

Función y desarrollo de los gráficos de control multivariantes

Function and development of multivariate control graphs

Edgar-Augusto Ruelas-Santoyo, Bertha-Laura Vargas-Rodríguez, José-Jovani Cardiel-Ortega y Geraldo Llamas-Perez
Instituto Tecnológico Superior de Irapuato (Méjico)

DOI: <http://dx.doi.org/10.6036/8674>

1. INTRODUCCIÓN

Los gráficos del control son una herramienta del control estadístico de procesos. Resultan útiles para lograr la estabilidad y mejorar la capacidad del proceso mediante la reducción de la variabilidad. El objetivo del gráfico de control es lograr una alta calidad del producto, reduciendo costos de producción y minimizando los defectos del producto. En cualquier proceso de fabricación no siempre se producen los mismos efectos, ya que los elementos que intervienen no siempre funcionan de forma exacta, dando lugar a cierta variabilidad, cuyas causas es preciso investigar. La variabilidad de un proceso puede deberse a causas no asignables y a causas asignables. Las primeras, también llamadas aleatorias son de naturaleza probabilística y forman parte de la variación propia del proceso. Por otro lado, las causas asignables dan lugar a variaciones irregulares no predecibles, que hay que eliminar o corregir. En esta etapa es donde los gráficos de control constituyen una excelente herramienta, pues permiten decidir cuándo hay que intervenir en el proceso para modificar una situación no deseada.

En los procesos de producción actuales, suelen existir características que influyen de forma conjunta e interrelacionada en la calidad final de los productos, cuando dicha condición ocurre las características de calidad deberían ser estudiadas conjuntamente ya que una anomalía en el proceso puede estar anunciada no sólo por la detección de una señal fuera de control en los gráficos individuales correspondientes a algunas de las variables, sino también en la relación entre las variables que se estudian. Una solución consiste en controlar todas de forma si-

multánea mediante técnicas de control estadístico multivariantes. De esta forma no sólo será analizado el efecto de cada una de las características sobre la calidad del producto, sino que también se toma en cuenta el efecto de las interacciones entre ellas.

Dentro de los procedimientos de control multivariante destacan: el gráfico T^2 de Hotelling, el gráfico de control multivariante de sumas acumuladas (MCUSUM) y el gráfico de control multivariante exponencial con peso promedio móvil (MEWMA). Actualmente la gráfica T^2 de Hotelling es el procedimiento de control con más investigación y aplicación en la industria para monitorear p características de calidad correlacionadas en un proceso de producción, en donde dichas características de calidad pueden ser modeladas por una distribución de probabilidad normal p -variante [1]. El estadístico T^2 es un escalar que combina información de medidas de dispersión (Σ) y posición (μ) de las variables. Cuando, como generalmente ocurre en los sistemas de producción, μ y Σ son parámetros desconocidos de una población normal p -variada con media μ y matriz de covarianzas Σ , estos son estimados por \bar{X} y S respectivamente, de acuerdo a las Ec. (1) y (2), [2].

Vector de medias (\bar{X}). Se agrupa en un vector el promedio de las mediciones realizadas para cada una de las p características de calidad correlacionadas.

$$\bar{X} = [\bar{X}_1, \bar{X}_2 \dots \bar{X}_p] \quad (1)$$

Matriz de covarianzas (S) La estructura de esta matriz, es de naturaleza simétrica y se agrupan las medidas de variación bivariada generadas en el proceso. En la diagonal principal se encuentran las varianzas y en el resto las covarianzas de cada una de las variables inspeccionadas en el sistema de producción.

$$S = \begin{bmatrix} S_{11}^2 & S_{12}^2 & \dots & S_{1p}^2 \\ S_{21}^2 & S_{22}^2 & \dots & S_{2p}^2 \\ \dots & \dots & \dots & \dots \\ S_{p1}^2 & S_{p2}^2 & \dots & S_{pp}^2 \end{bmatrix} \quad (2)$$

La variable aleatoria T^2 de Hotelling muestral, se expresa de acuerdo a la Ec. (3).

$$T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) \quad (3)$$

En este caso se calcula el límite control a partir de la Ec. (4):

$$\text{Límite de control} = \left[\frac{p(n+1)(n-1)}{n(n-p)} \right] F_{(\alpha, p, n-p)} \quad (4)$$

Dónde:

n = número de datos

$F_{(\alpha, p, n-p)}$: distribución F

α : probabilidad de falsa alarma del gráfico de control

De aquí en adelante, cuando se mencione a la T^2 se refiere al estadístico generado en la Ec. (3), donde los estadísticos \bar{X} y S son el vector de medias y la matriz de covarianzas respectivamente, de la distribución de referencia, y donde n es el número de datos. La gráfica controla cambios en el vector de medias suponiendo que la matriz de covarianzas del proceso es igual a la matriz de covarianzas de la distribución de referencia.

2. MATERIAL Y MÉTODOS

Para mostrar el funcionamiento de la carta de control T^2 de Hotelling como mecanismo de control en procesos multivariantes, se implementó en una línea de producción de transformadores eléctricos. El caso de estudio se limitó al análisis de transformadores del tipo monofásico con una potencia de 25 KVA (kilo-voltampere), debido a que es el que mayor volumen presenta dentro de la empresa donde se llevó a cabo la ejecución, pero no es una limitante para extender los resultados propuestos a otros tipos de transformadores que se fabriquen bajo el mismo esquema, u otros procesos de manufactura que requieran un control multivariante.

Una parte importante en la manufactura de un transformador eléctrico, es el ensamble de la bobina y el núcleo denominada parte activa. Para un ensamble adecuado es vital controlar las dimensiones del largo en la bobina (A), el ancho del núcleo (B) y largo del núcleo (C), por lo tanto $p=3$. Obsérvese la Fig. (1), se muestra un diseño del ensamble entre la bobina y los núcleos.

Para establecer un control estadístico del proceso adecuado, se deben de vigilar

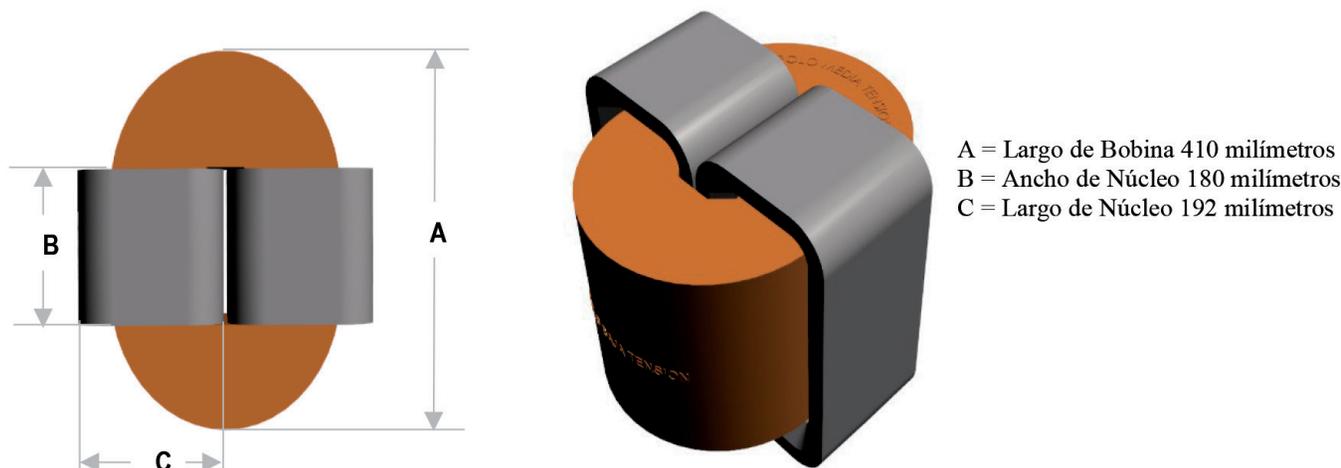


Fig. 1: Diseño de ensamble bobina-núcleo

las tres características de calidad del ensamble y no obviar la posible relación que pudiera existir entre las variables, como generalmente se lleva a cabo en el control estadístico tradicional. En términos estadísticos, se estaría obviando la información contenida en la estructura de covarianza de las características de calidad vigiladas en el producto.

2.1. GENERACIÓN DE DATOS

Para la construcción del gráfico de control multivariante T² de Hotelling como primer paso se procede a estimar el vector de medias y la matriz de covarianzas por medio de las Ec. (1) Ec. (2) respectivamente. Tanto el vector de medias y la matriz de covarianzas se calcularon a partir de 45 mediciones realizadas a transformadores para cada una de sus tres características de calidad críticas en la parte activa del transformador. Ver Tabla 1.

Vector de Medias calculado. Ver Ec. (5)

$$\bar{X} = \begin{bmatrix} 409.9279 \\ 179.6246 \\ 192.1582 \end{bmatrix} \quad (5)$$

Matriz de covarianzas calculada. Ver Ec. (6)

$$S = \begin{bmatrix} 1.7826 & -0.3891 & -0.4085 \\ -0.3891 & 1.1798 & -0.2816 \\ -0.4085 & -0.2816 & 1.6438 \end{bmatrix} \quad (6)$$

Como ejemplo, considere el cálculo del estadístico T² para la medición uno, utilizando notación matricial, se tiene que:

Se calcula otro valor del estadístico T², ahora con la medición cuarenta.

Para determinar si una observación está fuera de control, el valor del estadístico T² debe compararse con el límite de control superior de acuerdo a la Ec. (4):

Medición	A	B	C	Estadístico T ²
1	409.464	178.337	194.062	3.1188
2	409.213	179.711	192.496	0.3070
3	409.688	178.003	192.001	3.0175
4	409.539	178.988	192.887	0.6700
5	406.919	179.296	191.967	7.0009
6	410.421	180.335	192.345	0.9812
7	410.108	181.547	190.953	3.7084
8	410.082	180.619	190.961	1.4435
9	409.429	179.093	190.727	2.6945
10	411.141	180.372	192.334	2.2068
11	408.877	181.084	191.830	2.0177
12	411.213	180.276	191.499	1.7215
13	410.335	178.706	192.876	0.9190
14	410.124	180.089	192.653	0.6119
15	411.120	180.961	191.371	3.1262
16	407.821	180.359	193.699	3.4888
17	411.923	178.805	191.788	2.3691
18	409.508	179.708	191.300	0.7421
19	408.630	180.573	190.806	2.7731
20	409.400	179.802	192.048	0.1908
21	411.241	177.813	192.179	3.1178
22	412.658	179.420	192.687	5.4839
23	409.422	181.099	192.991	2.8181
24	410.737	177.121	193.595	5.8012
25	408.658	180.295	190.562	3.2054
26	410.359	178.170	191.954	2.0197
27	411.582	179.652	191.950	1.7370
28	410.180	179.965	192.679	0.5348
29	411.020	180.870	192.083	3.0023
30	410.717	179.376	191.918	0.3598
31	410.236	180.711	190.638	2.0560
32	409.942	179.919	190.809	1.1693
33	411.250	179.599	192.595	1.5260
34	411.578	178.858	191.521	1.7888
35	409.994	177.437	193.089	4.3467
36	408.895	178.270	192.849	2.8833
37	408.831	180.235	190.654	2.6654
38	410.575	178.714	192.026	0.8174
39	411.499	179.725	192.250	1.8187
40	406.261	180.592	197.889	24.7465
41	411.035	179.059	190.856	1.7250
42	409.597	178.972	191.280	1.5511
43	409.900	178.372	193.620	2.1908
44	406.935	181.398	190.496	8.8914
45	408.697	180.800	193.347	2.6349
Promedio	409.9279	179.6246	192.1582	

Tabla 1: Mediciones y cálculo del estadístico T²

$$T_1^2 = \left(\begin{bmatrix} 409.464 \\ 178.337 \\ 194.062 \end{bmatrix} - \begin{bmatrix} 409.9279 \\ 179.6246 \\ 192.1582 \end{bmatrix} \right)' \begin{bmatrix} 1.7826 & -0.3891 & -0.4085 \\ -0.3891 & 1.1798 & -0.2816 \\ -0.4085 & -0.2816 & 1.6438 \end{bmatrix}^{-1} \left(\begin{bmatrix} 409.464 \\ 178.337 \\ 194.062 \end{bmatrix} - \begin{bmatrix} 409.9279 \\ 179.6242 \\ 192.1582 \end{bmatrix} \right)$$

$$T_1^2 = [-0.4639 - 1.2876 \ 1.9038] \begin{bmatrix} 0.6690 & 0.2714 & 0.2127 \\ 0.2714 & 0.9938 & 0.2377 \\ 0.2127 & 0.2377 & 0.7019 \end{bmatrix} \begin{bmatrix} -0.4639 \\ -1.2876 \\ 1.9038 \end{bmatrix} = 3.1188$$

Se calcula otro valor del estadístico T^2 , ahora con la medición cuarenta T_{40}^2 .

$$T_{40}^2 = \left(\begin{bmatrix} 406.261 \\ 180.592 \\ 197.889 \end{bmatrix} - \begin{bmatrix} 409.9279 \\ 179.6242 \\ 192.1582 \end{bmatrix} \right)' \begin{bmatrix} 1.7826 & -0.3891 & -0.4085 \\ -0.3891 & 1.1798 & -0.2816 \\ -0.4085 & -0.2816 & 1.6438 \end{bmatrix}^{-1} \left(\begin{bmatrix} 409.464 \\ 178.337 \\ 194.062 \end{bmatrix} - \begin{bmatrix} 409.9279 \\ 179.6242 \\ 192.1582 \end{bmatrix} \right)$$

$$T_{40}^2 = [-3.6669 \ 0.9674 \ 5.7308] \begin{bmatrix} 0.6690 & 0.2714 & 0.2127 \\ 0.2714 & 0.9938 & 0.2377 \\ 0.2127 & 0.2377 & 0.7019 \end{bmatrix} \begin{bmatrix} -3.6669 \\ 0.9674 \\ 5.7308 \end{bmatrix} = 24.7465$$

$$\text{Límite de control} = \left[\frac{3(46)(44)}{45(42)} \right] F_{(0.0027, 3, 42)}$$

$$\text{Límite de control} = \left[\frac{3(46)(44)}{45(42)} \right] 5.5367 = 17.787$$

Los valores del estadístico T^2 calculado para cada una de las mediciones se grafican de forma conjunta y ordenada al igual que el límite de control para completar el gráfico de control T^2 de Hotelling. De acuerdo a la Fig. (2), donde se ejemplifica la carta de control multivariante aplicada al proceso de manufactura de transformadores, el estadístico T^2 calculado a partir de las mediciones realizadas al transformador número cuarenta T_{40}^2 se encuentra fuera de control estadístico debido a que supera el límite de control calculado y se deberá investigar la fuente que causa la señal fuera de control.

En el caso de los gráficos univariantes cuando existe una señal de fuera de control puede concluirse fácilmente sobre el problema y dar una solución. En el caso multivariante la solución no es tan directa. Cuando el valor del estadístico T^2 supera el límite de control se debe detectar a la(s) variable(s) responsable(s) lo cual es un tema de investigación que se encuentra en desarrollo y que se aborda en la siguiente sección. Para llevar a cabo un control estadístico de procesos bajo el enfoque multivariante se recomienda seguir el esquema presentado en la Fig. (3).

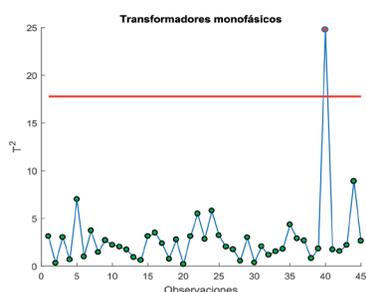


Fig. 2: Gráficos de control multivariantes T^2 de Hotelling

3. LIMITACIÓN Y AVANCES EN LOS GRÁFICOS DE CONTROL MULTIVARIANTES

Los gráficos de control estadístico multivariantes mencionados en la sección anterior son considerados paramétricos, ya que parten del supuesto de normalidad.

Además, las señales de fuera de control en las cartas de control multivariantes no revelan cuál variable (o combinación de variables) causó la señal, lo cual complica la interpretación y generación de estrategias para eliminar la causa asignable de la variación.

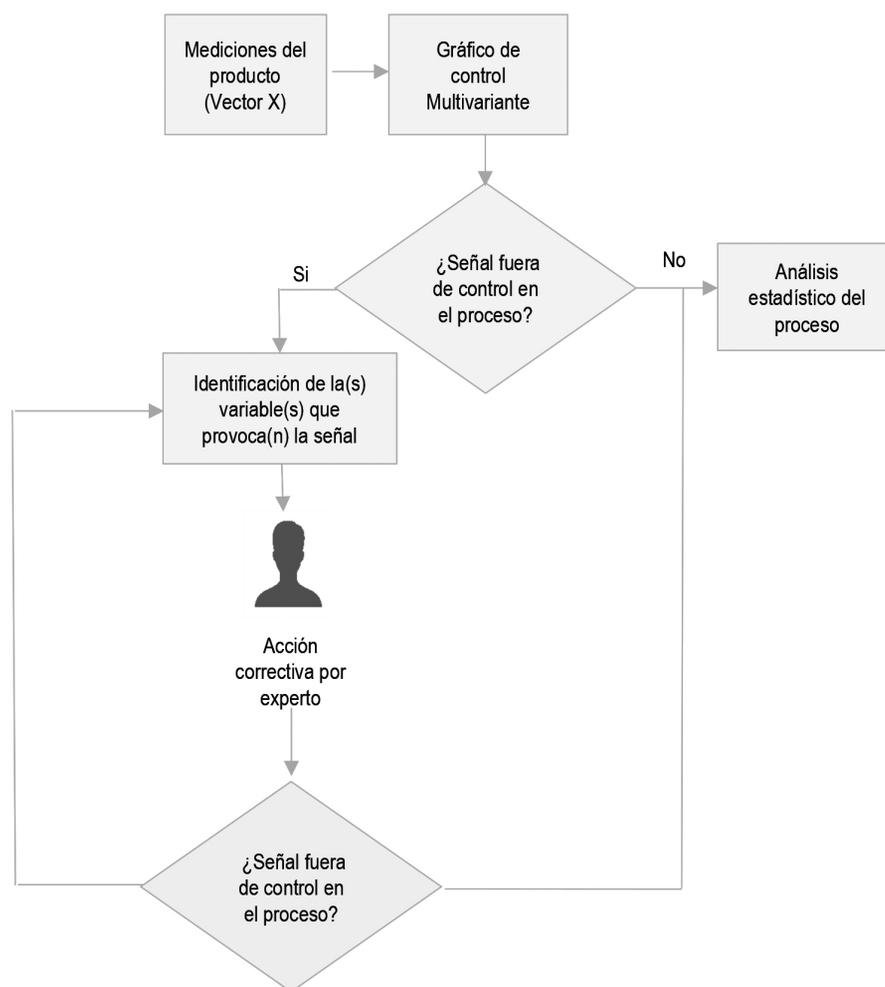


Fig. 3: Metodología para el control de procesos multivariantes

El problema de la interpretación de señales fuera de control en los gráficos multivariantes ha frenado el desarrollo de estas técnicas en la industria, ya que requiere un trabajo posterior nada simple para poder encontrar las variables que han cambiado, dado que una situación fuera de control puede deberse a una o varias variables fuera de control o a un cambio en la relación existente entre las variables.

Se han desarrollado técnicas para ayudar en la interpretación de las señales fuera de control, si bien la más utilizada consiste en analizar gráficos de control univariantes para cada una de las características de calidad. Este camino presenta ciertos inconvenientes: el primero es que cuando hay muchas variables, esta técnica puede resultar tediosa por el gran número de gráficos univariantes para analizar; y el segundo es que normalmente una señal fuera de control no es causada sólo por una variable, sino más bien por la relación existente entre variables, por lo que si son utilizadas gráficas de control univariadas separadas en una situación multivariada, el error de tipo I y la probabilidad de que un punto caiga correctamente dentro de los límites de control no es igual a sus valores esperados. La distorsión de estos valores aumenta con el número de variables de medición.

Para detectar la variable que ha causado la señal fuera de control también se puede recurrir a la descomposición del estadístico T^2 de forma que mida la influencia de cada una de las variables [3]. Si T^2 es el valor del estadístico y $T_{(i)}^2$ es el valor de ese estadístico para todas las variables del proceso excepto la i -ésima, es posible calcular un indicador de la contribución de la variable i -ésima sobre el conjunto mediante la Ec. (7).

$$d_{(i)} = T^2 - T_{(i)}^2 \quad (7)$$

Cuando aparece una situación fuera de control en un gráfico de control multivariante es conveniente calcular esta contribución para cada una de las variables y centrar la atención en aquellas variables cuya contribución sea superior. En el trabajo presentado por [4], se propone una descomposición del estadístico T^2 en dos componentes: T^2 no condicionada, que equivale al valor del estadístico T^2 para cada una de las variables y T^2 condicionada, que equivale al comportamiento de dicho estadístico basado en los residuos de la regresión de cada variable sobre las demás.

Siguiendo otro enfoque como el de componentes principales, actualmente dicha herramienta se ha utilizado con éxito

para detectar fallas en procesos multivariantes que cuentan con variables altamente correlacionadas. En [5], se presenta un método de monitoreo de procesos estadísticos para mejorar el rendimiento de la supervisión. El método se denomina análisis de componentes principales en movimiento, el cual plantea monitorear los cambios en la estructura de correlación de las variables de proceso, en lugar de los cambios en los puntajes de los componentes principales predefinidos. En el trabajo desarrollado por [6], se plantea un método de supervisión de procesos multivariantes a partir de un método basado en análisis de componentes independientes, el método demuestra tener una mejor eficiencia en comparación con el análisis de componentes principales para cumplir tareas de monitoreo y control de procesos.

Otra alternativa, más reciente, en la interpretación de gráficos de control multivariantes, consiste en el uso de redes neuronales artificiales que permiten automatizar la selección de las variables que han provocado el cambio en el proceso. En [7], se lleva a cabo una integración entre la carta de control multivariante de promedio móvil ponderado exponencialmente (MEWMA) y una arquitectura de red neuronal artificial con el fin de desarrollar un procedimiento que permita identificar la(s) causa(s) del descontrol estadístico ocurrido en un sistema de producción. Dentro de [8], se desarrolló un modelo basado en redes neuronales de cuantificación de vectores con el fin de monitorear y diagnosticar los cambios ocurridos en las medias de las características que se vigilan dentro de los procesos de fabricación. El trabajo presentado por [9], propone la aplicación de un sistema de clasificación para determinar la variable o variables que han causado la situación fuera de control detectada por el gráfico de control multivariante T^2 , en concreto, se estudió la aplicación de la red neuronal artificial Fuzzy ARTMAP como alternativa para detectar la fuente de variación asignable presente en procesos de producción multivariantes. En el estudio realizado por [10], se plantea un método para detectar fallas potenciales en turbinas de forma temprana e identificar las variables que probablemente contribuyen a la generación de dichas fallas durante un cierto período, así como en un momento específico. El método incorpora el uso de minería de datos para seleccionar las variables más importantes del sistema bajo supervisión además del uso de cartas de control multivariantes MEWMA como mecanismo para detección de señales fuera de control.

4. CONCLUSIONES

Es importante que los métodos estadísticos usados en el control estadístico de procesos se adapten a las nuevas necesidades de los procesos de producción, tales como las mediciones de múltiples variables de interés que muy posiblemente se relacionan y definen la calidad del producto manufacturado. Los gráficos de control multivariantes, como el T^2 de Hotelling satisfacen estos requerimientos. De acuerdo a lo estudiado en secciones anteriores, parece conveniente utilizar los gráficos de control multivariantes no de forma aislada, sino complementado con alguna técnica de análisis de señales fuera de control, obteniendo así una interpretación más clara de los resultados obtenidos al monitorear la calidad de los sistemas de producción bajo un enfoque multivariante. Otro punto importante, es que actualmente el uso de herramientas estadísticas pero sobre todo informáticas han venido a complementar las cartas multivariantes permitiendo realizar el análisis de grandes cantidades de información logrando para el caso de los valores fuera de control multivariantes identificar e interpretar los cambios ocurridos en las correlaciones entre las variables involucradas en los procesos de fabricación actuales, de esta forma los sistemas de producción son corregidos de forma más rápida y consiguen una mayor productividad y eficiencia.

PARA SABER MÁS

- [1] AIAG Statistical process control (SPC). Manual reference. Second edition Daimler Chrysler Corporation, Ford Motor Company and General Motors Corporation. 2005.
- [2] Montgomery, D. (2011). Control estadístico de la calidad. Edición 3. Limusa Noriega. ISBN: 9789681862343
- [3] Runger, G.C., Alt, F.B. Montgomery, D.C. (2007). Contributors to Multivariate Statistical Process control Signal. Communications in Statistics-Theory and Methods, Vol. 25, No.10. DOI: <https://doi.org/10.1080/03610929608831832>
- [4] Mason, R.L.; Young, J.C. (2002). Multivariate statistical process control with industrial applications. American Statistical Association and the Society for Industrial and Applied Mathematics (ASA-SIAM): Philadelphia.
- [5] Kano, M. Hasebe, S. Hashimoto, I. Ohno, H. (2001). A New Multivariate Statistical Process Monitoring Method Using Principal Component Analysis. Computers & Chemical Engineering. DOI: [https://doi.org/10.1016/S0098-1354\(01\)00683-4](https://doi.org/10.1016/S0098-1354(01)00683-4)
- [6] Kano, M. Hasebe, S. Hashimoto, I. Ohno, H. (2004). Evolution of multivariate statistical process control: application of independent component analysis and external analysis. Computers and Chemical Engineering. DOI: <https://doi.org/10.1016/j.compchemeng.2003.09.011>
- [7] Ahmadzadeh, F., Lundberg, J. & Strömberg, T. (2013). Multivariate process parameter change identification by neural network. The International Journal of Advanced Manufacturing Technology. 69: 2261. DOI: <https://doi.org/10.1007/s00170-013-5200-x>
- [8] Yang, W.A. (2015). Monitoring and diagnosing of mean shifts in multivariate manufacturing processes using two-level selective ensemble of learning vector quantization neural networks. Journal of Intelligent Manufacturing 26: 769. DOI: <https://doi.org/10.1007/s10845-013-0833-z>
- [9] Ruelas-Santoyo, E., Vazquez-Lopez, J., Jimenez-Garcia, J., Cruz-Dominguez, O., Vargas-Rodriguez, B. (2017). System for the monitoring and control of multivariate processes through the statistical T^2 hotelling and fuzzy artmap artificial neural network. DYNA Management, 5(1). [12 p.]. DOI: <http://dx.doi.org/10.6036/MN8206>
- [10] Yang, H. Huang, M. Lai, C. Jin, J. (2018). An approach combining data mining and control charts-based model for fault detection in wind turbines. Renewable Energy. DOI: <https://doi.org/10.1016/j.renene.2017.09.003>