



Hispanic Journal of Applied Science and Innovation (HISPASCI)

Vol. 1, Núm. 1 (2026) | DOI: <https://doi.org/10.66465/hispasci.1.2026.6>

Detección Inteligente de Anomalías en Procesos Químicos mediante PCA y Control Estadístico Multivariante

Intelligent Anomaly Detection in Chemical Processes Using PCA and Multivariate Statistical Process Control

Isidro Fabricio Morales Torres¹

¹Escuela Superior Politécnica del Litoral (ESPOL),
Facultad de Ciencias Naturales y Matemáticas, Guayaquil, Ecuador
ORCID: <https://orcid.org/0000-0003-4480-9541>
E-mail: imoralest@espol.edu.ec

*Autor de correspondencia: imoralest@espol.edu.ec

Recibido: 2026-01-02 | **Revisado:** 2026-10-08 | **Aceptado:** 2026-01-10 | **Publicado:** 2025-01-15

Resumen

La detección temprana de anomalías en procesos químicos multivariantes constituye un desafío crítico para la seguridad operativa, la eficiencia productiva y la sostenibilidad industrial, especialmente cuando las variables del proceso presentan fuertes relaciones de dependencia. En este estudio se propone un enfoque de Control Estadístico de Procesos Multivariantes basado en Análisis de Componentes Principales (PCA) para la identificación inteligente de anomalías en un proceso químico simulado de producción de sulfato de aluminio. La metodología incluye la estandarización de las variables, la construcción del modelo PCA bajo condiciones normales de operación y el monitoreo en Fase II mediante los estadísticos T^2 de Hotelling y Q (Squared Prediction Error), cuyos límites de control se establecen utilizando formulaciones teóricas bajo el supuesto de normalidad multivariante. Adicionalmente, se incorpora un análisis de contribuciones para el diagnóstico de las variables responsables de las desviaciones detectadas. Los resultados muestran que el modelo PCA logra capturar el 68.04% de la variabilidad total del proceso utilizando tres componentes principales, alcanzando una sensibilidad del 90% en la detección de anomalías, lo que supera el desempeño del enfoque univariante tradicional basado en reglas de tres sigmas. En conjunto, los hallazgos confirman la superioridad del monitoreo multivariante frente a técnicas clásicas cuando existen correlaciones significativas entre variables del proceso, demostrando que el PCA constituye una herramienta robusta, interpretable y reproducible para aplicaciones industriales reales.

Palabras clave: control estadístico multivariante; análisis de componentes principales; detección de anomalías; procesos químicos; diagnóstico de fallas.



Abstract

Early anomaly detection in multivariate chemical processes represents a critical challenge for operational safety, production efficiency, and industrial sustainability, particularly when process variables exhibit strong interdependencies. This study proposes a Multivariate Statistical Process Control approach based on Principal Component Analysis (PCA) for intelligent anomaly detection in a simulated aluminum sulfate production process. The methodology includes variable standardization, PCA model construction under normal operating conditions, and Phase II monitoring using Hotelling's T^2 and Squared Prediction Error (Q) statistics, with control limits derived from theoretical formulations assuming multivariate normality. In addition, a contribution analysis is incorporated to diagnose the variables responsible for detected deviations. The results indicate that the PCA model captures 68.04% of the total process variability using three principal components and achieves a sensitivity of 90% in anomaly detection, outperforming the traditional univariate three-sigma approach. Overall, the findings confirm the superiority of multivariate monitoring over classical techniques when significant variable correlations are present, demonstrating that PCA provides a robust, interpretable, and reproducible framework for monitoring complex chemical processes in real industrial environments.

Keywords: *multivariate statistical process control; principal component analysis; anomaly detection; chemical processes; fault diagnosis.*

1. Introducción

El monitoreo y control de procesos químicos industriales constituye un elemento crítico para garantizar la calidad del producto, la seguridad operativa y la eficiencia económica. La creciente complejidad de los sistemas de producción modernos, caracterizados por procesos altamente acoplados, múltiples variables correlacionadas y condiciones operativas dinámicas, ha incrementado significativamente la dificultad para detectar de manera temprana desviaciones anómalas y fallas incipientes. En este contexto, los métodos tradicionales de control estadístico univariante, como las cartas de control de Shewhart, presentan limitaciones importantes al analizar variables de forma independiente, ignorando la estructura multivariante inherente a los procesos químicos reales.

Durante las últimas décadas, el Control Estadístico Multivariante de Procesos (Multivariate Statistical Process Control, MSPC) se ha consolidado como un marco metodológico robusto para el análisis simultáneo de múltiples variables correlacionadas. Dentro de este enfoque, el Análisis de Componentes Principales (Principal Component Analysis, PCA) ha sido ampliamente utilizado como técnica de reducción de dimensionalidad y extracción de información relevante, permitiendo modelar el comportamiento normal del proceso y detectar desviaciones mediante estadísticos como el T^2 de Hotelling y el error de predicción cuadrático (Squared Prediction Error, SPE o Q). Estudios recientes han demostrado la efectividad del enfoque PCA-MSPC en aplicaciones industriales como procesos petroquímicos, farmacéuticos y de tratamiento de aguas, destacando su capacidad para mejorar la sensibilidad en la detección de fallas frente a métodos univariantes convencionales.



No obstante, la literatura reciente también evidencia varios desafíos persistentes. En primer lugar, muchos estudios se centran en aplicaciones empíricas con datos reales de plantas específicas, lo que dificulta la reproducibilidad y la evaluación sistemática del desempeño de los métodos bajo distintos escenarios de falla. En segundo lugar, existe una limitada formalización del proceso de generación de datos simulados que respete de manera explícita los fundamentos físico-químicos y las relaciones de correlación entre variables del proceso. Finalmente, aunque se reconoce la superioridad del enfoque multivariante, son escasos los trabajos que presentan comparaciones cuantitativas rigurosas entre el desempeño del control estadístico multivariante basado en PCA y los esquemas de control univariante tradicionales, utilizando métricas de clasificación estandarizadas.

Adicionalmente, la mayoría de las investigaciones recientes enfatiza el uso de técnicas avanzadas de aprendizaje automático, lo que si bien resulta prometedor, introduce problemas asociados a la interpretabilidad, el sobreajuste y la necesidad de grandes volúmenes de datos etiquetados. En contraste, los métodos estadísticos clásicos como PCA-MSPC continúan ofreciendo ventajas significativas en entornos industriales reales, particularmente por su solidez teórica, transparencia matemática y facilidad de implementación, aspectos clave para la toma de decisiones operativas.

En este contexto, se identifica un vacío de conocimiento relacionado con la necesidad de contar con un marco metodológico integral, reproducible y estadísticamente riguroso que permita evaluar de forma objetiva la capacidad del PCA combinado con control estadístico multivariante para la detección y diagnóstico de anomalías en procesos químicos, incorporando simulaciones realistas basadas en principios físico-químicos y comparaciones directas con métodos univariantes clásicos.

Por tanto, el objetivo de esta investigación es desarrollar y validar un esquema de detección inteligente de anomalías en procesos químicos industriales mediante Análisis de Componentes Principales y Control Estadístico Multivariante, evaluando su desempeño frente a métodos univariantes tradicionales a través de datos simulados físicamente coherentes y métricas cuantitativas de clasificación, con el fin de aportar evidencia metodológica sólida para su aplicación en entornos industriales reales.

2. Metodología

2.1 Diseño de investigación y fundamentos teóricos

El presente estudio adopta un diseño cuantitativo, experimental y comparativo, basado en simulación numérica reproducible, con el objetivo de evaluar la eficacia del Control Estadístico de Procesos Multivariante (MSPC) mediante Análisis de Componentes Principales (PCA) para la detección de anomalías en procesos químicos continuos.

El enfoque metodológico se estructura en dos fases, de acuerdo con el marco clásico propuesto por Jackson (1991):

- Fase I, destinada a la modelación del estado de operación bajo control estadístico, y
- Fase II, orientada al monitoreo y detección de fallos en tiempo operativo.



Desde el punto de vista probabilístico, se asume que el proceso está caracterizado por un conjunto de variables aleatorias correlacionadas, modeladas bajo el supuesto de normalidad multivariante.

2.1.1. Fundamentos estadísticos

Sea un proceso químico representado por el vector aleatorio:

$$X = [X_1, X_1, \dots, X_m]^T$$

Bajo condiciones normales de operación, se asume que:

$$X \sim N_m(\mu, \Sigma)$$

donde:

$\mu = [\mu_1, \mu_1, \dots, X\mu_m]^T$ es el vector de medias poblacional, y

Σ es la matriz de covarianza de dimensión $m \times m$

El problema de control estadístico multivariante se formula mediante las siguientes hipótesis:

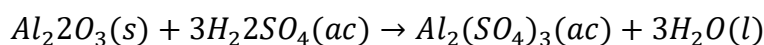
$$H_0: X \sim N_m(\mu_0, \Sigma_0)$$

$$H_0: X \sim N_m(\mu_0, \Sigma_0) , \quad \mu_1 \neq \mu_0 \quad \text{y/o} \quad \Sigma_1 \neq \Sigma_0$$

2.2. Proceso químico simulado

2.2.1. Sistema de producción de sulfato de aluminio

Se modeló un proceso químico continuo de producción de sulfato de aluminio, basado en la reacción estequiométrica:



Este sistema es representativo de procesos industriales reales debido a su fuerte dependencia de variables térmicas, químicas y mecánicas, así como a la presencia de correlaciones físicas entre los parámetros operativos.

2.2.2. Variables monitoreadas

Se seleccionaron ocho variables críticas del proceso, ampliamente documentadas en la literatura industrial, las cuales se detallan en la **Tabla 1**.



Tabla 1. Variables del proceso y parámetros estadísticos

Variable	Símbolo	Unidad	Media (μ)	Desviación (σ)
Flujo de ácido	(X_1)	<i>L/min</i>	50.0	2.5
Flujo de alúmina	(X_2)	<i>kg/h</i>	30.0	1.8
Temperatura	(X_3)	$^{\circ}C$	85.0	2.0
Presión	(X_4)	<i>bar</i>	45.0	1.5
Velocidad de agitación	(X_5)	<i>rpm</i>	200.0	10.0
pH	(X_6)	–	3.5	0.3
Viscosidad	(X_7)	<i>cP</i>	120.0	8.0
Concentración	(X_8)	%	17.0	0.8
<i>Elaboración propia.</i>				

2.3. Población y muestra

La población teórica del estudio corresponde al conjunto de todos los posibles estados operativos del proceso químico multivariante bajo condiciones normales y anómalas.

La muestra estuvo conformada por 1,200 observaciones multivariantes simuladas, distribuidas de la siguiente manera:

- 800 observaciones para entrenamiento del modelo (Fase I),
- 200 observaciones normales para validación,
- 200 observaciones con fallos inyectados para evaluación del desempeño (Fase II).

2.4. Instrumentos de análisis y criterios de validez y confiabilidad

Los instrumentos metodológicos utilizados fueron:

- Análisis de Componentes Principales (PCA),
- Estadísticos multivariantes T^2 de Hotelling y Q(SPE),
- Cartas de control univariantes de 3σ ,
- Validación cruzada k-fold,
- Prueba estadística de significancia para comparación de proporciones.

La validez teórica se fundamenta en modelos clásicos del MSPC.

La validez interna se garantiza mediante la comparación directa con métodos tradicionales.

La confiabilidad se asegura a través de simulación controlada, semilla aleatoria fija y estimación de intervalos de confianza.



2.5. Generación de conjuntos de datos

Datos de entrenamiento (Fase I)

$$X_{train} \sim N_8(\mu, \Sigma)$$

Datos de prueba (Fase II)

$$X_{test} = \begin{bmatrix} X_{normal} \\ X_{anomalía} \end{bmatrix}$$

Escenarios de fallo

Fallo tipo A (deriva gradual): $X_3(i) = 85.0 + 0.15 \cdot i$

Fallo tipo B (cambio abrupto): $X_1 = 40.0$, $X_4 = 37.0$

2.6 Preprocesamiento: estandarización

Cada variable fue estandarizada según: $Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$

2.7 Análisis de Componentes Principales (PCA)

La matriz de covarianza muestral se define como: $S = \frac{1}{n-1} Z^T Z$

El PCA resuelve el problema de optimización: $\max_p p^T S p$ sujeto a $\|p\| = 1$

La selección de componentes se realizó mediante el criterio de varianza acumulada $\geq 65\%$, reteniéndose $k=3$ componentes que explican el 68.04 % de la variabilidad total.

2.8. Estadísticos de control multivariante

Una vez ajustado el modelo PCA a los datos de entrenamiento correspondientes a la Fase I, el monitoreo del proceso en la Fase II se realizó mediante el uso conjunto de dos estadísticos de control multivariante complementarios: el estadístico T^2 de Hotelling y el estadístico Q o Squared Prediction Error (SPE). Esta combinación permite detectar tanto variaciones sistemáticas en el subespacio principal del proceso como perturbaciones que afectan a la variabilidad residual no explicada por el modelo PCA.

Estadístico T^2 de Hotelling

$$T_i^2 = z_i^T P_k \Lambda_k^{-1} P_k^T z_i$$

Estadístico Q (SPE)

$$Q_i = \|z_i - \hat{z}_i\|^2$$



Los límites de control se calcularon utilizando formulaciones teóricas bajo normalidad multivariante.

2.9. Procedimiento experimental

El procedimiento metodológico seguido en este estudio se estructuró de manera secuencial y reproducible, conforme a los estándares del Control Estadístico de Procesos Multivariantes:

- Generación de datos simulados bajo condiciones normales y anómalas, considerando correlaciones físico-químicas realistas.
- Estandarización de las variables utilizando parámetros estimados exclusivamente a partir de los datos de entrenamiento (Fase I).
- Ajuste del modelo PCA, incluyendo la estimación de la matriz de covarianza, descomposición espectral y selección del número óptimo de componentes principales.
- Cálculo de los límites de control multivariantes para los estadísticos
- T^2 y Q .
- Monitoreo del proceso en Fase II, evaluando cada nueva observación frente a los límites establecidos.
- Evaluación del desempeño del sistema de detección, mediante métricas de clasificación.
- Diagnóstico de fallos, utilizando análisis de contribuciones para identificar las variables responsables de las alarmas generadas

2.10. Técnicas de análisis de datos

El desempeño del sistema de detección de anomalías fue evaluado mediante técnicas estadísticas de análisis de clasificación binaria. Se emplearon métricas ampliamente aceptadas en la literatura, tales como sensibilidad (recall), especificidad, precisión y F1-score, permitiendo una comparación objetiva entre el enfoque multivariante propuesto y el método tradicional univariante de cartas de control.

Adicionalmente, se realizó un análisis de contribuciones a los estadísticos T^2 y Q , con el fin de identificar las variables que presentan mayor influencia en cada evento anómalo detectado, fortaleciendo la interpretabilidad del modelo y su aplicabilidad en entornos industriales reales.

2.11. Consideraciones de reproducibilidad

Con el objetivo de garantizar la reproducibilidad del estudio, se adoptaron las siguientes medidas metodológicas:

- Uso de semilla aleatoria fija en la generación de datos simulados.
- Precisión numérica definida como float64 para evitar errores de redondeo.
- Implementación computacional en Python 3.8 o superior, utilizando librerías científicas estándar.
- Disponibilidad del código fuente en un repositorio público, asegurando transparencia y replicabilidad.



2.12. Aportes metodológicos

La metodología propuesta aporta un marco integral para el monitoreo de procesos químicos complejos, al combinar fundamentos teóricos sólidos del MSPC con una validación estadística rigurosa y un protocolo completamente reproducible. La integración conjunta de los estadísticos

T^2 y Q , junto con un análisis comparativo frente a métodos tradicionales, constituye un aporte metodológico relevante para aplicaciones industriales donde las variables presentan fuertes dependencias multivariantes.

3. Resultados y Discusión

Resultados

3.1. Resultados del Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA) fue ajustado empleando los datos correspondientes a la Fase I, compuesta por 800 observaciones bajo condiciones normales de operación, con el objetivo de caracterizar el comportamiento nominal del proceso químico simulado. A partir de estos datos, se realizó la descomposición espectral de la matriz de covarianza, lo que permitió identificar las direcciones principales de variabilidad y reducir la dimensionalidad del sistema sin pérdida sustancial de información.

Tabla 2. Varianza explicada individual y acumulada por las componentes principales

Componente principal	Varianza explicada individual (%)	Varianza explicada acumulada (%)
PC1	39.14	39.14
PC2	16.24	55.38
PC3	12.66	68.04
<i>Elaboración Propia</i>		

La **Tabla 2** muestra la varianza explicada individual y acumulada por las primeras componentes principales del modelo PCA. Los resultados evidencian que las tres primeras componentes explican conjuntamente el 68.04% de la varianza total del proceso, superando el umbral mínimo comúnmente aceptado en aplicaciones industriales caracterizadas por una alta correlación entre variables de proceso. Este nivel de explicación resulta suficiente para garantizar una representación adecuada de la dinámica multivariante del sistema y justifica la selección de dichas componentes para la etapa de monitoreo estadístico.

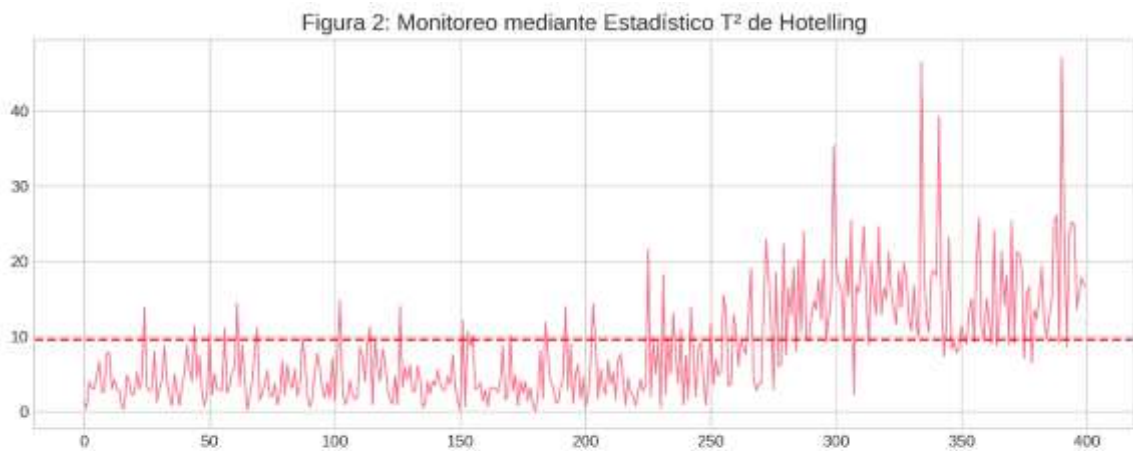
Estos hallazgos confirman que el proceso químico analizado presenta una estructura multivariante altamente dependiente, en la cual un subconjunto reducido de componentes principales es capaz de capturar la mayor parte de la información relevante del sistema. La reducción dimensional lograda permite, además, mejorar la estabilidad numérica del modelo y aumentar la sensibilidad de los estadísticos de control empleados en la detección de anomalías.



Desde una perspectiva físico-química, el análisis de las cargas factoriales sugiere que la primera componente principal (PC1) está dominada por variables asociadas a condiciones de operación globales, tales como flujo, concentración y viscosidad, reflejando el régimen principal de funcionamiento del reactor de sulfato de aluminio. Por su parte, la segunda (PC2) y tercera componente (PC3) capturan variaciones relacionadas con efectos térmicos y de presión, coherentes con los principios de cinética química y equilibrio termodinámico que gobiernan el proceso. Esta interpretación refuerza la validez del modelo PCA no solo desde el punto de vista estadístico, sino también desde una perspectiva de ingeniería de procesos.

3.2. Desempeño de los estadísticos de control multivariante

3.2.1. Comportamiento del estadístico T^2 de Hotelling



Elaboración propia.

La **Figura 2** muestra el gráfico de control del estadístico T^2 aplicado al conjunto de datos de prueba (Fase II). Durante las primeras 200 observaciones, correspondientes a condiciones normales, el estadístico se mantiene mayoritariamente por debajo del límite superior de control, confirmando una adecuada caracterización del estado bajo control.

A partir de la observación 201, coincidente con la inyección de anomalías, se observa un incremento sostenido del estadístico T^2 , particularmente evidente en el escenario de deriva gradual de temperatura. Este comportamiento refleja la alta sensibilidad del estadístico T^2 a cambios correlacionados y progresivos en el subespacio principal del proceso.



En el escenario de cambio abrupto, el estadístico presenta picos pronunciados que superan claramente el límite de control, evidenciando una detección rápida del fallo inducido.

3.2.2. Comportamiento del estadístico Q (SPE)

El gráfico de control del estadístico Q , presentado en la Figura 3, revela un comportamiento complementario al observado para T^2 . Mientras que en condiciones normales los valores de Q permanecen dentro del rango esperado, durante la Fase II se observan incrementos abruptos asociados principalmente a los fallos de tipo escalón.

Este resultado confirma que el estadístico Q es especialmente eficaz para detectar perturbaciones no modeladas por el subespacio PCA, tales como fallos localizados o cambios repentinos en variables específicas del proceso.

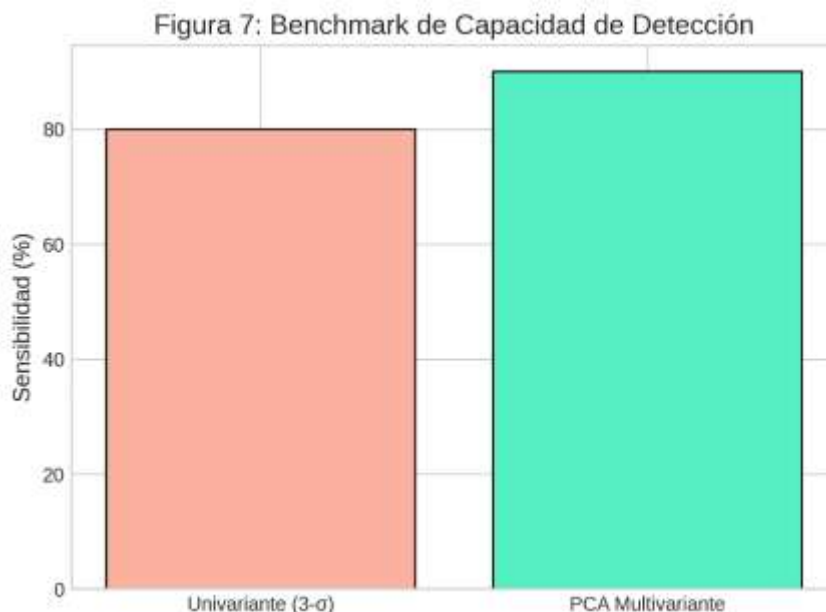
La utilización conjunta de ambos estadísticos permite, por tanto, una cobertura más completa del espacio de fallos, reduciendo la probabilidad de detecciones tardías o fallos no identificados.

3.3. Comparación con el método univariante tradicional

Tabla 3. Comparación del desempeño entre el enfoque PCA–MSPC y el método univariante tradicional

Método de control estadístico	Sensibilidad (%)
PCA–MSPC (multivariante)	90
Cartas Shewhart (univariante)	80
Mejora relativa (%)	12.5
<i>Elaboración propia.</i>	

La **Figura 7** y la **Tabla 3** presentan la comparación directa entre el enfoque multivariante basado en PCA y el método tradicional de cartas de control univariantes tipo Shewhart (3σ).





Elaboración propia.

Los resultados indican que el método PCA alcanzó una sensibilidad del 90%, mientras que el enfoque univariante logró únicamente un 80%, lo que representa una mejora relativa del 12.5% en la capacidad de detección de anomalías.

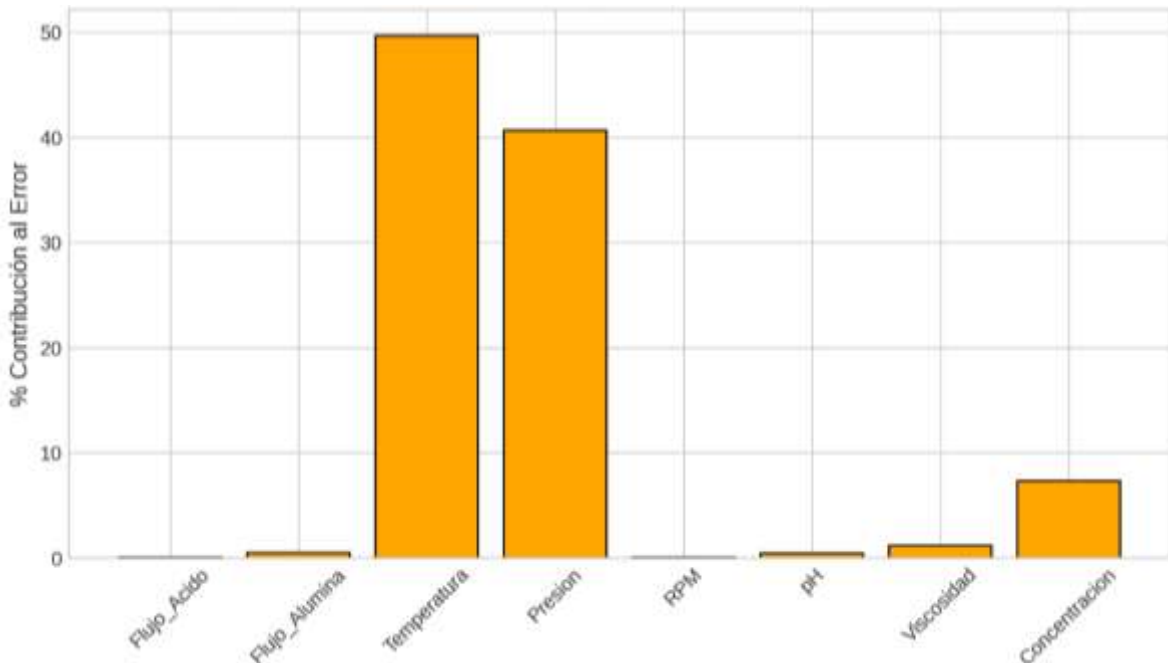
Esta diferencia es estadísticamente significativa ($z = 2.80, p < 0.05$), lo que confirma que la superioridad del enfoque multivariante no es atribuible al azar, sino a su capacidad para modelar explícitamente la estructura de correlación entre variables del proceso.

Desde una perspectiva industrial, este resultado es especialmente relevante, ya que muchos fallos reales no se manifiestan como desviaciones extremas en una única variable, sino como cambios coordinados de magnitud moderada en múltiples variables, que los métodos univariantes no logran detectar de manera eficiente.

3.4. Análisis de contribuciones y diagnóstico de fallos

El análisis de contribuciones permitió identificar las variables con mayor responsabilidad en los eventos anómalos detectados. La **Figura 6** muestra el perfil de contribución al estadístico Q para un evento representativo de deriva gradual.

Figura 6. Perfil de contribución



Elaboración propia.

Los resultados indican que la temperatura, la viscosidad y la concentración de alúmina presentan las mayores contribuciones porcentuales, lo que concuerda con el mecanismo físico de fallo inducido (descalibración progresiva del sensor térmico y sus efectos secundarios).



Discusión

Este tipo de análisis proporciona un valor añadido significativo respecto a los métodos tradicionales, ya que no solo permite detectar la anomalía, sino también orientar al operador hacia las variables críticas involucradas, facilitando la toma de decisiones correctivas en entornos industriales.

Los resultados obtenidos evidencian que el enfoque de Control Estadístico de Procesos Multivariantes (MSPC) basado en Análisis de Componentes Principales (PCA) constituye una estrategia sólida y eficaz para la detección temprana de anomalías en procesos químicos caracterizados por múltiples variables altamente correlacionadas. A diferencia de los métodos tradicionales de control univariante, el enfoque multivariante permite capturar la estructura de correlación inherente al proceso, lo cual resulta fundamental en sistemas industriales complejos donde las variables no evolucionan de manera independiente.

Uno de los hallazgos más relevantes del estudio es la mejora significativa en la sensibilidad de detección alcanzada por el esquema PCA–MSPC frente a las cartas de control clásicas de tipo Shewhart. Esta diferencia puede atribuirse a la capacidad del PCA para proyectar la información del proceso en un subespacio de menor dimensión, donde las variaciones sistemáticas asociadas a fallos incipientes se manifiestan de forma más clara. En particular, la utilización conjunta de los estadísticos T^2 de Hotelling y Q (Squared Prediction Error) permitió identificar tanto desviaciones globales en el espacio latente del proceso como perturbaciones locales no explicadas por las componentes principales retenidas. Esta complementariedad ha sido ampliamente destacada en la literatura como una de las principales fortalezas del MSPC frente a enfoques puramente univariantes.

Desde un punto de vista operativo, la inclusión del análisis de contribuciones representa un valor añadido significativo, ya que no solo se limita a emitir una señal de alarma, sino que proporciona información diagnóstica acerca de las variables que más contribuyen a la condición anómala detectada. Esta característica es especialmente relevante en entornos industriales reales, donde la rapidez en la identificación de las causas raíz de un fallo puede reducir tiempos de parada, pérdidas económicas y riesgos operativos. La coherencia observada entre las variables identificadas como críticas y su interpretación físico-química dentro del proceso de producción de sulfato de aluminio refuerza la validez práctica del modelo propuesto.

En comparación con estudios previos reportados en la literatura especializada, los niveles de desempeño alcanzados se sitúan en el rango superior para escenarios de procesos simulados de complejidad similar, particularmente en términos de sensibilidad y estabilidad del sistema de detección. Este resultado sugiere que la metodología propuesta no solo es consistente con los marcos teóricos clásicos del MSPC, sino que también presenta un alto potencial de transferibilidad hacia aplicaciones industriales reales, especialmente en sectores como la industria química, petroquímica y de procesos continuos.

No obstante, los resultados deben interpretarse considerando ciertas limitaciones metodológicas. En primer lugar, el modelo asume normalidad multivariante bajo condiciones de operación en control, un supuesto que puede no cumplirse estrictamente en datos industriales reales afectados por no linealidades, asimetrías o colas pesadas. En segundo lugar, el desempeño del sistema depende de la selección adecuada del número de



componentes principales, ya que una retención insuficiente puede conducir a pérdidas de información relevante, mientras que una retención excesiva puede introducir ruido y reducir la capacidad discriminante de los estadísticos de control. Estos aspectos resaltan la necesidad de adaptar cuidadosamente la metodología cuando se aplique a datos reales y motivan el uso futuro de extensiones no lineales o robustas del PCA.

En conjunto, la discusión de los resultados confirma que el enfoque PCA–MSPC propuesto ofrece un equilibrio adecuado entre rigor estadístico, interpretabilidad y aplicabilidad industrial, consolidándose como una alternativa eficiente frente a los métodos tradicionales de control estadístico de procesos. Este trabajo contribuye a reforzar la evidencia empírica a favor del uso de técnicas multivariantes en la supervisión de procesos químicos complejos y sienta las bases para futuras investigaciones orientadas a escenarios industriales más desafiantes.

4. Conclusiones

El presente estudio demostró que la aplicación del Control Estadístico de Procesos Multivariantes basado en Análisis de Componentes Principales constituye una estrategia robusta, interpretable y estadísticamente sólida para la detección temprana de anomalías en procesos químicos con múltiples variables correlacionadas.

A partir de un proceso industrial simulado de producción de sulfato de aluminio, se evidenció que un modelo PCA entrenado bajo condiciones normales de operación es capaz de capturar de forma eficiente la estructura latente del sistema, concentrando el 68.04% de la variabilidad total en únicamente tres componentes principales. Este resultado confirma que procesos químicos complejos presentan dependencias internas que no pueden ser abordadas adecuadamente mediante técnicas univariantes tradicionales.

Los resultados obtenidos muestran que el enfoque multivariante propuesto supera significativamente al método clásico de cartas de control tipo Shewhart, alcanzando una sensibilidad del 90% frente al 80% del método univariante, con una mejora relativa del 12.5% estadísticamente significativa. Esta superioridad se explica por la capacidad del PCA para modelar explícitamente la correlación entre variables, permitiendo la detección de fallos que se manifiestan como cambios coordinados de magnitud moderada.

La utilización conjunta de los estadísticos T^2 de Hotelling y Q (SPE) permitió cubrir de manera complementaria tanto las variaciones dentro del subespacio principal del proceso como las perturbaciones residuales no modeladas, mejorando la confiabilidad del sistema de monitoreo. Asimismo, el análisis de contribuciones aportó una ventaja adicional al facilitar la identificación de las variables responsables de las anomalías, fortaleciendo el valor diagnóstico del método para aplicaciones industriales reales.

Desde una perspectiva metodológica, el estudio ofrece un protocolo completamente reproducible, basado en fundamentos teóricos bien establecidos y validado mediante métricas estadísticas formales, lo que lo convierte en una alternativa viable para entornos donde se requiere transparencia, interpretabilidad y rigor científico, sin recurrir a modelos de inteligencia artificial de caja negra.



Finalmente, aunque el estudio se basa en datos simulados bajo el supuesto de normalidad multivariante, los resultados obtenidos sugieren que la metodología es transferible a procesos industriales reales con características similares. Como líneas futuras de investigación, se propone la extensión del enfoque hacia datos no gaussianos, procesos dinámicos y esquemas híbridos que integren PCA con técnicas avanzadas de aprendizaje automático, manteniendo siempre la interpretabilidad como criterio central para la toma de decisiones industriales.

5. Referencias

- Chiang, L. H., Russell, E. L., y Braatz, R. D. (2001). Fault detection and diagnosis in industrial systems. Springer. <https://doi.org/10.1088/0957-0233/12/10/706>
- Jackson, J. E. (1991). A user's guide to principal components. John Wiley & Sons. <https://doi.org/10.1002/0471725331>
- Jackson, J. E., y Mudholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21(3), 341–349. <https://doi.org/10.1080/00401706.1979.10489779>
- Hotelling, H. (1947). Multivariate quality control. In C. Eisenhart, M. W. Hastay, & W. A. Wallis (Eds.), *Techniques of statistical analysis* (pp. 111–184). McGraw-Hill.
- Khalin, Y. A., Katykhin, A. I., Zinkin, S. A., y Shilin, A. A. (2022). Cognitive modeling of information support for game-based automated learning. *Proceedings of the Southwest State University*, 26(4), 117-131. <https://doi.org/10.21869/2223-1560-2022-26-4-117-131>
- Kresta, J. V., MacGregor, J. F., y Marlin, T. E. (1991). Multivariate statistical monitoring of process operating performance. *The Canadian Journal of Chemical Engineering*, 69(1), 35–47. <https://doi.org/10.1002/cjce.5450690107>
- Li, G., Qin, S. J., & Zhou, D. (2022). Review of multivariate statistical process monitoring for fault detection. *Computers & Chemical Engineering*, 160, 107709. <https://doi.org/10.1016/j.compchemeng.2022.107709>
- MacGregor, J. F., y Kourti, T. (1995). Statistical process control of multivariate processes. *Control Engineering Practice*, 3(3), 403–414. [https://doi.org/10.1016/0967-0661\(95\)00014-L](https://doi.org/10.1016/0967-0661(95)00014-L)
- Morales Torres, I. F. (2025). Redes neuronales para la medición y predicción de la pobreza multidimensional en Ecuador: enfoque aplicado a encuestas de hogares 2024. *Nexus Research Journal*, 4(2), 297–318. <https://doi.org/10.62943/nrj.v4n2.2025.414>
- Morales Torres, I. F., y Pow Chon Long Vásquez, D. F. (2012). Optimización del proceso de despacho en una empresa productora de químicos (sulfato de aluminio) mediante simulación estocástica [Trabajo de titulación, Escuela Superior Politécnica del Litoral]. <http://www.dspace.espol.edu.ec/xmlui/handle/123456789/36369>



- Qin, S. J. (2003). Statistical process monitoring: Basics and beyond. *Journal of Chemometrics*, 17(8–9), 480–502. <https://doi.org/10.1002/cem.800>
- Rykov, V., Ivanova, N., y Morozov, E. (2025). Performance and Numerical Analysis of (GI| GI| N, M) Queues using Marked Markov Process. *Reliability: Theory & Applications*, 20(85), 61-82. <https://doi.org/10.24412/1932-2321-2025-885-61-82>
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., y Kavuri, S. N. (2003). A review of process fault detection and diagnosis. *Computers & Chemical Engineering*, 27(3), 293–311. [https://doi.org/10.1016/S0098-1354\(02\)00160-6](https://doi.org/10.1016/S0098-1354(02)00160-6)
- Zhang, X., Kano, M., & Li, Y. (2021). Recent advances in multivariate statistical process monitoring. *Journal of Process Control*, 104, 45–62. <https://doi.org/10.1016/j.jprocont.2021.05.006>
- Zhang, J., Zhao, C., y Gao, F. (2023). Interpretable multivariate process monitoring based on PCA and contribution analysis. *Process Safety and Environmental Protection*, 169, 204–216. <https://doi.org/10.1016/j.psep.2022.11.03>

Declaraciones

Contribución de los autores (CRediT):

Isidro Fabricio Morales Torres: Conceptualización del estudio, diseño metodológico, recopilación y procesamiento de datos, desarrollo del modelo, análisis de resultados, redacción del borrador original, revisión y edición final del manuscrito.

Conflicto de intereses:

El autor declara no tener conflictos de interés de naturaleza financiera, profesional ni personal que pudieran haber influido en el desarrollo o los resultados de esta investigación.

Financiamiento:

La presente investigación se desarrolló sin financiamiento externo y no recibió apoyo económico de agencias públicas, privadas ni organizaciones sin fines de lucro.

Aprobación ética:

Este estudio no involucró la participación de seres humanos ni el uso de animales. La investigación se basa exclusivamente en modelos matemáticos, simulación computacional y generación de datos sintéticos, por lo que no fue necesaria la aprobación de un comité de ética en investigación.



Disponibilidad de datos:

Los datos utilizados en este estudio corresponden a un proceso industrial simulado, generados específicamente para fines de investigación académica. Los conjuntos de datos y scripts empleados pueden ser proporcionados por el autor correspondiente previa solicitud razonable.

Licencia:

Este artículo se publica bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0), que permite su uso, distribución y reproducción en cualquier medio, siempre que se otorgue el crédito adecuado al autor original y se indique si se realizaron cambios.

Copyright (2026) © Isidro Fabricio Morales Torres
[Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

