

ỨNG DỤNG HỌC MÁY ĐỂ DỰ BÁO SỰ THAY ĐỔI CỦA TỶ SỐ CONDENSATE KHÍ THEO THỜI GIAN NHẪM PHỤC VỤ VIỆC LÊN KẾ HOẠCH VÀ QUẢN LÝ MỎ KHÍ CONDENSATE

Ngô Hữu Hải¹, Trịnh Xuân Vinh¹, Nguyễn Ngọc Tân¹, Hoàng Kỳ Sơn¹, Ngô Tuấn Anh¹
 Trần Ngọc Trung¹, Trần Vũ Tùng¹, Nguyễn Sỹ Tuấn²

¹Công ty Điều hành Dầu khí Biển Đông

²Trường Đại học Văn Lang

Email: sonhk@biendongpoc.vn

<https://doi.org/10.47800/PVSI.2024.01-07>

Tóm tắt

Một trong những thông số quan trọng nhất để đánh giá, dự báo và quản lý các vỉa khí condensate là tỷ số condensate khí của giếng theo thời gian khai thác. Thông số này có xu hướng giảm dần do áp suất vỉa suy giảm trong quá trình khai thác. Với cách tiếp cận truyền thống, mẫu chất lưu khí và condensate được lấy tại thời điểm ban đầu cũng như định kỳ trong quá trình khai thác để thực hiện phân tích thành phần chất lưu, xác định tỷ số condensate khí và phân tích tính chất chất lưu (thí nghiệm PVT). Tuy nhiên, việc lấy mẫu, vận chuyển và phân tích mẫu mất nhiều thời gian cũng như chi phí. Ngoài cách tiếp cận trên, phương pháp mô hình hóa thành phần và động thái chất lưu trong vỉa cũng thường được áp dụng. Phương pháp này chứa đựng nhiều yếu tố rủi ro vì phụ thuộc nhiều vào các thông số đầu vào giả định bao gồm cấu trúc vỉa, tương tác giữa các pha và các thông số tính chất khác của vỉa chứa. Do đó, ứng dụng học máy vào việc dự báo sự thay đổi của tỷ số condensate khí của giếng theo thời gian trong nghiên cứu này là một hướng đi mới hiệu quả để hỗ trợ cho các phương pháp truyền thống.

Từ khóa: Học máy, tỷ số condensate khí, dự báo khai thác.

1. Giới thiệu

Việc dự báo tỷ số condensate khí của giếng có ý nghĩa rất quan trọng trong việc quản lý khai thác các vỉa khí condensate. Whitson và các cộng sự đã chứng minh rằng, điểm khác biệt chính đối với vỉa khí condensate so với vỉa khí khô là cần phải dự báo được tỷ số condensate khí, từ đó tính toán được lượng condensate sẽ được khai thác lên bề mặt [1]. Do vậy, thông số này rất quan trọng và cần phải có dự báo đáng tin cậy.

Phương pháp truyền thống để dự báo tỷ số condensate khí qua quá trình suy giảm áp suất dựa vào mô hình chất lưu PVT (pressure - volume - temperature). Để xây dựng được mô hình này cần phải thu thập được mẫu chất lưu đại diện, sau đó thực hiện các thí nghiệm PVT để xác định các thông số tính chất chính của chất lưu. Quá trình này tốn kém nhiều chi phí và thời gian thực hiện. Sau khi đã có các thông số

tính chất chất lưu, mô hình phương trình trạng thái EOS (equation of state) sẽ được xây dựng để mô phỏng các đặc tính này của chất lưu trong vỉa. Nếu không có mẫu đại diện cùng các thông số tính chất chất lưu thu thập được qua quá trình phân tích mẫu, sẽ không thể xây dựng được mô hình trạng thái, còn nếu sử dụng tính chất chất lưu của vỉa khác thì độ chính xác của mô hình sẽ không đảm bảo.

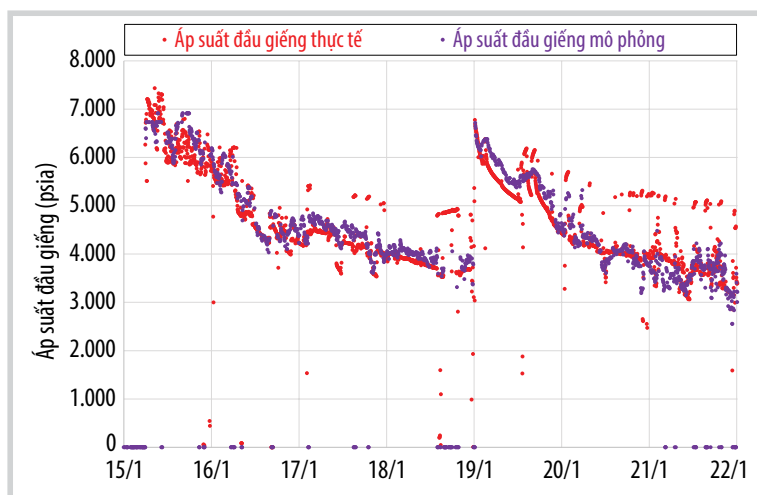
Một phương pháp truyền thống khác là sử dụng mô hình động. Tuy nhiên, với hầu hết các giếng thuộc mỏ Hải Thạch, việc dự báo tỷ số condensate khí bằng phương pháp này khá khó khăn dù kết quả khớp lịch sử khai thác của áp suất đầu giếng là tốt như trong Hình 1 và 2.

Có nhiều lý do khiến việc khớp lịch sử cho tỷ số condensate khí của các giếng Hải Thạch gặp nhiều khó khăn. Thứ nhất, mỏ Hải Thạch có 9 vỉa riêng biệt nhưng chỉ lấy được mẫu đại diện của 1 vỉa nên không có đủ dữ liệu để xây dựng mô hình tính chất chất lưu cho tất cả các vỉa cần phải mô phỏng. Đối với các vỉa không có mẫu

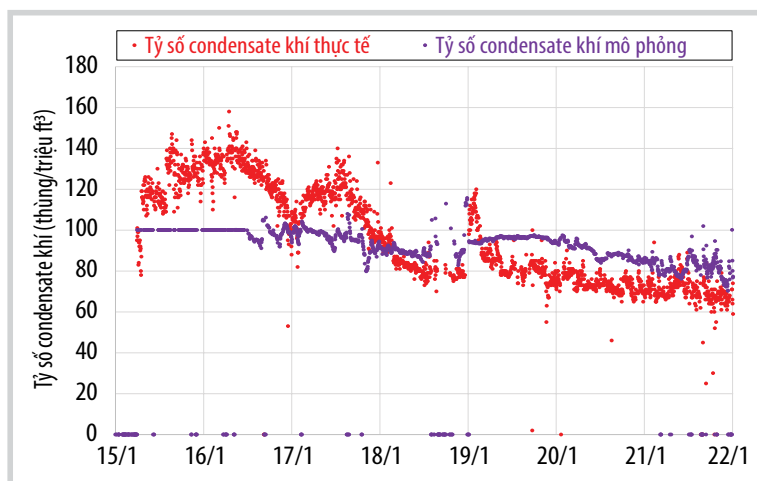


Ngày nhận bài: 21/7/2023. Ngày phản biện đánh giá và sửa chữa: 2/8/2023 - 23/1/2024.

Ngày bài báo được duyệt đăng: 23/1/2024.



Hình 1. Kết quả khớp lịch sử khá tốt cho áp suất đầu giếng HT-Y bằng mô hình động.



Hình 2. Kết quả khớp lịch sử cho tỷ số condensate khí của giếng HT-Y bằng mô hình động còn nhiều khó khăn.

đại diện, phải giả sử chúng có tính chất tương đồng với vỉa có mẫu đại diện, nên mô hình tính chất chất lưu cho các vỉa này sẽ có nhiều yếu tố không chắc chắn. Thứ hai, nhiều giếng Hải Thạch khai thác đồng thời từ nhiều tập vỉa và tỷ lệ đóng góp của các vỉa này thay đổi theo thời gian. Do đó việc khớp lịch sử cho tỷ số condensate khí bằng mô hình động là hết sức khó khăn và có sai lệch nhất định so với số liệu thực tế, dẫn đến việc dự báo tỷ số này trong tương lai cũng gặp nhiều khó khăn. Đối với dự báo ngắn hạn, sai số này gây ảnh hưởng lên các chuyển xuất bán condensate vì lượng khai thác dự báo được có thể ít hơn hoặc nhiều hơn so với thực tế, dẫn đến việc lập kế hoạch bán condensate có nhiều rủi ro với nguy cơ thiếu hàng hoặc lượng khai thác vượt quá khả năng của tàu chứa dẫn đến phải dừng giàn khai thác. Đối với dự báo dài hạn, sai số trong dự báo tỷ số condensate khí dẫn đến sản lượng condensate và tổng thu hồi condensate cho từng giếng bị sai lệch, ảnh hưởng đến kết quả tính toán về hiệu quả thương mại của dự án. Do đó, việc xây dựng một phương pháp hỗ trợ để dự báo tỷ số condensate khí là hết sức cấp thiết.

Ngày nay, học máy đã được sử dụng rộng rãi để giải quyết nhiều bài toán dự báo. Ứng dụng của phương pháp này vào việc dự báo

tính chất chất lưu của vỉa dầu khí cũng đã được nghiên cứu bởi một số nhóm tác giả, như áp dụng thuật toán học máy để ước tính áp suất điểm sương của vỉa khí condensate [2 - 5], tính toán tỷ số condensate khí [6, 7] và thành phần chất lưu [8]. Tuy nhiên, các nghiên cứu này chưa dự báo sự thay đổi của tỷ số condensate khí theo thời gian trong quá trình áp suất vỉa suy giảm khi khai thác.

Đối với giếng khai thác, quan sát thấy tỷ số condensate khí phụ thuộc nhiều vào giá trị áp suất đầu giếng. Ngoài ra, việc bắn vỉa bổ sung trong quá trình khai thác cũng có thể ảnh hưởng rất lớn đến tỷ số condensate khí. Do các giá trị áp suất đầu giếng, nhiệt độ đầu giếng và độ mở van khai thác đều được theo dõi và lưu đều đặn nên việc xác lập mối liên hệ giữa các dữ liệu trên với những dữ liệu thu thập khó khăn và tốn kém hơn như tỷ số condensate khí sẽ có nhiều giá trị thực tiễn. Do đó, việc sử dụng học máy để dự báo tỷ số condensate khí là một lựa chọn thay thế để khắc phục các khó khăn của các phương pháp truyền thống. Nghiên cứu này là một phần kết quả của đề tài cấp Nhà nước để nâng cao hiệu quả quản lý, khai thác mỏ khí condensate Hải Thạch - Mộc Tinh [9].

2. Phương pháp nghiên cứu

So với sự phức tạp của mô hình chất lưu hay mô hình động, học máy thực hiện dự báo với ít thông số đầu vào. Tuy nhiên, học máy cũng có những khó khăn nhất định trong việc dự báo theo thời gian, đặc biệt là thời gian tương đối dài trong tương lai. Lee và các cộng sự xây dựng mạng bộ nhớ ngắn dài LSTM (long short-term memories) được huấn luyện trên bộ dữ liệu của 300 giếng để dự báo sản lượng của 15 giếng, kết quả dự báo chính xác cao nhưng hạn chế là kết quả này chỉ cho dự báo ngắn hạn trong vòng 1 tháng [10]. Trong một nghiên cứu khác, Zhan và các cộng sự đã sử dụng dữ liệu từ hơn 300 giếng dầu phi truyền thống nhằm xây dựng 2 mô hình LSTM, một mô hình dự báo sự suy giảm của sản lượng và mô hình còn lại dự báo sản lượng cộng dồn [11]. Với mỗi giếng, nhóm tác giả sử dụng dữ liệu khai thác của 3 tháng đầu để huấn luyện mô hình và sử dụng 21 tháng còn lại cho việc thử nghiệm. Do các

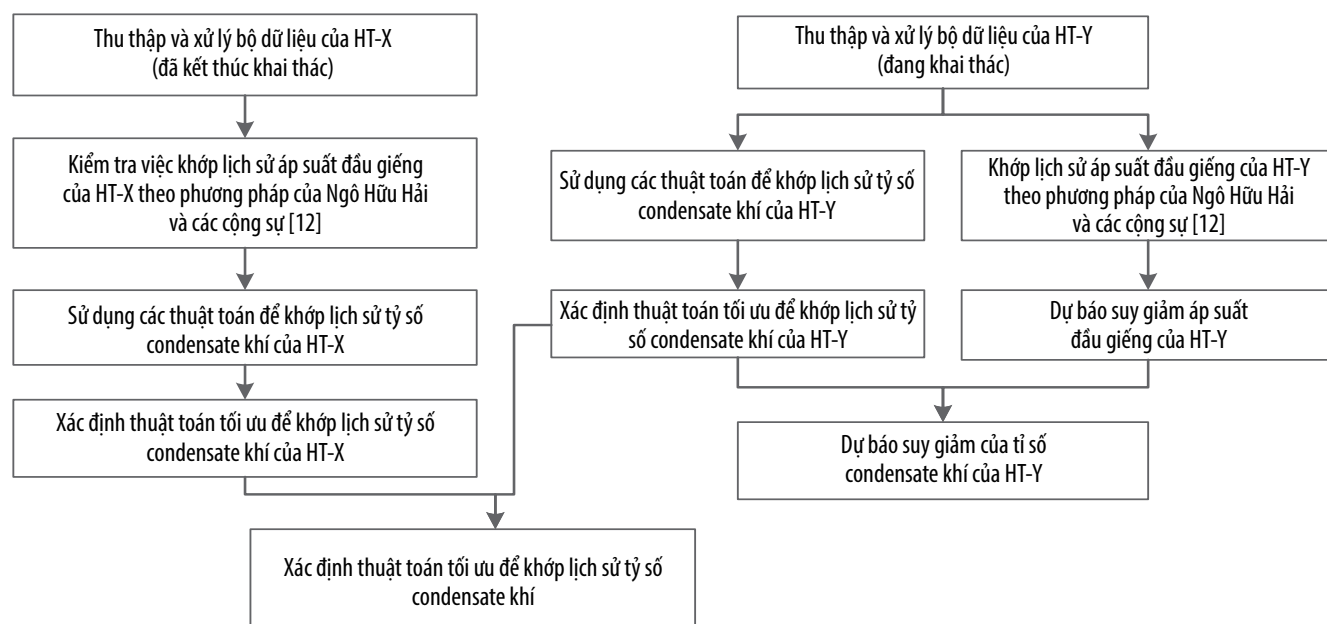
giếng dầu phi truyền thống có đặc điểm sụt giảm sản lượng nhanh chóng nên dữ liệu khai thác của 3 tháng đầu chứa rất nhiều thông tin để huấn luyện mô hình; điều này không đúng khi xây dựng mô hình cho các giếng dầu khí truyền thống. Để khắc phục khó khăn tích lũy sai số trong dự báo chuỗi thời gian và mô phỏng sự sụt giảm sản lượng nhanh chóng trong những tháng đầu, ngoài áp suất đầu giếng và sản lượng dầu của mỗi giếng, 12 giếng giống nhất với giếng đang được phân tích sẽ được chọn ra từ bộ dữ liệu và chuyển đổi thành các thông số đầu vào. Tuy nhiên, kết quả cho thấy mô hình LSTM dự báo sản lượng quá cao trong khi mô hình LSTM dự báo sản lượng cộng dồn quá thấp so với số liệu thực tế. Có thể thấy nhiều mô hình học máy có những khó khăn nhất định trong việc dự báo theo thời gian dài. Ngoài ra, việc nắm bắt những sự thay đổi bất thường trong dữ liệu lịch sử khai thác cũng là một khó khăn lớn cần được giải quyết cho các giếng có bản vỉa bổ sung. Với thuật toán hồi quy phân đoạn (piecewise regression) [12], nguyên tắc của thuật toán này là dùng thuật toán cây hồi quy (decision tree regressor) để chia nhóm các dữ liệu (bucketization), vì thế sẽ xử lý tốt các bộ dữ liệu có nhiều xu hướng khác nhau. Do đó, thuật toán hồi quy phân đoạn kết hợp với thuật toán hồi quy tuyến tính và XGboost đã được sử dụng để giải quyết bài toán dự báo này.

Tỷ số condensate khí là thông số phụ thuộc vào áp suất vỉa, do đó phụ thuộc vào áp suất đầu giếng. Vì vậy nghiên cứu này được chia thành 2 bước. Bước 1 là dự báo sự suy giảm của áp suất đầu giếng trong quá trình khai thác, theo phương pháp đã được công bố bởi Ngô Hữu

Hải và các cộng sự [12]. Bước 2 là dự báo tỷ số condensate khí theo sự suy giảm của áp suất đầu giếng.

Trong mỏ Hải Thạch, giếng HT-X bắt đầu khai thác từ năm 2015 với khả năng cho dòng tốt. Sau 5 năm khai thác, giếng HT-X đã cạn kiệt với áp suất đầu giếng suy giảm về sát với giá trị áp suất đầu vào hệ thống xử lý và tỷ số condensate khí suy giảm từ giá trị ban đầu 100 thùng/triệu ft^3 xuống còn 10 thùng/triệu ft^3 . Vì dữ liệu lịch sử về sự suy giảm của áp suất đầu giếng cũng như tỷ số condensate khí của HT-X đầy đủ cho đến khi đóng giếng, nên giếng này được sử dụng cho việc xây dựng mô hình học máy. Bộ số liệu áp suất đầu giếng của HT-X có 1.566 điểm dữ liệu từ lịch sử khai thác hàng ngày gồm thời gian, độ mở van khai thác và lưu lượng khí. Bộ số liệu tỷ số condensate khí của giếng HT-X có 52 điểm dữ liệu lịch sử từ kết quả kiểm tra động thái giếng gồm thời gian, độ mở van khai thác và áp suất đầu giếng. Do bộ dữ liệu lịch sử dùng để dự báo tỷ số condensate khí của giếng HT-X tương đối nhỏ nên độ rộng của chuỗi dữ liệu dùng để dự báo xu hướng (mean_leaf) sẽ ảnh hưởng nhiều tới kết quả dự báo. Sự kết hợp giữa thuật toán hồi quy phân đoạn và hồi quy tuyến tính hoặc hồi quy phân đoạn kết hợp với XGBoost lần lượt được sử dụng cho việc dự báo tỷ số condensate khí của HT-X với các tỷ lệ chia khác nhau của tập dữ liệu đào tạo và tập dữ liệu kiểm tra để tìm ra thuật toán tối ưu nhất cho việc dự báo sự suy giảm trong tương lai.

Quá trình trên sẽ được lặp lại cho giếng HT-Y, là đối tượng áp dụng chính của nghiên cứu này. Giếng HT-Y cũng bắt đầu khai thác từ năm 2015 như HT-X nhưng có



Hình 3. Lưu đồ nghiên cứu.

động thái khai thác tốt hơn và áp suất không suy giảm nhanh như HT-X nên sau 7 năm khai thác, HT-Y vẫn là giếng có đóng góp lớn nhất trong các giếng Hải Thạch và chưa có dấu hiệu bị cạn kiệt. Việc dự báo sự suy giảm của tỷ số condensate khí sẽ giúp quản lý khai thác tốt hơn cho giếng HT-Y. Bộ dữ liệu áp suất đầu giếng của HT-Y có 1.658 điểm dữ liệu từ lịch sử khai thác hàng ngày gồm thời gian, độ mở van khai thác và lưu lượng khí. Bộ dữ liệu tỷ số condensate khí có 132 điểm dữ liệu lịch sử từ kết quả kiểm tra động thái giếng gồm thời gian, độ mở van khai thác và áp suất đầu giếng. Đặc biệt, giếng HT-Y có khác biệt lớn so với HT-X là giếng này có bơm vỉa bổ sung, làm thay đổi động thái khai thác và sự kiện này sẽ được dùng để kiểm chứng khả năng của các thuật toán.

Lưu đồ nghiên cứu được thể hiện trong Hình 3.

3. Kết quả nghiên cứu

Bộ dữ liệu dùng để dự báo áp suất đầu giếng của HT-X được chia thành tập dữ liệu đào tạo và kiểm tra theo các

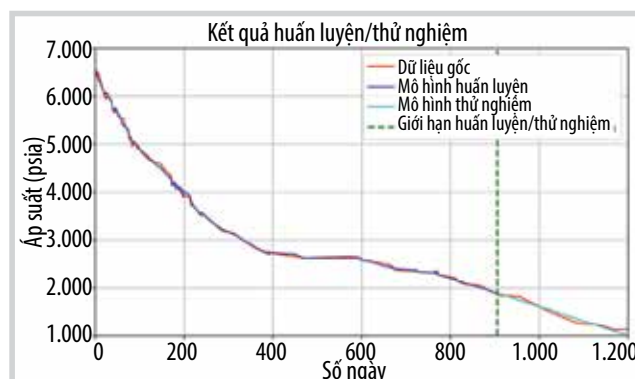
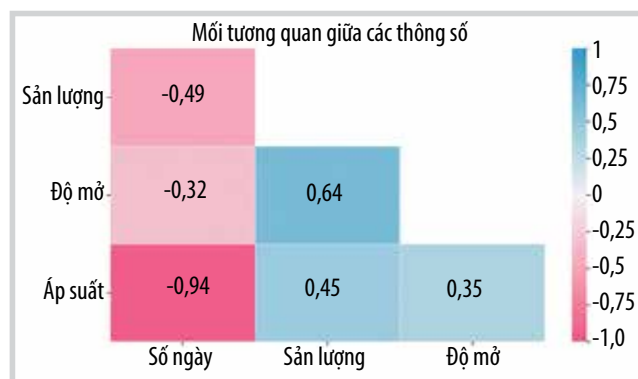
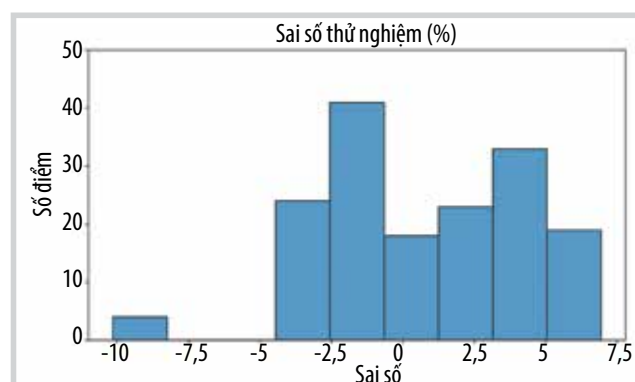
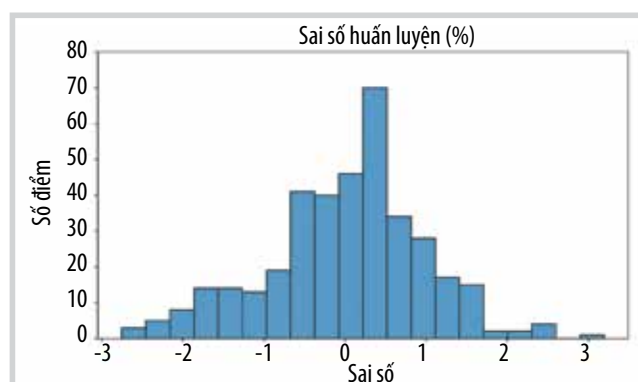
tỷ lệ khác nhau. Tham số mean_leaf được tối ưu hóa dựa trên điểm số cao nhất của hệ số tương quan khi so sánh kết quả dự báo với số liệu thực tế trên tập huấn luyện. Kết quả của quá trình tính toán được thể hiện trong các hình bên dưới.

Thuật toán hồi quy phân đoạn kết hợp với hồi quy tuyến tính được áp dụng để phục hồi lịch sử áp suất đầu giếng HT-X theo các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm khác nhau (Bảng 1) và kết quả đại diện trong Hình 4.

Kết quả kiểm tra cho thấy việc khớp lịch sử và dự báo áp suất đầu giếng theo phương pháp đề xuất bởi Ngô Hữu Hải và các cộng sự [12] là khả thi cho giếng HT-X khi tỷ lệ dữ liệu lịch sử/dự báo đạt 60/40 trở lên. Quá trình trên sẽ được lặp lại cho việc khớp lịch sử và dự báo sự suy giảm của áp suất đầu giếng HT-Y. Do áp suất đầu giếng của HT-Y thay đổi hoàn toàn sau khi bơm vỉa bổ sung nên các tỷ lệ chia 50/50, 60/40, 70/30 và 80/20 vẫn được áp dụng nhưng tính từ thời điểm bơm vỉa bổ sung. Kết quả dự

Bảng 1. Kết quả dự báo suy giảm áp suất đầu giếng HT-X bằng học máy theo các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm

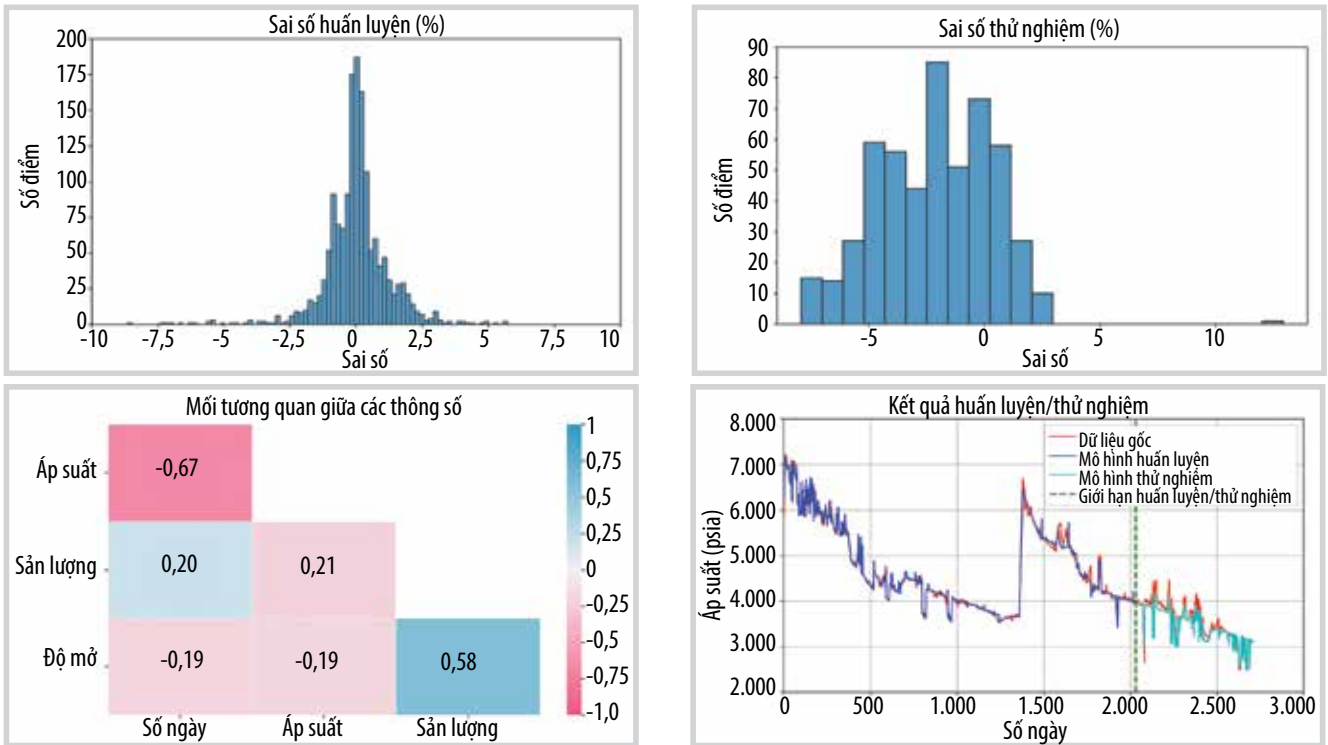
Tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm	Mean_leaf	Sai số huấn luyện	Sai số thử nghiệm	Kết quả dự báo
50/50	20	-1,5 đến 1,5	-15 đến 0	Chưa tốt
60/40	100	-15 đến 15	-7,5 đến 10	Khá tốt
70/30	50	-3 đến 3	-10 đến 7,5	Tốt
80/20	100	-8 đến 8	-2 đến 10	Tốt



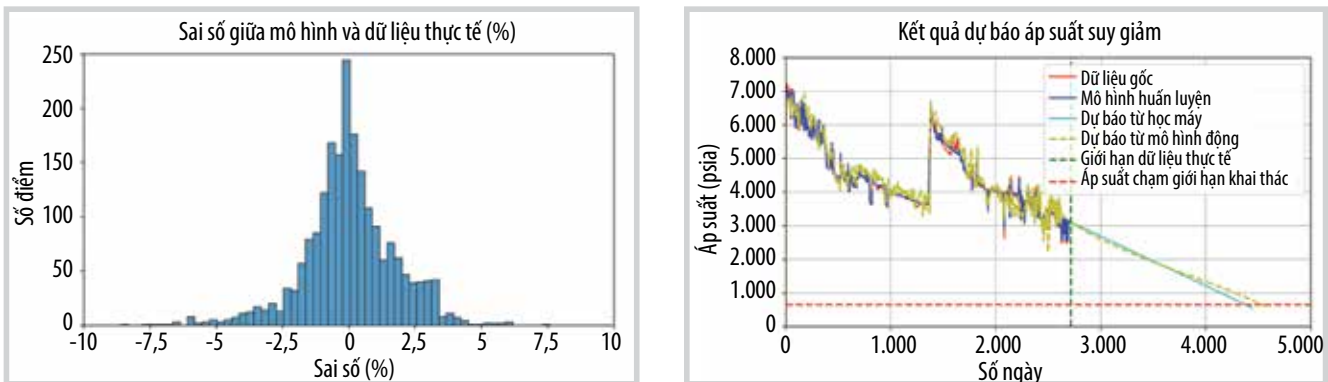
Hình 4. Kết quả dự báo suy giảm áp suất đầu giếng HT-X bằng học máy theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 70/30.

Bảng 2. Kết quả dự báo suy giảm áp suất đầu giếng HT-Y bằng học máy theo các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm

Tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm	Mean_leaf	Sai số huấn luyện	Sai số thử nghiệm	Kết quả dự báo
50/50	83	-5 đến 5	-8 đến 3	Tốt
60/40	199	-8 đến 8	-8 đến 4	Tốt
70/30	61	-8 đến 8	-12 đến 2	Tốt
80/20	184	-8 đến 8	-3 đến 6	Tốt



Hình 5. Kết quả dự báo sự suy giảm của áp suất đầu giếng HT-Y bằng học máy theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 50/50.



Hình 6. Kết quả dự báo sự suy giảm của áp suất đầu giếng HT-Y trong tương lai bằng học máy so sánh với kết quả dự báo bằng mô hình động.

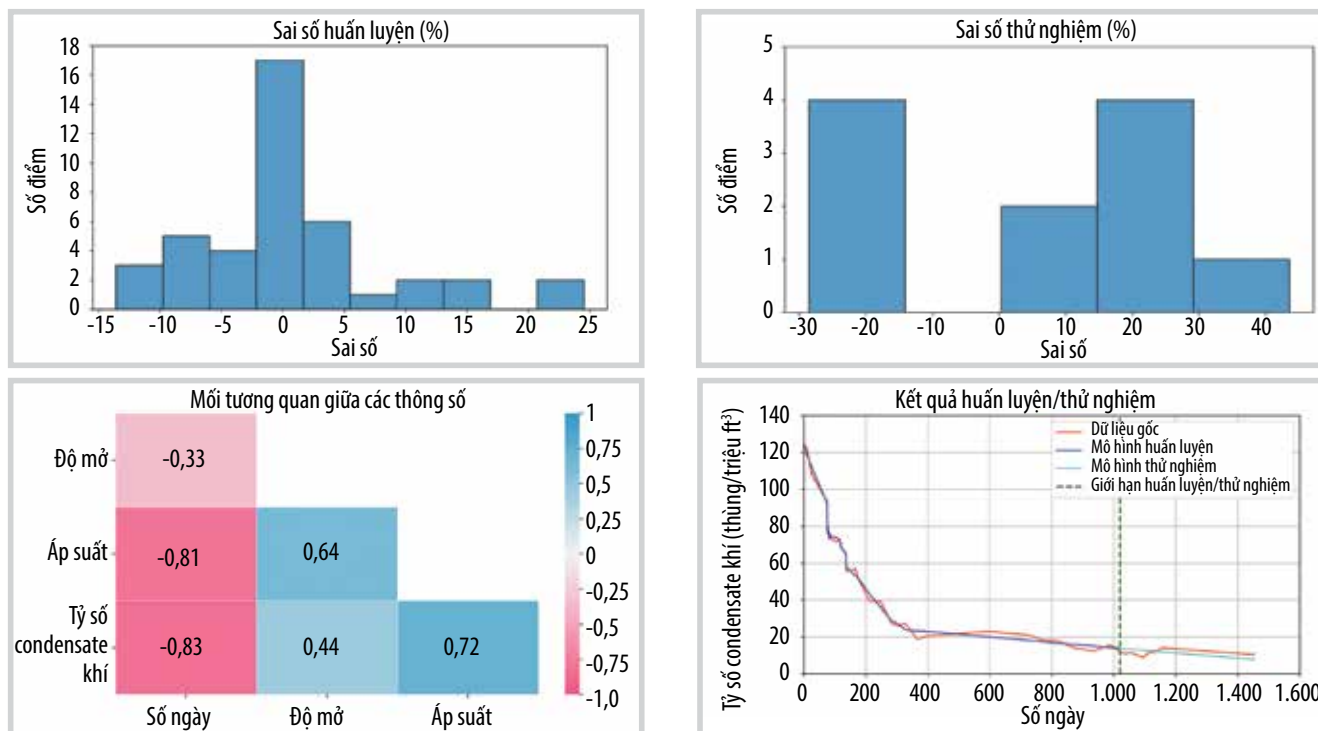
báo ở các tỷ lệ chia được tóm tắt trong Bảng 2 và kết quả đại diện được thể hiện trong Hình 5.

Kết quả dự báo áp suất đầu giếng cho HT-Y trong tương lai bằng học máy so với kết quả dự báo bằng mô hình động được thể hiện trong Hình 6. So sánh kết quả của 2 phương pháp dự báo cho thấy việc sử dụng học máy cho kết quả dự báo áp suất đầu giếng của HT-Y trong tương lai là hợp lý và chỉ có khác biệt nhỏ so với phương pháp sử dụng mô hình động.

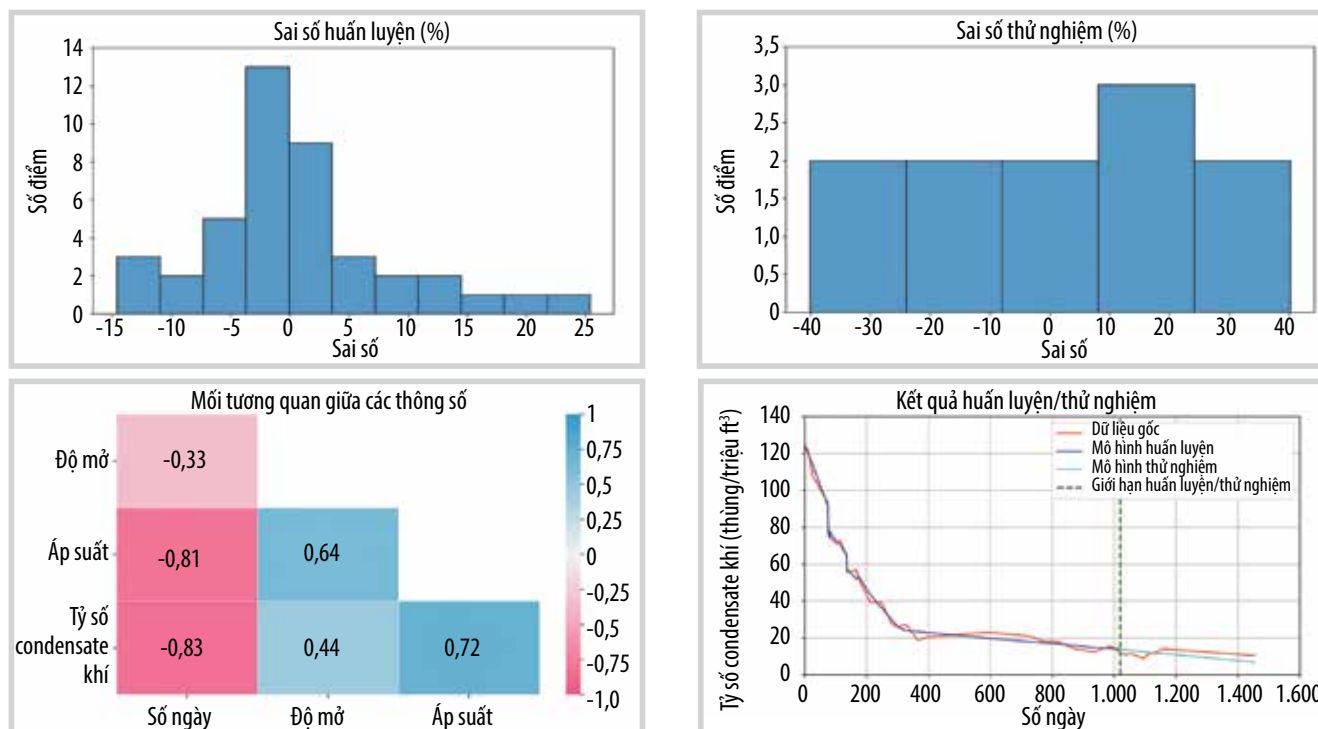
Sau khi kiểm chứng áp suất đầu giếng được dự báo với độ chính xác cao, trong bước tiếp theo tỷ số condensate khí cũng được dự báo bằng thuật toán học máy. Do bộ dữ liệu tỷ số condensate khí hạn chế so với bộ dữ liệu áp suất đầu giếng nên chỉ áp dụng các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm là 70/30 và 80/20. Kết quả cho giếng HT-X được tóm tắt trong Bảng 3 và kết quả đại diện được thể hiện trong Hình 7 và 8.

Bảng 3. Kết quả dự báo suy giảm tỷ số condensate khí của giếng HT-X bằng học máy theo các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm

Tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm	Thuật toán	Mean_leaf	Sai số huấn luyện	Sai số thử nghiệm	Kết quả dự báo
70/30	Hồi quy phân đoạn kết hợp với hồi quy tuyến tính	6	-10 đến 25	0 đến 70	Tốt
80/20	Hồi quy phân đoạn kết hợp với hồi quy tuyến tính	8	-15 đến 25	-30 đến 40	Tốt
70/30	Hồi quy phân đoạn kết hợp với XGBoost	6	-10 đến 25	0 đến 70	Tốt
80/20	Hồi quy phân đoạn kết hợp với XGBoost	8	-15 đến 25	-40 đến 40	Tốt



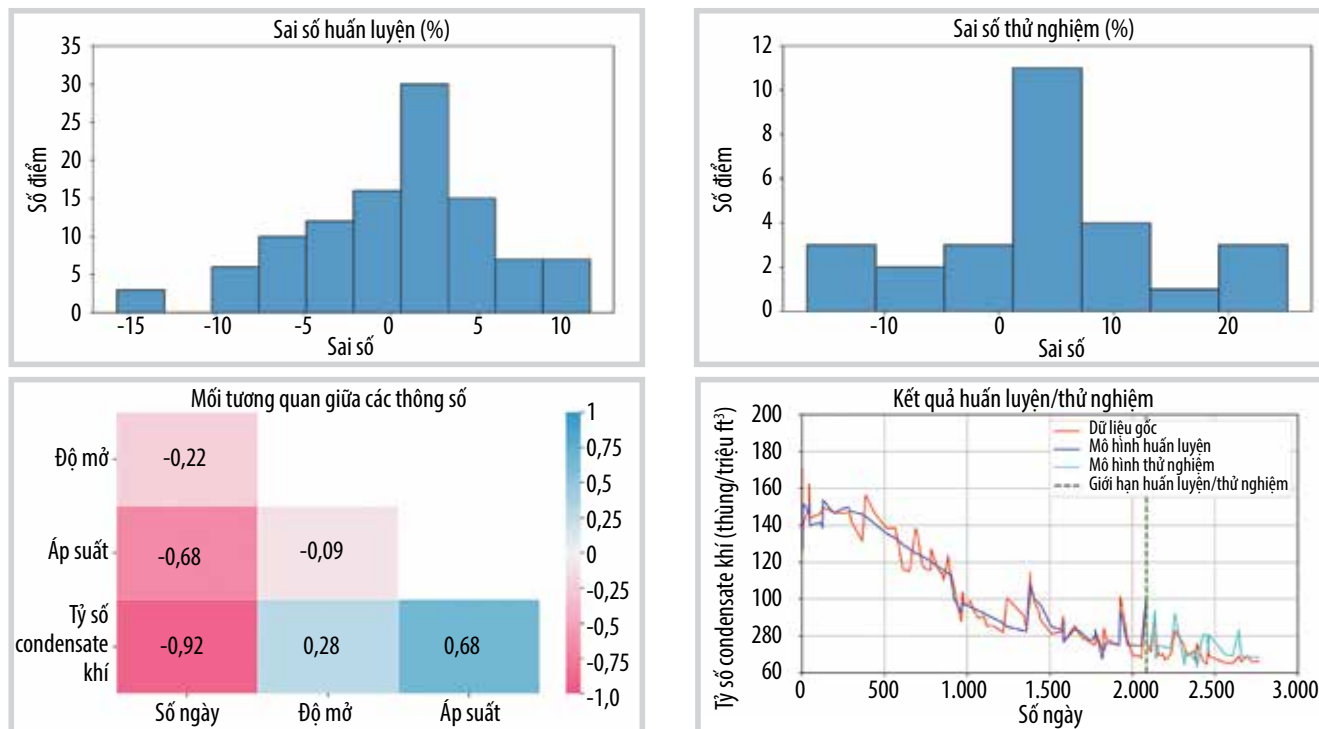
Hình 7. Kết quả dự báo tỷ số condensate khí của giếng HT-X bằng thuật toán hồi quy phân đoạn kết hợp với hồi quy tuyến tính theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 80/20.



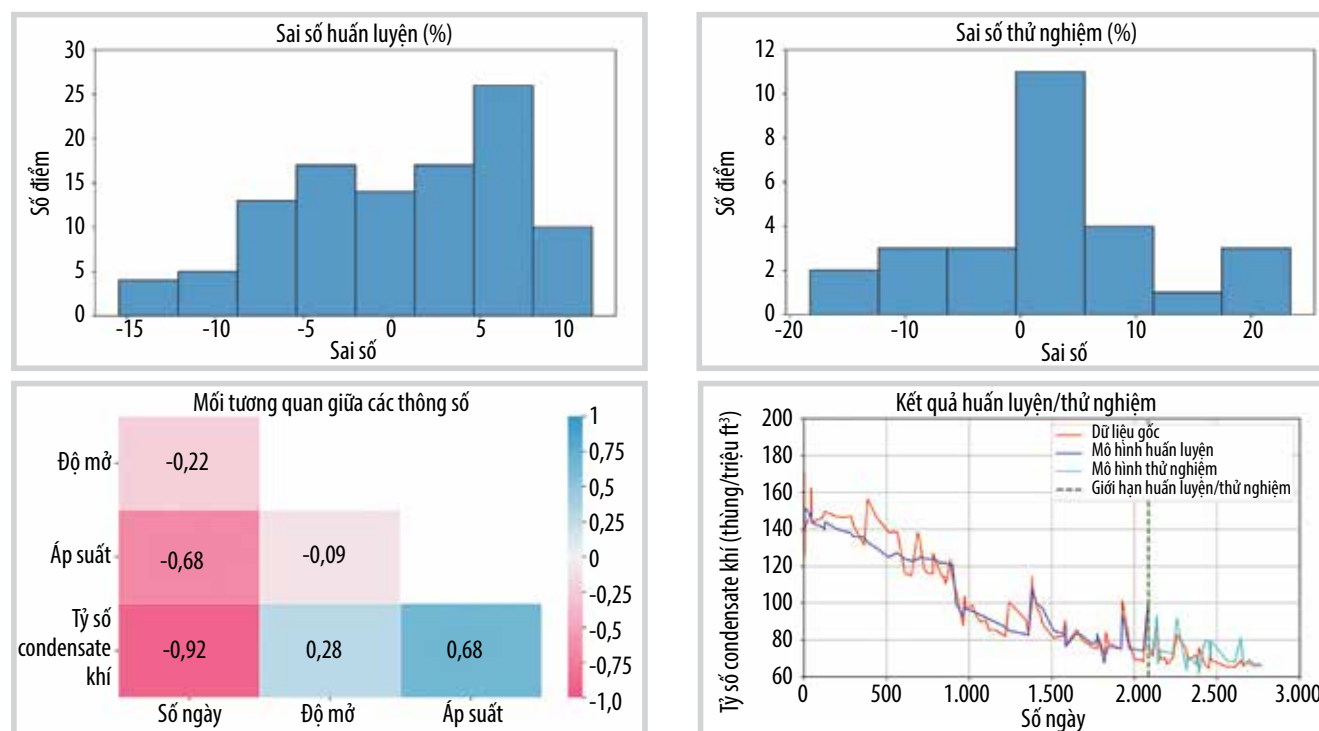
Hình 8. Kết quả dự báo tỷ số condensate khí của giếng HT-X bằng thuật toán hồi quy phân đoạn kết hợp với XGBoost theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 80/20.

Bảng 4. Kết quả dự báo suy giảm tỷ số condensate khí của giếng HT-Y bằng học máy theo các tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm

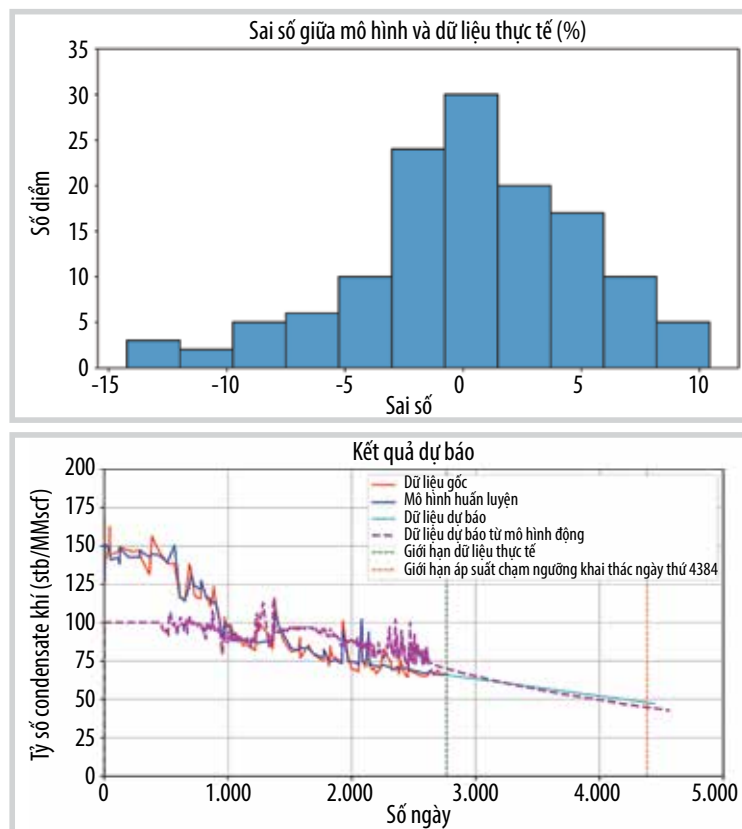
Tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm	Thuật toán	Mean_leaf	Sai số huấn luyện	Sai số thử nghiệm	Kết quả dự báo
70/30	Hồi quy phân đoạn kết hợp với hồi quy tuyến tính	20	-15 đến 10	-15 đến 30	Quá cao
80/20		30	-15 đến 10	-15 đến 25	Tốt
70/30	Hồi quy phân đoạn kết hợp với XGBoost	20	-15 đến 10	-20 đến 15	Tốt
80/20		30	-15 đến 10	-20 đến 20	Tốt



Hình 9. Kết quả dự báo tỷ số condensate khí của giếng HT-Y bằng thuật toán hồi quy phân đoạn kết hợp với hồi quy tuyến tính theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 80/20.



Hình 10. Kết quả dự báo tỷ số condensate khí của giếng HT-X bằng thuật toán hồi quy phân đoạn kết hợp với XGBoost theo tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm 80/20.



Hình 11. Kết quả dự báo tỷ số condensate khí của giếng HT-Y trong tương lai bằng học máy so sánh với kết quả dự báo bằng mô hình động.

Đối với việc dự báo tỷ số condensate khí của giếng HT-X, thuật toán hồi quy phân đoạn kết hợp với hồi quy tuyến tính cho kết quả dự báo khá tương đồng so với thuật toán hồi quy phân đoạn kết hợp với XGBoost. Do đó, tỷ số condensate khí của giếng HT-Y cũng được dự báo bằng cả 2 thuật toán với các tỷ lệ chia là 70/30 và 80/20 với kết quả được tóm tắt trong Bảng 4.

Thuật toán hồi quy phân đoạn kết hợp với XGBoost cho kết quả dự báo tốt hơn so với thuật toán hồi quy phân đoạn kết hợp với hồi quy tuyến tính với tỷ lệ chia 70/30.

Đối với tỷ lệ chia 80/20, kết quả dự báo khá tương đồng cho cả 2 thuật toán như trong Hình 9 và 10.

Như vậy, xét về tổng thể thì thuật toán hồi quy phân đoạn kết hợp với XGBoost sẽ cho kết quả dự báo tốt và ổn định hơn nên thuật toán này sẽ được sử dụng để dự báo cho tỷ số condensate khí trong tương lai của HT-Y, với kết quả học máy so với kết quả dự báo bằng mô hình động được thể hiện trong Hình 11. So sánh kết quả của 2 phương pháp dự báo (khác biệt tuyệt đối trung bình chỉ 2,5 thùng/triệu ft³ và khác biệt tương đối trung bình chỉ 4,6%) cho thấy việc sử dụng học máy cho kết quả dự báo tỷ số condensate khí trong tương lai là hợp lý với giếng HT-Y và có thể được sử dụng hỗ trợ cho dự báo mô hình động để quản lý khai thác giếng này.

4. Kết luận

Các kết luận chính của nghiên cứu có thể được tóm tắt như sau:

- Học máy đã được áp dụng thành công để dự báo sự thay đổi của tỷ số condensate khí theo thời gian khai thác, là một trong những thông số quan trọng nhất cho vỉa khí condensate nhưng rất khó dự báo bằng các phương pháp truyền thống;

- Về thuật toán học máy, thuật toán hồi quy phân đoạn kết hợp với XGBoost cho kết quả dự báo tốt và ổn định nhất cho tỷ số condensate khí. Kết quả xây dựng mô hình cho thấy kết quả dự báo tốt với tỷ lệ chia tập dữ liệu huấn luyện và thử nghiệm từ 70/30;

- Kết quả so sánh dự báo bằng học máy với dự báo bằng mô hình động cho thấy khác biệt tuyệt đối trung bình chỉ 2,5 thùng/triệu ft³ và khác biệt tương đối trung bình chỉ 4,6%;

- Việc dự báo thành công sự thay đổi của tỷ số condensate khí của giếng theo thời gian bằng học máy có ý nghĩa quan trọng trong việc hỗ trợ dự báo sản lượng condensate của giếng, qua đó giúp quản lý khai thác mỏ khí condensate tối ưu hơn.

Tài liệu tham khảo

[1] Curtis H. Whitson, Øivind Fevang, and Tao Yang, "Gas condensate PVT - what's really important and why?", *IBC Conference "Optimisation of Gas Condensate Fields", London, 28 - 29 January 1999.*

[2] Meisam Karbalaee Akbari, Farhang Jalali Farahani, and Yaser Abdy, "Dewpoint pressure estimation of gas condensate reservoirs, using artificial neural network (ANN)", *SPE Europec/EAGE Annual Conference and Exhibition, London, United Kingdom, 11 - 14 June 2007.* DOI: 10.2118/107032-MS.

[3] Seyed Mohammad Javad Majidi, Amin Shokrollahi, Milad Arabloo, Ramin Mahdikhani-Soleymanloo, and Mohsen Masihi, "Evolving an accurate model based on machine learning approach for prediction of dew-point pressure in gas condensate reservoirs", *Chemical Engineering Research and Design*, Volume 92, Issue 5, pp. 891 - 902, 2014. DOI: 10.1016/j.cherd.2013.08.014.

[4] Zhi Zhong, Siyan Liu, Mohammad Kazemi, and Timothy R. Carr, "Dew point pressure prediction based on mixed-kernels-function

support vector machine in gas-condensate reservoir", *Fuel*, Volume 232, pp. 600 - 609, 2018. DOI: 10.1016/j.fuel.2018.05.168.

[5] Princewill Ikpeka, Johnson Ugwu, Paul Russell, and Gobind Pillai, "Performance evaluation of machine learning algorithms in predicting dew point pressure of gas condensate reservoirs", *SN Applied Sciences*, Volume 2, 2020. DOI: 10.1007/s42452-020-03811-x.

[6] Sohrab Zendehboudi, Mohammad Ali Ahmadi, Lesley James, and Ioannis Chatzis, "Prediction of condensate-to-gas ratio for retrograde gas condensate reservoirs using artificial neural network with particle swarm optimization", *Energy & Fuels*, Volume 26, Issue 6, pp. 3432 - 3447, 2012. DOI: 10.1021/ef300443j.

[7] Mohammad Ali Ahmadi, Mohammad Ebadi, Payam Soleimani Marghmaleki, and Mohammad Mahboubi Fouladi, "Evolving predictive model to determine condensate-to-gas ratio in retrograded condensate gas reservoirs", *Fuel*, Volume 124, pp. 241 - 257, 2014. DOI: 10.1016/j.fuel.2014.01.073.

[8] Hana AlMatouq, Mohammed Alabbad, and Fatai Anifowose, "An artificial intelligence approach to predict molar compositions of reservoir fluid components", *SPE Gas & Oil Technology Showcase and Conference, Dubai, UAE, 21 - 23 October 2019*. DOI: 10.2118/198555-MS.

[9] Triệu Hùng Trường, Trần Vũ Tùng và nnk, "Nghiên cứu xây dựng bộ công cụ trí tuệ nhân tạo hỗ trợ đánh giá

phân tích, liên kết tài liệu địa chất, địa vật lý giếng khoan và số liệu khai thác để nâng cao hiệu quả quản lý, khai thác mỏ khí condensate Hải Thạch - Mộc Tinh Lô 05-2; 05-3, thuộc Biển Đông Việt Nam", Đề tài cấp Nhà nước thuộc "Chương trình khoa học và công nghệ trọng điểm cấp quốc gia phục vụ đổi mới, hiện đại hóa công nghệ khai thác và chế biến khoáng sản đến năm 2025", Mã số 077.2021.CNKK. QG/HĐKH-CN.

[10] Kyungbook Lee, Jungtek Lim, Daeung Yoon, and Hyungsik Jung, "Prediction of shale gas production at Duvernay formation using deep-learning algorithm", *SPE Journal*, Volume 24, Issue 6, pp. 2423 - 2437, 2019. DOI: 10.2118/195698-PA.

[11] Cheng Zhan, Sathish Sankaran, Vincent LeMoine, Jeremy Graybill, and Didi-Ooi Sher Mey, "Application of machine learning for production forecasting for unconventional resources", *Unconventional Resources Technology Conference, Denver, Colorado, USA, 22 - 24 July 2019*. DOI: 10.15530/urtec-2019-47.

[12] Ngo Huu Hai, Pham Hoang Duy, Nguyen Ngoc Tan, Hoang Ky Son, Tran Ngoc Trung và Tran Vu Tung, "Application of machine learning to decline curve analysis (DCA) for gas-condensate production wells with complex production history due to add-on perforation of new reservoirs", *Petrovietnam Science, Technology & Innovation*, Volume 2, pp. 4 - 9, 2023. DOI: 10.47800/PVSI.2023.02-01.

APPLICATION OF MACHINE LEARNING TO PREDICT THE TIME EVOLUTION OF CONDENSATE TO GAS RATIO FOR PLANNING AND MANAGEMENT OF GAS - CONDENSATE FIELDS

Ngô Hữu Hải¹, Trình Xuân Vinh¹, Nguyễn Ngọc Tân¹, Hoàng Kỳ Sơn¹, Ngô Tuấn Anh¹, Trần Ngọc Trung¹, Trần Vũ Tùng¹, Nguyễn Sỹ Tuấn²

¹Bien Dong Petroleum Operating Company (Bien Dong POC)

²Van Lang University

Email: sonhk@biendongpoc.vn

Summary

One of the most important parameters for evaluating, forecasting, and managing gas - condensate fields is the evolution of the condensate to gas ratio (CGR) over time. This parameter tends to decrease as reservoir pressure declines. Conventionally, gas and condensate samples are collected initially at the time starting production and periodically later to conduct laboratory analyses of fluid composition, properties and CGR. However, sampling, transporting and analysing samples take time and effort and, therefore, could be very expensive. To predict CGR over time, likewise, dynamic models are also frequently used. However, these models could include many uncertainties due to the assumption of input data, including reservoir structures, fluid phase interaction, and reservoir property distribution. Therefore, application of machine learning to predict the time evolution of CGR in this research is a new and effective approach to supplement conventional methods.

Key words: Machine learning, condensate to gas ratio, production forecast.