

## Modelamiento estadístico multivariado para el control de las variables críticas en el proceso productivo de quesos madurados

*Multivariate statistical modeling for the control of critical variables in the production process of matured cheeses*

**Gabriela Joseth Serrano Torres\***  
Universidad Nacional de Chimborazo  
Riobamba - Ecuador  
gabriela.serrano@unach.edu.ec  
<https://orcid.org/0009-0005-7448-7610>

**Klever Hernán Torres Rodríguez**  
Universidad Nacional de Chimborazo  
Riobamba - Ecuador  
ktorres@unach.edu.ec  
<https://orcid.org/0000-0002-3919-0812>

\*Correspondencia:  
gabriela.serrano@unach.edu.ec

**Cómo citar este artículo:**  
Serrano, G., & Torres, K. (2026). Modelamiento estadístico multivariado para el control de las variables críticas en el proceso productivo de quesos madurados. *Esprint Investigación*, 5(1), 402-416. <https://doi.org/10.61347/ei.v5i1.257>

**Recibido:** 5 de enero de 2026  
**Aceptado:** 6 de febrero de 2026  
**Publicado:** 12 de febrero de 2026

**Copyright:** Derechos de autor 2026 Gabriela Joseth Serrano Torres, Klever Hernán Torres Rodríguez.



Esta obra está bajo una licencia internacional Creative Commons Atribución-NonComercial 4.0.

**Resumen:** La producción de quesos madurados se caracteriza por una alta complejidad debido a la interacción simultánea de múltiples variables que influyen en la calidad y el rendimiento del producto final. En este contexto, el control estadístico tradicional basado en enfoques univariados resulta limitado para capturar la dinámica real del proceso. El objetivo de este estudio fue desarrollar e implementar un modelo de control estadístico multivariado para monitorear y optimizar las variables críticas del proceso productivo de quesos madurados en la empresa Productos Alimenticios LAFRES, con el fin de mejorar la estabilidad y la calidad del producto final. La metodología adoptó un enfoque cuantitativo de tipo aplicado, con un diseño no experimental y longitudinal, analizando datos correspondientes a treinta lotes consecutivos de producción. Se aplicaron técnicas de análisis multivariado, incluyendo el Análisis de Componentes Principales (PCA) para la reducción de la dimensionalidad, la regresión por Mínimos Cuadrados Parciales (PLS) para modelar el rendimiento del queso y el control estadístico multivariado mediante los estadísticos  $T^2$  de Hotelling y el Squared Prediction Error (SPE). Los resultados mostraron que las cinco primeras componentes principales explicaron más del 74 % de la variabilidad total, confirmando el carácter multifactorial del proceso. El modelo PLS presentó un coeficiente de determinación de  $R^2 = 0.499$ , identificando la acidez de la leche, el tiempo de cuajado y el tiempo de prensado como las variables de mayor influencia positiva sobre el rendimiento. Los gráficos de control multivariado evidenciaron una estabilidad global del proceso, con desviaciones puntuales de carácter no sistemático.

**Palabras clave:** Análisis de componentes principales, control estadístico multivariado, gráficos de control, mínimos cuadrados parciales, quesos madurados, rendimiento productivo.

**Abstract:** The production of ripened cheeses is characterized by high complexity due to the simultaneous interaction of multiple variables that influence the quality and yield of the final product. In this context, traditional statistical control based on univariate approaches is limited in its ability to capture the true dynamics of the process. The objective of this study was to develop and implement a multivariate statistical process control model to monitor and optimize the critical variables of the ripened cheese production process at Productos Alimenticios LAFRES, with the aim of improving the stability and quality of the final product. The methodology adopted a quantitative applied approach with a non-experimental, longitudinal design, analyzing data corresponding to thirty consecutive production batches. Multivariate analysis techniques were applied, including Principal Component Analysis (PCA) for dimensionality reduction, Partial Least Squares (PLS) regression to model cheese yield, and multivariate statistical control using Hotelling's  $T^2$  and Squared Prediction Error (SPE) statistics. The results showed that the first five principal components explained more than 74% of the total variability, confirming the multifactorial nature of the process. The PLS model yielded a coefficient of determination of  $R^2 = 0.499$ , identifying milk acidity, coagulation time, and pressing time as the variables with the greatest positive influence on yield. Multivariate control charts evidenced overall process stability, with occasional non-systematic deviations.

**Keywords:** Control charts, matured cheeses, multivariate statistical control, partial least squares, principal component analysis, productive performance.

## 1. Introducción

La producción de quesos madurados es un proceso industrial complejo que integra múltiples variables fisicoquímicas que interactúan dinámicamente a lo largo del tiempo, tales como el pH, la humedad, la temperatura de maduración, la actividad de agua, la concentración de sal y el tiempo de maduración (Bansal et al., 2025; López et al., 2024). Estas variables influyen de manera conjunta en la textura, el sabor y la seguridad microbiológica del producto final, generando un entorno de producción caracterizado por una variabilidad inherente y desafíos significativos para garantizar una calidad consistente (Anastasiou et al., 2022).

En el contexto de la ciencia y la práctica del control de procesos, el modelado y control estadístico multivariado emerge como una herramienta robusta para supervisar simultáneamente múltiples variables correlacionadas, basándose en métodos estadísticos que permiten sintetizar grandes conjuntos de datos para detectar desviaciones del comportamiento esperado de un proceso monitoreado (Ramos et al., 2021). Este marco multivariado supera las limitaciones de los métodos tradicionales univariados al considerar las correlaciones entre las características de calidad y los procesos, lo que permite una detección más sensible y holística de variaciones fuera de control (Ramos et al., 2025).

Los procesos productivos en sectores industriales modernos, incluidos los alimentarios, son sistemas en los que la calidad del producto y la eficiencia operativa dependen directamente de la gestión integrada de múltiples variables (Herrera-Vidal et al., 2025). La literatura de ingeniería de calidad industrial describe el control estadístico de procesos como una disciplina que busca vigilar la estabilidad de los procesos, detectar cambios de manera oportuna y guiar decisiones operativas mediante métodos cuantitativos (Mora et al., 2025). Distintas metodologías de control permiten reducir el tiempo de detección de cambios y mejorar la robustez del monitoreo en presencia de datos no normales o interdependientes, lo cual es común en entornos de producción con múltiples variables críticas (Ramos et al., 2025).

Por lo tanto, el modelado y control estadístico multivariado se contextualiza específicamente para abordar los desafíos del control de variables críticas en procesos productivos complejos (Dargère et al., 2025). En entornos donde múltiples variables de proceso interactúan simultáneamente, los enfoques multivariados como el Análisis de Componentes Principales (PCA, por sus siglas en inglés) y los gráficos de control basados en el estadístico  $T^2$  de Hotelling permiten integrar la información conjunta de varias características de calidad, facilitando la identificación de patrones de comportamiento anómalos y proporcionando una visión más comprensiva de la dinámica del proceso en comparación con los métodos univariados clásicos (Anwer et al., 2025).

Estudios previos han demostrado la utilidad de técnicas multivariadas para monitorear y mejorar la calidad de productos y procesos. Anwer et al. (2025) aplicaron un marco de control estadístico multivariado para supervisar un proceso de manufactura mediante un análisis integral de datos de producción usando PCA y Hotelling  $T^2$ , logrando identificar desviaciones del comportamiento normal y mejorar la toma de decisiones basada en datos. De manera similar, la investigación realizada por Sibono et al. (2025), enfocada en el uso de sensores espectroscópicos integrados con modelos multivariados, demostró que las técnicas de control estadístico basadas en PCA pueden detectar fallas y desviaciones en procesos alimentarios en etapas tempranas. Grassi et al. (2022) exploraron estrategias de análisis multivariante para modelar parámetros que afectan la calidad del producto, resaltando la importancia de estas técnicas en la gestión de procesos de producción alimentaria altamente variables.

En el contexto regional, Causado-Rodríguez et al. (2025) emplearon análisis estadísticos multivariados, como PCA y análisis de clúster, para caracterizar y clasificar las propiedades

fisicoquímicas del queso costeño, demostrando la utilidad de estas técnicas para reducir la dimensionalidad de datos complejos y detectar patrones de variación entre grupos de muestras de queso tradicional en el Caribe colombiano. A nivel nacional, Viera-Molina et al. (2024) examinaron la relación entre el volumen de producción y la proporción de productos defectuosos en una empresa láctea de Latacunga, identificando que defectos en la coagulación y problemas durante la maduración representan puntos críticos que impactan directamente la eficiencia y la calidad del proceso productivo, lo que resalta la necesidad de métodos de control y monitoreo más integrales en producciones locales.

Si bien existen esfuerzos por mejorar la calidad y la estandarización en entornos productivos lácteos en Ecuador y la región, la aplicación sistemática de técnicas multivariadas para el control y monitoreo de variables críticas permanece poco desarrollada (Flores & Rodríguez, 2025). En este sentido, la empresa Productos Alimenticios LAFRES enfrenta desafíos significativos para garantizar un control de calidad efectivo y oportuno en su proceso de producción de quesos madurados. El sistema de control de calidad actualmente implementado se basa en métodos univariados que analizan cada variable por separado, limitando la capacidad para identificar patrones multivariados de comportamiento anómalo y las relaciones estructurales entre variables críticas.

En consecuencia, esta aproximación segmentada dificulta la detección temprana de desviaciones que afectan la consistencia del producto y comprometen la eficiencia operativa, reduciendo la competitividad de LAFRES en el mercado objetivo. Por tanto, es imperativo desarrollar un modelo de control estadístico multivariado que integre y monitoree de manera conjunta las variables críticas del proceso. El presente estudio tiene como objetivo desarrollar un modelo de control estadístico multivariado que permita monitorear y optimizar las variables críticas del proceso productivo de quesos madurados en Productos Alimenticios LAFRES, utilizando técnicas de análisis multivariado y gráficos de control para mejorar la calidad y la estabilidad del producto final.

Para ello, se plantean las siguientes preguntas de investigación: ¿cuáles son las variables críticas del proceso productivo de quesos madurados que influyen significativamente en la calidad final del producto?, ¿cómo puede el análisis de componentes principales (PCA) facilitar la identificación y reducción de las variables más influyentes en el control de calidad del proceso?, y ¿en qué medida el modelo PLS explica y predice el rendimiento del queso en condiciones reales de producción agroindustrial?

## 2. Metodología

### Enfoque de investigación y diseño del estudio

La presente investigación adoptó un enfoque cuantitativo de tipo aplicado, con alcance descriptivo-explicativo, centrado en el análisis y modelamiento estadístico del proceso productivo de queso madurado. Se empleó un diseño no experimental y longitudinal, recolectándose datos directamente del proceso real en la empresa Productos Alimenticios LAFRES, sin manipulación de variables, y analizándose a través de treinta lotes de producción consecutivos. Esta aproximación garantizó la captura de la variabilidad natural bajo condiciones operativas auténticas.

### Unidad de análisis y recolección de datos

La unidad de análisis consistió en lotes completos de queso madurado (30 lotes), registrándose información sobre la materia prima, las etapas del proceso productivo, las condiciones de operación y las características del producto final. Los datos provinieron de los registros internos de producción y

control de calidad, asegurándose la trazabilidad y representatividad necesarias para reflejar el comportamiento industrial real.

## Variables de estudio y preprocesamiento de datos

Las variables se agruparon en tres categorías interconectadas: (a) materia prima, que incluyó temperatura, pH, acidez y densidad de la leche, variables que definieron las condiciones iniciales del proceso e influyeron directamente en la coagulación y el rendimiento; (b) proceso productivo, que comprendió la temperatura y el tiempo de pasteurización, dosis de calcio y cuajo, temperatura y tiempo de cuajado, tamaño del grano, tiempo de reposo y agitación, tiempo de desuerado, así como presión y tiempo de prensado, correspondientes a las operaciones unitarias clave; y (c) producto final, que abarcó el rendimiento (en kg y g), pH del queso madurado, unidades producidas, pérdida de masa y presencia de defectos, indicadores utilizados para evaluar la calidad y estabilidad final.

Previo al análisis multivariado, se ejecutó un riguroso proceso de preprocesamiento, depuración y exploración de datos, que incluyó la verificación de valores faltantes, el análisis de consistencia de unidades, la estadística descriptiva (media, desviación estándar y coeficiente de variación) y la evaluación de la variabilidad entre lotes. Esta etapa reveló la presencia de variables con valores constantes o baja variabilidad, tales como la dosis de calcio, la dosis de cuajo y la presión de prensado, las cuales fueron excluidas por no aportar información estadísticamente relevante. Esta decisión permitió evitar redundancias y mejoró la estabilidad del modelo, en concordancia con los principios estándar del análisis multivariado.

## Análisis multivariado y modelamiento estadístico

Con el conjunto de variables depurado, se aplicó el Análisis de Componentes Principales (PCA) con el propósito de reducir la dimensionalidad, identificar relaciones entre variables, determinar contribuciones clave al proceso y explorar la estructura interna de los datos. Este procedimiento evidenció un comportamiento multifactorial del proceso, en el cual ninguna variable individual dominó de manera aislada, lo que justificó el uso de técnicas multivariadas.

Posteriormente, se construyó un modelo de regresión por Mínimos Cuadrados Parciales (PLS), utilizándose el rendimiento del queso como variable respuesta y las once variables previamente seleccionadas como variables predictoras. El modelo fue validado mediante validación cruzada, evaluándose indicadores como el error cuadrático medio (RMSE), el error absoluto medio (MAE) y el coeficiente de determinación ( $R^2$ ). Los resultados obtenidos confirmaron una capacidad adecuada del modelo para describir el comportamiento del proceso productivo.

Con base en este modelo, se implementaron herramientas de control estadístico multivariado, entre ellas el estadístico  $T^2$  de Hotelling, empleado para detectar desviaciones globales y lotes fuera de control, y el Squared Prediction Error (SPE), utilizado para analizar los residuales no explicados por el modelo. Estos indicadores permitieron evaluar la estabilidad del proceso, identificar desviaciones puntuales y validar su dinámica multivariada.

## Herramientas Analíticas

Los análisis se realizaron en RStudio, empleándose paquetes como tidyverse, FactoMineR, factoextra y pls, los cuales facilitaron el procesamiento descriptivo de los datos, la ejecución del PCA, la regresión PLS y el control estadístico multivariado ( $T^2$  y SPE). Esta selección de herramientas aseguró la reproducibilidad, coherencia y precisión en el manejo de datos industriales complejos.

### 3. Resultados

#### Análisis descriptivo de las variables

A partir de los datos correspondientes a 30 lotes de producción de queso madurado, el análisis descriptivo evidenció estabilidad en las variables de materia prima, tales como pH, acidez y densidad de la leche, las cuales presentaron baja dispersión, reflejando una calidad consistente de los insumos. Las variables asociadas al proceso productivo, incluyendo los tiempos de cuajado, agitación y desuerado, mostraron una variabilidad controlada, característica de un proceso artesanal estandarizado. En cuanto al producto final, el rendimiento se mantuvo cercano a los valores esperados, con fluctuaciones atribuibles a variaciones naturales de la materia prima y a ajustes operativos menores.

La Tabla 1 presenta las variables registradas para las tres fases del sistema productivo, junto con el coeficiente de variación (CV%) empleado como medida adicional de dispersión. Este análisis descriptivo permitió documentar los rangos operativos y de calidad, identificar variables con baja y alta dispersión, y establecer una base cuantitativa del proceso, necesaria para las etapas posteriores del análisis multivariado.

**Tabla 1**

*Variables correspondientes a cada fase del proceso*

Fase 1: Materia prima		
Variable	CV%	Interpretación
Temp_Leche (°C)	8.59%	Variabilidad moderada, influencia potencial en cuajado
Acidez_Leche (°D)	5.13%	Variación controlada, relevante para coagulación
pH_Leche	1.08%	Alta estabilidad, indicador de calidad
Densidad_Leche	0.18%	Muy estable, control adecuado
Fase 2: Proceso		
Variable	CV (%)	Importancia
Reposo antes de agitar (min)	16.2%	Influye en la estructura del coágulo
Tiempo desuerado (min)	21.3%	Alta influencia en pérdida de humedad
Tiempo cuajado (min)	6.3%	Controla la firmeza del gel
Tamaño de grano (cm)	4.12%	Afecta drenaje y textura
Tiempo de prensado (h)	5.25%	Relacionado con humedad final
Tiempo de agitado (min)	3.81%	Controla sinéresis
Fase 3: Maduración		
Variable	CV (%)	Interpretación
Pérdida de masa (g)	60.3%	Muy alta variabilidad, indicador crítico
pH promedio del queso	10.5%	Relacionado con calidad y estabilidad
Rendimiento (kg)	3.82%	Indicador de eficiencia
Unidades producidas	4.29%	Refleja productividad
Peso promedio del queso	1.02%	Alta estandarización

## Análisis de Componentes Principales (PCA)

La selección del número de componentes principales se fundamentó en la minimización del error de reconstrucción de rango reducido. La aproximación del sistema mediante  $k$  componentes se expresó como:

$$\hat{\mathbf{Z}}_k = \mathbf{T}_k \mathbf{W}_k^T,$$

El error cuadrático total se determinó mediante la expresión:

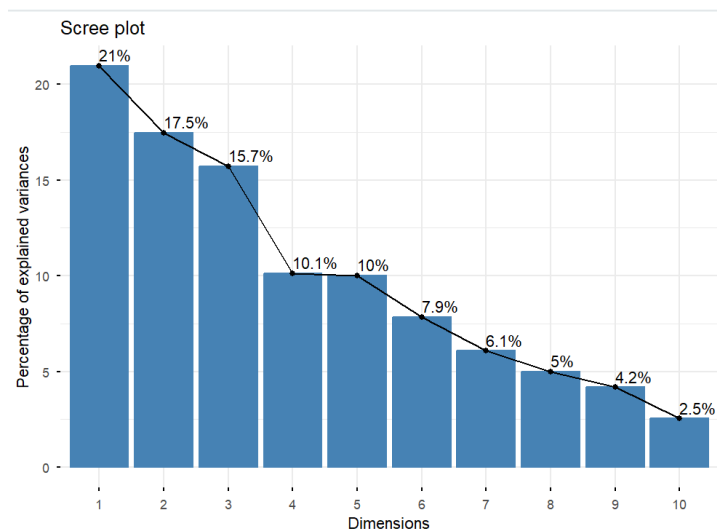
$$\|\mathbf{Z} - \hat{\mathbf{Z}}_k\|_F^2 = (n - 1) \sum_{j=k+1}^p \lambda_j,$$

Esta expresión demostró que la retención de los primeros componentes principales minimizó el error de reconstrucción global del sistema. En consecuencia, el PCA aplicado al conjunto de datos depurado redujo la dimensionalidad y permitió elucidar relaciones relevantes entre variables.

La primera componente principal (PC1) capturó el 21% de la variabilidad total, seguida de la PC2 con 17.5%, mientras que las primeras cinco componentes explicaron más del 74% de la varianza total, porcentaje considerado adecuado para la representación de un proceso productivo real. La Figura 1 ilustró gráficamente el aporte individual de cada componente principal.

**Figura 1**

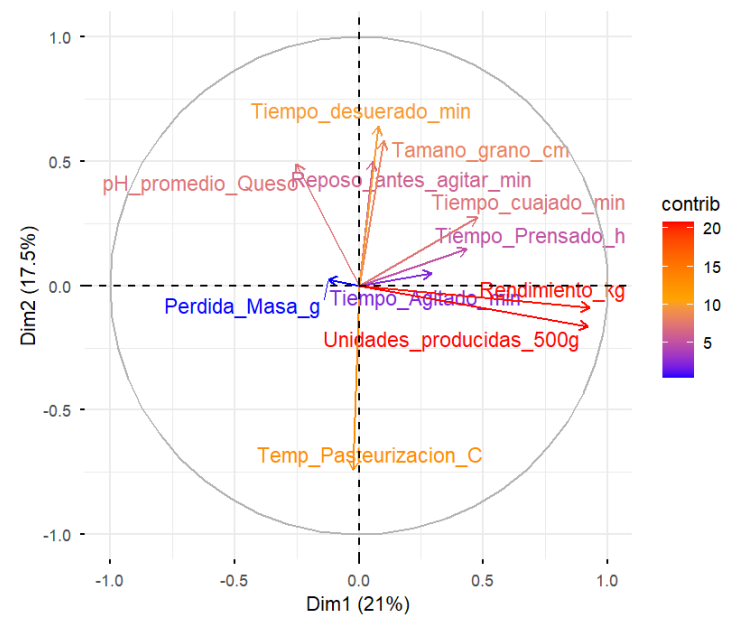
*Análisis de Componentes Principales (PCA)*



La Figura 2 representó el círculo de correlaciones entre PC1 y PC2, donde cada vector correspondió a una variable del proceso. La dirección indicó el tipo de relación, la longitud reflejó su contribución relativa y el color representó el peso de contribución. Se observó que el primer componente principal estuvo dominado por variables asociadas al rendimiento del proceso, tales como el tiempo de agitado, el tiempo de prensado y el rendimiento total, lo que permitió interpretarlo como un eje de eficiencia productiva. El segundo componente se asoció principalmente con variables fisicoquímicas, como el tamaño de grano, el tiempo de desuerado y el pH del queso, reflejando la dinámica estructural del producto durante la maduración.

**Figura 2**

Círculo de correlaciones PCA



La decisión final sobre el número de componentes seleccionados se sustentó en la evaluación de la proporción de varianza explicada individual y acumulada, definida como:

$$PVE_k = \frac{\lambda_k}{\sum_{j=1}^p \lambda_j}, \quad CPVE(K) = \sum_{k=1}^K PVE_k.$$

Bajo estos criterios, las primeras cinco componentes cumplieron simultáneamente con un porcentaje de varianza acumulada superior al 70% y con la regla de Kaiser ( $\lambda_k > 1$ ), justificando su selección para el análisis posterior.

### Control Estadístico Multivariado de Procesos (MSPC)

Una vez definido el subespacio latente a partir de las componentes principales retenidas, el monitoreo del comportamiento global del proceso se realizó mediante el estadístico  $T^2$  de Hotelling, el cual permitió evaluar la magnitud de cada observación en el espacio multivariado reducido.

Sea  $\mathbf{t} \in \mathbb{R}^k$  el vector de *scores* asociado a un lote de producción proyectado sobre las  $k$  componentes principales retenidas, y sea  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_k)$  la matriz diagonal de los autovalores correspondientes. El estadístico de Hotelling se definió como:

$$T^2 = \mathbf{t}^T \Lambda^{-1} \mathbf{t}.$$

Bajo el supuesto de normalidad multivariada en el espacio de componentes principales, el límite superior de control del estadístico  $T^2$  se aproximó mediante la distribución F, de acuerdo con:

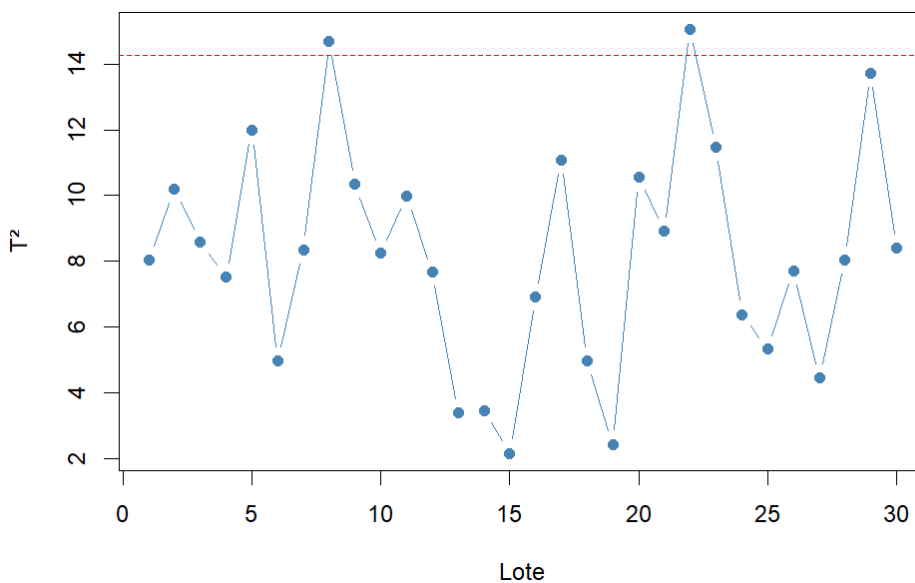
$$T_{\text{lim}}^2 = \frac{k(n-1)}{n-k} F_{k, n-k; \alpha},$$

Donde  $n$  representó el número de observaciones,  $k$  el número de componentes principales retenidas y  $\alpha$  el nivel de significancia.

En el análisis realizado, considerando  $k = 5$  componentes principales y  $n = 30$  lotes de producción, el límite de control calculado fue  $T_{lim}^2 = 15.097$ . Al comparar este umbral con los valores observados para cada lote (Figura 3), se constató que todos los puntos permanecieron por debajo del límite de control, lo que evidenció la ausencia de causas especiales de variación en el espacio latente y confirmó la estabilidad global del proceso productivo desde una perspectiva multivariada.

**Figura 3**

*Estadístico T2 de Hotelling*



De manera complementaria al control global proporcionado por el estadístico  $T^2$  de Hotelling, se evaluó la variabilidad no explicada por el modelo PCA mediante el estadístico SPE, también conocido como estadístico Q. Este indicador cuantificó la magnitud del error de reconstrucción de cada observación fuera del subespacio latente retenido.

Sea  $\mathbf{x}$  el vector original de variables estandarizadas y  $\hat{\mathbf{x}}$  su proyección reconstruida a partir de las componentes principales retenidas. El estadístico SPE se definió como:

$$Q = \|\mathbf{x} - \hat{\mathbf{x}}\|^2 = \mathbf{e}^T \mathbf{e}.$$

El límite superior del estadístico SPE se estimó mediante una aproximación basada en los momentos de la distribución residual, definida como:

$$Q_{lim} = \theta_1 \left[ \frac{z_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0},$$

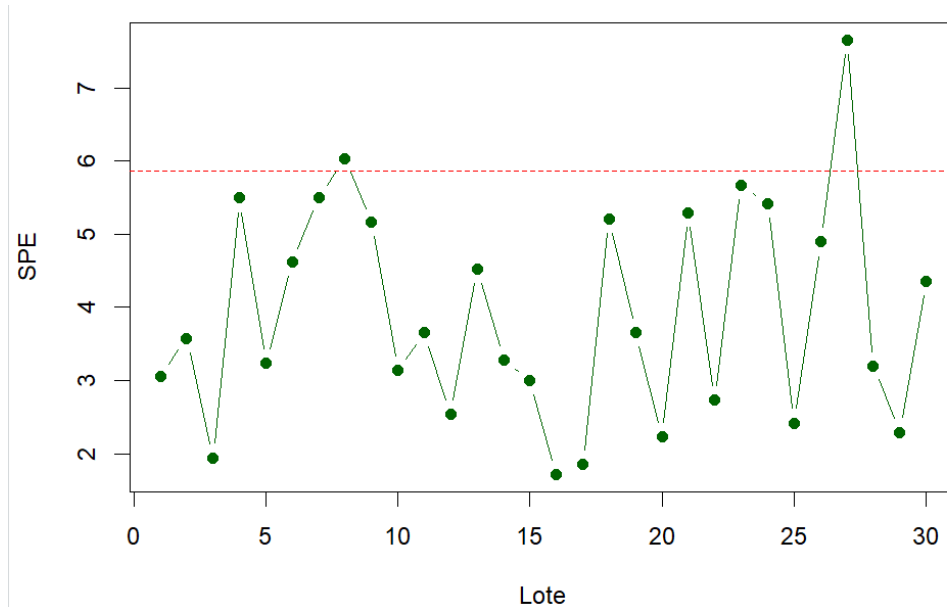
Donde los parámetros  $\theta_i$  dependieron de los autovalores asociados a las componentes descartadas y  $z_\alpha$  correspondió al cuantil de la distribución normal estándar.

El valor aproximado del umbral  $Q_{lim}$  fue 5.87. El análisis del estadístico SPE (figura 4) mostró que la mayoría de los lotes se ubicaron por debajo de este límite, identificándose un único lote con una desviación puntual, la cual no presentó un patrón sistemático ni recurrente. Este resultado indicó que

el modelo PCA capturó adecuadamente la estructura dominante del proceso y que la anomalía detectada correspondió a un evento aislado.

**Figura 4**

Estadístico SPE



## Modelo PLS e Importancia de Variables

La regresión por Mínimos Cuadrados Parciales (PLS) se fundamentó en la construcción de variables latentes que maximizaron la relación entre la matriz de predictores  $X$  y la variable respuesta  $y$ . A diferencia del Análisis de Componentes Principales, que buscó direcciones de máxima varianza interna de  $X$ , el modelo PLS determinó el vector de pesos  $w$  que maximizó la covarianza entre el score latente  $t = Xw$  y el rendimiento del queso, resolviendo el problema:

$$\max \text{Cov}(Xw, y).$$

Este criterio explicó que los objetivos de reducción dimensional del PCA y del PLS fueran distintos, ya que el primero describió la estructura interna del sistema productivo, mientras que el segundo identificó las combinaciones de variables con mayor capacidad explicativa de la respuesta.

La validación cruzada del modelo evidenció que el error predictivo mínimo se alcanzó utilizando un único componente latente, cumpliéndose que:

$$RMSEP(1) = 0.5018 < RMSEP(k), \forall k > 1,$$

La Figura 5 resumió el ajuste del modelo de regresión PLS, considerando 30 observaciones, 11 variables predictoras seleccionadas y una variable respuesta correspondiente al rendimiento. El análisis del error de predicción (RMSEP) mostró que el menor valor se obtuvo utilizando un solo componente latente (0.5018), lo que indicó que el comportamiento del rendimiento del queso pudo ser explicado adecuadamente mediante una estructura unidimensional del proceso. Adicionalmente, el modelo alcanzó un RMSE de 0.376, un MAE de 0.332 y un coeficiente de determinación  $R^2 = 0.499$  mediante validación cruzada, lo que indicó que el 49.9 % de la variabilidad del rendimiento fue explicada por las variables del proceso incluidas en el modelo.

**Figura 5***Resumen del ajuste del modelo*

```

> summary(pls_fit)
Data:  X dimension: 30 11
      Y dimension: 30 1
Fit method: kernelpls
Number of components considered: 11

VALIDATION: RMSEP
Cross-validated using 30 leave-one-out segments.
      (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps
CV          0.4647  0.5018  0.5498  0.5589  0.5477  0.5550  0.5582  0.5605
adjCV       0.4647  0.4998  0.5463  0.5551  0.5440  0.5512  0.5544  0.5566
      8 comps  9 comps 10 comps 11 comps
CV          0.5631  0.5638  0.5637  0.5637
adjCV       0.5592  0.5599  0.5597  0.5597

TRAINING: % variance explained
      1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps
X          15.38  26.23  39.78  47.40  55.88  62.31  70.68  79.00
rendimiento_kg  29.91  39.73  43.58  45.35  45.62  45.81  45.87  45.88
      9 comps 10 comps 11 comps
X          89.15  93.61 100.00
rendimiento_kg  45.88  45.88  45.88

```

El análisis de importancia de variables (VIP) del modelo PLS identificó la acidez, el pH y la temperatura de la leche como las variables de mayor influencia relativa, aunque sin dominancia individual, lo que indicó que no existió una única variable dominante y que el rendimiento dependió de una combinación de factores. Este resultado reforzó el carácter multifactorial del sistema y la ausencia de dependencias excesivas. La Tabla 2 presentó los valores de importancia asociados a cada variable analizada.

**Tabla 2***Importancia de las variables (VIP)*

Variable	VIP
Acidez de la leche (acidez_leche_d)	0.6070
pH de la leche (p_h_leche)	0.5015
Temperatura de la leche (temp_leche_c)	0.2651
Densidad de la leche (densidad_leche_g_ml)	0.2227
Tiempo de cuajado (tiempo_cuajado_min)	0.1772
Temperatura de pasteurización temp_pasteurizacion_c	0.0074
Reposo antes de agitar (reposo_antes_agitar_min)	0.0055
Tamaño del grano (tamano_grano_cm)	0.0033
Tiempo desuerado (tiempo_desuerado_min)	0.0012
Tiempo de agitado (tiempo_agitado_min)	0.0010
Tiempo de prensado (tiempo_prensado_h)	0.0003

## Modelo matemático final

A partir de las variables estandarizadas, el ajuste del modelo PLS mediante validación cruzada identificó un único componente latente como óptimo, a partir del cual se obtuvieron los coeficientes del modelo matemático final (Ecuación 1). Estos coeficientes representaron los pesos relativos derivados de la regresión:

$$\begin{aligned}
 \widehat{\text{Rendimiento}}_{kg} &= 0.1320(\text{Acidez\_leche}) + 0.1118(\text{Tiempo\_cuajado}) \\
 &+ 0.0704(\text{Tiempo\_prensado}) + 0.0577(\text{Densidad\_leche}) \\
 &- 0.0550(\text{pH\_leche}) + 0.0263(\text{Tiempo\_desuerado}) \\
 &+ 0.0232(\text{Temp\_leche}) - 0.0209(\text{Reposo\_antes\_agitar}) \\
 &+ 0.0125(\text{Tiempo\_agitado}) - 0.0051(\text{Tamaño\_grano}) \\
 &- 0.0044(\text{Temp\_pasteurización})
 \end{aligned}
 \tag{Ecuación 1}$$

Dado que las variables fueron estandarizadas, los coeficientes reflejaron efectos relativos. La acidez de la leche ( $\beta = 0.132$ ) destacó como la variable con mayor influencia positiva, seguida por el tiempo de cuajado ( $\beta = 0.112$ ) y el tiempo de prensado ( $\beta = 0.070$ ), los cuales favorecieron una coagulación eficiente y una adecuada expulsión del suero. La densidad de la leche ( $\beta = 0.058$ ) contribuyó positivamente al capturar sólidos útiles.

Entre las variables con efecto negativo se encontraron el pH de la leche ( $\beta = -0.055$ ), el reposo excesivo antes de agitar ( $\beta = -0.021$ ) y el tamaño de grano ( $\beta = -0.005$ ), los cuales redujeron el rendimiento al favorecer la retención de suero. Las variables con impacto reducido, como la temperatura de la leche, el tiempo de agitado, el tiempo de desuerado y la temperatura de pasteurización, evidenciaron una estabilidad operativa controlada del proceso.

## 4. Discusión

Debido a que el modelo fue estimado mediante regresión por Mínimos Cuadrados Parciales (PLS), la validación estadística no se realizó sobre coeficientes individuales mediante pruebas t, sino sobre el comportamiento global del modelo, evaluando su capacidad explicativa y predictiva. Este enfoque resulta consistente con la naturaleza del PLS, el cual opera sobre componentes latentes y es especialmente adecuado en contextos con multicolinealidad entre variables predictoras.

En este contexto, se planteó la hipótesis nula ( $H_0$ ): el modelo de regresión PLS no presenta capacidad explicativa significativa, por lo que la relación entre las variables del proceso y el rendimiento del queso es atribuible al azar; y la hipótesis alternativa ( $H_1$ ): el modelo de regresión PLS presenta capacidad explicativa significativa, indicando que al menos una de las variables del proceso influye de manera sistemática sobre el rendimiento del queso.

La hipótesis nula fue rechazada de manera estadísticamente significativa mediante una prueba de permutación, obteniéndose un p-valor de 0.041 ( $< 0.05$ ). En consecuencia, se aceptó la hipótesis alternativa, validando que el modelo explica el 49.9 % de la variabilidad del rendimiento del queso ( $R^2 = 0.499$ , RMSE = 0.376, MAE = 0.332). Estos indicadores evidencian un desempeño predictivo moderado pero consistente, lo cual permite inferir que el modelo es estadísticamente significativo y que sus resultados pueden generalizarse a la población bajo condiciones operativas similares. Este tipo de validación global es particularmente apropiado para modelos PLS, dada su capacidad para manejar estructuras latentes y correlaciones elevadas entre variables del proceso.

Los hallazgos obtenidos se alinean con investigaciones recientes en procesos lácteos analizados mediante enfoques multivariados. Grassi et al. (2019) demostraron que la aplicación de técnicas de control estadístico multivariante (MSPC), empleando métodos como  $T^2$  de Hotelling y el estadístico Q, permite detectar fallas en etapas tempranas del proceso de coagulación del queso, identificando patrones de variación asociados al pH, la temperatura y la composición de la leche. Estos resultados refuerzan la relevancia de integrar múltiples variables correlacionadas para el monitoreo efectivo de procesos lácteos complejos, tal como se evidenció en el presente estudio.

De manera similar, Rodríguez et al. (2025) aplicaron un enfoque multivariante al análisis de parámetros fisicoquímicos durante la maduración de quesos tradicionales, combinando PCA con estadísticas basadas en  $T^2$  de Hotelling, lo cual resultó útil para la exploración de agrupamientos y variaciones entre muestras. Asimismo, Tangorra et al. (2025) emplearon control estadístico multivariante complementado con técnicas de aprendizaje y modelos robustos, identificando que estos enfoques no solo permiten detectar desviaciones del comportamiento normal del proceso, sino también asociar dichas desviaciones con variables críticas interdependientes. La evidencia conjunta respalda la importancia de integrar múltiples variables correlacionadas para el monitoreo efectivo de procesos lácteos complejos, coherente con los resultados aquí presentados.

En este sentido, el modelo PLS desarrollado proporciona una herramienta práctica y aplicable para la empresa Productos Alimenticios LAFRES, ya que permite priorizar el control de variables clave como la acidez y el pH de la leche, así como los tiempos de cuajado y prensado, las cuales mostraron una influencia significativa sobre el rendimiento del queso. Adicionalmente, los gráficos de control  $T^2$  y SPE facilitan el monitoreo por lotes, permitiendo detectar lotes atípicos antes de completar la etapa de maduración, lo que contribuye a la reducción de pérdidas y a la mejora de la eficiencia del proceso productivo. La integración de PCA, el modelo PLS y el control estadístico multivariado constituye así una estrategia robusta para la estandarización de procesos artesanales de quesos madurados.

Entre las principales limitaciones del estudio se identificó el tamaño muestral relativamente reducido ( $n = 30$  lotes), el cual podría ser sensible a variaciones estacionales no capturadas en el análisis, así como un coeficiente de determinación moderado ( $R^2 = 0.499$ ), que indica que aproximadamente el 50 % de la variabilidad del rendimiento permanece sin explicar, posiblemente asociada a factores no medidos o de naturaleza no lineal. Además, el modelo PLS asume una relación lineal aproximada entre variables, lo que podría subestimar efectos no lineales presentes en procesos de maduración prolongada.

Futuras investigaciones podrían ampliar el tamaño muestral mediante estudios multisitio, incorporar variables microbiológicas o datos provenientes de sensores IoT, y explorar extensiones del modelo PLS, como SO-PLS, para capturar estructuras más complejas. Asimismo, la validación con datos prospectivos y la optimización mediante diseño de experimentos (DoE) permitirían maximizar el rendimiento del proceso, mientras que la integración con modelos de aprendizaje automático facilitaría la comparación de la robustez predictiva de PLS frente a enfoques no lineales en escenarios de alta dimensionalidad.

## 5. Conclusiones

Las variables críticas que influyeron significativamente en la calidad final del queso madurado, particularmente en su rendimiento, fueron identificadas de manera consistente mediante el enfoque multivariado aplicado. Entre ellas, la acidez de la leche se consolidó como el principal factor positivo, seguida por el tiempo de cuajado, el tiempo de prensado y la densidad de la leche, las cuales favorecieron la formación de una cuajada eficiente y la retención de sólidos útiles. En contraste, los efectos negativos más relevantes se asociaron al pH de la leche, al reposo prolongado antes del agitado

y al tamaño del grano, variables que tendieron a reducir el rendimiento al favorecer la retención de suero. Si bien el proceso exhibió estabilidad en variables operativas controladas, como la dosis de cuajo y la presión aplicada, el rendimiento dependió de interacciones multifactoriales, y no del control aislado de parámetros individuales.

El Análisis de Componentes Principales (PCA) facilitó la identificación de las variables más influyentes al reducir la dimensionalidad del sistema, permitiendo una interpretación clara de la estructura interna del proceso. Las dos primeras componentes principales explicaron aproximadamente el 38.5 % de la variabilidad total, mientras que las cinco primeras superaron el 74 %, evidenciando contribuciones elevadas de la acidez, el pH y la temperatura de la leche en el círculo de correlaciones. Esta técnica confirmó el carácter multifactorial del proceso productivo, eliminó redundancias entre variables altamente correlacionadas y sirvió como base metodológica para el modelado posterior, optimizando el control de calidad sin pérdida significativa de información relevante.

El estudio validó un modelo de regresión por Mínimos Cuadrados Parciales (PLS) robusto y estadísticamente significativo para el monitoreo y la mejora del proceso productivo en la empresa Productos Alimenticios LAFRES. Los resultados demostraron que la integración de técnicas multivariadas como PCA, PLS y gráficos de control basados en los estadísticos  $T^2$  de Hotelling y SPE constituye una estrategia eficaz, reproducible y adaptable a otros procesos queseros, previa consideración de ajustes contextuales y operativos. En conjunto, este enfoque multivariado apoyó de manera objetiva la toma de decisiones técnicas, contribuyendo a una mayor estabilidad del proceso, una reducción de pérdidas y un incremento potencial del rendimiento del queso madurado.

## Referencias

- Anastasiou, R., Kazou, M., Georgalaki, M., Aktypis, A., Zoumpopoulou, G., & Tsakalidou, E. (2022). Omics Approaches to Assess Flavor Development in Cheese. *Foods*, 11(2). <https://doi.org/10.3390/foods11020188>
- Anwer, S., Ahmad, R., Siddiqui, F., Ameen, R., & Rehman, W. (2025). Multivariate statistical techniques for quality control and process monitoring. *Journal of Media Horizons*, 6(6), 117–133. <https://jmhorizons.com/index.php/journal/article/view/938>
- Bansal, V., Singh, P., Goel, N., Deshmukh, G., Mishra, S., Sihag, M., & Chawla, R. (2025). Understanding pH dynamics in cheesemaking: Insights across different cheese varieties. *Journal of Food Science and Technology*, 62(12), 2233–2243. <https://doi.org/10.1007/s13197-025-06445-2>
- Causado-Rodríguez, E., Sanchez, J., & Fernández, A. (2025). Classification and Typification of Costeño Cheese in the Colombian Caribbean Through Multivariate Analysis of Physico-chemical Parameters: Base Study for the department of Magdalena. *Journal of Posthumanism*, 5(5), 1195–1217. <https://doi.org/10.63332/joph.v5i5.1433>
- Dargère, A., de Jesus, J., Ferrão, S., Santos, L., & Faria, P. (2025). A multivariate approach to verify correlations between sensory quality indices and analytical quality indices in traditional Brazilian cheese. *Journal of Food Science and Technology*, 62(9), 1764–1771. <https://doi.org/10.1007/s13197-024-06145-3>
- Flores, C., & Rodríguez, M. (2025). Aplicación de técnicas multivariadas en indicadores de productividad procedentes de la simulación de procesos de producción de una empresa láctea. *INNOVATION & DEVELOPMENT IN ENGINEERING AND APPLIED SCIENCES*, 7(2), 14–14. <https://doi.org/10.53358/ideas.v7i2.1248>

- López, N., Abarquero, D., Combarros-Fuertes, P., Prieto, B., Fresno, J., & Tornadijo, M. (2024). Influence of Salting on Physicochemical and Sensory Parameters of Blue-Veined Cheeses. *Dairy*, 5(1), 93–105. <https://doi.org/10.3390/dairy5010008>
- Grassi, S., Strani, L., Alamprese, C., Pricca, N., Casiraghi, E., & Cabassi, G. (2022). A FT-NIR Process Analytical Technology Approach for Milk Renneting Control. *Foods*, 11(1), 33. <https://doi.org/10.3390/foods11010033>
- Grassi, S., Strani, L., Casiraghi, E., & Alamprese, C. (2019). Control and Monitoring of Milk Renneting Using FT-NIR Spectroscopy as a Process Analytical Technology Tool. *Foods*, 8(9), 405. <https://doi.org/10.3390/foods8090405>
- Herrera-Vidal, G., Salcedo-Toro, G., Blanco-Camacho, J., Cabarcas-Sierra, A., & Doria-Orozco, N. (2025). Revisión de Literatura sobre Gestión de la Producción: Tendencias, Desafíos y Oportunidades para la Ingeniería Industrial. *Boletín de Innovación, Logística y Operaciones*, 7(1), 67–79. <https://doi.org/10.17981/bilo.7.1.2025.07>
- Mora, J., Guzmán, B., Suárez, J., & León, J. (2025). Aplicación del Control Estadístico de Procesos: Un caso de estudio. *Ciencia Latina Revista Científica Multidisciplinar*, 9(3), 440–462. [https://doi.org/10.37811/cl\\_rcm.v9i2.17639](https://doi.org/10.37811/cl_rcm.v9i2.17639)
- Ramos, M., Paredes, P., Morales, E., & Vistin, J. (2025). Evaluación comparativa de metodologías de control estadístico de procesos en la gestión de la calidad industrial. *Polo del Conocimiento*, 10(10), 727–744. <https://polodelconocimiento.com/ojs/index.php/es/article/view/10567>
- Ramos, M., Ascencio, J., Hinojosa, M., Vera, F., Ruiz, O., Jimenez-Feijoó, M., & Galindo, P. (2021). Multivariate statistical process control methods for batch production: A review focused on applications. *Production & Manufacturing Research*, 9(1), 33–55. <https://doi.org/10.1080/21693277.2020.1871441>
- Rodríguez, J., Suárez, P., Vázquez, L., Flórez, A. B., Vivar-Quintana, A., & Mayo, B. (2025). Metataxonomic profiling of microbial communities and metabolic analyses of the traditional Spanish raw cow's milk cheese 'Casín' from manufacture to ripening. *Frontiers in Microbiology*, 16. <https://doi.org/10.3389/fmicb.2025.1722502>
- Sibono, L., Tronci, S., Barsøe, M., Errico, M., & Grosso, M. (2025). Raman Spectroscopy Coupled with Multivariate Statistical Process Control for Detecting Anomalies During Milk Coagulation. *Processes*, 13(11), 3519. <https://doi.org/10.3390/pr13113519>
- Tangorra, F., Stojsavljevic, D., & Dzidic, A. (2025). Model-driven multivariate control chart and support vector machine as tools to detect variation in the milking process and monitor parlor performance. *Journal of Dairy Science*, 108(9), 9778–9787. <https://doi.org/10.3168/jds.2025-26612>
- Viera-Molina, G. (2024). Análisis de la Producción y Defectos en Queso Fresco: Estrategias para la Mejora de la Calidad y Eficiencia. *REVISTA DE INVESTIGACIÓN SIGMA*, 11(02). <https://journal.espe.edu.ec/ojs/index.php/Sigma/es/article/view/3537>

---

## Transparencia

### Conflicto de interés

Los autores declaran que no existen conflictos de interés de naturaleza alguna como parte de la presente investigación.

### Fuente de financiamiento

Los autores financiaron completamente la investigación.

### Contribución de autoría

Gabriela Joseth Serrano Torres: Conceptualización, metodología, software, validación, análisis formal, investigación, gestión de datos, visualización, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, administración del proyecto, recursos, supervisión.

Klever Hernán Torres Rodríguez: Conceptualización, metodología, validación, análisis formal, investigación, gestión de datos, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, recursos.

Los autores contribuyeron activamente en el análisis de los resultados, revisión y aprobación del manuscrito final.