

Analysis and Characterization of Power Quality using Data Mining

Análisis y Caracterización de la Calidad de Energía utilizando Minería de Datos

A.D. Mullo¹  0009-0009-6605-7134 J.M. Reinoso¹  0009-0006-0432-4143
M.S. Chamba¹  0000-0001-6843-7151 C.X. Lozada²  0000-0002-6036-3124

¹Pontificia Universidad Católica del Ecuador - Sede Esmeraldas, Esmeraldas, Ecuador.
E-mail: admulloa@pucese.edu.ec, jmreinoso@pucese.edu.ec, mschamba@pucese.edu.ec

²Operador Nacional de Electricidad - CENACE, Mejía, Ecuador
E-mail: clozada@cenace.gob.ec

Abstract

This work addresses the issue of power quality in electrical distribution networks, focusing on the identification and evaluation of harmonic distortions, which can affect equipment performance and regulatory compliance. To achieve this, a methodology was implemented that combines univariate analysis to verify compliance with the IEEE 519-2022 standard and the ARCONEL 009/2024 regulation, along with data mining techniques such as Principal Component Analysis (PCA) and the K-Means clustering algorithm, which classify harmonics based on their behavior within the electrical system. The methodology was validated through the analysis of historical harmonic data from a cement industry whose distribution network operates at 22 kV. The results made it possible to identify critical periods in which harmonic levels exceeded regulatory limits, mainly due to the operation of variable frequency drives, inverters, and rectifiers used in industrial processes such as raw material extraction, grinding, preheating, kiln operation, bagging, and dispatch. The developed model proved effective in processing large volumes of data, identifying the main sources of harmonic distortion, and segmenting behavior by time and day, thus facilitating the implementation of mitigation strategies and its adaptation to various industrial environments.

Resumen

El presente trabajo aborda la problemática de la calidad de energía en redes de distribución eléctrica, enfocándose en la identificación y evaluación de distorsiones armónicas, las cuales pueden afectar el funcionamiento de equipos y el cumplimiento normativo. Para ello, se implementó una metodología que combina el análisis univariante a fin de verificar el cumplimiento con las normas IEEE 519-2022 y ARCONEL 009/2024, junto con técnicas de minería de datos como el Análisis de Componentes Principales (PCA) y el algoritmo de clústeres K-Means, que permiten clasificar los armónicos según su comportamiento en el sistema eléctrico. La metodología fue validada a través del análisis del histórico de armónicos de una industria cementera, cuya red de distribución opera a 22 kV. Los resultados permitieron identificar periodos críticos en los que se superan los límites normativos, principalmente por el funcionamiento de variadores de frecuencia, inversores y rectificadores asociados a procesos industriales como extracción de materia prima, molienda, precalentamiento, calcinación, ensacado y despacho. El modelo desarrollado demostró ser eficaz para procesar grandes volúmenes de datos, detectar las principales fuentes de distorsión armónica y segmentar el comportamiento por horario y día, lo que facilita la implementación de estrategias de mitigación y su adaptación a diferentes entornos industriales.

Index terms— Quality, harmonics, clustering, mining, distribution systems.

Palabras clave— Calidad, armónicos, clúster, minería, sistemas de distribución

Recibido: 01-04-2025, Aprobado tras revisión: 04-07-2025

Forma sugerida de citación: Mullo, A.; Reinoso, J.; Chamba, M.; Lozada, C. (2025). "Análisis y Caracterización de la Calidad de Energía utilizando Minería de Datos". Revista Técnica "energía". No. 22, Issue I, Pp. 33-45.

ISSN On-line: 2602-8492 - ISSN Impreso: 1390-5074

Doi: <https://doi.org/10.37116/revistaenergia.v21.n2.2025.702>

© 2025 Autores



Esta publicación está bajo una licencia internacional Creative Commons Reconocimiento – No Comercial 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)



1. INTRODUCCIÓN

La calidad de la energía eléctrica es fundamental para garantizar que las cargas operen correctamente dentro de los límites establecidos, a fin de evitar la generación de perturbaciones que afecten el funcionamiento de los equipos en procesos industriales, comerciales y residenciales [1]. Entre las principales perturbaciones que afectan la calidad de la energía eléctrica se encuentran las distorsiones armónicas, generadas por variadores de frecuencia, convertidores electrónicos, equipos industriales y otras cargas no lineales [2].

Una de las principales consecuencias de los armónicos en la red eléctrica es el aumento de pérdidas por efecto Joule, provocando el sobrecalentamiento de transformadores, motores y conductores [3]. Por tal motivo, es importante identificar las principales fuentes generadoras, así como los períodos y condiciones en las que la amplitud armónica es más significativa, con el fin de implementar estrategias de mitigación y de esta forma, garantizar el cumplimiento de las normativas vigentes de calidad de energía, como el estándar IEEE 519-2022 [4] y la regulación ARCONEL 009/2024 para la calidad del servicio de distribución y comercialización de energía eléctrica en Ecuador [5].

Tradicionalmente, el análisis de armónicos en sistemas eléctricos se ha basado en técnicas univariantes, como el análisis de Distorsión Armónica Total (THD) y la Distorsión Total de Demanda (TDD), que cuantifican la distorsión armónica en la red eléctrica. Sin embargo, estas mediciones presentan limitaciones al momento de procesar grandes volúmenes de datos, lo que dificulta la caracterización del comportamiento armónico y la predicción de anomalías en el sistema eléctrico [6].

Ante esta problemática, la minería de datos se plantea como una herramienta eficaz para caracterizar y analizar la calidad de la energía. En este contexto, se emplean metodologías de procesamiento masivo de información, junto con algoritmos de aprendizaje automático, mediante los cuales se identifican patrones o comportamientos similares. Esto permite detectar irregularidades en el sistema eléctrico y en consecuencia, optimizar la toma de decisiones [7].

Entre estas técnicas, el Análisis de Componentes Principales PCA (Principal Component Analysis) permite identificar patrones, establecer correlaciones y reducir la dimensionalidad de los datos sin comprometer la variabilidad significativa. En [8], se presenta un estudio aplicado a redes industriales, en el cual se emplea el PCA para concentrar un conjunto elevado de variables en componentes principales que conservan la mayor parte de la información relevante. Esta estrategia facilitó la simplificación del análisis y permitió detectar fallas en la red eléctrica, contribuyendo al diagnóstico de distorsiones armónicas y comportamientos irregulares.

Por su parte, en [9] se combina el PCA con técnicas de Machine Learning para detectar y predecir el robo de

electricidad. La metodología propuesta permite analizar grandes volúmenes de datos e identificar con precisión patrones irregulares asociados a fraudes o pérdidas no técnicas, optimizando así la supervisión y control del sistema eléctrico.

Una vez reducidas las variables mediante PCA, se implementa el algoritmo de agrupamiento no supervisado K-Means, el cual permite clasificar los datos en distintos clústeres. Esta técnica resulta especialmente útil en sistemas eléctricos, ya que facilita la identificación de patrones operativos, la detección de comportamientos anómalos y la segmentación de datos en función de su impacto sobre la calidad de la energía [10].

En [10], se propone una metodología para la partición de redes de distribución eléctrica utilizando los algoritmos de agrupamiento K-Means y DBSCAN, demostrando que los nodos se agrupan eficazmente y logran una partición adecuada de la red de distribución. Por otro lado, en [11] se utiliza K-Means para clasificar clientes según el comportamiento energético, lo que permitió a las empresas distribuidoras optimizar estrategias de gestión de la demanda y diseñar políticas tarifarias más eficientes.

De manera similar, en [12] se utiliza el algoritmo K-Means para agrupar datos históricos de consumo eléctrico y mejorar la predicción de demanda a corto plazo, mediante minería de datos. Esta segmentación de clientes permitió optimizar los algoritmos de aprendizaje automático, reduciendo el margen de error y mejorando la gestión del sistema eléctrico.

A medida que el sector eléctrico avanza hacia una mayor digitalización, la aplicación de minería de datos en el análisis de la calidad de la energía continuará evolucionando. En este sentido, la integración de modelos predictivos permite anticipar fallas, analizar correlaciones y optimizar recursos [13]. En [14], se emplea el software Python para la predicción de precios de la energía, además se menciona que es un lenguaje eficiente en el procesamiento y análisis de grandes volúmenes de datos. Además, Python cuenta con bibliotecas especializadas como Spyder, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Scikit-learn, que permiten la manipulación, segmentación, reducción y visualización de datos.

Bajo esta perspectiva, el presente estudio tiene como objetivo analizar y caracterizar la calidad de la energía de una red de distribución eléctrica cementera, mediante una metodología orientada a la identificación y evaluación de armónicos.

Esta metodología emplea técnicas de minería de datos para el análisis de patrones, comportamiento y agrupación de datos eléctricos, permitiendo determinar las principales fuentes de generación de armónicos, identificar los períodos con mayor nivel de distorsión armónica. Finalmente, permite evaluar el cumplimiento normativo, analizar su impacto en la red de distribución



eléctrica y facilitar el desarrollo de estrategias de mitigación.

A continuación, en la segunda sección, se presenta la metodología, que abarca la evaluación del cumplimiento normativo y el análisis de armónicos usando minería de datos. Posteriormente, en el tercer acápite se discuten los principales resultados, además se analiza el cumplimiento normativo. Finalmente, se presentan las conclusiones y recomendaciones.

2. METODOLOGÍA

La caracterización de armónicos de la red de distribución eléctrica se la realiza en dos etapas:

1. Evaluación de cumplimiento normativo
2. Minería de datos

2.1 Evaluación del Cumplimiento Normativo

La Fig. 1 presenta la metodología diseñada para evaluar el cumplimiento de los límites establecidos en la normativa IEEE 519-2022 y la regulación ARCONEL 009/2024. Este proceso inicia con la adquisición de datos a partir del Power Quality Monitor (PQM), dispositivo que registra en tiempo real parámetros eléctricos como tensión, corriente, potencia activa, valores RMS y niveles de distorsión armónica [15].

Seguidamente, se exporta el histórico de armónicos de voltaje y corriente, el cual puede incluir valores atípicos (outliers) que deben procesarse antes de evaluar el cumplimiento normativo. Si todos los valores se encuentran dentro de los límites establecidos, el proceso finaliza. En caso contrario, se identifican los armónicos que generan mayor distorsión en el sistema de distribución eléctrica.

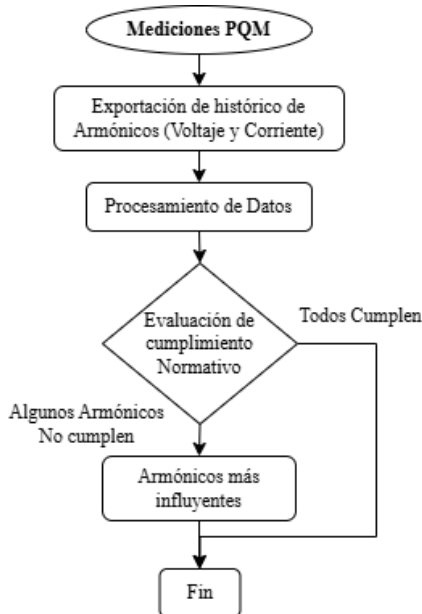


Figura 1: Metodología de evaluación normativa conforme a IEEE 519-2022 y ARCONEL 009/2024

2.1.1 Recopilación y Preprocesamiento de Datos

El análisis de armónicos requiere una adecuada recopilación de los datos obtenidos del dispositivo de medición (PQM), los cuales se exportan en formato CSV o se almacenan en bases de datos para su análisis posterior.

Para garantizar la calidad de los datos y evitar desviaciones en los resultados, es necesario aplicar un preprocesamiento que permita detectar y eliminar valores atípicos que pueden surgir por interferencias electromagnéticas, variaciones de tensión o corriente provocados por arranques de motores, conmutaciones de carga o descargas atmosféricas. Se consideran atípicos los datos que superan ± 3 desviaciones estándar respecto a la media o los límites definidos por normas como la IEEE 519-2022 [6].

2.1.2 Determinación de armónicos

La identificación de armónicos de voltaje y corriente se realiza a partir del historial de mediciones registradas de forma semanal, mensual o anual, el cual incluye los primeros 50 armónicos con un intervalo de muestreo de 10 minutos.

Este proceso se lleva a cabo, comparando la distorsión armónica individual DI (1) con los valores de referencia normativos establecidos en [4]. Este parámetro es la relación entre la corriente armónica de orden h y la corriente máxima de carga (I_L).

$$DI = \frac{I_h}{I_L} * 100 \quad (1)$$

Para armónicos en corriente, se evalúa la relación de cortocircuito SCR (2), que determina el porcentaje armónico admitido y con el cual se realizará la comparación. Esta relación se obtiene a partir de la corriente de cortocircuito máxima y la corriente máxima de carga a frecuencia fundamental [4].

$$SCR = \frac{I_{sc}}{I_L} \quad (2)$$

Los valores permisibles DI para cada categoría de SCR se presentan en la Tabla 1, lo que permite establecer los límites normativos aplicables en función del tipo de sistema y la capacidad de cortocircuito.

Adicional, la evaluación de la distorsión armónica se complementa con el cálculo de la Distorsión Armónica Total $ITHD$ (3) para corrientes y $VTHD$ (4) para voltajes, conforme a lo establecido en [4].

$$ITHD = \frac{\sqrt{\sum_{h=2}^{h=50} I_h^2}}{I_1} * 100 \quad (3)$$

$$VTHD = \frac{\sqrt{\sum_{h=2}^{h=50} V_h^2}}{V_1} * 100 \quad (4)$$



Tabla 1: Límites IEEE 519-2022 de distorsión de corriente para sistemas clasificados de 120 V a 69 Kv

SCR=ISC/IL	Límites armónicos $2 \leq h < 11$	Límites armónicos $11 \leq h < 17$	Límites armónicos $17 \leq h < 23$	Límites armónicos $23 \leq h < 35$	Límites armónicos $35 \leq h \leq 50$	TDD
<20c	4.0	2.0	1.5	0.6	0.3	5.0
20<50	7.0	3.5	2.5	1.0	0.5	8.0
50<100	10.0	4.5	4.0	1.5	0.7	12.0
100<1000	12.0	5.5	5.0	2.0	1.0	15.0
>1000	15.0	7.0	6.0	2.5	1.4	20.0

En entornos con alta variabilidad de carga, como instalaciones industriales, una métrica más adecuada es el TDD (5), que evalúa la distorsión armónica en relación con la corriente máxima de carga (I_L).

$$TDD = \frac{\sqrt{\sum_{h=2}^{h=50} I_h^2}}{I_L} * 100 \quad (5)$$

El análisis permite determinar el nivel de distorsión armónica, verificar el cumplimiento normativo IEEE 519-2022 y la resolución ARCONEL 009/2024. No obstante, este proceso puede complementarse con técnicas de minería de datos que permiten identificar los períodos, días y principales fuentes de generación responsables de afectar la calidad del suministro eléctrico.

2.2 Minería de datos

En la Fig. 2 se presenta el proceso de caracterización de armónicos mediante técnicas de minería de datos. El análisis comienza con el procesamiento de los datos registrados, donde se eliminan los valores atípicos (outliers). Luego, se aplican técnicas de reducción de dimensionalidad con el objetivo de conservar la mayor cantidad de información relevante.

Posteriormente, se ejecuta un análisis de clústeres, en el cual los armónicos se agrupan en función de la similitud de sus niveles de distorsión, permitiendo identificar aquellos más influyentes según el día, intervalo horario y relacionarlos con las principales fuentes de generación armónica.

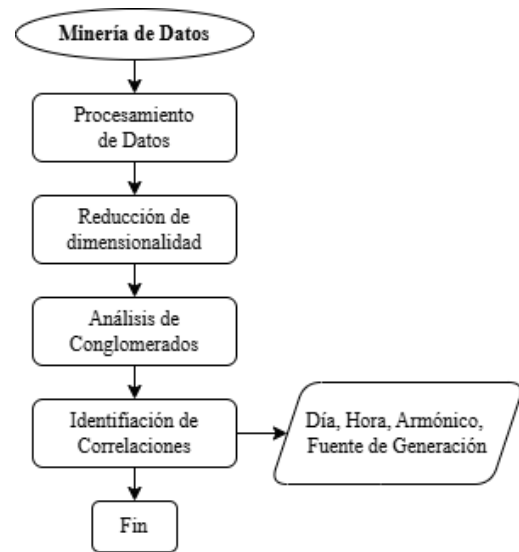


Figura 2: Caracterización de Armónicos usando Minería de Datos

2.2.1 Análisis de Componentes Principales

Existen diversas técnicas de reducción de dimensionalidad, como el Análisis de Componentes Principales (PCA), el Análisis Discriminante Lineal (LDA) y los métodos basados en Descomposición en Valores Singulares (SVD). En el presente estudio, se seleccionó la técnica de análisis PCA, debido a su capacidad para reducir y optimizar la dimensionalidad de datos sin perder información relevante [16].

La aplicación del PCA se desarrolló cumpliendo con la condición $n > p$, donde n corresponde al número de observaciones (intervalo de tiempo 10 minutos) y es superior al número de variables p . Los datos se organizan en la matriz X (6) de dimensión $n \times p$, donde cada fila n tendrá una observación temporal y cada columna las magnitudes de los 50 armónicos, junto con variables temporales correspondientes al día, la hora y la fecha de registro durante el periodo analizado [17].

$$X = \begin{Bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ x_{31} & x_{32} & \dots & x_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{Bmatrix} \quad (6)$$

Antes de aplicar el PCA, se debe garantizar que los datos sean coherentes y representativos, por lo que se realiza un procesamiento detallado de mediciones erróneas o valores atípicos. Posteriormente, se realiza una normalización de los datos basada en la media y la desviación estándar $X_{escaled}$ (7), debido a que las mediciones de los armónicos pueden presentar escalas muy distintas [17].

$$X_{escaled} = \frac{X - \mu}{\sigma} * 100 \quad (7)$$

Una vez estandarizados los datos, es decir, centrados respecto a la media y escalados por su desviación estándar, se calcula la matriz de covarianza C (8) [17].

$$C = \frac{1}{m-1} X^T X \quad (8)$$

Posteriormente, se realiza la descomposición en valores propios y vectores propios, lo que permite determinar la dirección de máxima variabilidad, representada por C_v (9) [17].

$$C_v = \lambda v \quad (9)$$

Finalmente, se seleccionan las componentes principales Z (10), las cuales corresponden a combinaciones lineales de las variables originales que capturan la mayor cantidad posible de varianza en los datos [17].

$$Z = XW \quad (10)$$

Las columnas de la variable W están ordenadas según los autovalores en orden descendente, lo que garantiza que las primeras componentes principales contengan la mayor varianza del sistema analizado. Al multiplicar la matriz de datos estandarizados X por W , se obtiene la matriz Z (10), cuyos componentes representan combinaciones lineales de las variables originales.

Para determinar cuántas componentes conservar, se analiza la varianza explicada acumulada, utilizando el método del codo (*scree plot*), el cual permite identificar el número óptimo de componentes principales al detectar el punto en el que la varianza explicada deja de aumentar significativamente. En la mayoría de los casos, las dos primeras (PC1 y PC2) capturan la mayor parte de la variabilidad de los datos analizados [18].

Esta transformación no solo optimiza la selección de componentes, sino que también reduce la complejidad del análisis, permitiendo una visualización más clara de

la distribución de los armónicos en el nuevo espacio definido por los componentes principales. De este modo, la interpretación del comportamiento armónico en la red eléctrica se vuelve más eficiente, facilitando la identificación de correlaciones entre armónicos, los períodos con mayor incidencia y las posibles fuentes generadoras [19].

2.2.2 Análisis de Clústeres

Existen diferentes métodos de análisis de clústeres, como K-Means, C-Means, DBSCAN o Gaussian Mixture Models. En este estudio, se optó por K-Means, una técnica de aprendizaje no supervisado de fácil implementación, que permite identificar patrones recurrentes en la distorsión armónica al agrupar observaciones con características temporales similares (como la hora y el día de medición), facilitando así la interpretación del comportamiento del sistema [20].

El número óptimo de clústeres se determina mediante dendrogramas generados con el método de Ward el cual minimiza la varianza interna de los grupos. El punto de corte se identificó en el tramo con mayor incremento de distancia entre fusiones consecutivas, lo que permite segmentar los datos de manera eficiente y sin pérdida significativa de información. [21].

Una vez determinado el número óptimo de clústeres, se emplean representaciones en el espacio tridimensional de las componentes principales (PC1, PC2, PC3), lo que permite visualizar concentraciones de datos asociadas a periodos donde los niveles de distorsión armónica superan los límites establecidos por la norma IEEE 519-2022 y regulación ARCONEL 009/2024 [22].

2.2.3 Fuentes generadoras de armónicos

Con el propósito de establecer la relación entre la presencia de armónicos y sus respectivas fuentes de generación, en la Tabla 2 se detallan los dispositivos identificados en estudios previos como fuentes de generación de perturbaciones. Esta correlación permite comprender la distribución de la distorsión armónica en la red eléctrica y orientar adecuadamente la implementación de estrategias de mitigación [23].

Tabla 2: Principales fuentes generadoras de armónicos

Rango de Armónicos	Fuentes Típicas de Generación
2° - 4°	Desequilibrios de carga, transformadores con carga asimétrica, cargas monofásicas no lineales, como luminarias LED, electrodomésticos, asimetrías en rectificadores.
5° - 7°	Cargas trifásicas no lineales como hornos de arco, variadores de velocidad, rectificadores trifásicos no controlados y de seis pulsos.
8° - 10°	Cargas electrónicas con conmutación rápida o ciclos irregulares, convertidores de potencia.
11° - 15°	Rectificadores de doce pulsos, sistemas de tracción eléctrica, fuentes conmutadas de telecomunicaciones, convertidores electrónicos.

Rango de Armónicos	Fuentes Típicas de Generación
17° - 19°	Inversores solares fotovoltaicos, equipos de soldadura industrial.
21° - 29°	Cargadores de vehículos eléctricos, sistemas de almacenamiento de energía.
31° - 50°	Electrónica avanzada de potencia, bancos de condensadores, filtros activos, conmutación de alta frecuencia en redes de distribución.

3. RESULTADOS Y DISCUSIÓN

En esta sección se presentan los resultados obtenidos al aplicar de la metodología propuesta. Para este análisis, se utilizó el histórico de armónicos de una red de distribución correspondiente a una industria cementera, cuyas especificaciones técnicas se detallan en la Tabla 3.

Tabla 3: Especificaciones técnicas del caso de estudio (Industria cementera)

Características Técnicas	Especificaciones
Industria	Cementera
Potencia	500 kVA
Tipo de transformador	Trifásico
Tensión línea a línea $L-L$	22 kV
Tensión línea a neutro $L-N$	12.7 kV
Corriente máxima de carga I_L	1.86 A
Relación de corto circuito (SCR)	>50

El modelo de caracterización de armónicos fue diseñado para ser adaptable a diferentes bases de datos, independientemente del intervalo temporal analizado. Por esta razón, se seleccionó el mes de octubre como periodo de referencia para demostrar su aplicabilidad.

A partir de la información recopilada, los datos fueron

procesados utilizando Python como entorno de programación. En este proceso se emplearon bibliotecas especializadas como *Pandas* para la manipulación de datos, *NumPy* para operaciones numéricas, *Matplotlib*, *Seaborn* y *Plotly* para la visualización gráfica, *Scikit-learn* para la reducción de dimensionalidad mediante PCA y la agrupación mediante K-Means, así como *SciPy* para el análisis jerárquico, todo ello implementado en el entorno de desarrollo Spyder. Este conjunto de herramientas permitió identificar los armónicos predominantes, evaluar el cumplimiento normativo IEEE 519-2022 y la regulación ARCONEL 009/2024, además de clasificar el comportamiento armónico en la red de distribución.

3.1 Minería de datos

Se analizaron los armónicos de corriente y voltaje registrados durante el mes de octubre de 2021, abarcando los órdenes desde H2 hasta H50, con un intervalo de muestreo de 10 minutos. Para el procesamiento de los datos se emplearon las bibliotecas *Pandas* para la gestión de series temporales, *NumPy* para cálculos numéricos, *Matplotlib* y *Seaborn* para la visualización gráfica, así como *Datetime* para el tratamiento de fechas y horas.

La Fig. 3 muestra el comportamiento dinámico de la amplitud de los armónicos durante el mes de octubre. Se observaron incrementos pronunciados en los niveles de distorsión armónica en franjas horarias comprendidas entre las 05:00 y 08:00, así como entre las 17:00 y 20:00, coincidiendo con los cambios de turno o picos de demanda. Además, los días 5, 12, 18 y 25 de octubre registraron valores elevados, lo que sugiere eventos anómalos o una mayor presencia de cargas no lineales. En contraste, los fines de semana presentaron mayor estabilidad, lo que reflejó una menor variabilidad en la carga de la red de distribución eléctrica.

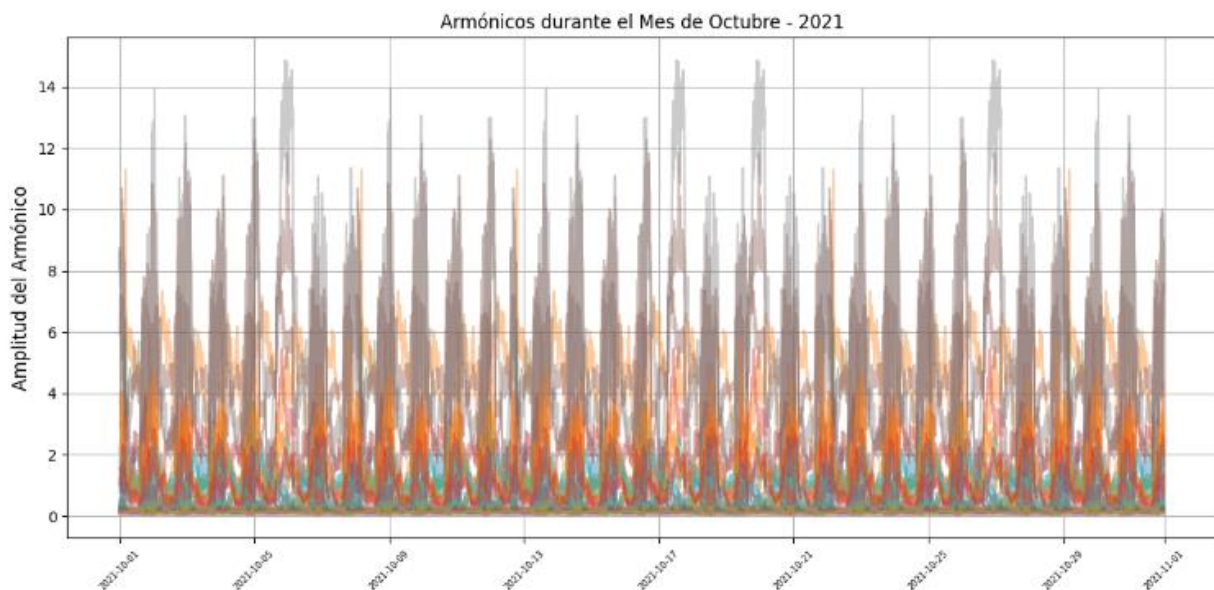


Figura 3: Comportamiento de los armónicos en corriente

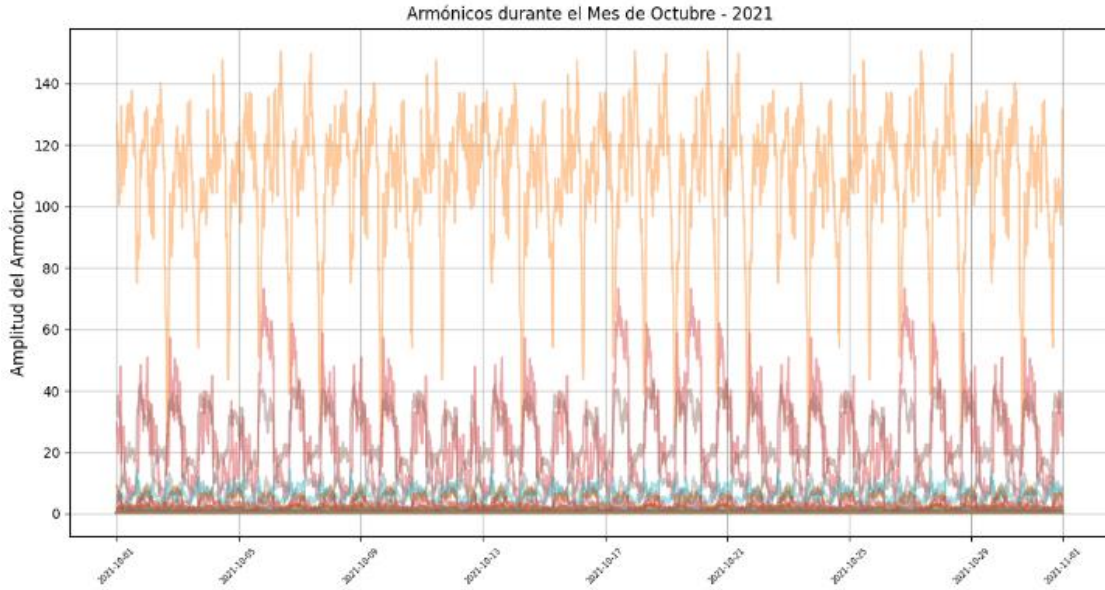


Figura 4: Comportamiento de los armónicos en voltaje

Por su parte, la Fig. 4 presenta la evolución de los armónicos de voltaje durante el mismo período. Algunos armónicos alcanzaron picos superiores a 140, evidenciando una inyección sostenida de distorsión. Este comportamiento mostró la presencia constante de cargas no lineales, especialmente asociadas a frecuencias armónicas altas.

Para una evaluación más detallada y la identificación de los armónicos más representativo durante el mes, se comparó la distorsión armónica individual DI con los límites establecidos en las normativas de calidad de

energía IEEE 519-2022 y la Regulación Nro. ARCONEL 009/2024. Esta comparación permitió verificar el cumplimiento de los valores máximos permisibles de distorsión armónica en redes de distribución eléctrica de media tensión.

La Fig. 5 presentan los armónicos en corriente, mientras que la Fig. 6 ilustra los armónicos en voltaje que excedieron los límites establecidos en las normativas vigentes. Estas figuras tomaron como referencia un día específico, permitiendo identificar los armónicos más influyentes en dicho período.

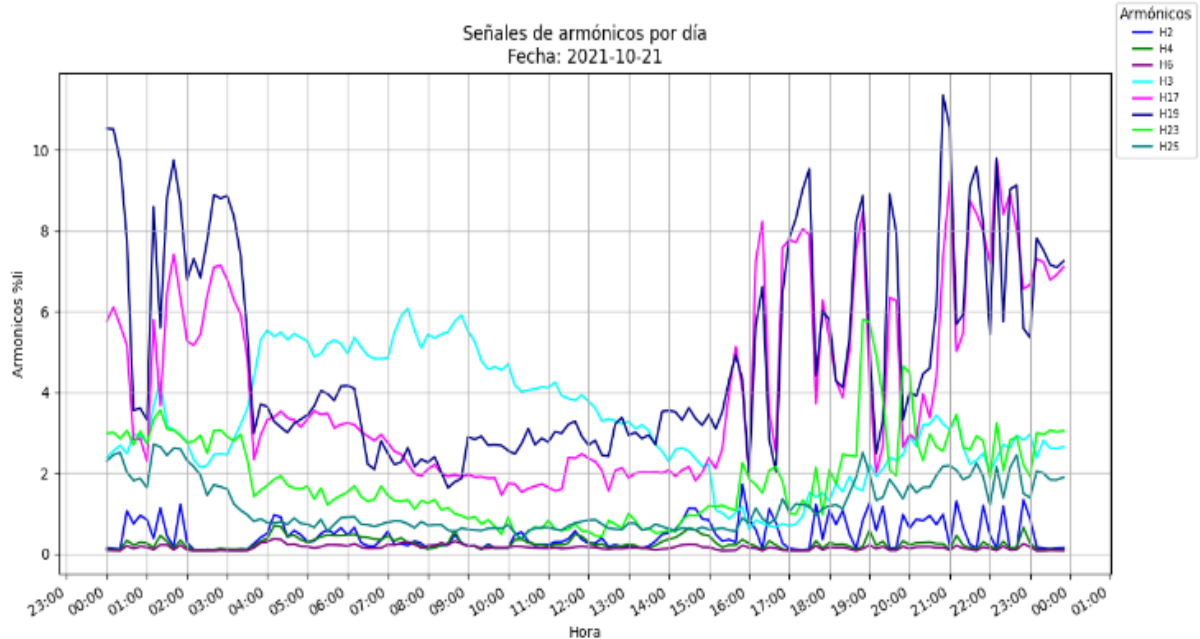


Figura 5: Armónicos individuales (H2–H50) en corriente que superan los límites establecidos por la normativa IEEE 519-2022 durante el día 2021-10-21

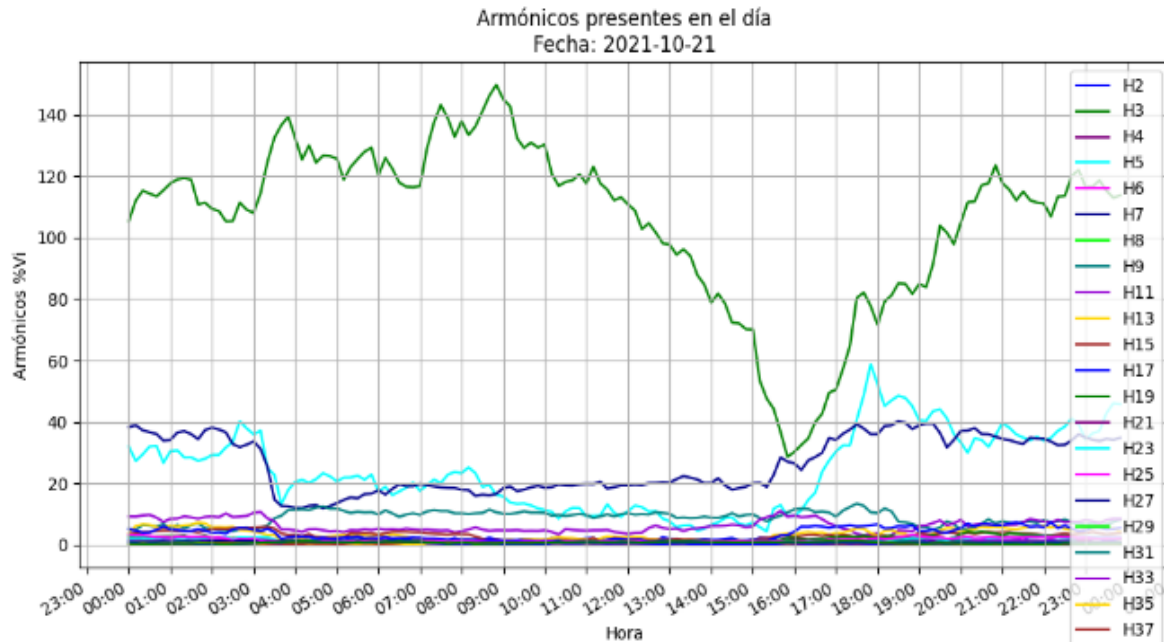


Figura 6: Armónicos individuales (H2–H50) en voltaje que superan los límites establecidos por la normativa IEEE 519-2022 durante el día 2021-10-21

Dado que los armónicos mostrados en la Fig. 5 presentaron variaciones horarias, se requirió un análisis más detallado. Por ejemplo, el armónico H17 superó el 4 % de distorsión armónica el 21 de agosto, entre las 16:00 y 04:00. Este comportamiento se repitió en otros días y horarios con distintos armónicos, lo que evidenció condiciones anómalas que podrían afectar la calidad del sistema eléctrico.

Para ello, se aplicaron técnicas de análisis multivariante como el Análisis de Componentes Principales (PCA) y el agrupamiento mediante K-Means, para reducir la dimensionalidad del conjunto de datos y segmentar comportamientos armónicos más influyentes.

3.2 Análisis de Componentes Principales

La presencia de múltiples armónicos con diferentes variaciones a lo largo del mes dificulta el análisis, por lo que se realizó una reducción de datos para entender el comportamiento armónico sin perder información relevante. Para este propósito, se utilizó la biblioteca *Scikit-learn* para aplicar técnicas de reducción de dimensionalidad mediante Análisis de Componentes Principales (PCA), lo que permitió identificar patrones dominantes en los datos. Asimismo, se emplearon *Pandas* y *NumPy* para la manipulación eficiente de los datos, y *Seaborn* junto con *Matplotlib* para la visualización de los resultados transformados.

A partir de la matriz X (5), se calculó la matriz de covarianza C (8), cuya descomposición permitió convertir los datos originales en un nuevo espacio de componentes principales. El número óptimo de PCA se determinó mediante el método del codo, mostrado en la Fig. 7.

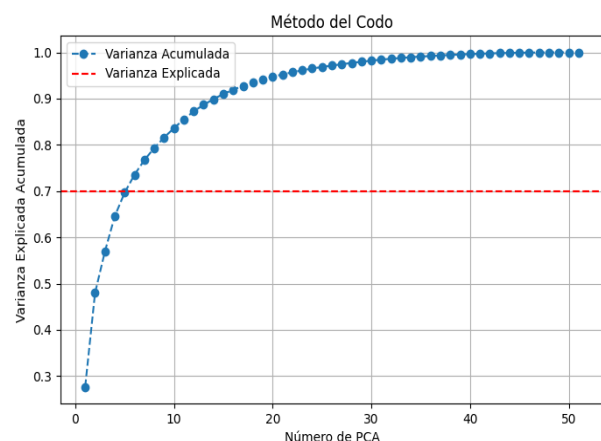


Figura 7: Número óptimo de PCA para corriente

La curva azul representa la varianza explicada, es decir, la proporción de la variabilidad total de los datos que logra ser representada por cada componente principal. En este análisis, las primeras cinco componentes acumulan cerca del 70 % de dicha varianza, lo que resultó suficiente para describir adecuadamente la estructura de los datos.

A partir de ese punto, la incorporación de componentes adicionales ya no aportó información relevante, lo que justifica su exclusión del modelo. Este comportamiento concuerda con lo analizado en [18], donde se consideró que las primeras cuatro componentes explicaban entre el 60 % y 70 % de la varianza total, umbral utilizado como criterio para seleccionar componentes significativas en el análisis de facturación eléctrica y reducir la dimensionalidad de los datos.

3.3 Análisis de Clústeres

Una vez obtenidas las componentes principales, se procedió a la aplicación de técnicas de análisis de clústeres para segmentar los armónicos en grupos con características similares, facilitando así la interpretación del comportamiento armónico. Se utilizaron las bibliotecas *SciPy* junto con *Matplotlib* para el análisis jerárquico y la visualización del dendrograma.

El número óptimo de clústeres se determinó mediante el enfoque jerárquico basado en dendrogramas, utilizando el método de Ward y la distancia euclídea. En la estructura jerárquica obtenida (Fig. 9), el eje x representa los intervalos temporales analizados, mientras que en el eje y muestra la disimilitud entre los grupos, proporcionando una medida cuantitativa de las diferencias entre observaciones.

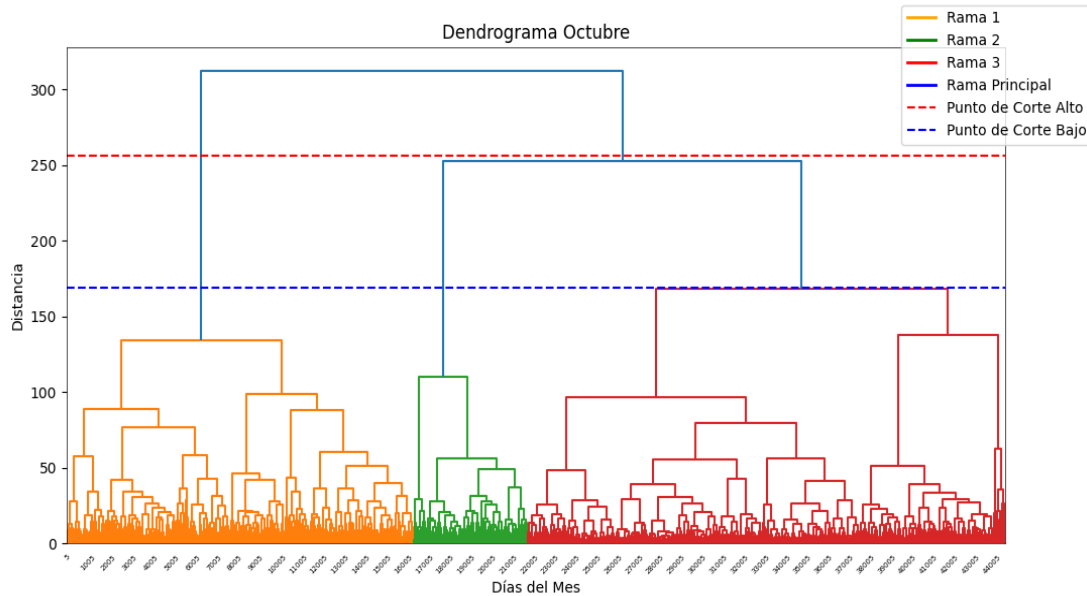


Figura 8. Estructura jerárquica de la Distorsión Armónica

En el dendrograma se distinguen tres ramas principales. La Rama 1 (color naranja) agrupa días con niveles de distorsión armónica similares y baja variabilidad. La Rama 2 (verde) presenta una disimilitud moderada, mientras que la Rama 3 (roja) reúne días con los mayores niveles de distancia, reflejando comportamientos armónicos más divergentes, posiblemente asociados a variaciones operativas significativas de la red.

Para definir el número óptimo de clústeres, se consideraron dos puntos de corte. El punto de corte alto, ubicado en una distancia aproximada de 250, sugirió la existencia de dos clústeres, uno formado por la Rama 1 y otro que agrupa las Ramas 2 y 3. En cambio, el punto de corte bajo cercano a 150, permitió diferenciar claramente tres clústeres distintos, cada uno correspondiente a una de las ramas.

En función del comportamiento observado y los criterios de segmentación aplicados por [11], quienes evaluaron distintas configuraciones con dos, tres, cuatro hasta cinco clústeres, concluyeron que dos clústeres ofrecía el mejor resultado para su caso de estudio. Considerando lo anterior, en el presente estudio se determinó que la segmentación de tres clústeres reflejó con mayor precisión las características del comportamiento armónico observado en la Fig. 9.

Una vez determinado el número de clústeres, se validó la segmentación utilizando el algoritmo K-Means, utilizando las bibliotecas *Scikit-learn*, *Pandas*, *NumPy* y *Matplotlib*. La representación tridimensional de los resultados se presenta en la Fig. 9, mientras que las similitudes entre clústeres, los días agrupados y las principales fuentes de generación armónica se resumen en la Tabla 4.

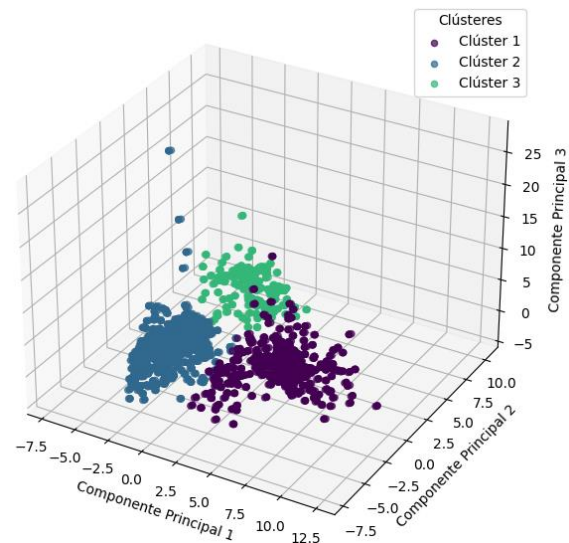


Figura 9: Distribución de clústeres en el espacio de CP

Tabla 4: Relación de Armónicos Dominantes y Frecuencia Temporal de cada clúster

Clúster	Armónicos	Día	H17	H19	H3	H7	Fuentes de Generación
1	H7, H17, H19	Martes, miércoles, domingo	12:40 - 16:40	12:50 - 16:50	-	11:30 - 15:30	Electrónica de Potencia, rectificadores, inversores
2	H3, H7, H19	Lunes, sábado	-	03:50 - 07:50	16:40 - 19:10	06:30 - 09:40	Cargas monofásicas, rectificadores, inversores
3	H7, H17, H19	Jueves, viernes	16:00 - 20:00	15:50 - 19:50	-	16:00 - 19:40	Electrónica de Potencia, rectificadores, inversores

El clúster 1 agrupa los armónicos H7, H17 y H19, con mayor incidencia durante los días martes, miércoles y domingo. El armónico H17, con el 26 % de incidencia se registró entre las 12:40 y 16:40, asociado a procesos de molienda y operación de hornos mediante variadores de frecuencia. El H19 aportó el 28 % y se presenta de 12:50 a 16:50, reflejando el funcionamiento de rectificadores en etapas de control térmico. Por su parte, el H7 alcanzó el 32 %, destacándose entre las 11:30 y 15:30, vinculado a sistemas de ventilación de hornos y procesos automatizados de despacho.

Por otro lado, el clúster 2 comprende los días lunes y sábado, con predominancia de los armónicos H3, H7 y H19. El H19 apareció con el 27 % de incidencia de 03:50 a 07:50, relacionado con el arranque temprano de bandas transportadoras. El H7 se registró entre 06:30 y 09:40, con el 30 % de incidencia en las etapas iniciales de molienda. El H3 con el 24% presente de 16:40 a 19:10, está asociado a cargas monofásicas no lineales, como iluminación y auxiliares de cierre de jornada. Aunque el sistema es trifásico, la aparición del H3 reveló desequilibrios causados por cargas monofásicas conectadas entre fase y neutro.

Por último, el clúster 3 corresponde a los días jueves y viernes, con presencia de H7, H17 y H19, pero en horarios más extendidos. El H7 con el 28% de incidencia de 16:00 a 19:40, el H17 con el 31% de 16:00 a 20:00 y H19 con el 34% de 15:50 a 19:50. La presencia de estos armónicos, al igual que el clúster 1, indican la operación de rectificadores, variadores de frecuencia, e inversores.

Como se evidencia en los clústeres, el armónico H3 superó el 10 %, mientras que H17 y H19 excedió el 4 %, sobrepasando los límites establecidos por la norma IEEE 519-2022 y la Regulación ARCONEL 009/2024. En contraste, el armónico H7 se mantuvo dentro del límite del 10 % permitido, lo que sugiere que su origen está relacionado con cargas de menor impacto en la red eléctrica. Estas distorsiones se registraron principalmente durante los turnos rotativos de operación de la cementera, organizados en tres jornadas de ocho horas, matutina (06:00–14:00), vespertina (14:00–22:00) y nocturna (22:00–06:00).

La aparición de estos armónicos se relaciona directamente con la operación simultánea de equipos como hornos de arco, variadores de frecuencia y sistemas

de control, lo que coincide con [24], donde se identifica a los variadores de frecuencia como la principal fuente de distorsión armónica en una planta de tratamiento de aguas residuales, superando incluso los límites establecidos por la norma IEEE 519.

Esta afectación revela la necesidad de implementar estrategias de mitigación dirigidas tanto a los períodos de mayor incidencia como a los equipos responsables de la generación de armónicos. En [24] se proponen soluciones como la instalación de filtros activos de compensación en tiempo real, el uso de algoritmos de predicción y la incorporación de sistemas de monitoreo continuo.

En este sentido, la metodología desarrollada en este estudio no solo permite identificar patrones armónicos y sus fuentes de generación, sino que también facilita la toma de decisiones y puede aplicarse a otras industrias con características similares, contribuyendo al cumplimiento normativo y a la mejora de la calidad del sistema eléctrico.

4. CONCLUSIONES

La metodología aplicada integró técnicas de análisis univariante y multivariante. El análisis univariante permitió evaluar el cumplimiento normativo, determinándose que los armónicos más influyentes y que superan los límites establecidos son el H3, con valores superiores al 10 %, y los armónicos H17 y H19, que exceden el 4 % permitido por la normativa IEE 519-2022 y Regulación ARCONEL 009/2024 para sistemas de media tensión. Por otro lado, el análisis multivariante, mediante minería de datos, identificó relaciones simultáneas entre múltiples armónicos que varían según el horario, el día de la semana o el tipo de carga conectada, lo cual facilitó la detección de las principales fuentes de generación de distorsión dentro del sistema, siendo los variadores de frecuencia los elementos con mayor incidencia.

La aplicación del Análisis de Componentes Principales (PCA) permitió reducir la dimensionalidad de los datos armónicos sin pérdida significativa de información, lo que facilitó su representación e interpretación. Posteriormente, mediante el uso del método K-Means, se agruparon los datos en clústeres bien definidos, permitiendo identificar períodos críticos con mayor concentración de distorsión armónica.

El modelo desarrollado en Python demostró ser una herramienta eficiente para la evaluación de la calidad de la energía, con capacidad de adaptarse a diferentes bases de datos y condiciones de medición en redes de distribución eléctrica. Su implementación se llevó a cabo utilizando bibliotecas como *Pandas*, *NumPy*, *Scikit-learn* y *Matplotlib*, dentro de un entorno Spyder, lo que facilita su reutilización y ajuste. Además, se evaluó la calidad de la agrupación generada por el algoritmo K-Means mediante métricas como la inercia y el análisis visual de los centroides. Esta metodología puede replicarse fácilmente en otros casos similares, para evaluar el comportamiento armónico y su impacto en la calidad del sistema eléctrico.

Para investigaciones futuras, se ampliará el análisis a un periodo anual y se explorará el impacto de la inyección de armónicos controlada, esto como estrategia para mitigar las perturbaciones en los sistemas de distribución eléctrica. Adicionalmente, se plantea realizar un análisis de impacto en la carga en la red considerando rotación de procesos, cuyos resultados serán de utilidad en alimentadores que comparten diferentes tipos de carga residencial, comercial e industrial.

5. GLOSARIO

PQM (Power Quality Monitor): Dispositivo que registra parámetros eléctricos como tensión, corriente y distorsión armónica.

THD (Distorsión Armónica Total): Porcentaje que representa la relación entre la suma de las componentes armónicas y la componente fundamental de una señal eléctrica.

TDD (Distorsión Total de Demanda): Porcentaje de distorsión armónica calculado respecto a la corriente de demanda máxima del sistema.

THD (Total Harmonic Distortion): Indicador del nivel de distorsión armónica en una señal eléctrica, expresado como porcentaje del valor fundamental. Se especifica como ITHD para corriente y VTHD para voltaje.

PCA (Análisis de Componentes Principales): Método estadístico que reduce la cantidad de variables conservando la mayor parte de la información original.

K-Means: Algoritmo de agrupamiento que divide los datos en grupos con características similares.

ARCONEL: Agencia de Regulación y Control de Electricidad.

Regulación ARCONEL 009/24: Regulación para la calidad del servicio de distribución y comercialización de energía eléctrica.

IEEE 519-2022: Norma internacional que establece límites recomendados para la distorsión armónica en sistemas eléctricos de potencia.

6. REFERENCIAS BIBLIOGRÁFICAS

- [1] M. M. Irfan, M. Alharbi, y C. H. H. Basha, «Artificial neural network controlled DSTATCOM for mitigating power quality concerns in solar PV and wind system», *Sci. Rep.*, vol. 15, n.o 1, 2025, doi: 10.1038/s41598-025-88540-0.
- [2] M. Pérez Martínez et al., «Desarrollo de aplicaciones en el software Scilab para el análisis de armónicos en sistemas industriales», *Ing. Energética*, vol. 43, n.o 1, pp. 24-34, abr. 2022.
- [3] L. Amaya-Vásquez, E. Inga-Ortega, L. Amaya-Vásquez, y E. Inga-Ortega, «Localización de Distorsiones Armónicas en Sistemas Eléctricos usando Sensado Comprimido», *Ing. Compet.*, vol. 24, n.o 1, jun. 2022, doi: 10.25100/iy.24i1.11037.
- [4] «IEEE Std 519-2022 (Revision of IEEE Std 519-2014)», *IEEE Std 519-2022 Revis.* IEEE Std 519-2014, pp. 1-31, ago. 2022, doi: 10.1109/IEEESTD.2022.9848440.
- [5] ARCONEL, Resolución Nro. ARCONEL-016/2024. 2024. [En línea]. Disponible en: <https://controlelectrico.gob.ec/wp-content/uploads/downloads/2024/10/Regulacion-ARCONEL-006-24-1-3.pdf>
- [6] J. C. Gonzales Arispe y A. Y. Marca Yucra, «Análisis de armónicos en sistemas eléctricos de potencia y normas regulatorias: impacto del horno de arco eléctrico de acerías», *Ingeniare Rev. Chil. Ing.*, vol. 32, pp. 0-0, 2024, doi: 10.4067/s0718-33052024000100201.
- [7] S. S. Cembranel, F. Lezama, J. Soares, S. Ramos, A. Gomes, y Z. Vale, «A Short Review on Data Mining Techniques for Electricity Customers Characterization», en *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*, mar. 2019, pp. 194-199. doi: 10.1109/GTDAAsia.2019.8715891.
- [8] B. L. Tarcsay, Á. Bárkányi, T. Chován, y S. Németh, «A Dynamic Principal Component Analysis and Fréchet-Distance-Based Algorithm for Fault Detection and Isolation in Industrial Processes», *Processes*, vol. 10, n.o 11, Art. n.o 11, nov. 2022, doi: 10.3390/pr10112409.
- [9] L. J. Lepolesa, S. Achari, y L. Cheng, «Electricity Theft Detection in Smart Grids Based on Deep Neural Network», *IEEE Access*, vol. 10, pp. 39638-39655, 2022, doi: 10.1109/ACCESS.2022.3166146.
- [10] J. Castro, P. Soto, R. Reategui, y T. Castillo, «Partición de una Red Eléctrica de Distribución Aplicando Algoritmos de Agrupamiento K-means y DBSCAN», *Rev. Téc. Energ.*, vol. 20, n.o 1, Art. n.o 1, jul. 2023, doi: 10.37116/revistaenergia.v20.n1.2023.572.

- [11] L. Marrero et al., «Uso de algoritmo K-means para clasificar perfiles de clientes con datos de medidores inteligentes de consumo eléctrico: Un caso de estudio», *Ingeniare Rev. Chil. Ing.*, vol. 29, n.o 4, pp. 778-787, dic. 2021, doi: 10.4067/S0718-33052021000400778.
- [12] C. Yajure, «Aplicación de la metodología de Ciencia de Datos para analizar datos de facturación de energía eléctrica. Caso de estudio: Uruguay 2000-2022», *ResearchGate*, doi: 10.15381/risi.v15i1.23544.
- [13] C. Yajure, «Uso de algoritmos de Machine Learning para analizar los datos de energía eléctrica facturada en la Ciudad de Buenos Aires durante el período 2010-2021», *ResearchGate*, oct. 2022, doi: 10.22206/cyap.2022.v5i2.pp7-37.
- [14] O. Yu. Maryasin y A. I. Lukashov, «A Python Application for Hourly Electricity Prices Forecasting Using Neural Networks», en 2020 International Russian Automation Conference (RusAutoCon), sep. 2020, pp. 138-143. doi: 10.1109/RusAutoCon49822.2020.9208035.
- [15] «Advanced Distribution Measurement Technologies and Data Applications for Smart Grids: A Review», *ResearchGate*, dic. 2024, doi: 10.3390/en13143730.
- [16] E. I. F. Cruz, «Análisis de Componentes Principales de la Tarifa de Electricidad en México – Caso de Estudio», *Cienc. Lat. Rev. Científica Multidiscip.*, vol. 8, n.o 4, Art. n.o 4, sep. 2024, doi: 10.37811/cl_rcm.v8i4.13173.
- [17] D. Garcia-Alvarez y M. J. Fuente, «Estudio comparativo de técnicas de detección de fallos basadas en el Análisis de Componentes Principales (PCA)», *Rev. Iberoam. Automática E Informática Ind. RIAI*, vol. 8, n.o 3, pp. 182-195, jul. 2011, doi: 10.1016/j.riai.2011.06.006.
- [18] J. Zhuang, M. Dong, J. Lin, M. Liu, y K. Lin, «Electricity charge anomaly detection based on PCA-IK-means», en 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), ene. 2023, pp. 1714-1718. doi: 10.1109/ICPECA56706.2023.10076038.
- [19] E. Pérez-Montalvo et al., «Model of monthly electricity consumption of healthcare buildings based on climatological variables using PCA and linear regression», *Energy Rep.*, vol. 8, pp. 250-258, nov. 2022, doi: 10.1016/j.egy.2022.06.117.
- [20] J. Wang y X. Li, «Abnormal Electricity Detection of Users Based on Improved Canopy-Kmeans and Isolation Forest Algorithms», *IEEE Access*, vol. 12, pp. 99110-99121, 2024, doi: 10.1109/ACCESS.2024.3429304.
- [21] G. Ruiz Aranibar, «Análisis de conglomerados», *Rev. Varianza*, vol. 16, oct. 2019, Accedido: 10 de marzo de 2025. [En línea]. Disponible en: http://revistasbolivianas.umsa.bo/scielo.php?script=sci_abstract&pid=&lng=en&nrm=iso&tlng=
- [22] M. Dimishkieh y A. H. Ramzi, «An Enhanced k-Means Clustering Algorithm for Pattern Discovery in Healthcare Data», *Sage J.*, jul. 2019, doi: <https://doi.org/10.1155/2015/6157>.
- [23] M. F. Abdullah, N. H. Hamid, Z. Baharudin, M. A. M. Ayob, y M. A. F. M. Hamim, «Characteristic of third harmonic from synchronous generator passing through transformer and rectifier», en 2012 IEEE International Conference on Power and Energy (PECon), dic. 2012, pp. 689-694. doi: 10.1109/PECon.2012.6450303.
- [24] A. R. Bonilla y H. T. Le, «Analysis and Mitigation of Harmonics for a Wastewater Treatment Plant Electrical System», *WSEAS Trans. Circuits Syst.*, vol. 23, pp. 1-13, 2024, doi: 10.37394/23201.2024.23.1.



Alex Dario Mullo. - Nació en Salcedo, Ecuador en 1996. Recibió el título de Ingeniero Electromecánico en la Universidad Técnica de Cotopaxi, Ecuador en el 2019. En el año 2025, obtuvo el título de Magister en Electricidad mención Energías Renovables y Eficiencia Energética en la Pontificia Universidad Católica del Ecuador. Actualmente ejerce como profesional independiente, desarrollando proyectos en ingeniería electromecánica, eléctrica y telecomunicaciones para sectores industriales y residenciales.



José Miguel Reinoso. - Nació en Quito, Ecuador en el 2000, Se graduó de Ingeniero Eléctrico en la Universidad Politécnica Salesiana en el 2023 y de Máster en Energías Renovables y Eficiencia Energética en el 2025. Actualmente se dedica a las instalación y mantenimiento de generadores eléctricos y tableros de transferencia.



Marlon Santiago Chamba. - Nació en Loja, Ecuador en 1982. Obtuvo el título de Ingeniero Eléctrico en la Escuela Politécnica Nacional, Ecuador en el 2007. En el año 2016, obtuvo el título de Doctor en Ingeniería Eléctrica en la Universidad Nacional de San Juan,

Argentina. Actualmente trabaja en la Agencia de Regulación y Control de Electricidad. Sus áreas de investigación son: Mercados de Energía, Confiabilidad, Calidad, Evaluación de la seguridad del SEP.



Carlos Xavier Lozada. - Nació en Quito en 1995, Recibió su título de Ingeniero Eléctrico de la Escuela Politécnica Nacional en el 2020; en el 2024 recibió su título de Magister en Electricidad Mención Redes Eléctricas. Actualmente se desempeña como Ingeniero de

Investigación y Desarrollo en la Subgerencia Nacional de Investigación y Desarrollo de CENACE. Sus áreas de interés son: Sistemas Eléctricos de Potencia, Optimización Aplicada y Machine Learning.