

# ỨNG DỤNG THUẬT TOÁN RỪNG NGẪU NHIÊN ĐỂ DỰ BÁO SÓNG VÀ ĐẬP KHÔNG KHÍ SINH RA DO NỔ Mìn TRÊN MỎ THAN ĐÈO NAI

NGUYỄN HOÀNG, BÙI XUÂN NAM,  
TRẦN QUANG HIẾU - Trường Đại học Mỏ-Địa chất  
Email: nguyenhoang.mdc@gmail.com

## 1. Giới thiệu

Nổ mìn là một trong những phương pháp phá vỡ đất đá trên các mỏ lộ thiên, mà hiệu quả của nó mang lại là điều không thể phủ nhận. Bên cạnh những ưu điểm của phương pháp phá vỡ đất đá bằng nổ mìn như khối lượng đất đá bị phá vỡ lớn, có thể phá vỡ bất kỳ loại đất đá nào, không phụ thuộc vào điều kiện thời tiết, khí hậu,...., thì các nhược điểm mà phương pháp nổ mìn phá vỡ đất đá đem lại là sinh ra sóng và đập không khí, sóng chấn động, đất đá bay,... làm rung động các công trình, nhà cửa, phá vỡ hệ thống cửa kính, ảnh hưởng trực tiếp tới sức khỏe người dân và tàn phá môi trường xung quanh.

Hầu hết các mỏ lộ thiên sử dụng phương pháp nổ mìn để phá vỡ đất đá đều đang phải đối mặt với việc giảm thiểu sóng chấn động và sóng va đập không khí sinh ra do nổ mìn. Một trong những cách đơn giản nhất để giảm thiểu những ảnh hưởng của sóng va đập không khí tới môi trường xung quanh là tăng khoảng cách an toàn và chất lượng, cấu trúc của các công trình [2], [14]. Các bức tường có tác dụng như những màn chắn nhằm giảm thiểu tác động của sóng chấn động cũng như sóng va đập không khí tới các công trình cũng đã được nghiên cứu và sử dụng [18]. Ngoài ra, một vài nghiên cứu tiếp cận theo phương pháp hiện đại hơn, sử dụng các thuật toán học máy như: giải thuật di truyền, các thuật toán hồi quy tuyến tính, hồi quy phi tuyến tính và hệ thống mạng nơ ron nhân tạo để dự báo sóng chấn động và sóng va đập không khí sinh ra do nổ mìn cũng đã được nghiên cứu và cho kết quả khả quan [1], [10], [15].

Tại Việt Nam, việc xác định sóng chấn động và sóng va đập không khí sinh ra do nổ mìn hiện nay chủ yếu được thực hiện bằng phương pháp đo

trực tiếp các vụ nổ ngoài hiện trường bằng các thiết bị như Blastmate hay Micromate. Công tác đo đạc trực tiếp ngoài hiện trường thường tốn kém về cả thời gian, chi phí và không được thực hiện một cách thường xuyên. Hơn nữa, công tác đo đạc ngoài hiện trường chỉ thực hiện được sau khi vụ nổ đã xảy ra và việc hạn chế ảnh hưởng của nó tới môi trường xung quanh là điều không thể. Do vậy, việc dự báo sóng va đập không khí và sóng chấn động sinh ra do nổ mìn trên các mỏ lộ thiên là một đề mới mẻ và có thể khắc phục được nhược điểm của phương pháp trên.

Bài báo trình bày một phương pháp tiếp cận sử dụng thuật toán rừng ngẫu nhiên để có thể sử dụng máy tính điện tử dự báo được sóng va đập không khí sinh ra do nổ mìn dựa trên các dữ liệu về sóng va đập không khí đo được bằng thiết bị Blastmate III từ 146 vụ nổ thu thập tại mỏ than Đèo Nai, Quảng Ninh.

## 2. Tổng quan về thuật toán rừng ngẫu nhiên

Thuật toán rừng ngẫu nhiên hay còn có tên gọi là Random Forests (RF) trong tiếng Anh, là một kỹ thuật phân lớp và hồi quy được đề xuất bởi Breiman năm 2001 [5] sử dụng nhiều cây phân loại hoặc hồi quy trong một nhóm. RF là một trong những thuật toán được xây dựng dựa trên mô hình cây quyết định. Mỗi cây đóng vai trò như một lá phiếu làm cơ sở ra quyết định cho thuật toán. Các phương pháp học nhóm kết hợp với các kết quả riêng lẻ của từng cây thường mang lại các kết quả tốt hơn [4]. RF là thuật toán được mở rộng dựa trên kỹ thuật đóng gói (bagging) hoặc tập hợp bootstrap sử dụng các mẫu ngẫu nhiên (có thay thế) của dữ liệu huấn luyện để tạo ra nhiều cây dữ liệu hồi quy không cần cắt tỉa và là tổng kết quả trung bình của chúng [12]. Trong kỹ thuật bagging, các cây được trồng bằng cách chọn

điểm chia tốt nhất tại mỗi nút cho tất cả các biến dự báo. Thuật toán RF chỉnh sửa phương pháp chia tách bằng cách lựa chọn điểm chia tốt nhất từ một tập hợp con ngẫu nhiên được lựa chọn của các biến dự báo [3], [12]. Việc lựa chọn các biến dự báo ngẫu nhiên sẽ tạo ra sự đa dạng hơn giữa các cây và làm giảm mối tương quan của chúng [17]. Tuy nhiên, do sử dụng các lựa chọn ngẫu nhiên nên RF mang lại kết quả hơi khác nhau mỗi khi nó được chạy. Việc kết hợp biến đầu ra từ các cây sẽ làm giảm đi sự khác nhau giữa các kết quả từ các cây và cho mô hình tổng quát hơn. Trong thực tế, khi cây được thêm vào, RF không làm hỏng chất lượng mô hình và hạn chế lỗi một cách tổng quát [5].

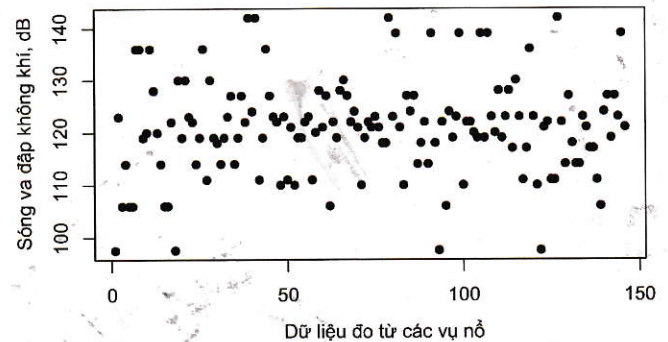
Ba đặc điểm hữu ích của thuật toán RF là đánh giá lỗi nội bộ, khả năng đánh giá tầm quan trọng của biến và khả năng xử lý các biến kém. Một trong những tính năng hấp dẫn của RF là nó có khả năng ước tính lỗi mà không cần bộ dữ liệu thử nghiệm riêng bằng cách sử dụng các ước tính lỗi ngoài [8], [11]. Ước tính lỗi ngoài được tạo ra từ dữ liệu không có trong mẫu bootstrap và được sử dụng cho sự phát triển của mỗi một cây [5]. Do không thể sử dụng một rừng ngẫu nhiên với hàng trăm cây để đánh giá vai trò của từng biến nên tầm quan trọng của mỗi biến có thể được ước tính bằng cách theo dõi lỗi dự báo thay đổi sau khi mỗi cây được xây dựng [5]. Tầm quan trọng của mỗi biến có thể được sử dụng để đạt được sự hiểu biết về các giá trị tương đối của các biến dự báo để đưa ra các giải pháp và có thể giảm số lượng các biến đầu vào.

Thuật toán RF đã được sử dụng rộng rãi trong rất nhiều lĩnh vực như: lập bản đồ không gian [8], [17]; dự báo dự phát triển của cây trồng [6]; dự báo sóng chấn động sinh ra do nổ mìn trên các mỏ lộ

thiên [13]; cải tạo phục hồi môi trường cho mỏ than [7]; dự báo sự ổn định bờ mỏ trên mỏ lộ thiên [20],... Trong phạm vi bài báo này, nhóm tác giả sử dụng RF để dự báo sóng va đập không khí sinh ra do nổ mìn trên mỏ than Đèo Nai, Quảng Ninh.

### 3. Số liệu sử dụng và phương pháp nghiên cứu

Số liệu sử dụng cho bài báo bao gồm các thông số như sóng va đập không khí (P), tần số f được đo bằng thiết bị Blastmate III và các thông số khác của hộ chiếu nổ mìn được thu thập từ 146 vụ nổ tại mỏ than Đèo Nai.



H.1. Dữ liệu thu thập được từ 146 vụ nổ tại mỏ than Đèo Nai

Ngoài ra, các dữ liệu cần thiết khác phục vụ cho nghiên cứu này bao gồm: Chiều dày lớp đất phủ B, khoảng cách giữa các hàng lỗ khoan (a), chỉ tiêu thuốc nổ (q), số hàng lỗ khoan nổ đồng thời (N), khoảng cách đặt thiết bị (R), tổng khối lượng thuốc nổ trong một lần nổ (Q), chiều cao tầng (H), chiều dài cột bua ( $L_b$ ) và độ ẩm không khí (W) được thu thập từ 146 vụ nổ của mỏ. Dữ liệu sử dụng cho nghiên cứu này được tóm tắt trong Bảng 1.

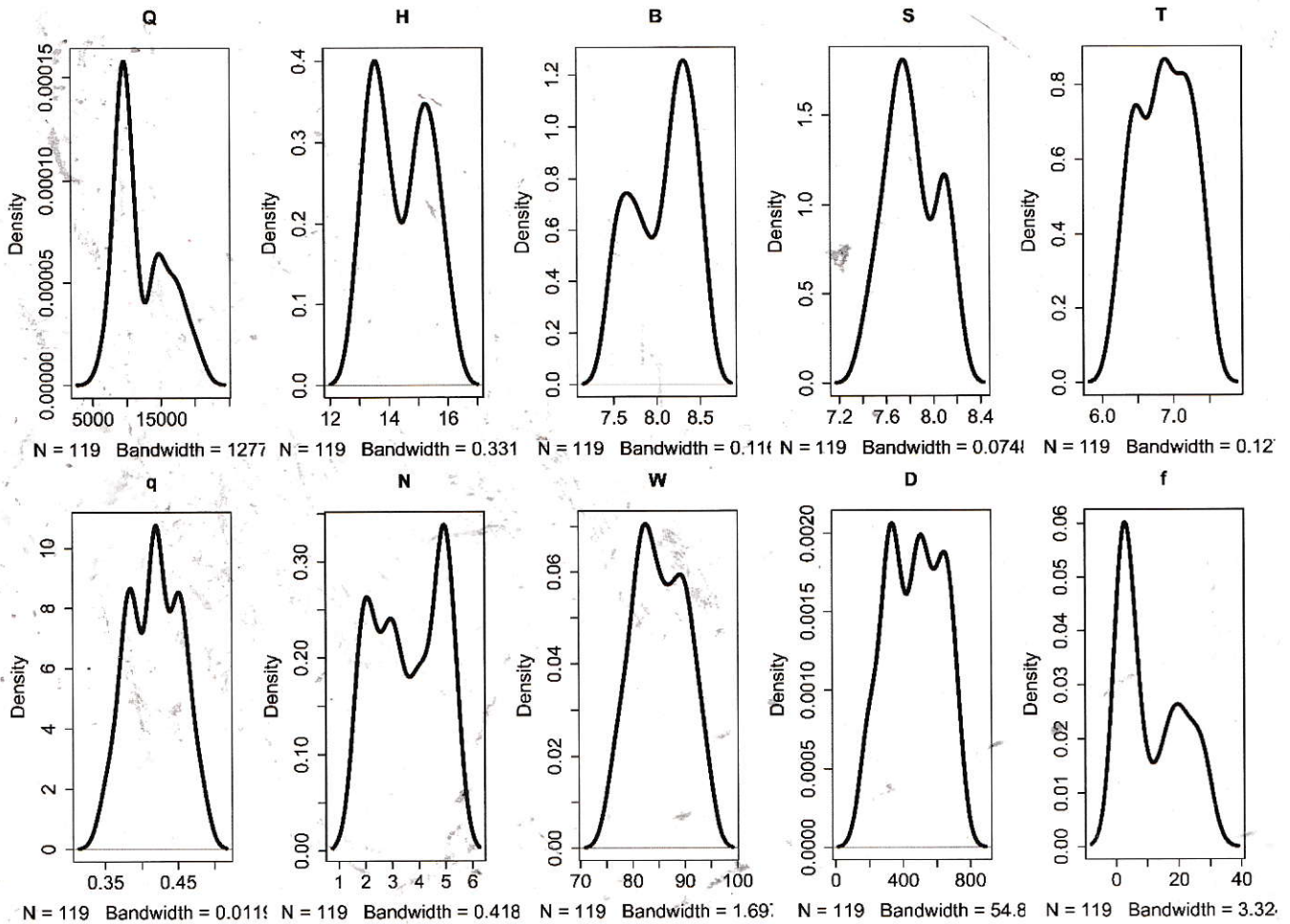
Bảng 1. Tóm tắt dữ liệu sử dụng cho nghiên cứu

Q	H	B	a	$L_b$	q
Min.: 6546	Min.:13.00	Min.:7.500	Min.:7.400	Min.:6.200	Min.:0.3500
P25 %: 9312	P25 %:13.50	P25 %:7.800	P25 %:7.700	P25 %:6.600	P25 %:0.3900
P50 %:12108	P50 %:14.37	P50 %:8.064	P50 %:7.814	P50 %:6.879	P50 %:0.4178
P75 %:14258	P75 %:15.00	P75 %:8.300	P75 %:8.000	P75 %:7.200	P75 %:0.4500
Max.:20426	Max.:16.00	Max.:8.500	Max.:8.200	Max.:7.500	Max.:0.4800
N	W	R	f	P	-
Min.:2.000	Min.:76.00	Min.:180	Min.: 2.00	Min.: 97.5	-
P25 %:2.000	P25 %:81.25	P25 %:338	P25 %: 2.50	P25 %:119.0	-
Median:3.000	Median:85.00	Median:495	Median: 5.20	Median:122.0	-
P50 %:3.486	P50 %:85.16	P50 %:469	P50 %:10.89	P50 %:121.6	-
P75 %:5.000	P75 %:89.00	P75 %:622	P75 %:19.50	P75 %:127.0	-
Max.:5.000	Max.:94.00	Max.:726	Max.:29.00	Max.:142.0	-

Các bước kiểm tra và xử lý dữ liệu trước khi tiến hành học máy sẽ được tiến hành. Để kiểm tra,

xử lý dữ liệu và thực hiện học máy, bộ dữ liệu bao gồm 146 quan sát sẽ được chia thành 2 phần. Trong

đó, 80 % (tương đương 119 quan sát) sẽ được sử dụng để huấn luyện thuật toán và 20 % (tương ứng 27 quan sát) sẽ được sử dụng để đánh giá kết quả của quá trình huấn luyện.



H.2. Đồ thị hàm mật độ xác định phân phối của dữ liệu đầu vào

Bảng 2. Kết quả phân tích mối tương quan giữa các biến đầu vào của dữ liệu

	Q	H	B	S	T	q	N	W	D	f
Q	1.00	0.71	0.47	0.77	-0.79	0.13	0.30	0.73	-0.90	0.22
H	0.71	1.00	0.24	0.58	-0.54	-0.03	0.31	0.48	-0.62	0.13
B	0.47	0.24	1.00	0.56	-0.56	0.32	0.13	0.53	-0.56	0.11
S	0.77	0.58	0.56	1.00	-0.76	0.31	0.17	0.67	-0.84	0.09
T	-0.79	-0.54	-0.56	-0.76	1.00	-0.23	-0.20	-0.80	0.89	-0.19
q	0.13	-0.03	0.32	0.31	-0.23	1.00	-0.02	0.12	-0.21	0.08
N	0.30	0.31	0.13	0.17	-0.20	-0.02	1.00	0.19	-0.29	-0.01
W	0.73	0.48	0.53	0.67	-0.80	0.12	0.19	1.00	-0.78	0.23
D	-0.90	-0.62	-0.56	-0.84	0.89	-0.21	-0.29	-0.78	1.00	-0.16
f	0.22	0.13	0.11	0.09	-0.19	0.08	-0.01	0.23	-0.16	1.00

Phân phối lý tưởng cho quá trình học máy là phân phối Gaussian. Tuy nhiên, H.2 cho thấy tất cả các biến đầu vào của dữ liệu đều gần với luật phân phối beta hoặc uniform. Điều này có thể không tốt cho quá trình huấn luyện và được gọi là độ nghiêng. Do vậy cần có phương pháp chuyển đổi năng lượng của các

biến đầu vào sao cho chúng gần với phân phối Gaussian nhất. Phương pháp chuyển đổi Box-Cox đã được sử dụng cho nghiên cứu này [16]. Ngoài việc kiểm tra phân phối của các biến đầu vào, sự tương quan giữa các biến đầu vào cũng là yếu tố cần thiết phải kiểm tra để thấy được mối liên hệ giữa các biến

đầu với với nhau. Một bảng phân tích giá trị tương quan giữa các biến đầu vào được xác định để minh họa sự điều này (Bảng 2). Các giá trị lớn hơn 0,75 và nhỏ hơn -0,75 là những giá trị khá hấp dẫn khi chúng cho thấy mối tương quan tích cực và mỗi tương quan tiêu cực giữa các biến đầu vào. Các giá trị 1 và -1 thể hiện mối tương quan tích cực và tiêu cực một cách tuyệt đối. Từ các phân tích số liệu và xử lý số liệu, nhóm tác giả sử dụng phương pháp học máy bằng cách huấn luyện máy tính điện tử học thuật toán Random Forest với các dữ liệu đầu vào đã được kiểm tra và xử lý, hiệu quả của quá trình học được đánh giá bằng các giá trị dự báo và so sánh với giá trị thực tế đo được bằng thiết bị Blastmate III.

**4. Kết quả nghiên cứu và thảo luận**

Từ các kết quả phân tích và số liệu của dữ liệu đầu vào sau khi được xử lý, nhóm tác giả tiến hành huấn luyện cho máy tính điện tử từ các dữ liệu đầu vào (từ 119 quan sát đã được chia ở trên) và tìm ra quy luật chung để dự báo sóng va đập không khí (dữ liệu đầu ra) bằng thuật toán Random Forests. Kết quả

mô hình dự báo bằng thuật toán Random Forests sẽ được đánh giá dựa trên bộ dữ liệu kiểm tra (bao gồm 27 quan sát đã được chia ở trên) để kiểm tra mức độ chính xác của mô hình dự báo. Một số tiêu chuẩn đánh giá bao gồm sai số toàn phương trung bình (RMSE) và hệ số xác định ( $R^2$ ) được sử dụng để đánh giá mức độ chính xác của mô hình. RMSE chỉ mức độ chính xác của mô hình dự báo và đạt giá trị bằng 0 là tuyệt đối,  $R^2$  chỉ mức độ phù hợp của dữ liệu và đạt giá trị bằng 1 là tuyệt đối, bằng 0 là tệ nhất.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{1}$$

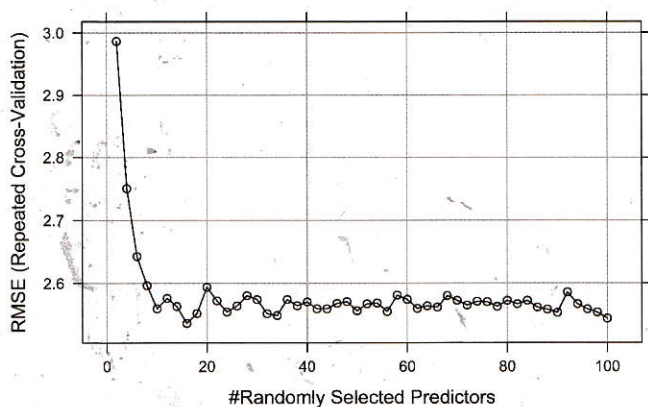
Trong đó: n - Số các quan sát;  $y_i$  - Giá trị được dự báo từ mô hình;  $\hat{y}_i$  - Giá trị thực tế của biến phụ thuộc.

$$R^2 = 1 - \frac{\sum (y_i - f_i)^2}{\sum (y_i - \bar{y})^2} \tag{2}$$

Trong đó:  $y_i$  - Giá trị được dự báo từ mô hình;  $f_i$  - Giá trị đích cần đạt được;  $\bar{y}$  - Giá trị trung bình của y.

Bảng 3. Mức độ chính xác và phù hợp của dữ liệu với mô hình Random Forest

mtry	RMSE	$R^2$	mtry	RMSE	$R^2$
2	2.986265	0.88485	52	2.567315	0.894845
4	2.750635	0.893074	54	2.56854	0.892419
6	2.6429	0.894318	56	2.554929	0.894721
8	2.596485	0.894204	58	2.581098	0.892318
10	2.559598	0.893229	60	2.574285	0.893206
12	2.576108	0.892303	62	2.559948	0.892775
14	2.563161	0.893651	64	2.563883	0.893504
16	2.536074	0.894842	66	2.561964	0.894226
18	2.551845	0.89449	68	2.580093	0.892505
20	2.593859	0.890837	70	2.572509	0.892696
22	2.572306	0.892494	72	2.56505	0.894293
24	2.554426	0.893622	74	2.571121	0.892822
26	2.563964	0.893752	76	2.570336	0.893102
28	2.580469	0.892757	78	2.562953	0.893443
30	2.574246	0.892932	80	2.572496	0.892078
32	2.551577	0.89468	82	2.567003	0.893322
34	2.54876	0.894071	84	2.572472	0.892984
36	2.574236	0.892497	86	2.561579	0.893106
38	2.564271	0.893033	88	2.55824	0.892667
40	2.57019	0.892844	90	2.55314	0.8948
42	2.559403	0.893539	92	2.58562	0.892015
44	2.559303	0.892828	94	2.56671	0.893613
46	2.568025	0.892865	96	2.558679	0.893951
48	2.570874	0.892473	98	2.553651	0.894474
50	2.556188	0.8933	100	2.543745	0.894449



H.3. Đồ thị xác định mô hình Random Forest tối ưu cho dữ liệu dự báo sóng và đập không khí tại mỏ than Đèo Nai

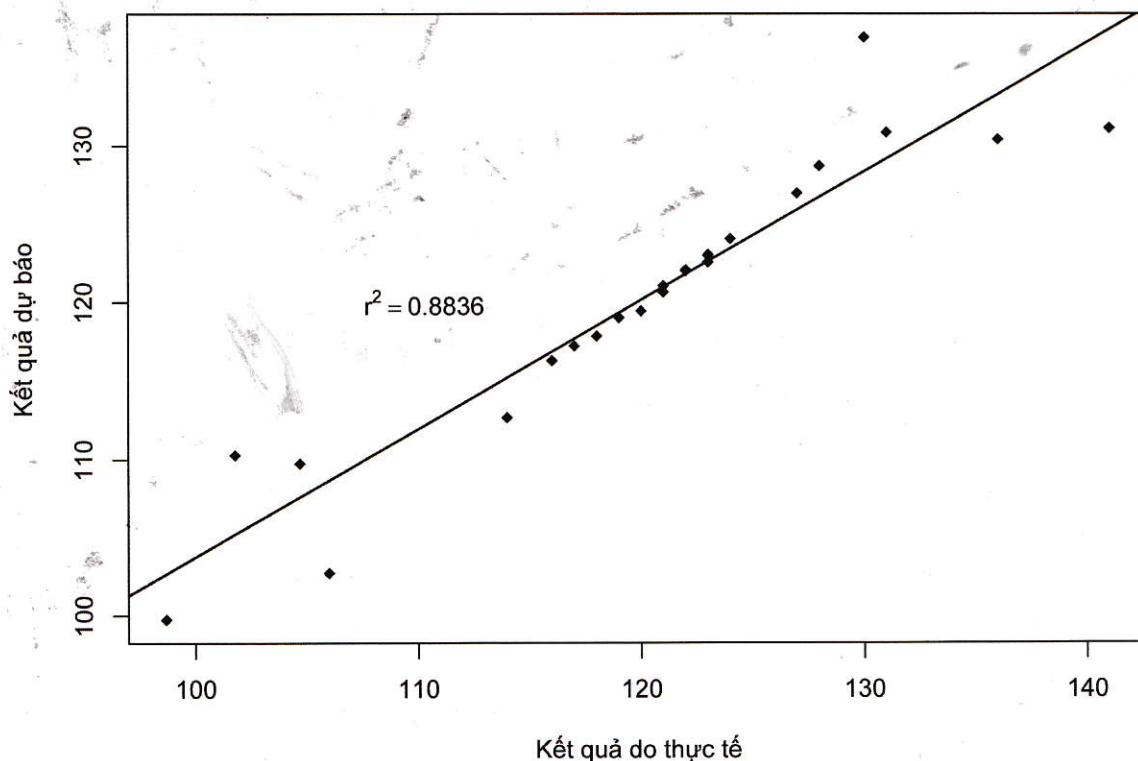
Để kiểm tra và xác định được mô hình dự báo tối ưu cho dữ liệu, nhóm tác giả tiến hành chạy 50

mô hình dựa trên thuật toán Random Forests với tham số "mtry" của thuật toán thay đổi từ 2 đến 100, bước nhảy là 2. Các kết quả của 50 mô hình được thể hiện trong Bảng 3.

Từ các kết quả huấn luyện 50 mô hình trong Bảng 3 và minh họa trong H.3 cho thấy: Mô hình Random Forests với mtry=16 là mô hình tối ưu nhất cho dự báo sóng và đập không khí tại mỏ than Đèo Nai với RMSE=2.536074 (0 là tốt nhất) và  $R^2=0.894842$  (1 là tốt nhất, 0 là tệ nhất) sau khi được huấn luyện. Tuy nhiên, hiệu suất của mô hình dự báo được kiểm tra và đánh giá bởi bộ dữ liệu đánh giá (bao gồm 278 quan sát đã được chia ở trên). Quá trình đánh giá mô hình sẽ cho thấy mô hình được chọn có phải là tối ưu hay không. Bảng 4 và H.4 minh họa kết quả của mô hình RF trên cả tập dữ liệu huấn luyện và dữ liệu đánh giá.

Bảng 4. Hiệu suất của mô hình dự báo RF được chọn

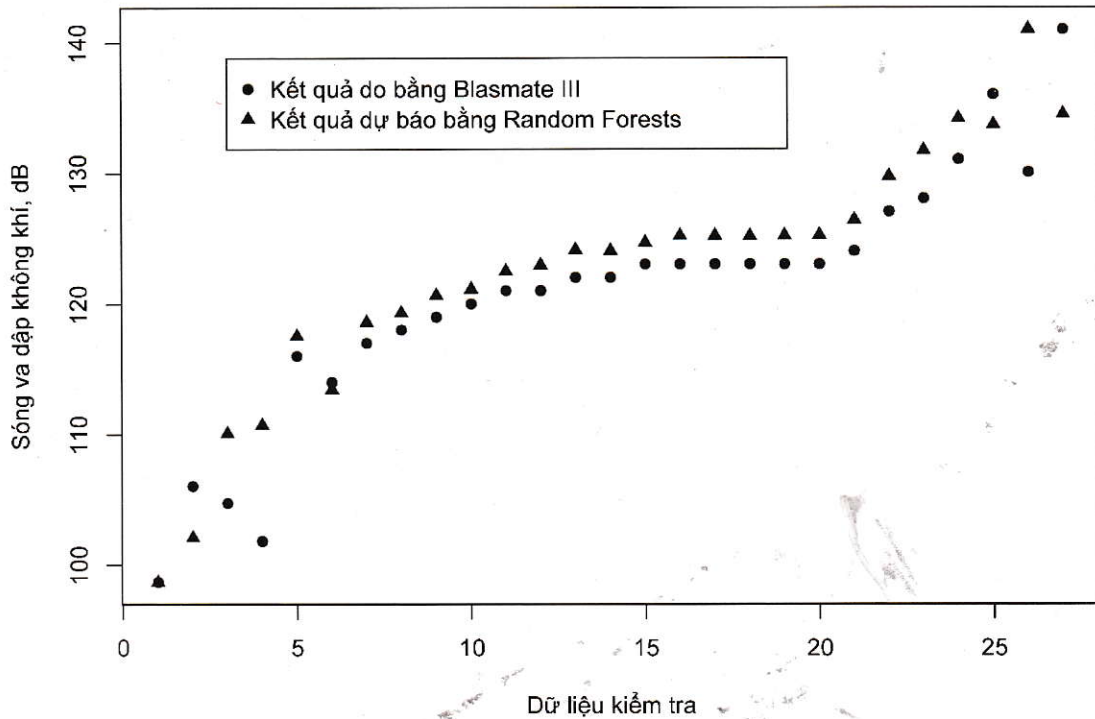
Mô hình	Dữ liệu huấn luyện		Dữ liệu đánh giá	
	RMSE	$R^2$	RMSE	$R^2$
Random Forests	2.536074	0.894842	3.26459	0.88361



H.4. Đồ thị hồi quy đánh giá mức độ phù hợp của dữ liệu kiểm tra

Các kết quả ở Bảng 4, H.4 cho thấy mô hình RF được chọn là một mô hình tốt với RMSE của tập dữ liệu đánh giá cũng khá nhỏ (3.26459) và  $R^2$  của tập dữ liệu đánh giá khá tốt (0.88361). Các kết quả này

rất phù hợp với bộ dữ liệu huấn luyện trước đó trong Bảng 2. Một đồ thị so sánh mức độ chính xác giữa các giá trị dự báo bằng mô hình được lựa chọn và các giá trị đo ngoài thực tế được thể hiện trong H.5.



H.5. Đồ thị so sánh mức độ chính xác giữa các kết quả đo thực tế với các kết quả dự báo bằng mô hình Random Forests

Các giá trị dự báo bằng mô hình Random Forests với  $mtry=16$  được thể hiện bằng các chấm hình tam giác màu xanh trong H.5 cho thấy mức độ chính xác của mô hình dự báo khá cao khi phần lớn các điểm màu xanh đều rất gần so với các điểm màu đỏ (giá trị đo thực tế).

So với các nghiên cứu trước đó trên thế giới [1], [10], [15] thì nghiên cứu này cho kết quả khá khả quan.

Ngoài ra, so với các nghiên cứu trước đó trên thế giới khi dự báo AOp [9], [19] thì nghiên cứu này đã đưa thêm thông số “Độ ẩm không khí” vào dữ liệu đầu vào để dự báo sóng và đập không khí sinh ra do nổ mìn trên mỏ than Đèo Nai.

Các kết quả tính toán trong Bảng 2 cho thấy độ ẩm không khí cũng có mối tương quan khá cao so với các thông số còn lại và đây cũng là một trong những biến đầu vào có ảnh hưởng lớn đến mức độ chính xác của mô hình dự báo sóng và đập không khí trong thực tế.

**5. Kết luận**

Phương pháp sử dụng các thuật toán học máy để dự báo sóng và đập không khí sinh ra trên các mỏ lộ thiên là một phương pháp hiện đại và mới mẻ, cho hiệu quả cao và kết quả áp dụng cho mỏ than Đèo Nai là một ví dụ điển hình trong nghiên cứu này. Kết quả nghiên cứu này đã chỉ ra rằng Random Forests là một thuật toán khá phù hợp

cho dự báo sóng và đập không khí tại mỏ than Đèo Nai với  $RMSE=3.26459$  và  $R^2=0.88361$ .

Kết quả này cũng là tài liệu tham khảo cho mỏ than Đèo Nai và các mỏ khác có điều kiện tương tự để điều chỉnh các thông số nổ mìn một cách hợp lý, đảm bảo an toàn về sóng và đập không khí sinh ra do nổ mìn.

Ngoài ra, việc sử dụng trí thông minh nhân tạo để dự báo sóng và đập không khí sinh ra do nổ mìn cũng giúp cho các mỏ tiết kiệm được nhiều thời gian và chi phí đo đạc thực tế, cập nhật những tiến bộ của khoa học-kỹ thuật, đặc biệt khi cuộc cách mạng công nghệ 4.0 đang tới gần. □

**TÀI LIỆU THAM KHẢO**

1. Danial Jahed Armaghani và các cộng sự. (2015), "Neuro-fuzzy technique to predict air-overpressure induced by blasting", Arabian Journal of Geosciences. 8(12), tr. 10937-10950.
2. US Army (1998), Technical manual design and analysis of hardened structures to conventional weapons effects, Army TM5-855-1, Washington DC.
3. L Breiman, "Manual setting up, using, and understanding random forests V4. 0. 2003 <http://oz.berkeley.edu/users/breiman>", Using\_random\_forests\_v4.0.pdf.
4. Leo Breiman (1996), "Bagging predictors", Machine learning. 24(2), tr. 123-140.

5. Leo Breiman (2001), "Random forests", *Machine learning*. 45(1), tr. 5-32.
6. Andrew G Bunn, Scott J Goetz và Gregory J Fiske (2005), "Observed and predicted responses of plant growth to climate across Canada", *Geophysical Research Letters*. 32(16).
7. Yamileth Domiguez-Haydar and Ambrecht (2011), "Response ants and their seed removal in rehabilitation areas and forests at El Cerrejon coal mine in Colombia", *Restoration Ecology*. 19(201), tr. 178-184.
8. Cesare Furlanello và các cộng sự. (2003), GIS and the random forest predictor: Integration in R for tick-borne disease risk assessment, *Proceedings of DSC*, tr. 2.
9. M Hajihassani và các cộng sự. (2014), "Prediction of airblast-overpressure induced by blasting using a hybrid artificial neural network and particle swarm optimization", *Applied Acoustics*. 80, tr. 57-67.
10. Mahdi Hasanipanah và các cộng sự. (2016), "Several non-linear models in estimating air-overpressure resulting from mine blasting", *Engineering with Computers*. 32(3), tr. 441-455.
11. Rick L Lawrence, Shana D Wood và Roger L Sheley (2006), "Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest)", *Remote Sensing of Environment*. 100(3), tr. 356-362.
12. Andy Liaw và Matthew Wiener (2002), "Classification and regression by randomForest", *R news*. 2(3), tr. 18-22.
13. Dong Longjun và các cộng sự. (2011), "Comparisons of random forest and support vector machine for predicting blasting vibration characteristic parameters", *Procedia Engineering*. 26, tr. 1772-1781.
14. R.P. Mayor và R. Flanders (1990), "Technical manual simplified computer model of air blast effects on building walls", US Department of State, Office of Diplomatic Security, Washington DC.
15. Edy Tonnizam Mohamad và các cộng sự. (2016), "Estimation of air-overpressure produced by blasting operation through a neuro-genetic technique", *Environmental Earth Sciences*. 75(2), tr. 174.
16. Jason W Osborne (2010), "Improving your data transformations: Applying the Box-Cox transformation", *Practical Assessment, Research & Evaluation*. 15(12), tr. 1-9.
17. Anantha M Prasad, Louis R Iverson và Andy Liaw (2006), "Newer classification and regression tree techniques: bagging and random forests for ecological prediction", *Ecosystems*. 9(2), tr. 181-199.
18. Alex M Remennikov và Timothy A Rose (2007), "Predicting the effectiveness of blast wall

barriers using neural networks", *International journal of impact engineering*. 34(12), tr. 1907-1923.

19. P Segarra và các cộng sự. (2010), "Prediction of near field overpressure from quarry blasting", *Applied Acoustics*. 71(12), tr. 1169-1176.

20. Wen Tingxin và Zhang Bo (2014), "Prediction model for open-pit coal mine slope stability based on random forest", *Science & Technology Review*. 32(4-5), tr. 105-109.

**Ngày nhận bài:** 25/07/2017

**Ngày gửi phản biện:** 19/8/2017

**Ngày nhận phản biện:** 20/09/2017

**Ngày chấp nhận đăng bài:** 10/12/2017

**Từ khóa:** sóng và đập không khí, học máy, Random Forest, thuật toán rừng ngẫu nhiên, mô hình dự báo, mỏ than Đèo Nai

## SUMMARY

Mechanization Blasting method is one of the methods to break rock in open-pit mine and its effects is undeniable. Beside the advantages of the blasting method such as the large volume of rock is broken, can break any type of rock, not depending on the climate, etc then the disadvantages of this method include air-blast overpressure, ground vibration and fly rock. They can make vibration of works, broken the glass of the doors and affecting of residential and environmental damages. Most of open-pit mine have facing to these problems. The paper presents an approach using Random Forest machine learning algorithm to predict the air-blast overpressure caused by blasting operation at Đèo Nai coal mine, Quảng Ninh. The data used in this case study were collected from 146 explosions of the mine. The results showed that Random Forest is a suitable algorithm for predicting air-blast overpressure at Đèo Nai coal mine with RMSE=1.519147 và  $R^2=0.941586$ . This study will help the blasters of Đèo Nai coal mine can adjust the parameters of blast fields and is the reference for the other open-pit mine, where the same condition with Đèo Nai coal mine.