

ÁP DỤNG BIỂU ĐỒ THỐNG KÊ ĐA BIẾN HOTELLING T² TRONG KIỂM SOÁT QUÁ TRÌNH SẢN XUẤT

Bùi Thu Hà, Nguyễn Thị Thanh Hiền, Ngô Thanh An*

Trường Đại học Công Thương Thành phố Hồ Chí Minh

*Email: annt@huit.edu.vn

Ngày nhận bài: 29/5/2023; Ngày chấp nhận đăng: 25/9/2023

TÓM TẮT

Biểu đồ thống kê đa biến là một công cụ phổ biến được áp dụng trong quản lý chất lượng để kiểm soát quá trình sản xuất. Trong bài báo này, biểu đồ thống kê đa biến Hotelling T² được sử dụng để đánh giá quá trình tuyến nổi chì - kẽm với dữ liệu tham khảo đã được công bố. Phương pháp xây dựng biểu đồ Hotelling T² được thể hiện qua từng giai đoạn. Kết quả các giá trị giới hạn kiểm soát và điểm ngoại lệ được xác định thông qua phương pháp thống kê Hotelling kết hợp với phương pháp phân tích thành phần chính PCA do phần mềm Minitab 21 thực hiện. Bên cạnh đó, các thông số chính gây ra điểm ngoại lệ cũng được phân tích bằng PCA và phương pháp phân tách T². Tương tự với biểu đồ thống kê đơn biến X, biểu đồ Hotelling T² cũng xác định “chiều cao lớp lỏng” và “giá trị pH” là hai thông số ảnh hưởng chính đến quá trình tuyến nổi. Tuy nhiên, số điểm ngoại lệ trên biểu đồ Hotelling T² ít hơn, thể hiện được sự tương quan của các thông số ảnh hưởng đến hiệu quả của quy trình. Qua đó cho thấy, nếu quá trình sản xuất có nhiều biến tác động và tương quan thì kiểm soát bằng biểu đồ đa biến Hotelling T² sẽ là công cụ hữu ích để đánh giá toàn diện hơn.

Từ khóa: Thống kê đa biến, Hotelling T², điểm ngoại lệ.

1. GIỚI THIỆU

Kiểm soát quy trình bằng thống kê (SPC) là một trong những công cụ kiểm soát chất lượng thông qua giám sát và đánh giá quy trình sản xuất bằng các phương pháp thống kê khác nhau. SPC sử dụng phép phân tích, thống kê các dữ liệu liên quan như nguyên liệu, điều kiện vận hành, thông số làm việc... trong sản xuất để xác định được các yếu tố ngoài tầm kiểm soát, hay các nguyên nhân có thể gây ra sự thay đổi trong quy trình và cảnh báo cho người vận hành. Từ đó, các lỗi phát hiện trong quá trình sản xuất được khắc phục trước khi sản xuất lượng lớn sản phẩm không phù hợp [1, 2].

Kỹ thuật kiểm soát quy trình phổ biến nhất là lập biểu đồ kiểm soát được phát triển bởi Walter Shewhart vào những năm 1920. Thời gian đầu, các biểu đồ kiểm soát đơn biến như: biểu đồ X (giá trị trung bình), biểu đồ R (khoảng biến đổi), biểu đồ S (độ lệch chuẩn) được áp dụng. Các biểu đồ này có ưu điểm đơn giản, dễ thực hiện nhưng có hạn chế là chỉ kiểm soát được một thông số vận hành hay một đặc tính chất lượng trên từng biểu đồ. Tuy nhiên, quy trình sản xuất phức tạp sẽ có nhiều thông số tác động đến hiệu quả của quy trình. Bên cạnh đó, công nghệ thu thập các dữ liệu của quá trình sản xuất ngày càng phát triển. Do đó, các phương pháp thống kê đa biến (MSPC) ra đời và phát triển mạnh mẽ, tiêu biểu như Hotelling T², MEWMA, MCUSUM [1-3]. Ưu điểm nổi bật của MSPC là phân tích đồng thời nhiều biến số, thể hiện được toàn cảnh về hệ thống sản xuất, xác định sự tương tác giữa các thông số và phát hiện lỗi trong hệ thống do thông số nào tác động. Điều này giúp cho người vận hành hiểu toàn

diện hơn quy trình, có thể đưa ra quyết định đúng đắn về sự cải tiến và điều chỉnh quy trình để đem hiệu quả tốt hơn.

Trong số các biểu đồ đa biến, Hotelling T² được biết đến nhiều nhất và ứng dụng phổ biến cho các quy trình công nghiệp do đơn giản và dễ thực hiện [1]. Thống kê này là phần mở rộng của kiểm định T-Student, được áp dụng để kiểm định giá trị trung bình của một tổng thể [2, 3]. Trên thế giới, các nghiên cứu ứng dụng Hotelling T² được công bố đa dạng và rộng rãi như trong phân tích dữ liệu hóa học, sinh học, điện hóa, sản xuất ván gỗ trong lâm nghiệp, hệ thống cung cấp nước [4-6]. Nhìn chung, các kết quả từ biểu đồ Hotelling T² đều cung cấp số liệu đáng giá cho việc tìm ra yếu tố gây bất lợi cho các quá trình. Thực tế là, trên biểu đồ Hotelling T² chỉ thể hiện vị trí điểm ngoại lệ mà không cung cấp được thông số trong quy trình gây ra điểm ngoại lệ này. Do đó, phương pháp phân tích thành phần chính PCA hay phương pháp phân tách T² thường được sử dụng kết hợp với biểu đồ kiểm soát Hotelling T² nhằm xác định thông số ảnh hưởng đến điểm ngoại lệ [5-8].

Trong khi đó, ở trong nước, việc áp dụng biểu đồ đa biến này trong thực tế còn khá mới, chưa có nhiều công bố so với biểu đồ đơn biến. Do đó, trong bài báo này, biểu đồ Hotelling T² được sử dụng để đánh giá dữ liệu quá trình sản xuất tham khảo là quá trình tuyển nổi [9]. Quá trình này được thực hiện trong bể tuyển nổi chì - kẽm bao gồm các thông số chính: tốc độ nhập liệu, giá trị pH, lượng CuSO₄, và chiều cao lớp lỏng. Các bước tiến hành, xây dựng biểu đồ Hotelling T² được trình bày rõ. Bên cạnh đó, phương pháp thành phần chính PCA và phân tách T² cũng được sử dụng để kết hợp với biểu đồ Hotelling T² trong việc xác định thông số gây điểm ngoại lệ. Hơn nữa, sự khác biệt giữa biểu đồ đa biến Hotelling T² và biểu đồ đơn biến X cũng được so sánh thông qua các vị trí và thông số của quy trình tuyển nổi thể hiện trên biểu đồ.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Lý thuyết phương pháp Hotelling T²

Biểu đồ kiểm soát Hotelling T² được xây dựng dựa trên đánh giá vector trung bình của một quá trình đa biến theo phương pháp thống kê của Hotelling. Trong đó, T² là khoảng cách có trọng số từ quan sát đến trung tâm của phân bố được tính toán theo công thức (1) [2, 3, 6, 10].

$$T^2 = n(X - \bar{X})' S^{-1} (X - \bar{X}) \quad (1)$$

Với n là kích thước mẫu; \bar{X} là vectơ trung bình mẫu; S là ma trận hiệp phương sai mẫu.

Biểu đồ kiểm soát Hotelling T² là một dạng của phương pháp MSPC có cách thức hoạt động tương tự nhau. Quá trình kiểm soát phân chia thành 2 giai đoạn. Giai đoạn 1, mô hình tham chiếu được xây dựng dựa trên dữ liệu khi quy trình hoạt động ổn định, có kiểm soát để ước lượng các tham số kiểm soát và chắc chắn mẫu mang tính thống kê. Giai đoạn 2, các giới hạn kiểm soát được xác định dựa trên tham số kiểm soát ở giai đoạn 1 sẽ được sử dụng để đánh giá dữ liệu vận hành (tương lai) và phát hiện điểm ngoại lệ [3, 10]. Từ các điểm ngoại lệ này sẽ xác định được thông số gây ra lỗi, tìm hiểu nguyên nhân sự cố để loại trừ.

Đối với các mẫu đơn lẻ $n = 1$, ở giai đoạn 1, giới hạn kiểm soát trên UCL được tính theo công thức (2), trong khi giới hạn kiểm soát dưới LCL bằng không

$$UCL = \frac{(m-1)^2}{m} \beta_{\alpha, p/2, (m-p-1)/2} \quad (2)$$
$$LCL = 0$$

Trong giai đoạn 2, giới hạn kiểm soát trên được tính theo công thức (3):

$$UCL = \frac{p(m+1)(m-1)}{m(m-p)} F_{\alpha,p,m-p} \quad (3)$$

Trong đó: m là số mẫu; p là số thông số; $F_{\alpha,p,m-p}$ là phân phối F ứng với mức xác suất α , p thông số, độ tự do $m-p$.

Đối với các mẫu theo nhóm, giới hạn kiểm soát trên trong giai đoạn 1 được tính theo công thức (4), giới hạn kiểm soát dưới cũng bằng không.

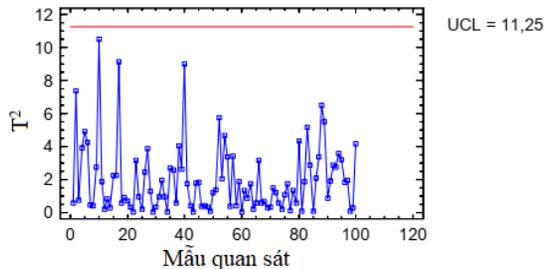
$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \quad (4)$$

$$LCL = 0$$

Trong giai đoạn 2, giới hạn kiểm soát trên được tính theo công thức (5).

$$UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \quad (5)$$

Hình 1 là một ví dụ về biểu đồ Hotelling T^2 với trục tung là giá trị T^2 và trục hoành là số mẫu. Đồ thị thể hiện vị trí của các mẫu (điểm) so với giá trị giới hạn trên UCL. Những điểm nằm dưới giá trị UCL được xem là ổn định và ngược lại là điểm ngoại lệ.



Hình 1. Biểu đồ kiểm soát đa biến Hotelling T^2 [2]

2.2. Xây dựng biểu đồ thống kê đa biến Hotelling T^2

Như đề cập ở trên, khi áp dụng biểu đồ đa biến Hotelling T^2 để kiểm soát một quá trình cần thực hiện hai giai đoạn riêng biệt. Giai đoạn 1, các tham số kiểm soát được ước tính trên dữ liệu lịch sử. Giai đoạn 2, giới hạn kiểm soát được tính dựa trên tham số ước lượng từ kết quả của giai đoạn 1 và áp dụng vào quy trình vận hành. Theo tài liệu [11], số lượng mẫu trong giai đoạn 1 lớn thì ước lượng giới hạn kiểm soát giai đoạn 2 hiệu quả.

Trong bài báo này, các thông số quá trình tuyến nổi được thu nhận từng thời điểm như là mẫu riêng lẻ và hai giai đoạn quy trình thực hiện như sau:

- Giai đoạn 1: Xây dựng biểu đồ tham chiếu bằng 150 số liệu lịch sử để thiết lập tập dữ liệu kiểm soát.

- Giai đoạn 2: Đánh giá quá trình đang vận hành bằng 50 số liệu được áp dụng minh họa cho phương pháp đánh giá, kiểm soát quá trình.

2.3. Phương pháp thực hiện

Toàn bộ dữ liệu được xử lý bằng phần mềm Minitab 21 [12] và phần mềm SPSS [13], trong đó bao gồm:

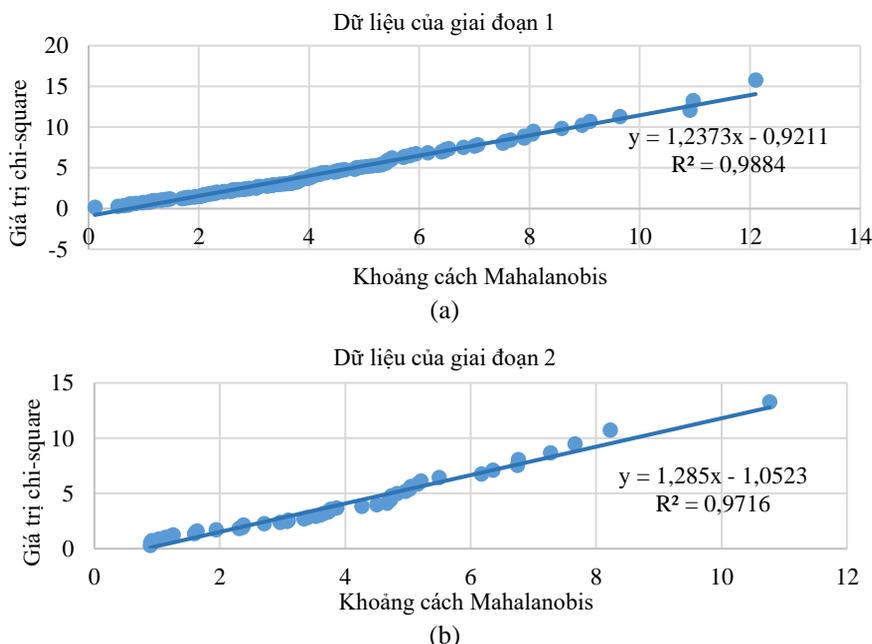
- Kiểm tra dữ liệu tuân theo phân phối chuẩn đa biến để tăng độ tin cậy của phương pháp kiểm soát bằng Hotelling T² [2];
- Phân tích thành phần chính PCA, xác định số thành phần chính sao cho có tỉ lệ tích lũy trên 80% [7];
- Xây dựng biểu đồ Hotelling T² của các thành phần chính được chọn theo hai giai đoạn.
- Xác định thông số tác động gây ra điểm ngoại lệ bằng phân tích thành phần chính thông qua giá trị p và phương pháp phân tách T².
- Xây dựng biểu đồ đơn biến X cho từng thông số để so sánh với biểu đồ đa biến Hotelling T².

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kiểm tra dữ liệu tuân theo phân bố chuẩn đa biến

Phương pháp kiểm tra dữ liệu tuân theo phân phối chuẩn được thực hiện bằng cách tính toán và sắp xếp theo thứ tự các giá trị khoảng cách Mahalanobis (MD) bằng phần mềm SPSS ứng với các số liệu của tập dữ liệu cần nghiên cứu. Sau đó, giá trị chi-square (CHISQ) tương ứng với các giá trị Mahalanobis được xác định. Đồ thị biểu diễn mối tương quan giữa MD và CHISQ tuân theo quy luật dạng đường thẳng thì dữ liệu được xem tuân theo phân phối chuẩn đa biến [2, 13].

Hình 2(a) và 2(b) thể hiện mối tương quan giữa MD và CHISQ của tập dữ liệu trong giai đoạn 1 và 2. Kết quả cho thấy có một mối quan hệ tuyến tính giữa các MD và CHISQ với hệ số tương quan R² cao (0,9884 và 0,9716). Như vậy dữ liệu đã tuân theo phân phối chuẩn đa biến.

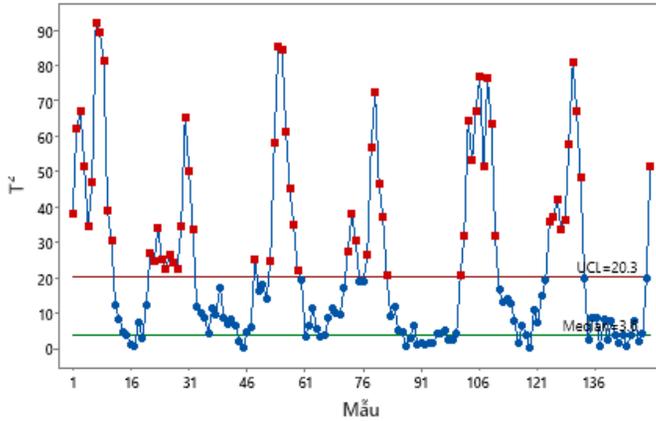


Hình 2. Biểu đồ mô tả sự tương quan giữa khoảng cách Mahalanobis và giá trị chi-square của tập dữ liệu trong: giai đoạn 1 (a) và giai đoạn 2 (b)

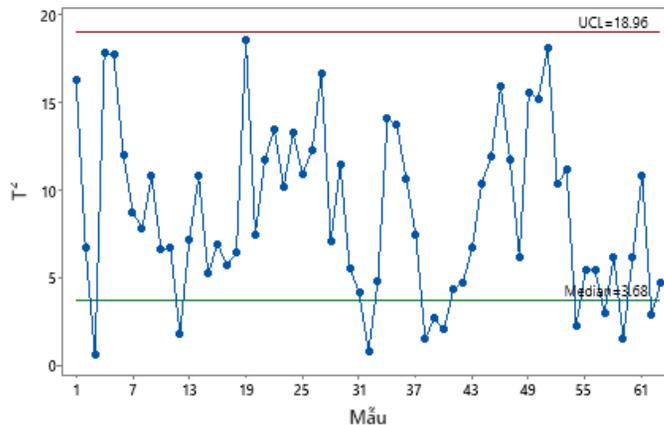
3.2. Xây dựng biểu đồ kiểm soát Hotelling T² trong giai đoạn 1

Với 150 dữ liệu lịch sử thu nhận ở từng thời điểm cho cả bốn thông số (tốc độ nhập liệu, giá trị pH, lượng CuSO₄, và chiều cao lớp lỏng) của quá trình tuyển nổi được sử dụng để phân tích thành phần chính PCA. Việc sử dụng phương pháp PCA nhằm giảm bớt các thông số không

cần thiết và đồng thời nhận biết thông số nào tác động lớn đến quy trình. Kết quả ba thành phần chính là PC1, PC2, PC3 với tỉ lệ tích lũy là 85,2% được chọn để xây dựng biểu đồ T^2 như Hình 3. Theo quan sát, 87 điểm ngoại lệ được xác định và được loại bỏ để hình thành biểu đồ tham chiếu. Hình 4 cho thấy biểu đồ T^2 tham chiếu đã được làm sạch, các điểm đều nằm dưới giá trị UCL. Như vậy dữ liệu trong giai đoạn này hoàn toàn được kiểm soát, và các tham số kiểm soát bộ dữ liệu này được dùng để đánh giá dữ liệu quy trình vận hành.



Hình 3. Biểu đồ Hotelling T^2 của 150 dữ liệu lịch sử



Hình 4. Biểu đồ Hotelling T^2 tham chiếu

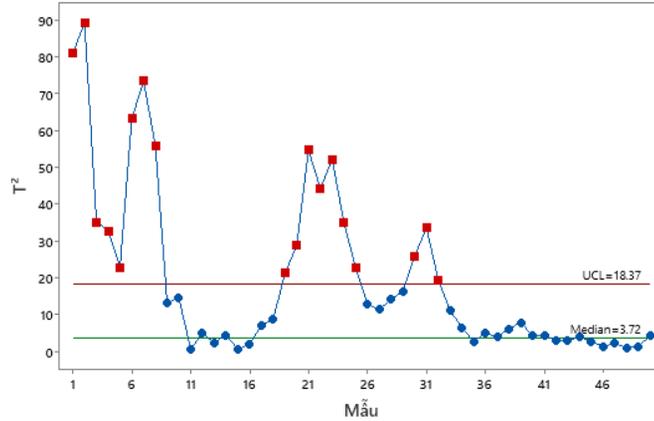
Bảng 1 tóm tắt các giá trị thu được giai đoạn 1. Ban đầu, 150 dữ liệu lịch sử bao gồm 87 điểm ngoại lệ (nằm trên giá trị UCL) và 63 điểm làm việc ổn định (nằm dưới giá trị UCL) có giá trị trung bình và UCL lần lượt là 3,60 và 20,30. Sau khi loại bỏ các điểm bất thường, bộ tham chiếu được xây dựng với giá trị trung bình tăng lên 3,68 và UCL giảm xuống còn 18,96.

Bảng 1. Tóm tắt các giá trị trên biểu đồ T^2 của giai đoạn 1

Số dữ liệu	Trung bình (Median)	Giới hạn trên (UCL)	Số điểm ngoại lệ được loại bỏ	Số điểm còn lại
Ban đầu-150	3,60	20,30	87	63
Đã kiểm soát- 63	3,68	18,96	0	63

3.3. Xây dựng biểu đồ kiểm soát Hotelling T^2 trong giai đoạn 2

Hình 5 là biểu đồ T^2 của các thành phần chính PC1, PC2, PC3 của 50 dữ liệu vận hành. Biểu đồ này có giá trị trung bình, giới hạn trên UCL lần lượt là 3,72 và 18,37 và xác định được 18 điểm ngoại lệ. Kết quả tính UCL được thu nhận dựa trên công thức (3) trên cơ sở số mẫu ước lượng có kiểm soát giai đoạn 1. Dữ liệu giai đoạn 2 được tóm tắt trong Bảng 2.



Hình 5. Biểu đồ T^2 của 50 dữ liệu trong giai đoạn 2

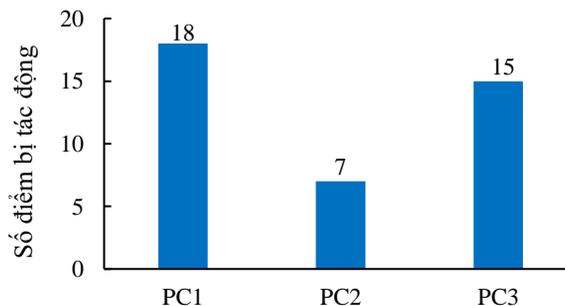
Bảng 2. Tóm tắt các thông số trên biểu đồ T^2 của giai đoạn 2

Số dữ liệu	Trung bình (Median)	Giới hạn trên (UCL)	Số điểm trong giới hạn kiểm soát	Số điểm ngoại lệ	Các điểm ngoại lệ
50	3,72	18,37	32	18	1-8, 19-25, 30-32

18 điểm ngoại lệ được xác định cụ thể là điểm 1-8, 19-25 và 30-32. Tuy nhiên, để có thông tin về thông số quy trình ảnh hưởng đến các điểm ngoại lệ này thì cần sử dụng các phương pháp phân tích thích hợp.

3.4. Xác định nguyên nhân gây ra điểm ngoại lệ trong giai đoạn 2

Hai phương pháp được lựa chọn để xác định thông số gây lỗi hệ thống là phân tích thành phần chính PCA và phân tách T^2 . Đối với phương pháp PCA, dựa trên giá trị p của các thành phần chính để đánh giá sự tác động của từng thành phần lên điểm ngoại lệ được thể hiện trong Bảng 3. Nếu giá trị p càng nhỏ thì ảnh hưởng của thành phần tương ứng càng lớn. Trong các biến thành phần chính PC1, PC2, PC3, có thể thấy PC1 ảnh hưởng nhiều nhất, tác động đến tất cả 18 điểm ngoại lệ, PC3 ảnh hưởng đến 15 điểm, còn PC2 chỉ gây ảnh hưởng đến 7 điểm (Hình 6).



Hình 6. Số lượng điểm ngoại lệ bị tác động bởi các thành phần chính PC1, PC2, PC3

Bảng 3. Kết quả các điểm ngoại lệ trên biểu đồ T² của giai đoạn 2

STT	Điểm ngoại lệ	Biến	Giá trị p	STT	Điểm ngoại lệ	Biến	Giá trị p		
1	1	PC1	0,0000	10	20	PC1	0,0000		
		PC2	0,0038			PC3	0,0024		
		PC3	0,0002	11	21	PC1	0,0000		
2	2	PC1	0,0000			12	22	PC3	0,0145
		PC3	0,0000	PC1	0,0000				
3	3	PC1	0,0000	13	23	PC2	0,0006		
		PC3	0,0000			PC3	0,0003		
4	4	PC1	0,0000	14	24	PC1	0,0000		
		PC2	0,0183			PC2	0,0012		
		PC3	0,0021			PC3	0,0000		
5	5	PC1	0,0000	15	25	PC1	0,0000		
		6	6			PC1	0,0000	PC2	0,0002
						PC2	0,0012	PC3	0,0275
7	7	PC3	0,0000	16	30	PC1	0,0000		
		PC1	0,0055			PC3	0,0045		
		PC3	0,0000	17	31	PC1	0,0000		
8	8	PC1	0,0000			18	32	PC3	0,0000
		PC3	0,0069	PC1	0,0000				
9	19	PC1	0,0000						
		PC3	0,0017						

Kết quả phân tích PCA cũng tương thích với kết quả phân tách T² được trình bày ở Bảng 4. Theo Mason và cộng sự (1995), việc phân tách T² thành các phần độc lập để biết sự tác động mỗi biến riêng lẻ. Nếu giá trị phân tách T² của biến nào có giá trị càng lớn thì ảnh hưởng của biến đó càng nhiều [8]. Từ đây có thể khẳng định PC1 là thành phần chính gây ảnh hưởng lớn nhất đến điểm ngoại lệ.

Bảng 4. Kết quả phân tách T² các thành phần chính PC1, PC2, PC3 của các điểm ngoại lệ trên biểu đồ T² trong giai đoạn 2

STT	Điểm ngoại lệ	Giá trị phân tách T ² các thành phần chính các điểm ngoại lệ		
		PC1	PC2	PC3
1	1	80,55867	7,77775	12,38282
2	2	87,40689	2,254636	15,40192
3	3	31,31556	1,568141	18,0757
4	4	32,15729	5,313961	8,641164
5	5	21,37128	2,545243	1,215407
6	6	61,90683	9,560443	18,72336

STT	Điểm ngoại lệ	Giá trị phân tích T^2 các thành phần chính các điểm ngoại lệ		
		PC1	PC2	PC3
7	7	73,18842	7,181156	18,71621
8	8	54,94598	2,962371	6,849419
9	19	20,10122	2,295914	8,971138
10	20	27,99014	0,884388	8,496134
11	21	53,10592	2,340924	5,681509
12	22	42,07555	10,45208	11,62513
13	23	49,64723	9,461037	18,69167
14	24	31,18705	11,92871	4,668583
15	25	21,79179	1,138242	1,987746
16	30	16,23047	2,513991	7,517966
17	31	30,74661	2,892281	16,32041
18	32	19,15785	1,124937	2,547596

Tuy nhiên, các thành phần chính là sự tương quan của nhiều thông số. Do đó, để đánh giá ảnh hưởng của thông số đến điểm ngoại lệ thì cần xem xét tải trọng hay mức đóng góp của thông số này đến các thành phần chính [2,10]. Kết quả phân tích tải trọng thành phần chính PCA ở giai đoạn 2 được trình bày trong Bảng 5. Trong đó, thành phần chính PC1 là biến ảnh hưởng nhiều nhất đến các điểm ngoại lệ nên tải trọng của các thông số đối với PC1 được quan tâm. Hai thông số “chiều cao lớp lỏng” và “giá trị pH” có tải trọng lớn hơn các thông số khác với giá trị lần lượt là 0,706 và 0,682. Do đó, có thể nói hai thông số này tác động lớn nhất đến lỗi quy trình nên cần được chú ý kiểm tra lại trong quy trình vận hành. Thực tế, theo các công bố [14,15], thông số pH ảnh hưởng đến điện tích trên bề mặt khoáng chất và đóng một vai trò quan trọng trong hấp phụ của các thuốc thử để tách các khoáng chất. Sự thay đổi pH sẽ tác động lớn đến hiệu quả quá trình tách. Ví dụ, đối với kẽm, pH thích hợp từ 9,5-10,5 và chì là 9,5, nếu tăng pH lên đến 11,5 thì hiệu quả tách giảm đáng kể. Trong khi đó, thông số chiều cao lớp lỏng liên quan đến việc duy trì mức bọt trong quá trình tuyển nổi. Nếu chiều cao lớp lỏng cao thì lớp bọt thấp và ngược lại. Quan trọng là lớp bọt sẽ quyết định hiệu suất và ổn định của việc thu hồi kim loại. Nếu lớp bọt cao thì nồng độ kim loại cao nhưng có thể giảm khả năng thu hồi [15]. Qua đó, có thể nói phương pháp thống kê này cho kết quả phù hợp với thực tế quy trình tuyển nổi.

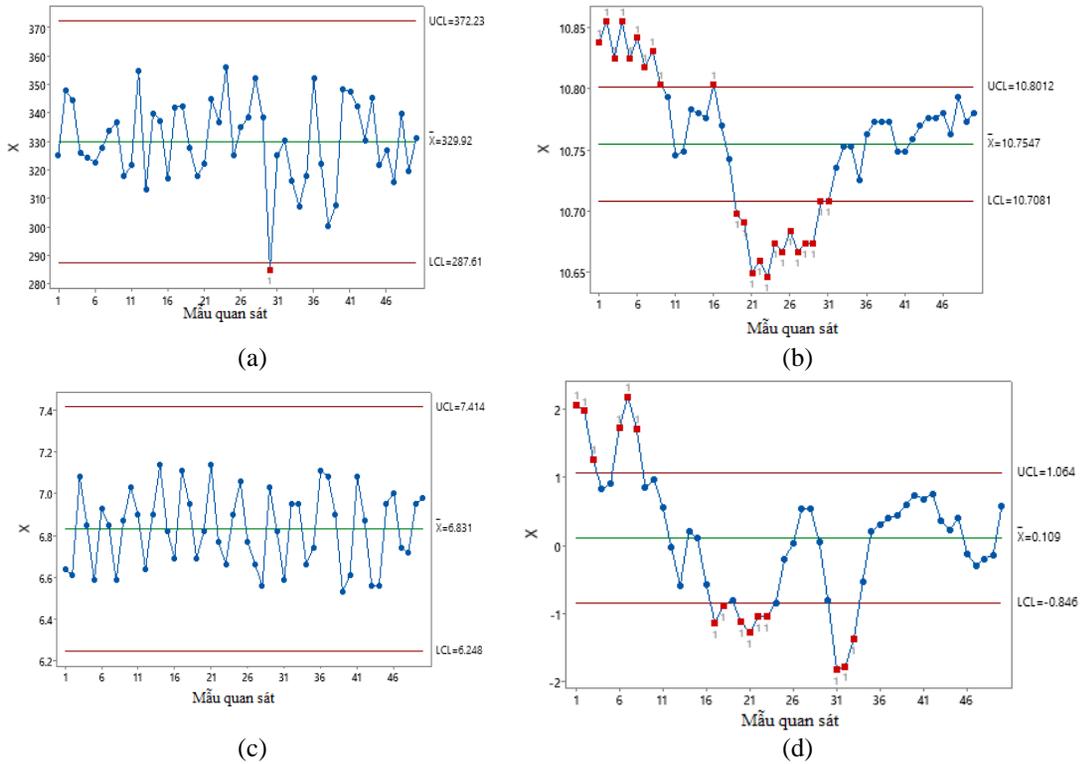
Bảng 5. Kết quả phân tích tải trọng của các thông số ban đầu lên các thành phần chính PC1, PC2, PC3 trong phân tích thành phần chính PCA trong giai đoạn 2

STT	Tham số ban đầu	PC1	PC2	PC3
1	Tốc độ nhập liệu	0,048	0,942	0,257
2	Giá trị pH	0,682	-0,231	0,159
3	Lượng CuSO_4	-0,186	-0,225	0,950
4	Chiều cao lớp lỏng	0,706	0,099	0,079

3.5. So sánh biểu đồ kiểm soát đa biến Hotelling T^2 với các biểu đồ kiểm soát đơn biến X

Các biểu đồ kiểm soát đơn biến X cũng là một công cụ kiểm soát quá trình hiệu quả. Tuy nhiên, trong trường hợp quy trình vận hành có nhiều biến tương quan với nhau như quá trình tuyển nổi thì việc áp dụng phương pháp kiểm soát đơn biến để tìm các điểm ngoại lệ cần được

cần nhắc. Để làm rõ hơn sự khác biệt hai biểu đồ, các điểm ngoại lệ của biểu đồ kiểm soát đa biến Hotelling T^2 với các biểu đồ kiểm soát đơn biến X được xác định và so sánh. Hình 7 cho thấy các điểm ngoại lệ tương ứng trên biểu đồ kiểm soát đơn biến của từng thông số. Cụ thể, thông số tốc độ nhập liệu có 1 điểm ngoại lệ, thông số giá trị pH tại đầu vào có 23 điểm ngoại lệ, thông số chiều cao lớp lông có 15 điểm ngoại lệ. Như vậy, trên biểu đồ X có tất cả 39 điểm ngoại lệ khác biệt với 18 điểm trên biểu đồ Hotelling T^2 . Các vị trí điểm ngoại lệ dựa trên biểu đồ đơn biến X được thể hiện cụ thể ở Bảng 6. Ngoài các điểm ngoại lệ trùng nhau từ 1-8, 19-25 và 30-32, một số điểm khác như 9, 16-18, 26-28, 33 xuất hiện trên biểu đồ đơn biến. Điều này cho thấy nếu một thông số độc lập khi vận hành có thể gây ra lỗi nhưng xét trong toàn bộ hệ thống, lỗi của thông số đó có thể không gây tác động đáng kể. Quan trọng là, cả hai biểu đồ đều thể hiện hai thông số “ giá trị pH tại đầu vào” và “chiều cao lớp lông” là thông số ảnh hưởng nhiều đến toàn bộ quy trình.



Hình 7. Các biểu đồ thị kiểm soát đơn biến: (a) Tốc độ nhập liệu; (b) Giá trị pH; (c) Lượng CuSO_4 ; (d) Chiều cao lớp lông

Bảng 6. Xác định các điểm ngoại lệ trên các biểu đồ kiểm soát đơn biến

STT	Tham số ban đầu	Số điểm ngoại lệ	Các điểm ngoại lệ
1	Tốc độ nhập liệu	1	30
2	Giá trị pH tại đầu vào	23	1, 2, 3, 4, 5, 6, 7, 8, 9, 16, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
3	Lượng CuSO_4	0	không có
4	Chiều cao lớp lông	15	1, 2, 3, 6, 7, 8, 17, 18, 20, 21, 22, 23, 31, 32, 33

4. KẾT LUẬN

Bài báo này đã giới thiệu biểu đồ kiểm soát thống kê đa biến Hotelling T² kết hợp phân tích thành phần chính PCA để kiểm soát quá trình tuyển nổi phân tách chì-kẽm. Cách thức để hình thành biểu đồ và phương pháp kiểm soát theo từng giai đoạn được thể hiện rõ. Bên cạnh đó, các kỹ thuật, phương pháp PCA và phân tách T² góp phần xác định được thông số gây ra lỗi của quá trình khi vận hành. So sánh với biểu đồ đơn biến X, biểu đồ Hotelling T² thể hiện ít điểm ngoại lệ hơn. Tuy nhiên, quá trình tuyển nổi có sự tương quan giữa các thông số, nên các điểm ngoại lệ của biểu đồ đơn biến chưa hẳn là ảnh hưởng đến toàn bộ quá trình. Ngược lại, điểm ngoại lệ trên biểu đồ Hotelling T² là tổng thể của các thông số tác động nên đánh giá chắc chắn. Cả hai biểu đồ đều thể hiện được hai thông số “chiều cao lớp lỏng” và “giá trị pH” ảnh hưởng nhiều đến lỗi của hệ thống phân tách tuyển nổi. Qua đây, có thể nói, biểu đồ Hotelling T² hoàn toàn có thể là công cụ hữu dụng trong việc kiểm soát thống kê quá trình sản xuất có nhiều thông số tác động.

TÀI LIỆU THAM KHẢO

1. Bersimis S., Panaretos J., Psarakis S. - Multivariate statistical process control charts and the problem of interpretation: A short overview and some applications in industry, Proceedings of the 7th Hellenic European Conference on Computer Mathematics and its Applications, Athens, Greece (2005).
2. Bersimis S., Psarakis S., Panaretos J.- Multivariate statistical process control charts: An Overview, Quality & Reliability Engineering International **23** (5) (2007) 517-543. <https://doi.org/10.1002/qre.829>
3. Montgomery D.C. - Introduction to Statistical Quality Control, (Sixth Edition), John Wiley & Sons, United States of America (2009).
4. Aslam M., Arif O.H.- Multivariate Analysis under Indeterminacy: An Application to Chemical Content Data Journal of Analytical Methods in Chemistry **2020**, Article ID 1406028. <https://doi.org/10.1155/2020/1406028>
5. Tiryaki S., Aydn A.- Multivariate Hotelling T² Control Chart for Monitoring Some Quality Characteristics in Medium Density Fiberboard Manufacturing Process, Drvna industrija **73** (1) (2022) 35-46. <https://doi.org/10.5552/drvind.2022.2046>
6. Costa M.T., Carolino E., Oliveira T.A.- Oliveira, Application of Hotelling's T² charts in Monitoring Quality Parameters in a Drinking Water Supply System, Proceedings of the International Conference on Numerical Analysis and Applied Mathematics (ICNAAM-2014), (2014). <https://doi.org/10.1063/1.4913046>
7. Hồ Thị Linh Trang - Ứng dụng phương pháp phân tích thành phần chính trong đánh giá tổng hợp trình độ phát triển kinh tế của các tỉnh thành Việt Nam, Tạp chí Kinh tế và Kinh doanh **16** (2021) 7-13.
8. Mason R.L., Tracy N.D., Young C.H. - Decomposition of T² for Multivariate Control Chart Interpretation, Journal of Quality Technology **27** (2) (1995) 99-108. <https://doi.org/10.1080/00224065.1995.11979573>
9. <https://openmv.net/info/flotation-cell>
10. Sullivan J.H., Woodall W.H.- A Comparison of Multivariate Control Charts for Individual Observations, Journal of Quality Technology **28** (4) (1996) 398-408. <https://doi.org/10.1080/00224065.1996.11979698>

11. Vining G.- Technical Advice: Phase I and Phase II Control Charts, *Quality Engineering*, **21** (4) (2009) 478-479. <https://doi.org/10.1080/08982110903185736>
12. MINITAB User's Guide: Data Analysis and Quality Tools, Vol. **2**, Minitab Incorporated, 2000.
13. Arifin W.N. - The Graphical Assessment of Multivariate Normality Using SPSS, *Education in Medicine Journal* **7** (2) (2015), 71-75. <https://doi.org/10.5959/eimj.v7i2.361>
14. Singh R., Banerjee B. and Srivastava J.P. - Effects of Process Parameters on Selective Flotation of Lead-Zinc Ore, In: *International Seminar on Mineral Processing Technology (MPT-2004)*, February 19-21, 2004. <http://eprints.nmlindia.org/6145>
15. Wills B.A., Finch J. - *Wills' Mineral Processing Technology*, (Eighth Edition), Butterworth-Heinemann (2016) 265-380.

ABSTRACT

APPLICATION OF HOTELLING'S T^2 MULTIVARIATE STATISTICS CHART IN PRODUCTION CONTROL

Bui Thu Ha, Nguyen Thi Thanh Hien, Ngo Thanh An*

Ho Chi Minh City University of Industry and Trade

*Email: annt@huit.edu.vn

A multivariate statistical process control chart is a popular tool applied in quality management to control the production process. This paper introduced a Hotelling T^2 multivariate statistical chart that is applied to evaluate the lead-zinc flotation process according to the announced data. The method of building the Hotelling T^2 chart is shown through two stages. The results of control limit values and outliers were determined by the Hotelling T^2 chart and the principal component analysis (PCA), performed by Minitab 21 software. Besides, the main parameters that affect the process were also analyzed by PCA and T^2 decomposition methods. Like the univariate graph X, the Hotelling T^2 chart also identified “pulp level” and “upstream pH value” as the two crucial factors affecting the flotation process. However, the number of outliers on the Hotelling T^2 graph is less than those of the X chart that displayed the correlation of the parameters affecting the process efficiency. It can be said that the Hotelling T^2 multivariate chart will be a useful tool for statistical control of the production process owning many correlated parameters.

Keywords: Multivariate statistical process control, Hotelling T^2 , outliers.