

# Forecasting of the Transformer Core Destruction Factor by means of Multivariate Methods for Data Analysis

G. A. Díaz, A. A. Romero, *Member, IEEE*, E. Mombello, *Senior Member, IEEE* and N. Furlan

**Abstract**— In designing and building of transformers, the core destruction factor is the superposition of all effects that cause a difference between the value of the core losses calculated during the design stage of the unit and the measured value after built. In this paper, two methods for forecasting the core destruction factor are proposed. One based on Mahalanobis distance and the other one based on cluster analysis. A comparison of the results obtained by conventional calculation procedure with respect to those obtained through the proposed methodologies is developed. Finally, the expected cost savings by applying the methods proposed in this article are estimated.

**Keywords**— transformer losses, forecasting, Mahalanobis distance, clustering.

## I. INTRODUCCIÓN

EL TRANSFORMADOR es un componente fundamental en los sistemas eléctricos de potencia y además es una de las máquinas que conceptualmente ha cambiado en menor medida desde su invención, y no se prevén grandes cambios en su concepto en el futuro [1]. Por lo tanto, cualquier aporte realizado para mejorar y optimizar los procesos actuales de construcción y operación constituye una contribución para los sistemas eléctricos actuales y del futuro.

A pesar de que el transformador es una máquina eficiente, el desarrollo de nuevos materiales y de métodos de cálculo sofisticados ha permitido incrementar más aún los niveles de eficiencia. Conceptualmente, la eficiencia es la capacidad que tiene el transformador de transferir energía de un circuito magnético a otro con la menor cantidad de pérdidas posible. Las pérdidas de energía se presentan principalmente en el cobre de las bobinas, en partes metálicas como: tanque, prensas, etc., y en el núcleo [2].

En este artículo, los esfuerzos han sido concentrados en la estimación de las pérdidas asociadas con el núcleo, también llamadas pérdidas de vacío. Su importancia radica en que a pesar de que son menores en orden de magnitud comparadas con las pérdidas en el cobre, las pérdidas en el núcleo no

dependen del nivel de carga del transformador, y por lo tanto estarán presentes como una constante a lo largo de la vida útil de la máquina. De hecho, estas pérdidas hacen parte de los costos fijos de operación de un sistema eléctrico.

En este contexto, se presentan dos metodologías, una basada en distancias de Mahalanobis y otra fundamentada en análisis de conglomerados, con el objetivo de pronosticar un *probable factor de destrucción*, a partir de núcleos similares ya construidos, el cual es útil para realizar una corrección del valor calculado de las pérdidas de vacío para un “*núcleo objetivo*” por construir o en etapa de diseño.

## II. PÉRDIDAS DE VACÍO EN TRANSFORMADORES

### A. Método Convencional para el Cálculo de las Pérdidas de Vacío Durante la Etapa de Diseño.

El valor de pérdidas de vacío es uno de los parámetros de diseño más sencillo de calcular en cuanto a la cantidad de variables involucradas y la complejidad asociada con los cálculos matemáticos. Sin embargo, tal valor de diseño, es también uno de los parámetros que más suele diferir con respecto al valor de pérdidas de vacío real que se verifica una vez construida la máquina.

En la práctica, el cálculo convencional de las pérdidas de vacío vienen dadas por

$$P_{0CA} = P_{0E} * N_P \quad (1)$$

donde  $P_{0CA}$  son las pérdidas de vacío calculadas con el método convencional en W,  $P_{0E}$  son las pérdidas específicas de vacío en W/kg, y  $N_P$  es el peso del núcleo en kg.

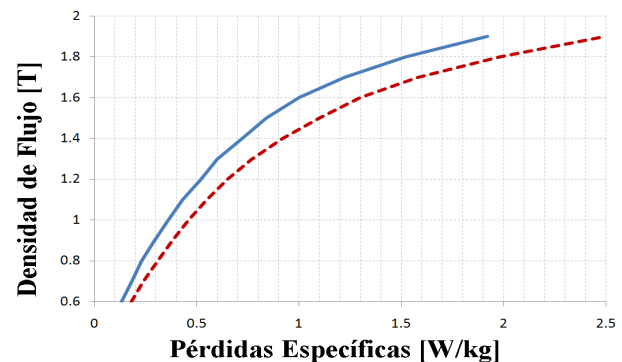


Figura 1. Curvas de pérdidas específicas de la lámina magnética de grado M4, línea a trazos 60Hz, línea continua 50 Hz.

Las pérdidas específicas de vacío,  $P_{0E}$ , son función de la densidad de flujo magnético, y pueden expresarse mediante un

This work was supported by the DAAD (Deutscher Akademischer Austauschdienst) and the CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), República Argentina. The database was kindly provided by Siemens Andina Transformers.

G. A. Díaz, Universidad Nacional de San Juan (UNSJ), San Juan Capital, Argentina, gdiaz@iee.unsj.edu.ar

N. Furlan, Universidad Nacional de San Juan (UNSJ), San Juan Capital, Argentina, nfulan@iee.unsj.edu.ar

A. A. Romero, Universidad Nacional de San Juan (UNSJ), San Juan Capital, Argentina, aromero@iee.unsj.edu.ar

E. E. Mombello, Universidad Nacional de San Juan (UNSJ), San Juan Capital, Argentina, mombello@iee.unsj.edu.ar

polinomio o una aproximación tipo spline, cuyos coeficientes dependen del grado de la lámina magnética y de la frecuencia de operación. Por ejemplo, para una lámina magnética de grado M4, las  $P_{0E}$  pueden ser determinadas a través de una gráfica como la de la Fig. 1.

El peso del núcleo,  $N_p$ , está en función de las dimensiones del mismo, de la sección transversal de la pierna,  $N_A$  en  $\text{cm}^2$ , y de la densidad volumétrica de la lámina magnética,  $\rho$  en  $\text{kg}/\text{dm}^3$ , (ver Fig. 2).

$$N_p = g(A, B, C, N_A, \rho) \quad (2)$$

donde  $A$  y  $B$  son la altura y el ancho de la ventana del núcleo, y  $C$  y  $D$  son el ancho y el diámetro de la pierna del núcleo.

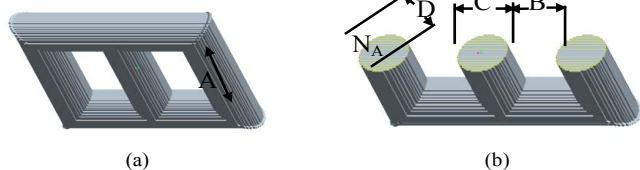


Figura 2. (a) Representación 3D del núcleo tipo apilado. (b) Corte transversal de las piernas del núcleo.

### B. Definición del Factor de Destrucción de la Lámina

El factor de destrucción (también conocido como factor de construcción, [3]) de la lámina representa la superposición de todos los efectos que hacen que el valor de las pérdidas de vacío calculadas ( $P_{0CA}$ ) durante la etapa de diseño, sea diferente del valor medido ( $P_{0Med}$ ) después de la construcción de la unidad. El factor de destrucción se define matemáticamente mediante la siguiente relación:

$$FD_{RA} = P_{0Med} / P_{0CA} \quad (3)$$

donde  $FD_{RA}$  es el factor de destrucción real de la lámina magnética estimado con la metodología actual en por unidad.

El factor de destrucción está principalmente asociado con el deterioro de las propiedades magnéticas de la lámina, relacionadas con la manipulación en el proceso productivo del transformador. Adicionalmente, existen otros factores de influencia, tales como la variabilidad en las características de los materiales y los procesos térmicos que los afectan, no existiendo aún una cuantificación clara para estos efectos.

### C. Hipótesis Asociadas con el Factor de Destrucción

La hipótesis principal sobre la cual se basa este artículo es que *el factor de destrucción puede ser caracterizado mediante variables dimensionales del núcleo*. Ésta hipótesis se soporta en el hecho de que la manipulación de la lámina magnética depende principalmente de las dimensiones de la misma.

## III. ANÁLISIS EXPLORATORIO INICIAL DE LOS DATOS

Para la elaboración del estudio se cuenta con una base de datos inicial de 98 transformadores de diferente diseño. La base de datos incluye los siguientes parámetros:  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $N_A$ ,  $N_p$  y  $FD_{RA}$ . Antes de proceder a aplicar los métodos de análisis, se realiza un análisis de la variabilidad del  $FD_{RA}$ .

### A. Variabilidad del Factor de Destrucción

Para apreciar el comportamiento global del factor de destrucción, se muestra en la Fig. 3 un histograma de la

variable  $FD_{RA}$ , donde se resalta la presencia de un dato atípico. Como se aprecia, la variabilidad es considerable, teniendo un valor mínimo de 0.83 y un máximo de 1.19. Evidentemente esta gran variabilidad está asociada con incertidumbre, ya que los valores mayores que 1 representan transformadores donde el valor medido superó el valor calculado con el riesgo de superar las tolerancias máximas de las pérdidas de vacío establecidas en las normas (con la posibilidad de incurrir en multas), mientras que valores por debajo de 1 indican que el valor medido fue inferior al calculado, indicando que el diseño pudo haberse realizado con menor cantidad de materiales.

De lo anterior se deduce que si fuese posible realizar un pronóstico del  $FD_{RA}$ , se podría por un lado reducir el riesgo, y por otro lado ahorrar gran cantidad de materiales, no sólo el asociado con el núcleo, sino también con el resto del transformador, ya que los demás componentes, como cobre de las bobinas, metalmecánica, cantidad de litros aceite, aislamientos, etc., son función del volumen del núcleo.

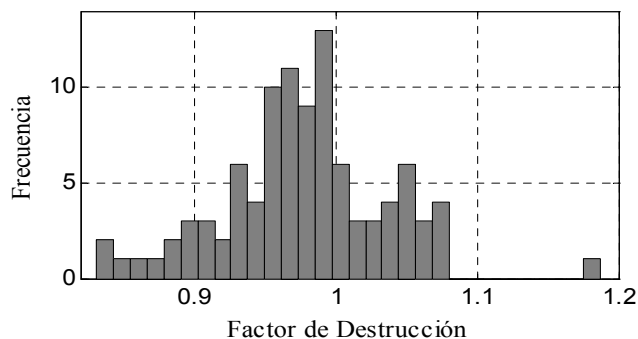


Figura 3. Histograma de la variable  $FD_{RA}$  para los 98 individuos.

En la tabla I se presenta un resumen del análisis descriptivo univariante del  $FD_{RA}$ .

TABLA I.  
ANÁLISIS DESCRIPTIVO DE  $FD_{RA}$  PARA LOS 98 INDIVIDUOS.

| Media  | Mediana | Desv. Típica | MEDA  | Coef. Var. | Coef. Var Robusto |
|--------|---------|--------------|-------|------------|-------------------|
| 0.9781 | 0.9779  | 0.0582       | 0.029 | 0.0595     | 0.0297            |

De la tabla I se puede concluir que la media es un buen indicador del centro de datos debido a que es muy similar a la mediana, lo cual indica que en promedio el valor medido es un 2.2% menor que el valor calculado. Sin embargo, la desviación típica y el coeficiente de variación indican que existe una importante variabilidad, y por lo tanto los cálculos de pérdidas de vacío no son del todo precisos, ya que en una situación “ideal” (de precisión en el cálculo) la media debería ser 1 y la desviación 0. Se debe observar que en la práctica es imposible lograr una desviación igual a 0, sin embargo sí es factible centrar en 1 la distribución y reducir la variabilidad a través de un adecuado pronóstico de las pérdidas, lo que precisamente representa el objetivo del este trabajo.

En este caso, ni la desviación típica ni el coeficiente de variación son indicadores robustos, debido a que difieren en casi el doble de la mediana y del coeficiente robusto (La mediana es la mediana de las desviaciones absolutas respecto de la mediana. El coeficiente de variación robusto se determina como el cociente entre la mediana y la mediana.) de variación

respectivamente; sin embargo ambos dan una idea de que la variabilidad de la variable  $FD_{RA}$  es relativamente grande.

### B. Detección de Variables Clave y Análisis de Dependencias

Como se mostró en la sección II.A, las pérdidas de vacío dependen de las variables geométricas  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $N_A$ ,  $N_P$ , y bajo la hipótesis propuesta de que el  $FD_{RA}$  depende de variables geométricas, se pretende determinar si existen dependencias lineales entre estas variables para así reducir la dimensión del problema [4].

Como se cuenta con variables que tienen diferentes unidades de medida (mm, cm<sup>2</sup>, kg y p.u.), se procede a realizar una transformación no lineal mediante la aplicación del logaritmo natural a la matriz de datos [5].

En la tabla II se muestra la matriz de correlación de la matriz de datos transformada.

TABLA II.  
MATRIZ DE CORRELACIÓN MATRIZ 98x7.

|           | $A$    | $B$     | $C$     | $D$     | $N_A$   | $N_P$   | $FD_{RA}$ |
|-----------|--------|---------|---------|---------|---------|---------|-----------|
| $A$       | 1      | -       | -       | -       | -       | -       | -         |
| $B$       | 0.4994 | 1       | -       | -       | -       | -       | -         |
| $C$       | 0.3699 | 0.7141  | 1       | -       | -       | -       | -         |
| $D$       | 0.3717 | 0.7278  | 0.9954  | 1       | -       | -       | -         |
| $N_A$     | 0.3713 | 0.7306  | 0.9951  | 0.9988  | 1       | -       | -         |
| $N_P$     | 0.5603 | 0.8076  | 0.9695  | 0.9733  | 0.9743  | 1       | -         |
| $FD_{RA}$ | 0.0678 | -0.1220 | -0.0669 | -0.0733 | -0.0797 | -0.0618 | 1         |

De la matriz de correlación se puede observar que las variables  $D$ ,  $N_A$ , y  $N_P$  tienen una alta dependencia lineal con la variable  $C$ , de modo similar  $N_A$  y  $N_P$  con  $D$ , y finalmente  $N_P$  con  $N_A$ . De hecho, el coeficiente de correlación entre estas cuatro variables es cercano a 1. Por lo tanto, se concluye que es posible reducir la dimensionalidad de la matriz de datos, manteniendo sólo a  $C$  como parámetro representativo. Los argumentos para mantener a  $C$  son: *i*) el parámetro  $D$  es una dimensión físicamente ficticia, ya que las secciones de las piernas de los núcleos no son totalmente circulares; *ii*) desde el punto de vista experimental, no se ha establecido una relación clara entre el peso,  $N_P$ , o la sección efectiva de la pierna,  $N_A$ , con el factor de destrucción, por tanto estas dos variables tampoco son apropiadas para continuar en el análisis; *iii*) por estandarización el parámetro  $C$  es discreto, esto es, varía en intervalos de 20 mm, además, físicamente es un indicador adecuado del tamaño de la pierna del núcleo.

### C. Detección de Datos Atípicos

En las secciones anteriores se ha presentado un panorama global del comportamiento del factor de destrucción y la estructura general de dependencia entre las variables. Hasta el momento se han realizado los análisis sobre una base de datos de 98 transformadores fabricados con dos tipos de lámina magnética diferente, M1H y M4, y adicionalmente con dos tipos de laminaciones diferentes las cuales denominaremos como L1 y L2. Por tanto, se tiene una mezcla de 4 grupos diferentes de transformadores los cuales serán denominados como M1H-L1, M1H-L2, M4-L1 y M4-L2.

Para contextualizar, en el tipo de laminación L1 se tienen transformadores con piernas construidas con una cantidad

reducida de paquetes, lo cual es ventajoso desde el punto de vista de la fabricación implicando tiempos más cortos de cortado y ensamblaje. En el tipo de laminación L2 se tienen núcleos con mayor cantidad de paquetes de menor espesor, lo cual tiene como ventaja el uso más eficiente del volumen de la pierna. Por otro lado, las láminas magnéticas M1H y M4 difieren en su comportamiento a nivel magnético debido a que tienen curvas de pérdidas específicas diferentes.

Buscando la mayor homogeneidad posible en la base de datos de trabajo, se han seleccionado de la muestra de 98 transformadores a los transformadores pertenecientes al grupo M4-L2, los cuales conforman el grupo de mayor tamaño (51 individuos) de los cuatro. Evidentemente los métodos propuestos en este artículo son extensibles para los otros grupos (M1H-L1, M1H-L2, M4-L1) en caso de tener disponible muestras de mayor tamaño.

De este modo, se tiene ahora una base de datos de tamaño 51 x 4, es decir, 51 individuos con las variables  $A$ ,  $B$ ,  $C$  y  $FD_{RA}$ . Para la detección de individuos atípicos, los cuales pueden distorsionar la información contenida en las matrices de correlación y de distancias, se han empleado dos de las técnicas propuestas en [5], una de ellas univariante basada en la mediana y en la meda como estimadores robustos, y otra multivariante realizando la proyección de la matriz de datos en las direcciones que maximizan la kurtosis.

Como resultado del análisis de atípicos, se han retirado 7 individuos de la base de 51 individuos, luego, el conjunto de datos final queda reducido a una matriz de 44 x 4.

## IV. ANÁLISIS PRELIMINAR DE LA MUESTRA SELECCIONADA

### A. Análisis Exploratorio

Para analizar cómo las variables  $A$ ,  $B$ , y  $C$  se relacionan con  $FD_{RA}$ , se presentan en la tabla III las matrices de correlación y correlación parcial respectivamente. En estas tablas queda en evidencia un hecho antes oculto o distorsionado en la tabla II, obtenida al estudiar el conjunto de datos completo incluyendo atípicos. Tal hecho es que la correlación más alta se presenta entre la variable  $A$  y  $FD_{RA}$ , lo cual refuerza la hipótesis planteada en la sección II. C.

Adicionalmente al quitar los efectos de las otras variables mediante la matriz de correlación parcial, se aprecia que la influencia de la variable  $B$  y  $C$  sobre  $FD_{RA}$  disminuye claramente, y se fortalece la idea de que la variable  $A$  afecta de manera importante al factor de destrucción.

TABLA III.  
MATRIZ DE CORRELACIÓN Y CORRELACIÓN PARCIAL, MATRIZ 44x4.

|           | Correlación |         |         |           | Correlación Parcial |         |         |           |
|-----------|-------------|---------|---------|-----------|---------------------|---------|---------|-----------|
|           | $A$         | $B$     | $C$     | $FD_{RA}$ | $A$                 | $B$     | $C$     | $FD_{RA}$ |
| $A$       | 1           | -       | -       | -         | 1                   | -       | -       | -         |
| $B$       | 0.3306      | 1       | -       | -         | 0.1078              | 1       | -       | -         |
| $C$       | 0.3100      | 0.7001  | 1       | -         | 0.1132              | 0.6631  | 1       | -         |
| $FD_{RA}$ | -0.4416     | -0.2674 | -0.2088 | 1         | -0.3875             | -0.1173 | -0.0151 | 1         |

### B. Distribución de Referencia

Antes de formular las metodologías de pronóstico del factor de destrucción, se presenta en la tabla IV un resumen

del análisis univariante para la variable  $FD_{RA}$  de la muestra reducida de 44 transformadores. Se hace esto con el objetivo de poder comparar los resultados obtenidos con las metodologías propuestas y poder determinar su eficacia.

TABLA IV.  
ANÁLISIS UNIVARIANTE  $FD_{RA}$  PARA LOS 44 INDIVIDUOS.

| Media  | Mediana | Desv. Típica | Meda   | Coef. Var. | Coef. Var Robusto |
|--------|---------|--------------|--------|------------|-------------------|
| 0.9614 | 0.9654  | 0.0311       | 0.0212 | 0.0323     | 0.0220            |

Como objetivo se espera que las metodologías propuestas logren reducir la variabilidad o como mínimo mantenerla constante, consiguiendo el centrar la distribución en 1, esto es, que la media del factor de destrucción sea lo más cercana a 1. Esto último significa que los valores de las pérdidas de vacío calculadas son más cercanos a los valores medidos y por tanto la precisión es más alta. De manera general se puede comentar que, de acuerdo a la tabla IV, el valor calculado es un 4% menor al valor medido, lo cual indica que se podría optimizar el diseño para lograr así un importante ahorro de materiales.

#### V. PRONÓSTICO DEL FACTOR DE DESTRUCCIÓN MEDIANTE DISTANCIAS DE MAHALANOBIS

##### A. Planteamiento de la Metodología

Partiendo de la hipótesis de que núcleos similares tienen un factor de destrucción similar, la metodología busca precisamente determinar cuáles son los núcleos más parecidos (sobre una base de datos preexistente) al núcleo al cual se le quiere pronosticar el factor de destrucción, el cual denominaremos como *núcleo objetivo*. Una vez identificados los núcleos más parecidos, a partir de éstos se procederá a estimar el factor de destrucción del *núcleo objetivo*.

Con este fin se aprovecha el potencial de la distancia de Mahalanobis, ya que esta distancia tiene en cuenta la estructura del conjunto de datos al estar basada en la matriz de varianzas covarianzas, y por tanto *tiene en cuenta la forma de cada uno de los individuos* a partir de la estructura de correlación, permitiendo así identificar los núcleos más parecidos entre sí. El procedimiento implementado se resume en los siguientes pasos:

1. Determinar la matriz de varianzas covarianzas de la matriz de datos. Es importante resaltar que como sólo es de interés encontrar similitudes en la geometría de los núcleos, la matriz de datos incluye únicamente a las variables  $A$ ,  $B$ , y  $C$ , omitiendo a la variable  $FD_{RA}$ .

2. Se determina la distancia de Mahalanobis entre cada individuo y los restantes mediante (4), siendo  $S$  la matriz de varianzas covarianzas de la matriz de  $44 \times 3$ , donde  $i = 1, \dots, 44$  y  $j = 1, \dots, 44$ . Este cálculo produce una matriz simétrica de  $44 \times 44$  con ceros en su diagonal, donde el elemento  $ij$  es la distancia de Mahalanobis entre el individuo  $i$  y el individuo  $j$ .

$$d_{ij} = \sqrt{(x_i - x_j)^T S^{-1} (x_i - x_j)} \quad (4)$$

3. Para cada individuo  $i$ , se organizan las distancias de Mahalanobis de los  $j-1$  individuos de menor a mayor, y se toman los  $FD_{RA}$  de los tres individuos más cercanos al

individuo  $i$ . Notar que en éste paso se asume que el  $FD_{RA}$  del individuo  $i$  (*núcleo objetivo*) no es conocido y se quiere estimar a partir de los 3 núcleos más parecidos.

4. Finalmente se aplica la media de los  $FD_{RA}$  de los tres núcleos más parecidos para obtener el factor de destrucción pronosticado mediante distancias de Mahalanobis,  $FD_{PM}$ , con el cual se corregirá el valor calculado de las pérdidas de vacío del núcleo objetivo  $i$ .

En la Fig. 4, mediante las caras de Chernoff [6], se ha ilustrado cómo el método busca a los núcleos más parecidos al *núcleo objetivo*. Las caras de la izquierda corresponden a un *núcleo objetivo* mientras que las caras de la derecha representan a los tres individuos más parecidos. A pesar de que los individuos de la misma fila no son exactamente iguales, guardan un cierto parecido físico entre sí. También es evidente la diferencia entre los individuos de las distintas filas.

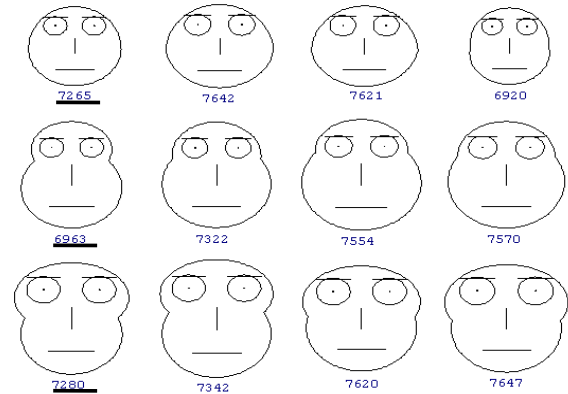


Figura 4. Caras de Chernoff de los 3 núcleos más parecidos al objetivo. Los números bajo las caras indican el número de identificación de cada núcleo.

##### B. Validación de la Metodología Propuesta

Con el factor de destrucción pronosticado mediante distancias de Mahalanobis,  $FD_{PM}$ , se realiza una corrección de las pérdidas de vacío calculadas con la metodología convencional para así obtener las pérdidas de vacío calculadas mediante el modelo de distancias de Mahalanobis,  $P_{OCM}$ , como

$$P_{OCM} = P_{OCA} * FD_{PM} \quad (5)$$

A partir de (5) se determina el factor de destrucción real debido al pronóstico de Mahalanobis,  $FD_{RM}$ :

$$FD_{RM} = P_{0Med} / P_{OCM} = P_{0Med} / (P_{OCA} * FD_{PM}) \quad (6)$$

Luego, sustituyendo (3) en (6) se reescribe  $FD_{RM}$  como:

$$FD_{RM} = FD_{RA} / FD_{PM} \quad (7)$$

En la tabla V se muestra un resumen del análisis univariante de la variable  $FD_{RM}$ . Como se muestra, la media sigue siendo un buen indicador del centro de los datos debido a que prácticamente es igual a la mediana. Se encontró que el factor de destrucción con la metodología de distancias de Mahalanobis es en promedio un 4% más preciso que el factor de destrucción actual.

La desviación estándar se ha reducido ligeramente, lo cual indica que la variabilidad se ha logrado reducir en alguna medida, pero ahora se tiene una precisión más alta.

TABLA V.  
ANÁLISIS UNIVARIANTE  $FD_{RM}$

| Media  | Mediana | Desv. Típica | Meda   | Coef. Var. | Coef. Var Robusto |
|--------|---------|--------------|--------|------------|-------------------|
| 1.0033 | 1.0033  | 0.0306       | 0.0202 | 0.0305     | 0.0201            |

A pesar del buen resultado desde el punto de vista de distribución, es necesario también analizar el comportamiento desde el punto de vista de cada individuo. Al aplicar la metodología se encontró que 31 de los individuos reportaron un valor más preciso de pérdidas de vacío comparado con el procedimiento convencional, lo cual corresponde a un 70% de la muestra. En el caso de los individuos que empeoraron la precisión con respecto a la metodología convencional, se encontró que la mayoría de ellos ya contaba con un nivel de precisión bastante bueno ( $FD_{RA}$  muy cercano a 1), y al aplicar la metodología se alejaron ligeramente del valor esperado.

Adicionalmente con esta metodología se logra reducir considerablemente el riesgo, ya que si se consideran como riesgosos individuos con un error mayor o igual a 7.5% en el factor de destrucción real con respecto al valor ideal (igual a 1), 6 casos se clasifican como riesgosos con la metodología actual (13.6% del total), valor que se reduce a 1 (2.3% del total) con la metodología propuesta.

#### VI. PRONÓSTICO DEL FACTOR DE DESTRUCCIÓN MEDIANTE ANÁLISIS DE CONGLOMERADOS

La metodología de análisis de conglomerados (clustering en inglés) busca determinar grupos de núcleos que tengan características comunes desde el punto de vista geométrico, de manera no supervisada y sistemática, y adicionalmente que el procedimiento sea influenciado lo menos posible por la subjetividad de la persona que realiza el estudio.

##### A. Selección del Algoritmo

Para el caso de estudio se ha seleccionado el algoritmo de k-Medias, ya que este método de k-Medias es computacionalmente eficiente cuando se tiene una cantidad de individuos elevado [7]. En éste caso se cuenta con una muestra de 44 individuos.

##### B. Planteamiento de la Metodología

Un aspecto clave del método de k-Medias es la determinación de la cantidad de grupos. Debido a que para la determinación de éste es necesario tener algún punto de referencia, se empleó el criterio probado satisfactoriamente en el método de distancias de Mahalanobis, el cual indicó que con 3 individuos similares a otro se obtiene un pronóstico apropiado del factor de destrucción. De éste modo, la metodología desarrollada es la siguiente:

1. Se empieza con una cantidad de grupos igual a la cantidad de individuos (44 con las variables  $A$ ,  $B$ , y  $C$ ) y se ejecuta el algoritmo de agrupación. Evidentemente se obtienen como resultado 44 grupos de un individuo cada uno.
2. Se reduce en uno la cantidad de grupos, y se verifica que cada uno de los grupos tenga como mínimo 4 individuos.
3. Si alguno de los grupos tiene menos de 4 individuos, se reduce en uno la cantidad de grupos actual y se ejecuta la

agrupación nuevamente.

4. Cuando cada uno de los grupos tenga 4 o más individuos, se detiene el procedimiento. Como resultado del procedimiento se encontró una cantidad de 7 grupos.

5. Denotando con  $FD_{PC}$  al factor de destrucción pronosticado mediante la metodología de conglomerados, para cada uno de los  $i$  individuos se toma la media del  $FD_{RA}$  de los individuos de su grupo. En ésta media no se incluye el  $FD_{RA}$  del individuo  $i$  (*núcleo objetivo*), ya que con propósitos de validación se hace el supuesto de que éste no se conoce y se busca estimar a partir de los factores de destrucción de los individuos de su grupo.

Para ilustrar como el algoritmo ha conformado los grupos, en la Fig. 5 se ha representado en el espacio cada uno de los 44 individuos y el grupo al que ha sido asignado. Se aprecia que la altura de la ventana  $A$  es un parámetro clave para la clasificación, notándose una distribución por capas de diferentes rangos de  $A$  para cada uno de los grupos.

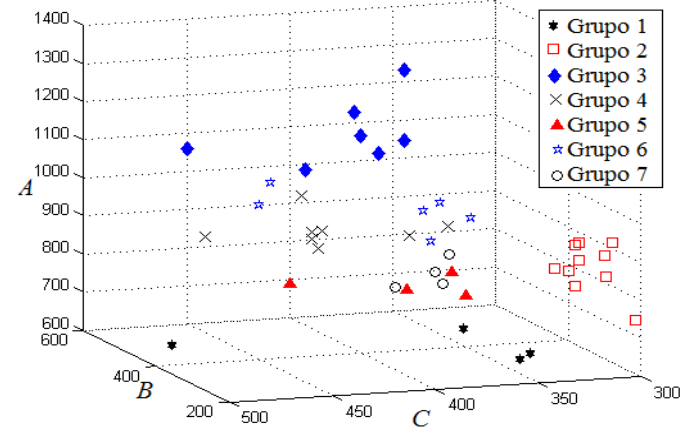


Figura 5. Representación de los 44 núcleos asociados en 7 grupos.

##### C. Validación de la Metodología Propuesta

De manera similar a lo mostrado en la metodología basada en distancias de Mahalanobis, se puede demostrar que el factor de destrucción real debido a aplicar el valor pronosticado de la metodología basada en conglomerados es

$$FD_{RC} = FD_{RA} / FD_{PC} \quad (8)$$

En la tabla VI se muestra el análisis univariante de la variable  $FD_{RC}$ . Se aprecia que en general los resultados son similares a los encontrados con la metodología basada en distancias de Mahalanobis, sin embargo en este caso la media es más cercana a 1 y la desviación estándar también se ha reducido ligeramente.

TABLA VI.  
ANÁLISIS UNIVARIANTE  $FD_{RC}$

| Media  | Mediana | Desv. Típica | Meda   | Coef. Var. | Coef. Var Robusto |
|--------|---------|--------------|--------|------------|-------------------|
| 1.0001 | 1.0015  | 0.0302       | 0.0207 | 0.0302     | 0.0207            |

Analizando desde el punto de vista de los individuos, se encontró que mediante la metodología de conglomerados el 73% de los individuos incrementó su precisión con respecto a la metodología convencional, y sólo 2 de los individuos tuvieron un error en el factor de destrucción real mayor al

7.5% (clasificados como riesgosos).

#### D. Conclusiones Derivadas del Método de Conglomerados

En la tabla VII se ha dispuesto un resumen con los resultados encontrados del análisis de conglomerados, de los cuales se puede concluir lo siguiente:

1. El grupo 2 contiene a los núcleos con las dimensiones de ancho de la pierna más pequeñas de toda la muestra, en éste caso todos los individuos tienen dimensión  $C=300$  mm y 320 mm, y adicionalmente el  $FD_{RA}$  promedio más alto.
2. El grupo 3 contiene a los núcleos altos, esto es, que tienen una altura de ventana elevada con respecto a los demás, y se evidencia que los factores de destrucción para estos transformadores son los más bajos de toda la muestra.

TABLA VII.  
RESUMEN DEL ANÁLISIS DE CLÚSTER.

| Clúste<br>r | Cant. de<br>individuos | A     |         | B     |         | C     |         | $FD_{RA}$ |         |
|-------------|------------------------|-------|---------|-------|---------|-------|---------|-----------|---------|
|             |                        | media | mediana | media | mediana | media | mediana | media     | mediana |
| 1           | 4                      | 650   | 657     | 324   | 328     | 385   | 350     | 0.9585    | 0.9618  |
| 2           | 10                     | 856   | 858     | 276   | 276     | 315   | 320     | 0.9816    | 0.9801  |
| 3           | 8                      | 1185  | 1161    | 393   | 402     | 410   | 400     | 0.9249    | 0.9175  |
| 4           | 8                      | 908   | 911     | 437   | 418     | 410   | 420     | 0.9592    | 0.9574  |
| 5           | 4                      | 772   | 765     | 400   | 386     | 380   | 370     | 0.9706    | 0.9716  |
| 6           | 6                      | 992   | 985     | 364   | 346     | 397   | 380     | 0.9690    | 0.9807  |
| 7           | 4                      | 851   | 838     | 298   | 301     | 385   | 380     | 0.9704    | 0.9667  |

#### VII. COMPARACIÓN DE LAS METODOLOGÍAS

Las distribuciones de probabilidad de las variables  $FD_{RA}$ ,  $FD_{RM}$  y  $FD_{RC}$ , se ajustan adecuadamente a la normal. En la Fig. 6 se ha representado la distribución de estas tres variables. Se observa que a través de los métodos propuestos se ha logrado centrar las distribuciones en 1. Adicionalmente las distribuciones obtenidas con las distancias de Mahalanobis y el método de conglomerados prácticamente se superponen, demostrando que el resultado es similar al aplicar estos métodos de pronóstico.

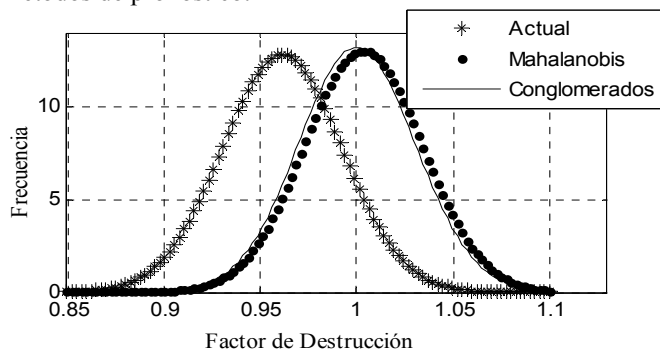


Figura 6. Comparación de las distribuciones de los métodos propuestos y la metodología convencional.

#### VIII. AHORROS ECONÓMICOS ESPERADOS

De acuerdo a lo encontrado al implementar las metodologías propuestas, se logró llevar el factor de destrucción de un valor promedio de 0.96 a 1 y además se redujo ligeramente la variabilidad actual del proceso, en otras palabras se puede decir que el incremento en la precisión del cálculo de las pérdidas de vacío fue de en promedio un 4%

con respecto a la metodología convencional (sin correcciones).

Para evaluar el ahorro obtenido al implementar las metodologías propuestas en algunos transformadores típicos, se encontró que es posible ahorrar un promedio 21.3 kg de lámina magnética por MVA. Aparentemente el ahorro es despreciable comparado con el peso de un núcleo, que puede ser de varias toneladas, sin embargo, proyectando los ahorros de material al total de MVA construidos por la fábrica en el año 2008, se observa un ahorro promedio anual de 23 toneladas de lámina magnética, que a un valor promedio de 4.04 USD/kg, corresponde a un ahorro anual de 92821USD.

Se debe notar que este cálculo de ahorro es conservativo, ya que sólo se evalúa el ahorro en lámina magnética, sin embargo existen ahorros adicionales en cobre, metalmecánica y aceite, que también deben ser considerados en el análisis.

#### IX. CONCLUSIONES

Se ha corroborado la validez de la hipótesis propuesta, que establece que el factor de destrucción puede caracterizarse en función de las dimensiones y la forma del núcleo, sin embargo, siguen existiendo otros efectos que causan variabilidad y que deben estudiarse en el futuro.

Las metodologías de pronóstico propuestas demostraron ser efectivas de acuerdo a las metas esperadas, reduciendo el error promedio de cálculo de las pérdidas de vacío en un 4% con respecto a la metodología convencional.

El análisis de conglomerados permitió mostrar que en núcleos altos se deteriora menos la lámina magnética (bajo factor de destrucción), mientras que en núcleos con ancho de pierna reducido la lámina se deteriora en mayor medida (alto factor de destrucción). Esta conclusión es importante ya que permitirá enfocar los esfuerzos de mejora en los procesos de manufactura para cierto tipo de transformadores donde la lámina es propensa al deterioro.

Una de las principales ventajas de las metodologías propuestas es que tienden a autocorregirse en la medida en que incrementa la cantidad de individuos de la muestra, de hecho con más muestras aumenta la probabilidad de encontrar núcleos similares al cual se desea estimar el factor de destrucción. De aquí surge el término de métodos "robustos".

#### REFERENCIAS

- [1] M. Kang, P. Enjeti, y I. Pitel, "Analysis and Design of Electronic Transformers for Electric Power Distribution System," *IEEE Transactions on Power Elect.*, vol. 14, no. 6, pp. 1133-1141, Nov 1999.
- [2] J. Heathcote, *The J & P Transformer Book*, Thirteenth Ed. Elsevier 2005
- [3] Qader, A.A.; Basak, A.;, "Building factor of a 100 KVA 3 phase distribution transformer core," *Magnetics, IEEE Transactions on*, vol.18, no.6, pp. 1487- 1489, Nov 1982
- [4] P. E. Green and J. D. Carroll, *Mathematical tools for applied multivariate analysis*, Acad. Press New York, New York, USA, 1976.
- [5] D. Peña, *Análisis de Datos Multivariantes*, McGraw – Hill /Interamericana de España 2002.
- [6] W. Härdle and L. Simar, *Applied multivariate statistical analysis*, Springer Verlag, 2011.
- [7] Y. Kay, T. Jayant, *Multivariate Statistical Methods in Quality Management*, McGraw – Hill Professional.



**Guillermo A. Díaz** nació en Bogotá, Colombia, en 1984. Recibió su título en Ingeniería Eléctrica en la Universidad de la Salle, Bogotá - Colombia en 2005. En 2006 inició sus actividades laborales para Siemens en área de Transformadores de Distribución desarrollando labores de investigación y desarrollo de software.

Actualmente tiene una beca de la Agencia Alemana de Intercambio Académico (Deutscher Akademischer Austauschdienst/DAAD) para estudios de doctorado en el Instituto de Energía Eléctrica, Facultad de Ingeniería de la Universidad Nacional de San Juan (IEE-UNSJ), Argentina. Sus intereses de investigación se concentran principalmente en el área de modelación electromagnética y optimización de equipos eléctricos. Email: [gdiatz@ice.unsj.edu.ar](mailto:gdiatz@ice.unsj.edu.ar).



**Andrés A. Romero** (IEEE M'06) nació en Bogotá, Colombia, el 9 de Julio de 1978. En el año 2002 recibió el grado de Ingeniero Electricista de la "Universidad Nacional de Colombia" y en el año 2009 el grado de Doctor en Ingeniería Eléctrica, en el "Instituto de Energía Eléctrica de la Universidad Nacional de San Juan (IEE-UNSJ)", Argentina. En el 2010 fue premiado con una

beca posdoctoral de dos años, por parte del Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, para trabajar en el proyecto de investigación: "Gestión Óptima de Activos en Sistemas Eléctricos de Potencia". Actualmente es Investigador del IEE-UNSJ y ha sido seleccionado de la convocatoria 2011, para ingresar a la Carrera del Investigador Científico y Tecnológico del CONICET. Sus áreas de interés son: gestión de activos, calidad del producto eléctrico y ensayos y técnicas de alta tensión.



**Enrique E. Mombello** (IEEE M'95-SM'00) nació en Buenos Aires, Argentina, en 1957. Se graduó como ingeniero electromecánico en 1982 y como doctor en ingeniería eléctrica en 1998. Cuenta con más de 20 años de experiencia en proyectos de investigación, incluyendo temas como transitorios eléctricos y procesos de resonancia en los transformadores como Investigador del

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, Argentina). Ha trabajado desde 1989 hasta 1991 en el Instituto de Alta Tensión de la RWTH, en Aachen, Alemania. Se desempeña desde 1982 en el Instituto de Energía Eléctrica de la Universidad de San Juan, Argentina, como profesor e investigador. Sus principales campos de interés son el diseño, modelado y diagnóstico de transformadores de energía, gestión de activos, gestión de transformador de la vida, transitorios electromagnéticos en máquinas eléctricas y de redes, modelos de equipos, las pérdidas corona en líneas de transmisión aéreas, campos electromagnéticos de baja frecuencia.



**Norma L. Furlan** nació en 1954 en Río Cuarto, Argentina. Se graduó como Licenciada en Matemáticas en la Universidad Nacional de San Luis, Argentina en 1979, y como Magister en Estadística e Investigación de Operaciones en la Universidad Nacional Autónoma de México en 1986. Se desempeña como Docente e Investigadora en el Instituto de Energía Eléctrica de la

Universidad Nacional de San Juan, Argentina, en el área de Matemáticas Aplicadas al Análisis de Sistemas de Potencia.