

## Mô phỏng sự phát triển của tảo xanh trên bề mặt vữa xi măng bằng phương pháp học máy

Estimating algal colonization on mortar surface using machine learning approach

Hoàng Nhật Đức<sup>a,\*</sup>, Nguyễn Trần Mộng Thùy<sup>b</sup>, Nguyễn Duy Tân<sup>c</sup>, Trần Thu Hiền<sup>d</sup>  
Nhat Duc Hoang, Mong Thuy Nguyen Tran, Duy Tan Nguyen, Thu Hien Tran

<sup>a</sup>*Viện Nghiên cứu phát triển Công nghệ Cao, Đại học Duy Tân, 03 Quang Trung, Đà Nẵng, Việt Nam*  
*Institute of Research and Development, Duy Tan University*

<sup>b</sup>*Công ty CP Tư vấn Đầu tư Xây dựng Hiệp Lực, Rạch Giá, Kiên Giang*  
*Hiep Luc Investment Consultation Joint-Stock Company, Rach Gia, Kien Giang*

<sup>c</sup>*Phòng Tài nguyên và Môi trường, Rạch Giá, Kiên Giang*  
*Department of Resource and Environment, Rach Gia city, Kien Giang province*

<sup>d</sup>*Khoa Xây Dựng, Đại học Duy Tân, 03 Quang Trung, Đà Nẵng, Việt Nam*  
*Faculty of Civil Engineering, Duy Tan University, 03 Quang Trung, Danang, Vietnam*

(Ngày nhận bài: 03/11/2018, ngày phản biện xong: 06/12/2018, ngày chấp nhận đăng: 20/1/2019)

### Tóm tắt

Sự phát triển của tảo xanh trên bề mặt công trình làm giảm hiệu suất năng lượng và gây ảnh hưởng lớn đến tính thẩm mỹ của tòa nhà. Vì vậy, sự phát triển của tảo phải được theo dõi và kiểm soát. Nghiên cứu của các tác giả áp dụng các mô hình tính toán dựa trên trí tuệ nhân tạo (AI) để dự báo quá trình phát triển của tảo trên bề mặt kết cấu công trình. Các kết quả thực nghiệm được thu thập từ 539 mẫu vữa bê tông có sự phát triển của tảo xanh. Các mô hình mạng nơ-ron nhân tạo (Artificial Neural Network - ANN) và mô hình phân tích hồi quy đa biến tự thích nghi (Multivariate adaptive regression splines - MARS) được sử dụng cho công tác mô phỏng. Kết quả tính toán cho thấy 2 mô hình AI là các công cụ hữu hiệu để dự báo mức độ phát triển của tảo xanh trên bề mặt vữa bê tông.

*Từ khóa:* Sự phát triển của tảo, bảo trì công trình, trí tuệ nhân tạo, mạng nơ-ron nhân tạo, mô hình phân tích hồi quy đa biến

### Abstract

The development of green algae on mortar surface degrades the energy efficiency and aesthetics of buildings. Thus, the algal development must be monitored and controlled. This study applies artificial intelligence approaches to forecast the development of green algae on building surface. The research data is established by collecting 539 mortar samples. Artificial Neural Network and Multivariate adaptive regression splines have been employed. Experimental result points out that the two AI models are useful tools for predicting algal development on mortar surface.

*Keywords:* Algal development, building maintenance, artificial intelligence, Artificial Neural Network, multivariate regression analysis.

### 1. Giới thiệu

Tảo xanh có thể được phát tán bằng gió như bào tử, nấm, phấn hoa, hoặc các dạng hạt khác

và khuếch tán trên mặt tiền tòa nhà được làm bằng vữa [1, 2]. Trong nhiều trường hợp, sự tăng trưởng sinh học gây ra nhiều tác hại khác nhau

cho công trình. Một hậu quả rõ ràng của việc xâm nhập của tảo xanh là nó tạo điều kiện cho các vi sinh vật khác sống trên bề mặt tòa nhà. Theo thời gian, các màng sinh học trở nên phù hợp cho các vi sinh vật đa dạng và phức tạp hơn, bao gồm các vi khuẩn dị dưỡng và nấm [3].

Những màng sinh học này là tác nhân chính cho sự thay đổi hóa học và vật lý của bề mặt tòa nhà và gây ra các hiệu ứng không mong muốn khác nhau cho cấu trúc tòa nhà. Trong nhiều trường hợp, các chất ngoại lai có thể hòa tan các thành phần của vật liệu xây dựng và làm hỏng bề mặt thông qua các chu trình làm ướt và hong khô lặp đi lặp lại; các vi sinh vật dạng sợi có thể thâm nhập vào vật liệu xây dựng và tách các cấu trúc hạt của vật liệu.

Ngoài ra, các vi sinh vật này tăng cường khả năng giữ nước của vật liệu bề mặt và dẫn đến giảm hiệu suất năng lượng của các tòa nhà. Sự xâm nhập của độ ẩm gây nên cả vấn đề sức khỏe thể chất và tinh thần cho người dân. Xét về tính thẩm mỹ của tòa nhà, sự tăng trưởng của tảo xanh có thể ảnh hưởng đáng kể đến bề mặt của mặt tiền bằng sự đổi màu và sự bám dính của các hạt bụi bẩn.

Sự xâm chiếm của tảo thường tạo điều kiện cho sự hiện diện của lớp vỏ và lớp vỏ cứng phức tạp mà rất khó loại bỏ [4]. Do đó, sự xuất hiện và phát triển của vi sinh vật trên bề mặt công trình có thể dẫn đến tổn thất kinh tế đáng kể cho chủ sở hữu do chi phí bảo trì và sửa chữa. Vì vậy, sự phát triển của tảo phải được theo dõi và kiểm soát nghiêm ngặt để giảm thiểu hậu quả có hại cho kết cấu của các tòa nhà.

Các phương pháp điều trị dự phòng thường liên quan đến việc sử dụng các chất chống thấm và chất diệt khuẩn để chống lại sự phát triển của vi sinh vật. Tuy nhiên, các hóa chất này có thể bị rửa trôi hoặc có thể bị phân hủy do vi khuẩn và vi sinh vật [5]. Vì các biện pháp khác nhau đòi hỏi các chi phí khác nhau và mỗi biện pháp thích hợp cho một số giai đoạn phát triển của tảo trên bề mặt tòa nhà, sự hiểu biết và dự báo tốt hơn về sự tăng trưởng của tảo có thể rất hữu ích cho việc xây dựng bảo trì.

Các nghiên cứu khác nhau đã nghiên cứu sự phát triển của tảo trên các bề mặt của kết cấu xây dựng. Bên cạnh các yếu tố môi trường, tỷ lệ phát triển của tảo đã được minh chứng là có phụ thuộc vào khả năng tiếp nhận sinh học của vật liệu [6]. Khả năng tiếp nhận sinh học của vật liệu đề cập đến khả năng của một vật liệu được xâm chiếm bởi các sinh vật sống; và nó liên quan đến các đặc điểm vật chất ảnh hưởng đến sự neo đậu và sinh trưởng của sinh vật.

Ngoài ra, các thông số vật liệu bao gồm độ nhám bề mặt, độ xốp bề mặt, độ pH bề mặt, độ ẩm và cấu trúc/kết cấu của vật liệu đã được chứng minh là ảnh hưởng đến các yếu tố xâm nhập của tảo. Tuy nhiên, các nghiên cứu định lượng liên quan đến quá trình xâm nhập của tảo, nhằm dự báo sự tăng trưởng của tảo, hiếm khi được thực hiện do tính chất phức tạp của vấn đề.

Gần đây, phương pháp học máy dựa trên trí tuệ nhân tạo có khả năng mô phỏng các quá trình phi tuyến, có thể cung cấp các giải pháp thay thế khả thi cho mô hình định lượng của quá trình phát triển của tảo. Trí thông minh nhân tạo sử dụng các kỹ thuật thống kê để cung cấp cho hệ thống máy tính khả năng tự tìm hiểu và khai phá từ dữ liệu mà không được lập trình một cách rõ ràng. Do các ưu điểm của chúng, các phương pháp trí tuệ nhân tạo để ước lượng tốc độ phát triển của tảo trên bề mặt công trình cần được xây dựng và kiểm chứng thêm.

Mục tiêu của nghiên cứu là xây dựng các mô hình dự báo quá trình phát triển của tảo trên bề mặt kết cấu dựa trên mạng nơron nhân tạo (Artificial Neural Network - ANN) và mô hình phân tích hồi quy đa biến tự thích nghi (Multivariate adaptive regression splines - MARS) dựa trên các kết quả thực nghiệm. Các kết quả thực nghiệm được thu thập từ 539 mẫu vữa bê tông được thu thập từ các nghiên cứu trước [7, 8]. Nghiên cứu sẽ đánh giá mức độ chính xác của mô hình, so sánh kết quả mô phỏng của các mô hình trí tuệ nhân tạo (ANN và MARS) với phương pháp truyền thống

là phân tích hồi quy đa biến (Multiple Linear Regression - MLR).

## 2. Dữ liệu thực nghiệm

Một bộ số liệu, bao gồm 539 thí nghiệm trên tảo xanh *Klebsormidium flaccidum*, đã được thu thập từ các nghiên cứu trước đây [8]. Sự phát triển của tảo trên bề mặt vữa là biến số được đo đạc và khảo sát. Bốn mươi chín hỗn hợp vữa được làm bằng xi măng, cát silic và phụ gia. Hỗn hợp vữa được đúc thành các khuôn nhựa có kích cỡ  $50 \times 50 \times 1$  cm, được lưu trữ ở  $21 \pm 1^\circ\text{C}$  và  $95 \pm 5\%$  độ ẩm tương đối trong 7 ngày trước khi được cắt thành các mẫu  $20 \times 8 \times 1$  (cm). Tất cả các mẫu vữa được thực hiện theo tiêu chuẩn NF-EN 196-1. Các mẫu vữa này sau đó được thử nghiệm cho

khả năng sinh sản của tảo. Trên mỗi loại vữa, ba mẫu thử được thực hiện để nâng cao độ tin cậy của kết quả thí nghiệm.

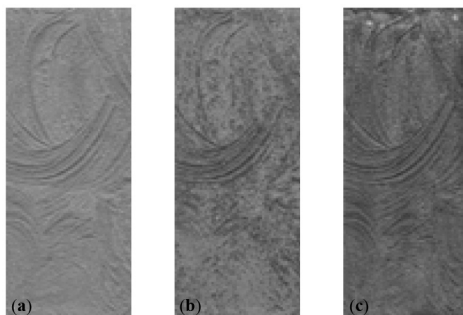
Trong quá trình thí nghiệm, bề mặt vữa đã được phun dung dịch từ một hệ thống treo tảo xanh. Lưu ý rằng trong suốt các thí nghiệm, các thông số bao gồm tốc độ phun dung dịch, cường độ ánh sáng, nhiệt độ, nồng độ ban đầu của huyền phù tảo, và độ nghiêng góc của mẫu được giữ cố định. Trong nghiên cứu này, sáu yếu tố ảnh hưởng (độ nhám bề mặt, độ xốp, độ pH bề mặt, độ thấm thấu, loại xi măng và giai đoạn phát triển của tảo) được sử dụng để mô hình hóa tốc độ sinh trưởng của tảo xanh *Klebsormidium flaccidum* trên bề mặt vữa.

Bảng 1. Mô tả thống kê của tập dữ liệu

Biến số		Đơn vị	Min	Trung bình	Trung vị	Độ lệch chuẩn	Max
X1	Độ nhám	( $\mu\text{m}$ )	23.50	104.30	61.75	75.70	281.25
X2	Độ rỗng	(%)	10.64	25.28	31.00	10.17	37.18
X3	pH bề mặt		8.70	10.02	9.05	1.24	12.15
X4	Hệ số thấm	(10-10m/s)	4.05	19.13	17.07	14.46	53.07
X5	Loại xi măng		0.00	0.33	0.00	0.47	1.00
X6	Diện tích phát triển	(%)	0.50	50.05	50.00	31.55	100.00
Y	Thời gian phát triển	(day)	2.00	27.00	20.00	20.00	96.00

Hình 1 thể hiện ba mẫu khi 0,5% (a), 50% (b) và 100% (c) bề mặt vữa được thuộc địa bởi tảo xanh *Klebsormidium flaccidum*. Hơn nữa, thời gian thuộc địa (Y) là khoảng thời gian (được đo trong ngày) cần thiết cho tảo để đạt được các giai

đoạn thuộc địa (từ 0,5% đến 100%), được sử dụng như đầu ra mô hình (Y). Các mô tả thống kê về thời gian thuộc địa cũng như sáu yếu tố ảnh hưởng nói trên trong tập dữ liệu được cung cấp trong Bảng 1.



Hình 1. Minh họa các bề mặt vữa có tảo xanh: (a) 0.5% độ che phủ bề mặt; (b) 50% độ che phủ bề mặt; (c) 100% độ che phủ bề mặt

### 3. Các phương pháp trí tuệ nhân tạo

#### 3.1. Mạng nơron nhân tạo

Mạng nơron nhân tạo (Artificial Neural Network), được gọi tắt là mạng nơron ANN, là một mô hình xử lý thông tin phỏng theo cách thức xử lý thông tin của các hệ nơron sinh học. Nó được tạo lên từ một số lượng lớn các phần tử (gọi là phần tử xử lý hay nơron) kết nối với nhau thông qua các liên kết (gọi là trọng số liên kết) làm việc như một thể thống nhất để giải quyết một vấn đề cụ thể nào đó [9]. Một mạng nơron nhân tạo được cấu hình cho một ứng dụng cụ thể (nhận dạng mẫu, phân loại dữ liệu,...) thông qua một quá trình học từ tập các mẫu huấn luyện [10]. Về bản chất, học chính là quá trình hiệu chỉnh trọng số liên kết giữa các nơron.

Nhiệm vụ đặt ra là huấn luyện mạng nơron để thực hiện xấp xỉ hàm số  $f: X \in R^D \rightarrow Y \in R^1$ ; trong đó,  $D$  là số thuộc tính đầu vào [11]. Một mô hình ANN, bao gồm các lớp đầu vào, lớp ẩn và lớp đầu ra.  $W1$  và  $W2$  biểu diễn tương ứng các ma trận trọng lượng của lớp ẩn và lớp đầu ra;  $N$  biểu thị số lượng nơron trong lớp ẩn;  $b_1 = [b_{11}, b_{12}, \dots, b_{1N}]$  biểu thị véc tơ phân cực của lớp ẩn;  $b_2$  là véc tơ phân cực của lớp đầu ra;  $f_A$  biểu thị hàm kích hoạt (ví dụ, log-sigmoid).

Cấu trúc ANN được sử dụng để phân tích hồi quy được biểu diễn như sau [7]:

$$f(X) = b_2 + W_2 \times (f_A(b_1 + W_1 \times X)) \quad (1)$$

Nói chung, các ma trận trọng lượng và vectơ phân cực của một ANN được thực hiện thông qua quá trình sử dụng quy tắc đẩy lùi lỗi. Hơn nữa, sai số quân phương (MSE) được sử dụng như một hàm mục tiêu để thực hiện cấu trúc ANN cho chức năng tính toán xấp xỉ hàm:

$$MSE = \min_{w_1, w_2, b_1, b_2} \frac{1}{M} \sum_{i=1}^M e_i^2 \quad (2)$$

#### 3.2. Mô hình phân tích hồi quy đa biến tự thích nghi (MARS)

MARS, đề xuất bởi Friedman [12], là một phương pháp để xây dựng các phương trình dùng

để mô hình hóa dữ liệu. Phương pháp này chia không gian các biến số đa chiều thành các miền con của các biến dự đoán và thiết lập mối quan hệ ảnh xạ giữa các biến dự đoán  $X$  và biến đầu ra  $Y$  [13].

MARS sử dụng các hàm tuyến tính từng phần để phù hợp với từng miền cục bộ và sử dụng một phương pháp thích ứng để xác định mô hình tổng hợp. Theo Friedman [12], MARS có thể được coi là sự tổng quát hóa hồi quy tuyến tính từng bước hoặc một biến thể của cây hồi quy với mục đích đạt được khả năng mô hình hóa tốt hơn so với phương pháp hồi quy thông thường. Các nghiên cứu trước đây đã chỉ ra MARS là một công cụ AI có khả năng mô tả các mối tương quan phức tạp giữa các biến số [14 - 16].

Một mô hình MARS được thể hiện thông qua một loạt các hàm cơ bản đơn giản mô tả mối quan hệ giữa các biến đầu vào và đầu ra. Một hàm cơ sở được hiển thị như sau:

$$b_m(x) = \max(0, C - x) \text{ hoặc } b_m(x) = \max(0, x - C) \quad (3)$$

trong đó  $b_m$  biểu thị một hàm cơ sở;  $x$  là biến đầu vào;  $C$  đại diện cho tham số ngưỡng được sử dụng để chia phạm vi ban đầu của  $x$  thành các không gian cục bộ. Dạng tổng quát của mô hình được thể hiện như sau:

$$f(x) = \alpha_0 + \sum_{m=1}^M \alpha_m b_m(x) \quad (4)$$

với  $\alpha_0, \alpha_1, \dots, \alpha_M$  là hệ số trọng số của mô hình MARS;  $f(x)$  đại diện cho đầu ra của mô hình.  $M$  là số lượng hàm cơ bản.

Việc thiết lập mô hình MARS được chia thành hai bước: bước 1 và bước 2. Trong bước 1, các hàm cơ sở được thêm vào mô hình để chúng có thể giúp giảm lỗi mô phỏng; quá trình này chấm dứt khi đạt đến số lượng hàm cơ bản tối đa. Bước thứ hai nhằm mục đích giảm hiện tượng 'overfitting' bằng cách cắt tỉa các hàm cơ bản dư thừa; mỗi mô hình con của MARS được đánh giá bằng chỉ số xác thực chéo tổng quát (GCV):

$$GCV = MSE / (1 - \frac{k + 0.5c(k-1)}{n})^2 \quad (5)$$

trong đó MSE là viết tắt của sai số trung bình của mô hình được tính toán với dữ liệu huấn luyện.

$k$  biểu thị số hàm cơ sở.  $n$  đại diện cho số quan sát trong dữ liệu huấn luyện.  $c$  là hệ số phạt; Friedman [12] and Jekabsons [17] khuyến nghị nên tìm kiếm tham số này trong phạm vi [2, 4].

**4. Kết quả mô phỏng sự phát triển của tảo xanh với ANN và MARS**

Để đánh giá sự chính xác của các mô hình hồi quy đã được xây dựng, nghiên cứu, bài báo sử dụng 3 chỉ số *RMSE* (căn bậc 2 của trung bình của tổng bình phương sai số), *MAPE* (giá trị trung bình của tỷ lệ giữa sai số và giá trị thực tế của biến số), và  $R^2$  (hệ số xác định) [18]. Chỉ số *RMSE* cho biết giá trị của độ lệch chuẩn của sai khác giữa giá trị của giá trị tỷ lệ diện tích vữa bao phủ thực tế và giá trị tỷ lệ diện tích vữa bao phủ được tính toán ra từ mô hình. Nếu *RMSE* càng nhỏ thì mô hình càng chính xác. Chỉ số *MAPE* cho biết giá trị trung bình của tỷ lệ sai số (giữa giá trị thực tế và giá trị tính toán) và giá trị thực tế. Thêm vào đó, để phục vụ cho quá trình xây dựng mô hình tính toán, 90% dữ liệu được dùng để huấn luyện các mô hình, 10% còn lại được sử dụng cho quá trình kiểm chứng.

Để phục vụ cho việc kiểm chứng tính ưu việt của các mô hình ANN và MARS, kết quả của 2 mô

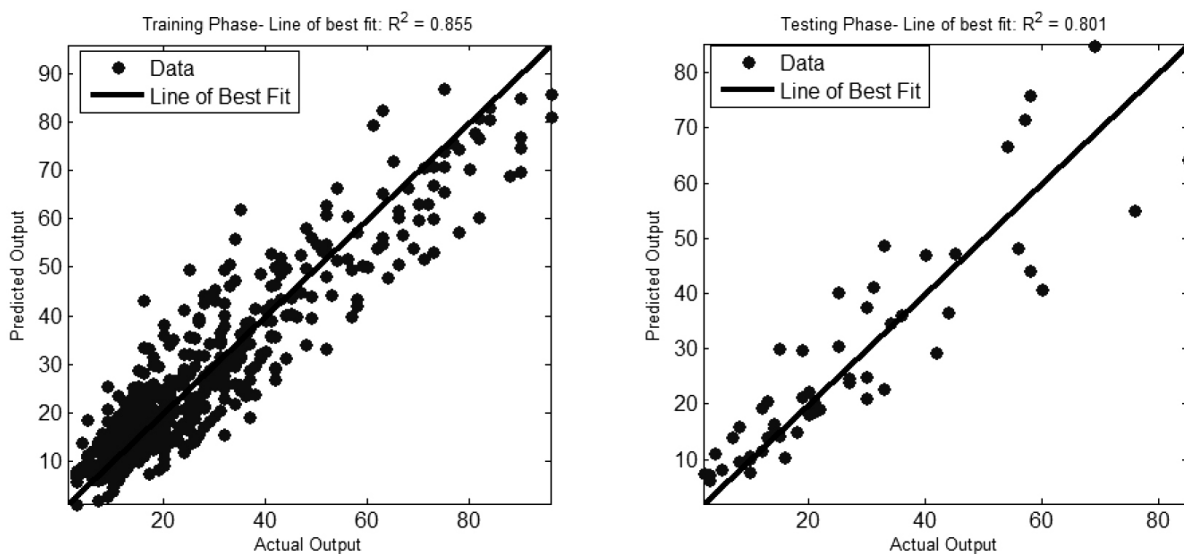
hình này được đem ra so sánh với mô hình phân tích hồi quy đa biến truyền thống. Kết quả của mô hình hồi quy đa biến sử dụng 6 biến dự báo (**X1, X2, X3, X4, X5, X6**) cho kết quả như sau  $RMSE = 13.203$ ,  $MAPE = 75.716\%$ ,  $MAE = 10.387$ , và  $R^2 = 0.581$ . Thêm vào đó, mô hình ANN đã được huấn luyện với các thông số được cho như sau:

- Số nơron đầu vào NI = 6
- Số nơron đầu ra NO = 1
- Số nơron trong lớp ẩn NR = 3 đến 9
- Hàm kích hoạt: Sigmoid

Bảng 2. Kết quả tính toán từ mô hình ANN

Chỉ số	ANN						
	Nr = 3	Nr = 4	Nr = 5	Nr = 6	Nr = 7	Nr = 8	Nr = 9
RMSE (%)	12.04	11.66	11.52	10.95	9.00	10.86	12.42
MAPE	37.24	41.98	47.50	41.22	37.71	35.24	46.97
MAE	8.22	8.60	8.72	7.74	6.84	7.69	8.58
$R^2$	0.64	0.66	0.68	0.70	0.80	0.71	0.62

Kết quả tính toán của mô hình ANN được cho trong Bảng 2 với số nơron trong lớp ẩn là 7 cho kết quả tốt nhất:  $RMSE = 9.00$ ,  $MAPE = 37.71\%$ ,  $MAE = 6.84$ , và  $R^2 = 0.80$ . Kết quả này cho thấy sự cải thiện về độ chính xác so với mô hình truyền thống MLR. So sánh giữa dữ liệu thực tế và dữ liệu tính toán từ mô hình ANN được minh họa trong Hình 2.



Hình 2. Kết quả tính toán từ mô hình ANN: dữ liệu thực tế và dữ liệu tính toán từ mô hình (nơron = 7)

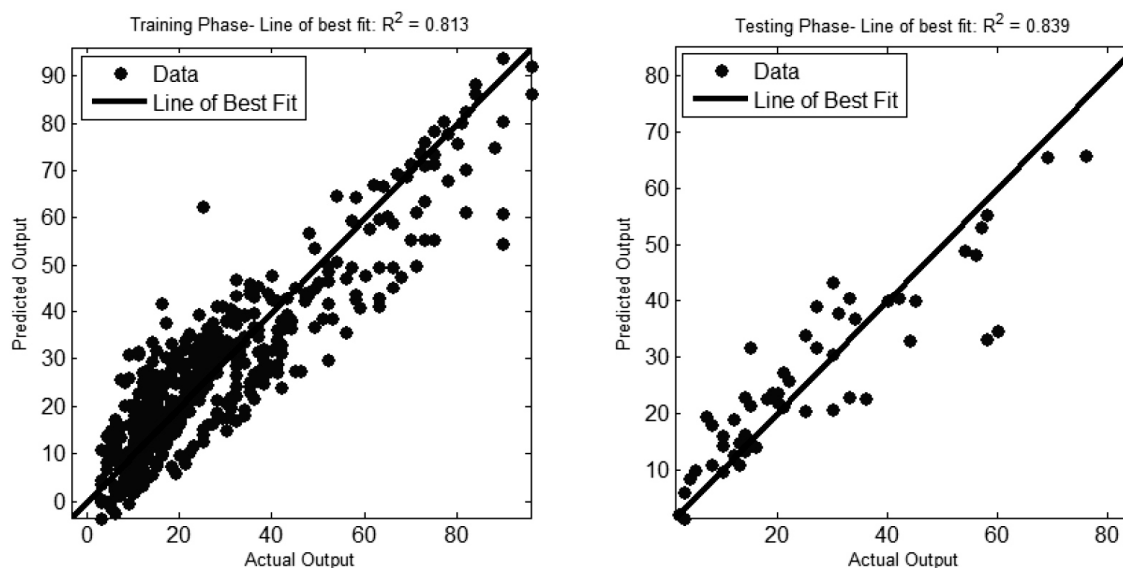
Mô hình MARS dùng cho quá trình tính toán sử dụng các hàm được viết bởi Jekabsons [17].

Kết quả của mô hình MARS trong quá trình huấn luyện kiểm chứng được tóm tắt trong Bảng 3

với số hàm cơ bản  $f_{max} = 5, 10, 15,$  và  $20$ . Mô hình MARS với  $f_{max} = 15$  cho kết quả tốt nhất:  $RMSE = 8.25, MAPE = 33.56\%, MAE = 7.31,$  và  $R^2 = 0.73$ . Kết quả này cho thấy sự cải thiện về độ chính xác so với mô hình MLR ( $RMSE = 13.203, MAPE = 75.716\%, MAE = 10.387$ ). So với mô hình ANN ( $RMSE = 9.00, MAPE = 37.71\%, MAE = 6.84,$  và  $R^2 = 0.80$ ), chỉ số  $RMSE$  và  $MAPE$  của MARS tốt hơn. Hình 3 mô tả kết quả tính toán từ mô hình MARS.

Bảng 3. Kết quả tính toán từ mô hình MARS

Chi số	MARS			
	$f_{max} = 5$	$f_{max} = 10$	$f_{max} = 15$	$f_{max} = 20$
RMSE (%)	14.29	10.41	8.25	8.26
MAPE	61.26	38.61	33.56	32.48
MAE	11.19	7.31	6.19	6.00
$R^2$	0.49	0.73	0.84	0.83



Hình 3. Kết quả tính toán từ mô hình MARS

### 5. Kết luận

Nghiên cứu trong bài báo đã áp dụng các mô hình tính toán dựa trên học máy để thực hiện dự đoán mức độ dự báo quá trình phát triển của tảo trên bề mặt kết cấu công trình, nhằm cung cấp cho các đơn vị bảo trì tòa nhà một công cụ hiện đại trong việc dự báo sự phát triển của tảo trên bề mặt vữa. Các kết quả thực nghiệm được thu thập từ 539 mẫu vữa bê tông được thu thập từ các nghiên cứu trước. Các mô hình ANN và MARS đã cho thấy tính ưu việt trong khả năng mô phỏng và dự báo so với phương pháp phân tích hồi quy đa biến truyền thống. Hướng nghiên cứu tiếp theo của đề tài là tiến hành các thí nghiệm để thu thập các dữ liệu về sự phát triển của tảo xanh tại môi trường Việt Nam đặc biệt là các vùng mưa nhiều và có khí hậu phù hợp cho sự phát triển của tảo trên bề mặt công trình.

### Tài liệu tham khảo

- [1] H. Barberousse, B. Ruot, C. Yéprémian, G. Boulon, An assessment of façade coatings against colonisation by aerial algae and cyanobacteria, *Build. Environ.*, 42 (2007) 2555-2561.
- [2] E. Dalod, A. Govin, R. Guyonnet, P. Grosseau, C. Lors, D. Damidot, Influence of the chemical composition of mortars on algal biofouling, In: *Proc. of the International Conference on Calcium Aluminate*, May 2014 Palais des Papes, Avignon, France, 2014, pp. 523 - 534.
- [3] C.C. Gaylarde, Morton, L.H.G., Deteriogenic biofilms on buildings and their control: a review, *Biofouling*, 14 ( 1999) 59-74.
- [4] C. Saiz-Jimenez, Deposition of anthropogenic compounds on monuments and their effect on airborne microorganisms, *Aerobiologia*, 11 (1995) 161-175.
- [5] M. Burkhardt, T. Kupper, S. Hean, R. Haag, P. Schmid, M. Kohler, M. Boller, Biocides used in building materials and their leaching behavior to

- sewer systems, *Water Sci. Technol.*, 56 (2007) 63-67.
- [6] T.-H. Tran, N.-D. Hoang, Predicting algal appearance on mortar surface with ensembles of adaptive neuro fuzzy models: a comparative study of ensemble strategies, *International Journal of Machine Learning and Cybernetics*, (2018).
- [7] T.-H. Tran, N.-D. Hoang, Predicting Colonization Growth of Algae on Mortar Surface with Artificial Neural Network, *Journal of Computing in Civil Engineering*, 30 (2016) 04016030.
- [8] T.-H. Tran, N.-D. Hoang, Estimation of algal colonization growth on mortar surface using a hybridization of machine learning and metaheuristic optimization, *Sādhanā*, 42 (2017) 929-939.
- [9] N.-D. Hoang, D. Tien Bui, GIS-Based Landslide Spatial Modeling Using Batch-Training Back-propagation Artificial Neural Network: A Study of Model Parameters, Springer International Publishing, Cham, 2018, pp. 239-254.
- [10] C.M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* Springer (April 6, 2011), ISBN-10: 03873107382011.
- [11] M.T. Hagan, H.B. Demuth, M.H. Beale, O.D. Jesús, *Neural Network Design (2nd Edition)*, Martin Hagan; 2 edition (September 1, 2014), ISBN-10: 09717321162014.
- [12] J.H. Friedman, Multivariate Adaptive Regression Splines, *Ann. Statist.*, 19 (1991) 1-67.
- [13] A. Parsaie, A.H. Haghiabi, M. Saneie, H. Torabi, Prediction of energy dissipation on the stepped spillway using the multivariate adaptive regression splines, *ISH J. Hydraul. Eng.*, 22 (2016) 281-292.
- [14] A. Gholampour, I. Mansouri, O. Kisi, T. Ozbakkaloglu, Evaluation of mechanical properties of concretes containing coarse recycled concrete aggregates using multivariate adaptive regression splines (MARS), M5 model tree (M5Tree), and least squares support vector regression (LSSVR) models, *Neural Computing and Applications*, (2018).
- [15] D.K. Roy, B. Datta, Multivariate Adaptive Regression Spline Ensembles for Management of Multilayered Coastal Aquifers, *Journal of Hydrologic Engineering*, 22 (2017) 04017031.
- [16] N.-D. Hoang, C.-T. Chen, K.-W. Liao, Prediction of chloride diffusion in cement mortar using Multi-Gene Genetic Programming and Multivariate Adaptive Regression Splines, *Measurement*, 112 (2017) 141-149.
- [17] G. Jekabsons, ARESLab: Adaptive Regression Splines toolbox for Matlab/Octave, Technical report, Riga Technical University, available at <http://www.cs.rtu.lv/jekabsons/>, (2016).
- [18] N.Đ. Hoàng, A.Đ. Phạm, Xây dựng mô hình hồi quy đa biến để tính toán cường độ chịu nén của hỗn hợp bê tông tính năng cao, *Journal of Science and Technology*, 2 (2016) 9-14.