



FORECASTING EARLY-AGE THERMAL CRACKING RISK IN STRUCTURAL CONCRETE WITH NOISY DATA

Nguyen Duy Tien¹, Do Anh Tu^{1*}, Hoang Viet Hai¹, Luu Hung Cuong²

¹University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam

²TEDI, No 278 Ton Duc Thang Street, Hanoi, Vietnam

ARTICLE INFO

TYPE: Research Article

Received: 06/05/2025

Revised: 29/05/2025

Accepted: 10/06/2025

Published online: 15/06/2025

<https://doi.org/10.47869/tcsj.76.5.7>

* Corresponding author

Email: doanhtu@utc.edu.vn

Abstract. Numerical methods have demonstrated significant advantages in simulating and predicting early-age thermal cracking in concrete (commonly in mass concrete), as traditional measurements are often time-consuming and costly. This study integrates numerical methods (to determine input parameters) with machine learning techniques to predict the early-age cracking potential of structural concrete blocks. A dataset of 759 samples on the early-age thermal cracking of bridge piers was collected by varying the height and width of the piers. Subsequently, this data with 10% noise was used to train the LR, SVR, RF, KNN, Adaboost, XGBoost, and CatBoost models. After training, the resulting models were used to predict the outcomes of the 759 samples with 0% noise. With a prediction accuracy of $R^2 = 0,9618$ for the cracking index, the CatBoost model performs best, even maintaining high accuracy when the training data contains up to 10% noise. This study contributes to demonstrating the applicability of machine learning methods in reducing computational time and enhancing prediction accuracy of early-age thermal cracking in concrete, even in cases where input data quality is low (noisy).

Keywords: Machine Learning, Early-Age Thermal Cracking, Numerical Simulation, Cracking Risk Prediction.

@2025 University of Transport and Communications



DỰ BÁO KHẢ NĂNG NÚT DO NHIỆT CỦA BÊ TÔNG KẾT CẤU Ở TUỔI SỚM KHI BỊ NHIỀU DỮ LIỆU

Nguyễn Duy Tiến¹, Đỗ Anh Tú^{1*}, Hoàng Việt Hải¹, Lưu Hùng Cường²

¹Trường Đại học Giao thông vận tải, Số 3 Cầu Giấy, Hà Nội, Việt Nam

²Tổng Công ty tư vấn thiết kế Giao Thông Vận tải, Số 278 Tôn Đức Thắng, Hà Nội

THÔNG TIN BÀI BÁO

CHUYÊN MỤC: Công trình khoa học

Ngày nhận bài: 06/05/2025

Ngày nhận bài sửa: 29/05/2025

Ngày chấp nhận đăng: 10/06/2025

Ngày xuất bản Online: 15/06/2025

<https://doi.org/10.47869/tcsj.76.5.7>

* Tác giả liên hệ

Email: doanhtu@utc.edu.vn

Tóm tắt. Phương pháp số đã chứng minh được tính ưu việt trong việc mô phỏng nhằm dự báo khả năng nứt nhiệt của bê tông ở tuổi sớm (thường là bê tông khối lớn) do công tác đo đạc thường tốn kém và mất thời gian. Bài báo này kết hợp phương pháp số (để xác định dữ liệu đầu vào) và sử dụng phương pháp học máy để dự báo khả năng nứt của khối bê tông kết cấu ở tuổi sớm. Dữ liệu 759 mẫu về khả năng nứt của trụ cầu được thu thập thông qua việc thay đổi chiều cao và chiều rộng của trụ cầu. Sau đó, dữ liệu đó được tạo nhiễu 10% và được sử dụng để huấn luyện các mô hình LR, SVR, RF, KNN, Adaboost, XGboost và Catboost. Sau khi được huấn luyện, mô hình thu được sử dụng để dự báo kết quả của 759 mẫu khi bị nhiễu 0%. Với độ chính xác $R^2=0,9618$ và $MAE = 0,032$ khi dự báo hệ số nứt, mô hình Catboost có độ chính xác lớn nhất, đồng thời thể hiện khả năng tối ưu trong việc dự báo kể cả khi kết quả đầu vào sử dụng để huấn luyện mạng có nhiễu lên đến 10%. Nghiên cứu góp phần chứng minh khả năng ứng dụng phương pháp học máy trong việc giảm thời gian tính toán và nâng cao độ chính xác trong dự báo khả năng nứt nhiệt của bê tông ở tuổi sớm trong trường hợp chất lượng dữ liệu đầu vào không cao (bị nhiễu).

Từ khóa: Học máy, Nứt nhiệt ở tuổi sớm, Mô phỏng số, Dự báo khả năng nứt.

@2025 Trường Đại học Giao thông vận tải

1. ĐẶT VẤN ĐỀ

Trong lĩnh vực xây dựng hạ tầng, đặc biệt với các công trình quy mô lớn như đập thủy điện, móng, trụ cầu hay các bộ kết cấu bê tông cốt thép khối lớn, hiện tượng sinh nhiệt do quá trình thủy hóa xi măng trong giai đoạn đầu đóng rắn là một vấn đề kỹ thuật then chốt. Sự gia

tăng nhiệt độ tại lõi khối bê tông, kết hợp với quá trình mất nhiệt nhanh tại bề mặt tiếp xúc với môi trường ngoài, tạo ra sự chênh lệch nhiệt độ đáng kể. Hệ quả là xuất hiện ứng suất nhiệt nội sinh, và nếu không được kiểm soát hiệu quả, có thể dẫn đến nứt sớm trong bê tông – gây suy giảm độ bền lâu, tính toàn vẹn kết cấu và tuổi thọ của công trình [1-3].

Dự báo nguy cơ nứt do nhiệt trong giai đoạn đầu đóng rắn giúp các kỹ sư đưa ra các giải pháp hợp lý như kiểm soát tốc độ thủy hóa, bổ sung vật liệu khoáng, tối ưu phương pháp đổ và dưỡng hộ bê tông. Tuy nhiên, dữ liệu thực tế thu thập từ công trường thường chịu ảnh hưởng mạnh bởi các yếu tố nhiễu như sai số đo đạc, điều kiện môi trường biến động và tính không đồng nhất của vật liệu xây dựng. Điều này làm cho các mô hình truyền thống dựa trên phân tích cơ học thuần túy hoặc hồi quy tuyến tính đơn giản khó đáp ứng yêu cầu về độ tin cậy và tính ổn định dự báo.

Trong những năm gần đây, các thuật toán Machine Learning (ML) đã mở ra những hướng tiếp cận mới trong việc xử lý các bài toán dự báo phức tạp và phi tuyến trong kỹ thuật xây dựng. Nhiều nghiên cứu đã sử dụng ML để dự đoán cường độ của vật liệu bê tông [4, 5], nhiệt thủy hóa của xi măng [6], độ mở rộng vết nứt trong bê tông [7], hoặc hư hỏng trong kết cấu [8, 9]. Các mô hình như Linear Regression, Random Forest, AdaBoost, XGBoost và CatBoost đã được chứng minh hiệu quả trong việc dự đoán các chỉ số kỹ thuật của vật liệu và kết cấu [10-12]. Linear Regression, với ưu điểm đơn giản và dễ giải thích, phù hợp với các quan hệ tuyến tính; trong khi đó, Random Forest và AdaBoost có khả năng khai thác tốt các mối quan hệ phi tuyến, đồng thời hạn chế hiện tượng quá khớp (overfitting). Đặc biệt, XGBoost và CatBoost – những thuật toán boosting tiên tiến – đã cho thấy ưu thế vượt trội về độ chính xác và tốc độ huấn luyện trong nhiều bài toán kỹ thuật [12].

Trong nghiên cứu này, chúng tôi triển khai và so sánh hiệu quả của một số mô hình học máy tiêu biểu – bao gồm Linear Regression, Random Forest, AdaBoost, XGBoost và CatBoost – trong dự báo giá trị cực đại của hệ số nứt (η_{max}) của bê tông trong giai đoạn đầu đóng rắn. Bộ dữ liệu sử dụng gồm 759 mẫu dữ liệu mô phỏng số, thu được từ mô phỏng số, phản ánh sự phát triển nhiệt độ và ứng suất theo thời gian dưới nhiều điều kiện biên hình học và môi trường khác nhau.

Việc sử dụng mô hình học máy nhằm mục tiêu không chỉ tăng cường độ chính xác trong dự báo mà còn đảm bảo tính ổn định của mô hình trước các biến động và nhiễu loạn dữ liệu, từ đó góp phần hỗ trợ kỹ sư trong việc ra quyết định sớm tại công trường – nơi dữ liệu đo đạc thực tế luôn tiềm ẩn nhiều bất định khó lường.

2. MÔ HÌNH SỐ KẾT HỢP PHƯƠNG PHÁP HỌC MÁY

2.1. Ứng dụng mô hình số

Các phương pháp số bao gồm phương pháp phần tử hữu hạn (Finite Element Method - FEM), sai phân hữu hạn (Finite Difference Method - FDM) hay thủy động lực học hạt mịn (Smoothed Particle Hydrodynamics - SPH) là các công cụ mạnh mẽ trong phân tích nhiệt–cơ học (thermo-mechanical analysis), cho phép mô phỏng quá trình phát triển nhiệt độ, ứng suất và biến dạng trong quá trình thành hình kết cấu từ vật liệu [13, 14]. Trong nghiên cứu này, phương pháp số được sử dụng để xây dựng bộ dữ liệu mô phỏng, từ đó xác định các chỉ số nguy cơ nứt do nhiệt tại các thời điểm cụ thể trong quá trình đóng rắn của bê tông.

- Phân tích nhiệt (Thermal analysis):

Mục tiêu của phân tích này là xác định phân bố nhiệt độ theo thời gian và không gian trong

khối bê tông. Phương trình dẫn nhiệt được biết đến rộng rãi thể hiện trong công thức sau:

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q_h \quad (1)$$

Trong đó: T là nhiệt độ ($^{\circ}\text{C}$); ρ là khối lượng riêng (kg/m^3); c là nhiệt dung riêng ($\text{J}/\text{kg}\cdot\text{K}$); k là hệ số dẫn nhiệt ($\text{W}/\text{m}\cdot\text{K}$); Q_h là nguồn nhiệt do phản ứng thủy hóa (W/m^3).

Điều kiện biên nhiệt gồm: Tại mặt ngoài: trao đổi nhiệt đối lưu với hệ số h ($\text{W}/\text{m}^2\cdot\text{K}$) và nhiệt độ môi trường T_{am} ; Tại đáy/mặt tiếp xúc cốt-pha: có thể coi là cách nhiệt hoặc dẫn nhiệt thấp tùy vật liệu làm ván khuôn.

• Phân tích ứng suất:

Sau khi xác định trường nhiệt độ, ứng suất nhiệt được tính từ trường biến dạng nhiệt được xác định theo công thức (2):

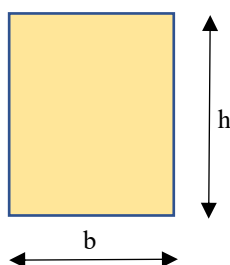
$$\sigma = E \cdot (\varepsilon_{\text{total}} - \varepsilon_{\text{th}}) \quad \varepsilon_{\text{th}} = \alpha \cdot (T - T_{\text{ref}}) \quad (2)$$

Trong đó: σ là ứng suất nhiệt (MPa), E là mô đun đàn hồi của bê tông (MPa); α : hệ số giãn nở nhiệt ($1/^{\circ}\text{C}$); T_{ref} : nhiệt độ tham chiếu (thường lấy bằng 23°C); ε_{th} : biến dạng nhiệt.

Phương pháp số gồm FEM và FDM nói trên đã được tích hợp vào chương trình tính "EACTSA" [14] phục vụ tính toán nhiệt độ và ứng suất nhiệt của bê tông ở tuổi sớm sẽ được sử dụng ở mục tiếp theo.

2.2. Thu thập dữ liệu

Để thiết lập bộ dữ liệu đầu vào phục vụ cho quá trình huấn luyện mô hình học máy, nghiên cứu này sử dụng chương trình tính toán "EACTSA" – một công cụ phân tích ứng suất nhiệt trong bê tông ở tuổi sớm [14]. Chương trình tính EACTSA cho phép mô phỏng quá trình phát triển nhiệt độ và ứng suất nhiệt trong kết cấu trụ cầu bê tông với nhiều tổ hợp tham số đầu vào khác nhau. Độ tin cậy của chương trình đã được kiểm chứng thông qua các nghiên cứu trước đây [14]. Quy trình phân tích gồm hai bước chính: đầu tiên là mô phỏng quá trình truyền nhiệt trong khối bê tông để xác định phân bố nhiệt độ theo thời gian; tiếp theo, dữ liệu nhiệt độ theo lịch sử thời gian tại từng nút trong mô hình được sử dụng làm tải trọng nhiệt cho giai đoạn phân tích ứng suất nhiệt. Một đặc điểm nổi bật của chương trình EACTSA là khả năng xét đến sự biến đổi theo thời gian của các thông số cơ học của bê tông, cũng như tính đến hiệu ứng đàn nhót trong giai đoạn đóng rắn và phát triển cường độ sớm của vật liệu.



Hình 1. Minh họa mặt cắt ngang trụ cầu kích thước $b \times h$.

Tham số đầu vào gồm: mặt cắt ngang trụ cầu bê tông với bề rộng và chiều cao thay đổi từ

1 – 5 m (minh họa trên Hình 1). Nhiệt độ môi trường xung quanh được giả định nằm trong dải nhiệt độ môi trường đặc trưng ở Việt Nam (từ 14°C đến 36°C). Vật liệu bê tông trụ là bê tông thông thường cấp 35 MPa được sử dụng ở Việt Nam, với các đặc trưng nhiệt và cơ học (cường độ) được thí nghiệm tại Trường Đại học Giao thông vận tải được trình bày trong Bảng 1 và 2. Các tham số về nhiệt được lấy theo mô hình 3 tham số đã được trình bày kỹ trong các nghiên cứu trước đây [14, 15].

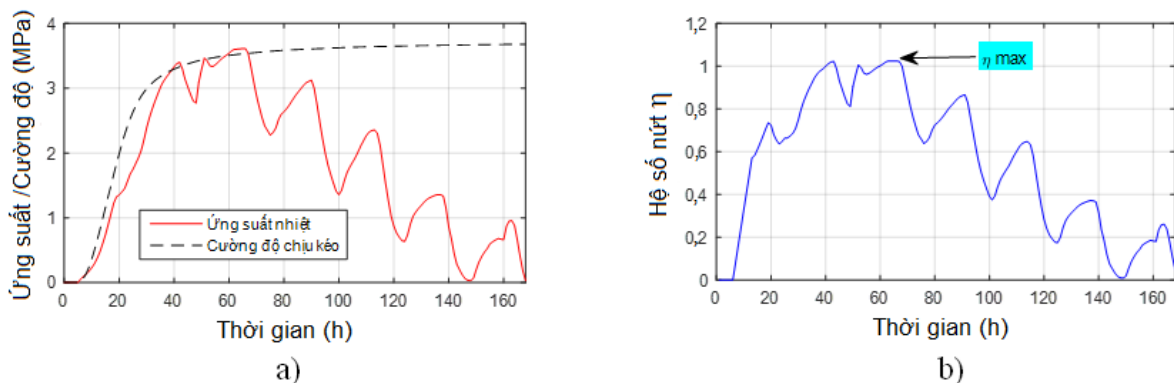
Bảng 1. Các chỉ số thống kê của tập dữ liệu

	Kí hiệu	Số lượng mẫu	Trung bình	Độ lệch chuẩn	min	max
b (m)	X1	759	2,46	0,766	1	3,2
h (m)	X2	759	2,8458	1,2076	1	5
T_{am} (°C)	X3	759	24,49	5,13	14	36
T_{ini} (°C)	X4	759	24,51	3,225	20	30
η_{max}	Y1	759	0,77	0,21	0,26	1,11
t (°C)	Y2	759	56,65	19,32	19,81	103,83

Bài báo này nhằm nghiên cứu mô hình dự đoán để dự báo chỉ số nứt tới hạn η_{max} và bằng các thuật toán học máy. Thông thường quá trình mô phỏng bằng phương pháp phần tử hữu hạn và chạy mô hình mất nhiều thời gian, do đó việc dự đoán việc áp dụng các thuật toán học máy, lựa chọn các thuật toán có khả năng dự đoán chính xác lực η_{max} là đối tượng rất đáng được quan tâm. Cơ sở dữ liệu gồm có 759 mẫu, với đầu vào (input) là 4 thông số khác nhau gồm kích thước mặt cắt ngang thân trụ: bề rộng ($b = X1$), chiều cao ($h = X2$), nhiệt độ trung bình của môi trường ($T_{am} = X3$) và nhiệt độ bê tông tươi ban đầu ($T_{ini} = X4$). Phân tích thống kê cơ sở dữ liệu được thể hiện trong Bảng 1 và Hình 2. Đầu ra (output) bao gồm giá trị lớn nhất của hệ số nứt η và thời gian (t) tương ứng. Hệ số nứt η được định nghĩa như sau [13]:

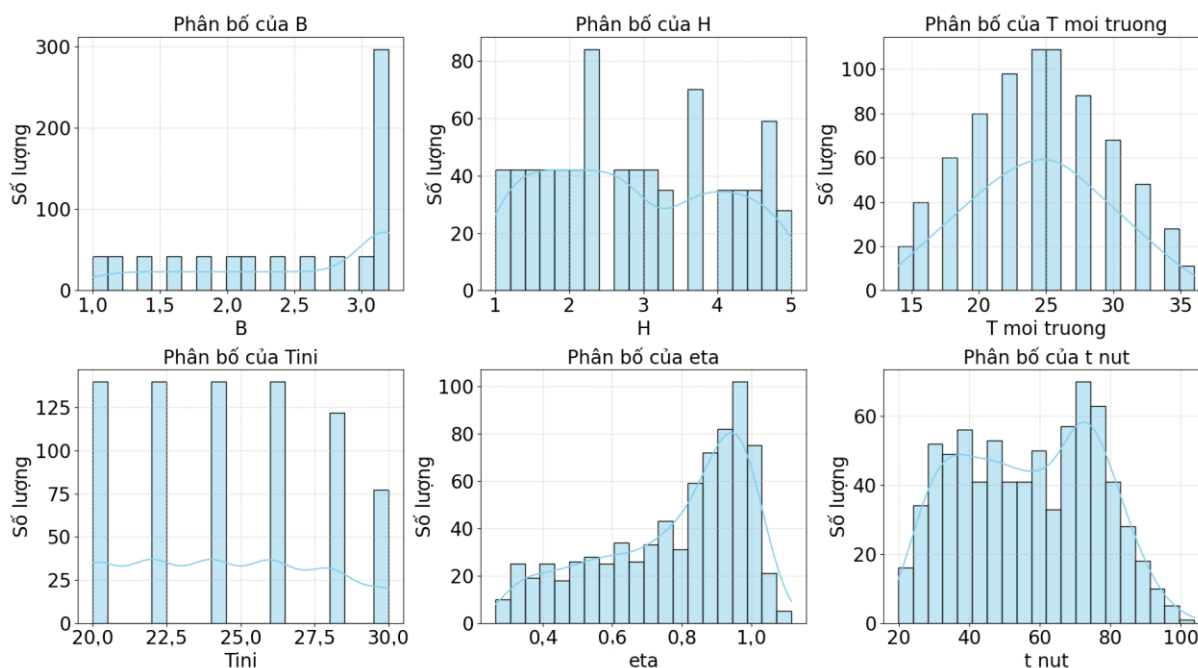
$$\eta = \left[\frac{\sigma(t)}{f_{ct}(t)} \right]_{\max} \quad (3)$$

trong đó: t là thời gian (h); $\sigma(t)$ là ứng suất nhiệt chính kéo (lớn nhất) theo thời gian, minh họa trên Hình 2; $f_{ct}(t)$ là cường độ chịu kéo của bê tông tại điểm đang xét theo thời gian.



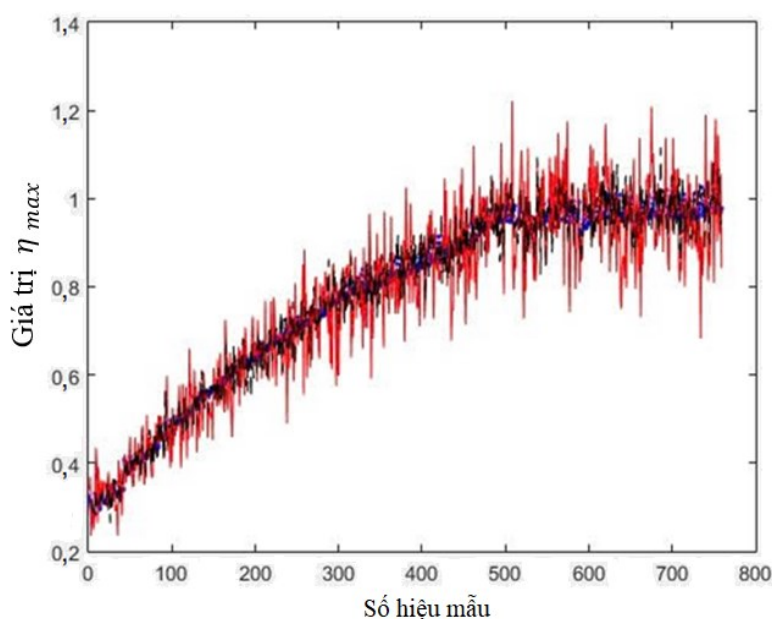
Hình 2. Minh họa kết quả phân tích từ chương trình “EACTSA”: ứng suất nhiệt và cường độ chịu kéo của bê tông (a); hệ số nứt (b).

Hệ số nứt η càng lớn thì khả năng nứt của bê tông càng cao. Theo một số nghiên cứu như của Riding và cộng sự [16], khi $\eta < 0,7$ thì khả năng nứt thấp, $0,7 \leq \eta < 1$ thì khả năng nứt cao, $\eta \geq 1$ thì khả năng nứt rất cao.



Hình 3. Biểu đồ phân phối của các thuộc tính dựa trên số liệu mô phỏng

Trong nghiên cứu này ngoài 759 mẫu được xác định từ chương trình tính toán “EACTSA”, bài báo giả định trường hợp đo đạc, tính toán giá trị bị lệch 10%. Nghiên cứu tìm ra mô hình học máy (ML) có hiệu suất cao nhất trong việc dự báo kết quả. Do đó, 759 mẫu khác được khởi tạo có giá trị η lệch ngẫu nhiên tối đa đến 10%, như mô tả trong Hình 4.



Hình 4: Biểu diễn giá trị η_{max} và giá trị η_{max} có độ lệch tối đa 10% (nét liền màu đỏ).

2.3. Các mô hình dự báo

Trong những năm gần đây, bên cạnh cách mô hình truyền thống: Mô hình rừng cây ngẫu nhiên RF (Random forest), mô hình tăng cường thích ứng (Adaboost), mô hình máy véc-tơ hỗ trợ (SVM), mô hình tăng cường trong máy học (XGBoost), mô hình người láng giềng gần nhất (KNN), mô hình Catboost,... Mỗi một thuật toán có ưu và nhược điểm riêng căn cứ vào bộ dữ liệu sử dụng để huấn luyện đặc tính, đặc điểm các tham số trong thuật toán [13]. Trong nghiên cứu này, việc dự báo là trị η_{max} đạt được giữa các thuật toán được so sánh với nhau lựa chọn và đề xuất ra các thuật toán tối ưu. Phần tiếp theo sẽ giới thiệu sơ lược về cơ sở lý thuyết của các thuật toán được sử dụng trong nghiên cứu này.

2.3.1. Thuật toán Random Forest

Random Forest là một thuật toán học máy thuộc nhóm ensemble learning, hoạt động bằng cách xây dựng nhiều cây quyết định trên các tập dữ liệu con ngẫu nhiên và tổng hợp kết quả dự đoán từ các cây. Phương pháp này giúp giảm thiểu overfitting, nâng cao độ chính xác và ổn định trong các bài toán phân loại và hồi quy. Nhờ khả năng xử lý dữ liệu phức tạp và đo lường tầm quan trọng của đặc trưng, Random Forest được ứng dụng rộng rãi trong nhiều lĩnh vực kỹ thuật và khoa học dữ liệu.

2.3.2. Thuật toán véc-tơ hỗ trợ (SVM)

Thuật toán Véc-tơ Hỗ trợ (SVM) là một phương pháp học máy mạnh mẽ dùng để phân loại và hồi quy. Mục tiêu của SVM là tìm siêu phẳng phân chia dữ liệu sao cho khoảng cách giữa các lớp là lớn nhất. SVM sử dụng các vector hỗ trợ (support vectors) để xác định vị trí của siêu phẳng này. Đối với bài toán phi tuyến, SVM sử dụng hàm nhân (kernel) để chuyển đổi không gian đặc trưng. Thuật toán này hiệu quả trong không gian đặc trưng lớn và giúp tránh quá khớp thông qua điều chỉnh tham số C và kernel. SVM được ứng dụng rộng rãi trong nhận dạng mẫu, phân loại văn bản, và nhiều lĩnh vực khác.

2.3.3. Thuật toán Adaboost

AdaBoost (Adaptive Boosting) là một thuật toán boosting mạnh mẽ, kết hợp nhiều mô hình học máy yếu (weak learners), thường là cây quyết định nhỏ, để tạo thành một mô hình mạnh mẽ. Mỗi mô hình trong AdaBoost được huấn luyện để cải thiện các dự đoán sai của các mô hình trước đó. Thuật toán này điều chỉnh trọng số của các điểm dữ liệu, tăng trọng số cho các mẫu bị phân loại sai và giảm trọng số cho các mẫu đúng. AdaBoost có khả năng giảm overfitting và làm việc hiệu quả với các mô hình đơn giản. Thuật toán này đặc biệt hữu ích trong các bài toán phân loại và có thể sử dụng với nhiều loại mô hình học máy yếu.

2.3.2. Thuật toán KNN

KNN (K-Nearest Neighbors) là thuật toán học máy đơn giản và dễ hiểu dùng cho phân loại và hồi quy. Khi dự đoán cho một điểm dữ liệu mới, KNN tìm K điểm dữ liệu gần nhất trong tập huấn luyện và gán nhãn hoặc giá trị dự đoán dựa trên đa số nhãn (cho phân loại) hoặc giá trị trung bình (cho hồi quy) của các điểm gần nhất. Khoảng cách giữa các điểm thường được tính bằng Euclidean hoặc các đo lường khác. KNN không yêu cầu giai đoạn huấn luyện, nhưng có thể trở nên chậm khi xử lý dữ liệu lớn. Thuật toán dễ triển khai và hiệu quả trong các bài toán với dữ liệu không có giả thuyết phân phối cụ thể [17].

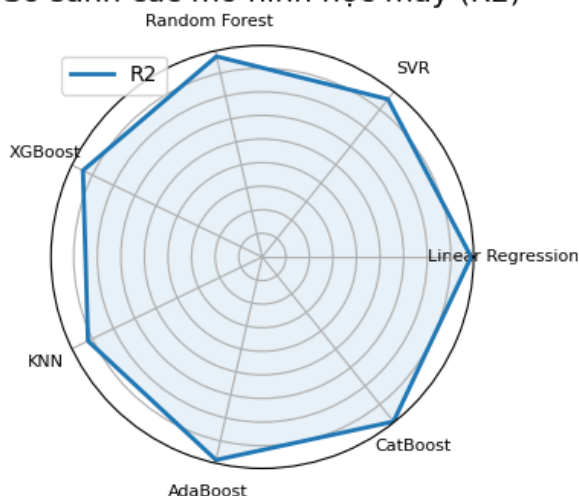
2.3.2. Thuật toán Catboost

CatBoost là một thuật toán học máy dựa trên cây quyết định, được phát triển bởi nhóm Yandex [18]. Thuật toán này nổi bật với khả năng xử lý hiệu quả dữ liệu rời rạc, dữ liệu thưa

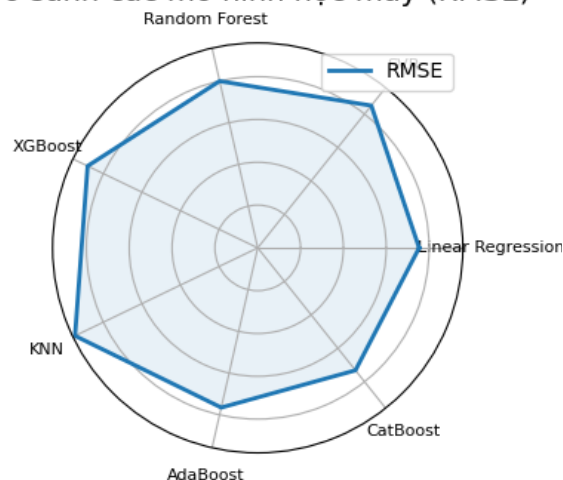
và dữ liệu có giá trị thiếu, giúp cải thiện độ chính xác trong các bài toán phân loại và dự báo. CatBoost kết hợp kỹ thuật tăng tốc gradient cùng với cơ chế đánh giá mức độ quan trọng của các đặc trưng, từ đó xây dựng mô hình dự đoán mạnh mẽ và ổn định.

2.3.2. Thuật toán XGBoost

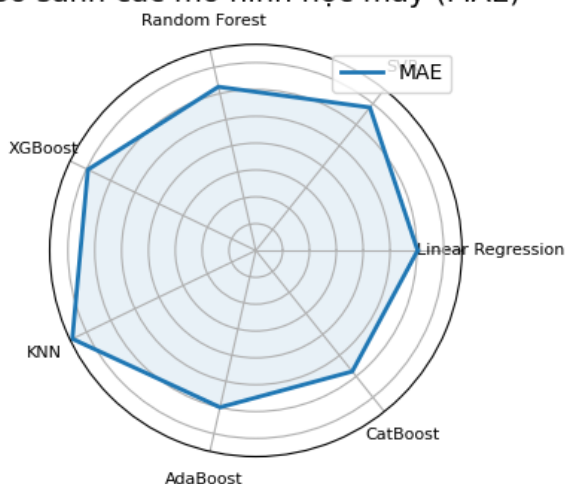
So sánh các mô hình học máy (R2)



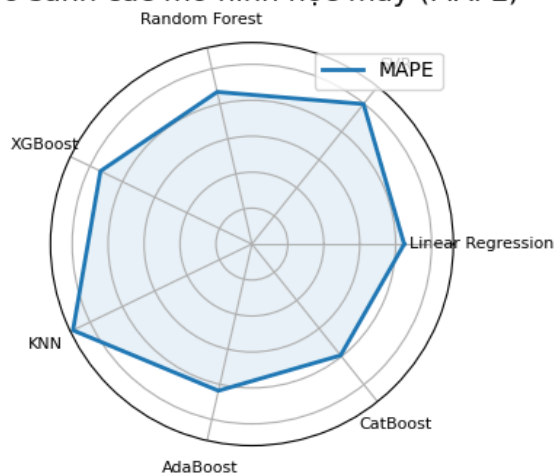
So sánh các mô hình học máy (RMSE)



So sánh các mô hình học máy (MAE)



So sánh các mô hình học máy (MAPE)



Hình 5. Biểu đồ so sánh R2, RMSE, MAE và MAPE của các mô hình trong việc dự báo giá trị η_{max} .

XGBoost (Extreme Gradient Boosting) là thuật toán boosting mạnh mẽ cho phân loại và hồi quy, kết hợp nhiều cây quyết định yếu để tạo thành mô hình mạnh [19]. Sử dụng phương pháp gradient boosting, XGBoost giảm thiểu sai số qua từng bước học. Thuật toán này áp dụng regularization và pruning để giảm overfitting, đồng thời hỗ trợ tính toán song song và xử lý dữ liệu thiếu. XGBoost được biết đến nhờ hiệu suất cao và khả năng tối ưu hóa tham số, thường được sử dụng trong các cuộc thi học máy và ứng dụng thực tế.

Trong mỗi lần tính toán, các hàm số được tạo ra một cách ngẫu nhiên, sau đó, hàm số có độ chính xác cao nhất được lưu và đánh giá với tập dữ liệu kiểm chứng. Để đánh giá độ chính xác của mô hình, bốn hệ số đánh giá độ chính xác là R2, RMSE, MAE và MAPE được sử dụng như sau:

$$R_2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

3. KẾT QUẢ DỰ BÁO VÀ THẢO LUẬN

Các Mô hình được sử dụng để dự báo giá trị η_{max} khi dữ liệu bị lệch 10%. Kết quả so sánh các mô hình được thể hiện ở hình 5 và bảng 2.

Độ chính xác của mô hình được đánh giá thông qua bốn hệ số R2, RMSE, MAE và MAPE. Có thể thấy rằng thuật toán Catboost có độ chính xác cao nhất với R2=0,892, RMSE = 0,0731; MAE = 0,0576; MAPE = 7,918 khi dự báo giá trị η_{max} (xem Bảng 2, Hình 5). Giá trị R2 lớn nhất sát giá trị 1 và giá trị RMSE, MAE, MAPE càng nhỏ chứng tỏ khả năng dự báo của mô hình càng chuẩn xác.

Bảng 2. So sánh thông số của các thuật toán dự báo.

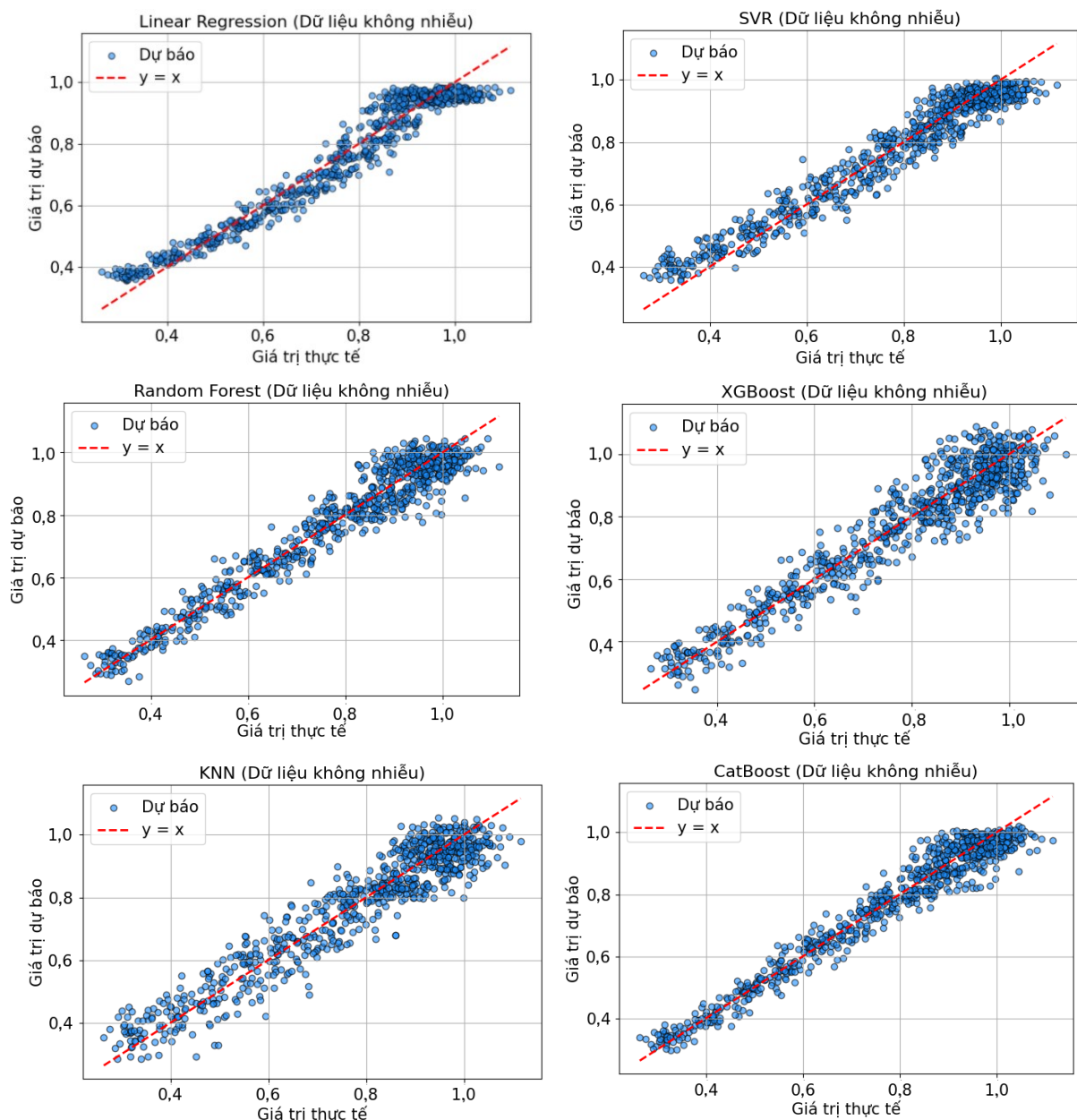
Mô hình	Tham số dự báo	R2	RMSE	MAE	MAPE
LR	η_{max}	0,886	0,075	0,06	8,476
SVR	η_{max}	0,855	0,085	0,0681	9,964
RF	η_{max}	0,872	0,0799	0,0625	8,681
XGBoost	η_{max}	0,844	0,088	0,0694	9,353
KNN	η_{max}	0,8207	0,094	0,0759	11,052
AdaBoost	η_{max}	0,882	0,076	0,0599	8,366
Catboost	η_{max}	0,892	0,0731	0,0576	7,918

Sau khi tiến hành sử dụng các mô hình để huấn luyện, mạng được huấn luyện được sử dụng để dự báo hệ số nứt và đối chiếu hệ số nứt dự báo với hệ số nứt thu được từ chương trình tính "EACTSA" (ứng với trường hợp không có nhiễu). Kết quả so sánh được thể hiện ở Hình 6.

Bảng 3. So sánh các mô hình khi đối chiếu giá trị dự báo và giá trị tính toán bằng "EACTSA"

Mô hình	R2	MAE
LR	0,9521	0,0374
SVR	0,949	0,0386
RF	0,943	0,0392
XGBoost	0,923	0,0458
KNN	0,921	0,0483
AdaBoost	0,9568	0,0353
Catboost	0,9618	0,032

Hình 6 và Bảng 3 thể hiện so sánh giữa kết quả dự báo và kết quả tính toán từ chương trình "EACTSA". Đường chéo là đường $y = x$ đánh giá độ chụm của dữ liệu. Kết quả cho thấy hai thuật toán Catboost và AdaBoost là hai thuật toán có khả năng dự báo tốt nhất với độ chính xác R2 lần lượt là 0,9618 và 0,9568 sát với giá trị 1 và giá trị MAE của hai mô hình này lần lượt là 0,0353 và 0,032 nhỏ nhất.



Hình 6. So sánh giá trị dự báo bằng học máy và giá trị thực tế (tính toán bằng "EACTSA").

4. KẾT LUẬN

Nghiên cứu đã đề xuất và kiểm chứng một quy trình tích hợp giữa phương pháp số và các thuật toán học máy nhằm tối ưu hóa việc dự báo nguy cơ nứt do nhiệt của bê tông ở tuổi sớm trong điều kiện dữ liệu có nhiều. Thông qua việc đánh giá hiệu quả của các mô hình LR, SVR, RF, KNN, AdaBoost, XGBoost và CatBoost trên tập dữ liệu có mức nhiễu 10%, kết quả cho thấy CatBoost đạt độ chính xác cao nhất với hệ số tương quan $R^2 = 0,9618$ và giá trị MAE = 0,032. Điều này cho thấy CatBoost có khả năng dự báo ổn định và tin cậy ngay cả khi dữ liệu đầu vào không hoàn toàn chính xác.

Kết quả nghiên cứu cho thấy tiềm năng ứng dụng các thuật toán học máy trong việc rút ngắn thời gian mô phỏng, nâng cao hiệu quả dự báo nứt do nhiệt và hỗ trợ thiết kế kết cấu bê

tổng an toàn, hợp lý hơn trong thực tiễn.

Trong các nghiên cứu tiếp theo, cần mở rộng đánh giá mô hình với các loại cấp phối bê tông khác nhau nhằm tăng tính tổng quát, làm phong phú hơn khả năng dự đoán của phương pháp đề xuất và đồng thời xác định mô hình dự báo phù hợp hơn cho từng điều kiện vật liệu và môi trường cụ thể.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Trường Đại học Giao thông vận tải (ĐH GTVT) trong đề tài mã số T2024-CT-003TĐ.

TÀI LIỆU THAM KHẢO

- [1]. ACI 207.1R-05, Guide to Mass Concrete, Farmington Hills, MI, USA: ACI 207.1R-05, American Concrete Institute, 2005.
- [2]. M. Tia, A. Lawrence, C. Ferraro, T. A. Do, and Y. Chen, Pilot project for maximum heat of mass concrete, Florida. Dept. of Transportation, 2013
- [3]. T.A. Do, T.H. Nguyen, T.X. Vu, T.T. Hoang, T.D. Tran, T.T. Bui, Adiabatic temperature rise and thermal analysis of high-performance concrete bridge elements, In ICSCEA 2019: Proceedings of the International Conference on Sustainable Civil Engineering and Architecture, 2020, Springer, Singapore, 313-423.
- [4]. M.Y. Mansour, M.U.R.A.T. Dicleli, J.Y. Lee, J.J.E.S. Zhang, Predicting the shear strength of reinforced concrete beams using artificial neural networks, Engineering Structures, 26(6) (2004) 781-799.
- [5]. S. Lai, M. Serra, Concrete strength prediction by means of neural network, Construction and Building Materials, 11(2) (1997) 93-98.
- [6]. R. Cook, T. Han, A. Childers, C. Ryckman, K. Khayat, H. Ma, J. Huang, A. Kumar, Machine learning for high-fidelity prediction of cement hydration kinetics in blended systems, Materials & Design, 208 (2021). <https://doi.org/10.1016/j.matdes.2021.109920>
- [7]. A.A. Elshafeyet, N. Dawood, H. Marzouk, M. Haddara, Crack width in concrete using artificial neural networks, Engineering Structures, 52 (2013) 676-686. <https://doi.org/10.1016/j.engstruct.2013.03.020>
- [8]. D. Maity, A. Saha, Damage assessment in structure from changes in static parameter using neural networks, Sadhana, 29(3) (2004) 315-327.
- [9]. H.N. Tran, S. Khatir, G. De Roeck, T.B. Tien, M.A. Wahab, An efficient artificial neural network for damage detection in bridges and beam-like structures by improving training parameters using cuckoo search algorithm, Engineering Structures, 199 (2019).
- [10]. P.N. Huynh, Development of an artificial neural network based-prediction model for bond strength of FRB bars in concrete, Transport Communications Science Journal, 75 (2024) 1502-1517. <https://doi.org/10.47869/tcsj.75.4.3>
- [11]. M. Mohtasham Moein, A. Saradar, K. Rahmati, S.H. Ghasemzadeh Mousavinejad, J. Bristow, V. Aramali, M. Karakouzian, Predictive models for concrete properties using machine learning and deep learning approaches: A review, Journal of Building Engineering, 63 (2023). <https://doi.org/10.1016/j.jobbe.2022.105444>
- [12]. H.T. Thai, Machine learning for structural engineering: A state-of-the-art review, Structures, 38 (2022) 448-491. <https://doi.org/10.1016/j.istruc.2022.02.003>
- [13]. H.T. Nguyen, B. Cosson, M.F. Lacrampe, T.A. Do, Numerical simulation of reactive polymer flow during rotational molding using smoothed particle hydrodynamics method and experimental verification, International Journal of Material Forming, 11 (2018), 583-592.

- [14].T.A. Do, H.L. Chen, G. Leon, T.H. Nguyen, A combined finite difference and finite element model for temperature and stress predictions of cast-in-place cap beam on precast columns, *Construction and Building Materials*, 217 (2019) 172-184.
- [15].Y. Lin, H.L. Chen, Thermal analysis and adiabatic calorimetry for early-age concrete members, *Journal of Thermal Analysis and Calorimetry*, 122(2) (2015) 937-945.
- [16].K.A. Riding, J.L. Poole, A.K. Schindler, M.C.G. Juenger, K.J. Folliard, Statistical determination of cracking probability for mass concrete, *Journal of materials in civil engineering*, 26(9) (2014).
- [17].V.H. Hoang, N.L. Nguyen, T.T. Bui, N.H. Tran, A Two-stage Method for Damage Detection in Z24 Bridge Based on K-nearest Neighbor and Artificial Neural Network, *Periodica Polytechnica Civil Engineering*, 3 (2024). <https://doi.org/10.3311/PPci.23884>
- [18].J.T. Hancock, T.M. Khoshgoftaar, CatBoost for big data: an interdisciplinary review, *J Big Data* 7, 94 (2020). <https://doi.org/10.1186/s40537-020-00369-8>
- [19].T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, ACM, (2016) 785-794. <https://doi.org/10.1145/2939672.2939785>