


# Các chương trình giám sát dựa trên kNN được cải tiến để phát hiện lỗi trong hệ thống PV

Fouzi Harrou , Thành viên, IEEE, Bilal Taghezouit và Ying Sun 

**Tóm tắt**– Bài báo này trình bày một phương pháp phát hiện dị thường dựa trên mô hình để giám sát phía dòng điện một chiều (dc) của các hệ thống quang điện (PV). Để đạt được mục đích này, một khuôn khổ kết hợp các lợi ích của k-nearest neighbors (kNN) với các phương pháp giám sát đơn biến đã được đề xuất. Cụ thể, các lược đồ Shewhart dựa trên kNN và trung bình động có trọng số theo hàm mũ (EWMA) với ngưỡng tham số và không tham số đã được giới thiệu để phát hiện lỗi trong các hệ thống PV một cách phù hợp. Việc lựa chọn phương pháp kNN để tách các đặc điểm bình thường và bất thường được thúc đẩy bởi khả năng xử lý các đặc điểm phi tuyến tính của nó và không đưa ra giả định về phân phối dữ liệu cơ bản. Ngoài ra, vì phương pháp EWMA nhạy trong việc phát hiện những thay đổi nhỏ. Đầu tiên, một mô hình mô phỏng cho mảng PV được kiểm tra để xây dựng. Sau đó, các phần dư được tạo ra từ mô hình này được sử dụng làm đầu vào cho các lược đồ dựa trên kNN để phát hiện dị thường.

Ngưỡng tham số và không tham số sử dụng ước tính mật độ hạt nhân đã được sử dụng để phát hiện lỗi. Hiệu quả của các quy trình dựa trên kNN được xác minh bằng cách sử dụng các phép đo thực tế từ hệ thống kết nối lưới điện 9,54 MW ở Algeria. Kết quả công bố hiệu quả của chiến lược đề xuất để giám sát phía dc của hệ thống PV.

Thuật ngữ chỉ mục –Phát hiện lỗi, k-láng giềng gần nhất (kNN), che bóng một phần, hệ thống quang điện (PV), biểu đồ giám sát thống kê.

## I. GIỚI THIỆU

Năng lượng tái tạo đang quay trở lại được quốc tế công nhận trên toàn thế giới (IEA), năng lực sản xuất năng lượng tái tạo trên thế giới đang tăng nhanh chóng hàng năm để ghi nhận mức tăng 178 GW vào năm 2017 [1]. Hơn nữa, các hệ thống quang điện mặt trời (PV) đã dẫn đầu cho các kỷ lục mới về tích hợp năng lượng tái tạo để sản xuất điện, đạt khoảng 98 GW. Với sự tiến bộ nhanh chóng trong công nghệ PV, PV mặt trời trở thành một

Bản thảo nhận ngày 27 tháng 11 năm 2018; sửa đổi ngày 13 tháng 1 năm 2019; chấp nhận ngày 25 tháng 1 năm 2019. Công trình này được hỗ trợ bởi Đại học Khoa học và Công nghệ Vua Abdullah (KAUST), Văn phòng Nghiên cứu được Tài trợ (OSR) theo Giải thưởng OSR-2015-CRG4-2582. (Tác giả liên hệ: Fouzi Harrou.)

F. Harrou và Y. Sun làm việc tại Phân khoa Khoa học Máy tính, Điện và Toán học và Kỹ thuật (CEMSE), Đại học Khoa học và Công nghệ Quốc vương Abdullah (KAUST), Thuwal 23955-6900, Ả Rập Xê Út (email: fouzi.harrou@kaust.edu.sa; ying.sun@kaust.edu.sa).

B. Taghezouit làm việc tại Trung tâm Phát triển Năng lượng tái tạo (CDER), Algiers 16340, Algeria, và Phòng thí nghiệm Thiết bị truyền thông và Chuyển đổi quang điện, Trường Bách khoa Quốc gia Algiers, Algiers 16200, Algeria (e-mail: b.taghezouit@gmail.com).

Phiên bản màu của một hoặc nhiều hình ảnh trong bài báo này có sẵn trực tuyến tại <http://ieeexplore.ieee.org>.

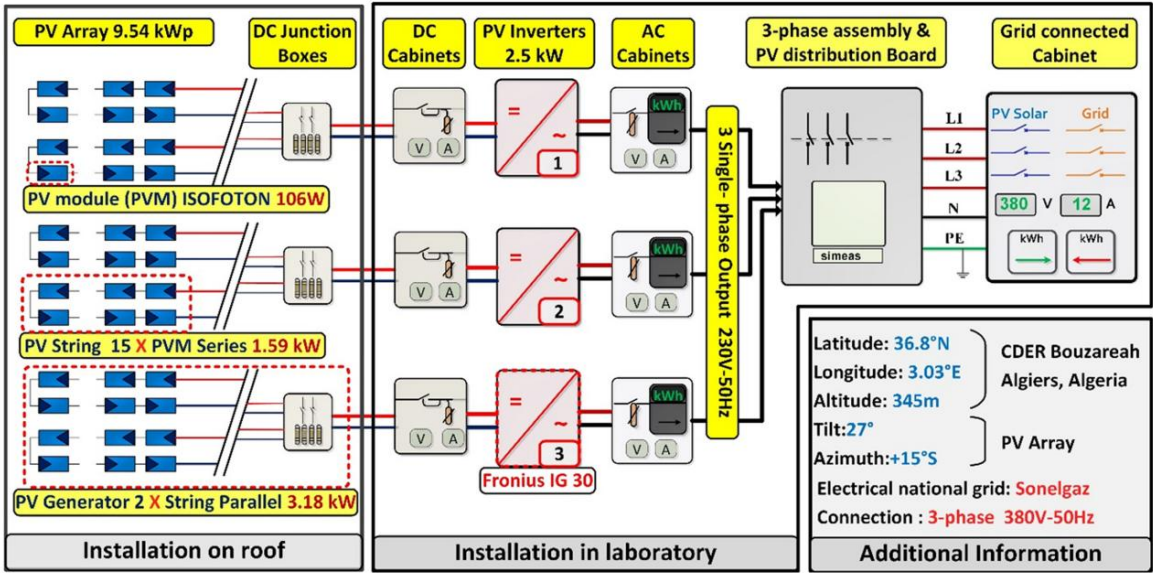
Mã số đối tượng kỹ thuật số 10.1109/JPHOTOV.2019.2896652

và nguồn năng lượng ít tốn kém hơn, có thể dẫn đến tương lai năng lượng bền vững.

Ngay cả với các công nghệ tiên tiến trong sản xuất điện PV, các nhà máy PV vẫn liên tục phải chịu các sự cố và xuống cấp có thể ảnh hưởng đến năng suất của chúng. Trên thực tế, có nhiều lỗi không thể đoán trước, bao gồm mạch hở, mạch ngắn, đứt dây-dây đứt dây, lỗi hồ quang, điểm nóng, lỗi MPPT, lỗi công tắc bộ biến đổi và bất thường của lưới điện, có thể ảnh hưởng đến hiệu suất mong muốn của các hệ thống PV [2], [3]. Ví dụ, các bất thường trên phần DC của hệ thống PV có thể dẫn đến mất năng lượng tiềm ẩn và thậm chí là các vấn đề an toàn thảm khốc [2], [4]. Để biết thêm thảo luận chi tiết về các loại lỗi phổ biến khác nhau trong hệ thống PV, hãy xem [2], [5] và [3]. Theo đó, việc giám sát thời gian thực các hệ thống PV đang trở nên quan trọng về cơ bản để đảm bảo an toàn, hoạt động đáng tin cậy và hiệu suất tối ưu của chúng.

Phát hiện lỗi trong các nhà máy PV quy mô tiện ích là điều cơ bản để tăng độ tin cậy, hiệu quả, hiệu suất tuổi thọ dịch vụ và độ an toàn của chúng. Do đó, để duy trì hoạt động tối ưu của các hệ thống PV, nhiều phương pháp giám sát khác nhau đã được thiết kế trong tài liệu [6]. Trong [7], một bản tóm tắt về các phương pháp phát hiện lỗi phổ biến trong các hệ thống PV đã được cung cấp. Gần đây, trong [8], một phương pháp dựa trên các gói sóng con để tăng cường phát hiện lỗi đứt dây-dây đứt dây trong một nhà máy PV bằng cách chỉ sử dụng các phép đo điện áp và dòng điện của mảng đã được đề xuất. Các phép biến đổi gói sóng con của điện áp mảng được lấy mẫu, năng lượng của điện áp mảng và năng lượng của sự thay đổi trở kháng được tính toán và so sánh với các ngưỡng quyết định được xác định trước để phát hiện lỗi. Phương pháp này cho thấy khả năng phù hợp để xác định lỗi đứt dây-dây đứt dây mà không cần sử dụng thêm cảm biến. Tuy nhiên, việc trích xuất các gói sóng con không phải là một nhiệm vụ dễ dàng để phát hiện lỗi trực tuyến và cần các nền tảng phần mềm bổ sung. Trong [9], một phương pháp làm mịn hàm mũ đa biến được áp dụng cho các vectơ dư từ một mô hình mô phỏng PV để phát hiện và xác định loại lỗi được phát hiện. Tuy nhiên, cách tiếp cận này dựa trên giả định chuẩn của phần dư. Trong [10], việc phát hiện lỗi hồ quang song song trong phần DC của nhà máy PV đã được giải quyết thông qua tiêu chí hỗn hợp kết hợp các tiêu chí trong miền thời gian và tần số. Cụ thể, người ta đã chỉ ra rằng bằng cách phân tích dòng điện, dải tần số đặc trưng của lỗi hồ quang song song nằm trong dải tần số 126-250 kHz.

Tuy nhiên, cách tiếp cận này đòi hỏi tốc độ lấy mẫu cao. Trong [11], một phương pháp để phát hiện và định vị các lỗi hồ quang nối tiếp và song song trong một nhà máy PV dựa trên các tụ điện song song đơn đã được đề xuất. Biến đổi wavelet rời rạc đã được thực hiện trên dòng điện tụ điện và phân phối biên độ, phổ tần số và cực tính của chúng được báo cáo tại đây [11].



Hình 1. Cấu trúc tổng quan của nhà máy điện mặt trời CDER.

Phương pháp này có thể phân biệt lỗi hồ quang với các hệ thống khác nhiều loạn. Gần đây hơn, trong [12], một phát hiện lỗi mạnh mẽ cách tiếp cận để giải quyết các phép đo nhiễu được đề xuất cho PV giám sát hệ thống. Cách tiếp cận này kết hợp dựa trên wavelet biểu diễn đa độ phân giải và làm mịn theo cấp số nhân để phát hiện lỗi. Nó cho thấy khả năng phát hiện lỗi tốt trong các tình huống nhiễu bằng cách giảm ảnh hưởng của tiếng ồn và bảo động sai. Tuy nhiên, cách tiếp cận này là một phương pháp hàng loạt, nghĩa là nó đòi hỏi toàn bộ dữ liệu có sẵn trước. Do đó, nó không thể được áp dụng trực tiếp để giám sát trực tuyến các hệ thống PV. Trong một nghiên cứu khác, [13] đã giới thiệu một quy trình để phát hiện lỗi hồ quang song song của phần DC sử dụng biến đổi Fourier thời gian ngắn (STFT) dựa trên dòng điện đầu ra của tấm pin PV. Kết quả của bài báo này cho thấy rằng thật khó để phân biệt giữa trạng thái hồ quang và không hồ quang trạng thái bằng cách sử dụng độ lớn của thành phần tần số ở trên 50 kHz. Theo đó, tổng các nội dung tần số dưới 50 kHz với giá trị tăng dần trong mỗi cửa sổ thời gian được giám sát để phát hiện lỗi. Tuy nhiên, phương pháp STFT này có thể đạt được hiệu quả phát hiện phù hợp bằng cách lựa chọn một cách thích hợp chiều dài cửa sổ. Trong [14], một phương pháp phát hiện lỗi dựa trên trên Thuật toán rừng ngẫu nhiên được giới thiệu để giám sát mảng PV. Phương pháp này chỉ sử dụng điện áp mảng PV và chuỗi dòng điện để phát hiện lỗi. Tuy nhiên, như một phương pháp học máy có giám sát, Rừng ngẫu nhiên cần các mẫu lỗi được gắn nhãn để phát hiện lỗi. Trong [15], một phương pháp để phát hiện các bất thường và che bóng một phần trong các nhà máy PV bằng cách sử dụng hai tính năng mới được liên kết để đo lường bức xạ và nhiệt độ được đề xuất. Điều này phương pháp có thể phân biệt giữa bóng râm một phần, khỏe mạnh và điều kiện bất thường. Người quyết định cho mỗi biến có đã được xác định và sử dụng để phát hiện lỗi theo các cách khác nhau tình trạng mây. Các nhà nghiên cứu khác tập trung vào việc sử dụng không dây giám sát cảm biến tự cấp nguồn cho từng tấm pin riêng lẻ để giám sát và chẩn đoán nhà máy PV [16]. Tuy nhiên, việc sử dụng của một số cảm biến để giám sát mảng PV dẫn đến các phương pháp tiếp cận tốn kém. Dựa trên chuỗi thời gian tự hồi quy vectơ (VAR) phương pháp mô hình hóa và tỷ lệ khả năng tổng quát, Chen et al.

[17] đề xuất một phương pháp phát hiện lỗi PV. Ý tưởng là sử dụng mô hình VAR để mô tả sự tiến triển theo thời gian của sản lượng PV biến và quy trình tỷ lệ khả năng tổng quát cho lỗi phát hiện [17]. Thật vậy, các biến đầu ra PV đã đưa ra một mùa chu kỳ có chu kỳ 24 giờ do tính theo mùa của đầu vào dữ liệu (tức là, các phép đo khí tượng). Tuy nhiên, VAR mô hình không tính đến tính theo mùa. Nói chung, khi một chuỗi thời gian thể hiện tính theo mùa, một mô hình chuỗi thời gian theo mùa có thể được sử dụng để nắm bắt tính theo mùa có tác dụng trong dữ liệu và cải thiện độ chính xác của phát hiện. Trong [18], một cơ chế phát hiện lỗi PV sử dụng mô hình hồi quy đa thức và thuật toán phân loại mờ được giới thiệu. Cơ chế này cho thấy hiệu suất đáng tin cậy trong việc phát hiện lỗi ngắn mạch và điểm nóng trong hệ thống PV. Tuy nhiên, điều này phương pháp không thể phát hiện lỗi xảy ra trong diode bypass và trong bộ biến tần dc/ac [18]. Trong [19], một quy trình cây quyết định C4.5 được đề xuất để giám sát hệ thống PV. Trong công trình này, Sandia mô hình được sử dụng để dự đoán năng lượng được sản xuất từ PV theo điều kiện không có lỗi. Sau đó, thuật toán cây quyết định có giám sát được sử dụng để phân biệt giữa bốn lớp, cụ thể là lỗi không có lỗi, lỗi chuỗi, lỗi ngắn mạch và lỗi đứt dây-dùng dây. Tuy nhiên, như một thuật toán học có giám sát, phương pháp này cần dữ liệu, một tập hợp của các ví dụ đào tạo để xây dựng một bộ phân loại có khả năng phân biệt cả ví dụ đào tạo và ví dụ kiểm tra. Hơn nữa, một số nghiên cứu tập trung vào việc phát hiện che bóng một phần trong hệ thống PV [20], [21]. Trong [20], phân tích thành phần chính (PCA) đã được áp dụng để xác định khả năng che bóng trong hệ thống PV bị cô lập bằng cách sử dụng các đường cong I-V. Ưu điểm của phương pháp này là khả năng cảm nhận bóng râm chỉ bằng cách sử dụng PV dòng điện và điện áp. Tuy nhiên, PCA, là một mô hình tĩnh (thời gian bất biến) không mô tả được sự phụ thuộc nối tiếp theo thời gian dữ liệu chuỗi, được sử dụng trong giám sát ngoại tuyến. Trong [22], di chuyển cửa sổ PCA (MWPCA) được sử dụng để giám sát thích ứng của PV hệ thống. Trong cách tiếp cận này, một bộ lọc dựa trên logic mờ được áp dụng để MWPCA giám sát các chỉ số để cải thiện hơn nữa khả năng phát hiện hiệu suất. Phương pháp MWPCA với chiều dài cố định của

cửa sổ cho thấy hiệu suất tốt hơn so với phư ơ ng pháp PCA. Tuy nhiên, trong MWPCA, số lư ợng mẫu đư ợc chọn trong cửa sổ đóng vai trò cốt lõi để có đư ợc hiệu suất phát hiện mong muốn, trong khi trong [21], các tác giả đã sử dụng hình ảnh phản xạ nhiệt để cải thiện hiệu quả MPPT của hệ thống PV trong tình huống che bóng một phần. Một nhữ ợc điểm lớn của phư ơ ng pháp này là cần phải phân tích định kỳ hệ thống PV thông qua camera ảnh nhiệt đất tiền. Sử dụng hệ thống máy bay không ngư ời lái, Li et al. [23] đã đề xuất một phư ơ ng pháp để giám sát các hệ thống PV quy mô lớn bằng cách phân tích hình ảnh PV trên không. Sơ đồ này phù hợp để xác định đư ờng đi của ốc sên và bóng bụi trong nhà máy PV. Tuy nhiên, các sơ đồ dựa trên tầm nhìn nhạy cảm với tiếng ồn và tốn kém về mặt tính toán (tức là cần xử lý trư ớc để trích xuất thông tin có liên quan từ hình ảnh).

Giám sát các hệ thống PV là chìa khóa để đảm bảo một môi trư ờng bền vững và đáp ứng hiệu suất và yêu cầu mong muốn. Theo đó, đảm bảo hiệu suất an toàn và mong muốn của các hệ thống PV là điều tối quan trọng đối với lư ời điện và ngư ời dùng cuối. Trong bài báo này, để phát triển một chiến lư ợc phát hiện lỗi sáng tạo để giám sát phía DC của hệ thống PV và che bóng tạm thời, các đặc điểm mong muốn của phư ơ ng pháp KNN và biểu đồ She-whart và trung bình động có trọng số theo hàm mũ (EWMA) [24] đã đư ợc sử dụng rộng rãi [25], cũng đã đư ợc sử dụng trong khuôn khổ phát hiện lỗi [26]. Trong [27], để giám sát hệ thống PV, quy tắc KNN truyền thống đư ợc áp dụng. Trong phư ơ ng pháp này, độ rọi ( $W/m^2$ ), nhiệt độ mô-đun ( $^{\circ}C$ ), điện áp (VMPP), dòng điện (IMPP) và công suất (PMPP) đã đư ợc sử dụng làm đầu vào cho thuật toán KNN để phát hiện và phân loại lỗi, trong khi đó, bài báo này xử lý phát hiện lỗi bằng cách áp dụng các thuật toán dựa trên KNN cho các phần đư thu đư ợc từ mô hình mô phỏng PV. Đáng chú ý, các phư ơ ng pháp đề xuất khai thác khả năng lớn hơn của KNN để tách các đặc điểm bình thư ờng khỏi bất thư ờng và độ nhạy của biểu đồ giám sát để phát hiện lỗi tốt hơn trong hệ thống PV [26]. Hơn nữa, trong bài báo này, một phư ơ ng pháp phi tham số sử dụng ư ớc tính mật độ hạt nhân đã đư ợc sử dụng để đặt ngư ỡng cho thống kê quyết định của các phư ơ ng pháp Shewhart và EWMA dựa trên KNN và so sánh với các phư ơ ng pháp tham số thông thư ờng. Các phần đư từ mô hình mô phỏng dựa trên một mô hình diode đư ợc sử dụng làm đầu vào cho các lư ợc đồ dựa trên KNN để phát hiện dị thư ờng. Hiệu suất của các thuật toán dựa trên KNN đư ợc đánh giá dựa trên các phép đo thực tế từ hệ thống kết nối lư ời điện 9,54 kWp.

Nhà máy PV đư ợc nghiên cứu đư ợc trình bày trong Phần II. Sau đó, mô hình mô phỏng đư ợc tóm tắt trong Phần III. Trong Phần IV, các phư ơ ng pháp tiếp cận dựa trên KNN đư ợc đề xuất (tham số và không tham số) và việc sử dụng chúng trong giám sát hệ thống PV đư ợc mô tả. Trong Phần V, hiệu quả phát hiện của các quy trình đã phát triển đư ợc đánh giá và các kết luận đư ợc trình bày trong Phần VI.

II. MÔ TẢ HỆ THỐNG PV

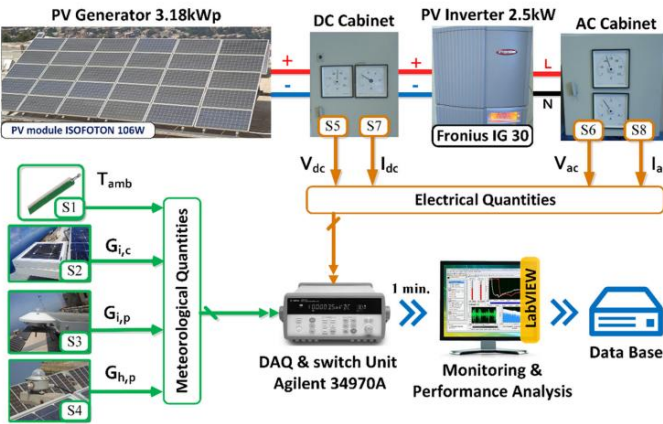
Hệ thống PV kết nối lư ời (GCPVS) đư ợc thể hiện trong Hình 1 bao gồm một mảng PV gồm 90 mô-đun ISOFOTON I106-12 công nghệ đơ n tinh thể Si. Nhà máy PV đư ợc hình thành từ ba máy phát điện PV tư ơ ng tự nhau có công suất 3,18 kWp (hai chuỗi song song gồm 15 mô-đun PV nối tiếp) và mỗi máy phát điện PV đư ợc kết nối với

BẢNG I  
ĐẶC ĐIỂM ĐIỆN CHÍNH CỦA MODULE PV VÀ MÁY PHÁT ĐIỆN

Parameters	$V_{OC} (V)$	$I_{SC} (A)$	$V_{MPP} (V)$	$I_{MPP} (A)$	$P_{MPP} (W)$
PV Module	21.6	6.54	17.4	6.1	106
PV Generator	324	13.08	261	12.2	3180

BẢNG II  
ĐẶC ĐIỂM KỸ THUẬT CHÍNH CỦA PV INVERTER FRONIUS IG 30

Parameters	Nominal $P_{AC} (W)$	$V_{MPP}$ range (V)	Efficiency (%)	$V_{AC}$ range (V)	Frequency range (Hz)
Value	2500	150-400	92.7-94.3	195-253	49.8-50.2



Hình 2. Hệ thống PV một pha và quy trình giám sát dữ liệu.

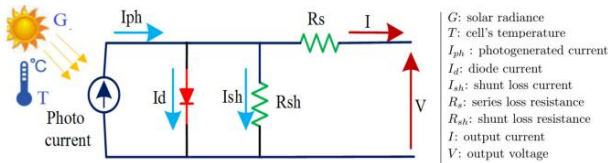
BẢNG III  
ĐẶC ĐIỂM CỦA CÁC CẢM BIẾN ĐƯ ỢC SỬ DỤNG ĐỂ GIÁM SÁT KHÍ TỰ ỢNG VÀ CÁC PHÉP ĐO ĐIỆN

Measured parameters			Symbol	N <sup>o</sup> Sensor	Sensor type	Sensor Reference	Max. Measured values.
Meteorological	Temperature	Ambient	$T_{amb}$	S1	Thermocouple K	-	50 °C
	Global Irradiance	Inclined_27°	$G_{1,L}$	S2	PV reference cell	Isofoton	1400 W/m <sup>2</sup>
		Horizontal	$G_{1,P}$	S3	Pyranometer	Kipp & zonen CM 11	
			$G_{H,P}$	S4			
Electrical	Voltage	DC	$V_{DC}$	S5	Voltage divider	-	340 V
		AC	$V_{AC}$	S6	Voltage transformer	-	260 V
	Current	DC	$I_{DC}$	S7	Hall effect	F.W. BELL	14 A
		AC	$I_{AC}$	S8	Closed-loop	CLSM-50S	12 A

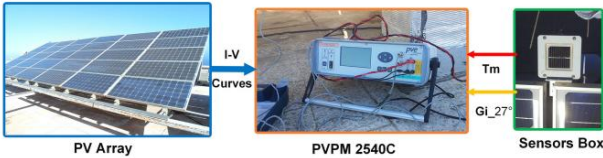
biến tần, tạo thành hệ thống PV một pha. Ba hệ thống PV một pha đư ợc nhóm thành ba pha và kết nối với lư ời điện quốc gia.

Các đặc tính điện của máy phát điện PV và bộ biến tần PV đư ợc sử dụng trong bài báo này lần lư ợt đư ợc báo cáo trong Bảng I và Bảng II. Hình 1 cho thấy sơ đồ điện toàn cầu của hệ thống PV đang đư ợc nghiên cứu.

Hệ thống giám sát cho hệ thống con thứ hai đư ợc thể hiện ở Hình 2, đư ợc sử dụng để theo dõi hệ thống con này bằng cách thu thập, hiển thị và phân tích các phép đo của cảm biến. Các phép đo khí tự ợng và điện đư ợc thu thập thông qua các cảm biến (xem Bảng III). Ở đây, thời gian lấy mẫu đư ợc cố định là 1 phút, cung cấp 1440 phép đo mỗi ngày và quá trình giám sát đư ợc thực hiện thông qua LabVIEW [28].



Hình 3. Sơ đồ tư ơng đư ơng của SDM của một tế bào quang điện.



Hình 4. Minh họa về đư ơng cong PVPM I-V .

III. MÔ HÌNH MẢNG PV

Mô hình diode đơ n (SDM), thư ờng đư ợc sử dụng để mô hình hóa pin mặt trời PV, đư ợc thể hiện trong Hình 3 và công thức phân tích mô tả đặc điểm I-V là [29]

$$I = I_{ph} - I_0 \exp\left(\frac{q(V + R_s I)}{n k_B T}\right) - \frac{V + R_s I}{R_{sh}} \quad (1)$$

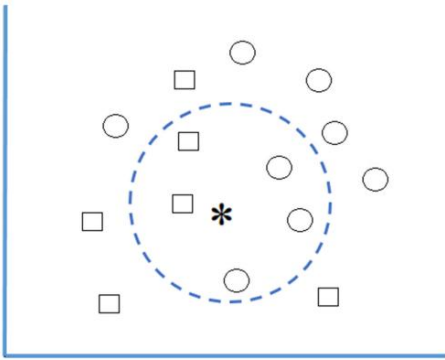
trong đó  $I_0$  biểu diễn dòng bão hòa tối,  $n$  biểu diễn hệ số lý tư ơng của diode,  $k_B$  là hằng số Boltzmann và  $q$  biểu diễn điện tích electron. Độ chính xác của SDM đư a ra bởi (1) phần lớn phụ thuộc vào các giá trị đã chọn của năm tham số chưa biết:  $I_{ph}$ ,  $I_0$ ,  $n$ ,  $R_s$  và  $R_{sh}$ .

Tóm lại, phư ơng pháp phát hiện dị thư ờng đư ợc đề xuất đư ợc chia thành các giai đoạn sau: 1) xác định các tham số của SDM; 2) xây dựng và xác thực mô hình mô phỏng của hệ thống PV bằng các tham số đã xác định; và 3) áp dụng các quy trình dựa trên KNN trên các giá trị đư o mô hình mô phỏng tạo ra để phát hiện các dị thư ờng trong hệ thống PV có giám sát.

Cụ thể hơn, trong bư ớc đầu tiên, các giá trị của các tham số đư ợc chọn sao cho sai số bình phư ơng trung bình căn bậc hai (RMSE) giữa các đư ờng cong I-V đo đư ợc và dự đoán thông qua SDM đư ợc giảm thiểu [30]. Ở đây, thiết bị PVPM 2540C (xem Hình 4) đư ợc sử dụng để đo đư ờng cong I-V của mô-đun PV tham chiếu và quy trình đàn ong nhân tạo (ABC) [30], [31] đư ợc áp dụng để xác định năm tham số chưa biết của SDM. Sau khi các tham số SDM đư ợc xác định, bư ớc tiếp theo tập trung vào việc lập mô hình toàn bộ mảng PV dựa trên dữ liệu thực phản ánh các điều kiện lành mạnh của nhà máy PV đư ợc kiểm tra. Bư ớc này đư ợc thực hiện bằng cách sử dụng mô phỏng đồng thời giữa Matlab/Simulink™ và PSIM™ dựa trên các tham số SDM đã tính toán. Sau đó, mô hình mô phỏng đư ợc xác minh bằng cách so sánh đỉnh công suất dự đoán Pmpsim của nó với đỉnh công suất đo đư ợc Pmpmeas.

IV. ĐỀ XUẤT CÁC KẾ HOẠCH GIÁM SÁT DỰA TRÊN KNN

Sơ đồ KNN là một phư ơng pháp phân biệt phi tham số đư ợc áp dụng rộng rãi để đo sự khác biệt giữa các phép đo thực tế và các phép đo đào tạo tư ơng ứng của chúng [32], [33]. Một lý do chính cho sự phổ biến của



Hình 5. Thuật toán KNN.

Sơ đồ KNN là phép tính đơ n giản và tính linh hoạt của nó bằng cách không đư a ra giả định về phân phối dữ liệu cơ bản. Điều này làm cho nó rất hữu ích khi các phép đo không phân phối theo chuẩn Gauss hoặc không thể phân biệt tuyến tính [25]. Nói chung, trong KNN, khi khoảng cách giữa phép đo mới và KNN của các phép đo không có lỗi gần bằng 0, thì bộ phát hiện KNN sẽ báo hiệu các phép đo này là bình thư ờng. Khoảng cách Euclidean thư ờng đư ợc sử dụng để định lượng độ gần trong các phư ơng pháp dựa trên KNN. Ở đây, các giá trị đư từ mô hình mô phỏng đư ợc đư a vào KNN để phát hiện dị thư ờng. Khoảng cách đáng kể là một chỉ báo tốt về các lỗi tiềm ẩn và do đó đư ợc sử dụng để phát hiện lỗi. Trong nghiên cứu này, hai phư ơng pháp phát hiện lỗi dựa trên khoảng cách KNN đã đư ợc giới thiệu.

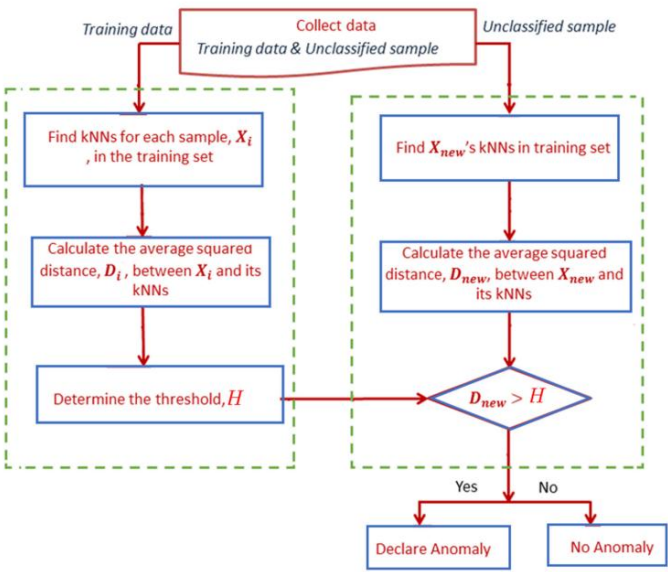
Hình 5 cung cấp một ví dụ cho thấy khái niệm cơ bản của phư ơng pháp tiếp cận KNN. Trong Hình 5, các hình vuông và hình tròn là dữ liệu từ hai lớp (tức là không có lỗi và bị lỗi); biểu thị điểm dữ liệu cần phân loại. Ví dụ, nếu số lượng hàng xóm gần nhất  $k = 1$ , thì điểm " " đư ợc phân loại là lớp hình vuông. Ở đây, khoảng cách Euclid đư ợc sử dụng để tìm các hàng xóm gần nhất. Nếu  $k = 5$ , điểm " " nằm trong lớp hình tròn. Điều này là do lớp đa số trong năm điểm gần nhất (nằm trong các vòng tròn đứt nét) là lớp hình tròn. Thật vậy, phát hiện lỗi dựa trên KNN, không có giai đoạn đào tạo rõ ràng, phép tính chính bao gồm việc tìm KNN của điểm dữ liệu thử nghiệm mới trực tuyến.

Ý tư ơng chính của thuật toán KNN, một bộ phát hiện không giám sát, để phát hiện lỗi là đánh giá sự khác biệt giữa dữ liệu thử nghiệm mới và dữ liệu không có lỗi (đào tạo). Thuật toán này chỉ cần không có lỗi trong quá trình đào tạo mà không cần gắn nhãn dữ liệu. Thật vậy, trong quá trình phát hiện lỗi dựa trên KNN, không có giai đoạn đào tạo rõ ràng và phép tính chính bao gồm việc tìm KNN của điểm dữ liệu thử nghiệm mới trực tuyến. Trong giai đoạn đào tạo, ngư ỡng quyết định KNN H đư ợc tính toán dựa trên dữ liệu không có lỗi. Trong giai đoạn thử nghiệm, khoảng cách giữa dữ liệu mới và dữ liệu đào tạo  $D_{new}$  đư ợc tính toán và so sánh với ngư ỡng đã tính toán để phát hiện lỗi. Khái niệm cơ bản về phư ơng pháp phát hiện lỗi dựa trên KNN đư ợc thể hiện trong Hình 6.

A. Phư ơng pháp tiếp cận KNN-Shewhart hỗn hợp

Phư ơng pháp này kết hợp các biện pháp KNN và sơ đồ Shewhart [34]. Trong phư ơng pháp này, sơ đồ Shewhart đư ợc áp dụng cho khoảng cách KNN để phát hiện lỗi. Thật vậy, đối với mỗi quan sát mới xi, hãy tính khoảng cách Euclidean của nó đến kth gần nhất





Hình 6. Sơ đồ khối của thuật toán phát hiện dị thường dựa trên KNN.

hàng xóm trong bộ đào tạo Di

$$C\grave{u}a = \sum_{j=1}^n \frac{1}{n} \cdot \text{ch\grave{a}nh sai} \tag{2}$$

trong đó dij là khoảng cách từ quan sát đến jth gần nhất của nó hàng xóm. Hiệu suất hệ thống có thể đư ợc đánh giá bằng cách so sánh khoảng cách KNN Di với ngư ỡng tham số đư ợc đư a ra đư i đây và biểu đồ kiểm soát kết quả đư ợc gọi là KNN-Shewhart biểu đồ

$$H = \mu D + 3\sigma D \tag{3}$$

trong đó  $\mu D$  và  $\sigma D$  là giá trị trung bình và độ lệch chuẩn của KNN khoảng cách trong điều kiện không có dị thường. Trong KNN-Shewhart thủ tục, khoảng cách Euclidean  $D$  của phần dư của dữ liệu thử nghiệm tới hàng xóm gần nhất của nó trong tập huấn luyện đư ợc tính toán và so sánh với ngư ỡng thống kê. Phư ơng pháp KNN-Shewhart báo hiệu lỗi khi giá trị của  $D$  nằm ngoài ngư ỡng. Các bư ớc chính trong phư ơng pháp phát hiện lỗi dựa trên KNN là đư ợc tóm tắt sơ đồ trong Hình 6.

B. Phư ơng pháp tiếp cận KNN-EWMA hỗn hợp

Ở đây, quy trình phát hiện lỗi KNN-EWMA đư ợc ghép nối là trình bày. Các chuỗi còn lại từ mô hình mô phỏng là đư ợc đánh giá bằng thuật toán KNN-EWMA. Ở đây, KNN đư ợc sử dụng để định lư ợng khoảng cách giữa quan sát dư thực tế và các mẫu dư đào tạo bình thường. Nói một cách trực quan, trong điều kiện danh nghĩa, khoảng cách KNN trở nên gần hơn đến 0, trong khi các giá trị khoảng cách KNN lớn hơn n ạt đư ợc trong điều kiện lỗi dẫn đến sự hiện diện của các sự kiện bất thường. Hãy định nghĩa vectơ khoảng cách KNN là  $D = [d_1, \dots, d_j, \dots, d_n]$ . Thống kê

KNN-EWMA là đư ợc định nghĩa là

$$z_{kNN}^t = v d_t + (1 - v) z_{kNN}^{t-1} \tag{4}$$

trong đó  $z_{kNN}^t$  là giá trị trung bình không có bất thường của chuỗi KNN khoảng cách  $\mu_0$ .  $v$  ( $0 < v \leq 1$ ) biểu thị tham số làm mịn. Thông thường, việc lựa chọn các giá trị nhỏ của  $v$  là phù hợp để tăng độ nhạy với độ lệch nhỏ trong khoảng cách KNN, trong khi việc sử dụng các giá trị lớn của  $v$  là phù hợp để phát hiện những thay đổi lớn hơn trong khoảng cách KNN [24]. KNN-EWMA cách tiếp cận cung cấp tín hiệu lỗi nếu thống kê KNN-EWMA  $z_{kNN}^t$  đang vượt quá ngư ỡng h đư ợc xác định là

$$h = \mu_{kNN} + k \sigma_{kNN} \sqrt{\frac{N}{(1-v)2t}} \tag{5}$$

trong đó  $k$  là hệ số nhân của độ lệch chuẩn của khoảng cách KNN  $\sigma_{kNN} = \sigma_{kNN} \sqrt{\frac{N}{(2-v)[1-(1-v)2t]}}$ .

Trong (4), có thể thấy rằng thống kê KNN-EWMA sử dụng tất cả thông tin từ các mẫu quá khứ và thực tế trong quy tắc quyết định, khiến nó nhạy cảm với những thay đổi nhỏ. Trong (4), phần trư ớc này thông tin  $z_{kNN}^{t-1}$  đư ợc kết hợp trong thống kê KNN-EWMA. Để thể hiện rõ điểm này, KNN-EWMA có thể đư ợc viết lại đê quy như sau:

$$z_{kNN}^t = v d_t + (1-v) z_{kNN}^{t-1} = v d_t + v(1-v) d_{t-1} + (1-v)^2 z_{kNN}^{t-2} \tag{6}$$

Sử dụng (4) đê quy, ta có

$$z_{kNN}^t = v d_t + v(1-v) d_{t-1} + v(1-v)^2 d_{t-2} + \dots + n(1-n)^{n-1} d_1 + (1-v) \text{ và } d_0. \tag{7}$$

Phư ơng trình (7) cũng có thể đư ợc biểu thị ở dạng nhỏ gọn như

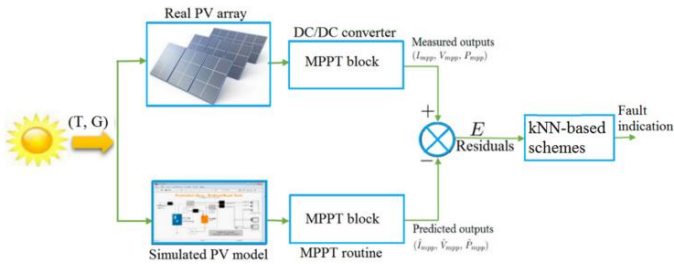
$$z_{kNN}^t = \sum_{t=1}^N \text{m\grave{o}t} (1 - \text{s\grave{o}})^{\text{kh\grave{o}ng}} d_t + (1-v) n d_0 \tag{8}$$

trong đó  $v(1-v)^{n-t}$  biểu thị trọng số cho  $d_t$ , giảm theo cấp số nhân đối với các mẫu trư ớc đó. Giá trị của  $v$  xác định tác động của các mẫu trong quá khứ lên giá trị hiện tại. Thông thường, việc lựa chọn các giá trị nhỏ của  $v$  là thích hợp để tăng độ nhạy với độ lệch nhỏ trong khoảng cách KNN, trong khi sử dụng của các giá trị lớn của  $v$  là thích hợp để phát hiện những thay đổi lớn hơn trong khoảng cách KNN. Giá trị của  $v$  giữa 0,2 và 0,3 là đư ợc khuyến nghị để phát hiện những thay đổi nhỏ [24].

C. Các lư ợc đồ dựa trên KNN với ngư ỡng không tham số

Cần lưu ý rằng ngư ỡng tham số trong Shewhart và EWMA đư ợc tính toán dựa trên giả định chuẩn của dữ liệu đầu vào. Khi giả định chuẩn mực không hợp lệ, kết quả phát hiện sẽ không phù hợp. Như một biện pháp khắc phục, phân phối khoảng cách KNN có thể đư ợc ước tính bằng cách sử dụng ước tính mật độ hạt nhân [35]. Các bư ớc chính của KNN dựa trên sơ đồ phát hiện dị thường với ngư ỡng không tham số là đư ợc trình bày tiếp theo.

Bư ớc 1: Đối với mọi phần dư của quan sát xi trong quá trình đào tạo tập dữ liệu, tìm khoảng cách Euclidean của nó đến gần nhất hàng xóm trong tập huấn luyện Di, dựa trên đó có thể thu đư ợc phân phối mẫu khoảng cách.



Hình 7. Sơ đồ khối của bộ phát hiện dị thường dựa trên KNN dự đoán.

Bước 2: Từ phân phối của Di, ngưỡng phi tham số của phương pháp KNN-Shewhart được định nghĩa là phân vị thứ  $(1 - \alpha)$  của phân phối ước tính của khoảng cách KNN thu được bởi KDE. Tương tự, ngưỡng phi tham số của phương pháp KNN-EWMA được định nghĩa là phân vị thứ  $(1 - \alpha)$  của phân phối ước tính của thống kê KNN-EWMA thu được bởi KDE.

Bước 3: Khi KNN-Shewhart hoặc KNN-EWMA vượt quá ngưỡng quyết định, lỗi được tuyên bố.

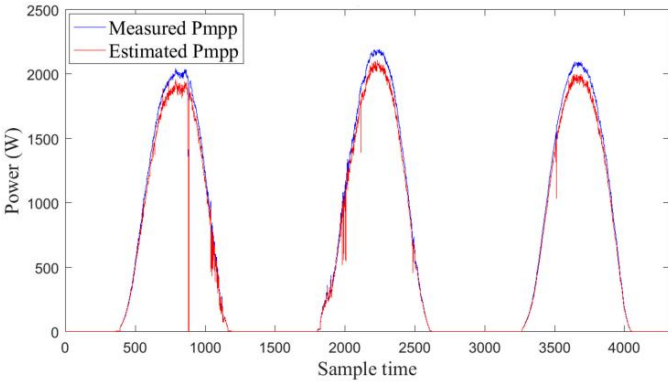
D. Giám sát hệ thống PV sử dụng các chương trình dựa trên KNN

Ở đây, khái niệm chính của các chương trình phát hiện lỗi dựa trên KNN được phát triển được mô tả ngắn gọn. Các phép đo điện từ một mảng PV và các phép đo môi trường là cần thiết để liên tục theo dõi hiệu suất của một mảng PV. Chiến lược phát hiện dị thường dựa trên KNN được đề xuất bao gồm ba phần: 1) xác định các tham số mô-đun PV; 2) xây dựng mô hình mô phỏng của mảng PV bằng cách sử dụng mô phỏng đồng thời PSIMTM /MatlabTM ; và 3) phát hiện lỗi bằng cách sử dụng các lược đồ giám sát KNN. Hình 7 cho thấy sơ đồ khối của hệ thống phát hiện lỗi đã phát triển.

Như đã đề cập ở trên, trước tiên, phương án tối ưu hóa ABC được áp dụng để xác định các tham số SDM của mô-đun PV. Sau đó, mô hình của mảng PV được nghiên cứu trong điều kiện vận hành lành mạnh được mô phỏng và xác minh. Sau đó, mô hình tham chiếu đã phát triển được sử dụng để dự đoán các biến MPP và tạo ra các giá trị dự của dữ liệu mới. Thật vậy, các giá trị dự biểu thị sự khác biệt giữa các giá trị tính toán của các biến MPP từ mô hình đã phát triển (dòng điện MPP  $I_t$ , điện áp MPP  $V_t$  và công suất đỉnh  $P_t$ ) và giá trị đo được (dòng điện MPP  $I$ , điện áp  $V$  và công suất đỉnh  $P$ ) thu thập được từ hệ thống PV. Khi phép đo mới được thu thập trong điều kiện vận hành hệ thống bình thường, các giá trị dự được tạo ra sẽ gần bằng không. Tuy nhiên, nếu phép đo mới được thu thập trong bất kỳ bất thường nào, các giá trị dự sẽ luôn lệch khỏi không. Phương pháp tiếp cận KNN được áp dụng cho các giá trị dự thu được từ mô hình mô phỏng để phát hiện bất thường. Ở đây, lược đồ dựa trên KNN được sử dụng để định lượng khoảng cách giữa quan sát dự thực tế và các mẫu dự tạo bình thường. Trong điều kiện bình thường, giá trị khoảng cách KNN dao động gần bằng không, trong khi trong điều kiện lỗi, khoảng cách KNN lệch đáng kể so với không. Trong nghiên cứu này, để thiết lập ngưỡng quyết định, hai phương pháp hợp nhất KNN vào biểu đồ giám sát EWMA và Shewhart đã được giới thiệu. Những phương pháp này

BẢNG IV  
XÁC ĐỊNH CÁC THAM SỐ SDM SỬ DỤNG PHƯƠNG PHÁP ABC

	$I_{ph}$ [A]	$I_0$ [A]	$n$	$R_s$ [ $\Omega$ ]	$R_{sh}$ [ $\Omega$ ]	RMSE
ISOFOTON 106-12	6.54	1.11e-0.5	1.66	0.1474	202.6	0.014



Hình 8. Công suất MPP thu thập và ước tính của mảng PV được nghiên cứu.

Các phương pháp kết hợp (tham số và phi tham số) được áp dụng cho khoảng cách KNN để phát hiện lỗi trong hệ thống PV một cách phù hợp.

V. KẾT QUẢ MÔ PHỎNG VÀ THẢO LUẬN

Phần này trình bày mô hình mảng PV và phát hiện dị thường kết quả phân tích bằng cách sử dụng các chương trình dựa trên KNN được đề xuất.

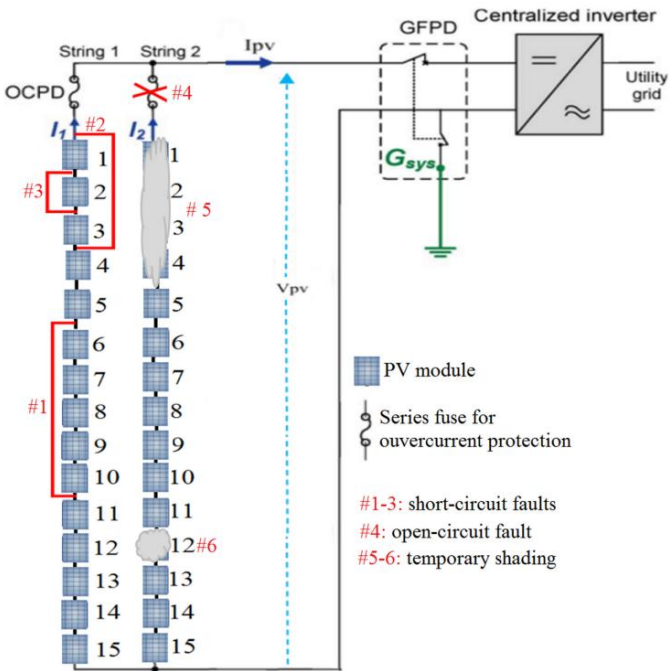
A. Mô hình hóa mảng PV

Như đã mô tả tóm tắt trong Phần III, các thông số mô-đun PV được xác định bằng cách sử dụng các phép đo đường cong thực (I-V) từ loại mô-đun PV Isofoton I-106/12, công nghệ đơn tinh thể, của mảng PV đang nghiên cứu. Các đường cong thực đo được (I-V) được thu thập dưới cường độ chiếu xạ ở mức 755 W/m<sup>2</sup> và nhiệt độ ở mức 27 °C. Các thông số SDM được xác định bằng cách sử dụng quy trình tối ưu hóa ABC được đưa ra trong Bảng IV.

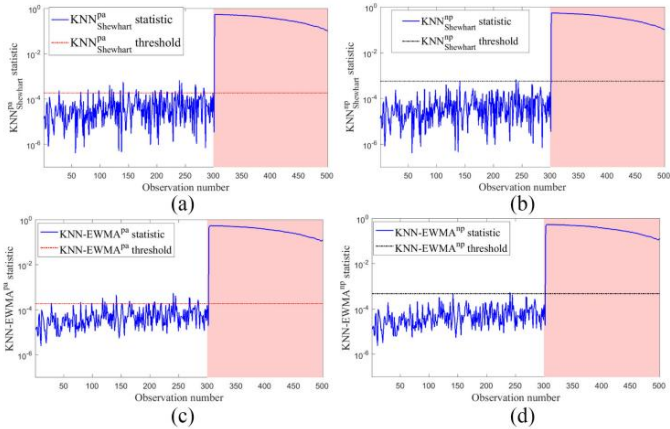
Hiện nay, dựa trên các thông số SDM đã xác định (xem Bảng IV), một mô hình mô phỏng toàn bộ mảng PV trong điều kiện bình thường được phát triển bằng Matlab/SimulinkTM và PSIMTM. Để đánh giá mô hình mô phỏng, ba hồ sơ ngày khí hậu thực tế về nhiệt độ và độ chiếu xạ được sử dụng làm đầu vào của mô hình này. Như thể hiện trong Hình 8, các phép đo công suất cực đại ( $P_{mppmeas}$ ) khớp tốt với giá trị dự đoán ( $P_{mppsim}$ ).

B. Kết quả phát hiện

Trong tiểu mục này, khả năng phát hiện của các lược đồ dựa trên KNN để phát hiện ra các bất thường trong dc của các mảng PV được xác minh. Các phép đo thực tế từ hệ thống GCPV 9,54 kWp (xem Phần II) được sử dụng để chứng minh hiệu quả của các lược đồ được đề xuất. Hiệu suất phát hiện lỗi của các lược đồ được đề xuất đã được mô tả bằng cách sử dụng số lượng dự đoán tính thực (TP), dự đoán tính giả (FP), âm tính thực (TN), âm tính giả (FN), độ chính xác, độ chính xác và diện tích dưới đường cong (AUC).



Hình 9. Những bất thường được xem xét trong mảng PV CDER.



Hình 10. (a)  $KNN_{Shewhart}^{pa}$ , (b)  $KNN_{Shewhart}^{np}$ , (c)  $KNN-EWMA_{pa}$ , và (d)  $KNN-EWMA_{np}$  khi năm mô-đun bị đoản mạch (lỗi số 1 trong Hình 9).

C. Trư ờng hợp A–Ví dụ về ngắn mạch

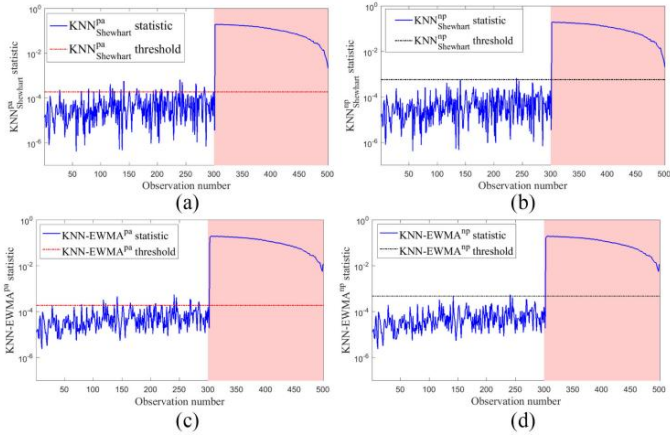
Ba kịch bản được trình bày ở đây để nghiên cứu rõ ràng độ nhạy của các phương pháp dựa trên KNN (tham số và không tham số): 1) năm mô-đun bị ngắn mạch; 2) ba mô-đun bị ngắn mạch; và 3) một mô-đun bị ngắn mạch trong chuỗi thứ hai (lỗi #1-3 trong Hình 9) từ số quan sát 300 đến 500.

Kết quả của bốn thuật toán trong trư ờng hợp có năm mô-đun bị ngắn mạch được thể hiện trong Hình 10. Số lượng hàng xóm gần nhất để tính toán khoảng cách K NN được chọn là  $k = 1$ . Như mong đợi, tất cả các phương pháp đều đánh giá sự bất thường này vì lỗi, trong trư ờng hợp này, là khá lớn. Hơn nữa, phương pháp tiếp cận tham số (EWMA dựa trên KNN và Shewhart) dựa trên giả định rằng các phần dư tuân theo phân phối chuẩn khi mảng PV đang được đề cập đang hoạt động đúng. Các ngưỡng của

BẢNG V

THỰC HIỆN CÁC CHƯƠNG TRÌNH GIÁM SÁT TRONG TRƯ ỜNG HỢP CỦA NĂM MODULE NGẮN MẠCH

Approach	TP	FP	FN	TN	TPR	FPR	Accuracy	Precision	AUC
KNN-EWMA <sup>np</sup>	199	0	3	299	0.985	0.000	0.994	1.000	0.993
KNN-EWMA <sup>pa</sup>	199	10	3	289	0.985	0.033	0.974	0.952	0.976
KNN-Shewhart <sup>np</sup>	200	1	2	298	0.990	0.003	0.994	0.995	0.993
KNN-Shewhart <sup>pa</sup>	200	29	2	270	0.990	0.097	0.938	0.873	0.947
Shewhart	197	0	4	300	0.980	0	0.992	1	0.990
EWMA	199	2	2	298	0.990	0.006	0.992	0.990	0.991



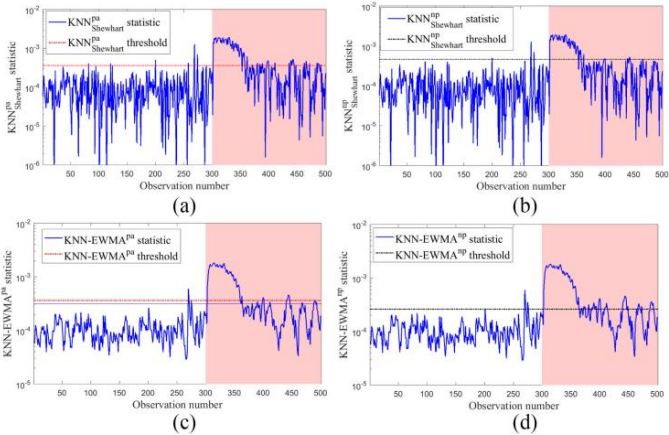
Hình 11. (a)  $KNN_{Shewhart}^{pa}$ , (b)  $KNN_{Shewhart}^{np}$ , (c)  $KNN-EWMA_{pa}$ ; EWMA<sup>np</sup> khi ba mô-đun bị đoản mạch (lỗi số 2 trong Hình 9).

EWMA và Shewhart dựa trên KNN tham số được tính toán với giả định rằng phần dư được phân phối chuẩn. Ngoài ra, chúng tôi đã so sánh các phương pháp tiếp cận dựa trên KNN với các phương pháp Shewhart và EWMA thông thường. Việc phát hiện hiệu suất của sáu thủ tục được thể hiện trong Bảng V. Bảng V cho thấy hiệu quả phát hiện được tăng cường đáng kể bằng cách sử dụng các phương pháp tiếp cận dựa trên KNN phi tham số. Nó nên được chỉ ra rằng sử dụng các phương pháp tiếp cận phi tham số, AUC là 0,99 đạt được, điều đó cho thấy khả năng cao trong việc phát hiện lỗi xảy ra. Biểu đồ tham số (tức là Shewhart dựa trên KNN và EWMA) phát hiện lỗi ngắn mạch này, nhưng có một số lỗi sai báo động và phát hiện bị bỏ sót (xem Bảng V). Kết quả phát hiện chỉ ra rõ ràng sự vượt trội của KNN dựa trên phi tham số cách tiếp cận.

Bây giờ, trong kịch bản thứ hai, khả năng phát hiện của các phương pháp đề xuất được đánh giá khi có ba mô-đun bị ngắn mạch trong mảng PV được kiểm tra (lỗi số 2 trong Hình 9). Kết quả trong Hình 11 và Bảng VI khẳng định rõ ràng tính ưu việt của các phương pháp tiếp cận Shewhart và EWMA dựa trên KNN không tham số so với các phương pháp tiếp cận KNN-EWMA và KNN-Shewhart tham số và các phương pháp Shewhart và EWMA thông thường, và cho thấy rằng các phương pháp tiếp cận phi tham số mang lại kết quả cao nhất độ chính xác và AUC so với độ chính xác tham số. Thật vậy, phương pháp phi tham số dựa trên KDE tránh mọi giả định trong phân phối dữ liệu và đã chứng minh khả năng phát hiện được cải thiện kết quả. Trong kịch bản thứ ba, khả năng phát hiện của đề xuất cách tiếp cận được đánh giá khi có một mạch ngắn

BẢNG VI
HIỆU SUẤT CỦA CÁC CHỨC NĂNG TRÌNH GIÁM SÁT KHI BA
CÁC MODULE NGẮN MẠCH

Table with 10 columns: Approach, TP, FP, FN, TN, TPR, FPR, Accuracy, Precision, AUC. Rows include KNN-EWMA, KNN-EWMA, KNN-Shewhart, KNN-Shewhart, Shewhart, and EWMA.

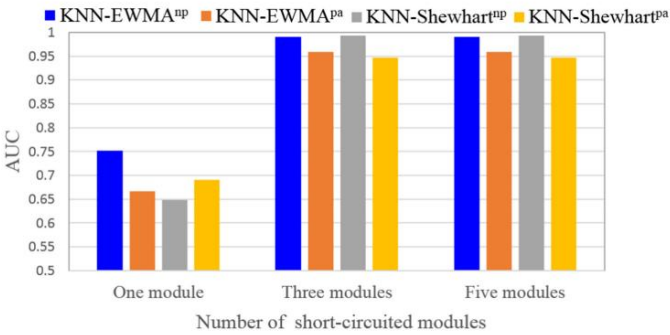


Hình 12. (a) kNNpha Shewhart, (b) kNNpha Shewhart, (c) kNN-EWMApha và (d) kNN-EWMApha trong có một mô-đun bị ngắn mạch (lỗi số 3 trong Hình 9).

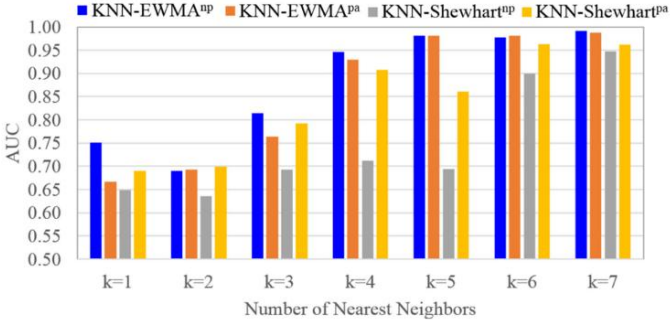
BẢNG VII
THỰC HIỆN CÁC CHỨC NĂNG TRÌNH GIÁM SÁT TRONG TRƯỜNG HỢP
CỦA MỘT MODULE NGẮN MẠCH

Table with 10 columns: Approach, TP, FP, FN, TN, TPR, FPR, Accuracy, Precision, AUC. Rows include KNN-EWMA, KNN-EWMA, KNN-Shewhart, KNN-Shewhart, Shewhart, and EWMA.

mô-đun trong một mảng PV (xem lỗi số 3 trong Hình 9), từ thời điểm tức thời 300-500. Kết quả giám sát của bốn phương pháp tiếp cận để kiểm tra dữ liệu được hiển thị trong Hình 12(a)-(d). Hình 12(a) và (b) cho thấy các phương pháp Shewhart dựa trên KNN (tham số và không tham số) dẫn đến một số phát hiện bị bỏ sót và bảo động giả (xem Bảng VII). Mặt khác, cốt truyện trong Hình 12(d) cho thấy hiệu quả phát hiện được cải thiện trong việc phát hiện lỗi nhỏ này khi sử dụng phương pháp phi tham số sơ đồ kNN-EWMA. Bảng VII tóm tắt hiệu suất của các phương pháp tiếp cận tham số và phi tham số, và độc lập Phương pháp Shewhart và EWMA. Từ Bảng VII, có vẻ như biểu đồ Shewhart không phù hợp để sử dụng trong trường hợp này. Phương pháp Shewhart độc lập không hiệu quả trong việc phát hiện tương đối lỗi nhỏ vì nó chỉ sử dụng quan sát thực tế trong quyết định và bỏ qua thông tin từ dữ liệu trong quá khứ. cách tiếp cận EWMA độc lập cung cấp kết quả phát hiện được cải thiện so với biểu đồ Shewhart. Điều này là do



Hình 13. AUC của các phương pháp dựa trên KNN theo chức năng của một số mô-đun bị ngắn mạch.



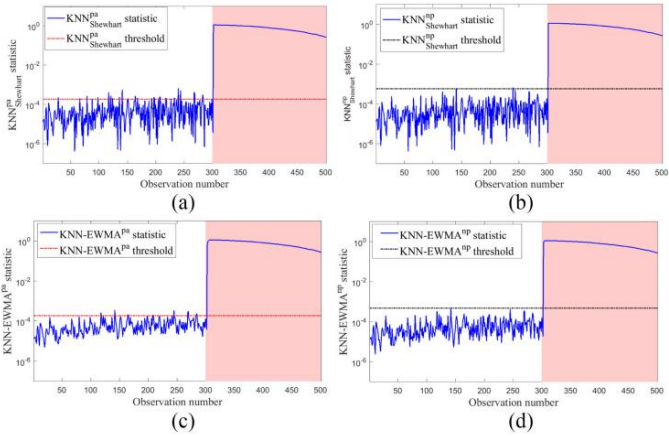
Hình 14. So sánh hiệu suất các chức năng dựa trên KNN khi thay đổi số lượng hàng xóm gần nhất.

thực tế là cách tiếp cận EWMA sử dụng tất cả dữ liệu quan sát được, điều này làm cho nó nhạy cảm với các lỗi nhỏ. Bảng VII cũng cho thấy rằng sự cải thiện đạt được với chức năng kNN-EWMA phi tham số so với việc sử dụng các phương pháp tiếp cận khác bằng cách tăng AUC từ 0,667 lên 0,752. Kết quả này thu được bởi KNN với khoảng cách Euclidian và số lượng hàng xóm k = 1.

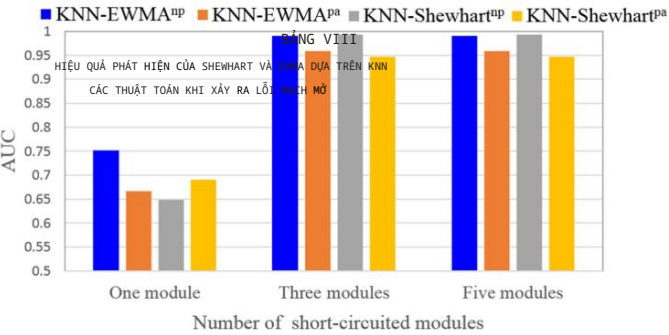
Hình 13 tóm tắt các giá trị AUC của bốn phương pháp tiếp cận (tham số và phi tham số) theo hàm số của các mô-đun bị ngắn mạch trong mảng PV được kiểm tra. Hình 13 cho thấy rằng kNN-EWMA phi tham số cung cấp chất lượng phát hiện tốt hơn so với các phương pháp khác, đặc biệt là đối với các lỗi (tức là một mô-đun bị ngắn mạch). Ngoài ra, có thể nhận thấy rằng các phương pháp tiếp cận không tham số hoạt động tốt hơn trong việc phát hiện lỗi ngắn mạch. Điều này chủ yếu là do tính linh hoạt của biểu đồ phi tham số trong việc tính toán ngưỡng quyết định dựa trên về ước tính mật độ hạt nhân.

Các kết quả trình bày ở trên được thu được khi số lượng của những người hàng xóm được sử dụng để tính toán khoảng cách KNN được chọn để là k = 1. Ở đây, bốn cách tiếp cận được đánh giá khi có một mô-đun bị ngắn mạch khi thay đổi số lượng hàng xóm. Nói chung, việc sử dụng một số lượng lớn hàng xóm sẽ làm tăng khả năng chống nhiễu và cải thiện khả năng phát hiện chất lượng. Ngoài ra, nó có thể làm tăng độ nhạy của máy dò bằng cách đưa thêm thông tin vào quy tắc quyết định. Người ta đã báo cáo trong tài liệu rằng các giá trị lớn hơn của k giảm tác động của tiếng ồn lên khả năng phân biệt [36]. Hình 14 cho thấy các giá trị AUC của bốn phương pháp tiếp cận khi số lượng hàng xóm được sử dụng để tính toán sự thay đổi khoảng cách KNN



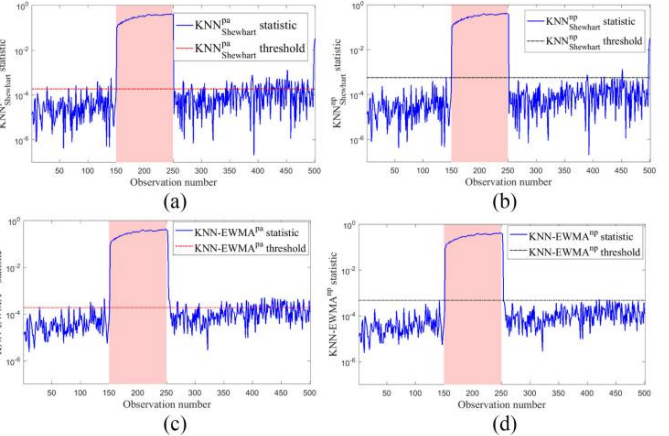


Hình 15. (a)  $kNN_{Shewhart}^{pa}$ , (b)  $kNN_{Shewhart}^{np}$ , (c)  $kNN_{EWMA}^{pa}$ ;  $kNN_{EWMA}^{np}$  khi xảy ra lỗi mạch hở.

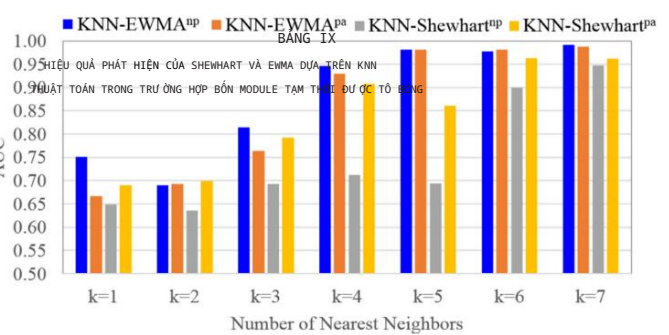


từ 1 đến 7. Kết quả trong Hình 14 cho thấy việc sử dụng các giá trị lớn hơn của các láng giềng gần nhất cải thiện hiệu suất phát hiện của tất cả các phương pháp tiếp cận. Ngoài ra, có thể thấy rằng các phương pháp tiếp cận không tham số cung cấp kết quả tốt hơn trong việc phát hiện lỗi ngắn mạch này trong mảng PV. So sánh các phương pháp tiếp cận dựa trên KNN hiệu suất với giá trị  $k$  lớn (ví dụ:  $k = 6$ ) (xem Hình 14), đối với các phương pháp Shewhart và EWMA độc lập (xem Bảng VII), kết quả khẳng định hiệu quả phát hiện vượt trội của các phương pháp tiếp cận được đề xuất so với các phương pháp độc lập. Tóm lại, việc sử dụng các giá trị  $k$  lớn hơn cho phép cải thiện chất lượng hiệu suất của máy dò bằng cách tăng khả năng phát hiện tỷ lệ và giảm báo động giả. Ngoài ra, KNN-EWMA kế hoạch sử dụng tất cả các thông tin có sẵn từ quá khứ và hiện tại để đo trong thống kê quyết định, khiến nó nhạy cảm với các lỗi nhỏ. Hơn nữa, bằng cách tính toán quyết định ngay lập tức, nó không tham số làm cho máy dò linh hoạt hơn và hiệu quả hơn các phương pháp tiếp cận tham số, dựa trên phân phối phân bố được xác định trước.

1) Trùng hợp B-Ví dụ về mạch hở: Mục tiêu trong tình huống này là để điều tra năng lực của Shewhart dựa trên KNN và các phương pháp tiếp cận EWMA (tham số và phi tham số) trong việc xác định lỗi mạch hở trong một mảng PV. Để làm như vậy, một mạch hở lỗi trong mảng PV được đưa vào (lỗi số 4 trong Hình 9) trong thời gian tức thời 300-500. Hình 15(a)-(d) cho thấy kết quả phát hiện của các thuật toán dựa trên KNN dựa trên các giá trị dư thừa của công suất MPP. Trong Ngoài ra, chúng tôi đã so sánh hiệu suất phát hiện của đề xuất cách tiếp cận với Shewhart và EWMA độc lập phương pháp (xem Bảng VIII). Như thể hiện trong Hình 15 và Bảng VIII,

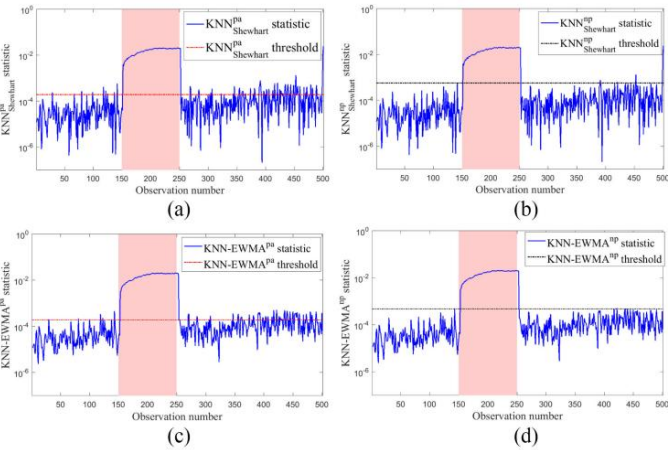


Hình 16. (a)  $kNN_{Shewhart}^{pa}$  và (b)  $kNN_{Shewhart}^{np}$  trong trùng hợp bốn mô-đun PV được che bóng tạm thời.



hiệu quả phát hiện được tăng cường đáng kể bằng cách sử dụng đề xuất Các phương pháp tiếp cận dựa trên KNN phi tham số. Có thể thấy từ Bảng VIII rằng các phương pháp tiếp cận dựa trên KNN phi tham số cung cấp hiệu suất phát hiện tốt hơn so với các phiên bản tham số của các phương pháp tiếp cận dựa trên KNN. Thật vậy, trong trùng hợp khi phân phối dữ liệu không có lỗi không phải là Gaussian, lỗi có thể được phát hiện hiệu quả hơn bằng các phương pháp giám sát phi tham số (hoặc không phân phối). Thật vậy, các phương pháp phi tham số vượt trội hơn các phương pháp khác (xem Bảng VIII) bằng cách đạt được AUC là 0,993. So với Shewhart và EWMA độc lập, phương pháp đề xuất đạt được hiệu suất phát hiện tương đương (xem Bảng VIII). Điều này là do lỗi lớn và dễ phát hiện.

2) Trùng hợp C-Các tình huống có bóng râm tạm thời: Ở đây, Mục đích là kiểm tra tính khả thi của các phương pháp được đề xuất để cảm nhận bóng râm tạm thời trong một mảng PV. Trong ví dụ đầu tiên, bốn mô-đun được che bóng tạm thời (lỗi số 5 trong Hình 9) từ thời gian tức thời 150-250. Bốn biểu đồ giám sát được hiển thị trong Hình 16(a)-(d). Hiệu suất của các phương pháp tiếp cận được nghiên cứu là so với biểu đồ Shewhart và EWMA độc lập (xem Bảng IX). Trong kịch bản này, Shewhart và EWMA dựa trên KNN các chương trình với người không tham số có thể được cho là máy dò tốt nhất, dẫn đến giá trị AUC cao khoảng 0,98 mặc dù các phương pháp tiếp cận tham số có thể phát hiện ra điều này tạm thời che bóng với AUC khoảng 0,90 với nhiều cảnh báo sai (xem Hình 16(a) và Bảng IX). Tất nhiên, những kết quả này xác nhận sự vượt trội của các phương pháp phi tham số so với phương pháp tiếp cận tham số.



Hình 17. (a) kNN<sup>na</sup> Shewhart, (b) kNN<sup>np</sup> Shewhart, (c) kNN-EWMA<sup>na</sup> ; EWMA<sup>np</sup> trong trường hợp một mô-đun PV bị che bóng tạm thời.

BẢNG X

HIỆU QUẢ PHÁT HIỆN CỦA SHEWHART VÀ EWMA DỰA TRÊN KNN THUẬT TOÁN TRONG TRƯỞNG HỢP MỘT MODULE TẠM THỜI ĐƯỢC TỔ BÓNG

Approach	TP	FP	FN	TN	TPR	FPR	Accuracy	Precision	AUC
KNN-EWMA <sup>np</sup>	98	6	3	394	0.970	0.015	0.982	0.942	0.978
KNN-EWMA <sup>na</sup>	98	71	3	329	0.970	0.178	0.852	0.580	0.896
KNN-Shewhart <sup>np</sup>	99	8	2	392	0.980	0.020	0.980	0.925	0.980
KNN-Shewhart <sup>na</sup>	99	77	2	323	0.980	0.193	0.842	0.563	0.894
Shewhart	0	0	100	400	0	0	0.8	0	0.5
EWMA	98	28	2	373	0.980	0.069	0.940	0.777	0.955

Bây giờ, chúng ta hãy xem xét tình huống khi có một mô-đun PV tạm thời bị che bóng trong hệ thống PV (xem lỗi số 6 trong Hình 9), từ thời điểm 150-250. Kết quả giám sát của bốn cách tiếp cận được thể hiện trong Hình 17. So sánh hiệu suất với các phương pháp Shewhart và EWMA độc lập là tóm tắt trong Bảng X. Shewhart độc lập không hiệu quả trong việc phát hiện lỗi nhỏ (xem Bảng X). EWMA độc lập phương pháp cung cấp hiệu suất phát hiện phù hợp với AUC khoảng 0,95. Thực tế này là do độ nhạy của Phương pháp EWMA đối với các lỗi nhỏ. Dựa trên kết quả này, các phương pháp không tham số cho thấy kết quả phát hiện tốt hơn với Giá trị AUC khoảng 0,98 mạnh mẽ so với tham số các cách tiếp cận mà họ đạt được AUC khoảng 0,89 (xem Bảng X).

Tóm lại, các phương pháp Shewhart và EWMA dựa trên KNN đã được đề xuất với ngưỡng tham số và không tham số. Đối với việc giám sát hệ thống PV, các phần dư được tạo ra từ mô hình mô phỏng được sử dụng làm đầu vào cho các máy dò dựa trên KNN. Sau đó, các biểu đồ Shewhart và EWMA thống nhất được áp dụng cho khoảng cách dựa trên KNN để phát hiện lỗi. Từ kết quả số được mô tả ở trên, có thể kết luận rằng phương pháp dựa trên KNN với ngưỡng không tham số đáng tin cậy hơn khi sử dụng so với biểu đồ có ngưỡng tham số, và chúng cũng khá hiệu quả trong nhiều trường hợp khác nhau được xem xét bên trên. Giữa các chương trình KNN-Shewhart và KNN-EWMA, Có vẻ như chương trình KNN-EWMA nhìn chung có hiệu quả hơn.

VI. KẾT LUẬN

Trong quá trình xem xét, các phương pháp tiếp cận dựa trên mô hình sáng tạo đã được giới thiệu để phát hiện các lỗi tiềm ẩn ở phía DC của PV hệ thống. Một mô hình mô phỏng được phát triển để mô phỏng thực tế đặc điểm của mảng PV được giám sát và tạo ra các phần dư để phát hiện dị thường. Các chương trình phát hiện mới tích hợp lợi ích của phương pháp KNN và các biểu đồ giám sát đơn biến để phát hiện lỗi một cách thích hợp. Bốn phương pháp dựa trên khoảng cách KNN đã được thiết kế cho lỗi phát hiện, các lược đồ Shewhart và EWMA dựa trên KNN (tham số và không tham số). Kết quả phát hiện được cải thiện đã được đạt được bằng cách sử dụng ngưỡng không tham số được tính toán thông qua KDE, so với việc sử dụng ngưỡng tham số KNN. Phương pháp đã phát triển được xác minh bằng cách sử dụng các phép đo thực tế từ hệ thống PV 9,54 kWp ở Algeria. Hiệu quả phát hiện của các phương pháp phát hiện lỗi dựa trên KNN tham số và phi tham số đã được so sánh. Kết quả cho thấy các phương pháp tiếp cận EWMA và Shewhart dựa trên KNN với phi tham số ngưỡng đã chứng minh khả năng phát hiện vượt trội trong việc phát hiện lỗi hở mạch, lỗi ngắn mạch và bóng râm tạm thời.

Lưu ý rằng thuật toán được đề xuất có khả năng phát hiện một phần bóng mờ, nhưng không thể phân biệt nó với các lỗi xảy ra trong phía dc của một mảng PV. Để cải thiện hơn nữa khả năng phát hiện bất thường và chẩn đoán trong hệ thống PV, trong các công trình tương lai, một hệ thống tự động phương pháp giám sát có thể được thiết kế bằng cách phân tích các phần còn lại để phân biệt giữa các loại lỗi khác nhau và xác định vị trí thành phần lỗi trong mảng PV được kiểm tra. Nó được thiết kế để xem xét các thông số khác, chẳng hạn như điện áp mạch hở, dòng điện ngắn mạch và hệ số lấp đầy, vì bóng râm ảnh hưởng đến ba thông số này các thông số đáng kể.

TÀI LIỆU THAM KHẢO

[1] IEA PVPS, “Bức tranh tổng quan về thị trường quang điện toàn cầu,” IEA PVPS, St. Urten, Thụy Sĩ, Đại diện IEA PVPS T1-33:2018, 2018.

[2] DS Pillai và N. Rajasekar, “Một đánh giá toàn diện về bảo vệ thách thức và chẩn đoán lỗi trong hệ thống PV,” Renewable Sustain. Năng lượng Rev., tập 91, trang 18-40, 2018.

[3] MK Alam, F. Khan, J. Johnson và J. Flicker, “Một đánh giá toàn diện về các lỗi thâm khốc trong mảng PV: Các loại, phát hiện và giảm thiểu kỹ thuật,” IEEE J. Photovolt., tập 5, số 3, trang 982-997, tháng 5 năm 2015.

[4] MJ Albers và G. Ball, “Đánh giá so sánh về việc giảm thiểu lỗi DC kỹ thuật trong các hệ thống PV lớn,” IEEE J. Photovolt., tập 5, số 4, trang 1169-1174, tháng 7. 2015.

[5] SR Madeti và SN Singh, “Một nghiên cứu toàn diện về các loại khác nhau về các lỗi và kỹ thuật phát hiện lỗi trong hệ thống quang điện mặt trời,” Sol. Năng lượng, tập 158, trang 161-185, 2017.

[6] S. Silvestre, “Các chiến lược phát hiện và chẩn đoán lỗi của hệ thống PV,” trong Những tiến bộ trong Năng lượng tái tạo và Công nghệ điện. Elsevier, 2018, tr. 231-255. [Trực tuyến]. Có sẵn: <https://doi.org/10.1016/B978-0-12-812959-3.00007-1>

[7] A. Mellit, GM Tina và SA Kalogirou, “Phát hiện và chẩn đoán lỗi phương pháp cho hệ thống quang điện: Một đánh giá,” Renewable Sustain. Năng lượng Rev., tập 91, trang 1-17, 2018.

[8] BP Kumar, GS Ilango, MJB Reddy và N. Chilakapati, “Lỗi trực tuyến phát hiện và chẩn đoán trong hệ thống quang điện sử dụng các gói sóng nhỏ,” IEEE J. Photovolt., tập 8, số 1, trang 257-265, tháng 1 năm 2018.

[9] F. Harrou, Y. Sun, B. Taghezouit, A. Saidi, và M.-E. Hamlati, “Đánh tin cậy phát hiện và chẩn đoán lỗi của hệ thống quang điện dựa trên thống kê “các phương pháp tiếp cận giám sát”, Năng lượng tái tạo, tập 116, trang 22-37, 2018.

[10] C. He, L. Mu và Y. Wang, “Phát hiện lỗi hở quang song song trong hệ thống quang điện dựa trên tiêu chí hỗn hợp,” IEEE J. Photovolt., tập 7, số 6, trang 1717-1724, 2017.

[11] Q. Xiong et al., “Phát hiện và định vị lỗi hồ quang trong hệ thống quang điện bằng cách sử dụng bản đồ phân bố đặc điểm của dòng điện tự điện song song,” IEEE J. Photovolt., số 99, trang 1-8, 2018.

[12] F. Harrou, B. Taghezouit và Y. Sun, “Chiến lược mạnh mẽ và linh hoạt để phát hiện lỗi trong hệ thống quang điện dự ợc kết nối với lưu trữ điện,” Energy Convers. Quản lý, tập 180, trang 1153-1166, 2019.

[13] S. Chen, X. Li, và J. Xiong, “Xác định lỗi hồ quang nối tiếp cho hệ thống quang điện dựa trên phân tích miền thời gian và miền thời gian-tần số,” IEEE J. Photovolt., tập 7, số 4, trang 1105-1114, tháng 7 năm 2017.

[14] Z. Chen et al., “Chẩn đoán lỗi thông minh dựa trên rừng ngẫu nhiên cho các mảng PV sử dụng điện áp mảng và dòng điện chuỗi,” Energy Convers. Manage., tập 178, trang 250-264, 2018.

[15] R. Hariharan, M. Chakkarapani, GS Ilango và C. Nagamani, “Một phương pháp phát hiện lỗi mảng quang điện và che bóng một phần trong hệ thống quang điện”, IEEE J. Photovolt., tập 6, số 5, trang 1278-1285, tháng 9 năm 2016.

[16] P. Guerriero et al., “Giám sát và chẩn đoán các nhà máy điện mặt trời bằng cảm biến tự cấp nguồn không dây cho từng tấm pin riêng lẻ”, IEEE J. Photovolt., tập 6, số 1, trang 286-294, tháng 1 năm 2016.

[17] L. Chen, S. Li và X. Wang, “Phát hiện lỗi nhanh nhất trong hệ thống quang điện”, IEEE Trans. Smart Grid, tập 9, số 3, trang 1835-1847, tháng 5 năm 2018.

[18] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales và P. Mather, “Thuật toán phát hiện lỗi quang điện dựa trên mô hình dự ờng cong lý thuyết và hệ thống phân loại mờ,” Năng lượng, tập 140, trang 276-290, 2017.

[19] R. Benkercha và S. Moulahoum, “Phát hiện và chẩn đoán lỗi dựa trên thuật toán cây quyết định c4. 5 cho hệ thống PV dự ợc kết nối lưu trữ điện,” Sol. Energy, tập 173, trang 610-634, 2018.

[20] S. Fadhel et al., “Phát hiện và phân loại lỗi che bóng PV dựa trên dự ờng cong IV bằng cách sử dụng phân tích thành phần chính: Ứng dụng cho hệ thống PV bị cò lập,” Sol. Energy, tập 179, trang 1-10, 2019.

[21] Y. Hu, W. Cao, J. Wu, B. Ji và D. Holliday, “Sơ đồ MPPT ảo dựa trên nhiệt độ để cải thiện hiệu suất năng lượng PV trong điều kiện che bóng một phần”, IEEE Trans. Power Electron., tập 29, số 11, trang 5667-5672, tháng 11 năm 2014.

[22] M. Ammiche, A. Kouadri, LM Halabi, A. Guichi và S. Mekhilef, “Phát hiện lỗi trong hệ thống quang điện dự ợc kết nối lưu trữ điện bằng phương pháp ngưỡng thích ứng,” Sol. Energy, tập 174, trang 762-769, 2018.

[23] X. Li, Q. Yang, Z. Chen, X. Luo và Y. Wenjung, “Phát hiện khuyết tật có thể nhìn thấy dựa trên kiểm tra dựa trên UAV trong các hệ thống quang điện quy mô lớn,” IET Renewable Power Gener., tập 11, số 10, trang 1234-1244, 2017.

[24] DC Montgomery, “Giới thiệu về Kiểm soát chất lượng thống kê. Hoboken, NJ, Hoa Kỳ: Wiley, 2007.

[25] X. Wu et al., “10 thuật toán hàng đầu trong khai thác dữ liệu,” Knowl. Inf. Syst., tập 14, số 1, trang 1-37, 2008.

[26] QP He và J. Wang, “Phát hiện lỗi bằng quy tắc k-gần nhất cho quy trình sản xuất chất bán dẫn,” IEEE Trans. Semicond. Manuf., tập 20, số 4, trang 345-354, tháng 11 năm 2007.

[27] SR Madeti và SN Singh, “Mô hình hóa hệ thống PV dựa trên dữ liệu thực nghiệm để phát hiện lỗi bằng phương pháp knn,” Sol. Energy, tập 173, trang 139-151, 2018.

[28] A. Chouder, S. Silvestre, B. Taghezouit và E. Karatepe, “Giám sát, mô hình hóa và mô phỏng hệ thống PV bằng LabVIEW,” Sol. Energy, tập 91, trang 337-349, 2013.

[29] W. Zhou, H. Yang và Z. Fang, “Một mô hình mới để dự đoán hiệu suất của mảng quang điện,” Appl. Energy, tập 84, số 12, trang 1187-1198, 2007.

[30] E. Garoudja et al., “Phát hiện lỗi thống kê trong hệ thống quang điện,” Sol. Energy, tập 150, trang 485-499, 2017.

[31] D. Karaboga và B. Basturk, “Về hiệu suất của thuật toán đàn ong nhân tạo (ABC),” Appl. Soft Comput., tập 8, số 1, trang 687-697, 2008.

[32] J. Han, J. Pei và M. Kamber, Khai thác dữ liệu: Khái niệm và kỹ thuật. Waltham, Hoa Kỳ: Nhà xuất bản Morgan Kaufmann, 2011.

[33] T. Cheng, F. Harrou, Y. Sun và TO Leiknes, “Giám sát các phép đo dòng chảy vào tại cơ sở phục hồi tài nguyên nước bằng cách sử dụng phương pháp cảm biến mềm dựa trên dữ liệu,” IEEE Sensors J., tập 19, số 1, trang 1-11, tháng 1 năm 2019.

[34] F. Kadri, F. Harrou, S. Chaabane, Y. Sun và C. Tahon, “Biểu đồ SPC theo mùa dựa trên ARMA để phát hiện dị thường: Ứng dụng vào hệ thống khoa cấp cứu,” Neurocomputing, tập 173, trang 2102-2114, 2016.

[35] EB Martin và AJ Morris, “Giới hạn tin cậy phi tham số cho biểu đồ giám sát hiệu suất quy trình,” J. Process Control, tập 6, số 6, trang 349-358, 1996.

[36] BS Everitt, S. Landau, M. Leese và D. Stahl, “Các phương pháp phân cụm khác nhau,” trong Phân tích cụm. Wiley, 2011, trang 215-255.

Ảnh và tiểu sử của tác giả không có sẵn tại thời điểm xuất bản.