

# Phân loại sự cố mô-đun quang điện dựa trên điểm công suất cực đại và kỹ thuật máy học

Nguyễn Phước Hoàng Khang<sup>1</sup>, Nguyễn Chí Nhân<sup>1,2,\*</sup>



Use your smartphone to scan this QR code and download this article

<sup>1</sup>Khoa Vật lý - Vật lý Kỹ thuật, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

<sup>2</sup>Phòng thí nghiệm Thiết kế vi mạch, Trường Đại học Khoa học Tự Nhiên, ĐHQG-HCM

## Liên hệ

**Nguyễn Chí Nhân**, Khoa Vật lý - Vật lý Kỹ thuật, Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

Phòng thí nghiệm Thiết kế vi mạch, Trường Đại học Khoa học Tự Nhiên, ĐHQG-HCM

Email: ncnhan@hcmus.edu.vn

## Lịch sử

- Ngày nhận: 12-8-2022
- Ngày chấp nhận: 20-12-2022
- Ngày đăng: 31-12-2022

DOI: 10.32508/stdjns.v6i4.1221



## Bản quyền

© ĐHQG Tp.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



## TÓM TẮT

Mô-đun quang điện (Photovoltaic - PV) là thành phần quan trọng trong hệ thống năng lượng mặt trời. Việc phân loại sự cố đối với mô-đun quang điện là cần thiết để đảm bảo an toàn, hiệu quả và độ tin cậy của hệ thống quang điện. Khi xảy ra lỗi của mô-đun quang điện, đặc tuyến cường độ dòng điện-điện áp (đặc tuyến I-V) của mô-đun quang điện thay đổi dẫn tới sự dịch chuyển về cường độ dòng điện và điện áp của điểm công suất cực đại (Max Power Point - MPP). Trong bài báo này, việc phân loại sự cố mô-đun quang điện dựa trên điểm MPP và kỹ thuật máy học được đề xuất. Thông qua việc phân tích vị trí của điểm MPP trong miền không gian tham số I-V, các sự cố của mô-đun PV trong quá trình vận hành có thể được xác định và phân loại. Các thông số về điện áp và cường độ dòng điện của mô-đun PV trong các sự cố (sự cố dây – dây, bóng che, hở mạch một phần, hở mạch toàn phần, ngắn mạch trong các điều kiện cường độ bức xạ khác nhau) được thu thập thông qua mô phỏng trên Simulink/MATLAB. Sau đó, các tham số này sẽ được xử lý và xây dựng thành một tập dữ liệu. Tập dữ liệu này sẽ được đưa vào mô hình phân loại sự cố mô-đun PV bằng cách sử dụng các thuật toán máy học bao gồm Máy vectơ hỗ trợ (Support Vector Machine - SVM) và k-means. Kết quả mong muốn của các mô hình máy học là phân nhóm chính xác các sự cố bằng cách sử dụng tham số điện áp và cường độ dòng điện của điểm MPP và khảo sát chẩn đoán sự cố thông qua khả năng phân tách ngưỡng của mô hình. Kết quả cho thấy khả năng phân loại sự cố và độ chính xác của mô hình máy học sử dụng thuật toán SVM (98,9%) tốt hơn so với thuật toán k-mean (61,1%).

**Từ khóa:** tấm pin quang điện, phân loại sự cố, đặc tuyến I-V, SVM, k-means, máy học

## GIỚI THIỆU

Như mọi hệ thống khác, trong quá trình hoạt động, một hệ thống điện năng lượng mặt trời (NLMT) có khả năng xảy ra những sự cố không mong muốn. Những sự cố này xuất hiện do hỏng hóc, suy giảm từ các thành phần bên trong hệ thống. Các trường hợp sự cố trong dây pin quang điện được khảo sát trong đề tài này bao gồm:

- Dây-Dây (Line-Line): Sự cố do chạm dây kết nối hai tấm pin quang điện khác nhau.
- Bóng che (Partial Shading): Sự cố do cường độ bức xạ mặt trời chiếu lên một hay nhiều tấm pin trong dây không giống nhau do các yếu tố bên ngoài.
- Hở mạch (Open Circuit): Sự cố do dây tấm pin bị hở mạch.
- Ngắn mạch (Short Circuit): Sự cố do chạm mạch ở 2 đầu dây ngõ ra của dây.

Khi xảy ra sự cố, đường đặc tuyến I-V hoạt động của tấm pin sẽ thay đổi. Sự thay đổi này dẫn tới việc điểm MPP sẽ dịch chuyển khi có sự cố xảy ra. Do đó, ta có thể thực hiện nhận diện và phân loại sự cố dựa trên sự chênh lệch giữa điểm MPP hoạt động (được đảm bảo nhờ bộ MPPT tích hợp trong hệ thống điện NLMT)

và điểm MPP tính toán lý thuyết trong cùng một điều kiện cường độ bức xạ.

Phân tích tình trạng hoạt động của tấm pin quang điện được thực hiện từ các thông tin về dòng điện, điện áp và công suất hoạt động của tấm pin. Các phương pháp phân tích tình trạng hoạt động tấm pin quang điện bao gồm: Phương pháp phân tích dựa trên mô hình tấm pin (MBDM); Phương pháp phân tích thời gian thực (RDM); Phương pháp phân tích tín hiệu ngõ ra (OSA); Phương pháp phân tích dựa trên mô hình máy học (MLT)<sup>1</sup>. Các phương pháp thông dụng hơn cả là MBDM sẽ sử dụng thông số môi trường thu thập được gồm cường độ bức xạ mặt trời và nhiệt độ để đưa vào các mô hình tấm pin quang điện được xây dựng dựa trên cơ sở lý thuyết, sau đó so sánh ngõ ra của mô hình tấm pin quang điện với dữ liệu thông số hoạt động thu thập thực tế nhằm đưa ra các nhận định về tình trạng hoạt động. Quá trình so sánh có thể được thực hiện thông qua việc đánh giá sai lệch giữa các thông số điện áp, dòng điện, công suất của mô hình lý thuyết với giá trị thực tế như trong các nghiên cứu<sup>2-5</sup> hoặc đặt ngưỡng cho các tỉ lệ đặc trưng giữa kết quả mô hình và giá trị thực tế như ở

**Trích dẫn bài báo này:** Khang N P H, Nhân N C. **Phân loại sự cố mô-đun quang điện dựa trên điểm công suất cực đại và kỹ thuật máy học.** *Sci. Tech. Dev. J. - Nat. Sci.*; 2022, 6(4):2389-2400.

ngghiên cứu<sup>6</sup>. Khác với MBDM, trong phương pháp RDM, các thông số thu thập từ tấm pin quang điện được phân tích và đặt ngưỡng cảnh báo thông qua các quy luật và kiến thức được lập trình trước. Các quy luật về so sánh và đặt ngưỡng cảnh báo có thể vẫn được lấy từ kết quả phân tích mô hình tấm pin lý thuyết<sup>7,8</sup> hoặc từ các đặc trưng trong quá trình vận hành thực tế<sup>9,10</sup>. OSA là các phương pháp biến đổi các tín hiệu thông số tấm pin bằng các phép biến đổi toán học để trích xuất đặc trưng của tín hiệu và đưa ra đánh giá dựa trên những đặc trưng được phân tích. Các phương pháp biến đổi có thể sử dụng như phân phối chuẩn<sup>11</sup>, mô hình xác suất lượng tử<sup>12</sup>, biến đổi Wavelet<sup>13</sup>, ... Việc đánh giá phân tích tình trạng tấm pin sẽ được thực hiện thông qua việc so sánh và xét ngưỡng các giá trị đặc trưng của tín hiệu sau khi biến đổi với giá trị trong trạng thái ổn định. Các phương pháp MLT áp dụng các mô hình trí tuệ nhân tạo hay máy học để huấn luyện và học nhận diện các đặc trưng của quá trình hoạt động của tấm pin xuyên suốt các tình trạng ổn định và sự cố. Các thông số được sử dụng để đưa vào các thuật toán máy học nhằm nhận diện sự cố có thể là các đặc tính I-V của hệ thống khi vận hành<sup>14,15</sup> hay sử dụng ảnh hồng ngoại kiểm tra tình trạng nhiệt của tấm pin quang điện<sup>16</sup>. Việc huấn luyện mô hình được thực hiện thông qua dữ liệu thu thập thực tế hay kết quả của các phép tính mô phỏng mô hình tấm pin quang điện.

Từ các nghiên cứu trên, có thể thấy được rằng khi xảy ra sự cố, hệ thống sẽ mang những thông số hoạt động về điện áp và cường độ dòng điện đặc trưng cho loại sự cố xảy ra. Trong thực tế, các thông số hoạt động này tương ứng với điểm công suất cực đại trên đường đặc tuyến I-V tại điều kiện cường độ bức xạ hoạt động. Điều này được đảm bảo nhờ các thuật toán MPPT được tích hợp trong các bộ chuyển đổi và điều khiển sạc năng lượng mặt trời hiện có trên thị trường. Khi đưa những thông số này lên miền không gian tham số I-V, những trường hợp sự cố giống nhau thì sẽ có những điểm hoạt động nằm gần nhau tạo thành một nhóm. Từ đó, việc sử dụng các thuật toán phân cụm dữ liệu (clustering) trong kỹ thuật học máy để phân loại các trường hợp sự cố khác nhau với vùng tập trung khác nhau trên miền không gian tham số I-V là hoàn toàn khả thi. Trong đó, SVM và k-means là hai thuật toán máy học thông dụng nhất được sử dụng trong giải quyết các bài toán phân cụm dữ liệu (clustering).

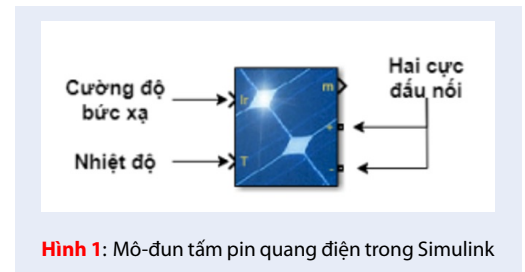
Trong bài báo này, các trường hợp sự cố của một dãy nhiều tấm pin quang điện sẽ được phân loại dựa vào việc huấn luyện các mô hình máy học sử dụng thuật toán SVM và k-means. Dữ liệu huấn luyện được thu thập từ chương trình mô phỏng gồm thông số

cường độ dòng điện, điện áp của dãy tấm pin trong các trường hợp sự cố khác nhau. Các dữ liệu này sau đó sẽ được phân tích và biến đổi để tiếp tục đưa vào mô hình phân loại sử dụng các thuật toán máy học được xây dựng trên nền tảng Google Colab bằng ngôn ngữ Python và các thư viện hỗ trợ.

## PHƯƠNG PHÁP NGHIÊN CỨU

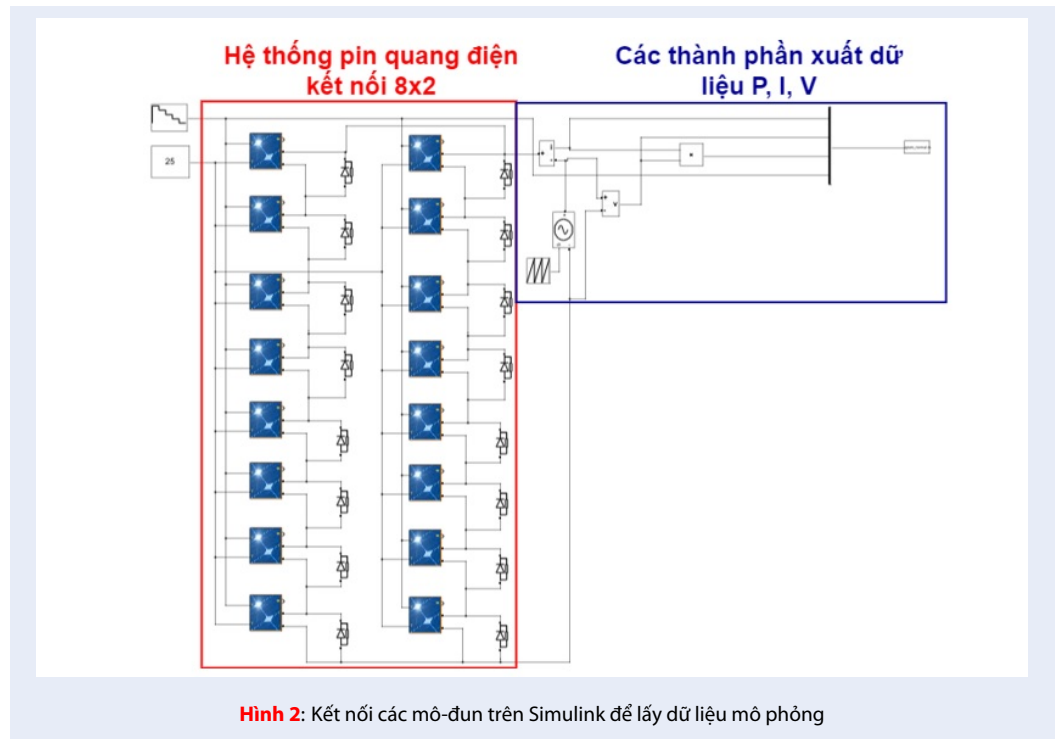
### MÔ PHỎNG CÁC SỰ CỐ CỦA MÔ-ĐUN QUANG ĐIỆN

Để tạo dữ liệu về các sự cố của mô-đun quang điện, một hệ thống tấm pin quang điện kết nối với cấu trúc 8x2 (2 hàng, 8 tấm nối tiếp) được mô phỏng trên phần mềm MATLAB. Mô-đun mô phỏng tấm pin quang điện trong công cụ Simulink của phần mềm MATLAB được trình bày trong Hình 1. Mô-đun có hai ngõ vào để đưa thông số cường độ bức xạ và nhiệt độ cell, hai ngõ ra là hai cực âm dương để đấu nối với các khối mô-đun khác trong hệ thống mô phỏng. Bảng 1 trình bày các thông số kỹ thuật của tấm pin quang điện (SunPower SPR-E20-245) mà được sử dụng trong mô phỏng. Để có thể thu được thông số cường độ dòng điện và điện áp của hệ thống, mô-đun tấm pin quang điện sẽ được kết nối với các khối chức năng đo đặc tương ứng được cung cấp bởi Simulink. Hình 2 trình bày kết nối hệ thống trong Simulink/MATLAB gồm các mô-đun tấm pin được kết nối thành một hệ 8x2 và các khối chức năng thu thập và xuất dữ liệu về điện áp, cường độ dòng điện và công suất của hệ thống.



Hình 1: Mô-đun tấm pin quang điện trong Simulink

Với hai trường hợp sự cố ngắn mạch và hở mạch, hệ thống được mô phỏng lần lượt ở điều kiện điện áp bằng không và cường độ dòng điện bằng không ở các điều kiện cường độ bức xạ trên. Trong trường hợp hở mạch một dãy, hệ thống coi như chỉ bao gồm 8 tấm pin nối tiếp và loại bỏ một dãy bị hở mạch khỏi hệ thống. Các sự cố còn lại sẽ được thêm vào hệ thống mô phỏng bằng cách thêm những kết nối hoặc các thông số bất thường. Các sự cố được khảo sát với số lượng tấm pin gặp sự cố khác nhau trong các điều kiện cường độ bức xạ khảo sát gồm 200, 400, 600, 800, 1000 ( $W/m^2$ ).



**Bảng 1:** Thông số tấm pin sử dụng trong mô phỏng

Thông số kỹ thuật	Giá trị
Công suất định danh	245W
Điện áp hở mạch	48,8V
Dòng điện ngắn mạch	6,43A
Điện áp điểm công suất lớn nhất*	40,5
Dòng điện điểm công suất lớn nhất*	6,05
Hệ số thay đổi dòng điện	0,07 (%/°C)
Số cell nối tiếp	72

\* Trong điều kiện STC – 1000W/m2, 25°C

### THUẬT TOÁN K-MEANS

Thuật toán k-means thực hiện phân nhóm dữ liệu bằng cách tìm điểm trọng tâm của nhóm dữ liệu sao cho tổng khoảng cách từ các điểm trọng tâm tới các điểm trong nhóm dữ liệu tương ứng là nhỏ nhất. Việc tính toán điểm trọng tâm của thuật toán k-means được thực hiện theo các bước<sup>17</sup>:

Bước 1: Khởi tạo các giá trị trọng tâm ngẫu nhiên

Bước 2: Xác định phân nhóm dựa trên khoảng cách của các điểm dữ liệu tới điểm trọng tâm gần nhất.

Bước 3: Tính toán lại điểm trọng tâm mới theo công thức

$$c_k = \frac{\sum_{d_i \in S_k} d_i}{S_k} \quad (1)$$

Với:

$c_k$ : điểm trọng tâm của phân nhóm thứ k

$d_i$ : các điểm trong tập dữ liệu được xét

$S_k$ : số lượng các điểm được phân vào nhóm thứ k

Bước 4: Lặp lại bước 2 và bước 3 cho tới khi giá trị điểm trọng tâm không thay đổi

Sau khi huấn luyện, khi thực hiện phân nhóm một điểm mới, mô hình sẽ phân điểm mới này vào nhóm mà khoảng cách từ điểm tới điểm trọng tâm của nhóm là ngắn nhất.

### THUẬT TOÁN SVM

Các bài toán phân loại có thể được tổng quát hóa thành tìm đường phân cách các nhóm dữ liệu trên không gian vector. Tuy nhiên, để phân tách nhóm dữ liệu thì có thể sử dụng rất nhiều đường phân cách khác nhau. Vấn đề đặt ra là đường phân cách nào thì tốt nhất để khi đưa một điểm mới vào thì mô hình vẫn thực hiện phân loại chính xác. Thuật toán SVM giải quyết vấn đề này bằng cách lựa chọn đường phân cách sao cho khoảng dự trữ giữa hai nhóm dữ liệu là lớn nhất. Đường phân cách có thể được miêu tả như phương trình với n là số chiều không gian tham số. Việc phân chia nhóm dữ liệu phụ thuộc vào dấu của phương trình khi thế các giá trị tham số. Khoảng dự trữ được định nghĩa là khoảng cách gần nhất từ một điểm trong nhóm dữ liệu tới đường phân cách. Khoảng cách này được định nghĩa như phương trình.

Nhiệm vụ của thuật toán SVM là tìm bộ số sao cho phương trình này lớn nhất.

$$\omega^T x + b = \sum_{i=0}^n \omega_i x_i + b \quad (2)$$

$$\text{margin} = \min \frac{y_i (\omega^T x_i + b)}{\|\omega\|_2} \quad (3)$$

Với:

$$y_i = \text{sign}(\omega^T x_i + b) \quad (4)$$

$$\|\omega\|_2 = \sqrt{\sum_{i=1}^n \omega_i^2} \quad (5)$$

Về mặt toán học, việc giải quyết bài toán này là rất phức tạp<sup>18</sup>. Tuy nhiên, nhờ sự phát triển của các thư viện hỗ trợ máy học sử dụng ngôn ngữ Python như scikit-learn<sup>19</sup>, việc nghiên cứu các ứng dụng của thuật toán SVM được giảm tải các vấn đề thiên hướng toán học và có thể tập trung hơn vào hiệu quả của thuật toán.

## KẾT QUẢ VÀ THẢO LUẬN

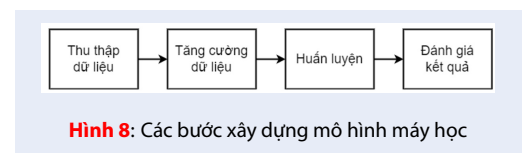
### MÔ PHỎNG CÁC TRƯỜNG HỢP SỰ CỐ

Hình 3 và Hình 4 thể hiện đường đặc tuyến I-V của hệ thống trong các trường hợp sự cố dây – dây và bóng che ở điều kiện cường độ bức xạ 1000 W/m<sup>2</sup> với số lượng n tấm pin gặp sự cố lần lượt là 1, 6. Khi số lượng tấm pin gặp sự cố nhỏ với n = 1, sự thay đổi ở đường I-V gần như không đáng kể ở trường hợp sự cố dây – dây. Sự thay đổi diễn ra rõ hơn khi vị trí lỗi bao hàm nhiều tấm pin hơn, khi đó ta nhìn thấy rõ sự tổn hao điện áp so với trạng thái bình thường trong hệ thống. Trong trường hợp sự cố bóng che, khi số lượng tấm pin gặp sự cố nhỏ, sự thay đổi điểm MPP khác nhau với các cường độ bức xạ bóng che khác nhau. Khi cường độ bức xạ bóng che gần với cường độ bức xạ hiện trạng (800 W/m<sup>2</sup> so với 1000 W/m<sup>2</sup>), đường I-V ít bị biến dạng và điểm MPP bị mất một lượng cường độ dòng điện. Ngược lại, khi cường độ bức xạ bóng che chênh lệch lớn (400 W/m<sup>2</sup> so với 1000 W/m<sup>2</sup>), đường I-V bị biến dạng nhiều hơn nhưng điểm MPP lại dịch chuyển theo hướng giảm điện áp. Khi số lượng tấm pin bị bóng che gia tăng, điểm MPP trong trường hợp sự cố này có xu hướng giảm cường độ dòng điện và duy trì mức điện áp, độ chênh lệch giữa cường độ bóng che và cường độ bức xạ càng lớn, cường độ dòng điện càng giảm nhiều. Sự thay đổi về đường I-V và điểm MPP trong các trường hợp sự cố dây – dây và bóng che ở các cường độ bức xạ khác nhau được trình bày trong Hình 5 và Hình 6. Các trường hợp sự cố ngắn mạch và hở mạch do cố định một trong hai thông số điện áp hay cường độ dòng điện nên không có dòng I-V mà chỉ có điểm làm

việc cố định, các điểm này được biểu diễn trong phần tổng hợp toàn bộ các trường hợp sự cố ở Hình 7.

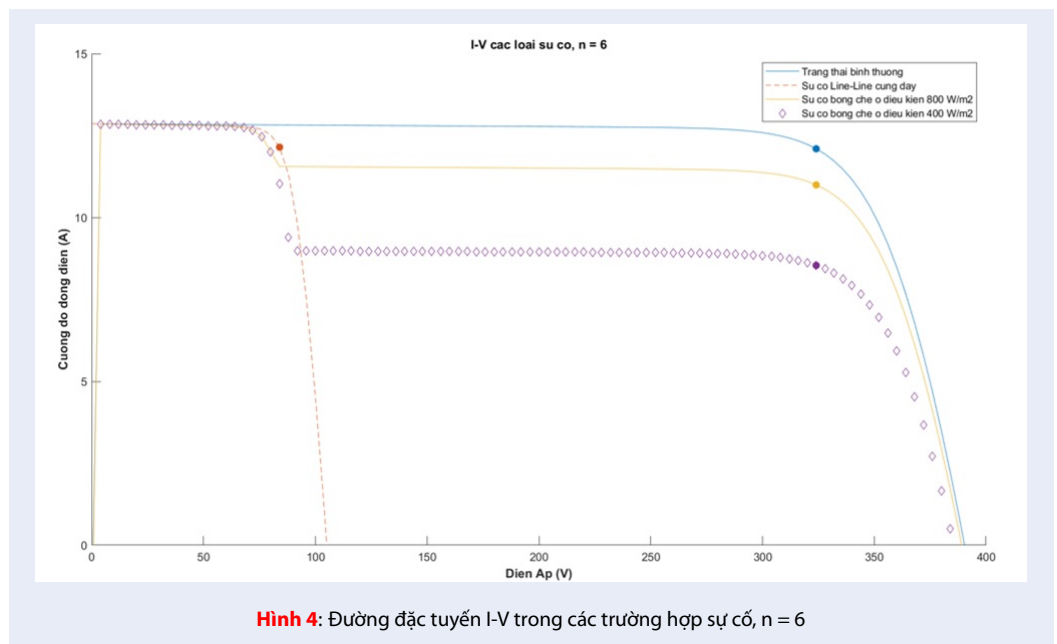
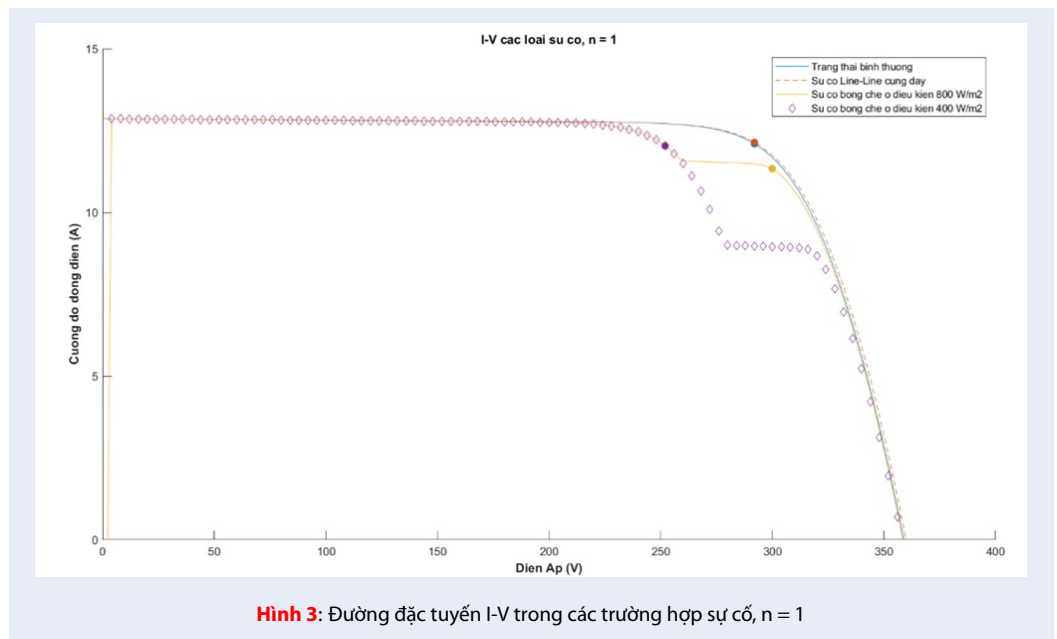
Như Hình 7(a), mặc dù các điểm MPP của các trường hợp sự cố có nằm ở những vùng khác nhau trên miền không gian I-V đối với từng loại sự cố, các điểm MPP vẫn nằm rải rác trên một vùng rộng và phải nhận diện theo giá trị cường độ bức xạ tại thời điểm đang xét. Các điểm sự cố ngắn mạch do điện áp bằng không nên nằm bám sát trục tung của đồ thị, còn các điểm sự cố hở mạch có cường độ dòng điện bằng không nằm liền kề nhau ở trục hoành ở các vị trí điện áp hở mạch. Các trường hợp sự cố hở mạch một dây do mất một dây hoạt động so với bình thường nên điểm làm việc có mức điện áp chỉ bằng một nửa so với bình thường. Các trường hợp sự cố dây – dây và bóng che có sự thay đổi đã được đề cập ở trên nằm phân bố ở những vùng khác nhau nhưng có trùng lặp ở một vài điểm do khi cường độ bức xạ ở vị trí bóng che chênh lệch nhiều thì diode bảo vệ trên tấm pin sẽ bỏ qua tấm pin đó và tương đương với trường hợp sự cố dây – dây trên cùng một dây. Tuy nhiên, khi sử dụng giá trị điện áp và cường độ dòng điện của điểm làm việc bình thường làm mốc chuẩn hóa, các điểm trong Hình 7(a) chuyển sang như Hình 7(b), loại bỏ phụ thuộc vào giá trị cường độ bức xạ cũng như các điểm của từng trường hợp sự cố nằm phân tách rõ ràng hơn trong miền không gian tham số I-V. Dựa vào sự phân tách theo đặc tính I-V này của trường hợp sự cố, khả năng phân loại của các thuật toán máy học SVM và k-means sẽ được khảo sát.

### MÔ HÌNH MÁY HỌC PHÂN LOẠI SỰ CỐ MÔ-ĐUN QUANG ĐIỆN



Hình 8: Các bước xây dựng mô hình máy học

Các bước xây dựng mô hình máy học được trình bày như trong Hình 8. Quá trình thu thập dữ liệu được thực hiện thông qua mô phỏng như đã trình bày ở phần trước. Việc tăng cường dữ liệu được thực hiện gồm: gia tăng số lượng điểm của từng trường hợp sự cố và thêm một trường hợp đại diện cho các trường hợp sự cố chưa được xét tới. Việc gia tăng số lượng điểm của từng trường hợp sự cố được thực hiện bằng cách thêm một sai số ngẫu nhiên đối với tham số cường độ dòng điện hoặc điện áp của các điểm ban đầu thu thập từ mô phỏng. Các sai số được thêm vào này đại diện cho hai trường hợp: một là thay thế cho



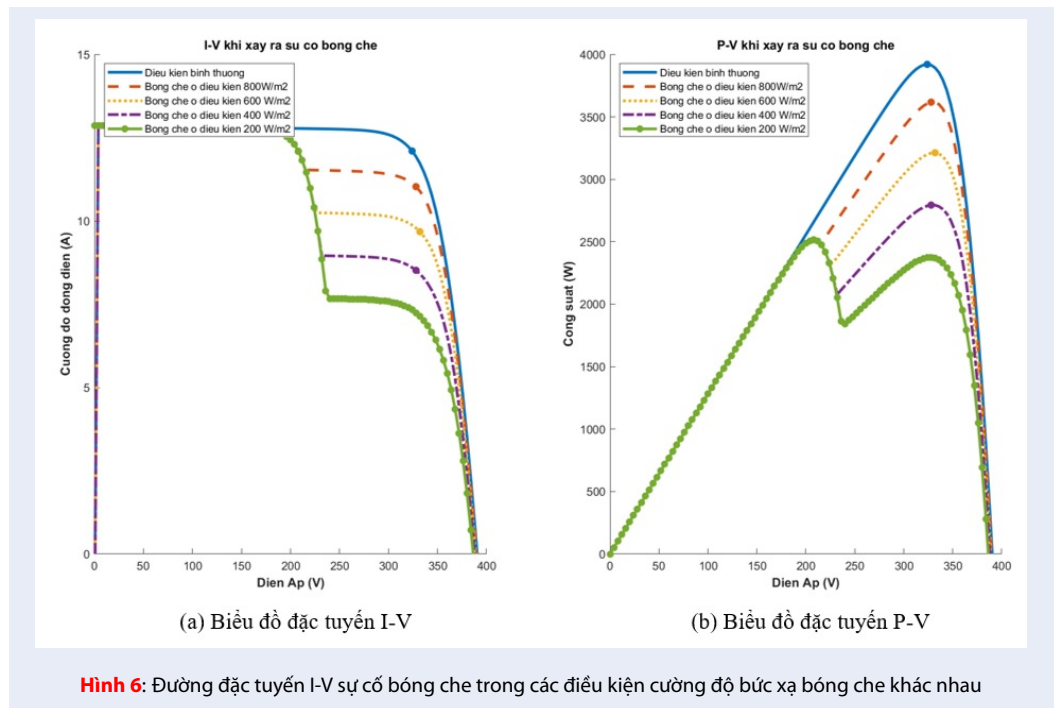
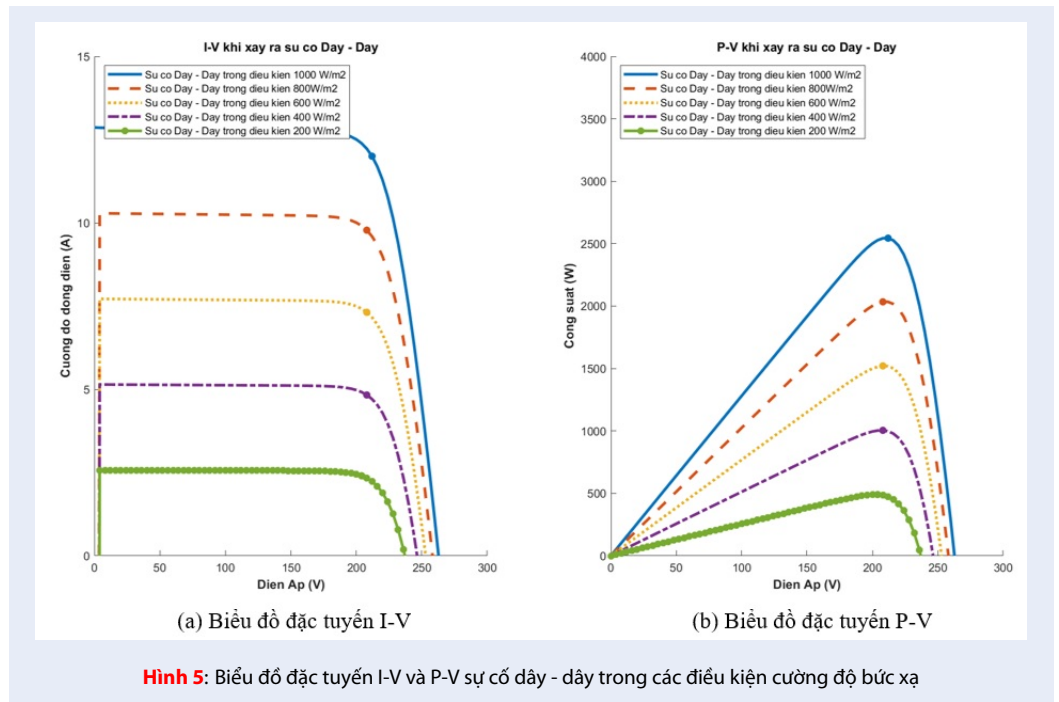
các điểm sự cố tương tự ở các điều kiện cường độ bức xạ lân cận với các mức đã được xét tới, hai là các sai số có khả năng xảy ra từ quá trình thu thập dữ liệu trong thực tế hoặc từ các kết quả tính toán của mô hình tấm pin khi ứng dụng mô hình phân loại sự cố trong các phương pháp MBDM hay RDM. Mỗi điểm sự cố ban đầu sẽ được thay đổi giá trị cường độ dòng điện hay điện áp với một giá trị ngẫu nhiên trong khoảng  $[0, \epsilon]$  theo phương trình (6) và (7) với  $\epsilon$  là mức sai số chấp nhận để tạo thêm nhiều điểm sự cố mới. Bước tăng

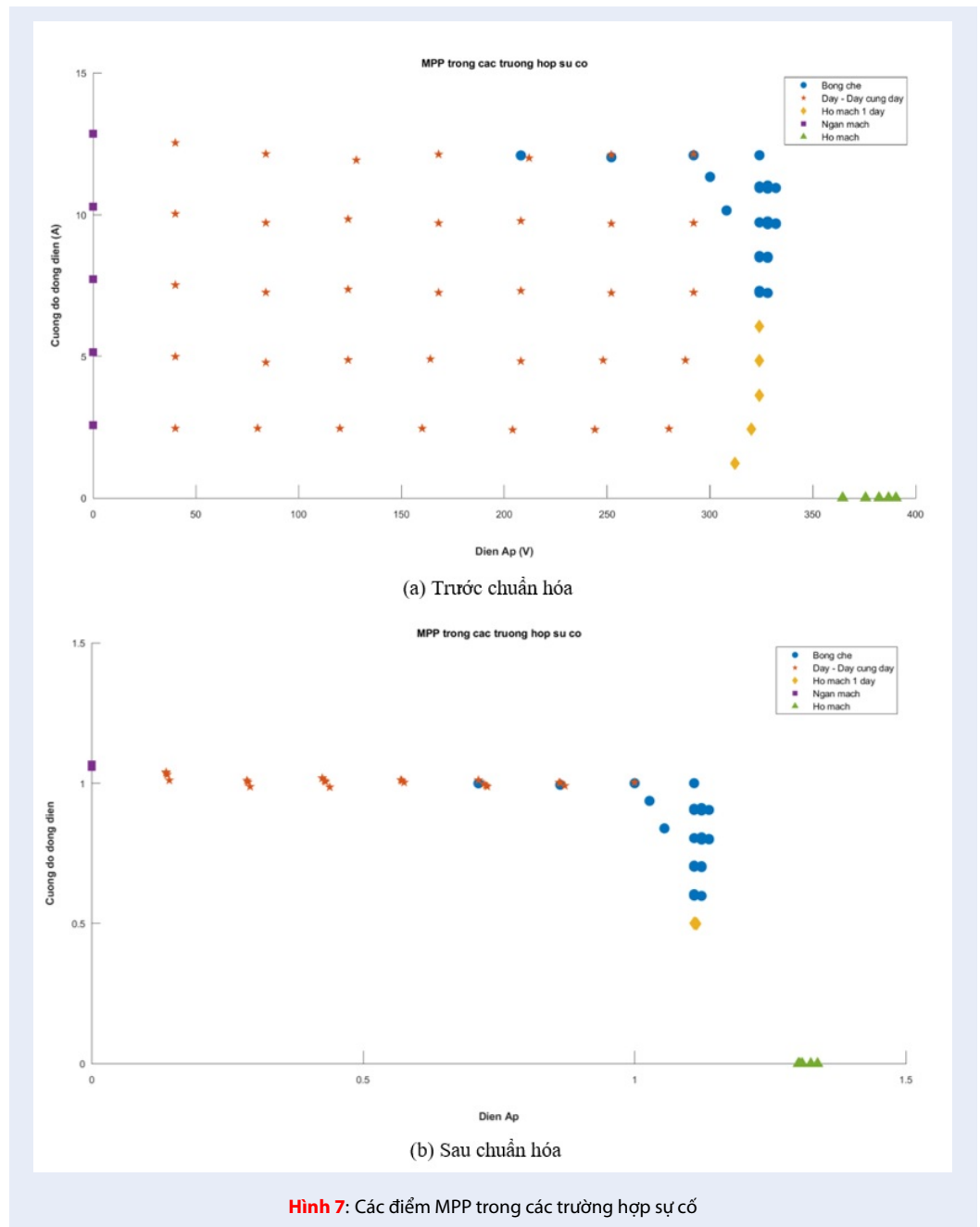
cường dữ liệu thứ hai là bổ sung các điểm vào khoảng trống trong khoảng  $0 - 0,9$  trong miền I-V chuẩn hóa để thể hiện cho các trường hợp sự cố khác chưa được đề cập tới trong bài báo này. Thông tin về số lượng cũng như các ký hiệu của các trường hợp sự cố trong bộ dữ liệu huấn luyện được trình bày trong Bảng 2.

$$v = V \pm rand([0, \epsilon]) \tag{6}$$

$$i = I \pm rand([0, \epsilon]) \tag{7}$$







Hai mô hình phân loại ứng dụng thuật toán máy học k-means và SVM được xây dựng trên nền tảng Google Colab bằng ngôn ngữ Python và sử dụng các thư viện hỗ trợ nghiên cứu các thuật toán máy học scikit-learn. Thư viện cung cấp các công cụ để xây dựng mô hình, huấn luyện và trình bày kết quả sau huấn luyện để đánh giá hiệu suất phân loại của mô hình.

### KẾT QUẢ MÔ HÌNH MÁY HỌC DỰA TRÊN THUẬT TOÁN K-MEANS

Hình 9 trình bày các kết quả phân loại của mô hình máy học dựa trên thuật toán k-means. Hình 9(a) trình bày bộ dữ liệu thu được sau khi tăng cường và Hình 9(b) thể hiện kết quả phân loại sau khi huấn luyện. Kết quả huấn luyện cho thấy rằng thuật toán k-means phân loại các trường hợp sự cố huấn luyện không tốt, việc phân loại các sự cố lẫn nhau vào vùng các điểm sự cố khác cũng như có nhiều nhầm lẫn

**Bảng 2:** Thông tin bộ dữ liệu huấn luyện

CÁC SỰ CỐ CỦA PV MÔ-ĐUN	SỐ LƯỢNG MẪU	KÝ HIỆU SỰ CỐ	NHÂN SỰ CỐ
Bóng che	2000	F1	0
Hở mạch một dây	500	F2	1
Dây – Dây	3500	F3	2
Ngắn mạch	1000	F4	3
Hở mạch	1000	F5	4
Các lỗi khác	1500	Others	5

giữa các sự cố khác nhau với nhau. Hình 9(c) trình bày kết quả dự đoán của mô hình k-means trên miền không gian tham số I-V đã chuẩn hóa, qua đó thấy được đường phân tách các trường hợp sự cố nằm xa so với vị trí phân tách thật sự dẫn tới sự nhầm lẫn khi phân loại. Hình 9(d) trình bày ma trận phân loại sự cố của mô hình máy học dựa trên thuật toán k-means. Độ chính xác tổng quan của mô hình chỉ đạt 61,1%. Hai trường hợp mà mô hình dự đoán chính xác cao nhất là sự cố ngắn mạch và hở mạch (với độ chính xác là 100%), tuy nhiên đối với sự cố dây – dây và hở mạch một dây thì mô hình dự đoán nhầm lẫn giữa hai trường hợp này. Các trường hợp sự cố hở mạch một dây cũng có độ chính xác khá cao với 84,20%, nhưng đồng thời có tới 48,50% các sự cố bóng che bị dự đoán nhầm sang trường hợp sự cố hở mạch một dây. Các trường hợp sự cố còn lại đều có độ chính xác thấp chỉ đạt trong khoảng 40-60%.

Lý do thuật toán k-mean có kết quả cao trong phân loại tình trạng ngắn mạch và hở mạch so với hai trường hợp sự cố còn lại có thể giải thích bởi việc phân bố các điểm hoạt động trong hai trường hợp sự cố này tập trung trong một vùng hẹp trên miền không gian I-V đã chuẩn hóa (tập trung quanh điểm [0 ; 1.1] trong trường hợp ngắn mạch và [1.2 ; 0] trong trường hợp hở mạch). Các trường hợp khác thì các điểm hoạt động trải dài trên một vùng rộng hơn dẫn đến việc tính toán điểm trọng tâm bao phủ toàn bộ các điểm dữ liệu là không thể.

### KẾT QUẢ MÔ HÌNH MÁY HỌC DỰA TRÊN THUẬT TOÁN SVM

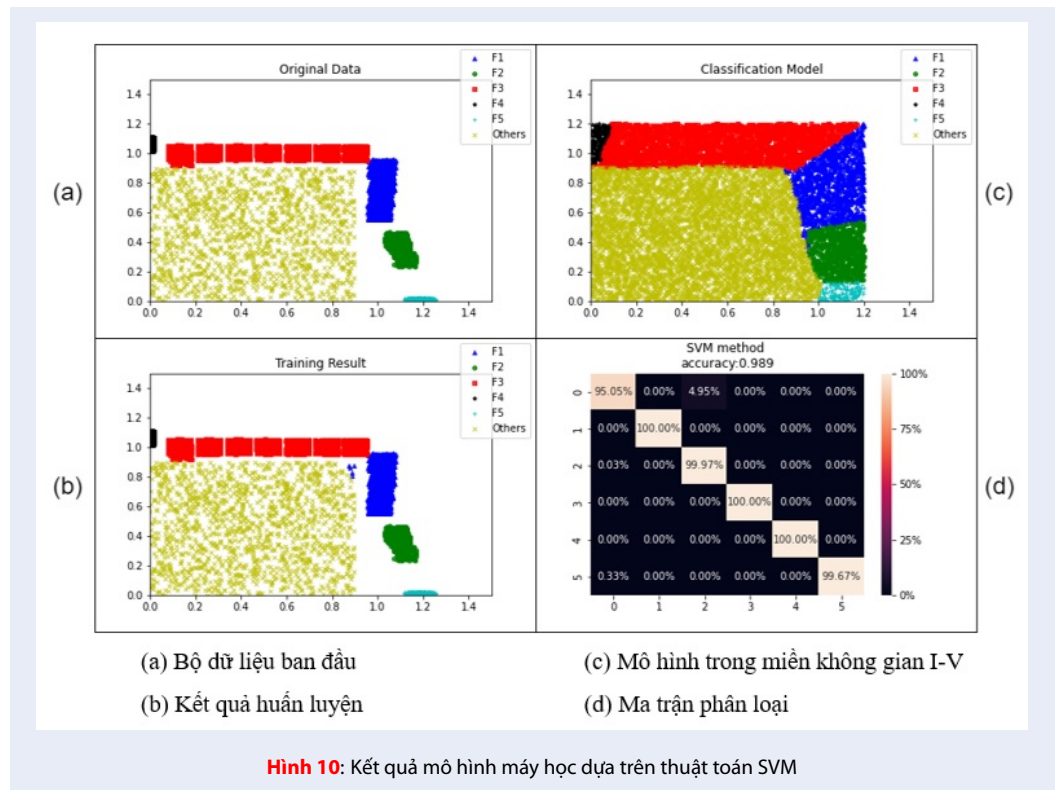
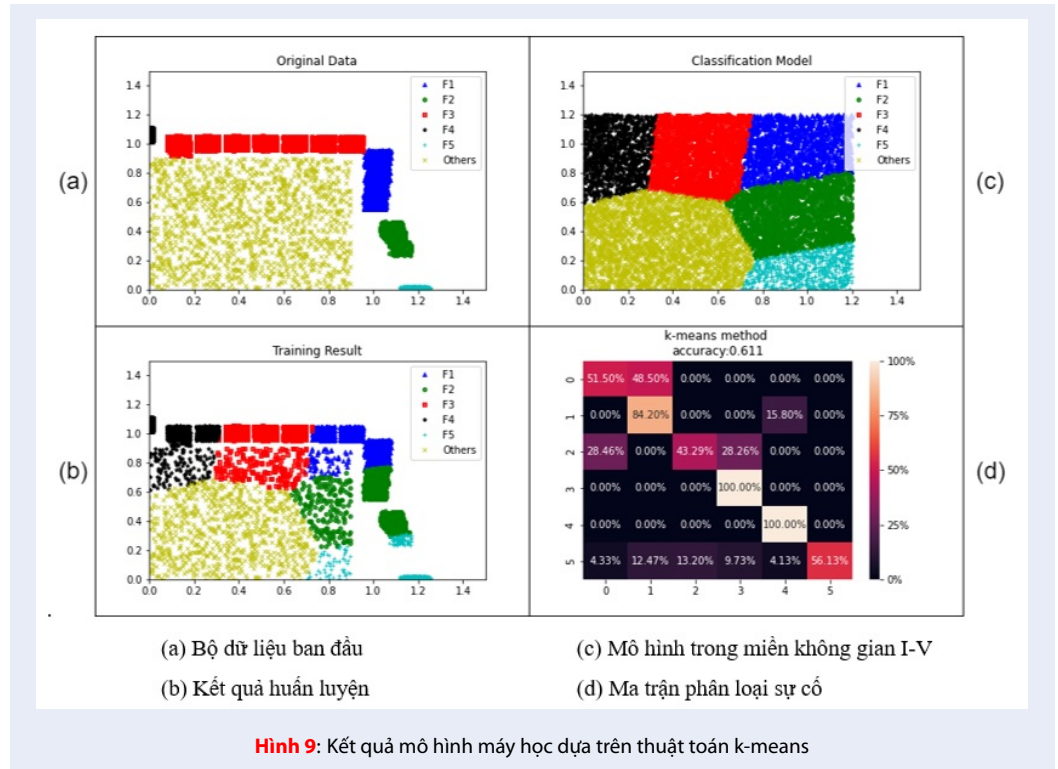
Tương tự, Hình 10 trình bày các kết quả của mô hình máy học dựa trên thuật toán SVM. Hình 10(a) trình bày bộ dữ liệu ban đầu và Hình 10(b) kết quả huấn luyện của mô hình. Kết quả cho thấy rằng mô hình phân loại ứng dụng thuật toán SVM cho ra kết quả tương tự và bám sát bộ dữ liệu ban đầu. Đồng thời, Hình 10(c) trình bày kết quả phân loại của mô hình

trên miền không gian, kết quả này cho thấy đường phân tách giữa các nhóm trường hợp sự cố bám sát với các điểm dữ liệu biên trong tập dữ liệu, do đó tạo ra ngưỡng phân loại các trường hợp sự cố tốt hơn so với thuật toán k-means. Hình 10(d) trình bày ma trận phân loại sự cố dựa trên thuật toán SVM, với độ chính xác tổng quan cả mô hình đạt 98,9%. Kết quả phân loại của tất cả các trường hợp sự cố đều đạt trên 95% và hai trường hợp sự cố ngắn mạch và hở mạch đạt 100%.

Bảng 3 trình bày việc so sánh độ chính xác và thông số f1-score theo loại sự cố của hai mô hình máy học dựa trên thuật toán k-means và SVM. Kết quả cho thấy việc phân loại theo từng loại sự cố và độ chính xác của mô hình máy học dựa trên thuật toán SVM (98,9%) cao hơn so với thuật toán k-means (61,1%). Sự vượt trội này xảy ra do đặc tính phân bố trên miền tham số I-V sau chuẩn hóa của các trường hợp sự cố không nằm trong một vùng tỏa rộng mà thiên hướng tập trung vào những vùng có dạng chữ nhật. Đặc tính tạo đường phân nhóm sao cho mô hình có biên độ dữ trữ lớn nhất của thuật toán SVM tạo khả năng phân nhóm tốt với các trường hợp sự cố khác nhau cũng như hứa hẹn tránh sự thiên vị một nhóm sự cố so với việc đặt ngưỡng thủ công qua phân tích bằng mắt thông thường.

Bảng 4 trình bày kết quả nghiên cứu của bài báo với kết quả của các nghiên cứu trước đây<sup>14,15</sup>. Phương pháp 8paras + SVM<sup>14</sup> sử dụng 8 thông số được cung cấp bởi nhà sản xuất của tấm pin quang điện để tái tạo đường đặc tuyến I-V và phân tích, so sánh với đường I-V thu được từ tấm pin quang điện để đưa ra kết quả phân loại sự cố qua phương pháp máy học SVM. Mặt khác, phương pháp GADF + ANN<sup>14</sup> lại sử dụng mạng nơ-ron nhân tạo với ngõ vào là biến đổi Gramian Angular để cho ra nhận định loại sự cố. Phương pháp SVM-GRB<sup>15</sup> sử dụng mô hình SVM với kernel GRB để đưa ra kết quả phân loại từ các thông số hoạt động của tấm pin gồm cường độ bức xạ, nhiệt độ, cường độ dòng điện và điện áp.





**Bảng 3:** So sánh kết quả hai mô hình máy học dựa trên thuật toán k-means và SVM

		k-means (%)	SVM (%)
Độ chính xác của mô hình		61,1	98,9
Độ chính xác của từng trường hợp sự cố	Bóng che	50	97
	Hở mạch một dây	41	100
	Dây-dây	58	99
	Ngắn mạch	64	100
	Hở mạch	93	100
	Các lỗi khác	72	100

**Bảng 4:** So sánh kết quả mô phỏng của phương pháp đề xuất với các công bố trước đây

Các phương pháp nghiên cứu	Giá trị ngõ vào	Các loại sự cố	Độ chính xác
Phương pháp đề xuất (SVM)	G, I, V, Impp, Vmpp	Bóng che, hở mạch một dây, dây – dây, hở mạch, ngắn mạch, các sự cố khác.	98,90%
8paras + SVM <sup>14</sup>	G, T, Vmpp, Impp, Voc, Isc, Rs, Rsh	Bóng che, suy giảm do Rs, suy giảm do Rsh	94,83%
GADF + ANN <sup>14</sup>	Gramian Angular Difference Field	Bóng che, suy giảm do Rs, suy giảm do Rsh	100,00%
SVM-GRB <sup>15</sup>	G, T, VPV, and IPV	Chập điện, dây – dây, hở mạch, bóng che, MPPT Failure	89,80%

Thông qua việc sử dụng giá trị tính toán của điểm công suất lớn nhất MPP từ mô hình tương đương tấm pin, phương pháp đề xuất đã gia tăng độ chính xác của việc phân loại sự cố sử dụng mô hình SVM từ 94.83% và 89.80% lần lượt của phương pháp 8paras và SVM-GRB lên 98.9%. Mặc dù kết quả của mô hình đề xuất không cao được so với phương pháp GADF + ANN, nhưng phương pháp thực hiện lại đơn giản hơn do không cần thu thập dữ liệu của đường đặc tuyến I-V cũng như biến đổi toán học phức tạp.

## KẾT LUẬN

Trong bài báo này, mô hình máy học phân loại sự cố mô-đun quang điện đã được xây dựng dựa trên đặc tuyến I-V của pin quang điện, điểm công suất cực đại (MPP) và hai thuật toán máy học phân nhóm (k-means và SVM). Mô hình máy học này có thể phân loại được các trường hợp sự cố như: sự cố dây – dây, bóng che, hở mạch một phần, hở mạch toàn phần, ngắn mạch trong các điều kiện cường độ bức xạ khác nhau. Bộ dữ liệu đầu vào của mô hình máy học được xây dựng dựa trên các mô phỏng trên phần mềm Simulink/MATLAB. Kết quả cho thấy mô hình máy học phân loại sự cố của mô-đun quang điện dựa trên thuật toán SVM có độ chính xác cao hơn thuật toán k-means. Các mô hình phân loại ứng dụng các thuật

toán máy học này hứa hẹn sẽ hỗ trợ phát triển với các thuật toán giám sát và đánh giá tình trạng tấm pin hiện đang ứng dụng trong các mô hình MBDM hay RDM, qua đó giúp quá trình vận hành các hệ thống NLMT dùng tấm pin quang điện diễn ra hiệu quả và gia tăng tốc độ bảo trì bảo dưỡng khi có sự cố xảy ra.

## LỜI CẢM ƠN

Nghiên cứu được tài trợ bởi Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM trong khuôn khổ Đề tài mã số T2021-06.

## DANH MỤC TỪ VIẾT TẮT

- PV (Photovoltaic): Quang điện
- MPP (Max Power Point): Điểm làm việc công suất cực đại
- SVM (Support Vector Machine): máy véc-tơ hỗ trợ
- MPPT(Max Power Point Tracking): Thuật toán dò điểm công suất lớn nhất
- NLMT: Năng lượng mặt trời
- MBDM (Model Based Difference Measurement): Phương pháp phân tích dựa trên mô hình tấm pin
- RDM (Real-time Difference Measurement): Phương pháp phân tích thời gian thực

- OSA (Output Signal Analysis): Phương pháp phân tích tín hiệu ngõ ra
- MLT (Machine Learning Technique): Phương pháp phân tích dựa trên mô hình máy học
- STC (Standard Testing Condition): Điều kiện thí nghiệm tiêu chuẩn

## XUNG ĐỘT LỢI ÍCH

Các tác giả cam kết không có bất kỳ xung đột lợi ích nào trong công bố bài báo.

## ĐÓNG GÓP CỦA TÁC GIẢ

- Nguyễn Phước Hoàng Khang: Xây dựng chương trình mô phỏng thu thập, tăng cường dữ liệu, tổng hợp số liệu hình ảnh.
- Nguyễn Chí Nhân: Xây dựng chương trình thuật toán máy học, phân tích đánh giá kết quả.

## TÀI LIỆU THAM KHẢO

1. Pillai DS, Blaabjerg F, Rajasekar N. A comparative evaluation of advanced fault detection approaches for PV systems. *IEEE J Photovoltaics*. 2019;9(2):513-27; Available from: <https://doi.org/10.1109/JPHOTOV.2019.2892189>.
2. Davarifar M, Rabhi A, El Hajjaji AE, Dahmane M. New method for fault detection of PV panels in domestic applications. In: 3rd International Conference on Systems and Control; 2013. p. 727-32; Available from: <https://doi.org/10.1109/ICoSC.2013.6750940>.
3. Tadj M, Benmouiza K, Chekneane A, Silvestre S. Improving the performance of PV systems by faults detection using GISTEL approach. *Energy Convers Manag*. 2014;80:298-304; Available from: <https://doi.org/10.1016/j.enconman.2014.01.030>.
4. Chine W, Mellit A, Pavan AM, Lughfi V. Fault diagnosis in photovoltaic arrays. In: International Conference on Clean Electrical Power (ICCEP). Vol. 2015; 2015. p. 67-72; Available from: <https://doi.org/10.1109/ICCEP.2015.7177602>.
5. Harrou F, Sun Y, Taghezouit B, Saidi A, Hamlati M-E. Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. *Renew Energy*. 2018;116:22-37; Available from: <https://doi.org/10.1016/j.renene.2017.09.048>.
6. Platon R, Martel J, Woodruff N, Chau TY. Online fault detection in PV systems. *IEEE Trans Sustain Energy*. 2015;6(4):1200-7; Available from: <https://doi.org/10.1109/TSTE.2015.2421447>.
7. Garoudja E, Kara K, Chouder A, Silvestre S, Kichou S. Efficient fault detection and diagnosis procedure for photovoltaic systems; 2016; Available from: <https://doi.org/10.1109/ICMIC.2016.7804233>.
8. Xu X, Wang H, Xu X, Zuo Y. Method for diagnosing photovoltaic array fault in solar photovoltaic system. In: *Power A-P*, editor. Vol. 2011; 2011. p. 1-5; Available from: <https://doi.org/10.1109/APPEEC.2011.5747701>.
9. Silvestre S, Silva MAD, Chouder A, Guasch D, Karatepe E. New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. *Energy Convers Manag*. 2014;86:241-9; Available from: <https://doi.org/10.1016/j.enconman.2014.05.008>.
10. Alajmi M, Abdel-Qader I. Fault detection and localization in solar photovoltaic arrays using the current-voltage sensing framework; 2016. p. 307-12; Available from: <https://doi.org/10.1109/EIT.2016.7535257>.
11. Zhao Y, Lehman B, Ball R, Mosesian J, de Palma Jd. Outlier detection rules for fault detection in solar photovoltaic arrays. In: *Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*. Vol. 2013; 2013. p. 2913-20; Available from: <https://doi.org/10.1109/APEC.2013.6520712>.
12. Georgijevic NL, Jankovic MV, Srdic S, Radakovic Z. The detection of series arc fault in photovoltaic systems based on the arc current entropy. *IEEE Trans Power Electron*. 2016;31(8):5917-30; Available from: <https://doi.org/10.1109/TPEL.2015.2489759>.
13. Yi Z, Etemadi AH. Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems. *IEEE Trans Smart Grid*. 2017;8(3):1274-83; Available from: <https://doi.org/10.1109/TSG.2016.2587244>.
14. Li B, Delpha C, Migan-Dubois A, Diallo D. Fault diagnosis of photovoltaic panels using full I-V characteristics and machine learning techniques. *Energy Convers Manag*. 2021;248:114785; Available from: <https://doi.org/10.1016/j.enconman.2021.114785>.
15. Badr M, Hamad Ms, Abdel-Khalik A, Hamdy R, Ahmed S, Hamdan E. Fault identification of photovoltaic array based on machine learning classifiers. *IEEE Access*. 2021;9:1-20; Available from: <https://doi.org/10.1109/ACCESS.2021.3130889>.
16. Kurukuru VSB, Haque A, Khan MA, Tripathy AK. Fault classification for Photovoltaic Modules Using thermography and Machine Learning Techniques. In: *International Conference on Computer and Information Sciences (ICCI)*. Vol. 2019; 2019. p. 1-6; Available from: <https://doi.org/10.1109/ICCI.2019.8716442>.
17. Khan SS, Ahmad A. Cluster center initialization algorithm for K-means clustering. *Pattern Recognit Lett*. 2004;25(11):1293-302; Available from: <https://doi.org/10.1016/j.patrec.2004.04.007>.
18. Widodo A, Yang B-S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech Syst Signal Process*. 2007;21(6):2560-74; Available from: <https://doi.org/10.1016/j.ymsp.2006.12.007>.
19. scikit-learn Machine Learning in Python; Available from: <https://scikit-learn.org/stable/>.

# Fault classification for photovoltaic module based on maximum power point and machine learning techniques

Nguyen Phuoc Hoang Khang<sup>1</sup>, Nguyen Chi Nhan<sup>1,2,\*</sup>



Use your smartphone to scan this QR code and download this article

## ABSTRACT

Photovoltaic (PV) module is the key component in the solar energy system. Fault classification for photovoltaic module is necessary for safety, efficiency and reliability of photovoltaic systems. When faults of PV module occurred, the current-voltage characteristics (I-V curves) of PV module are changed leads to the shift in voltage and current of the maximum power point (MPP). In this paper, fault classification for photovoltaic module based on MPP and machine learning techniques is proposed. Through analysing the location of the MPP in I-V parameters plane, faults of PV module during performance of could be identified and classified. The parameters of voltage and current of PV module in the faults (line-to-line fault, shading, open circuit in one array, open circuit, short circuit during different irradiation levels) are collected by simulations in Simulink/MATLAB software. Then, these parameters would be processed and built into a dataset. This dataset would be fed into fault classification model for PV module using machine learning algorithms included Support Vector Machine (SVM) and k-means. The desired results of machine learning models would be accurately classified faults using parameters of voltage and current of the MPP and survey the fault diagnosis through the borderline of each classification group of the models. The result shows that the fault classification capability and accuracy of the machine learning model using SVM algorithm (98,9%) are better than the k-means algorithm (61,1%).

**Key words:** photovoltaic panel, fault diagnosis, I-V characteristic, SVM, k-means, Machine Learning

<sup>1</sup>Faculty of Physics and Engineering Physics, University of Science, VNU-HCM

<sup>2</sup>Integrated Circuits Design Laboratory, University of Science, VNU-HCM

## Correspondence

**Nguyen Chi Nhan**, Faculty of Physics and Engineering Physics, University of Science, VNU-HCM

Integrated Circuits Design Laboratory, University of Science, VNU-HCM

Email: ncnhan@hcmus.edu.vn

## History

- Received: 12-8-2022
- Accepted: 20-12-2022
- Published: 31-12-2022

DOI : 10.32508/stdjns.v6i4.1221



## Copyright

© VNUHCM Press. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



**Cite this article :** Khang N P H, Nhan N C. **Fault classification for photovoltaic module based on maximum power point and machine learning techniques.** *Sci. Tech. Dev. J. - Nat. Sci.*; 2022, 6(4):2389-2400.