

SISTEMA DE CONTROL MULTIVARIADO NO PARAMÉTRICO DE PROCESOS

M.C. Federico Zertuche Luis¹, Dr. Mario Cantú Sifuentes² y
Dr. Manuel R. Piña Monarrez³

¹ Estudiante de doctorado del programa PICYT con opción Terminal en Ingeniería Industrial y de Manufactura. Centro de Investigación COMIMSA. (el artículo está basado sobre su Disertación Doctoral). ² Programa de posgrado en Estadística. Universidad Autónoma Agraria Antonio Narro. ³ Profesor Investigador del programa PICYT. Centro de Investigación COMIMSA.

RESUMEN

Debido a que el control de procesos multivariado clásico como lo es la gráfica T^2 de Hotelling, cuando el supuesto de normalidad no se cumple, ha mostrado ineficiencia en la detección e identificación de las variables fuera de control. En este trabajo se presenta un sistema de control multivariado no paramétrico que identifica la señal fuera de control y determina la(s) variable(s) que la causa(n) (ya sea que se cumpla o no el supuesto de normalidad), a través de una gráfica de contribuciones, la cual es la principal aportación del artículo. Una aplicación en la industria automotriz se presenta en la sección 3.

Palabras Claves: Gráfica de comparación de rangos, bootstrap, componentes principales gráficas de contribuciones

1. INTRODUCCIÓN.

Cuando las características de calidad de un producto están correlacionadas y se aplica el control estadístico univariado, como lo es la gráfica de control de medias y rangos, este detecta un gran número de señales falsas que identifican al proceso como fuera de control (Cook, 1998).

Dentro de los procedimientos de control multivariado las gráficas con más investigación y aplicación en la industria son (Stoumbos, 2000): La gráfica T^2 de Hotelling, la gráfica de control multivariada de sumas acumuladas (MCUSUM) y La gráfica de control multivariada exponencial con peso promedio móvil (MEWMA). Siendo la gráfica T^2 de Hotelling el procedimiento de control más familiar para monitorear vectores de medias de varias características de calidad correlacionadas en un proceso de producción (AIAG, 2005).

Lamentablemente, el supuesto básico para su implementación, es que el conjunto de las características a controlar tienen un comportamiento normal multivariado, supuesto que generalmente no se cumple, por lo que no es razonable modelar el comportamiento aleatorio de las características de calidad de interés mediante esta distribución, ya que conduce a un mal desempeño de las gráficas (Stoumbos, 2000).

Un enfoque de control no paramétrico para los procesos multivariados, fue inicialmente propuesto por Liu (1995), ella propone una carta de control, llamada gráfica de comparación de rangos (r chart), basada en el concepto de profundidad de datos. La figura 1, representa el esquema propuesto por Liu (2000) para el control del proceso utilizando la r chart.

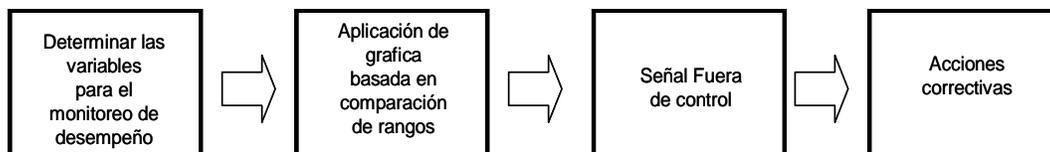


FIGURA 1. Esquema de monitoreo aplicando la gráfica de comparación de rangos.

Una aplicación reciente del esquema de control propuesto por Liu es presentada por Messaoud (2006), la aplicación fue realizada en un proceso de manufactura donde se controlan dos variables correlacionadas que afectan la profundidad de taladrado de un producto denominado BTA. En este trabajo se resalta que es necesaria investigación encaminada a identificar las variables que causan una señal fuera de control en este tipo de gráficas.

Chakraborty (2001) y Stoumbous (2000) mencionan que los métodos de control multivariados no paramétricos han sido poco aplicados en la industria y enfatizan que no existen métodos para interpretar las señales fuera de control en términos de la identificación de las variables causantes de estas. Esto es particularmente cierto en la industria automotriz (AIAG, 2005).

Por otra parte dentro de las investigaciones en el campo del control multivariado, se ha encontrado que cuando las variables a controlar no pueden modelarse razonablemente mediante una distribución normal multivariada, las gráficas de control basadas en la profundidad de datos son más eficientes que las basadas en la T^2 de Hotelling y cuando siguen un comportamiento normal

multivariado, la eficiencia es similar (Liu, 1999 y Zertuche, 2004). En este sentido, el principal aporte del artículo, se centra en el desarrollo de un sistema de control multivariado no paramétrico, basado en la profundidad de Mahalanobis y en componentes principales, que identifica cuando un proceso de producción está fuera de control y determina las variables causantes de esta anomalía. El sistema propuesto, se aplica mediante un programa computacional en Matlab 7.0, a un proceso de producción de la industria automotriz.

2. SISTEMA DE CONTROL MULTIVARIADO NO PARAMÉTRICO

Para los objetivos del trabajo $X = (X_1, X_2, \dots, X_p)$ denota un p -vector aleatorio que colecta las p características de calidad a controlar. El sistema que se propone se esquematiza en la figura 2 y extiende al propuesto por Liu (2000) figura 1, siendo la diferencia principal, que el sistema propuesto realiza un análisis para identificar la variable o variables que son la causa de una señal de alarma en una gráfica de control multivariada basada en la comparación de rangos.

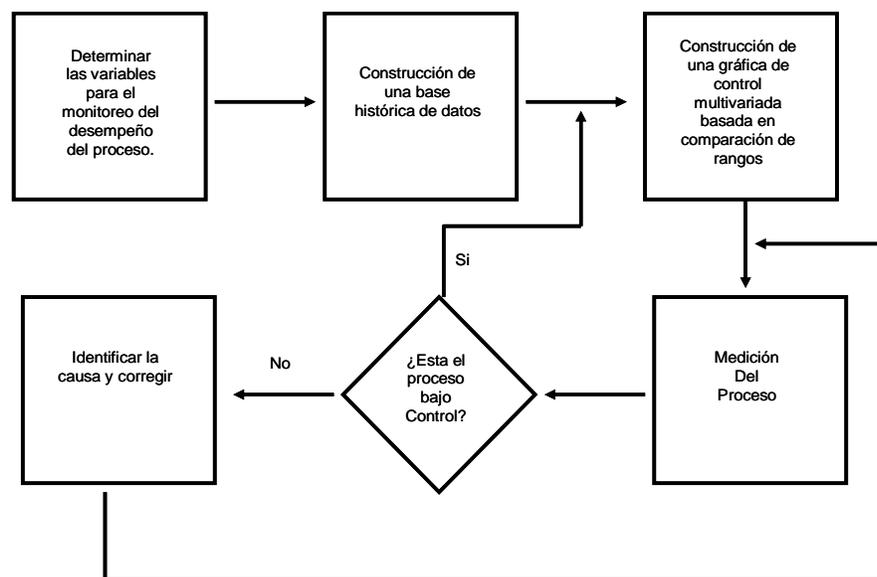


FIGURA 2. Esquema del sistema de control de procesos propuesto

El sistema mostrado en la figura 2, es presentado como un procedimiento de pasos los cuales se discuten a continuación:

Paso 1: Determinar las variables para el monitoreo del desempeño del proceso.-La selección de las variables a controlar en todo proceso de producción es hecha por el cliente. Generalmente se envía a la empresa el diseño del producto con las tolerancias de diseño. De estas, el personal de ingeniería de calidad determina cuales son las variables críticas; estas son las que hay que controlar (Macgregor, 1995).

Paso 2. Seleccione una base histórica de datos.- Antes de construir un sistema de control es necesario contar con datos de referencia, llamados base histórica de datos, la cual está constituida por mediciones de características que cumplen con los requerimientos de calidad especificados por el cliente (paso 1) (Macgregor, 1995).

Paso 3. A partir de la distribución de referencia construya una gráfica de control basada en rangos.-Uno de los principales pasos del sistema de control es la aplicación de un método multivariado no paramétrico, entre los que se encuentran los métodos de profundidad de datos. Estos empezaron prácticamente con el artículo de Liu (1992), siendo esta, la primera aplicación de la metodología en el control de calidad. El único requerimiento para usar

este método, es contar con una distribución de referencia denotada como F , la cual describe una distribución k ($k \geq 1$) dimensional, en la práctica esta suposición se traduce en contar con una muestra Y_1, \dots, Y_m , llamada muestra de referencia de vectores k dimensionales (Parelius, 1999).

La noción de profundidad de datos se basa en el hecho de que cualquier densidad de probabilidades distingue entre puntos “centrales” y “periféricos”. Una función de profundidad asigna a cada $y \in R^k$, un rango no negativo, el cual puede interpretarse como su localización en la nube de datos. Las profundidades más grandes corresponden al centro de la distribución, mientras que las más pequeñas corresponden a regiones externas. Este sistema de clasificación por rangos es usado para determinar si una observación es diferente de aquellas que conforman el conjunto original. Las funciones de profundidad deben satisfacer un número de propiedades (Zuo, 2000): invariante afín, maximalidad al centro y desvanecimiento al infinito. Existen funciones que cumplen con estas propiedades, entre otras la profundidad de Mahalanobis, la profundidad simplicial y la profundidad de Tukey. De estas la más usada, quizá debido a su facilidad de cómputo e interpretación, es la profundidad de Mahalanobis denotada por $MD_F(y)$, y definida como (Liu, 1993):

$$MD_F(y) = \frac{1}{[1 + (y - \mu_F)' \Sigma_F^{-1} (y - \mu_F)]} \quad (1)$$

Donde, μ_F es el vector de medias y Σ_F^{-1} es la matriz de covarianzas de la distribución de referencia F . Si los parámetros de la distribución de referencia no son conocidos, la versión muestral de la profundidad de Mahalanobis dada por la ecuación (2) es:

$$MD_{F_m}(y) = \frac{1}{[1 + (y - \bar{Y})' S^{-1} (y - \bar{Y})]} \quad (2)$$

Donde \bar{Y} es el vector de medias muestrales de los datos Y_1, \dots, Y_m y S es la matriz de covarianzas de la muestra de referencia F_m . Para construir una gráfica de control basada en la profundidad de Mahalanobis, llamada comúnmente gráfica de clasificación por rangos (o r-chart en el idioma inglés), primero se debe determinar que F_m será la distribución de referencia y G la distribución

de monitoreo las cuales a su vez serán distribuciones continuas. Una vez determinadas las distribuciones F_m y G , se calculan la matriz de covarianzas y el vector de medias de los datos que contiene F_m para de esta manera, obtener la profundidad de todos los datos de las variables contempladas en F_m por medio de la ecuación (2) denotando las profundidades obtenidas como $MD_{F_m}(y_i)$ $i=1, 2, \dots, m$. Estas

profundidades son colocadas en orden ascendente y se utiliza $y_{[j]}$ para denotar el valor de la muestra asociada con el j -ésimo valor de profundidad, siendo de esta manera que $y_{[1]}, y_{[2]}, \dots, y_{[m]}$ son los ordenes estadísticos de F_m .

Después de obtener los ordenes estadísticos se obtienen nuevas observaciones X_1, X_2, \dots de las variables que forman la distribución G , para cada X_i

se obtiene su profundidad aplicando la ecuación (2), utilizando para ello el vector de medias y la matriz de covarianzas de F_m y denotando las profundidades obtenidas mediante $MD_{F_m}(x_i) \quad i=1,2,\dots$. Para cada nuevo vector de observaciones de la distribución G se calcula el estadístico de clasificación por rangos, $r(\cdot)$, mediante:

$$r(x_i) = \frac{\#\{y_j \mid MD_{F_m}(y_j) \leq MD_{F_m}(x_i), j = 1, \dots, m\}}{m+1} \quad (3)$$

Donde $r(x_i)$ es el i -ésimo estadístico de clasificación estimado correspondiente al i -ésimo vector de observaciones de la distribución G , donde y_j es el j -ésimo valor del orden estadístico que se cumple cuando la $MD_{F_m}(y_j)$ (profundidad de datos de la distribución de referencia) es menor o igual a $MD_{F_m}(x_i)$ (profundidad de datos del i -ésimo vector de la distribución de monitoreo). El estadístico se obtiene al dividir el valor resultante del y_j entre el tamaño de muestra de la distribución de referencia mas uno ($m+1$). Al obtener los estadísticos de clasificación por rangos de cada x_i , estos se grafican contra el tiempo, con un límite central de control $CL=0.5$ y un límite inferior de control $LIC=\alpha$, donde α es llamada la proporción de alarma con un valor común de aplicación de 0.05, siendo la gráfica resultante llamada de clasificación de rangos, donde se visualizará que si $r(x_i)$ se grafica por debajo del LIC, entonces la observación X_i se declara fuera de control, siendo esta anomalía una señal de alerta para los ingenieros del proceso. La gráfica de clasificación por rango, contrasta a su vez sucesivamente las hipótesis:

Ho: La nueva observación tienen la misma distribución que la distribución de referencia.

Ha: Existe un cambio en la ubicación o dispersión de la distribución de la nueva observación con respecto a la distribución de referencia.

Donde si $X_i \leq \alpha (0.05)$ *Ho* es rechazada y se llega a la conclusión de que se detecta un deterioro en la calidad del producto y se está presentando un cambio en la ubicación o dispersión de la nueva observación con respecto a F_m .

Paso 4. Monitoree el proceso.-Para cada producto terminado, obtenga las mediciones de las características de calidad determinadas en el paso 1 y

alimente al sistema con ellas. El sistema mostrará una gráfica, llamada gráfica de control de comparación de rangos, como la mostrada en la figura 4 de la sección 3.

Paso 5. Si el proceso está bajo control, continúe monitoreando, de otra forma tome acciones correctivas (Paso 6).-Cuando un producto terminado cumple con las especificaciones de calidad determinadas por las observaciones en la base histórica de datos el proceso se declara bajo control, de otra forma, el sistema dará una señal de alarma, la cual se manifiesta como una observación bajo la línea punteada en la carta de control (observaciones 6, 10 y 17 de la figura 4.).

Cuando exista una señal de alarma en la gráfica de comparación de rangos, se deberá de realizar, una gráfica de contribuciones (la cual es la principal contribución del sistema propuesto). Para la realización de la gráfica de contribuciones uno de los pasos es realizar un análisis de componentes principales de la distribución de referencia la principal ventaja de este análisis es la reducción de la dimensionalidad. Utilizando el método de extracción promedio (Jackson, 1991) se pueden escoger los componentes principales más importantes, dado que los primeros dos o tres componentes principales usualmente explican la mayoría de la variabilidad en un proceso, pueden ser usados para propósitos de interpretación en lugar de todo el conjunto de variables. Es obvio que el uso de las primeras d componentes principales excluye información, pero debido a la correlación entre las variables, las primeras d componentes contienen gran información de la variabilidad del proceso (Jackson, 1991).

La típica forma de un modelo de componentes principales (PCA) es la siguiente: $Z_k = u_{1k}X_1 + u_{2k}X_2 + u_{3k}X_3 + \dots + u_{pk}X_p$. Donde Z_k es

la k componente principal, $(u_{1k}, u_{2k}, u_{3k}, \dots, u_{pk})$ es el correspondiente eigenvector del componente k y X_1, \dots, X_p son las variables del proceso. Uno de los resultados más importantes para la realización de las

gráficas de contribución utilizando componentes principales es la obtención de la contribución del vector x_i en el componente principal k por medio de la ecuación dada en (4).

$$Y_{ki} = u_{1k}x_{i1} + u_{2k}x_{i2} + u_{3k}x_{i3} + \dots + u_{pk}x_{ip} \quad (4)$$

Donde Y_{ki} representa la contribución del i -ésimo vector de observaciones en el k -ésimo componente principal. Hay que hacer notar que el cálculo de las contribuciones de las variables elimina la crítica que existe hacia componentes principales, por lo difícil de su interpretación, ya que en forma gráfica mostrará la influencia de las variables en una señal fuera de control, lo cual será muy práctico en su aplicación dentro del ramo industrial.

Después de realizar un análisis de componentes principales, se realiza el cálculo de las proporciones para todas las variables de la distribución de referencia, utilizando la ecuación (5). Debido a que se seleccionaron $d \leq p$ componentes principales basado en el método de extracción promedio (Jackson, 1991), el método de control propuesto en este caso está basado en proporciones de la forma.

$$r_{ki} = \frac{(u_{k1} + u_{k2} + \dots + u_{kd})x_{ki}}{Y_{1i} + Y_{2i} + \dots + Y_{di}} \quad (5)$$

Donde x_{ki} es el i -ésimo valor de la variable X_k bajo observación, $Y_{ji}, j=1, \dots, d$, es el score del i -ésimo vector de observaciones de las p variables bajo observación en el j -ésimo componente principal. En esta proporción el numerador corresponde a la suma de contribuciones de la i -ésimo valor de variable X_k en las primeras d componentes, mientras que en el denominador es la suma de contribuciones de i -ésimo vector de observaciones en las primeras d componentes. Como se puede observar la importancia de este paso es la de realizar el cálculo de cada una de las p variables en la señal fuera de control por medio de su contribución.

obtenidos en una muestra, con remplazamiento, para generar una estimación empírica de la distribución muestral completa de un estadístico.

La idea básica del bootstrap es tratar la muestra como si fuera la población, y aplicar el muestreo Montecarlo para generar una estimación empírica de la distribución muestral del estadístico. La verdadera estimación Montecarlo requiere un conocimiento total de la población, pero por supuesto este no está generalmente disponible en la investigación aplicada.

Una parte fundamental en el sistema propuesto es la aplicación del método bootstrap, este método descansa en la analogía entre la muestra y la población de la cual la muestra es extraída. De acuerdo a Davison (1997) dada una muestra con n observaciones el estimador no paramétrico de máxima verosimilitud de la distribución poblacional es la función de densidad de probabilidad que asigna una masa de probabilidad de $1/n$ a cada una de las observaciones. La idea central es que muchas veces puede ser mejor extraer conclusiones sobre las características de la población estrictamente a partir de la muestra que se maneja, que haciendo asunciones quizás poco realistas sobre la población. El bootstrap implica remuestreo de los datos

El bootstrap supone que se tiene una situación, en la cual se quiere realizar un análisis de los datos de la siguiente forma: una muestra aleatoria $x=(x_1, x_2, \dots, x_n)$ de una distribución de probabilidad desconocida F ha sido observada, y se requiere estimar un parámetro de interés $\theta=t(F)$, basándose en x . Los métodos bootstrap dependen de la noción de remuestreo. Dado que \hat{F} es una distribución empírica, donde la probabilidad de cada uno de los valores observados $x_i, i=1,2,\dots,n$ es $1/n$. Una muestra bootstrap es definida como una muestra aleatoria de tamaño n delineando a \hat{F} , de tal manera que:

$$\hat{F} \rightarrow (x_1^*, x_2^*, \dots, x_n^*) \quad (6)$$

La notación inicial indica que x^* no es el actual conjunto de datos x , más bien sería una versión de remuestreo de x . Es decir que los puntos de datos bootstrap $x_1^*, x_2^*, \dots, x_n^*$ son una muestra aleatoria de tamaño n obtenida con remplazamiento de la población de n objetos (x_1, x_2, \dots, x_n) . En consecuencia se puede tener $x_1^* = x_7, x_2^* = x_3, x_3^* = x_3, x_4^* = x_2, \dots, x_n^* = x_7$. El conjunto de datos bootstrap consiste de miembros del conjunto original de datos (x_1, x_2, \dots, x_n) , donde algunos no aparecerán ninguna vez, otros una vez, otros dos veces, etc. La aplicación del método de remuestreo bootstrap (Davison, 1997) se realiza a la distribución de referencia para obtener un remuestreo de datos de las variables de F_m y aplicando la ecuación (5) se pueden calcular las proporciones de la muestra de tamaño n con reemplazo, las cuales simularían el comportamiento poblacional de las proporciones de contribución de cada una de las variables bajo monitoreo, al obtener un remuestreo de proporciones se puede determinar el cuantil para $\alpha_1=0.025$ ($\alpha/2$) (Límite inferior de la gráfica de contribución) y el cuantil para $\alpha_2=0.975$ ($1-\alpha/2$) (Límite superior de la gráfica de contribución), estos cuantiles servirán de límite inferior y superior de la gráfica de contribución para cada una de las variables bajo estudio.

Por último se calculan las proporciones para todas las variables de la distribución de monitoreo, considerando el estudio de componentes

principales de F_m y utilizando la ecuación (5), dichas proporciones se representan en la gráfica de contribuciones de cada una de las variables, donde anteriormente ya estaban señalados los límites de control y de esta manera se podrá observar si la proporción obtenida, cae dentro de los límites. Por ejemplo para la observación 6, se puede notar en la figura 5 de la sección 3 que las variables causantes de la señal de alarma es la variable 1.

Paso 6. Tomar acciones correctivas.- Si algún producto es declarado fuera de control, una vez ejecutado el paso 5, (se han identificado las variables causantes de la anomalía), el responsable del proceso de calidad deberá ajustar el proceso conforme a su manual de operación. Una vez ajustado el proceso, siga con el paso 4.

3.- APLICACIÓN DEL SISTEMA DE CONTROL PROPUESTO

En esta sección se presentan los resultados de la aplicación del sistema propuesto en la sección 2, a un proceso de manufactura en la industria automotriz. El proceso está diseñado para fabricar convertidores de torque; un dispositivo que en los vehículos de transmisión automática, transfiere la potencia que genera el motor a la caja de transmisión. En la figura 3 se muestra una sección transversal del convertidor de torque.

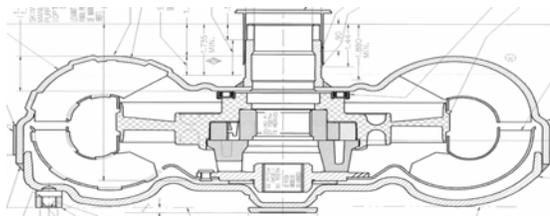


FIGURA 3. Sección transversal del convertidor de torque.

Como un primer paso en la implementación del sistema, los ingenieros de planta determinaron controlar el proceso de ensamble del convertidor; las características de calidad consideradas fueron: holgura interna (variable 1), altura de piloto a orejas (variable 2) y paralelismo (variable 3). En seguida,

se determinó una distribución de referencia de 100 mediciones provenientes de producción controlada bajo especificaciones de calidad. A partir de esta, se calcularon los parámetros necesarios para la construcción de la carta de control. Con esto da cumplimiento al paso 2.

En el paso 3, usando los datos provenientes del paso anterior, se determina que si X es el vector observado cuyos elementos son las variables de calidad de un artículo, tiene un rango menor que 0.05, el proceso se declarará fuera de control y bajo control de otra forma. Por facilidad de exposición, los pasos 4, 5, y 6 se describirán simultáneamente.

El proceso se monitorea al alimentar el sistema con mediciones obtenidas de los productos terminados y al aplicar los pasos 4 y 5 se obtienen a través de MATLAB 7.0, la

gráfica de comparación de rangos (figura 4) y la gráfica de proporciones (figura 5), mostrando la evolución del proceso con referencia a la base.

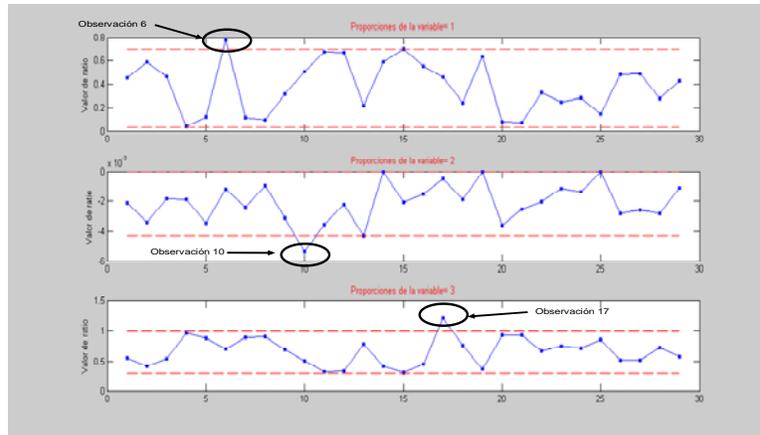


FIGURA 4. Gráfica r, para los datos del convertidor de torque.

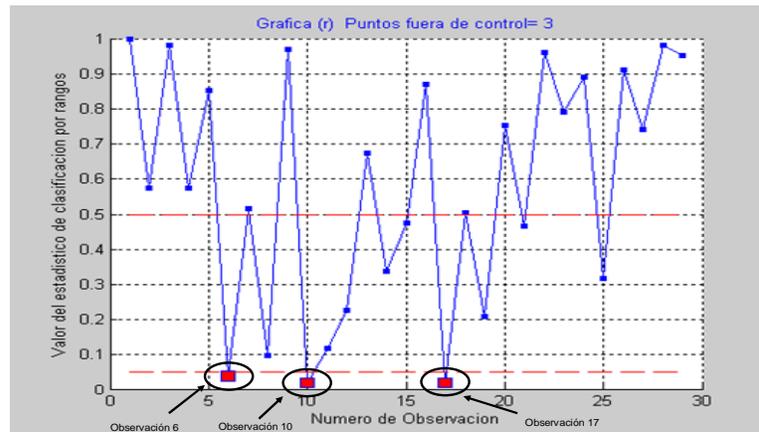


FIGURA 5. Gráfica de contribuciones de las variables del proceso.

Se puede observar en la figura 4 que, conforme a las mediciones de los primeros cinco convertidores, el proceso estaba dentro de control. Hubo cambios en el proceso tales que la observación seis provocó una señal de alarma; mientras que en la gráfica de contribuciones (figura 5) se muestra que la variable 1 fue la causante. Los ingenieros de control

ajustaron el proceso y continuaron monitoreando; la observación 10 nuevamente dispara una señal de alarma, causada ahora por la variable 2; se reajusta el proceso y después de seis observaciones se dispara nuevamente la señal de alarma; ahora debida a la variable 3. Después de un nuevo ajuste, se puede notar que el proceso está controlado. Se encontró

que las señales de alerta efectivamente localizaron productos fuera de control debido a las variables determinadas por el sistema.

4. DISCUSIÓN Y CONCLUSIONES

Actualmente en la industria existen normas concretas para la aplicación de métodos univariados en el control de procesos. En años recientes en la industria automotriz se ha considerado la aplicación de cartas de control multivariadas basadas en el estadístico T^2 de Hotelling, a pesar de que éste se basa en la suposición de normalidad, la cual como lo enfatiza Liu, (1995), difícilmente puede sostenerse en la mayoría de los casos. Así, la principal contribución del artículo, como se mostró en la sección 2, el sistema de control multivariado no paramétrico propuesto, a diferencia de la gráfica T^2 de Hotelling, es capaz de detectar la señal fuera de control y determinar la(s) variable(s) causante(s) de la señal, ya sea que el conjunto de variables siga o no un comportamiento normal multivariado, lo cual lo hace más eficiente que el control de procesos paramétrico tradicional cuando su comportamiento es no normal multivariado, y de eficiencia similar si el comportamiento es normal multivariado.

REFERENCIAS

AIAG Statistical process control (SPC) (2005). Manual reference. Second edition Daimler Chrysler corporation, Ford Motor company and General Motors corporation.

Chakraborty S., Van der Laan P. and Bakir S. T. (2001). Non parametric control charts: An overview and some results. Journal of Quality Technology . Vol. 33 pp 304-315.

Cook D. F., Chiu C. (1998). Using radial basis function neural network to recognize shift in manufacturing process. IIE Transactions. Vol. 30, pp. 227-234.

Davison A.C., Hinkley D.V. (1997). Bootstrap Methods and Their Application, Cambridge Series in Statistical and Probabilistic Mathematics Press, No 1. ISBN-10: 0521574714.

Jackson, J.E. (1991). A User Guide to Principal Components; John Wiley & Sons: New York,.

Liu R. (1992). Data Depth and multivariate rank test. Statistics and Related Methods. pp. 279-294.

Liu R. (1993). A quality Index based on data depth and multivariate rank test. Journal American Statistics Association. Vol. 91, pp. 266-277.

Liu R. (1995). Control chart for Multivariate Processes. Journal American Statistics Association. Vol. 90, pp. 209-125.

Liu R., Parelius M. J. and Singh Kesar. (1999). Multivariate analysis by data depth: Descriptive Statistics and Inference. The annals of statistics Vol. 27, pp. 753-858.

Liu Regina, Cheng Y.A. and Luxhoj T. J. (2000). Monitoring multivariate aviation safety by data depth: Control chart and threshold system. IIE transactions pp. 861-872.

MacGregor, J.F. and Kourti, T. (1995). Statistical Process Control of Multivariate Processes, Control Engineering Practice, Vol. 3, pp. 403-414.

Messaoud A., Theis W. and Hering F. (2006). Application and use of multivariate control charts in BTA deep hole drilling process. Unpublished paper, available from the first author.

Parelius Jesse M. (1999). Multivariate Analysis By Data Depth: Descriptive Statistics, Graphics and Inference. The Annals of Statistics, Vol. 27, No. 3, pp. 783-858.

Stoumbous Zachary G. (2000). The state of statistical process control as we proceed into the 21st century. Journal of the American Statistical Association, Vol. 95, pp. 992- 998.

Zertuche L.F., Cantu S. M. and Galvan M. D. (2004). A Comparison of Two Multivariate Control Processes: Data Depth and Hotelling T^2 , Proceedings of the 34th Conference on Computers and Industrial Engineering, pp. 339-344.

Zuo Yijun & Serfling Robert. (2000). Structural properties and converge results for contours of sample statistical depth functions. The annals of statistics, Vol. 28, No. 2, pp. 483-499.