



**Article info**

**Type of article:**

Original research paper

**DOI:**

<https://doi.org/10.58845/jstt.utt.2026.vn.6.1.30-47>

**\*Corresponding author:**

Email address:

[giangth@utt.edu.vn](mailto:giangth@utt.edu.vn)

**Received:** 28/01/2026

**Received in Revised Form:**

06/03/2026

**Accepted:** 09/03/2026

## Evaluating the Effect of Polypropylene Fibers on the Rutting Resistance of Asphalt Binder via Artificial Intelligence Modeling

Hoang Thi Huong Giang

Research group on Industry 4.0 in Transportation (I4T group), University of Transport Technology, Trieu Khuc, Thanh Liet, Hanoi 100000, Vietnam

**Abstract:** This study presents an advanced artificial intelligence-based approach for predicting the rutting factor ( $G^*/\sin\delta$ ) of Polypropylene (PP) fiber-modified asphalt, a critical parameter governing the plastic deformation resistance of pavements at elevated temperatures. Utilizing an experimental dataset of 132 samples synthesized from reputable publications, the CatBoost algorithm was implemented and fine-tuned using two hyperparameter optimization strategies- Grid Search and Random Search, integrated with 5-fold cross-validation. The results demonstrate that the optimized CatBoost model achieves superior predictive accuracy, yielding a coefficient of determination ( $R^2$ ) of 0.9967 and a Root Mean Square Error (RMSE) of 0.5720 on the independent testing set. Model-agnostic interpretability analysis via Partial Dependence Plots identified the initial rutting factor of the base asphalt and the fiber content as the two most influential features affecting material performance. Notably, the study delineated an optimal interaction zone between the mixing temperature (160-180°C) and PP fiber dosage (3-5%) to achieve the most effective reinforcement network structure. This research not only validates the efficacy of the CatBoost algorithm in capturing the complex non-linear relationships of composite materials but also provides a robust predictive tool to optimize design protocols and minimize experimental costs in modified asphalt technology.

**Keyword:** Polypropylene fiber (PP), machine learning, rutting factor ( $G^*/\sin\delta$ ), modified asphalt.



Thông tin bài viết  
Dạng bài viết:  
Bài báo nghiên cứu

DOI:  
<https://doi.org/10.58845/jstt.utt.2026.vn.6.1.30-47>

\*Tác giả liên hệ:  
Địa chỉ Email:  
[giangth@utt.edu.vn](mailto:giangth@utt.edu.vn)

Ngày nộp bài: 28/01/2026  
Ngày nộp bài sửa: 06/03/2026  
Ngày chấp nhận: 09/03/2026

## Đánh giá ảnh hưởng của sợi Polypropylene đến khả năng kháng lún của nhựa đường thông qua mô hình trí tuệ nhân tạo

Hoàng Thị Hương Giang

Nhóm nghiên cứu Ứng dụng công nghệ 4.0 trong Giao thông vận tải (I4T), Trường Đại học Công nghệ Giao thông vận tải, 54 Triều Khúc, Thanh Liệt, Hà Nội, Việt Nam

**Tóm tắt:** Nghiên cứu này trình bày một phương pháp tiếp cận tiên tiến dựa trên trí tuệ nhân tạo để dự báo hệ số hần lún ( $G^*/\sin\delta$ ) của nhựa đường biến tính sợi Polypropylene (PP), một chỉ tiêu quan trọng quyết định khả năng kháng biến dạng dẻo của mặt đường ở nhiệt độ cao. Dựa trên bộ dữ liệu thực nghiệm gồm 132 mẫu thu thập từ các công bố uy tín, thuật toán CatBoost đã được sử dụng và tối ưu hóa thông qua hai chiến lược tìm kiếm siêu tham số là Grid Search và Random Search kết hợp xác thực chéo 5 lần. Kết quả cho thấy mô hình CatBoost tối ưu đạt độ chính xác vượt trội với hệ số xác định  $R^2$  đạt 0.9967 và sai số RMSE chỉ 0.5720 trên tập dữ liệu kiểm tra. Phân tích kỹ thuật giải thích độc lập với mô hình đã xác định hệ số hần lún của nhựa ban đầu và hàm lượng sợi là hai yếu tố có tầm ảnh hưởng lớn nhất đến hiệu suất vật liệu. Ngoài ra, nghiên cứu đã chỉ ra vùng tương tác tối ưu giữa nhiệt độ khuấy (160-180°C) và hàm lượng sợi PP (3-5%) để đạt được cấu trúc mạng lưới gia cường hiệu quả nhất. Nghiên cứu này không chỉ khẳng định năng lực của thuật toán CatBoost trong việc xử lý các mối quan hệ phi tuyến phức tạp của vật liệu hỗn hợp mà còn cung cấp một công cụ dự báo đáng tin cậy, giúp tối ưu hóa quy trình thiết kế và giảm thiểu chi phí thực nghiệm trong công nghệ nhựa đường biến tính.

**Từ khóa:** Sợi Polypropylene (PP), học máy, hệ số hần lún ( $G^*/\sin\delta$ ); nhựa đường biến tính.

### 1. Giới thiệu

Sự gia tăng nhanh chóng của lưu lượng giao thông và tải trọng trục xe đã đặt ra những thách thức lớn đối cho độ bền lâu dài mặt đường bê tông nhựa [1–4]. Đặc biệt tại Việt Nam, các dự án giao thông trọng điểm như cao tốc Bắc - Nam đang đòi hỏi những tiêu chuẩn khắt khe về vật liệu để thích ứng với tải trọng nặng và nhiệt độ mặt đường cao. Việc ứng dụng nhựa đường biến tính đã và đang trở thành xu hướng chủ đạo trong các dự án này

nhằm đảm bảo tính bền vững của kết cấu. Trong đó, hiện tượng hần lún bánh xe ở nhiệt độ cao là một trong những hư hỏng điển hình ảnh hưởng đến hành trình và gia tăng chi phí bảo trì đường bộ. Để khắc phục vấn đề này, việc ứng dụng các loại sợi gia cường, đặc biệt là sợi Polypropylene (PP), đã trở thành một giải pháp đầy hứa hẹn nhờ khả năng hình thành mạng lưới không gian ba chiều giúp tăng cường độ cứng và cải thiện đặc tính nhót đàn hồi của chất kết dính [5–7]. Tuy

nhien, hiệu suất cải thiện sợi PP trong nhựa đường phụ thuộc vào một tổ hợp phức tạp các yếu tố như đặc tính nhựa đường ban đầu, hàm lượng sợi, chiều dài sợi và quy trình chế tạo nhiệt độ cao.

Mặc dù vai trò của sợi gia cường đã được khẳng định qua nhiều nghiên cứu thực nghiệm, việc xác định chính xác hệ số hàn lún ( $G^*/\sin\delta$ ) của nhựa đường biến tính hiện nay vẫn tồn tại nhiều hạn chế mặt kỹ thuật và chi phí [8–12]. Các phương pháp thí nghiệm truyền thống đòi hỏi quy trình chuẩn bị mẫu khá phức tạp, thiết bị đo hiện đại như máy đo cắt động (DSR) và tốn kém thời gian. Hạn chế này gây khó khăn cho việc tối ưu hóa nhanh cấp phối trong thực tế sản xuất. Hơn nữa, các mối quan hệ tương tác phi tuyến và sự phân tán không đồng nhất của sợi trong nhựa đường thường dẫn đến những sai số khó dự đoán khi sử dụng các mô hình hồi quy truyền thống. Chính sự hạn chế về khả năng dự báo nhanh và chi phí thí nghiệm cao đã tạo ra khoảng trống đáng kể trong việc ứng dụng các vật liệu biến tính sợi vào các dự án giao thông với quy mô lớn.

Trong những năm gần đây, sự phát triển mạnh mẽ của trí tuệ nhân tạo đã mở ra một kỷ nguyên mới trong việc mô phỏng và dự báo tính chất của vật liệu xây dựng với độ chính xác cao [13–16]. Tuy nhiên, các nghiên cứu hiện hữu về nhựa đường biến tính thường ứng dụng các mô hình hộp đen truyền thống như mạng nơ-ron nhân tạo (ANN) hoặc rừng ngẫu nhiên (RF) vẫn gặp nhiều khó khăn trong việc giải quyết các đặc trưng phân loại hoặc quá khớp khi làm việc với bộ dữ liệu thực nghiệm có kích thước trung bình [17, 18]. Đặc biệt, một khoảng trống nghiên cứu quan trọng là thiếu sự kết hợp giữa khả năng dự báo chính xác và tính giải thích được của mô hình, khiến cho các kỹ sư vật liệu khó có thể hiểu rõ cơ chế tương tác bên trong các thành phần bên trong cấu trúc nhựa đường từ kết quả của học máy. Để giải quyết các khoảng trống nêu trên, nghiên cứu này đề xuất thuật toán CatBoost - một kỹ thuật tăng cường độ dốc tiên tiến nhằm xây dựng mô hình dự báo hệ số hàn lún của nhựa đường biến tính sợi PP với độ tin cậy cao. CatBoost được lựa chọn nhờ giải quyết

tốt các mối quan hệ phi tuyến và độ ổn định vượt trội so với các thuật toán học máy thông thường. Bên cạnh việc tập trung vào độ chính xác dự báo, nghiên cứu này còn tiên phong ứng dụng kỹ thuật giải thích độc lập với mô hình để khai thác hộp đen mô hình, làm rõ tầm quan trọng và sự tương tác giữa các biến số như hàm lượng sợi, nhiệt độ trộn và đặc tính nhựa đường ban đầu. Bằng cách tích hợp 132 dữ liệu thực nghiệm đa dạng từ sáu nghiên cứu uy tín và sức mạnh tính toán của thuật toán CatBoost, nghiên cứu hướng tới việc thiết lập một công cụ thông minh trong thiết kế vật liệu. Kết quả không chỉ cung cấp một phương pháp dự báo hệ số hàn lún chính xác mà còn mang lại những hiểu biết sâu sắc về cơ chế biến tính vật liệu dưới góc nhìn của trí tuệ nhân tạo có thể giải thích (XAI). Đây là cơ sở khoa học quan trọng trong tối ưu hóa quy trình chế tạo và thúc đẩy ứng dụng các loại vật liệu tiên tiến trong hạ tầng giao thông hiện đại.

## 2. Cơ sở dữ liệu

Để huấn luyện mô hình, bộ dữ liệu về hệ số hàn lún của nhựa đường sử dụng sợi PP đã được tổng hợp từ các nghiên cứu thực nghiệm uy tín trên thế giới. Tổng cộng có 132 mẫu dữ liệu được thu thập, đảm bảo tính đa dạng và đại diện cho nhiều điều kiện chế tạo khác nhau. Chi tiết nguồn gốc và số lượng mẫu được tổng hợp ở Bảng 1.

Việc lựa chọn các biến đầu vào có ý nghĩa quyết định đến khả năng tổng quát hóa và độ chính xác của các mô hình trí tuệ nhân tạo. Trong nghiên cứu này, 7 tham số đầu vào đã được lựa chọn để dự báo hệ số hàn lún ( $G^*/\sin\delta$ , ký hiệu O) của nhựa đường biến tính PP, bao gồm các đặc tính của sợi gia cường, thông số trộn trong quá trình chế tạo chất kết dính, và đặc tính của nhựa ban đầu. Cụ thể các biến đầu vào là cường độ kéo của sợi (I1, MPa), hàm lượng sợi (I2, %), chiều dài sợi (I3, mm), nhiệt độ khuấy (I4, °C), thời gian khuấy (I5, phút), vận tốc khuấy (I6, vòng/phút), và hệ số hàn lún  $G^*/\sin\delta$  của nhựa đường ban đầu (I7, MPa). Dù các đặc tính vi mô có ảnh hưởng đến tương tác bề mặt, nghiên cứu này ưu tiên sử dụng các biến số vĩ mô để đảm bảo tính thực tiễn. Trong đó, hệ số hàn lún nhựa ban đầu đóng vai trò là tham số hợp

nhất các đặc tính cơ lý nền tảng. Việc tập trung vào các thông số hình học sợi và điều kiện chế tạo giúp mô hình dễ dàng tiếp cận dữ liệu thực tế tại hiện

trường. Điều này là cần thiết để xây dựng một công cụ dự báo nhanh, chi phí thấp và phù hợp với yêu cầu thực tế của ngành giao thông hiện nay.

**Bảng 1.** Tổng hợp dữ liệu được thu thập từ các nghiên cứu

STT	Nguồn thu thập	Số lượng mẫu
1	Yeh và cộng sự., 2005 [8]	20
2	Guo và cộng sự., 2022 [9]	24
3	Zhang và cộng sự., 2024 [10]	4
4	Li và cộng sự., 2023 [11]	12
5	Kathari và cộng sự., 2016 [19]	42
6	Xiaoming và cộng sự., 2019 [20]	30
<b>Tổng</b>		<b>132</b>

Phân tích đặc điểm thống kê dữ liệu tại Bảng 2 cho thấy các biến số bao phủ một dải giá trị rộng, từ các thông số hình học của sợi đến điều kiện nhiệt độ và vận tốc khuấy trộn, đảm bảo độ tin cậy

và phổ quát cho việc mô hình hóa. Cuối cùng, dữ liệu được chia thành 2 phần với 70% được sử dụng để xây dựng các mô hình, 30% còn lại được sử dụng kiểm tra hiệu suất mô hình.

**Bảng 2.** Phân tích thống kê bộ dữ liệu sử dụng trong nghiên cứu

Tên biến	Cường độ kéo của sợi	Hàm lượng sợi	Chiều dài sợi	Nhiệt độ khuấy	Thời gian khuấy	Vận tốc khuấy	G*/sinδ ban đầu	G*/sinδ
Ký hiệu	I1	I2	I3	I4	I5	I6	I7	O
Đơn vị	MPa	%	mm	°C	phút	Vòng/phút	kPa	kPa
Trung bình	438.636	2.034	4.046	137.197	44.048	1547.159	8.114	12.793
Độ lệch chuẩn	237.706	1.648	3.508	60.141	30.346	1053.795	29.716	39.094
Nhỏ nhất	0.000	0.000	0.000	25.000	0.000	0.000	0.100	0.100
10%	0.000	0.000	0.000	25.000	0.000	0.000	0.200	0.494
20%	0.000	0.000	0.000	25.000	0.000	0.000	0.300	0.793
30%	537.500	1.000	0.100	160.000	30.000	500.000	0.493	1.130
40%	537.500	1.320	3.500	160.000	40.000	2200.000	0.750	1.533
50%	537.500	2.000	3.500	165.000	40.000	2200.000	1.100	2.387
60%	537.500	2.500	6.000	170.000	60.000	2200.000	1.500	3.188
70%	537.500	3.000	6.000	170.000	60.000	2200.000	2.413	4.978
80%	625.000	3.500	6.000	174.000	60.000	2500.000	4.040	9.384
90%	625.000	5.000	10.000	180.000	90.000	2500.000	17.280	21.620
Lớn nhất	625.000	5.000	10.000	180.000	90.000	2575.000	220.000	320.000

Ngoài ra, mối liên hệ giữa các biến đầu vào và đầu ra, cùng với đặc trưng phân phối của tập dữ liệu thực nghiệm, được thể hiện chi tiết thông qua ma trận tương quan (Hình 1) và biểu đồ phân phối tần suất (Hình 2).

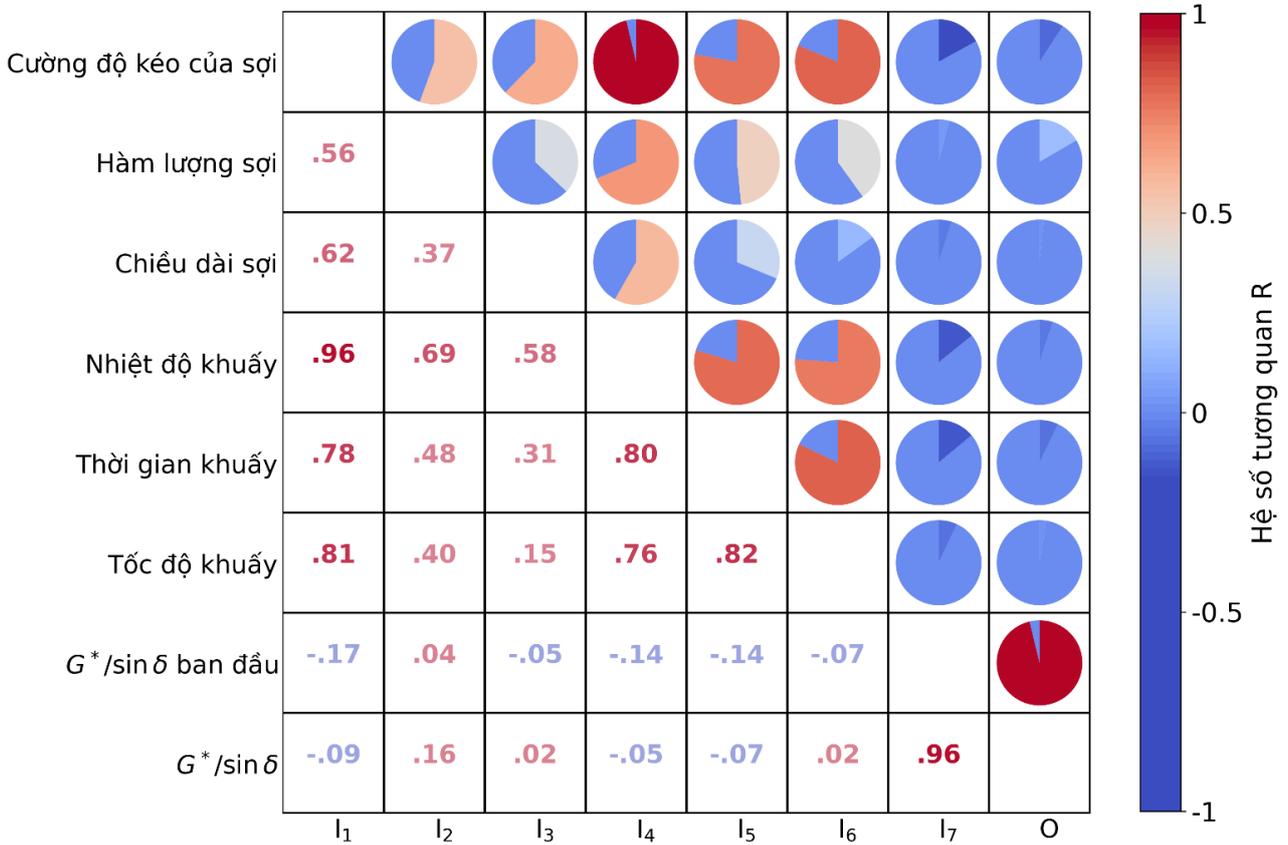
Kết quả từ Hình 1 cho thấy biến G\*/sinδ ban đầu (I7) có hệ số tương quan thuận mạnh nhất đối với giá trị đầu ra G\*/sinδ (O) của nhựa đường biến tính sợi PP với hệ số tương quan là 0.96. Điều này phản ánh quy luật vật lý khách quan khi đặc tính cơ lý của nhựa nền đóng vai trò tiên quyết đến hiệu

suất của hỗn hợp sau khi biến tính bằng PP. Bên cạnh đó, các thông số chuẩn bị mẫu như nhiệt độ khuấy (I4), thời gian khuấy (I5) và vận tốc khuấy (I6) thể hiện sự độc lập tương đối với hệ số tương quan thấp. Điều này đảm bảo tính khách quan, phổ quát và độ tin cậy cho quá trình huấn luyện mô hình học máy.

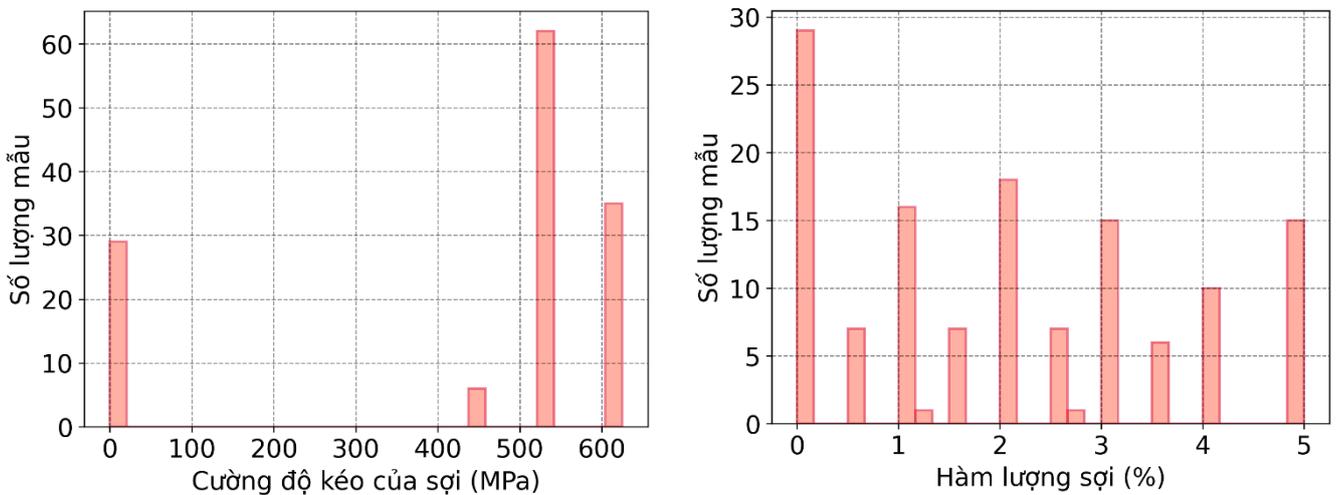
Phân tích sâu hơn về cấu trúc dữ liệu tại Hình 2, có thể thấy các biến đầu vào được phân bố trải rộng trên nhiều dải giá trị thực nghiệm khác nhau. Cụ thể, các thông số kỹ thuật như hàm lượng sợi

(12) hay nhiệt độ khuấy (14) không chỉ tập trung ở các giá trị phổ biến mà còn bao phủ cả các điểm dữ liệu ở biên. Đối với biến đầu ra (Hình 2h), phân phối giá trị cho thấy một dải đo rộng từ 0.1 đến 320 kPa với mật độ tập trung cao ở vùng giá trị thấp. Đặc điểm lệch phải của phân phối này là một thách

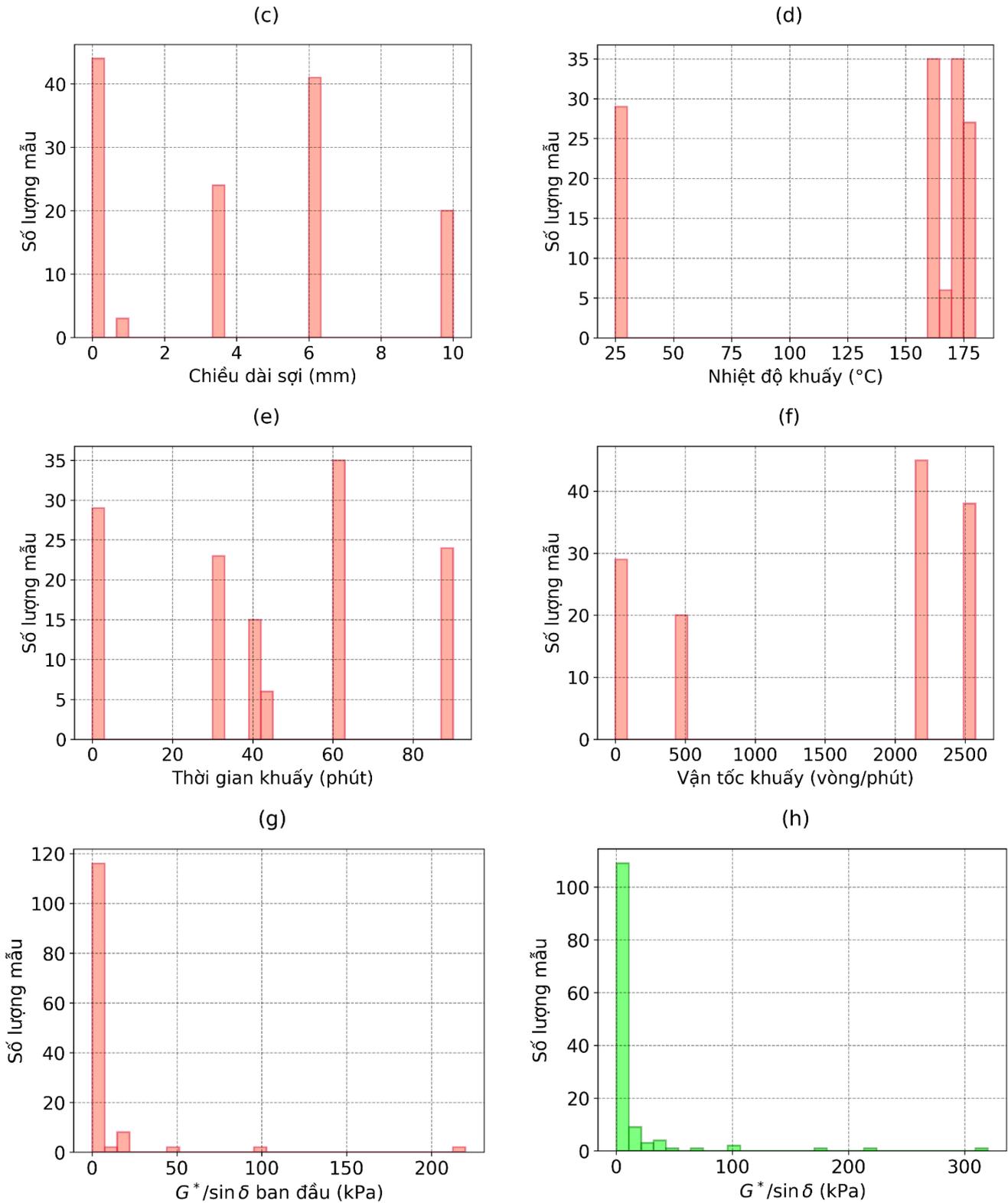
thức lớn trong việc dự báo, đòi hỏi mô hình học máy có khả năng xử lý tốt vấn đề phi tuyến, đảm bảo độ chính xác trên toàn bộ dữ liệu. Những phân tích định lượng nêu trên là cơ sở quan trọng để thực hiện các bước xây dựng và tối ưu hóa mô hình trong phần tiếp theo của nghiên cứu.



Hình 1. Tương quan của các biến đầu vào, đầu ra trong bộ dữ liệu (a) (b)



Hình 2. Phân phối dữ liệu các biến đầu vào, đầu ra của bộ dữ liệu



Hình 2 (tiếp). Phân phối dữ liệu các biến đầu vào, đầu ra của bộ dữ liệu

### 3. Phương pháp học máy

#### 3.1. Thuật toán CatBoost

Thuật toán CatBoost là một bước tiến hiện đại trong dòng thuật toán tăng cường độ dốc dựa trên cây quyết định. Đặc điểm nổi bật nhất của

thuật toán này là khả năng xử lý trực tiếp các biến phân loại và dữ liệu hồi quy mà không cần các bước tiền xử lý phức tạp, đồng thời khắc phục hiện tượng dịch chuyển dự báo và giảm thiểu rủi ro quá khớp [21]. Đối với tập dữ liệu thực nghiệm về nhựa

đường biến tính, thuật toán này thể hiện khả năng xử lý các mối quan hệ phi tuyến mạnh mẽ giữa các thông số chế tạo và đặc tính cơ lý của vật liệu. Cấu trúc cây đối xứng đặc trưng của CatBoost không chỉ giúp tăng tốc độ dự báo mà còn đảm bảo tính ổn định của mô hình, đặc biệt với các bộ dữ liệu vừa và nhỏ.

### 3.2. Xác thực chéo

Xác thực chéo (Cross-validation - CV) đóng vai trò là một cơ chế kiểm chứng cơ bản trong lĩnh vực học máy đối với dữ liệu chưa từng tiếp cận [22]. Quy trình chuẩn thường bao gồm việc phân chia tập dữ liệu thành các tập con riêng biệt dành cho huấn luyện, xác thực và kiểm chứng. Trong các kịch bản mà dữ liệu ban đầu đã được chia thành hai phần huấn luyện và kiểm tra, quy trình CV sẽ được thực hiện duy nhất trên tập huấn luyện. Tập hợp này được chia nhỏ thành  $k$  phần rời rạc với kích thước xấp xỉ bằng nhau. Quá trình này sau đó được thực hiện lặp đi lặp lại: tại mỗi vòng lặp,  $k-1$  phần được dùng để huấn luyện mô hình, trong khi phần còn lại đóng vai trò là tập xác thực để kiểm tra hiệu suất. Điều quan trọng là tập kiểm tra được để riêng trong suốt giai đoạn này nhằm đảm bảo kết quả đánh giá cuối cùng về mô hình trên dữ liệu mới là khách quan. Kết quả dự báo tổng thể của mô hình được định lượng bằng cách tính trung bình cộng các chỉ số hiệu suất thu được qua tất cả  $n$  lần lặp. Để đảm bảo đánh giá chắc chắn về khả năng tổng quát hóa, nghiên cứu này áp dụng chiến lược xác thực chéo 5 lần.

### 3.3. Tìm kiếm lưới

Tìm kiếm lưới (Grid search) là một kỹ thuật tối ưu hóa tham số có hệ thống, hoạt động bằng cách thiết lập một không gian tìm kiếm dưới dạng lưới tọa độ dựa trên các giá trị định sẵn của từng tham số. Phương pháp này sẽ thực hiện đánh giá hiệu suất của mô hình trên tất cả các kết hợp có thể có trong không gian này. Ưu điểm nổi bật của tìm kiếm lưới là tính triệt để, đảm bảo không bỏ sót bất kỳ tổ hợp tham số nào trong phạm vi được xác định. Tuy nhiên, chi phí tính toán của tìm kiếm lưới tỉ lệ thuận với số lượng tham số và các mức giá trị, dẫn đến thời gian tìm kiếm kéo dài khi phạm vi tìm

kiếm mở rộng [23].

### 3.4. Tìm kiếm ngẫu nhiên

Khác với tìm kiếm lưới, phương pháp tìm kiếm ngẫu nhiên (Random search) không thực hiện thử nghiệm mọi kết hợp mà chọn lọc giá trị tham số một cách ngẫu nhiên từ không gian tìm kiếm được phân phối. Hiệu suất tìm kiếm theo phương pháp này hiệu quả hơn về mặt thời gian và tài nguyên tính toán, đặc biệt khi không gian tìm kiếm lớn. Bằng cách lấy mẫu ngẫu nhiên, thuật toán có khả năng khám phá các vùng không gian tiềm năng mà phương pháp tìm kiếm lưới có thể bỏ qua do giới hạn về bước nhảy của lưới nhưng vẫn đạt được độ chính xác tương đương với tìm kiếm lưới.

### 3.5. Biểu đồ phụ thuộc một phần

Biểu đồ phụ thuộc một phần (PDP) là một kỹ thuật giải thích độc lập với mô hình, nhằm mục đích minh họa bằng đồ họa tác động biên trung bình của các đặc điểm đầu vào cụ thể đối với kết quả dự đoán của một mô hình học máy đã được huấn luyện [24]. PDP một chiều (PDP1D) được dùng để kiểm tra tác động biên của một đặc điểm duy nhất, tạo ra một biểu đồ đường thể hiện bản chất của sự phụ thuộc; trong khi PDP hai chiều (PDP2D) rất quan trọng để kiểm tra tác động tương tác giữa hai đặc điểm cụ thể. Quá trình tính toán cho cả hai phương pháp đều bao gồm việc thiết lập có hệ thống các đặc điểm đang được nghiên cứu thành các giá trị từ một lưới được xác định trước, tính trung bình các dự đoán thu được cho mỗi trường hợp dữ liệu và trình bày đầu ra 2D dưới dạng bản đồ nhiệt, trong đó các sắc độ khác nhau thể hiện độ lớn của kết quả dự đoán.

### 3.6. Các tiêu chí đánh giá

Để đưa ra đánh giá toàn diện về khả năng dự báo của các mô hình học máy đã phát triển, nghiên cứu này sử dụng một bộ chỉ số đánh giá đa dạng. Khung đánh giá bao gồm: Hệ số xác định ( $R^2$ ), sai số bình phương trung bình căn (RMSE), sai số tuyệt đối trung bình (MAE), và sai số phần trăm tuyệt đối trung bình (MAPE).

$R^2$  đóng vai trò là một chỉ số cơ bản giúp đánh giá mức độ phù hợp của mô hình. Mặc dù giá

trị  $R^2$  có giá trị [0; 1] với giá trị cao cho thấy mức độ phù hợp tốt hơn, nhưng một chỉ số  $R^2$  cao đơn thuần không đảm bảo khả năng tổng quát của mô hình. Các mô hình có thể gặp hiện tượng quá khớp, dẫn đến điểm  $R^2$  cao trên tập dữ liệu đã được học nhưng lại thất bại trên dữ liệu mới, điều này cho thấy sự cần thiết của việc sử dụng nhiều chỉ số bổ trợ khác nhau. RMSE được tính toán bằng căn bậc hai của trung bình bình phương sai lệch giữa giá trị dự đoán và giá trị thực tế. Chỉ số này trừng phạt mạnh mẽ các độ lệch lớn, khiến nó đặc biệt nhạy cảm với các giá trị ngoại lai. RMSE càng thấp biểu thị mô hình càng phù hợp và có độ chính xác cao. MAE là trung bình cộng của các giá trị tuyệt đối giữa sai lệch giữa giá trị dự báo và giá trị thực tế. Ngược lại với RMSE, MAE cung cấp một điểm số tuyến tính trong đó các sai số có trọng số như nhau, phản ánh trực tiếp mức độ lớn bình quân của sai số. Giá trị này dao động từ 0 (phù hợp hoàn hảo) đến dương vô cùng.

Cuối cùng, sai số tuyệt đối trung bình MAPE là một chỉ số không phụ thuộc vào quy mô, dùng

để so sánh hiệu suất mô hình trên các tập dữ liệu hoặc các bậc quy mô khác nhau. Tỷ lệ phần trăm MAPE càng thấp phản ánh độ chính xác dự báo càng cao. Công thức xác định các chỉ số hiệu suất trên có thể tham khảo trong tài liệu trích dẫn [12].

#### 4. Kết quả và thảo luận

##### 4.1. Xây dựng mô hình CatBoost

Quá trình xây dựng mô hình CatBoost bắt đầu bằng việc thiết lập không gian tìm kiếm cho các siêu tham số cốt lõi nhằm tối ưu hóa khả năng dự báo hệ số hàn lún của nhựa đường biến tính sợi PP. Ba tham số quan trọng được tối ưu bao gồm: tốc độ học (learning\_rate), độ sâu của cây (depth) và số vòng lặp (iteration) với các khoảng giá trị tìm kiếm được chi tiết trong Bảng 3. Trong đó, tốc độ học được giới hạn từ 0.05-0.3 để cân bằng giữa tốc độ hội tụ và độ chính xác của mô hình, độ sâu của cây thiết lập (3-8) nhằm tránh cấu trúc cây quá phức tạp gây hiện tượng quá khớp, và số vòng lặp dao động từ 50 đến 500 cho phép thuật toán đủ số bước lặp để tối ưu hàm mất mát.

**Bảng 3.** Không gian tìm kiếm các siêu tham số mô hình

STT	Ký hiệu	Tham số	Giá trị tìm kiếm
1	L	learning_rate	0.05-0.3
2	D	depth	3-8
3	I	iteration	50-500

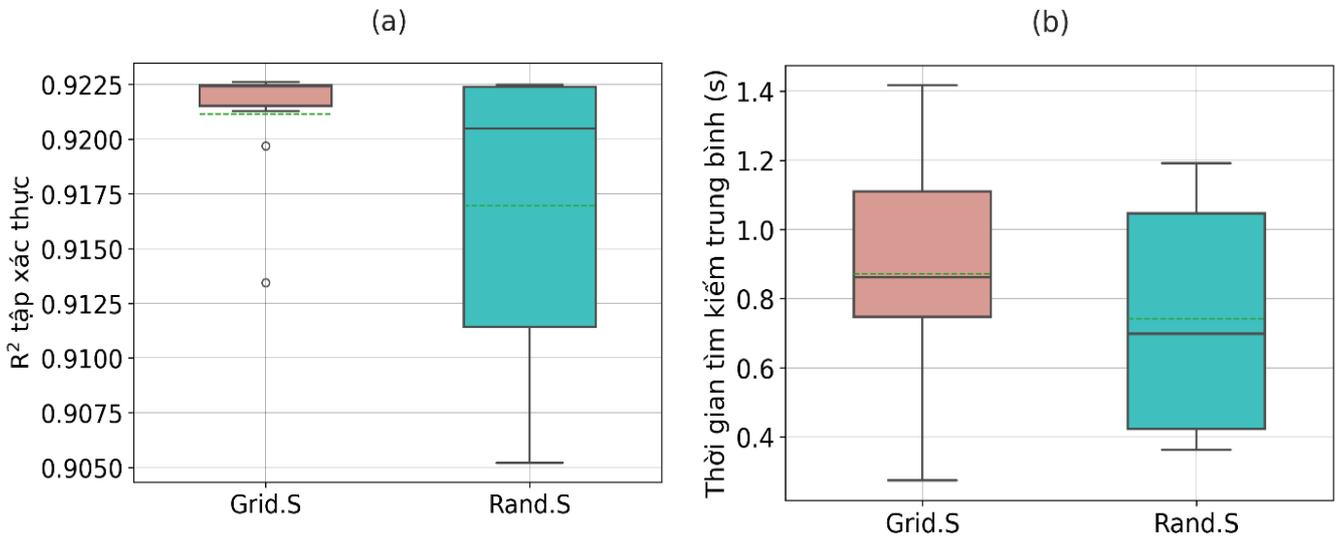
Để tìm ra bộ tham số tối ưu, nghiên cứu này sử dụng hai chiến lược tìm kiếm. Kết quả tại Bảng 4 cho thấy cả hai chiến lược tìm kiếm đều đạt được hiệu suất dự báo rất cao với giá trị  $R^2$  lớn nhất đạt 0.9226 trên tập xác thực. Cụ thể, mô hình tối ưu (xếp hạng 1) được xác lập với cấu hình tham số: L = 0.1; D = 4 và I = 500. Đáng chú ý, các mô hình xếp hạng cao nhất của cả hai phương pháp đều có sự tương đồng lớn về chỉ số hiệu suất, cho thấy sự ổn định của mô hình CatBoost. Hiệu suất của hai chiến lược tìm kiếm được phân tích trực quan qua Hình 3. Qua biểu đồ so sánh giữa chỉ số  $R^2$  và thời gian tìm kiếm, có thể nhận thấy phương pháp tìm kiếm ngẫu nhiên thể hiện ưu thế vượt trội về mặt tốc độ xử lý khi tìm ra các vùng không gian tham số hiệu quả nhanh hơn đáng kể so với tìm kiếm

lưới, trong khi vẫn đảm bảo độ chính xác tương đương. Tuy nhiên, cần lưu ý rằng dù tìm kiếm lưới và tìm kiếm ngẫu nhiên đã cung cấp các bộ tham số đạt độ chính xác cao cho mô hình CatBoost trong nghiên cứu này, việc ứng dụng các thuật toán tối ưu hóa như tối ưu hóa bầy đàn (PSO) hay thuật toán tối ưu hóa sói xám (GWO) hứa hẹn sẽ mang lại khả năng hội tụ nhanh và chính xác hơn ở các không gian tham số phức tạp. Đây là hướng nghiên cứu tiềm năng để lai tạo các mô hình CatBoost tối ưu, giúp tự động hóa hoàn toàn quy trình tinh chỉnh cấu trúc cây đối xứng và nâng cao năng lực tổng quát hóa của mô hình đối với các tập dữ liệu vật liệu nhựa đường quy mô lớn hơn trong tương lai.

Sự ổn định và chính xác của mô hình tiếp tục được khẳng định qua Hình 4 và Bảng 5 thông qua

hai tiêu chí then chốt là RMSE và  $R^2$ . Với tập huấn luyện, cả 5 mô hình tiêu biểu đều đạt giá trị  $R^2$  xấp xỉ tuyệt đối (0.9999) và sai số RMSE rất thấp (từ 0.2230 đến 0.4738). Hơn nữa, mô hình còn được đánh giá khả năng tổng quát hóa trên tập kiểm tra khi mô hình 1 đạt giá trị  $R^2 = 0.9967$  và RMSE =

0.5720. Các chỉ số MAE (0.4210) và MAPE (0.3764) của mô hình 1 trên tập kiểm tra đều ở mức thấp nhất trong số các mô hình, minh chứng cho độ tin cậy vượt trội khi dự báo hệ số hàn lún của nhựa đường biến tính sợi PP cho các dữ liệu thực nghiệm mới.



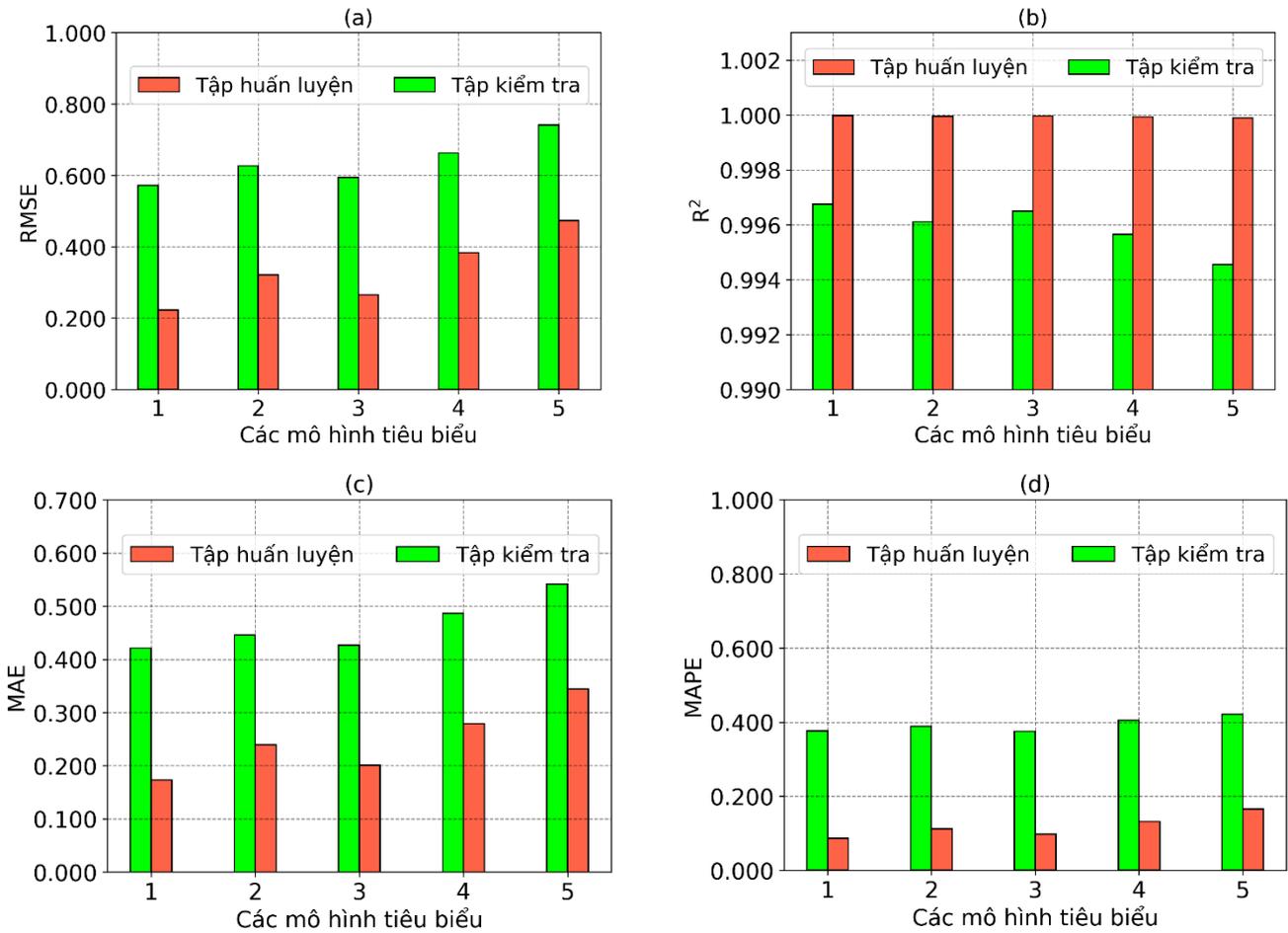
**Hình 3.** So sánh hiệu suất tìm kiếm của Grid.S và Rand.S theo tiêu chí  $R^2$  và thời gian tìm kiếm 10 mô hình tốt nhất trên tập dữ liệu xác thực

**Bảng 4.** Kết quả tìm kiếm siêu tham số mô hình theo hai phương pháp trên tập dữ liệu xác thực

Phương pháp	Tìm kiếm lưới (Grid.S)				Tìm kiếm ngẫu nhiên (Rand.S)			
	Xếp hạng mô hình	L	D	I	$R^2$	L	D	I
1	0.1	4	500	0.9226	0.1	4	500	0.9226
2	0.1	4	400	0.9225	0.1	4	400	0.9225
3	0.1	4	450	0.9225	0.1	4	300	0.9224
4	0.1	4	350	0.9225	0.05	3	500	0.9222
5	0.1	4	300	0.9224	0.1	8	300	0.9213
6	0.1	4	250	0.9224	0.1	5	400	0.9200
7	0.05	3	500	0.9222	0.1	8	500	0.9134
8	0.1	8	300	0.9213	0.1	3	500	0.9108
9	0.1	8	350	0.9197	0.1	8	200	0.9100
10	0.1	8	500	0.9134	0.1	3	250	0.9052

**Bảng 5.** Tổng hợp giá trị metric của 5 mô hình tiêu biểu theo bốn tiêu chí

Tiêu chí	$R^2$	RMSE	MAE	MAPE	$R^2$	RMSE	MAE	MAPE
Mô hình	Tập huấn luyện				Tập kiểm tra			
1	0.9999	0.2230	0.1728	0.0865	0.9967	0.5720	0.4210	0.3764
2	0.9999	0.3209	0.2391	0.1117	0.9961	0.6272	0.4459	0.3889
3	0.9999	0.2651	0.2008	0.0973	0.9965	0.5947	0.4271	0.3752
4	0.9999	0.3832	0.2791	0.1322	0.9957	0.6627	0.4865	0.4052
5	0.9999	0.4738	0.3439	0.1656	0.9946	0.7418	0.5417	0.4217



**Hình 4.** Hiệu suất dự báo của 5 mô hình tiêu biểu theo các tiêu chí: (a) RMSE, (b) R<sup>2</sup>; (c) MAE; (d) MAPE

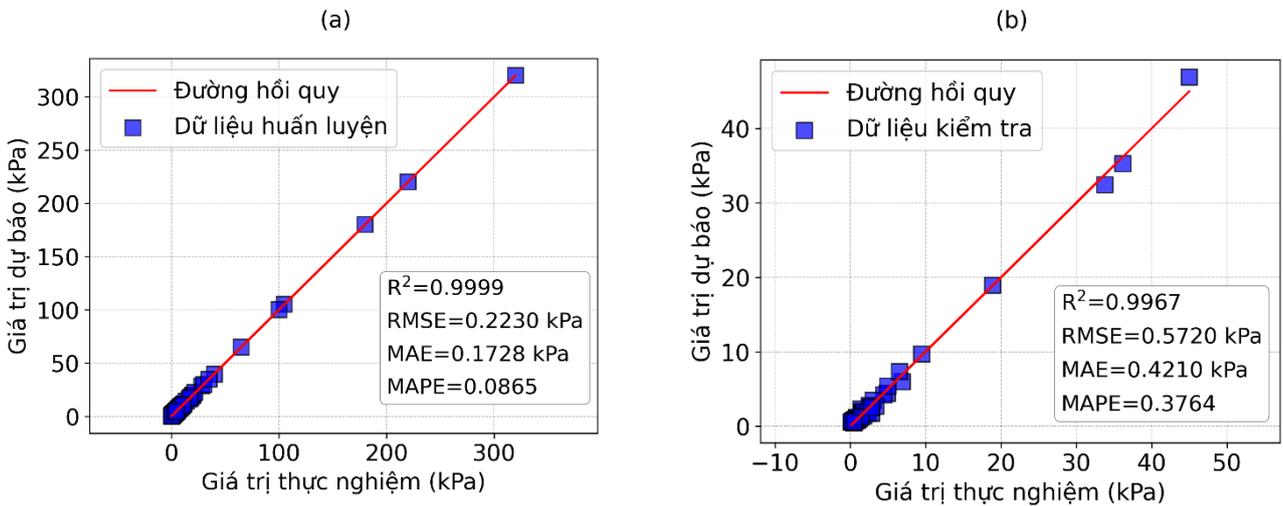
**4.2. Kết quả dự báo mô hình tiêu biểu**

Phần này trình bày hiệu suất dự báo của mô hình CatBoost tối ưu được đánh giá thông qua biểu đồ hồi quy và biểu đồ sai số. Hình 5 thể hiện sự tương quan giữa giá trị dự báo từ mô hình và giá trị thực nghiệm, được trình bày chi tiết cho cả hai tập huấn luyện (Hình 5a) và tập kiểm tra (Hình 5b). Dễ dàng nhận thấy, các dữ liệu phân bố rất sát đường chéo (đường lý tưởng), minh chứng cho khả năng mô hình hóa chính xác các quy luật biến thiên của vật liệu. Với hệ số xác định R<sup>2</sup> đạt giá trị 0.9999 trên tập huấn luyện và duy trì ở mức cao 0.9967 trên tập kiểm tra, mô hình CatBoost không chỉ thể hiện khả năng học sâu các đặc trưng của dữ liệu cũ mà còn khẳng định năng lực tổng quát hóa xuất sắc đối với các mẫu dữ liệu mới. Khoảng cách rất nhỏ giữa các đường hồi quy thực tế và đường lý tưởng là minh chứng cho việc mô hình đã thu giữ được gần như toàn bộ các tương quan

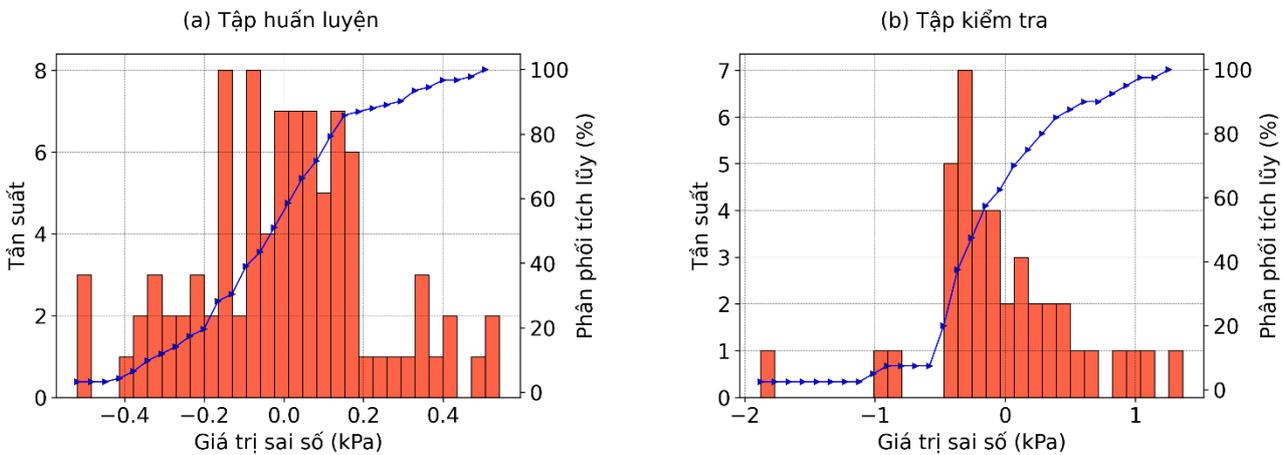
phi tuyến giữa thông số đầu vào và đầu ra. Các kết quả này là minh chứng quan trọng về độ tin cậy của mô hình khi áp dụng vào bài toán dự báo hệ số đàn hồi cho nhựa đường biến tính sợi PP. Ngoài ra, phân tích chi tiết biểu đồ sai số tại Hình 6 cho thấy mô hình CatBoost đạt độ chính xác rất cao với các giá trị sai số tập trung mật độ lớn xung quanh đường tiệm cận 0. Sự hội tụ sai số ấn tượng khi hầu hết các dữ liệu đều nằm trong dải sai số hẹp. Cụ thể, trên tập huấn luyện, phần lớn các mẫu có sai số nằm trong khoảng từ - 0.4 đến + 0.4 kPa, có khoảng 8 mẫu trong tổng số 92 mẫu có sai số nằm ngoài khoảng trên. Với tập kiểm tra, mặc dù độ biến thiên có tăng nhẹ nhưng tính ổn định vẫn được duy trì khi số lượng mẫu nằm ngoài khoảng sai số [-1; 1] kPa là rất ít, khoảng 3 mẫu trong tổng số 40 mẫu. Nguyên nhân của các sai số này chủ yếu xuất phát từ tính không đồng nhất của vật liệu trong quá trình chế tạo mẫu thực nghiệm, đặc biệt

là hiện tượng phân tán sợi không đều tại một số điều kiện biên. Điều này cho thấy mô hình mặc dù có độ chính xác rất cao nhưng vẫn phản ánh đúng những thách thức khách quan trong thực nghiệm vật liệu nhựa đường biến tính. Đặc biệt, khi xét đến phân phối sai số tích lũy, mô hình thể hiện năng lực dự báo vượt trội với trên 80% các sai số nằm trong

khoảng  $[-0.3; 0.3]$  kPa đối với tập huấn luyện, và trên 80% sai số nằm trong khoảng từ  $[-0.5; 0.5]$  kPa. Với sai số rất thấp, mô hình CatBoost được đánh giá tốt khi dự báo hệ số đàn hồi của nhựa đường biến tính PP. Mô hình này có thể được áp dụng vào thực tế để giúp các kỹ sư vật liệu tiết kiệm thời gian và chi phí trước khi làm thí nghiệm.



**Hình 5.** Tương quan giữa giá trị thực nghiệm và giá trị dự báo hệ số đàn hồi cho (a) tập huấn luyện; (b) tập kiểm tra



**Hình 6.** Sai số dự báo giữa giá trị thực nghiệm và giá trị dự báo hệ số đàn hồi cho (a) tập huấn luyện; (b) tập kiểm tra

**4.3. Phân tích độ nhạy**

Phần này trình bày biểu đồ ảnh hưởng từng phần nhằm cung cấp sâu hơn về sự ảnh hưởng của từng biến đầu vào đơn lẻ đối với giá trị dự báo hệ số đàn hồi. Kết quả từ Hình 7g và Bảng 6 cho thấy biến  $G^*/\sin\delta$  của nhựa gốc (17) thể hiện biên độ tác động lớn nhất (từ 2.0 đến 200 kPa), khẳng định tính chất của nhựa đường ban đầu là khung xương quyết định khả năng kháng biến dạng của

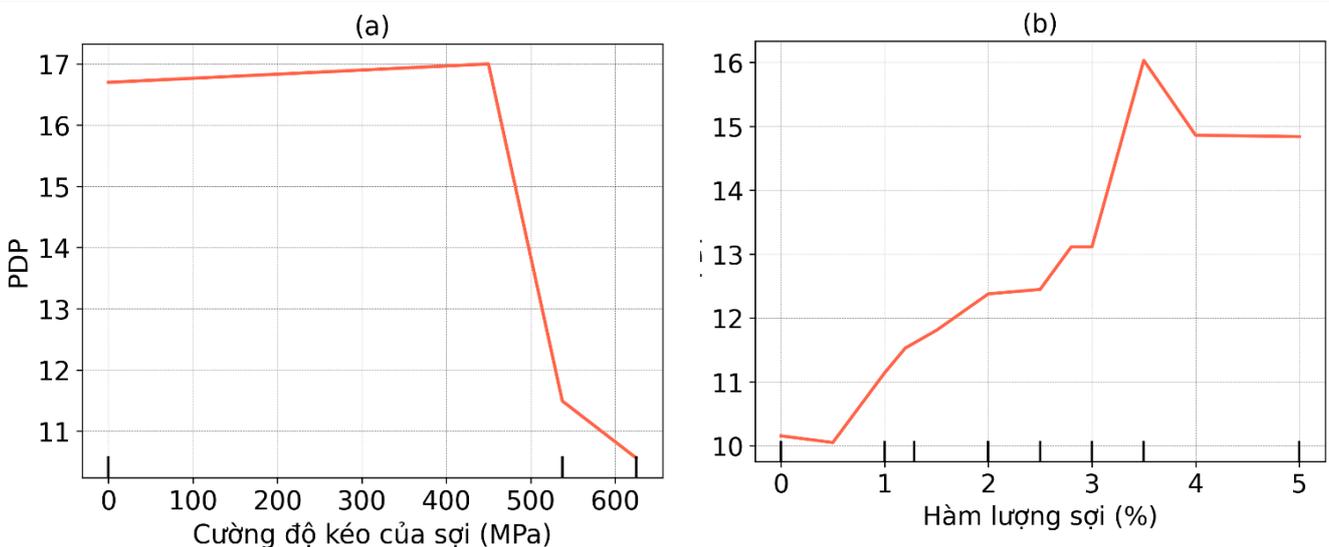
nhựa đường biến tính PP. Kết quả này hoàn toàn tương đồng với các nghiên cứu của Hoàng và cộng sự, khi cho rằng đặc tính nhớt đàn hồi của nhựa đường ban đầu thiết lập nền tảng cho mọi sự cải thiện hiệu suất sau này [25]. Đối với hàm lượng sợi I2 (Hình 7b), biểu đồ cho thấy xu hướng tăng dần của hệ số đàn hồi khi nồng độ sợi tăng từ 0% đến 5%. Điều này có thể giải thích, sự hiện diện của sợi PP tạo ra một mạng lưới không gian ba chiều trong

nhựa đường, giúp hấp thụ ứng suất và hạn chế sự dịch chuyển tương đối giữa các phân tử nhựa ở nhiệt độ cao. Hiện tượng này làm tăng độ cứng và khả năng hồi phục đàn hồi, từ đó cải thiện đáng kể chỉ số hần lún. Bên cạnh đó, cường độ kéo của sợi I1 (Hình 7a) xếp vị trí thứ ba về tầm quan trọng với hiệu số tác động là 5.5. Điều này giải thích rằng các loại sợi có khả năng chịu lực cao hơn sẽ cung cấp sự gia cường tốt hơn cho cấu trúc chất kết dính khi chịu tải trọng bánh xe lặp lại. Ngược lại, nhiệt độ khuấy I4 (Hình 7d) và thời gian khuấy I5 (Hình 7e) cho thấy sự tác động có ngưỡng. Khi

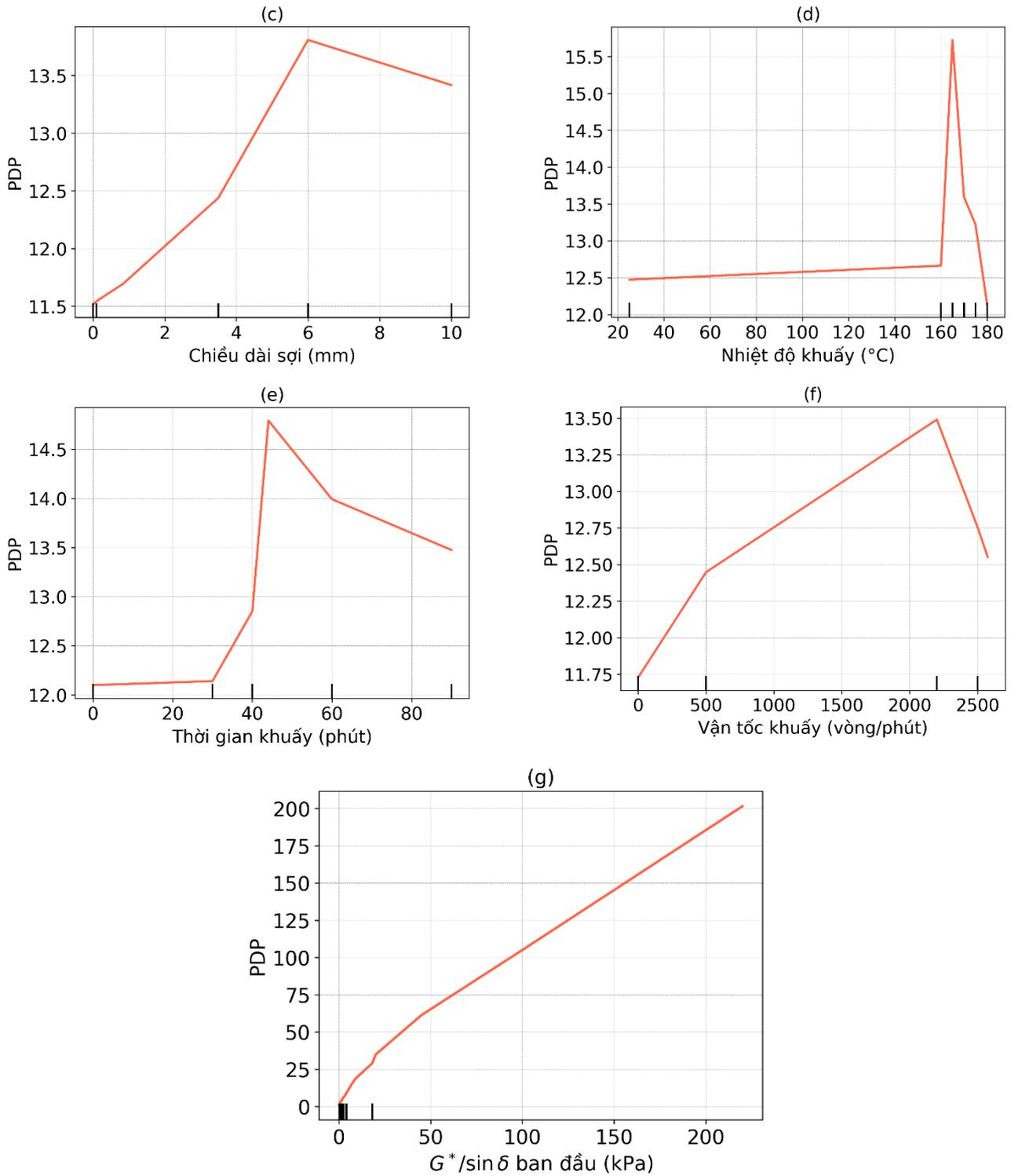
nhiệt độ và thời gian đạt tới một giá trị tối ưu, hệ số hần lún có xu hướng ổn định. Tuy nhiên, nếu vượt quá ngưỡng này, nhiệt độ cao có thể gây ra hiện tượng hóa già sớm hoặc phân hủy nhiệt một phần của sợi PP, làm giảm hiệu quả biến tính. Các biến vận tốc khuấy (I6) và chiều dài sợi (I3) có mức độ ảnh hưởng thấp nhất (xếp hạng 6 và 7) gợi ý rằng trong dải thực nghiệm khảo sát, các yếu tố này chủ yếu đóng vai trò đảm bảo sự phân tán đồng nhất của sợi. Điều này hoàn toàn phù hợp với những nhận định trong nghiên cứu [26] khi sử dụng nhựa đường biến tính vật liệu nano.

**Bảng 6.** Tổng hợp tác động và thứ tự ưu tiên của các biến đầu vào đối với hệ số hần lún dựa trên phân tích PDP

Biến đầu vào	PDP 1D			
	Lớn nhất	Nhỏ nhất	Hiệu số	Xếp hạng
Cường độ kéo của sợi	17.0	11.5	5.5	3
Hàm lượng sợi	16.1	10.0	6.1	2
Chiều dài sợi	13.8	11.5	2.3	6
Nhiệt độ khuấy	15.7	12.5	3.2	4
Thời gian khuấy	14.8	12.2	2.6	5
Vận tốc khuấy	13.5	11.8	1.7	7
G*/sinδ ban đầu	200	2.0	198	1



**Hình 7.** Biểu đồ ảnh hưởng từng phần (PDP 1D) thể hiện tác động của các biến đầu vào đơn lẻ đến hệ số hần lún



**Hình 7 (tiếp).** Biểu đồ ảnh hưởng từng phần (PDP 1D) thể hiện tác động của các biến đầu vào đơn lẻ đến hệ số hàn lún

Để làm rõ cơ chế tác động phức tạp của các thành phần vật liệu và điều kiện chế tạo, Hình 8 trình bày các biểu đồ ảnh hưởng riêng phần hai chiều cho các cặp biến quan trọng nhất trong 4 biến quan trọng nhất, bao gồm hệ số hàn lún của nhựa đường ban đầu, hàm lượng sợi, cường độ

kéo của sợi và nhiệt độ khuấy. Sự tương tác giữa các biến này không chỉ mang tính cộng dồn đơn thuần mà còn thể hiện các mối quan hệ phi tuyến sâu sắc, phản ánh bản chất phức tạp của cấu trúc vật liệu hỗn hợp chất kết dính. Phân tích sự tương tác giữa hệ số hàn lún của nhựa đường ban đầu

và hàm lượng sợi (Hình 8a) cho thấy một vùng hiệu suất tăng vọt khi cả hai thông số này đồng thời đạt giá trị cao. Về mặt vật liệu, khi nhựa nền có độ cứng ban đầu lớn, việc bổ sung sợi PP với hàm lượng tối ưu từ 3-5% sẽ tạo ra một mạng lưới gia cường ổn định, giúp neo giữ các phân tử nhựa và hạn chế sự trượt dẻo ở nhiệt độ cao. Kết quả này hoàn toàn phù hợp với các nghiên cứu [27] khi khẳng định rằng sợi gia cường đóng vai trò là khung xương thứ cấp, phát huy hiệu quả tối đa khi kết hợp với nhựa đường có chất lượng gốc tốt.

Hình 8b là sự kết hợp giữa hệ số đàn hồi của nhựa gốc và cường độ kéo của sợi cho thấy một quy luật tác động hỗ tương rõ rệt. Khi nhựa đường ban đầu có chỉ số  $G^*/\sin\delta$  cao kết hợp với loại sợi có cường độ kéo lớn, bề mặt đáp ứng đạt tới điểm cực đại về khả năng kháng biến dạng. Điều này có thể giải thích, nhựa đường ban đầu đóng vai trò là pha liên tục truyền ứng suất, trong khi sợi PP đóng vai trò là pha gia cường chịu ứng suất kéo phát sinh. Sự tương tác này phản ánh rằng các loại nhựa đường có độ cứng cao sẽ tạo ra môi trường bám dính tốt hơn, cho phép sợi PP đảm bảo khả năng chịu kéo để ngăn chặn sự trượt dẻo của hỗn hợp.

Bên cạnh đó, nhiệt độ khuấy thể hiện vai trò là yếu tố kiểm soát quá trình phân tán và liên kết bề mặt. Hình 8c chỉ ra rằng có một vùng nhiệt độ tối ưu (160-180°C) để phát huy tác dụng của sợi gia cường. Ở dải nhiệt độ này, độ nhớt của nhựa đường đủ thấp để sợi PP phân tán đồng nhất, đồng thời năng lượng nhiệt giúp cải thiện khả năng bám dính giữa bề mặt sợi và nhựa. Lý do kỹ thuật quan trọng của việc chọn dải nhiệt độ này nằm ở điểm nóng chảy của sợi PP (khoảng 165°C). Nhiệt độ này vừa đủ để sợi đạt trạng thái mềm hóa, thúc đẩy quá trình tương tác hóa-lý tại giao diện sợi-nhựa mà không làm biến tính cấu trúc polymer của sợi. Vùng tương tác nhiệt độ tối ưu này từ mô hình CatBoost hoàn toàn tương đồng với các quan sát vi mô trong các nghiên cứu thực nghiệm trước đây [8, 19]. Đồng thời, tại hàm lượng 3-5%, các sợi PP phân tán đồng nhất và đan xen tạo thành mạng lưới không gian ba chiều ổn định trong nền nhựa

đường, giúp hấp thụ và phân tán ứng suất hiệu quả [20]. Tuy nhiên, sự tương tác với các biến khác cũng cảnh báo rằng nếu nhiệt độ vượt quá ngưỡng giới hạn, các đặc tính của nhựa nền có thể bị suy giảm do hiện tượng hóa già nhiệt, từ đó làm giảm tổng hiệu suất kháng hàn lún của hệ thống.

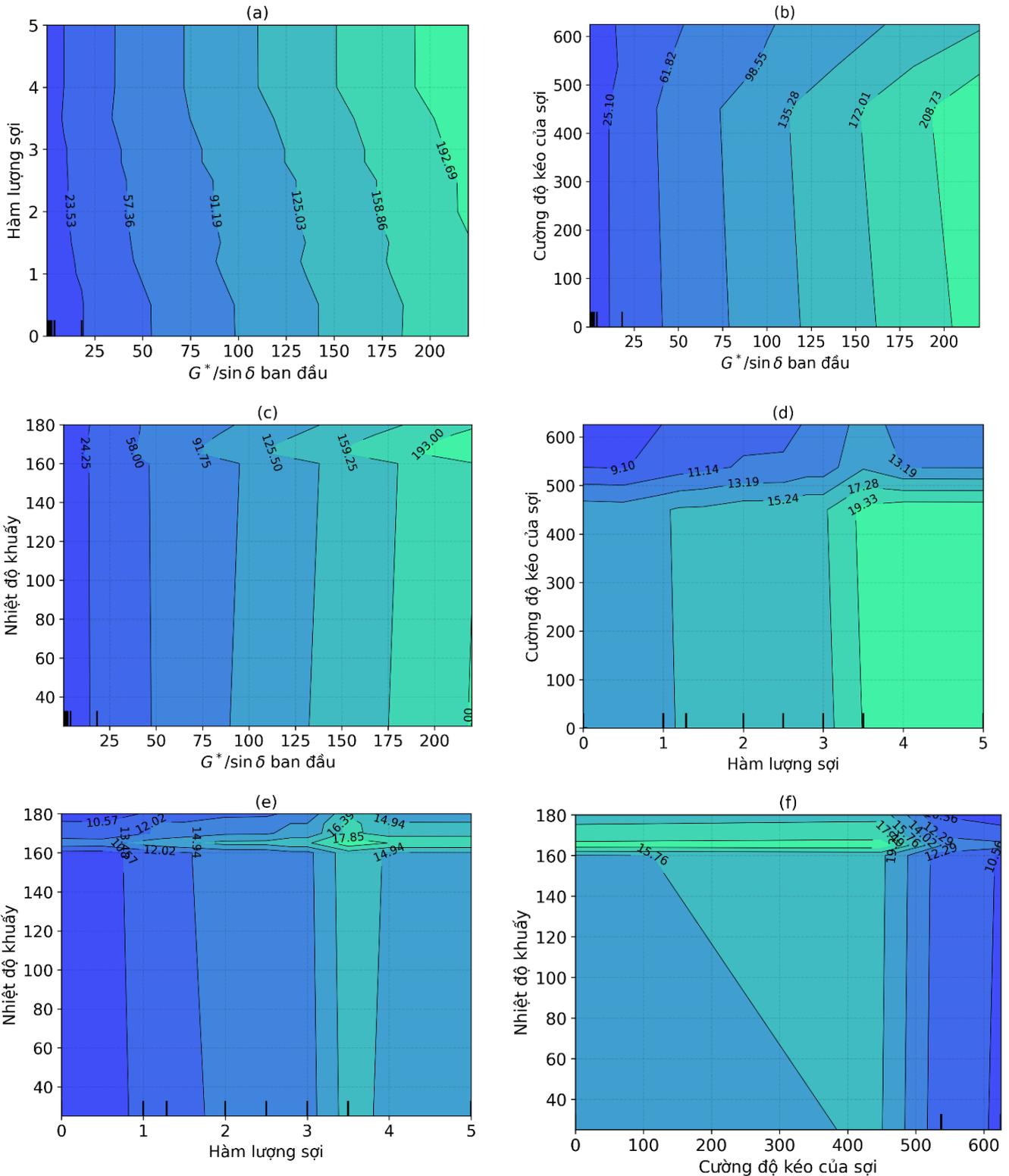
Sự kết hợp giữa cường độ kéo của sợi và hàm lượng sợi (Hình 8d) cũng tạo nên những ảnh hưởng rõ rệt đến đặc tính kháng hàn lún. Khi nồng độ sợi tăng lên, mật độ các điểm giao cắt trong mạng lưới không gian ba chiều dày đặc hơn. Đồng thời, nếu các sợi này có cường độ kéo cao, chúng sẽ ngăn chặn hiệu quả sự hình thành và lan truyền của các vết nứt vi mô bên trong chất kết dính. Điều này giúp duy trì tính nguyên vẹn của cấu trúc nhựa đường dưới tác động lặp lại của tải trọng bánh xe, một cơ chế đã được chỉ ra trong nghiên cứu của [28] về tương tác giữa vật liệu nano hoặc sợi và nhựa đường.

Sự tương tác giữa hàm lượng sợi và nhiệt độ khuấy (Hình 8e) cho thấy hiệu suất của nhựa đường biến tính không tăng tuyến tính theo nhiệt độ mà đạt ngưỡng tối ưu trong khoảng 160-175°C khi hàm lượng sợi nằm trong khoảng 2.5 - 4.0%. Nếu nhiệt độ quá thấp, độ nhớt của nhựa đường tăng cao khiến các sợi PP có xu hướng vón cục, làm giảm diện tích tiếp xúc bề mặt và dẫn đến sự phân tán không đồng nhất. Ngược lại, việc duy trì nhiệt độ trong dải tối ưu giúp nhựa đường bao phủ hoàn toàn bề mặt sợi, tạo ra sự liên kết hóa - lý bền vững, giúp tăng cường độ ổn định nhiệt cho hệ thống chất kết dính.

Cuối cùng, Hình 8f hiển thị tương tác giữa cường độ kéo của sợi và nhiệt độ khuấy nhằm làm rõ thêm ảnh hưởng của năng lượng chế tạo đến đặc tính cơ lý của sợi gia cường. Tại vùng nhiệt độ thấp dưới 160°C, sự thay đổi cường độ kéo của sợi ít gây ra biến động lớn đến hệ số đàn hồi tổng thể do lực bám dính tại giao diện sợi-nhựa chưa đủ mạnh để truyền tải trọng hiệu quả. Tuy nhiên, khi nhiệt độ khuấy tăng lên mức thích hợp, bề mặt tương tác cho thấy xu hướng cải thiện hiệu suất rõ rệt hơn đối với các loại sợi có cường độ kéo cao. Điều này minh chứng rằng quy trình chế tạo nhiệt

độ cao không chỉ hỗ trợ phân tán mà còn thúc đẩy quá trình tương thích bề mặt, cho phép đặc tính cường độ của sợi PP tham gia trực tiếp vào việc nâng cao khả năng kháng hần lún cho nhựa

đường. Các kết quả phân tích PDP 2D này một lần nữa khẳng định tính đúng đắn của việc tối ưu hóa đồng thời cả thành phần vật liệu và thông số công nghệ để đạt được hiệu quả biến tính cao nhất.



## 5. Kết luận

Nghiên cứu này đã xây dựng thành công mô hình dự báo hệ số hần lún của nhựa đường biến tính sợi PP bằng thuật toán CatBoost. Qua quá trình tối ưu hóa và đánh giá định lượng, mô hình tiêu biểu đã thể hiện sự hội tụ cao với  $R^2 = 0.9967$  và  $RMSE = 0.5720$ , minh chứng cho tính nhất quán giữa dự báo học máy và kết quả thực nghiệm. Tầm quan trọng của các biến thông qua biểu đồ phụ thuộc một phần đã làm sáng tỏ cơ chế gia cường vật lý, trong đó sợi PP đóng vai trò là khung xương thứ cấp giúp hấp thụ ứng suất và hạn chế sự dịch chuyển dẻo của nhựa đường dưới tải trọng bánh xe. Kết quả nghiên cứu khẳng định rằng việc thiết kế tối ưu giữa các đặc tính của nhựa gốc, cường độ sợi và điều kiện chế tạo nhiệt độ cao là yếu tố tiên quyết để nâng cao khả năng kháng hần lún của chất kết dính. Nghiên cứu ứng dụng thành công kỹ thuật giải thích thông qua PDP đã giúp mở hộp đen của mô hình học máy, đảm bảo các logic dự báo hoàn toàn phù hợp với nguyên lý vật lý và các nghiên cứu quốc tế đã công bố. Đây là cơ sở khoa học quan trọng để triển khai các ứng dụng học máy trong việc dự báo nhanh và thiết kế cấp phối các loại vật liệu đường bộ thông minh, góp phần nâng cao tuổi thọ công trình và tối ưu hóa nguồn lực trong ngành giao thông vận tải. Hơn nữa, tính linh hoạt của cấu trúc mô hình CatBoost cho phép mở rộng khả năng dự báo sang các hệ vật liệu sử dụng các loại sợi gia cường khác, khẳng định giá trị của phương pháp nghiên cứu trong việc phát triển các dòng vật liệu nhựa đường biến tính thế hệ mới.

### Tài liệu tham khảo

- [1] C. Fan, K. Hu, T. Zhang, J. Zhou, W. Zhang, G. Su. (2025). Evaluation of silicone resin enhanced recycled concrete aggregates for asphalt mixture application. *Construction and Building Materials*, 463, 140030. <https://doi.org/10.1016/j.conbuildmat.2025.140030>
- [2] B. Lei, W. Li, Z. Luo, V.W. Tam, W. Dong, K. Wang. (2020). Performance enhancement of permeable asphalt mixtures with recycled aggregate for concrete pavement application. *Frontiers in Materials*, 7, 253. <https://doi.org/10.3389/fmats.2020.00253>
- [3] Z. Deng, W. Li, W. Dong, Z. Sun, J. Kodikara, D. Sheng. (2023). Multifunctional asphalt concrete pavement toward smart transport infrastructure: Design, performance and perspective. *Composites Part B: Engineering*, 265, 110937. <https://doi.org/10.1016/j.compositesb.2023.110937>
- [4] G. Qian, D. Yao, X. Gong, H. Yu, N. Li. (2019). Performance evaluation and field application of hard asphalt concrete under heavy traffic conditions. *Construction and Building Materials*, 228, 116729. <https://doi.org/10.1016/j.conbuildmat.2019.116729>
- [5] S. Tapkın. (2008). The effect of polypropylene fibers on asphalt performance. *Building and Environment*, 43(6), 1065–1071. <https://doi.org/10.1016/j.buildenv.2007.02.011>
- [6] M. Latifi, Ö.B. Altun, A. Mardani. (2022). Effect of the addition of polypropylene fiber on concrete properties. *Journal of Adhesion Science and Technology*, 36(4), 345–369. [https://doi.org/10.1080/01694243.2021.1922221?urlappend=%3Futm\\_source%3Dresearchgate.net%26utm\\_medium%3Darticle](https://doi.org/10.1080/01694243.2021.1922221?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle)
- [7] L.M.G. Klinsky, K.E. Kaloush, V.C. Faria, V.S.S. Bardini. (2018). Performance characteristics of fiber modified hot mix asphalt. *Construction and Building Materials*, 176, 747–752. <https://doi.org/10.1016/j.conbuildmat.2018.04.221>
- [8] P. Yeh, Y. Nien, J. Chen, W. Chen, J. Chen. (2005). Thermal and rheological properties of maleated polypropylene modified asphalt. *Polymer Engineering & Science*, 45(8), 1152–1158. <https://doi.org/10.1002/pen.20386>
- [9] L. Guo, W. Xu, Y. Zhang, W. Ji, S. Wu. (2022). Selecting the best performing modified asphalt based on rheological properties and microscopic analysis of RPP/SBS modified

- asphalt. *Materials*, 15(23), 8616. [https://doi.org/10.3390/ma15238616?urlappened=%3Futm\\_source%3Dresearchgate.net%26utm\\_medium%3Darticle](https://doi.org/10.3390/ma15238616?urlappened=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle)
- [10] Q. Zhang, D. Hou, Z. Li, H. Wang, S. Dong. (2024). Evaluation of the thermal stability and micro-modification mechanism of SBR/PP-modified asphalt. *Polymers*, 16(4), 456. <https://doi.org/10.3390/polym16040456>
- [11] C. Li, H. Liu, Y. Xiao, J. Li, T. Wang, L. Peng. (2023). Modification and enhancing contribution of fiber to asphalt binders and their corresponding mixtures: A study of viscoelastic properties. *Materials*, 16(16), 5727. [https://doi.org/10.3390/ma16165727?urlappened=%3Futm\\_source%3Dresearchgate.net%26utm\\_medium%3Darticle](https://doi.org/10.3390/ma16165727?urlappened=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle)
- [12] H.-G.T. Hoang, T.-A. Nguyen, H.-L. Nguyen, H.-B. Ly. (2025). An Effective Extreme Gradient Boosting Approach to Predict the Physical Properties of Graphene Oxide Modified Asphalt. *Intenational Journal of Pavement Research and Technology*. <https://doi.org/10.1007/s42947-025-00636-y>
- [13] H.-B. Ly, L.M. Le, H.T. Duong, T.C. Nguyen, T.A. Pham, T.-T. Le, V.M. Le, L. N. Ngoc, B.T. Pham. (2019). Hybrid artificial intelligence approaches for predicting critical buckling load of structural members under compression considering the influence of initial geometric imperfections. *Applied Sciences*, 9(11), 2258. [https://doi.org/10.3390/app9112258?urlappend=%3Futm\\_source%3Dresearchgate.net%26utm\\_medium%3Darticle](https://doi.org/10.3390/app9112258?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle)
- [14] H.-B. Ly, B.T. Pham, D.V. Dao, V.M. Le, L.M. Le, T.-T. Le. (2019). Improvement of ANFIS model for prediction of compressive strength of manufactured sand concrete. *Applied Sciences*, 9, 3841. <http://dx.doi.org/10.3390/app9183841>
- [15] J.A. Crowder, J.N. Carbone. (2011). Occam learning through pattern discovery: Computational mechanics in AI systems, in: Proceedings on the International Conference on Artificial Intelligence (ICAI). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 1-5.
- [16] S.-A. Boukabara, V. Krasnopolsky, S.G. Penny, J.Q. Stewart, A. McGovern, D. Hall, J.E. T. Hoeve, J. Hickey, H.-L. Allen Huang, J.K. Williams. (2021). Outlook for exploiting artificial intelligence in the earth and environmental sciences, *Bulletin of the American Meteorological Society*, 102(5), E1016–E1032. <https://doi.org/10.1175/BAMS-D-20-0031.1>
- [17] K. Zhong, Q. Meng, M. Sun, G. Luo. (2022). Artificial neural network (ANN) modeling for predicting performance of SBS modified asphalt. *Materials*, 15(23), 8695. [https://doi.org/10.3390/ma15238695?urlappened=%3Futm\\_source%3Dresearchgate.net%26utm\\_medium%3Darticle](https://doi.org/10.3390/ma15238695?urlappened=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle)
- [18] A.S. Hosseini, P. Hajikarimi, M. Gandomi, F.M. Nejad, A.H. Gandomi. (2021). Optimized machine learning approaches for the prediction of viscoelastic behavior of modified asphalt binders. *Construction and Building Materials*, 299, 124264. <https://doi.org/10.1016/j.conbuildmat.2021.124264>
- [19] P.M. Kathari. (2016). Rheological Properties of Polypropylene Reinforced Asphalt Binder, *Transportation Infrastructure Geotechnology*, 3, 109-126. <https://doi.org/10.1007/s40515-016-0033-3>
- [20] H. Xiaoming, I.B. Eldouma. (2019). Experimental study to determine the most preferred additive for improving asphalt performance using polypropylene, crumb rubber, and tafpack super in medium and high-temperature range. *Applied Sciences*, 9(8), 1567.
- [21] J.T. Hancock, T.M. Khoshgoftaar. (2020). CatBoost for big data: an interdisciplinary review. *Journal of Big Data*, 7, 94. <https://doi.org/10.1186/s40537-020-00369-8>
- [22] R. Kohavi. (1995). A study of cross-

- validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14<sup>th</sup> international conference on Artificial Intelligence*, 2, 1137-1143. doi/10.5555/1643031.1643047
- [23] J. Kim. (1997). Iterated grid search algorithm on unimodal criteria. Virginia Polytechnic Institute and State University. <https://search.proquest.com/openview/56ed5c094a247c5b93535a1d35798db2/1?pq-origsite=gscholar&cbl=18750&diss=y> (accessed January 16, 2026).
- [24] J.H. Friedman. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- [25] H.-G.T. Hoang, H.-V.T. Mai, H.L. Nguyen, H.-B. Ly. (2024). Application of extreme gradient boosting in predicting the viscoelastic characteristics of graphene oxide modified asphalt at medium and high temperatures. *Frontiers of Structural and Civil Engineering*, 18, 899–917. <https://doi.org/10.1007/s11709-024-1025-y>
- [26] H.-G.T. Hoang, T.-A. Nguyen, H.-L. Nguyen, H.-B. Ly. (2022). Neural network approach for GO-modified asphalt properties estimation, *Case Studies in Construction Materials*, 17, e01617. <https://doi.org/10.1016/j.cscm.2022.e01617>
- [27] S.M. Abtahi, M. Sheikhzadeh, S.M. Hejazi. (2010). Fiber-reinforced asphalt-concrete – A review, *Construction and Building Materials*, 24(6), 871–877. <https://doi.org/10.1016/j.conbuildmat.2009.11.009>
- [28] N. Bala, M. Napiah, I. Kamaruddin. (2018). Effect of nanosilica particles on polypropylene polymer modified asphalt mixture performance. *Case Studies in Construction Materials*, 8, 447–454