

Estimation of Foamed Concrete Compressive Strength and Relationship Identification with Input Factors Using Support Vector Machine

Hai-Bang Ly*

University of Transport Technology, 54 Trieu Khuc , Thanh Xuan, Hanoi

Article info

Type of article:

Original research paper

*Corresponding author:

E-mail address:

banglh@utt.edu.vn

Received:

03/12/2021

Accepted:

20/01/2022

Published:

09/02/2022

Abstract: Foamed concrete (FC) is a building material with many advantages and is widely used in construction to reduce the load on the structure. Compressive strength is an important mechanical characteristic of concrete in general and of FC in particular. Therefore, it is important to quickly and accurately estimate this quantity. This study proposes the application of support vector machine (SVM) model to predict the FC compressive strength and construct a relationship with the input variables for the purpose of optimizing FC design process. A database of 220 test results is collected and used to build and verify the predictive performance of the proposed SVM model. The input factors of the problem are FC density, water/cement ratio, and sand/cement ratio. The results show that SVM is a good predictor of compressive strength of FC with performance evaluation criteria such as root mean square error, RMSE = 3.475 MPa, mean absolute error MAE = 2.816 MPa and coefficient of determination $R^2 = 0.930$. Finally, 2-Dimensional Partial Dependence Plot (PDP) is developed to correlate the three input variables with the compressive strength of FC, which is useful for material engineers in the design phase of FC.

Keywords: Foamed Concrete; Artificial Intelligence; Compressive Strength; Support Vector Machine (SVM); Partial Dependence Plot (PDP).

Nghiên cứu dự báo cường độ bê tông xốp và xác định tương quan với các tham số đầu vào bằng máy vec tơ hỗ trợ

Lý Hải Bằng*

Trường Đại học Công nghệ Giao thông Vận tải, Số 54 Triều Khúc, Hà Nội
100000

Thông tin bài viết

Dạng bài viết:

Bài báo nghiên cứu

*Tác giả liên hệ:

Địa chỉ E-mail:

banglh@utt.edu.vn

Ngày nộp bài:

03/12/2021

Ngày chấp nhận:

20/01/2022

Ngày đăng bài:

09/02/2022

Tóm tắt: Bê tông xốp (BTX) là một loại vật liệu xây dựng với nhiều ưu điểm và được sử dụng rộng rãi trong xây dựng công trình để giảm tải trọng cho kết cấu. Cường độ nén là một đặc trưng cơ học quan trọng của bê tông nói chung và của BTX nói riêng, nên việc ước tính nhanh, chính xác đại lượng này rất quan trọng. Nghiên cứu này đề xuất ứng dụng mô hình máy vec tơ hỗ trợ (SVM) để dự báo cường độ BTX và xây dựng mối tương quan giữa các yếu tố đầu vào của bài toán với mục đích tối ưu hóa quá trình thiết kế BTX. Bộ cơ sở dữ liệu gồm 220 kết quả thí nghiệm cường độ BTX được thu thập và dùng để xây dựng, kiểm chứng năng lực dự báo của mô hình SVM. Các yếu tố đầu vào của bài toán được xét tới là tỷ trọng BTX, tỷ lệ nước/xi măng, và tỷ lệ cát/xi măng. Kết quả cho thấy SVM là một công cụ dự báo rất tốt cường độ nén của BTX với các tiêu chí đánh giá hiệu suất như căn của sai số toàn phương trung bình $RMSE = 3.475$ MPa, sai số tuyệt đối trung bình $MAE = 2.816$ MPa và hệ số xác định $R^2 = 0.930$. Cuối cùng, biểu đồ phụ thuộc được phát triển để xây dựng tương quan giữa ba biến đầu vào nói trên với cường độ nén của BTX, giúp ích cho các kỹ sư vật liệu trong công tác thiết kế cấp phối.

Từ khóa: Bê tông xốp; trí tuệ nhân tạo; cường độ nén; máy vec tơ hỗ trợ; biểu đồ phụ thuộc từng phần.

1. Đặt vấn đề

Bê tông là một loại vật liệu được sử dụng rộng rãi trong lĩnh vực xây dựng công trình nhờ vào rất nhiều ưu điểm nổi trội của nó. Tuy nhiên, sử dụng bê tông cũng tồn tại nhiều hạn chế. Thứ nhất, việc sản xuất bê tông gây nên nhiều ảnh hưởng nghiêm trọng tới môi trường do cốt liệu được lấy từ nguồn tài nguyên thiên nhiên, hoặc do quá trình sản xuất xi măng và các loại phụ gia. Thứ hai, thời gian xây dựng công trình bị ảnh hưởng đáng kể do phải chờ bê tông các cấu kiện đạt được đủ cường độ. Cuối cùng, giá thành sản xuất

bê tông cũng có thể gây ảnh hưởng không nhỏ các chủ đầu tư, nhà thầu xây dựng hoặc các doanh nghiệp vật liệu xây dựng. Trong những năm gần đây, việc nghiên cứu và sản xuất các loại vật liệu mới thay thế cho bê tông truyền thống và giảm thiểu các nhược điểm của nó đang thu hút được sự quan tâm của các nhà nghiên cứu. Trong số đó, bê tông siêu nhẹ, hay bê tông xốp (BTX) được coi là một phương án hữu hiệu.

Bê tông siêu nhẹ, hay BTX, được tạo ra chủ yếu từ xi măng, cát và chất phụ gia tạo bọt. Loại vật liệu này đã được chứng minh là có khả năng

cách âm, cách nhiệt rất tốt [1,2]. Hơn nữa, do đặc tính nhẹ và bền nên loại bê tông này thường được sử dụng trong xây dựng công trình ở các vị trí như tường ngăn, tường bao cho các chung cư cao tầng, thay thế cho các vùng không ổn định, và đặc biệt là giảm tải trọng cho các lớp lấp phía trên các kết cấu ngầm, hoặc lớp đệm cho các bể ngầm, đường ống [3,4]. Việc sử dụng BTX sẽ giúp giảm thời gian xây dựng công trình, giảm giá thành cho các phần như kết cấu móng hoặc khung từ 10-20%. Nếu xét tới quá trình chế tạo thì BTX được làm khác với bê tông sử dụng cốt liệu truyền thống. Tuy rằng BTX cũng có những thành phần cơ bản như cốt liệu mịn hoặc một phần nhỏ cốt liệu thô, nước, xi măng, nhưng thành phần chủ yếu của BTX là vô số các bọt khí tạo ra bằng các tác động cơ học từ một số loại dung dịch đặc biệt nhất định. Điều này giúp người sử dụng có thể điều chỉnh được tỷ trọng của bê tông (dao động trong khoảng 300 đến 2000 kg/m³). Sự xuất hiện và phân bố khá đồng đều của các bọt khí nhỏ này giúp tạo ra một loại bê tông tươi, dễ dàng sử dụng và có thể dùng bơm bê tông để bơm. Hơn nữa, bọt khí được sử dụng để thay thế cho cốt liệu nhẹ trong bê tông xốp chính là lý do giải thích tính kinh tế cao của loại bê tông này.

Với những ưu điểm nổi bật đã nêu trên thì việc tối ưu hóa thiết kế cấp phối BTX là rất quan trọng, trong đó, cường độ nén là một đặc tính cơ học cần được đặc biệt quan tâm. Rất nhiều nghiên cứu trên thế giới đã tiến hành xây dựng tương quan giữa thành phần BTX và cường độ nén. Cụ thể, Asadzadeh và Khoshbayan [5] sử dụng phương pháp thiết kế và phân tích thử nghiệm (DOE) để tối ưu hóa rất nhiều đặc tính cơ học của BTX. Phương pháp đáp ứng bề mặt (RSM) cũng được áp dụng để tìm mối tương quan giữa các thành phần chế tạo và cường độ của BTX sử dụng tro bay trong nghiên cứu [6]. Tuy nhiên, kết quả của các phương pháp trên đều chỉ đạt độ chính xác tương đối, với hệ số xác định $R^2 = 0.60$. Với phương pháp tiếp cận thực nghiệm, Bing và cộng sự [7] hoặc Liu và cộng sự [8] đã tiến hành thử nghiệm với 16 cấp phối khác nhau để nghiên cứu cường độ nén của BTX, và ảnh hưởng của chất tạo bọt tới cường độ này [7]. Tuy rằng đã có rất nhiều những kết quả khả quan ban đầu, thời gian

và kinh phí thí nghiệm là hai rào cản lớn nhất cho những cách tiếp cận trên. Bên cạnh đó, các phương trình thực nghiệm cũng được đưa ra để dự báo cường độ nén của BTX. Tuy nhiên, do số lượng mẫu thí nghiệm hạn chế, khoảng giá trị của các yếu tố cấu thành BTX và cường độ nằm trong phạm vi nhất định nên độ chính xác của các phương trình thực nghiệm cũng bị giới hạn theo [9]. Đặc biệt, việc đề xuất các phương trình thực nghiệm đa số gặp nhiều khó khăn vì mối quan hệ giữa các thành phần với cường độ của BTX là phi tuyến và rất phức tạp [9-13]. Chính vì vậy, việc phát triển một phương pháp mới có tính tổng quan và chính xác hơn hiện đang là một nhu cầu tất yếu.

Trong khoảng 4 thập kỷ trở lại đây, phương pháp sử dụng trí tuệ nhân tạo (AI) hoặc các mô hình máy học (ML) đang thu hút được sự quan tâm của cộng đồng nghiên cứu thế giới. Phương pháp tiếp cận này đã được chứng minh là đặc biệt hiệu quả trong nhiều lĩnh vực của cuộc sống, đặc biệt là giải quyết các bài toán khó trong xây dựng công trình. Những bài toán phức tạp trong kết cấu công trình [14-16], khoa học vật liệu [17,18], địa kỹ thuật [19,20], khoa học trái đất [21,22] đều được giải quyết một cách thuyết phục khi sử dụng phương pháp tiếp cận AI-ML. Điều này chứng minh rằng rất nhiều thuật toán ML có thể được sử dụng để xây dựng mối tương quan phức tạp cho các bài toán kỹ thuật.

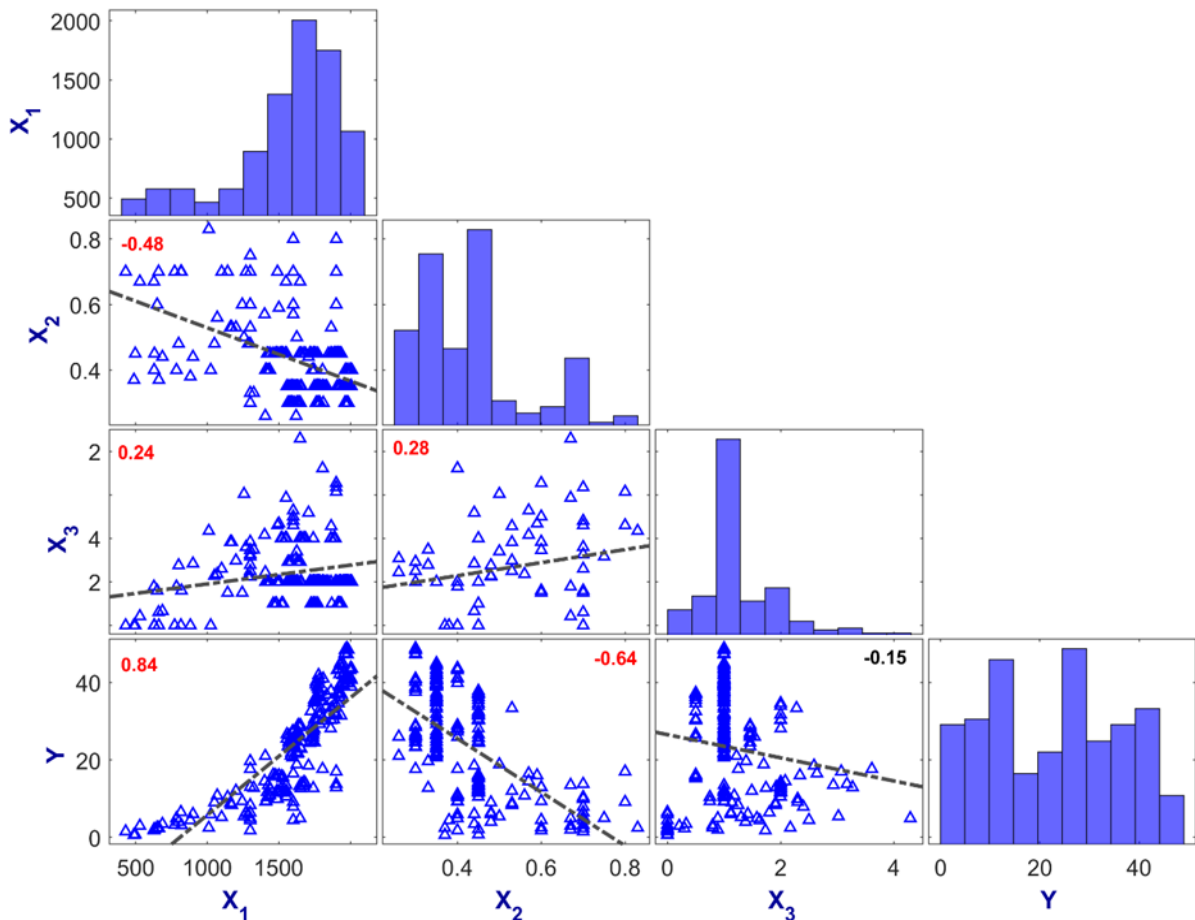
Trong nghiên cứu này, một thuật toán máy học có tên máy vec tơ hỗ trợ (support vector machine – SVM) được sử dụng để nghiên cứu và dự báo cường độ nén của BTX. Với mục đích trên, một bộ dữ liệu gồm 220 thí nghiệm nén BTX đã được tập hợp và sử dụng [23]. Các tham số đầu vào của bài toán là tỷ trọng BTX, tỷ lệ nước/xi măng, và tỉ lệ cát/xi măng, còn hàm mục tiêu là cường độ nén của bê tông. Quá trình phát triển công cụ SVM được chia thành 2 giai đoạn: giai đoạn đầu là huấn luyện sử dụng 70% dữ liệu, và giai đoạn kiểm chứng, sử dụng 30% dữ liệu còn lại. Các tiêu chí đánh giá hiệu suất dự báo của SVM sử dụng trong nghiên cứu này bao gồm hệ số xác định (R^2), căn của sai số toàn phương trung bình (RMSE) và sai số tuyệt đối trung bình (MAE). Cuối cùng, nghiên cứu phát triển một biểu đồ phụ

thuộc để đánh giá tương quan giữa các tham số đầu vào và cường độ nén của BTX.

2. Cơ sở dữ liệu

Như đã trình bày ở phần trước, nghiên cứu này sử dụng 220 dữ liệu thí nghiệm cường độ nén của BTX để phát triển công cụ SVM. Bộ dữ liệu này được thu thập từ 10 công trình công bố quốc tế uy tín khác nhau, và được tổng hợp ở tài liệu [23]. Hàm mục tiêu của nghiên cứu là cường độ nén BTX ở 28 ngày tuổi (ký hiệu là Y , MPa), còn các biến đầu vào được xét tới là tỷ trọng BTX (viết

gắn gọn cho “tỷ trọng khô”, ký hiệu là X_1 , kg/m³), tỷ lệ nước/xi măng (ký hiệu là X_2), và tỷ lệ cát/xi măng (ký hiệu là X_3). Tỷ trọng của BTX có giá trị trong khoảng khá rộng, từ 430 đến 2009 (kg/m³), tỷ lệ nước/xi măng có giá trị từ 0.26 đến 0.83, còn tỷ lệ cát/xi măng có giá trị từ 0 đến 4.29. Điều này có nghĩa là có một vài mẫu BTX không sử dụng cát trong bộ dữ liệu này. Ngoài ra, cường độ nén BTX bao phủ một khoảng rộng từ 0.60 đến 48.88 (MPa).



Hình 1. Biểu đồ phân bố và mối tương quan giữa các tham số đầu vào và đầu ra được xét tới trong nghiên cứu này

Hình 1 biểu diễn đồ thị phân bố của X_1 , X_2 , X_3 , Y và mối tương quan giữa các tham số đó. Hệ số tương quan Pearson được tính toán và ghi chú trong mỗi cặp tham số. Có thể thấy rằng Y không có tương quan lớn nào trực tiếp giữa Y và các tham số khác, ngoại trừ X_1 . Mặt khác, tuy rằng này được thu thập từ 10 công trình công bố quốc tế uy

tín khác nhau, và được tổng hợp ở tài liệu [23]. Hàm mục tiêu của nghiên cứu là cường độ nén BTX ở 28 ngày tuổi (ký hiệu là Y , MPa), còn các biến đầu vào được xét tới là tỷ trọng BTX (viết gắn gọn cho “tỷ trọng khô”, ký hiệu là X_1 , kg/m³), tỷ lệ nước/xi măng (ký hiệu là X_2), và tỷ lệ cát/xi măng (ký hiệu là X_3). Tỷ trọng của BTX có giá trị trong

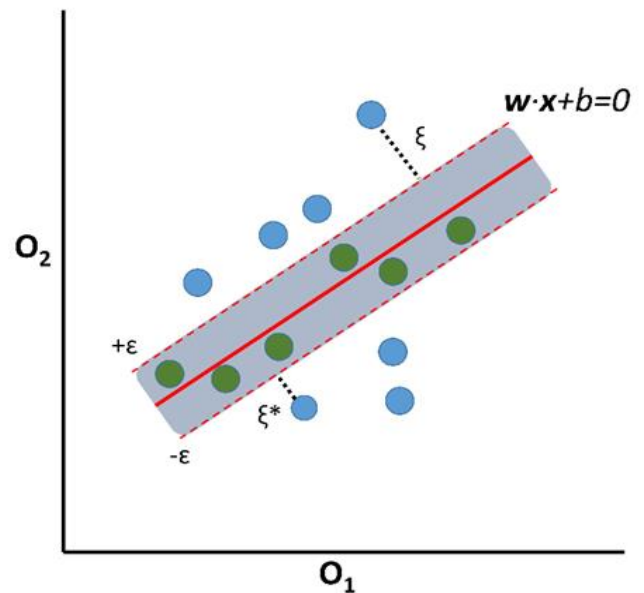
khoảng khá rộng, từ 430 đến 2009 (kg/m³), tỷ lệ nước/xi măng có giá trị từ 0.26 đến 0.83, còn tỷ lệ cát/xi măng có giá trị từ 0 đến 4.29. Điều này có nghĩa là có một vài mẫu BTX không sử dụng tương quan giữa Y và X₁ là 0.84 nhưng có thể thấy trên biểu đồ rằng đây là hàm phụ thuộc dạng hàm số mũ, tuy tương quan lớn nhưng các sai số RMSE hoặc MAE cũng sẽ rất lớn. Có thể kết luận rằng các biến được lấy trong nghiên cứu này đều là các biến độc lập và sẽ được sử dụng để xây dựng mối tương quan với Y. Ngoài ra, tuy rằng bọt tạo khí (foaming agent) là một nhân tố không thể thiếu trong thành phần cấu thành của BTX nhưng biến này đã được ẩn bên trong (là biến phụ thuộc) của biến X₁ – tỷ trọng BTX. Sau quá trình thu thập, phân tích và đánh giá thì chỉ 3 biến độc lập X₁, X₂, X₃ được xét tới trong nghiên cứu này. Cuối cùng, để giảm thiểu sai số phát sinh trong quá trình mô phỏng bằng SVM, bộ dữ liệu này được chuẩn hóa về khoảng giá trị 0-1. Đây là một phương pháp được sử dụng rất phổ biến trong các bài toán trí tuệ nhân tạo để hạn chế các lỗi do mô phỏng số tạo ra, ví dụ như độ lệch khá lớn giữa tỷ trọng (đơn vị hàng nghìn) và tỷ lệ nước/xi măng (giá trị nhỏ hơn 1). Quá trình này, kèm theo các hình vẽ trong nghiên cứu được thực hiện bằng việc lập trình các đoạn mã sử dụng phần mềm Matlab.

3. Phương pháp nghiên cứu

3.1. Máy vec tơ hỗ trợ – SVM

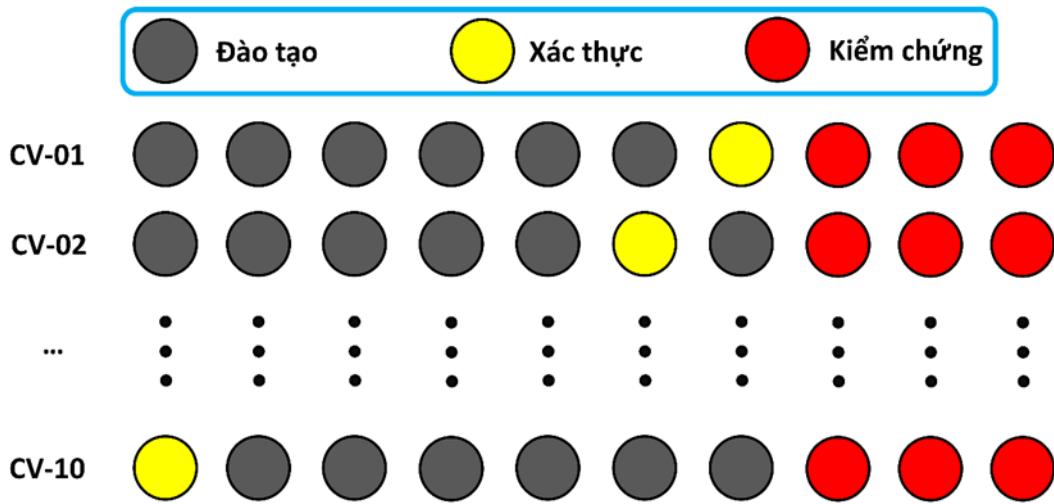
Máy vec tơ hỗ trợ (Support Vector Machine – SVM) là một phương pháp máy học được sử dụng rất phổ biến trong các bài toán kỹ thuật [24]. SVM thường được sử dụng để giải quyết các bài toán phân loại với hai hay nhiều lớp. Ngoài ra, việc sử dụng SVM để giải quyết các bài toán hồi quy (giống như nghiên cứu này) hoặc các bài toán phân cụm cũng được nghiên cứu nhiều trên thế giới. Nguyên lý làm việc của SVM dựa trên việc tìm ra được một (đối với bài toán phân loại 2 lớp)

hoặc nhiều “siêu phẳng” (hyper-plane) đối với bài toán phân loại đa lớp. Với các dữ liệu trong không gian hai chiều hay đa chiều, mô hình SVM hiển thị các dữ liệu đó và lựa chọn ranh giới – chính là các siêu phẳng – sao cho khoảng cách từ các điểm gần nhất thuộc 2 lớp dữ liệu đó tới ranh giới là xa nhất có thể. Trong không gian Euclid 3 chiều, siêu phẳng chính là mặt phẳng 2 chiều, còn trong không gian Euclid 2 chiều thì siêu phẳng là 1 đường thẳng. Đối với bài toán hồi quy, SVM có một tham số gọi là epsilon (ε), nếu dữ liệu nào nằm trong khoảng ±ε thì sẽ được thuật toán đánh giá là dự báo đúng, ngược lại thuật toán sẽ đánh giá sai.

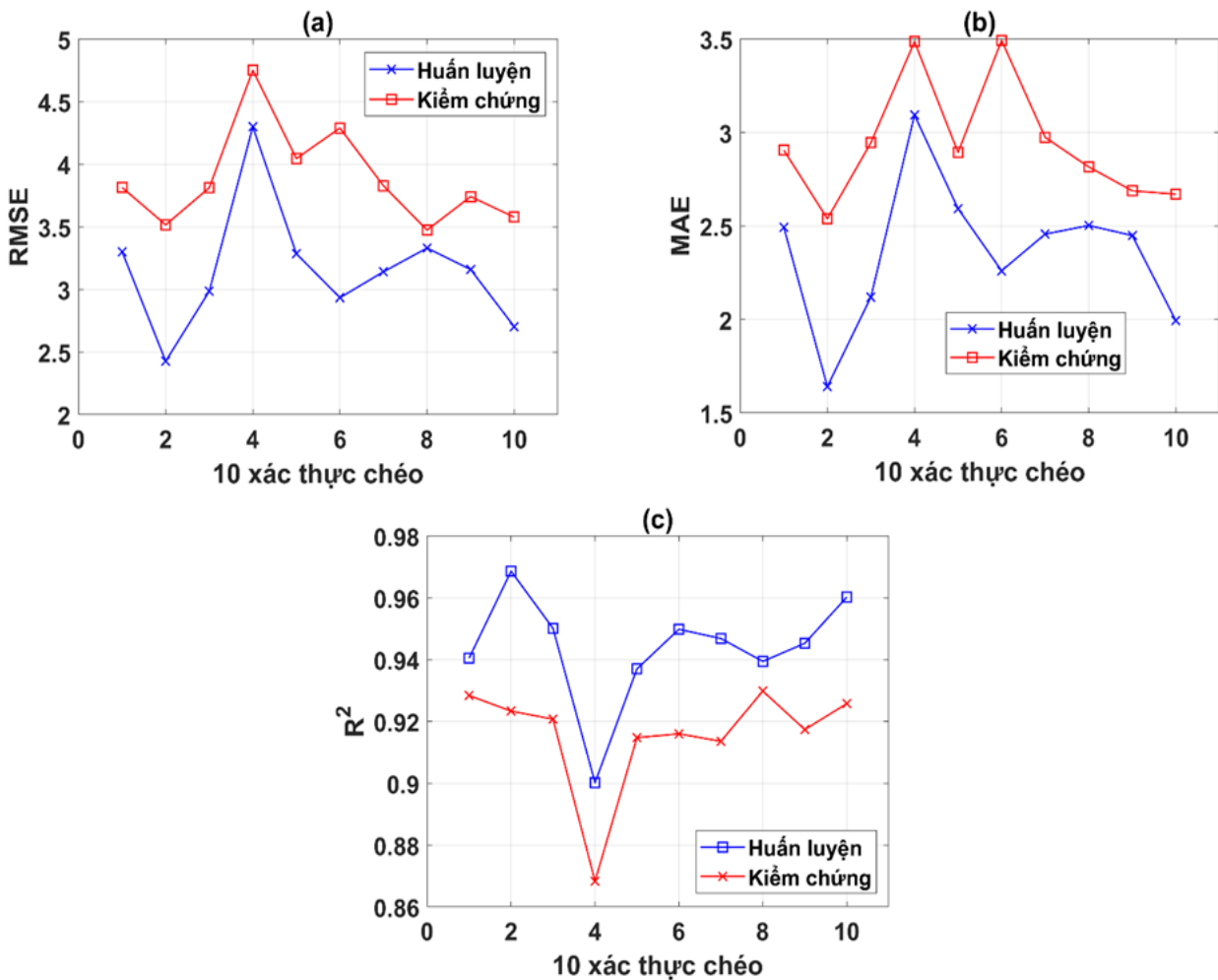


Hình 2. Mô hình máy vec tơ hỗ trợ (SVM)

Trong thuật toán SVM hồi quy, các điểm nằm ngoài siêu phẳng chính là các vec tơ hỗ trợ. Đại lượng ksi (ξ) trong thuật toán được sử dụng để xác định khoảng cách từ các siêu phẳng tới những điểm nằm ngoài siêu phẳng. Đây là một đại lượng sẽ được điều chỉnh bởi người sử dụng, đại diện cho sai số mà họ chấp nhận được đối với bài toán tương ứng. Nói một cách đơn giản (tham khảo Hình 2), nhiệm vụ của thuật toán SVM là đưa ra được các siêu phẳng dưới dạng phương trình (đường đỏ ở giữa) với các đường biên ±ε để giảm thiểu sai số nhiều nhất có thể và đưa được nhiều dữ liệu vào trong miền đó nhất có thể, song song với việc giảm ξ. Một mô hình đơn giản của SVM được minh họa trên Hình 2.



Hình 3. Minh họa kỹ thuật xác thực chéo sử dụng trong bài báo



Hình 4. Kết quả quá trình đào, xác thực và kiểm chứng của mô hình SVM sau 10 lần xác thực chéo (10 CVs) dựa trên các tiêu chí đánh giá hiệu suất khác nhau: (a) RMSE; (b) MAE; và (c) R^2

3.2. Xác thực chéo

Xác thực chéo (cross validation – CV) là một kỹ thuật được sử dụng rất nhiều trong các bài toán được giải bằng phương pháp máy học, trí tuệ

nhân tạo. CV là một giải pháp cung cấp khả năng ước tính, hoặc giúp tổng quát hóa hiệu suất của mô hình máy học đối với các dữ liệu mà mô hình đó chưa được biết đến (được học) trong quá trình đào tạo. CV được thực hiện bởi hai bước chính:

chia nhỏ dữ liệu thành các tập hợp con (được gọi là các nếp gấp - fold), sau đó luân phiên đào tạo và xác thực giữa chúng. Kỹ thuật chia tách thường có các đặc điểm là: (i) mỗi nếp gấp có kích thước xấp xỉ nhau, (ii) dữ liệu có thể được chọn ngẫu nhiên theo từng nếp gấp hoặc phân tầng, (iii) tất cả các nếp gấp được sử dụng để đào tạo mô hình ngoại trừ một nếp gấp được sử dụng để xác thực, lần gấp xác thực đó phải được xoay cho đến khi tất cả các nếp gấp trở thành nếp gấp xác thực một lần và chỉ một lần duy nhất. Thông thường, nếp gấp (k) sẽ được chọn là 5 hoặc 10 để đảm bảo khả năng đánh giá tổng quan hiệu suất của mô hình. Trong nghiên cứu này, giá trị $k=10$ được lựa chọn. Dữ liệu kiểm chứng trong bộ dữ liệu ban đầu được tách ra thành một phần riêng biệt, và quá trình huấn luyện (bao gồm đào tạo và xác thực mô hình) không hề sử dụng dữ liệu kiểm chứng đó. Minh họa kỹ thuật xác thực chéo CV được biểu thị trên hình 3 với 10 CV và 3 bộ dữ liệu gồm đào tạo, xác thực và kiểm chứng.

3.3. Biểu đồ phụ thuộc một phần (PDP)

Biểu đồ phụ thuộc một phần là một phương thức phổ biến dùng để diễn giải các mô hình học máy. Kết quả của PDP không những làm sáng tỏ được những quy luật tương quan giữa các tham số đầu vào và hàm dự báo, PDP còn cho phép người ta hiểu được cơ chế hoạt động của các mô hình máy học khác nhau, và cho phép người sử dụng biết được rằng chúng ta có được những thông tin gì từ cơ sở dữ liệu có sẵn, và làm sao để tái hiện lại những tương quan này trên thực tế.

Về bản chất, PDP là giá trị trung bình của các ICE (Individual Conditional Expectation), tạm dịch là giá trị kỳ vọng từng phần. ICE là đồ thị một đường biểu thị giá trị dự báo của mỗi mẫu trong cơ sở dữ liệu khi thay đổi mỗi tham số đầu vào. Giá trị các điểm trên đường ICE được tính bằng cách cố định tất cả các tham số còn lại, chỉ thay đổi tham số được nghiên cứu, và tính toán tham số đầu ra sử dụng công cụ học máy được tạo ra trước đó. Kết quả thu được khi tính toán với tất cả các mẫu trong dữ liệu sẽ tạo thành rất nhiều điểm giá trị và được nối với nhau thành một đường trên đồ thị. Sau đó, khi tính toán giá trị trung bình của tất cả các ICE ở các giá trị tham số đầu vào tương ứng, sẽ thu được đường trung bình PDP [16].

Nghiên cứu này sử dụng đồ thị PDP hai chiều, tương ứng với tập hợp rất nhiều đường PDP được tính cho 2 tham số đầu vào lần lượt thay đổi.

3.4. Đánh giá năng lực dự báo mô hình máy học

Việc đánh giá hiệu suất dự báo mô hình là bước quan trọng nhất sau khi xây dựng mô hình ML. Điều này quyết định tính sẵn sàng và độ tin cậy của mô hình được tạo thành. Các chỉ tiêu đánh giá hiệu suất được sử dụng trong nghiên cứu này bao gồm căn của sai số toàn phương trung bình (RMSE), sai số tuyệt đối trung bình (MAE) và hệ số xác định (R^2). MAE được cho là một tiêu chí đánh giá hồi quy đơn giản nhất, được tính bằng cách lấy giá trị trung bình của các phần dư giữa giá trị thực tế và giá trị dự đoán. Mỗi phần dư được lấy các giá trị tuyệt đối giúp cho các phần dư dương và phần dư âm không triệt tiêu lẫn nhau. MSE là thước đo mức độ xa/gần của đường hồi quy với các điểm. Do MSE lấy bình phương các phần dư, MSE sẽ luôn lớn hơn MAE, nên không thể so sánh MSE với MAE. Tuy nhiên, chúng ta có thể so sánh các giá trị này với chính chỉ số của một mô hình hồi quy khác, hoặc so sánh với cùng một mô hình nhưng với nhiều lần mô phỏng khác nhau. Đáng chú ý, sai số tăng theo bậc hai trong tiêu chí MSE. Điều này có nghĩa là MSE sẽ có tổng sai số cao hơn tiêu chí MAE nếu xuất hiện các kết quả dự báo có giá trị đặc biệt lớn hơn so với thực tế. Sự khác biệt lớn giữa giá trị dự đoán và giá trị thực tế sẽ được phản ánh rõ hơn khi sử dụng MSE so với MAE. Do MSE có giá trị cao hơn vì bản chất đang lấy bình phương sai số, nên nghiên cứu này sử dụng RMSE để đưa giá trị này về cùng thang điểm của sai số dự đoán. Điều này làm cho việc giải thích các sai số dễ dàng hơn. Đáng chú ý, cả MSE và RMSE đều lấy bình phương các phần dư, nên ảnh hưởng của hai chỉ tiêu này là tương đồng đối với các giá trị ngoại lai. Cuối cùng, R^2 là hệ số xác định, được xác định bởi tổng bình phương của số dư chia cho tổng của tổng trung bình. Giá trị của R^2 nằm trong khoảng từ 0 đến 1 và càng gần 1 thì sự phù hợp giữa dự đoán và giá trị thực càng tốt. Giá trị R^2 cao không phải lúc nào cũng đảm bảo hiệu suất dự báo tốt của một mô hình ML. Một mô hình có hiệu suất dự báo tốt có

thể có R^2 thấp và một mô hình có hiệu suất dự báo không tốt nhưng lại có thể có R^2 cao hơn. Điều này khẳng định trong các nghiên cứu về ML thì việc đánh giá nhiều chỉ tiêu khác nhau là đặc biệt quan trọng. Công thức tính toán ba chỉ tiêu trên có thể được tham khảo trong các tài liệu [25–27].

4. Kết quả và thảo luận

Trong phần này, quá trình xây dựng mô hình SVM được thực hiện. Cụ thể, quá trình này bao gồm hai giai đoạn: giai đoạn huấn luyện, là quá trình đào tạo mô hình đi kèm với xác thực chéo với 10 nếp gấp. Ở giai đoạn sau, khi công cụ SVM đạt được hiệu suất dự báo tối ưu trên bộ dữ liệu huấn luyện thì tiến hành đánh giá trên bộ dữ liệu kiểm chứng. Bộ dữ liệu huấn luyện (chiếm 70% số lượng mẫu) được chia thành 10 phần để tiến hành xác thực chéo. Với 10 lần mô phỏng như vậy, các chỉ tiêu đánh giá dự báo sẽ được lấy giá trị trung bình. Bộ dữ liệu kiểm chứng (chiếm 30% dữ liệu còn lại) được sử dụng để kiểm chứng khả năng dự báo của mô hình đối với những dữ liệu chưa được biết đến. Kết quả đánh giá hiệu suất dự báo mô hình SVM đối với cả hai tập dữ liệu được thể hiện ở Hình 4.

Có thể thấy được khi dữ liệu đào tạo bị thay đổi thì năng lực dự báo của SVM cũng thay đổi theo. Đối với tập dữ liệu huấn luyện, các tiêu chí đánh giá hiệu suất đều thay đổi trong những khoảng nhất định, nhưng biên độ dao động được đánh giá là khá ổn định: RMSE dao động quanh 3, với lần chạy tốt nhất có giá trị RMSE=2.5, còn lần chạy kém nhất là RMSE= 4.4, tương ứng với các CV theo thứ tự là CV2 và CV4. Nhận định tương tự cũng được kiểm chứng bởi sai số MAE, khi giá trị của MAE dao động trong khoảng từ 2-2.5, và các lần mô phỏng tốt nhất ở CV2 (MAE=1.6) và kém nhất ở CV4 (MAE=3.1). Tiêu chí đánh giá R^2 cũng đưa ra những nhận định tương tự, khi R^2 đạt giá trị quanh 0.95 và tốt nhất ở CV2 ($R^2= 0.97$), kém nhất ở CV4 ($R^2= 0.90$). Điều này có nghĩa là với tập dữ liệu huấn luyện, mô hình SVM được đào tạo có được năng lực dự báo rất tốt, có thể được lựa chọn để đưa ra thử nghiệm trên bộ dữ liệu kiểm chứng.

Xét bộ dữ liệu kiểm chứng gồm 66 mẫu thí nghiệm BTX hoàn toàn không được biết tới trong

quá trình đào tạo và xác thực. Điều đầu tiên có thể nhận thấy mô hình SVM đề xuất với 10 lần xác thực chéo đã có năng lực dự báo rất tốt, hơn nữa, không xuất hiện hiện tượng quá khớp (overfitting) vì năng lực của SVM ở trên tập huấn luyện là tốt hơn tập dữ liệu kiểm chứng. Mô hình SVM khi dự báo các dữ liệu mới đưa ra kết quả khá tốt, với RMSE đạt giá trị khoảng 3.9, MAE khoảng 3.0 và R^2 khoảng 0.92. Tuy nhiên, dữ liệu kiểm chứng xuất hiện sự khác biệt giữa các tiêu chí đánh giá hiệu suất mô hình. Cụ thể, RMSE và R^2 đưa ra CV8 là lần chạy tốt nhất, trong khi MAE lại kết luận lần chạy CV2 mới là tốt nhất. Có thể thấy sự khác biệt ở đây là không nhiều, và xét tổng thể thì mô hình SVM vẫn đưa ra được một hiệu suất dự báo cao ($R^2 = 0.92$). Với độ chính xác cao như vậy, mô hình này hoàn toàn có thể được sử dụng bởi các kỹ sư vật liệu giúp dự báo nhanh tính chất cơ học (cường độ nén) của BTX.

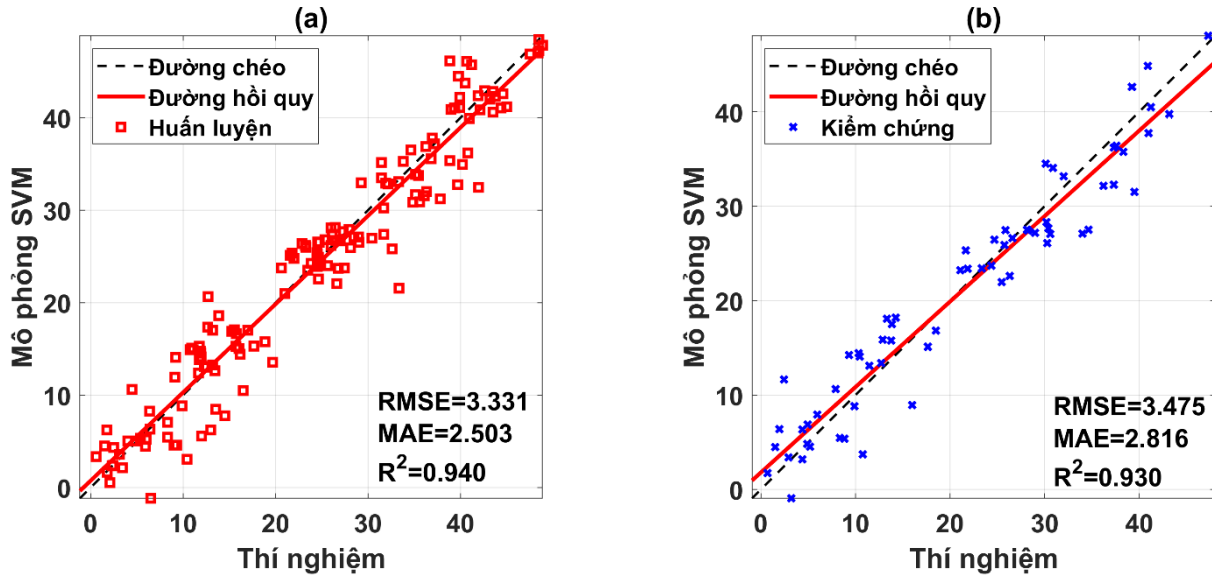
Tiếp theo, kết quả dự báo điển hình của CV8 được lựa chọn để trình bày, kết quả này được lựa chọn theo hai tiêu chí có phần nghiêm khắc hơn trong việc xét khả năng dự báo của mô hình (RMSE và R^2). Hình 5 trình bày phân tích hồi quy cho tập dữ liệu huấn luyện (Hình 5a) và cho tập dữ liệu kiểm chứng (Hình 5b). Trong mỗi hình, đường chéo được vẽ bằng nét đứt màu đen, thể hiện một tương quan lý tưởng cho bài toán ($R^2=1$). Bên cạnh đó, đường hồi quy cũng được thể hiện trên các hình vẽ bằng đường màu đỏ, và thường lệch đôi chút so với đường chéo hồi quy lý tưởng. Đối với mỗi trường hợp, các chỉ tiêu dự báo đều được tính và thể hiện: RMSE=3.331, MAE=2.503 và $R^2=0.940$ cho dữ liệu huấn luyện, còn RMSE=3.475, MAE=2.816 và $R^2=0.930$ cho dữ liệu kiểm chứng.

Ngoài ra, để phân tích cụ thể hơn, Hình 6 thể hiện biểu đồ phân bố sai số cho các dữ liệu huấn luyện và kiểm chứng tương ứng. Có thể nhận thấy khoảng 95% sai số giữa thí nghiệm cường độ BTX và kết quả từ mô phỏng SVM đều nằm trong khoảng ± 4.8 (MPa). Xét tới giá trị thực nghiệm cao nhất của cường độ BTX trong bộ dữ

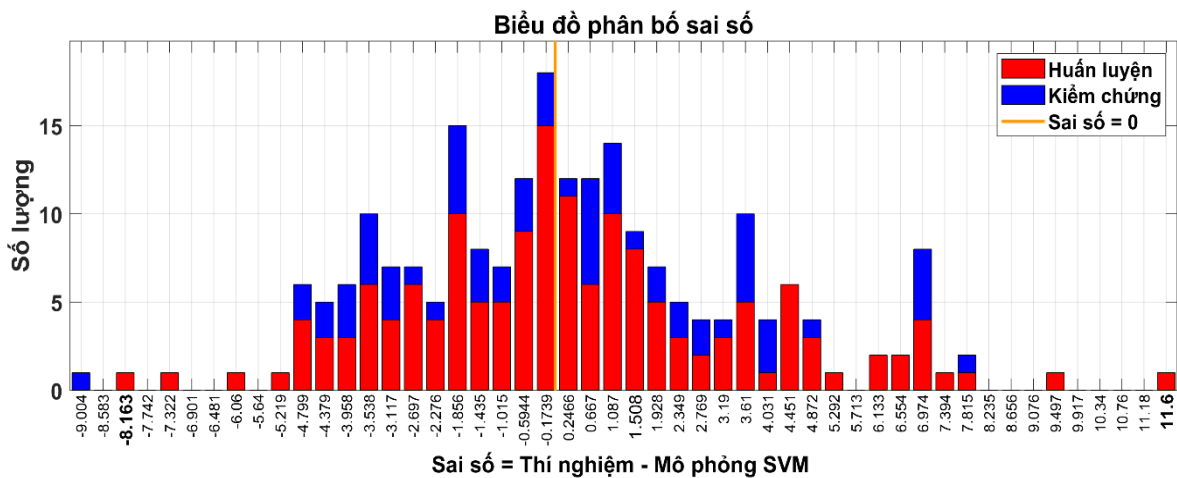
liệu (khoảng 50 MPa) thì sai số này nhỏ hơn 10%.

Với tiêu chí R^2 rất cao thu được ở cả hai bộ dữ liệu huấn luyện và kiểm chứng, có thể kết luận rằng công cụ mô phỏng số dựa trên SVM đạt hiệu suất tốt và hoàn toàn có thể được sử dụng cho những hướng nghiên cứu sâu hơn về BTX. Cùng

với các sai số RMSE và MAE được đánh giá là tương đối thấp, công cụ SVM đề xuất trong nghiên cứu này hoàn toàn có khả năng giúp cho việc thiết kế cấp phối vật liệu được hiệu quả hơn.



Hình 5. Biểu đồ hồi quy giữa các giá trị thí nghiệm và giá trị mô phỏng tính toán bởi SVM được xét tới trong nghiên cứu này cho: (a) phần huấn luyện; và (b) phần kiểm chứng



Hình 6. Biểu đồ phân bố và mối tương quan giữa các tham số đầu vào và đầu ra được xét tới trong nghiên cứu này

Cuối cùng, các biểu đồ phụ thuộc PDP 2 chiều được đề xuất cho bài toán cường độ nén BTX trên Hình 7. Trước tiên xem xét mối quan hệ giữa tỷ trọng và tỷ lệ nước/xi măng tại Hình 7a. Có thể thấy rằng khi tỷ trọng tăng thì cường độ nén của BTX cũng tăng dần. Nhưng giá trị cường độ nén đa phần nằm trong khoảng từ 15-35 MPa khi tỷ trọng thay đổi từ 400 tới 1700 kg/m³. Mặt khác, cường độ nén cũng phụ thuộc vào tỷ lệ nước/xi

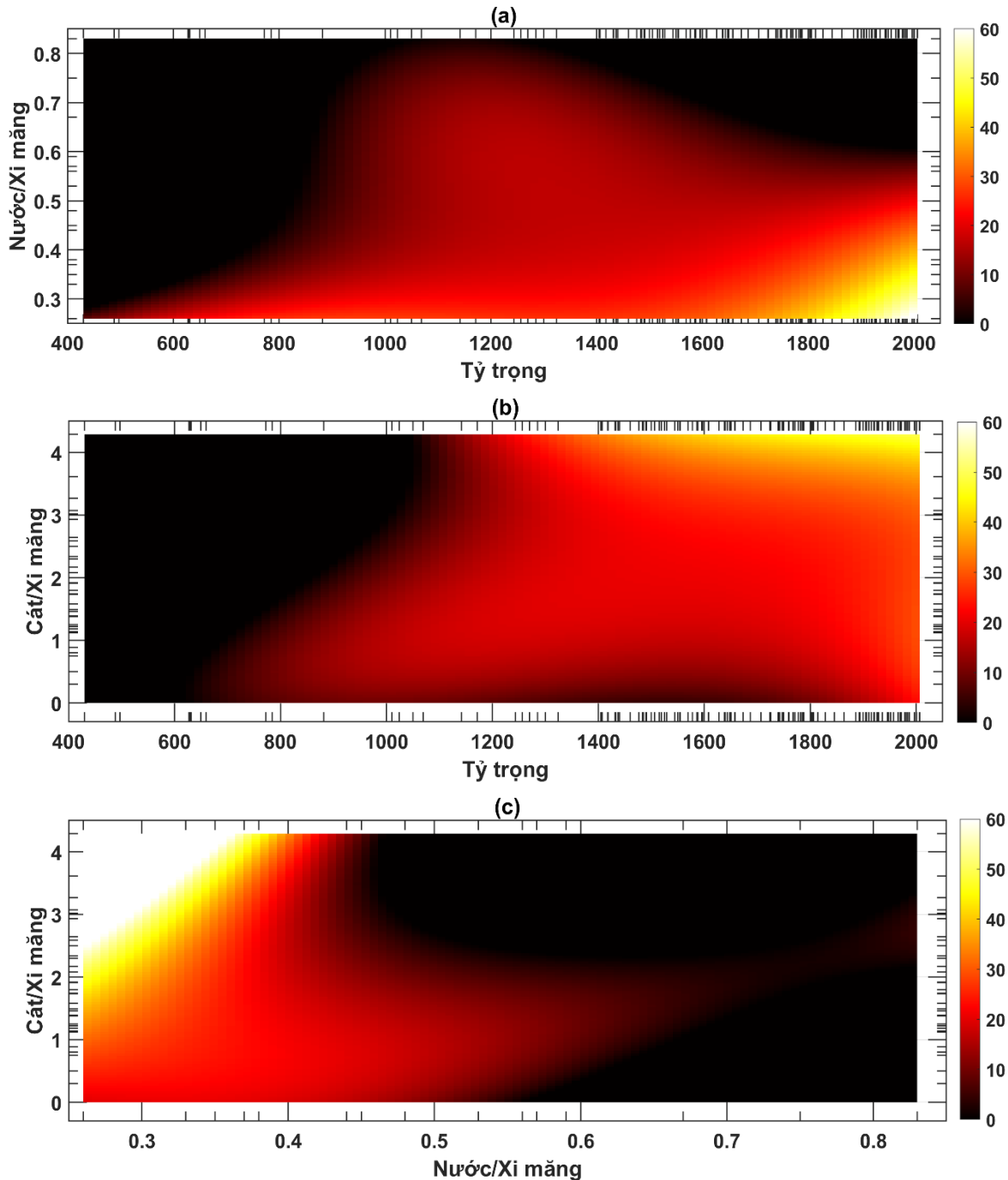
măng, và thường thì tỉ lệ nước/xi măng thấp sẽ cho cường độ cao hơn. Cường độ nén cao nhất có thể đạt được khi tỷ trọng đạt giá trị gần 2000 và tỷ lệ nước/xi măng thấp (từ 0.25 đến 0.35).

Hình 7b đưa ra tương quan tương tự nhưng với tỷ lệ cát/xi măng. Có thể thấy rằng tỷ lệ này ít quan trọng hơn so với nước/xi măng, vì những thay đổi của cường độ nén khi tỷ lệ này thay đổi từ 0 đến 4 là rất ít. Với tỷ trọng khoảng từ 1100 kg/m³

trở lên thì cường độ nén của BTX sẽ đạt từ 15 đến 30 MPa. Nếu muốn đạt tới cường độ cao hơn 50 MPa thì tỷ trọng phải đạt trên 1600 kg/m³, đồng thời tỷ lệ cát/xi măng phải đạt xấp xỉ 4.

Cuối cùng, Hình 7c cho thấy tương quan giữa tỷ lệ nước/xi măng với cát/xi măng và ảnh hưởng của chúng tới cường độ nén của BTX. Có

thể thấy rằng để đạt được cường độ cao thì tỷ lệ nước/xi măng phải thấp hơn 0.6, và nếu tỷ lệ này thấp hơn 0.4 thì không có BTX cường độ thấp (nhỏ hơn 25 MPa). Ngoài ra, nếu muốn thiết kế BTX cường độ cao thì tỷ lệ nước/xi măng phải ở quanh giá trị 0.3 còn tỷ lệ cát/xi măng nên nằm trong khoảng từ 3 đến 4.



Hình 7. Biểu đồ phân bố và mối tương quan giữa các tham số đầu vào và đầu ra được xét tới trong nghiên cứu này

Với ba dạng biểu đồ nhiệt được phát triển dựa vào công cụ SVM trong nghiên cứu này, các kỹ sư vật liệu có thể coi chúng như một công cụ trợ giúp hữu ích khi tính toán, thiết kế cấp phối cho

vật liệu BTX. Sử dụng các biểu đồ nhiệt này, họ có thể định lượng trước được các thành phần cấp phối để có thể thiết kế được BTX với cường độ theo mục đích sử dụng của công trình.

5. Kết luận

Nghiên cứu này đề xuất ứng dụng mô hình máy vec tơ hỗ trợ (SVM) để dự báo cường độ nén của BTX và xây dựng mối tương quan giữa các yếu tố đầu vào với cường độ nén. Bộ cơ sở dữ liệu gồm 220 kết quả thí nghiệm cường độ BTX được thu thập và dùng để xây dựng mô hình SVM. Để kết quả mô phỏng được tin cậy hơn và tìm ra được mô hình có năng lực dự báo tổng quan hơn, nghiên cứu này áp dụng kỹ thuật xác thực chéo với 10 nếp gấp trên tập dữ liệu huấn luyện. Kết quả của nghiên cứu chỉ ra công cụ SVM dự báo rất tốt cường độ nén của BTX với các chỉ số đánh giá hiệu suất như $RMSE = 3.475$ MPa, $MAE = 2.816$ MPa và $R^2 = 0.930$. Cuối cùng, biểu đồ phụ thuộc một phần hai chiều được phát triển để xây dựng tương quan giữa ba biến đầu vào nói trên với cường độ nén của BTX. Kết quả chỉ ra rằng nếu tỷ trọng BTX đạt 1600 trở lên, tỷ lệ nước/xi măng thấp hơn 0.4 và tỷ lệ cát/xi măng hơn 3 sẽ giúp BTX đạt cường độ cao nhất. Nghiên cứu này đặc biệt giúp ích cho các kỹ sư vật liệu trong công tác thiết kế cấp phối BTX.

Tài liệu tham khảo

- [1] K. Ramamurthy, E.K. Nambiar, G.I.S. Ranjani, A classification of studies on properties of foam concrete, *Cem. Concr. Compos.* 31 (2009) 388–396.
- [2] D.K. Panesar, Cellular concrete properties and the effect of synthetic and protein foaming agents, *Constr. Build. Mater.* 44 (2013) 575–584.
- [3] Z. Pan, F. Hiromi, T. Wee, Preparation of high performance foamed concrete from cement, sand and mineral admixtures, *J. Wuhan Univ. Technol.-Mater Sci Ed.* 22 (2007) 295–298.
- [4] P.J. Tikalsky, J. Pospisil, W. MacDonald, A method for assessment of the freeze–thaw resistance of preformed foam cellular concrete, *Cem. Concr. Res.* 34 (2004) 889–893.
- [5] S. Asadzadeh, S. Khoshbayan, Multi-objective optimization of influential factors on production process of foamed concrete using Box-Behnken approach, *Constr. Build. Mater.* 170 (2018) 101–110.
- [6] N.U. Kockal, T. Ozturan, Optimization of properties of fly ash aggregates for high-strength lightweight concrete production, *Mater. Des.* 32 (2011) 3586–3593.
- [7] C. Bing, W. Zhen, L. Ning, Experimental research on properties of high-strength foamed concrete, *J. Mater. Civ. Eng.* 24 (2011) 113–118.
- [8] M.Y.J. Liu, U.J. Alengaram, M.Z. Jumaat, K.H. Mo, Evaluation of thermal conductivity, mechanical and transport properties of lightweight aggregate foamed geopolymer concrete, *Energy Build.* 72 (2014) 238–245.
- [9] T. Nguyen, A. Kashani, T. Ngo, S. Bordas, Deep neural network with high-order neuron for the prediction of foamed concrete strength, *Comput.-Aided Civ. Infrastruct. Eng.* 34 (2019) 316–332.
- [10] A.A. Sayadi, J.V. Tapia, T.R. Neitzert, G.C. Clifton, Effects of expanded polystyrene (EPS) particles on fire resistance, thermal conductivity and compressive strength of foamed concrete, *Constr. Build. Mater.* 112 (2016) 716–724.
- [11] E.P. Kearsley, P.J. Wainwright, The effect of porosity on the strength of foamed concrete, *Cem. Concr. Res.* 32 (2002) 233–239.
- [12] Z.M. Yaseen, R.C. Deo, A. Hilal, A.M. Abd, L.C. Bueno, S. Salcedo-Sanz, M.L. Nehdi, Predicting compressive strength of lightweight foamed concrete using extreme learning machine model, *Adv. Eng. Softw.* 115 (2018) 112–125.
- [13] A. Ashrafian, F. Shokri, M.J.T. Amiri, Z.M. Yaseen, M. Rezaie-Balf, Compressive strength of Foamed Cellular Lightweight Concrete simulation: New development of hybrid artificial intelligence model, *Constr. Build. Mater.* 230 (2020) 117048.
- [14] T.-A. Nguyen, H.-B. Ly, V.Q. Tran, Investigation of ANN Architecture for Predicting Load-Carrying Capacity of Castellated Steel Beams, *Complexity*. 2021 (2021).
- [15] T.-A. Nguyen, H.-B. Ly, H.-V.T. Mai, V.Q. Tran, On the Training Algorithms for Artificial Neural Network in Predicting the Shear Strength of Deep Beams, *Complexity*. 2021 (2021).
- [16] H.-B. Ly, T.-T. Le, H.-L.T. Vu, V.Q. Tran, L.M. Le, B.T. Pham, Computational Hybrid Machine Learning Based Prediction of Shear

Capacity for Steel Fiber Reinforced Concrete Beams, Sustainability. 12 (2020) 2709. <https://doi.org/10.3390/su12072709>.

[17] H.-B. Ly, T.-A. Nguyen, V.Q. Tran, Development of deep neural network model to predict the compressive strength of rubber concrete, Constr. Build. Mater. 301 (2021) 124081.

[18] H.-B. Ly, M.H. Nguyen, B.T. Pham, Metaheuristic optimization of Levenberg–Marquardt-based artificial neural network using particle swarm optimization for prediction of foamed concrete compressive strength, Neural Comput. Appl. (2021) 1–21.

[19] H.-B. Ly, T.-A. Nguyen, B.T. Pham, Estimation of Soil Cohesion Using Machine Learning Method: A Random Forest Approach, Adv. Civ. Eng. 2021 (2021).

[20] T.-A. Nguyen, H.-B. Ly, B.T. Pham, Backpropagation neural network-based machine learning model for prediction of soil friction angle, Math. Probl. Eng. 2020 (2020).

[21] H.-B. Ly, P.G. Asteris, T.B. Pham, Accuracy assessment of extreme learning machine in predicting soil compression coefficient, Vietnam J. Earth Sci. 42 (2020) 228–336.

[22] T. Van Phong, H.-B. Ly, P.T. Trinh, I. Prakash, P. BTJVJOES, Landslide susceptibility mapping using Forest by Penalizing Attributes (FPA) algorithm based machine learning approach, Vietnam J Earth Sci. 42 (2020).

[23] D.V. Dao, H.-B. Ly, H.-L.T. Vu, T.-T. Le, B.T. Pham, Investigation and Optimization of the C-ANN Structure in Predicting the Compressive Strength of Foamed Concrete, Materials. 13 (2020) 1072. <https://doi.org/10.3390/ma13051072>.

[24] H.-B. Ly, B.T. Pham, Prediction of Shear Strength of Soil Using Direct Shear Test and Support Vector Machine Model, Open Constr. Build. Technol. J. 14 (2020). <https://benthamopen.com/EPUB/BMS-TOBCTJ-2019-HT6-957-1> (accessed May 4, 2020).

[25] B.T. Pham, M.D. Nguyen, H.-B. Ly, T.A. Pham, V. Hoang, H. Van Le, T.-T. Le, H.Q.

Nguyen, G.L. Bui, Development of Artificial Neural Networks for Prediction of Compression Coefficient of Soft Soil, in: C. Ha-Minh, D.V. Dao, F. Benboudjema, S. Derrible, D.V.K. Huynh, A.M. Tang (Eds.), CIGOS 2019 Innov. Sustain. Infrastruct., Springer Singapore, 2020: pp. 1167–1172.

[26] B.T. Pham, S.K. Singh^o, H.-B. Ly, Using Artificial Neural Network (ANN) for prediction of soil, Vietnam J. Earth Sci. 42 (2020) 311–319.

[27] T.-A. Nguyen, H.-B. Ly, A. Jaafari^o, B.T. Pham, Estimation of friction capacity of driven piles in clay using, Vietnam J. Earth Sci. 42 (2020) 265–275.