỚI PHÁT TRIỂN BỀN VỮNG

Hà Nội, 12 - 11 - 2020

ERSD 2020



NHÀ XUẤT BẢN GIAO THÔNG VẬN TẢI

MỤC LỤC

TIỂU BAN KỸ THUẬT DẦU KHÍ TÍCH HỢP

Nghiên cứu và thử nghiệm bơm ép diesel vào vùng nón nước nhằm giảm hàm lượng nước tại các giếng khai thác dầu nặng vỉa Mioxen trung, bể trầm tích Cửu Long <i>Nguyễn Hải An, Trần Anh Quân, Nguyễn Hoàng Đức, Nguyễn Thế Vinh, Nguyễn Văn Thịnh</i>	1
Đặc điểm hình thành gas hydrate, dự báo gradient địa nhiệt khu vực Đông Bắc vùng biển Quy Nhơn <i>Lê Ngọc Ánh</i>	8
Interpretation of mud diapirs using 2D seismic attributes and Unsupervised Neural Network: A case study of the Song Hong Basin <i>Le Ngọc Ánh, Ha Quang Man</i>	14
Nhận diện trầm tích gợn sóng đáy biển sâu trên tài liệu địa chấn 3D <i>Lê Ngọc Ánh</i>	21
Nghiên cứu việc nâng cấp bình tách hai pha thành bình tách ba pha trên giàn cố định mỏ Bạch Hổ <i>Hoàng Anh Dũng</i>	27
Phân loại thạch học từ các tham số vật lý trong tài liệu giếng khoan 1143, chương trình khoan đại dương tại biển Đông bằng mạng trí tuệ nhân tạo <i>Vũ Hồng Dương, Kiều Duy Thông</i>	32
Nghiên cứu lựa chọn hệ polymer bơm ép cho tầng Mioxen hạ mỏ Bạch Hổ <i>Lê Quang Duyên, Lê Đức Vinh, Lê Văn Nam, Phạm Trường Giang, Đinh Đức Huy</i>	39
Sử dụng phương pháp đường cong suy giảm dự báo khai thác các giếng tại cấu tạo Bò Câu Trắng <i>Lê Quang Duyên</i>	46
Ứng dụng mô hình địa cơ lựa chọn trọng lượng riêng cho các giếng khoan phức tạp tại bồn trũng Nam Côn Sơn <i>Lê Quang Duyên, Hoàng Thanh Tùng</i>	52
Nâng cao hiệu quả xác định các đới đập vỡ, nứt nẻ và hang karst bằng phương pháp biểu diễn 3D kết quả đo sâu điện 2D <i>Phạm Ngọc Đạt, Kiều Duy Thông, Phạm Ngọc Kiên, Lại Hợp Phòng, Dương Thị Ninh</i>	59
Nghiên cứu tạo nano đồng từ dung dịch $Cu(NO_3)_2$ bằng dịch chiết vỏ trái chuối hột <i>Nguyễn Huỳnh Đức, Nguyễn Thị Thanh Bảo, Nguyễn Thanh Bình, Hoàng Thu Hà, Tống Thị Thanh Hương</i>	65
Phân tích thuộc tính địa chấn nghiên cứu phân bố đá chứa Miocen dưới khu vực Đông Bắc mỏ Bạch Hổ <i>Nguyễn Minh Hòa, Nguyễn Duy Mười, Bùi Thị Ngân</i>	71
Đặc điểm trầm tích cacbonat qua phân tích tài liệu mẫu lõi và tài liệu địa vật lý giếng khoan khu vực đông nam bể Nam Côn Sơn <i>Nguyễn Thị Minh Hồng, Nguyễn Duy Mười, Phạm Thị Hồng</i>	76
Phương pháp minh giải kết quả nghiên cứu thủy động lực của giếng khoan ngang dầu khí <i>Đông Văn Hoàng, Nguyễn Minh Hòa</i>	82

Phân loại thạch học từ các tham số vật lý trong tài liệu giếng khoan 1143, chương trình khoan đại dương tại biển đông bằng mạng trí tuệ nhân tạo

Vũ Hồng Dương^{1,*}, Kiều Duy Thông¹

¹ Trường Đại học Mở - Địa chất

TÓM TẮT

Phân loại thạch học là một khâu quan trọng trong công tác nghiên cứu, phục hồi lịch sử phát triển các quá trình địa chất cũng như tìm kiếm thăm dò các khoáng sản có ích. Với các phương pháp truyền thống, thạch học được phân loại dựa theo kinh nghiệm của người minh giải, kết quả thường mang tính chủ quan và đòi hỏi nhiều thời gian. Do đó, việc sử dụng các kỹ thuật trí tuệ nhân tạo để thực hiện công việc này sẽ cho kết quả nhanh và mang tính khách quan hơn. Chúng tôi áp dụng các kỹ thuật trí tuệ nhân tạo khác nhau trên bộ tài liệu của giếng khoan 1143, chương trình khoan đại dương (ODP) tại Biển Đông. Các phương pháp học không giám sát phân tích phân cụm mờ c-trung bình (FCM). Các phương pháp học có giám sát, mạng Ne-ron nhân tạo để dự đoán thạch học. Các kết quả đạt được sẽ được so sánh với nhau để xác định kỹ thuật phù hợp cho tập số liệu này và có thể áp dụng cho vùng lân cận tại Biển Đông.

Từ khóa: Thạch Học; Trí tuệ nhân tạo; địa chất.

1. Giới thiệu

Thông tin về cấu trúc địa chất dưới biển, như khu vực biển đông thường được cung cấp bởi các kết quả xử lý và phân tích các tài liệu địa vật lý (Braitenberg et al., 2006; Ding et al., 2013; Ding et al., 2016; Gozzard et al., 2018; Huang and Wang, 2006; Trung et al., 2004). Các thông tin về địa chất được minh giải từ các tài liệu địa vật lý biển thường phải được đối sánh với các tài liệu trong giếng khoan, do vậy phân tích các tài liệu địa vật lý trong giếng khoan để đưa ra những thông tin hữu ích về địa chất là rất quan trọng (Wu et al., 2017). Tuy nhiên công việc này thường tốn nhiều thời gian và trong nhiều trường hợp mang tính chủ quan và thiếu định lượng. Để vượt qua các thách thức này, hiện nay các công cụ tính toán như học máy (machine learning) và trí tuệ nhân tạo (AI) đang được sử dụng ngày càng nhiều trong các lĩnh vực xử lý số liệu, trong đó có xử lý tài liệu địa vật lý. Cụ thể việc áp dụng các công cụ này cho việc phân tích các số liệu giếng khoan đã được công bố bởi các công trình của Bosch et al. (2013); Dekkers et al. (2014), trong khu vực Biển Đông cũng có các công trình của Karmakar et al. (2018); Tse et al. (2019). Các kết quả đạt được của các tác giả này cho thấy hiệu quả của việc áp dụng các công cụ tính toán này. Do vậy, trong nghiên cứu này chúng tôi áp dụng các thuật toán học máy và AI vào trong xử lý tài liệu giếng khoan 184-1143 A, B, và C (Hình 1).

Phân loại thạch học dựa vào tài liệu địa vật lý giếng khoan và mẫu lõi là một trong những công tác quan trọng. Việc xác định tương thạch học ở những khu vực có cấu trúc địa chất phức tạp, nếu chỉ sử dụng những phương pháp minh giải địa vật lý giếng khoan thông thường đem lại kết quả không khả quan và chính xác khi so sánh với tài liệu mẫu lõi thu thập được. Ví dụ cụ thể như ở khu vực Biển Đông nơi có rất nhiều các tàn tích núi lửa với những dạng thù hình, việc minh giải và xác định thạch học cho khu vực này cực kỳ khó khăn khi chỉ sử dụng các phương pháp minh giải dựa trên các đường log cơ bản của địa vật lý giếng khoan.

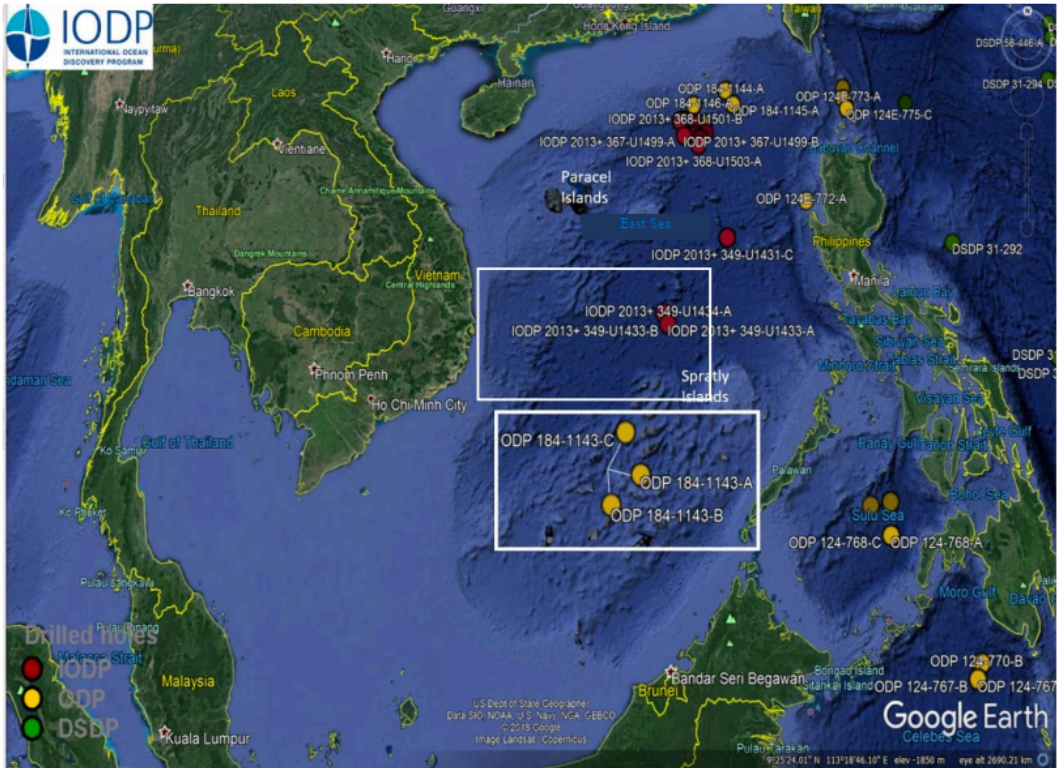
2. Nguồn số liệu

Trong nghiên cứu này chúng tôi sử dụng nguồn số liệu từ các chương trình khám phá đại dương quốc tế và khoan sâu đại dương (<http://web.iodp.tamu.edu/>). Tại vị trí này có 3 giếng khoan OPD-184-1143 A, B, và C (Hình 1), để ngắn gọn trong bài viết này chúng tôi đặt tên là GKA, GKB và GKC.

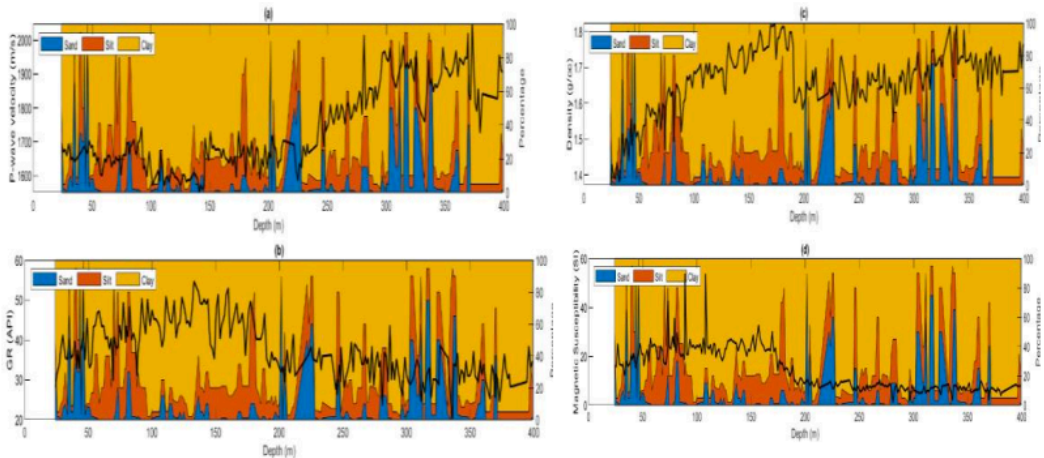
Số liệu bao gồm các tham số vật lý và thạch học đo trong giếng khoan và từ mẫu vận tốc sóng dọc (V_p), gamma ray, mật độ, độ từ cảm, tỉ lệ phần trăm các thành phần sand, silt và clay trong mẫu.

* Tác giả liên hệ

Email: vuhongduong@humg.edu.vn



Hình 1. Vị trí của ba giếng khoan OPD-184-1143 A, B, và C (đánh dấu trong khung màu trắng)



Hình 2. Phân bố của các mẫu theo tỉ lệ của sand, silt và clay theo chiều sâu trong GKA. Đường màu đen thể hiện các tham số vật lý vận tốc sóng dọc (a), gamma ray (b), mật độ (c) và độ từ cảm (d)

3. Dự đoán thạch học từ các tham số vật lý

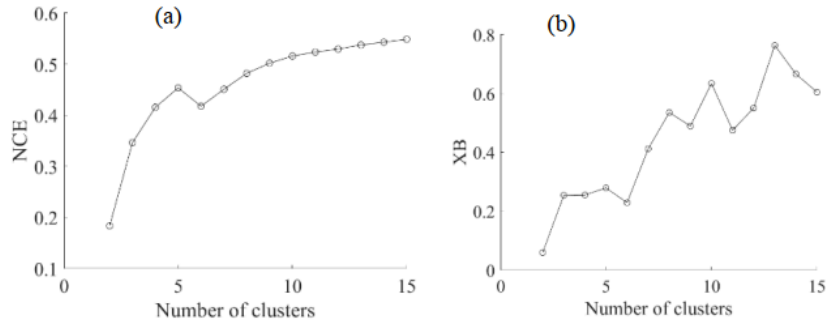
a. Phân cụm mờ c-trung bình (FCM)

Thuật toán fuzzy c-means (Bezdek et al., 1984) chia các phần tử của số liệu đầu vào thành các nhóm dựa vào sự tương đồng của các tính chất của các phần tử. Khác với phương pháp K trung bình mỗi phần tử chỉ thuộc về một nhóm, trong phương pháp này các phần tử có thể thuộc về nhiều nhóm với hệ số thành viên khác nhau. Quá trình này được thực hiện bằng việc cực tiểu hóa hàm sau:

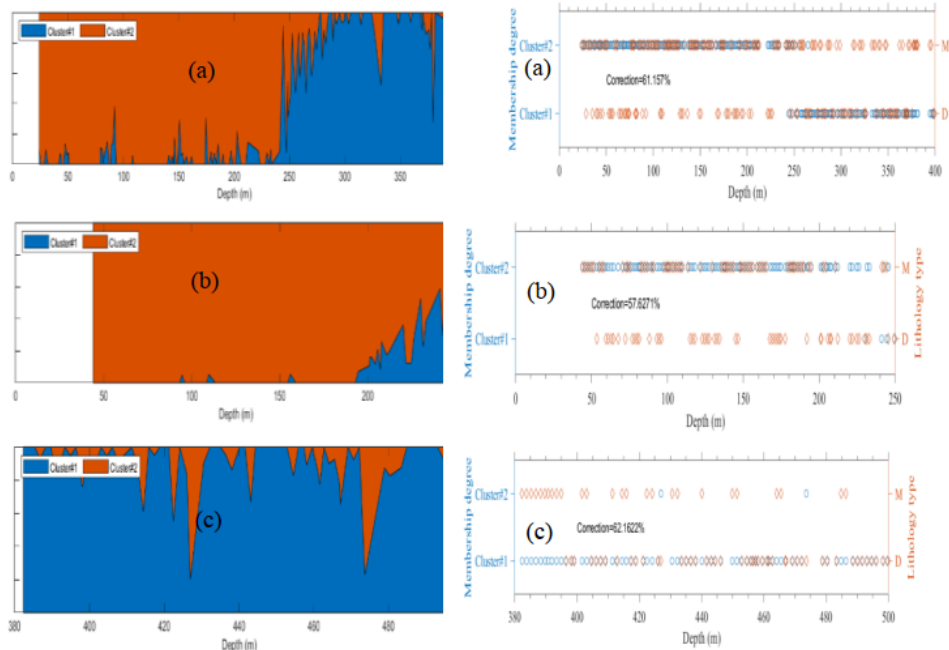
$$\Phi_{FCM} = \sum_{j=1}^N \sum_{k=1}^C u_{jk}^q \|z_j - v_k\|_2^2, \quad (1)$$

trong đó, N là số lượng phần tử của số liệu z , z có thể gồm nhiều thuộc tính. C là số lượng nhóm, q ($q > 1$) là tham số 'mờ' (fuzziness), q thường được lấy giá trị bằng 2 (Bezdek et al., 1984), v_k là giá trị trung tâm của nhóm thứ k , u_{jk} là hệ số thành viên (membership degree) của phần tử thứ j đối với nhóm thứ k , với điều kiện $\sum_{k=1}^C u_{jk} = 1$.

Một trong nhiệm vụ quan trọng nhất trong phân tích cụm mờ trung bình là xác định số cụm. Hiện nay, số cụm thường được xác định dựa vào một số chỉ số. Trong nghiên cứu này chúng tôi sử dụng hai chỉ số NCE (Roubens, 1982) và XB (Xie and Beni, 1991). Hai chỉ số này càng nhỏ chứng tỏ quá trình phân tích cụm càng tốt. Kết quả phân tích xác định rằng số cụm bằng hai là tối ưu cho tập số liệu này (Hình 3).



Hình 3. Các chỉ số NCE và XB sử dụng để xác định số cụm tối ưu



Hình 4. Hệ số thành viên của phân tích cụm mờ c - trung bình của các giếng GKA (a), GKB (b) và GK (c). Có thể thấy rõ ranh giới phân chia giữa phần trên và dưới trong giếng khoảng GKA

Hình 5. So sánh kết quả phân chia thạch học dựa vào phân tích cụm mờ c - trung bình và mô tả mẫu lõi trong các giếng GKA (a), GKB (b) và GK (c) với độ chính xác lần lượt là 61.57%; 57.62% và 62.16%

Kết quả phân tích cho thấy, nhóm 1 (Cluster#1) tương ứng với loại thạch học D, và nhóm 2 (Cluster#2) tương ứng với loại thạch học M. Độ chính xác so với mẫu lõi nằm trong khoảng 60%, tỉ lệ này không quá cao nhưng có thể chấp nhận được do các tham số thạch học và các tham số vật lý có mức độ tương quan thấp, như đã phân tích ở phần trên. Hơn nữa, phân loại mẫu lõi mang tính chủ quan, và có thể có sai số, bằng chứng là so với tỉ lệ clay thì vẫn có rất nhiều điểm nằm lẫn trong hai vùng của hai loại thạch học.

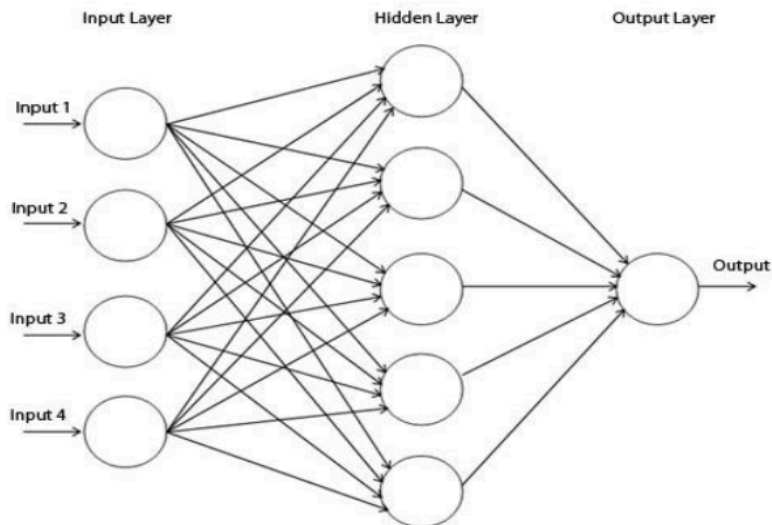
b. Sử dụng mạng Neuron nhân tạo để dự đoán thạch học

Dựa trên tài liệu địa chất khu vực, nhóm tác giả nhận định cột địa tầng khu vực này chủ yếu là các lớp cát và sét xen kẽ, do đó thông số hàm lượng sét (Vsh) sẽ được sử dụng làm tiêu chuẩn để phân chia ra các tướng thạch học cơ bản, cụ thể ở đây là 3 nhóm:

- i. Nhóm cát sạch: Các lớp đất đá có hàm lượng sét $Vsh < 20\%$
- ii. Nhóm cát pha sét: Các lớp đất đá có hàm lượng sét $20\% < Vsh < 80\%$
- iii. Nhóm sét: Các lớp đất sét có hàm lượng $Vsh > 80\%$

Nhóm tác giả quyết định sử dụng mạng neuron lan truyền ngược (back-propagation neural network) để dự báo hàm lượng sét dựa trên tài liệu mẫu lõi của 3 giếng khoan với dữ liệu đầu vào input layer gồm 4 neuron là các tham số thạch học thu được từ mẫu lõi: Vận tốc sóng dọc Pwave, Cường độ phóng xạ tự nhiên GR, Mật độ đất đá Density và Độ từ cảm Magnetic Suscept; lớp ẩn của mạng gồm 5 neuron và lớp đầu ra gồm 1 neuron chính là giá trị Vsh cần dự báo (Hình 6). Tài liệu mẫu lõi của 2 giếng A và B sẽ được đưa vào phục vụ quá trình đào tạo mạng, sau đó sẽ sử dụng mạng neuron thu được để dự báo hàm lượng sét của giếng C. Trong quá trình huấn luyện, tập dữ liệu mẫu sử dụng bao gồm tài liệu mẫu lõi của 2 giếng A, B gồm cả giá trị đầu vào input data (Pwave, GR, Dens, Magnetic) lẫn giá trị đầu ra output (Vsh), các dữ liệu sẽ được tính toán và cho ra kết quả output data là hàm lượng sét Vsh, kết quả này sẽ được so sánh với tài liệu mẫu lõi 2 giếng A, B để rút ra sai số. Sai số này được lan truyền ngược lại tới output data và các neuron trong lớp ẩn để điều chỉnh lại giá trị các trọng số. Quá trình lan truyền ngược này lặp đi lặp lại cho đến khi sai số đạt tới giá trị cực tiểu thỏa mãn giá trị nhóm tác giả cho phép, đây là quá trình học có giám sát. Các trọng số được xác định sao cho sai số E là cực tiểu:

$$E_k = \frac{1}{2} \sum_{k=1}^{n_0} \left(y_{i,k} - \hat{y}_{i,k} \right)^2$$



Hình 6. Cấu trúc mạng neuron dự báo hàm lượng sét cho 3 giếng GKA, GKB, và GKC

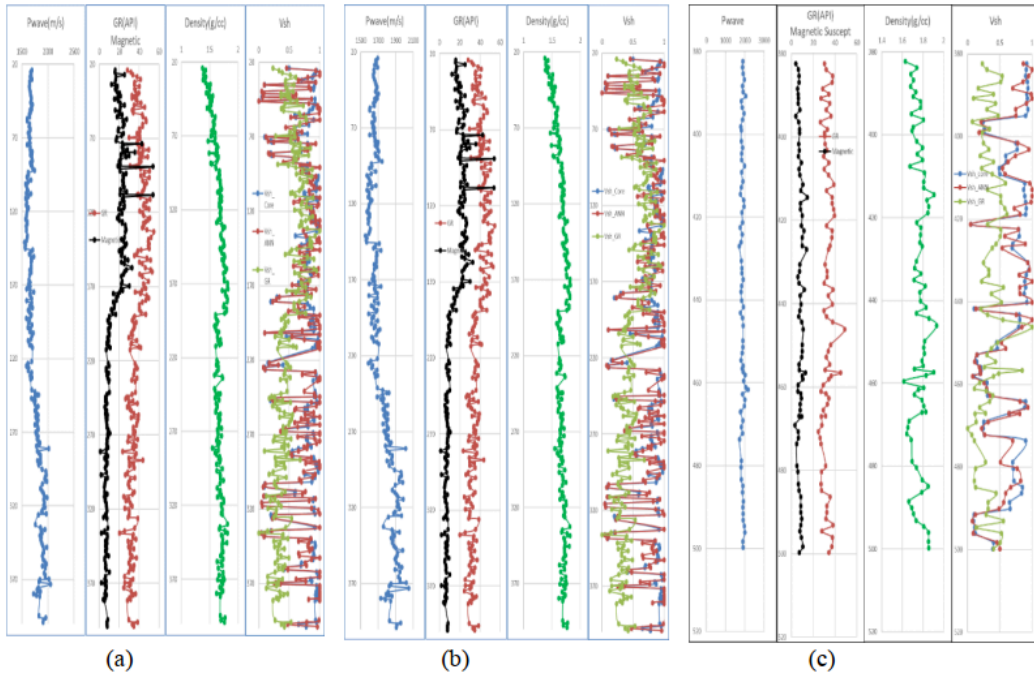
Kết quả dự báo hàm lượng sét của 3 giếng khoan A, B, C được biểu diễn cùng với hàm lượng sét được tính theo đường cong GR thông thường và giá trị hàm lượng sét đo được trên mẫu lõi của 3 giếng GKA, GKB, và GKC (Hình 7).

Độ chính xác của mô hình mạng neuron dự báo hàm lượng sét phụ thuộc vào các yếu tố: Cách trung bình hóa các đường cong ĐVLGK; tính chính xác trong việc dịch chuyển chiều sâu giữa mẫu lõi và tài liệu ĐVLGK; số lượng mẫu lõi càng nhiều thì khả năng xây dựng mạng neuron cho kết quả chính xác càng cao; độ tin cậy của kết quả đo ghi mẫu lõi; hàm lượng sét mẫu được đo trong phòng thí nghiệm còn các đường cong ĐVLGK lại đo trong điều kiện vỉa; cấu trúc mạng neuron; thuật toán.

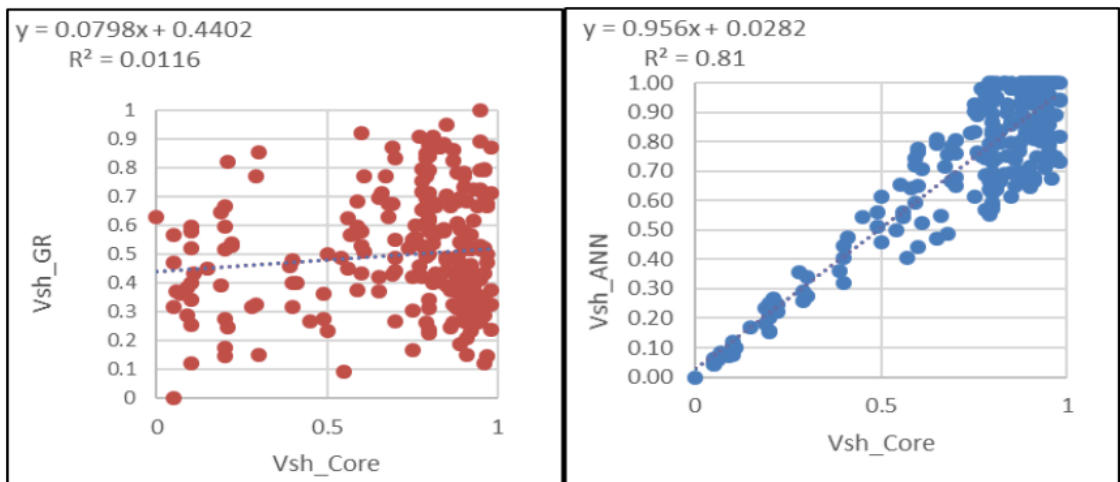
Có thể thấy hàm lượng sét dự báo bằng mạng ANN phù hợp với hàm lượng sét thu được từ tài liệu mẫu lõi hơn với kết quả tính toán thông thường chỉ dựa vào đường GR (Hình 14, 15, 16)

4. Kết luận

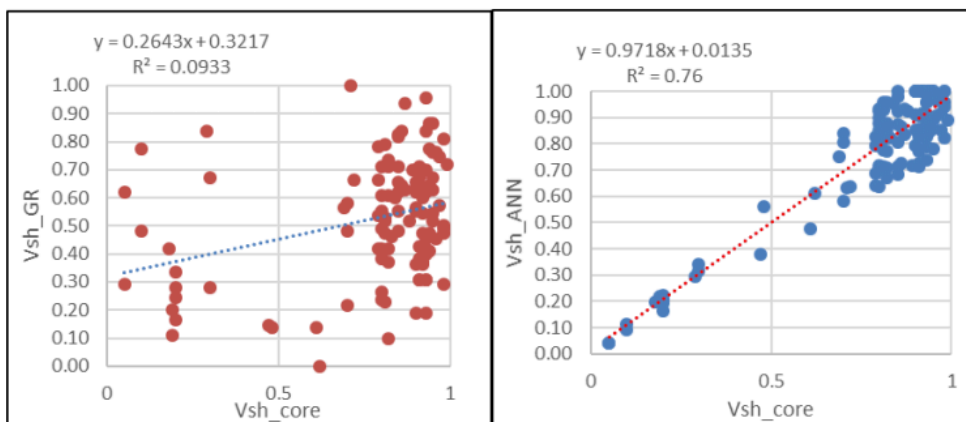
Trong nghiên cứu này chúng tôi sử dụng các thuật toán học có giám sát và không có giám sát để dự báo thạch học từ các tham số vật lý của đất đá. Tham số hàm lượng sét (Vsh) được sử dụng rất hiệu quả để làm tiêu chuẩn phân chia ra các tướng thạch học cơ bản, cụ thể ở đây là 3 nhóm: cát, sét và cát pha sét. Kết quả phân loại thạch học dựa trên các thuật toán học có giám sát và mạng ANN là chính xác hơn khi so với phương pháp truyền thống. Các kết quả bước đầu là khá quan, tuy nhiên để nâng cao độ chính xác nhóm tác giả đề xuất cần có các nghiên cứu kỹ hơn.



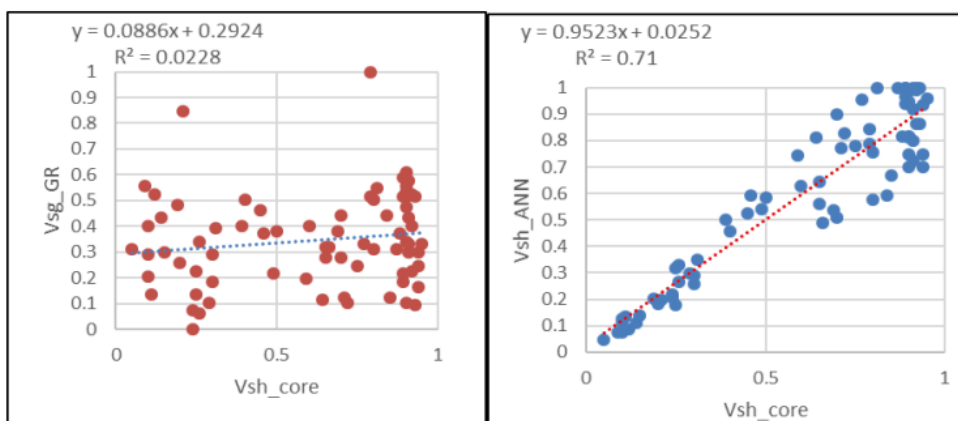
Hình 13. Kết quả dự báo từ mạng ANN, tính từ GR và đo được từ mẫu lõi giếng GKA(a), GKB(b), GK(C)



Hình 14. Hệ số tương quan giữa hàm lượng sét đo được của mẫu lõi với hàm lượng sét tính từ GR và dự báo từ mạng ANN của GKA



Hình 15. Hệ số tương quan giữa hàm lượng sét đo được của mẫu lõi với hàm lượng sét tính từ GR và dự báo từ mạng ANN giếng GKB



Hình 16. Hệ số tương quan giữa hàm lượng sét đo được của mẫu lõi với hàm lượng sét tính từ GR và dự báo từ mạng ANN giếng GKB

Tài liệu tham khảo

- Bezdek, J. C., Ehrlich, R., and Full, W., 1984, FCM: The fuzzy c-means clustering algorithm: Computers & Geosciences, v. 10, no. 2-3, p. 191-203.
- Bosch, D., Ledo, J., and Queralt, P., 2013, Fuzzy Logic Determination of Lithologies from Well Log Data: Application to the KTB Project Data set (Germany): Surveys in Geophysics, v. 34, no. 4, p. 413-439.
- Braitenberg, C., Wienecke, S., and Wang, Y., 2006, Basement structures from satellite-derived gravity field: South China Sea ridge: Journal of Geophysical Research: Solid Earth, v. 111, no. B5.
- Dekkers, M. J., Heslop, D., Herrero-Bervera, E., Acton, G., and Krasa, D., 2014, Insights into magmatic processes and hydrothermal alteration of in situ superfast spreading ocean crust at ODP/IODP site 1256 from a cluster analysis of rock magnetic properties: Geochemistry, Geophysics, Geosystems, v. 15, no. 8, p. 3430-3447.
- Ding, W., Franke, D., Li, J., and Steuer, S., 2013, Seismic stratigraphy and tectonic structure from a composite multi-channel seismic profile across the entire Dangerous Grounds, South China Sea: Tectonophysics, v. 582, p. 162-176.
- Ding, W., Li, J., and Clift, P. D., 2016, Spreading dynamics and sedimentary process of the Southwest Sub-basin, South China Sea: Constraints from multi-channel seismic data and IODP Expedition 349: Journal of Asian Earth Sciences, v. 115, p. 97-113.
- Gozzard, S., Kuszniir, N., Franke, D., Cullen, A., Reemst, P., and Henstra, G., 2018, South China Sea crustal thickness and oceanic lithosphere distribution from satellite gravity inversion: Petroleum Geoscience, p. petgeo 2016-2162.

Huang, W., and Wang, P., 2006, Sediment mass and distribution in the South China Sea since the Oligocene: *Science in China Series D: Earth Sciences*, v. 49, no. 11, p. 1147-1155.

Karmakar, M., Maiti, S., Singh, A., Ojha, M., and Maity, B. S. J. M. G. R., 2018, Mapping of rock types using a joint approach by combining the multivariate statistics, self-organizing map and Bayesian neural networks: an example from IODP 323 site, v. 39, no. 3, p. 407-419.

Roubens, M., 1982, Fuzzy clustering algorithms and their cluster validity: *European Journal of Operational Research*, v. 10, no. 3, p. 294-301.

Trung, N. N., Lee, S.-M., and Que, B. C., 2004, Satellite Gravity Anomalies and Their Correlation with the Major Tectonic Features in the South China Sea: *Gondwana Research*, v. 7, no. 2, p. 407-424.

Tse, K. C., Chiu, H.-C., Tsang, M.-Y., Li, Y., and Lam, E. Y. J. F. o. E. S., 2019, Unsupervised learning on scientific ocean drilling datasets from the South China Sea, v. 13, no. 1, p. 180-190.

Wu, H., Shi, M., Zhao, X., Huang, B., Zhang, S., Li, H., Yang, T., and Lin, C., 2017, Magnetostratigraphy of ODP Site 1143 in the South China Sea since the Early Pliocene: *Marine Geology*, v. 394, p. 133-142.

Xie, X. L., and Beni, G., 1991, A validity measure for fuzzy clustering: *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 13, no. 8, p. 841-847.

ABSTRACT

Classification of lithology from well log data using machine learning and artificial intelligence techniques: an example from odp 1143 site

Hong Duong Vu¹, Duy Thong Kieu¹
¹*Hanoi University of Mining and Geology*

Lithological classification is an important stage in the research, restoration of the history of geological processes development as well as exploration of useful minerals. Lithology is usually classified based on the experience of the interpreter, the results are often subjective and time consuming. Therefore, the use of artificial intelligence techniques to perform this task will result in faster and more objective. We apply various artificial intelligence techniques on the 1143 data set, the ocean drilling program (ODP) in the South China Sea. Un-supervised methodologies c-mean fuzzy (FCM) clustering analysis. Supervised learning methods, artificial neural network to predict lithology. The results obtained will be compared with each other to determine the appropriate technique for this dataset and can be applied to the vicinity in the South China Sea.

Keywords: Lithology, artificial intelligence, geology.