

ỨNG DỤNG PHƯƠNG PHÁP PHÂN TÍCH THÀNH PHẦN CHÍNH PHÁT HIỆN BẤT THƯỜNG TRONG QUY TRÌNH RANG KẼM

APPLICATION OF PRINCIPAL COMPONENT ANALYSIS METHOD TO DETECTE ABNORMALITY IN THE ZINC ROASTING PROCESS

Đỗ Văn Vang^{1,2,*},
Nguyễn Thị Thương Duyên¹, Trần Thị Thu Lan¹

DOI: <https://doi.org/10.57001/huih5804.2023.165>

TÓM TẮT

Rang kẽm là quy trình đầu tiên quan trọng nhất của quy trình sản xuất kẽm. Quy trình sản xuất kẽm ngày càng có quy mô lớn, độ phức tạp cao với hàng loạt các phản ứng vật lý, hóa học kết hợp. Các điều kiện vận hành bình thường có thể chuyển sang các điều kiện bất thường như phân hủy quá mức, oxy hóa kém, thiếu kết tủa sôi do sự dao động của nguyên liệu thô và sự rò rỉ của hệ thống nước làm mát tuần hoàn. Sự hoạt động ổn định và an toàn của quá trình rang có ý nghĩa quan trọng trong việc đảm bảo chất lượng kẽm đầu ra, giảm thiểu ô nhiễm công nghiệp và tiêu thụ năng lượng. Để thực hiện vận hành an toàn, ổn định, tránh các rủi ro về an toàn, thiệt hại về kinh tế thì việc giám sát phát hiện các bất thường trong quy trình là rất quan trọng. Một phương pháp phân tích thành phần chính được đề xuất để giám sát phát hiện hai điều kiện bất thường điển hình trong quy trình rang kẽm. Phương pháp phân tích sử dụng dữ liệu công nghiệp thu thập được từ quy trình rang kẽm thực tế.

Từ khóa: Phân tích thành phần chính, giám sát quá trình, quy trình rang kẽm.

ABSTRACT

Roasting is the first and most crucial process in the zinc production process. The zinc production process is becoming increasingly large in scale and highly complex, involving a series of combined physical and chemical reactions. Normal operating conditions can transition into abnormal conditions such as excessive decomposition, poor oxidation, and layer boiling agglomeration due to fluctuations in raw materials and leakage in the circulating cooling water system. The stable and safe operation of the roasting process is of great significance in ensuring the quality of the output zinc, minimizing industrial pollution, and reducing energy consumption. To ensure safe and stable operation, and to mitigate safety risks and economic losses, it is crucial to monitor and detect any abnormalities in the process. One proposed method for monitoring and detecting two typical abnormal conditions in the roasting process is Principal Component Analysis (PCA). The analysis method utilizes industrial data collected from real roasting processes.

Keywords: Principal component analysis, process monitoring, zinc roasting process.

¹Trường Đại học Công nghiệp Quảng Ninh

²Trường Đại học Trung Nam, Trung Quốc

*Email: Vangdkh1@qui.edu.vn

Ngày nhận bài: 30/5/2023

Ngày nhận bài sửa sau phản biện: 22/9/2023

Ngày chấp nhận đăng: 15/10/2023

1. GIỚI THIỆU

Ngày nay, với sự phát triển của khoa học kỹ thuật, các quy trình công nghiệp đang dần chuyển đổi thành những quy trình thông minh. Nhiều quy trình công nghiệp hiện đại hóa được trang bị một số cảm biến thiết kế tinh vi để thu thập dữ liệu liên quan đến quy trình, nhằm phát hiện các lỗi tồn tại hoặc phát sinh trong quy trình, cũng như theo dõi trạng thái quy trình [1]. Các ngành công nghiệp chế biến chủ yếu bao gồm thép, kim loại màu, hóa dầu và vật liệu xây dựng là nền tảng kinh tế của nhiều quốc gia. Quy trình sản xuất quan trọng nhất của một nhà máy luyện kẽm là chuyển đổi kẽm sunfua cô đặc thành oxit kẽm. Vì quy trình rang kẽm là cung cấp nguyên liệu thô cho quá trình thủy luyện hạ nguồn, tính ổn định của nó là tiền đề để nâng cao hiệu quả của lò luyện kẽm [2].

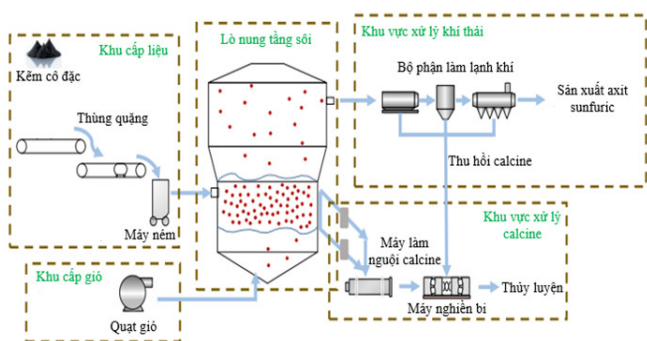
Trong quá trình vận hành, do sự đa dạng của thành phần nguyên liệu bao gồm thành phần tạp chất, cấp độ cô đặc, độ ẩm, thao tác sai của nhân viên vận hành và lỗi thiết bị, các điều kiện làm việc bất thường đôi khi xảy ra trong quy trình rang. Hơn nữa, sự thay đổi trong điều kiện làm việc sẽ làm tăng nồng độ sunfua dioxit (SO₂) của khí thải gây ô nhiễm môi trường nghiêm trọng. Những bất thường này có thể dẫn đến ngừng trệ sản xuất, ô nhiễm môi trường, rủi ro an toàn và các vấn đề về chất lượng [3]. Do đó, việc phát hiện các điều kiện làm việc bất thường của một quy trình công nghiệp là rất cần thiết. Trong bài báo này, nghiên cứu một phương pháp phân tích thành phần chính PCA (Principal Component Analysis) dựa trên dữ liệu công nghiệp để giám sát quá trình các điều kiện làm việc bất thường trong quy trình rang kẽm.

2. PHÂN TÍCH QUY TRÌNH VÀ THUẬT TOÁN PHÁT HIỆN BẤT THƯỜNG

2.1. Mô tả quy trình

Quy trình rang kẽm là một hệ thống sản xuất liên tục, phức tạp, chứa máy rang kẽm và một số hệ thống con phụ trợ như: hệ thống cấp liệu, máy phun khí, xử lý khí và làm mát calcine [4]. Kẽm cô đặc, từ thùng cô đặc được trộn hoàn toàn trong bộ nạp đĩa và ném vào máy rang kẽm thông qua

hai máy ném đai. Trong khi đó, không khí được bơm vào hộp gió của máy rang kềm thông qua máy thổi khí. Máy rang kềm là một máy rang tầng sôi điển hình, có thể được chia thành thành ba phần: ván tự do, giường tầng sôi, và hộp gió. Hầu hết các hạt rắn phản ứng với không khí trong tầng sôi, trong khi một số hạt mịn trong chất cô đặc bị không khí cuốn vào và phản ứng trong bảng tự do. Tầng sôi là vùng phản ứng chính, nơi hầu hết các nguyên tố kim loại bị oxy hóa thành oxit kim loại và lưu huỳnh được chuyển thành sunfua dioxide [5]. Sau một loạt các phản ứng, calcine được làm mát thông qua bộ làm mát sôi hóa và bộ làm mát xi lanh và sau đó được lưu trữ trong thùng calcine. Khí thải được làm sạch trong hệ thống loại bỏ bụi và gửi đến hệ thống lọc khí để tạo ra axit sunfuric. Trong quá trình này, nhiệt của khí thải được lò hơi sử dụng để tạo ra hơi nhiệt độ cao. Toàn bộ quy trình được thể hiện như trong hình 1 [6].



Hình 1. Quy trình nung kềm

Các biến quy trình được sử dụng để theo dõi các điều kiện làm việc của quy trình rang kềm được liệt kê trong bảng 1. Tất cả các biến được lấy mẫu mỗi phút một lần và được lưu trữ trong hệ thống DCS. Vì các chỉ số chất lượng được lấy mẫu ngoại tuyến và phân tích, cứ sau mười hai giờ thì chúng không được sử dụng để theo dõi các điều kiện làm việc bất thường trong quá trình rang. Bộ dữ liệu quy trình được sử dụng để phân tích dựa trên 6000 mẫu dữ liệu đã được chuẩn hóa.

Bảng 1. Các biến quy trình được sử dụng trong giám sát quá trình

Hệ thống con trong quy trình	Số	Biến quy trình (đơn vị)	Số	Biến quy trình (đơn vị)
Máy rang kềm	1	Nhiệt độ lò (°C)	2	Áp suất hộp gió (kPa)
	3	Nhiệt độ trên các tầng sôi (°C)	4	Nhiệt độ đáy của các tầng sôi (°C)
	5	Nhiệt độ trung bình của tầng sôi (°C)	6	Nhiệt độ khí thải (°C)
	7	Áp suất khí thải (kPa)	8	Lưu lượng nước tuần hoàn của ống nhúng (m³/h)
Hệ thống cấp liệu	9	Tốc độ nạp liệu của máy ném 1 (t/h)	10	Tốc độ nạp liệu của máy ném 2 (t/h)
Hệ thống khí nén	11	Khối lượng khí nén (Nm³/h)	12	Nhiệt độ khí nén (Nm³/h)
	13	Áp suất khí nén (kPa)		

Hệ thống xử lý khí	14	Chất lỏng ở các mức trong nồi hơi (m)	15	Lưu lượng cấp liệu nồi hơi (m³/s)	
	16	Nhiệt độ cấp liệu nồi hơi (°C)	17	Áp suất cấp liệu nồi hơi (kPa)	
	18	Áp suất lò hơi nhiệt thải (kPa)	19	Áp suất đầu vào bơm tuần hoàn (kPa)	
	20	Áp suất đầu ra bơm tuần hoàn (kPa)	21	Nhiệt độ đầu ra của bơm tuần hoàn (°C)	
	22	Áp suất đầu ra ống hơi (kPa)	23	Nhiệt độ đầu ra ống hơi (°C)	
	24	Tốc độ dòng chảy đầu ra của ống hơi (m³/s)	25	Tốc độ dòng chảy tuần hoàn của bơm tuần hoàn (m³/s)	
	26	Áp suất đầu ra của bơm cao áp A (kPa)	27	Áp suất đầu ra của bơm cao áp B (kPa)	
	28	Nhiệt độ khí thải đầu vào của lò hơi nhiệt thải (°C)	29	Áp suất đầu vào của Cottrell (kPa)	
	30	Nhiệt độ đầu vào Cottrell (°C)	31	Áp suất đầu ra của Cottrell (kPa)	
	32	Nhiệt độ đầu ra của Cottrell (°C)	33	Nhiệt độ đầu vào của cyclone (°C)	
	34	Áp suất đầu vào của cyclone (kPa)	35	Nhiệt độ đầu ra của quạt gió (°C)	
	36	Áp suất đầu ra của quạt gió (kPa)	37	Nồng độ sunfur dioxit của khí nhiên liệu (vol%)	
	Hệ thống làm nguội calcine	38	Nhiệt độ đầu vào của bộ làm mát xi lanh (°C)	39	Áp suất đầu vào của bộ làm mát xi lanh (kPa)
		40	Nhiệt độ đầu ra của bộ làm mát xi lanh (°C)	41	Nhiệt độ đầu ra của bộ làm mát sôi hóa 1 (°C)
		42	Nhiệt độ đầu ra của bộ làm mát sôi hóa 2 (°C)	43	Áp suất không khí của máy làm mát tầng sôi (kPa)
		44	Lưu lượng khí nén của bộ làm mát sôi hóa 1 (Nm³/h)	45	Lưu lượng khí nén của bộ làm mát sôi hóa 2 (Nm³/h)

2.2. Những bất thường điển hình trong quy trình rang kềm

Rang kềm là quy trình sản xuất đầu tiên nên trạng thái vận hành của nó quyết định tính ổn định của toàn nhà máy luyện kềm. Khi quy trình rang chạy ổn định, lượng khí thổi không đổi và tốc độ cấp liệu được điều chỉnh để giữ sự ổn định của nhiệt độ rang. Do sự thay đổi thường xuyên trong thành phần nguyên liệu như: loại cô đặc, hàm lượng silic, hàm lượng chì, độ ẩm và tình trạng hoạt động bất thường của một số thiết bị, quy trình rang không thể duy trì sự ổn định lâu dài, dẫn đến giảm sản xuất, các vấn đề về chất lượng, ô nhiễm môi trường [3]. Trong một số điều kiện nghiêm trọng, quy trình rang sẽ ngừng sản xuất, gây thiệt hại lớn về kinh tế. Dựa trên điều tra quy trình thực tế cho thấy, hai điều kiện làm việc bất thường điển hình của quy trình thường gặp là:

1. Thiêu kết tầng sôi: thiêu kết là một điều kiện làm việc bất thường phổ biến trong các quá trình luyện kim. Các tạp chất có điểm nóng chảy thấp trong kẽm nung (như silicon dioxide và chì sunfua) sẽ hợp nhất thành các nhóm với các vật liệu khác ở nhiệt độ nung cao. Khi có một lượng lớn tạp chất có điểm nóng chảy thấp trong tinh quặng kẽm, chất thiêu kết sẽ lắng đọng ở đáy tầng sôi và chặn các vòi trên tấm phân phối, dẫn đến trạng thái hóa lỏng kém của máy rang kẽm.

2. Rò rỉ hệ thống nước làm mát tuần hoàn: hệ thống nước làm mát tuần hoàn được sử dụng để hấp thụ nhiệt phản ứng của lò nung, cung cấp nhiệt cho nồi hơi. Khi rò rỉ xảy ra trong hệ thống nước làm mát tuần hoàn, lưu lượng nước làm mát tuần hoàn sẽ giảm và nhiệt độ nung sẽ tăng lên, dẫn đến giảm sản lượng và thiêu kết.

Hai rủi ro này cũng rất phổ biến trong các quy trình công nghiệp khác. Quá trình thiêu kết sẽ làm tắc nghẽn cổng xả của các lò phản ứng. Do đó, công nhân vận hành phải chấp nhận rủi ro về nhiệt độ cao để mở cổng xả. Sự rò rỉ sẽ gây ra sự gia tăng nhiệt độ nung. Một mặt, nó mang đến những rủi ro về an toàn cho hệ thống nồi hơi. Mặt khác, công nhân phải sửa chữa đường ống ở nhiệt độ cao dẫn đến một số nguy cơ ảnh hưởng đến sức khỏe. Do đó, việc phát hiện hai điểm bất thường này có thể làm tăng độ an toàn của quy trình và giảm thiểu rủi ro.

2.3. Giới thiệu thuật toán phát hiện bất thường PCA

PCA là một phương pháp giảm kích thước dữ liệu trong khi vẫn giữ được phần lớn các thông tin của dữ liệu ban đầu, phương pháp PCA trích xuất các thành phần chính trực giao theo tiêu chí phương sai tối đa, do đó làm giảm sự dư thừa dữ liệu [7]. Giả sử rằng ma trận dữ liệu của quy trình là $X \in R^{n \times m}$ với n mẫu và m biến. Trước khi phân tách PCA, cần chuẩn hóa dữ liệu để có được ma trận dữ liệu chuẩn hóa \bar{X} . Khi đó, ma trận hiệp phương sai của \bar{X} được tính như sau [8]:

$$\sum \frac{1}{n} \bar{X}^T \bar{X} \tag{1}$$

Cuối cùng, mô hình PCA có thể thu được bằng cách phân tách riêng của ma trận hiệp phương sai:

$$\bar{X} = TP^T + E \tag{2}$$

$$T = \bar{X}P \tag{3}$$

$$E = \bar{X}(I - PP^T) \tag{4}$$

Trong đó, $T \in R^{n \times p}$ và $P \in R^{m \times p}$ lần lượt là ma trận điểm và ma trận tải trọng, $E \in R^{n \times m}$ là ma trận dư và p là số lượng thành phần chính (PCs) với $p \leq m$. Số lượng PCs được xác định bởi tỷ lệ phần trăm phương sai tích lũy (CPV) và được tính toán theo công thức (5), mang lại tỷ lệ phần trăm phương sai chiếm bởi thành phần chính đầu tiên (p-PC). CPV càng lớn thì p-PC đầu tiên càng dễ giải thích thông tin về toàn bộ quá trình.

$$p = \operatorname{argmin}_{1 \leq p \leq m} \left(\frac{\sum_{j=1}^p \lambda_j}{\sum_{i=1}^m \lambda_i} \cdot 100\% \right) \geq 85\% \tag{5}$$

Trong quá trình giám sát, số liệu thống kê T^2 và thống kê SPE (hay còn được gọi là thống kê Q) được sử dụng trong thuật toán PCA để giám sát xem có xảy ra bất thường hay không, hai loại thống kê này giả định dữ liệu bình thường của quá trình tuân theo phân bố Gaussian, khi đó giới hạn kiểm soát của thống kê T^2 và SPE được tính toán như sau:

$$T^2 = x^T P \Lambda P^T x \leq T_\alpha^2 = \frac{p(n-1)}{n-p} F_\alpha(p, n-p) \tag{6}$$

$$SPE = x^T (I - PP^T) x \leq SPE_\alpha = g \chi_{n,\alpha}^2 \tag{7}$$

Trong đó:

$\Lambda = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_p\}$ bao gồm giá trị riêng p đầu tiên của Σ , $F_\alpha(p, n-p)$ là giá trị tới hạn của phân phối F với bậc tự do là p và (n-p). Trong phương trình (7) $\chi_{n,\alpha}^2$ là giá trị tới hạn của phân phối Chi-square với bậc tự do của $n = \frac{2a^2}{b}$ và

trọng lượng của $g = \frac{a}{2b}$ với a và b lần lượt là giá trị trung bình và phương sai của thống kê SPE, α là độ tin cậy của hai phân phối.

T_α^2 là giới hạn kiểm soát của thống kê T^2 .

SPE_α là giới hạn kiểm soát của thống kê SPE.

Thống kê SPE dùng để đo sự thay đổi của ma trận dữ liệu mẫu trong không gian dư, miêu tả mức độ thay đổi mối tương quan giữa các biến quy trình thông thường.

Vậy, để đánh giá xem quy trình làm việc bình thường hay bất thường thì ta dựa vào các điều kiện sau để đánh giá:

Nếu $\begin{cases} T^2 < T_\alpha^2 \\ SPE < SPE_\alpha \end{cases}$ thì quy trình làm việc là bình thường

Nếu $\begin{cases} T^2 < T_\alpha^2 \\ SPE > SPE_\alpha \end{cases}$ thì quy trình làm việc là bất thường

Nếu $\begin{cases} T^2 > T_\alpha^2 \\ SPE < SPE_\alpha \end{cases}$ thì quy trình làm việc là bình thường

Nếu $\begin{cases} T^2 > T_\alpha^2 \\ SPE > SPE_\alpha \end{cases}$ thì quy trình làm việc là bất thường

3. KẾT QUẢ VÀ THẢO LUẬN

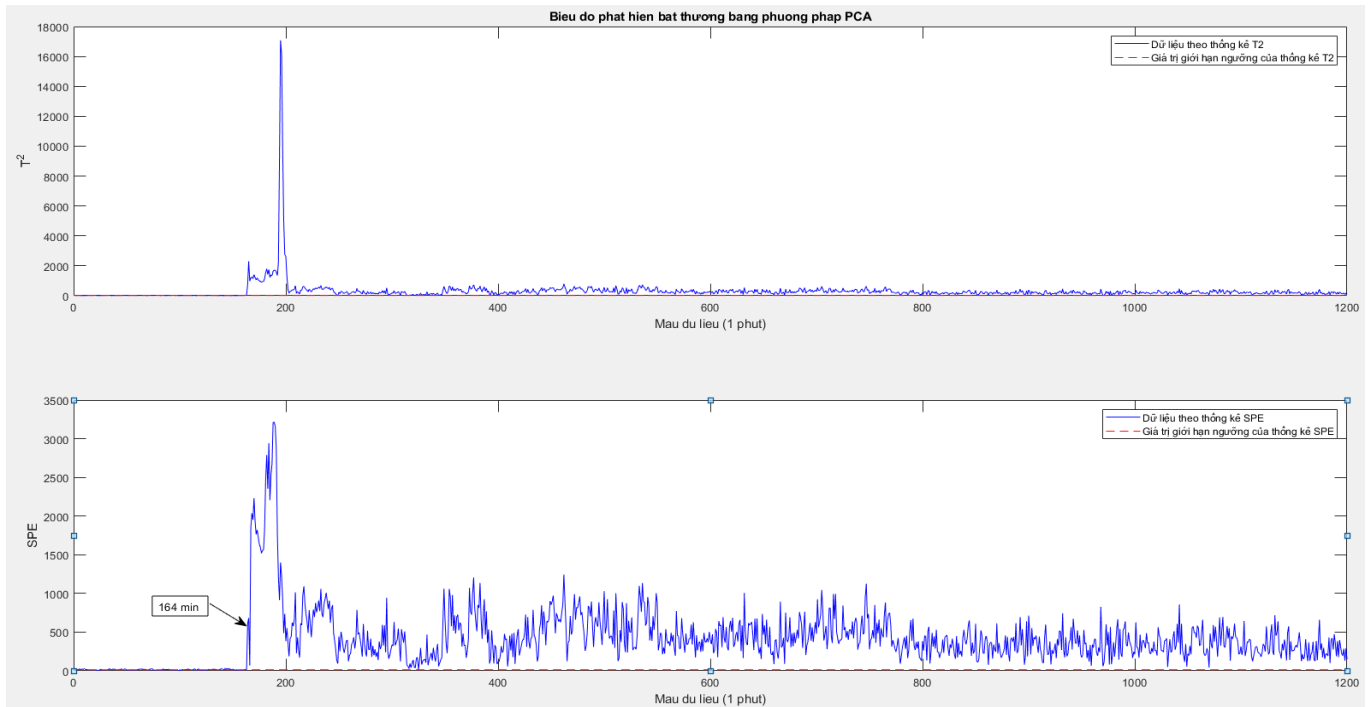
3.1. Ứng dụng matlab simulink xây dựng chương trình phát hiện bất thường

Căn cứ vào mẫu dữ liệu thực tế thu thập được thông qua các cảm biến và các công thức trong thuật toán PCA ta xây dựng chương trình phát hiện bất thường trên M-File. Nội dung chương trình M-File được thể hiện trong phần phụ lục.

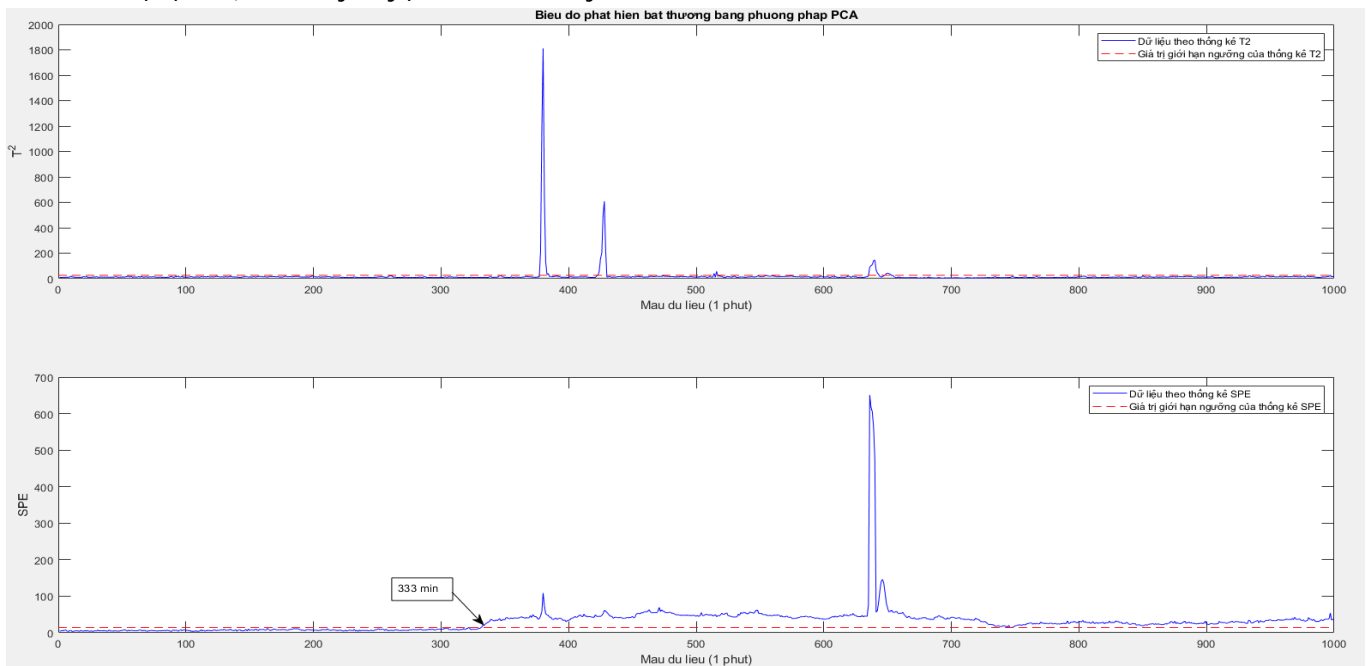
3.2. Kết quả phát hiện bất thường dựa trên phân tích dữ liệu

Phát hiện bất thường của quá trình thiêu kết tầng sôi

Các mẫu dữ liệu sử dụng trong nghiên cứu này, được thu thập thực tế tại nhà máy luyện kẽm, thông qua các thiết bị cảm biến của quy trình rang kẽm, thống kê lại các mẫu bình thường và mẫu lỗi của quy trình.



Hình 2. Kết quả phát hiện bất thường trong quá trình thiêu kết tăng sôi



Hình 3. Kết quả phát hiện bất thường trong rò rỉ hệ thống nước làm mát tuần hoàn

Để xác minh hiệu quả của phương pháp PCA trong việc giám sát quá trình thiêu kết tăng sôi, 2600 mẫu trước khi xảy ra thiêu kết được chọn làm tập huấn luyện và 1200 mẫu chứa dữ liệu lỗi được chọn làm tập kiểm tra, trong đó lỗi thiêu kết xảy ra rõ nhất tại mẫu thứ 164. Dựa vào quy tắc CPV, số lượng thành phần chính của phương pháp PCA được chọn là 10, có thể giải thích 86,59% thông tin của toàn bộ quá trình. Giới hạn tin cậy được đặt là 0,99 khi xác định giới hạn kiểm soát của PCA. Các kết quả giám sát được thể hiện trong hình 2.

Phát hiện bất thường của rò rỉ hệ thống nước làm mát tuần hoàn

Tương tự, để xác minh hiệu quả của phương pháp PCA trong việc giám sát rò rỉ của hệ thống nước làm mát tuần hoàn, 6000 mẫu trước khi xảy ra lỗi rò rỉ được chọn làm tập huấn luyện và 1000 mẫu chứa dữ liệu lỗi được chọn làm tập kiểm tra, dựa trên phân tích thống kê T^2 và SPE đã phát hiện lỗi ở vị trí mẫu thứ 333. Dựa trên quy tắc CPV, số thành phần chính của phương pháp PCA được chọn là 11, có thể giải thích 85,96% thông tin của toàn bộ quá trình. Giới hạn tin cậy được đặt là 0,99 khi xác định giới hạn kiểm soát của PCA. Các kết quả giám sát được thể hiện trong hình 3.

Từ việc phân tích thành phần chính của các mẫu thử nghiệm được thể hiện trong hình 2 và 3, có thể nhận thấy thành phần chính 10 và 11 có sự thay đổi rõ rệt khi xảy ra hiện tượng thiêu kết tầng sôi và rò rỉ hệ thống nước làm mát tuần hoàn. Hay nói một cách khác, dựa vào điều kiện đánh giá quy trình làm việc bình thường hay bất thường, chúng ta quan sát thấy tại mẫu 164 dữ liệu theo thống kê SPE lớn hơn giá trị giới hạn ngưỡng của thống kê SPE. Tương tự, tại mẫu 333 của hệ thống nước làm mát tuần hoàn cũng cho chúng ta thấy dữ liệu thống kê SPE lớn hơn giá trị giới hạn ngưỡng của thống kê SPE.

4. KẾT LUẬN

Phương pháp phân tích thành phần chính (PCA) là một phương pháp thống kê đa biến để kiểm tra mối tương quan giữa các biến. PCA xây dựng một không gian mới ít chiều hơn nhưng lại có khả năng đại diện cho bộ dữ liệu gốc, bảo toàn lượng thông tin nhiều nhất có thể, giúp cho việc phân tích đánh giá xử lý dữ liệu được đơn giản và thuận tiện hơn. Dựa trên quy tắc CPV có thể xác định được các thành phần chính của phương pháp và tỷ lệ đóng góp tích lũy của các thành phần chính (tỷ lệ phần trăm lượng thông tin mà các thành phần chính đó bao hàm). Phạm vi chấp nhận được của tỷ lệ đóng góp tích lũy của các thành phần chính là $\geq 85\%$. Bên cạnh những ưu điểm kể trên của phương pháp PCA thì nó cũng có hạn chế nhất định, kết quả phát hiện bất thường sẽ phụ thuộc vào việc lựa chọn điều kiện dữ liệu ban đầu nghĩa là hai thống kê T^2 và SPE giả định dữ liệu của quá trình tuân theo phân bố Gaussian.

Thống kê T^2 và SPE của phương pháp PCA có thể phát hiện sự xuất hiện của các điều kiện bất thường trong quá trình rang kẽm. Việc phát hiện lỗi kịp thời, chính xác có thể giúp cho người quản lý có những kế hoạch vận hành an toàn và ổn định, đảm bảo chất lượng sản phẩm và giảm thiểu ô nhiễm của nhà máy trong quá trình sản xuất. Trong tương lai, nhóm tác giả sẽ tiếp tục nghiên cứu sâu hơn các thuật toán giám sát có thể phát hiện sớm, chính xác các điều kiện bất thường của quy trình sản xuất kẽm.

Phụ lục. Chương trình M-File

```
%%Nhập du lieu
load('data1.mat')
a=Train_1(1:end,:);
b=Test_1(1:1000,:);
%Xu ly du lieu huan luyen
a0=a(1:end,3:17);
X_mean1 = mean(a0); %Tinh gia tri trung binh cua Xtrain theo cac cot
X_std1 = std(a0); %Tinh do lech chuan
[coeff,score,latent,tsquared,explained] = pca(zscore(a0));
Xtrain=[score(1:end,1:3) a(1:end,1:2) a(1:end,18:end)];
%Xu ly du lieu thu nghiem
b0=b(1:end,3:17);
```

```
[m_b0,n_b0]=size(b0);
b00=(b0-
repmat(X_mean1,m_b0,1))./repmat(X_std1,m_b0,1);
b01=b00*coeff;
Xtest =[b01(1:end,1:3) b(1:end,1:2) b(1:end,18:end)];
Xtest = double(Xtest);
%%
%Xu ly chuan hoa
X_mean = mean(Xtrain); %Tinh gia tri trung binh Xtrain theo cot
X_std = std(Xtrain); %Tinh do lech chuan
[X_row,X_col] = size(Xtrain); %Tim so hang va so cot cua ma tran Xtrain
Xtrain=(Xtrain-
repmat(X_mean,X_row,1))./repmat(X_std,X_row,1);
%%
%Tim ma tran hiep phuong sai
sigmaXtrain = cov(Xtrain);
[T,lamda] = eig(sigmaXtrain);
D = flipud(diag(lamda));
num_pc = 1;
while sum(D(1:num_pc))/sum(D) < 0.85
num_pc = num_pc + 1;
end
%Tim vecto dac trung tuong ung voi cac tri rieng lamda
P = T(:,X_col-num_pc+1:X_col);
%%
% Gioi han kiem soat thong ke T2 la 95% voi do tin cay 99%
T2UCL1=num_pc*(X_row-
1)*(X_row+1)*finv(0.95,num_pc,X_rownum_pc)/(X_row*(X_row - num_pc));
for i = 1:3
theta(i) = sum((D(num_pc+1:X_col)).^i);
end
h0 = 1 - 2*theta(1)*theta(3)/(3*theta(2)^2);
ca = norminv(0.95,0,1);
QUCL = theta(1)*(h0*ca*sqrt(2*theta(2)))/theta(1) + 1 + theta(2)*h0*(h0 - 1)/theta(1)^2^(1/h0);
%%
%Su dung uoc tinh mat do hat nhan de xac dinh gioi han kiem soat
c=0.99;
T0= [];
Q0 = [];
[x1,y1]=size(P*P');
l1=eye(x1,y1);
```

```

for i = 1:X_row
    T0=[T0
        Xtrain(i,:)*P*inv(lamda(X_col-
num_pc+1:X_col,X_col-num_pc+1:X_col))*P*Xtrain(i,:)];
    Q0=[Q0 Xtrain(i,:)*(I1 - P*P)*Xtrain(i,:)];
end
% Gioi han kiem soat thong ke T2 voi do tin cay 99%
[T2UCL1,T2xi]=ksdensity(abs(T0),c,'function','icdf');
% Gioi han kiem soat thong ke SPE voi do tin cay 99%
[QUCL,SPExi]=ksdensity(Q0,c,'function','icdf');
%%
n = size(Xtest,1);
Xtest=(Xtest-repmat(X_mean,n,1))./repmat(X_std,n,1);
[r,y] = size(P*P');
I = eye(r,y);
T2 = [];
Q = [];
for i = 1:n
    T2=[T2
        Xtest(i,:)*P*inv(lamda(X_col-
num_pc+1:X_col,X_col-num_pc+1:X_col))*P*Xtest(i,:)];
    Q=[Q Xtest(i,:)*(I - P*P)*Xtest(i,:)];
end
%%
%Ve do thi
Figure
subplot(2,1,1);
plot(1:n,T2,'b-');
title('Bieu do phat hien bat thuong bang phuong phap
PCA');
xlabel('Mau du lieu (1 phut)');
ylabel('T^2');
hold on;
line([0,n],[T2UCL1,T2UCL1],'LineStyle','--','Color','r');
subplot(2,1,2);
plot(1:n,Q,'b-');
xlabel('Mau du lieu (1 phut)');
ylabel('SPE');
hold on;
line([0,n],[QUCL,QUCL],'LineStyle','--','Color','r');
% figure
% subplot(2,1,1)
% plot(Train_1(:,22))
% subplot(2,1,2)
% plot(Train_1)
%
```

```

% figure
% subplot(2,1,1)
% plot(Train_1(:,22))
% subplot(2,1,2)
% plot(Train_1)
%
```

TÀI LIỆU THAM KHẢO

- [1]. L. Ming, J. Zhao, 2017. *Review on chemical process fault detection and diagnosis*. in 2017 6th International Symposium on Advanced Control of Industrial Processes (AdCONIP), IEEE, pp. 457–462. doi: 10.1109/ADCONIP.2017.7983824.
- [2]. B. Sun, C. Yang, Y. Wang, W. Gui, I. Craig, L. Olivier, 2020. *A comprehensive hybrid first principles/machine learning modeling framework for complex industrial processes*. J Process Control, vol. 86, pp. 30–43, doi: 10.1016/j.jprocont.2019.11.012.
- [3]. B. Zhang, C. Yang, H. Zhu, P. Shi, W. Gui, 2018. *Controllable-Domain-Based Fuzzy Rule Extraction for Copper Removal Process Control*. IEEE Transactions on Fuzzy Systems, vol. 26, no. 3, pp. 1744–1756, doi: 10.1109/TFUZZ.2017.2751000.
- [4]. J. Fan, Y. Wang, 2014. *Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis*. Inf Sci (N Y), vol. 259, pp. 369–379, doi: 10.1016/j.ins.2013.06.021.
- [5]. X. Wang, U. Kruger, G. W. Irwin, 2005. *Process Monitoring Approach Using Fast Moving Window PCA*. Ind Eng Chem Res, vol. 44, no. 15, pp. 5691–5702, doi: 10.1021/ie048873f.
- [6]. Liang Huiping, et al., 2021. *A hybrid first principles and data-driven process monitoring method for zinc smelting roasting process*. IEEE Transactions on Instrumentation and Measurement 70: 1–14.
- [7]. Harrou Fouzi, et al., 2016. *Ozone measurements monitoring using data-based approach*. Process Safety and Environmental Protection 100: 220–231
- [8]. Feng Zhenxiang, et al., 2022. *Process monitoring of abnormal working conditions in the zinc roasting process with an ALD-based LOF-PCA method*. Process Safety and Environmental Protection 161: 640–650.

AUTHORS INFORMATION

Do Van Vang^{1,2}, Nguyen Thi Thuong Duyen¹, Tran Thi Thu Lan¹

¹Quang Ninh University of Industry, Vietnam

²Central South University, China