



UNIVERSIDAD NACIONAL DE COLOMBIA
SEDE MANIZALES

TESIS DE MAESTRÍA

ANÁLISIS COMPARATIVO DE METODOLOGÍAS DE
CÁLCULO DE ENERGÍA INCIDENTE APLICABLES A
NIVELES DE TENSIÓN SUPERIORES A 15 kV

Alejandro Villegas Ruiz

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura
Departamento de Ingeniería Eléctrica, Electrónica y Computación
Manizales, Colombia
2025

ANÁLISIS COMPARATIVO DE METODOLOGÍAS DE CÁLCULO DE ENERGÍA INCIDENTE APLICABLES A NIVELES DE TENSIÓN SUPERIORES A 15 kV

Alejandro Villegas Ruiz

Tesis presentada como requisito parcial para optar al título de:

Magister en Ingeniería Eléctrica

Director:

Dr. Ing. Eduardo Antonio Cano Plata

Codirector:

Dr. Ing. Armando Jaime Ustariz Farfán

Grupo de Investigación:

Redes de Distribución y Potencia “GREDyP”

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura
Departamento de Ingeniería Eléctrica, Electrónica y Computación
Manizales, Colombia
2025

COMPARATIVE ANALYSIS OF INCIDENT ENERGY CALCULATION METHODOLOGIES APPLICABLE TO VOLTAGE LEVELS ABOVE 15 kV

Alejandro Villegas Ruiz

Thesis submitted as a partial requirement for the degree of:

Master of Science in Electrical Engineering

Advisor:

Ph. D. Eduardo Antonio Cano Plata

Co-advisor:

Ph. D. Armando Jaime Ustariz Farfán

Research Group:

Power Lines Distribution and Power System “GREDyP”

Universidad Nacional de Colombia
Faculty of Engineering and Architecture
Department of Electrical, Electronic and Computer Engineering
Manizales, Colombia
2025

AGRADECIMIENTOS

A mi mamá, Yhara Bibiana, por ser incondicional y siempre confiar y creer en mí. Gracias por todo el amor y apoyo que me has brindado en este camino.

A mis hermanas, Valentina y Vanessa, por su cariño y apoyo en todo momento.

A mi papá, Jorge Iván, quién siempre ha depositado su plena confianza en que lo iba a lograr.

A mi primo, Sebastián, a quién considero como un hermano, por su apoyo constante.

A mi abuela, Mariela, tías y tíos, Diana, Edna, Daliana, David y Andrés, por siempre estar conmigo y darme palabras de aliento.

A mis grandes amigos, Juan David, Edilson, Mario Andrés, Maria Alejandra, Camilo Augusto, Luis Bernardo, Antonio, Lorena, Juan Fernando, Jose David, Alexandra, Jeison y Jesús. Quienes siempre me han impulsado a ser mejor.

A Juan Carlos, quien fue mi mentor en arco eléctrico durante mi práctica, por siempre brindarme su confianza.

A mi equipo de trabajo, APOE, por brindarme su confianza y aportarme tanto en mi proceso.

Al Dr. Eduardo Antonio Cano, por ser mi director, quien me ayudó en enrutarme en diversas ocasiones durante la realización de esta tesis.

Al Dr. Armando Jaime Ustariz, por ser mi codirector, quien me ayudó a estructurar la tesis y me motivó a realizar un mejor trabajo.

Por último, a la Universidad Nacional de Colombia y al Departamento de Ingeniería Eléctrica, Electrónica y Computación por brindarme tanto en mi proceso como profesional.

RESUMEN

En esta tesis se presenta un análisis comparativo entre distintas metodologías para el cálculo de la energía incidente en niveles de tensión superiores a 15 kV. Como referencia principal se utilizó el método establecido en la norma IEEE Std. 1584-2018, el cual se contrastó con metodologías alternativas derivadas de los estudios desarrollados por EPRI y Terzija/Koglin. El objetivo de este análisis se basó en la identificación de las variables más relevantes para la predicción de factores de ajuste que permitan incrementar los valores de energía incidente, con el fin de representar adecuadamente la energía calorífica generada durante un arco eléctrico trifásico en encerramientos. Estos factores de ajuste se proponen considerando la principal limitación de las metodologías alternativas, las cuales fueron desarrolladas originalmente para estimar la energía incidente de arcos monofásicos al aire libre.

Se desarrolló un Modelo de Ajuste de Energía Incidente (MAEI) orientado a la predicción de factores de ajuste, seleccionado a partir del entrenamiento y validación de tres modelos de aprendizaje automático. El proceso de entrenamiento se realizó con las variables más relevantes identificadas durante el análisis comparativo entre metodologías, donde se examinó el comportamiento general de un conjunto de datos de entrada general y se construyó una matriz de correlaciones para evaluar su relación con los factores de ajuste propuestos, tanto en la metodología EPRI como en la de Terzija/Koglin. El conjunto de datos de entrada fue generado a partir del planteamiento de distintos escenarios obtenidos de subestaciones operando a niveles de tensión inferiores a 15 kV y considerando las variables requeridas por cada una de las metodologías de interés. Por su parte, la validación de los modelos se llevó a cabo utilizando métricas orientadas a la evaluación del desempeño en modelos de regresión aplicados a enfoques predictivos. Este proceso permitió analizar la dispersión y variabilidad de las predicciones en relación con los valores originales.

Finalmente, se evaluó el desempeño del MAEI frente a nuevos escenarios de estudio que abarcan niveles de tensión desde 4.16 kV hasta 35 kV. Los resultados obtenidos fueron comparados con los proporcionados por metodologías tradicionales y referencias de la literatura técnica. Esta evaluación tuvo como objetivo analizar la viabilidad del uso de metodologías alternativas en configuraciones con encerramiento, independientemente del nivel de tensión, mediante la aplicación de factores de ajuste estimados a través de modelos predictivos.

Palabras clave: NFPA 70E, IEEE 1584, arco eléctrico, energía incidente, equipo de protección personal, EPRI, Terzija Koglin, factor de ajuste, Ralph Lee, distancia de trabajo, gap entre electrodos, configuración de electrodos, arco en encerramientos.

ABSTRACT

This thesis presents a comparative analysis of various methodologies for calculating incident energy at voltage levels above 15 kV. The primary reference used is the method established in IEEE Std. 1584-2018, which is contrasted with alternative approaches derived from studies conducted by EPRI and Terzija/Koglin. The objective of this analysis is to identify the most relevant variables for predicting adjustment factors that increase incident energy values in order to appropriately represent the thermal energy generated during a three-phase arc flash within enclosures. These adjustment factors are proposed considering the main limitation of the alternative methodologies, which were originally developed to estimate incident energy from single-phase open-air arc flashes.

An Incident Energy Adjustment Model (MAEI) was developed for the prediction of these adjustment factors, selected based on the training and validation of three machine learning models. The training process used the most relevant variables identified in the comparative analysis between methodologies, where the general behavior of a comprehensive input dataset was examined, and a correlation matrix was built to assess the relationship between these variables and the proposed adjustment factors, both for the EPRI and Terzija/Koglin methods. The input dataset was generated based on different scenarios derived from substations operating at voltage levels below 15 kV, considering the specific variables required by each methodology of interest. Model validation was conducted using performance metrics suitable for regression models in predictive applications. This evaluation allowed the assessment of prediction dispersion and variability in relation to the original values.

Finally, the performance of the MAEI was assessed using new study scenarios covering voltage levels from 4.16 kV to 35 kV. The results were compared with those obtained through traditional methods and technical literature references. This evaluation aimed to assess the feasibility of applying alternative methodologies within enclosure configurations, regardless of voltage level, by using adjustment factors estimated through predictive models.

Keywords: NFPA 70E, IEEE 1584, arc flash, incident energy, personal protective equipment, EPRI, Terzija Koglin, adjustment factor, Ralph Lee, working distance, electrode gap, electrode configuration, enclosed arc.

CONTENIDO

1. INTRODUCCIÓN	1
1.1 Definiciones.....	2
1.2 Planteamiento del problema.....	3
1.3 Objetivos del presente trabajo.....	3
2. FUNDAMENTOS TEÓRICOS	5
2.1 Riesgos asociados al arco eléctrico.....	5
2.2 Metodologías de Cálculo de Energía Incidente	6
2.2.1 IEEE 1584 – 2002 y 2018.....	6
2.2.2 Método de Cálculo de Ralph Lee	8
2.2.3 Método a partir de ecuaciones Terzija/Koglin	8
2.2.4 Método a partir de ecuaciones de EPRI	11
2.3 Modelos de Machine Learning	13
2.3.1 Modelos de regresión aplicados.....	14
2.3.2 Escalado de datos.....	16
2.3.3 Evaluación de los modelos de Machine Learning	17
3. METODOLOGÍA PROPUESTA	18
3.1 Datos de entrada.....	21
3.1.1 Escenarios de entrenamiento y validación implementados.....	22
3.1.2 Sensibilidades de variables	23
3.1.3 Coeficiente de encerramiento y gradiente de tensión B.....	23
3.1.4 Variables calculadas.....	24
3.1.5 Métodos iterativos propuestos.....	24
3.1.6 Definiciones adicionales	27
3.2 Segmentación de casos.....	28
3.3 Análisis de variables.....	28
3.3.1 Análisis general de los conjuntos de datos.....	29
3.3.2 Análisis de las correlaciones entre variables	33
3.4 Selección de variables relevantes.....	38
3.5 Generación de modelos de Machine Learning.....	38
3.6 Entrenamiento de modelos.....	40
	XI

3.7	Validación de modelos y selección de MAEI	41
4.	IMPLEMENTACIÓN Y EVALUACIÓN DEL MAEI	49
4.1	Propuesta de implementación.....	49
4.2	Evaluación de desempeño del MAEI en niveles de tensión inferiores a 15 kV.....	50
4.3	Evaluación de desempeño del MAEI en niveles de tensión superiores a 15 kV.....	53
4.3.1	Validación con subestaciones de 33 kV	53
4.3.2	Validación con resultados de literatura técnica	54
5.	CONCLUSIONES Y FUTUROS DESARROLLOS.....	57
5.1	Conclusiones generales	57
5.2	Aportes	58
5.3	Futuros desarrollos	59
6.	REFERENCIAS BIBLIOGRÁFICAS.....	61

LISTA DE FIGURAS

Figura 1. Ejemplo de red neuronal tomada de [30]	16
Figura 2. Proceso metodológico para obtener MAEI	18
Figura 3. Distribución de atributos del conjunto de datos entrenamiento.....	32
Figura 4. Distribución de atributos del conjunto de datos validación.....	32
Figura 5. Mapa de calor – Matriz de correlaciones conjunto de datos entrenamiento	37
Figura 6. Mapa de calor – Matriz de correlaciones conjunto de datos validación	37
Figura 7. Evaluación de conjuntos de datos – Métrica MAE - ElasticNet	42
Figura 8. Evaluación de conjuntos de datos – Métrica RMSE - ElasticNet.....	43
Figura 9. Evaluación de conjuntos de datos – Métrica R2 - ElasticNet.....	43
Figura 10. Evaluación de conjuntos de datos – Métrica MAE - Polynomial Grade 2.....	44
Figura 11. Evaluación de conjuntos de datos – Métrica RMSE - Polynomial Grade 2.....	45
Figura 12. Evaluación de conjuntos de datos – Métrica R2 - Polynomial Grade 2.....	45
Figura 13. Evaluación de conjuntos de datos – Métrica MAE – Red Neuronal.....	46
Figura 14. Evaluación de conjuntos de datos – Métrica RMSE – Red Neuronal	47
Figura 15. Evaluación de conjuntos de datos – Métrica R2 – Red Neuronal.....	47
Figura 16. Propuesta de implementación de MAEI.....	49
Figura 17. Comparación entre metodologías para tensiones menores a 15 kV.....	52
Figura 18. Comparación entre metodologías para tensiones mayores a 15 kV.....	54
Figura 19. Comparación entre metodologías con datos de literatura técnica	55

LISTA DE TABLAS

Tabla 1. Escenarios implementados para el análisis.....	22
Tabla 2. Impedancias de secuencia de cada escenario.....	27
Tabla 3. Estadísticas descriptivas conjunto de datos entrenamiento	29
Tabla 4. Estadísticas descriptivas conjunto de datos validación.....	30
Tabla 5. Matriz de correlación conjunto de datos entrenamiento.....	34
Tabla 6. Matriz de correlación conjunto de datos validación.....	35
Tabla 7. Escenarios de estudio - MAEI en niveles de tensión inferiores a 15 kV.....	51
Tabla 8. Resultados - MAEI en niveles de tensión inferiores a 15 kV.....	52
Tabla 9. Escenarios de estudio - MAEI en niveles de tensión superiores a 15 kV.....	53

1. INTRODUCCIÓN

En Colombia, el Reglamento Técnico de Instalaciones Eléctricas (RETI) exige la identificación del riesgo por arco eléctrico mediante etiquetas de advertencia, así como la especificación del equipo de protección personal (EPP) adecuado para cada caso. Para cumplir con esta exigencia, es necesario realizar un análisis de energía incidente utilizando metodologías empíricas que permitan estimar este valor en función de las condiciones eléctricas evaluadas.

La IEEE Std. 1584-2018 es el estándar con la metodología de cálculo más aceptada para la estimación de energía incidente, dado que sus ecuaciones están basadas en datos experimentales y ofrecen alta confiabilidad en rangos de tensión entre 208 V y 15 kV. Sin embargo, su limitación para rangos de tensión mayores a 15 kV restringe su aplicación directa en celdas de 33 o 34.5 kV, situación común en subestaciones de media tensión. En estos casos, la NFPA 70E menciona el método de Ralph Lee [1], basado en el teorema de máxima transferencia de potencia. Sin embargo, las ecuaciones de este método generan resultados que tienden a sobredimensionar el EPP, llegando incluso a especificar aquellos con la máxima categoría de protección. De acuerdo con Ammerman, Gammon, Sen y Nelson en su artículo [2], sobreestimar el riesgo de arco eléctrico podría inducir al trabajador a generar los incidentes de arco a causa de la movilidad limitada o el agotamiento por calor asociado con el EPP voluminoso. Es por ello por lo que autores como Stokes y Sweeting son críticos frente al uso del teorema de máxima transferencia de potencia de Ralph Lee, al punto de recomendar, en sus publicaciones [3] [4], que la IEEE 1584 incluya una advertencia de que el valor contenido en este teorema no es adecuado para el análisis o la predicción del arco.

Ante estas limitaciones, se han desarrollado metodologías alternativas, tales como EPRI, Terzija/Koglin, Arc Pro y Duke HFC. Algunas de estas metodologías son utilizadas en software especializado, proporcionando valores de energía incidente ampliamente aceptados en la industria. Entre ellas, la metodología de Arc Pro destaca como una de las más sobresalientes; sin embargo, su acceso está limitado a través de una licencia de pago, lo que dificulta su uso de manera extensiva en diversas ocasiones.

Uno de los obstáculos principales de las metodologías alternativas antes mencionadas es que los cálculos resultantes corresponden a fallas monofásicas en entornos al aire libre, lo cual las hace inaplicables para celdas de media tensión. Por esta razón, en estos cálculos se incorporan factores de ajuste que consideran el confinamiento de la celda y la naturaleza trifásica del arco que suele generarse en espacios encerrados a causa de la corta separación entre electrodos. Sin embargo, los factores de ajuste mencionados carecen de un método de cálculo claro, lo que introduce un alto grado de incertidumbre en los estudios de energía incidente.

En este contexto, el presente trabajo plantea un análisis comparativo entre las metodologías IEEE 1584-2018, EPRI y Terzija/Koglin, con el fin de identificar las variables que tienen un impacto representativo en la estimación de factores de ajuste. Estas variables servirán como base para el desarrollo de un Modelo de Ajuste de Energía Incidente (MAEI) basado en técnicas estadísticas y de Machine Learning, con el objetivo de mejorar la precisión del cálculo de energía incidente en celdas de media tensión.

1.1 Definiciones

Arco eléctrico: Fenómeno que se produce como resultado de una descarga cuando la tensión entre dos puntos supera el límite de rigidez dieléctrica de un gas interpuesto (generalmente aire), en ciertas condiciones se forma un plasma que conduce la corriente eléctrica hasta que interviene la protección aguas arriba [5].

Celda tipo AIS: Celda aislada en aire (Air Insulated Switchgear).

Choque eléctrico: Contacto o aproximación a conductores o partes de un circuito energizado que produce el paso de corriente a través del cuerpo [6].

Configuración de electrodos: Orientación y disposición de los electrodos utilizados en las pruebas realizadas para el desarrollo del modelo de la IEEE Std. 1584–2018 [7].

Convección: La convección es una de las tres formas de transferencia de calor y se caracteriza porque se produce por intermedio de un fluido (aire, agua) que transporta el calor entre zonas con diferentes temperaturas [8].

Corriente de arco: Corriente eléctrica que se presenta cuando se produce un fenómeno de arco eléctrico.

Corriente de falla: Corriente entregada en un punto de falla a causa de un cortocircuito [6].

Corriente de falla atornillada: La corriente de falla atornillada (Bolted fault current) corresponde a la máxima corriente de cortocircuito cuando la impedancia es prácticamente nula [7].

DataFrame: Estructura de datos con dos dimensiones en la cual se puede guardar datos de distintos tipos (como caracteres, enteros, valores de punto flotante, factores y más) en columnas [8].

Dataset: Es un conjunto organizado de datos que se utiliza para realizar análisis o alimentar modelos de aprendizaje automático [9].

Distancia de trabajo: La distancia entre la cara y el área del pecho de una persona y una fuente potencial de arco [6].

Energía incidente: La cantidad de energía térmica expuesta sobre una superficie a una cierta distancia de la fuente generada durante un evento de arco eléctrico. La energía incidente generalmente es expresada en calorías sobre centímetro cuadrado (cal/cm²) [6].

EPP: Equipo de Protección Personal.

EPRI: Electric Power Research Institute.

Gap: Distancia o espacio entre electrodos o conductores energizados.

IAC AFLR: Internal Arc Classification. Protección contra arco eléctrico interno en una celda de media tensión. A: solamente a personal autorizado. FLR: Protección en las direcciones Front, Lateral, Rear [10].

IED: Intelligent Electronic Device.

IEEE: Institute of Electrical and Electronics Engineers.

Metrala: Fragmentos de materiales, usualmente metálicos, que son disparados como proyectiles a causa de un arco eléctrico.

NESC: National Electrical Safety Code.

NFPA 70E: Norma de la National Fire Protection Association para la Seguridad Eléctrica en Lugares de Trabajo [6].

Radiación: El fenómeno de la radiación consiste en la propagación de energía en forma de ondas electromagnéticas o partículas subatómicas a través del vacío o de un medio material [11].

Resistencia de arco: Oposición al flujo de corriente de arco cuando se genera un evento de arco eléctrico.

Resistencia de falla: Oposición al flujo de corriente de cortocircuito cuando se genera una falla en conductores.

RETI: Reglamento técnico de instalaciones eléctricas [12].

Script: Instrucciones escritas en un lenguaje de programación interpretado, cuya ejecución es llevada a cabo por otro programa en lugar de directamente por el procesador de la computadora [13].

Tensión de arco: Diferencia de potencial entre los electrodos fallados por arco eléctrico.

Tiempo de despeje: Temporización configurada de acuerdo con estudios de coordinación de protecciones para la apertura de interruptores que permitan aislar el circuito fallado.

1.2 Planteamiento del problema

¿Qué variables y condiciones deben considerarse para estimar factores de ajuste que permitan calcular con mayor precisión la energía incidente en celdas con tensión superior a 15 kV mediante metodologías alternativas?

1.3 Objetivos del presente trabajo

Objetivo general

Realizar un análisis comparativo entre las metodologías de cálculo de energía incidente IEEE 1584 – 2018, EPRI y Terzija/Koglin en celdas AIS, empleando herramientas estadísticas y modelos de Machine Learning para la identificación de variables de mayor relevancia y predicción de factores de ajuste.

Objetivos específicos

1. Establecer las ecuaciones y variables de entrada utilizadas para el cálculo de energía incidente en las metodologías IEEE Std. 1584 – 2018, Ralph Lee, EPRI y Terzija/Koglin.
2. Definir e implementar una metodología basada en técnicas de Machine Learning para el desarrollo de un Modelo de Ajuste de Energía Incidente (MAEI), orientado a la predicción de factores de ajuste.
3. Determinar escenarios para el entrenamiento y validación de los modelos a partir de subestaciones con celdas tipo AIS a niveles de tensión inferiores a 15 kV.
4. Analizar el comportamiento entre las variables planteadas en los escenarios de entrenamiento y validación mediante evaluación de correlaciones.
5. Entrenar y evaluar modelos de Machine Learning para la predicción de factores de ajuste y seleccionar el de mejor desempeño como MAEI.
6. Validar el desempeño del MAEI en niveles de tensión inferiores a 15 kV, comparándolo con la metodología IEEE Std. 1584-2018.
7. Validar el desempeño del MAEI en niveles de tensión superiores a 15 kV, contrastándolo con la metodología de Ralph Lee y resultados reportados en la literatura.



2. FUNDAMENTOS TEÓRICOS

En este capítulo se presentan los fundamentos teóricos que sustentan el desarrollo de la presente tesis. Se abordan, en primer lugar, los riesgos inherentes al fenómeno del arco eléctrico, incluyendo sus causas y consecuencias. Luego se describen las principales metodologías utilizadas para el cálculo de la energía incidente, detallando sus ecuaciones, rangos de aplicación y limitaciones. Finalmente, se introduce en términos generales sobre los modelos de Machine Learning empleados en este estudio, proporcionando una visión general de su funcionamiento y evaluación.

2.1 Riesgos asociados al arco eléctrico

El arco eléctrico genera grandes riesgos para las personas que están cerca a este fenómeno. Algunos efectos mencionados en literatura de ABB [5] son:

- Daños en la retina del ojo.
- Daños en el oído.
- Inhalación de gases tóxicos.
- Lesiones debido a la metralla sobrecalentada.
- Quemaduras.

Daños en la retina del ojo: Cuando se produce un arco eléctrico se genera un medio plasmático con una gran intensidad lumínica, la cual puede provocar daños en la retina del ojo.

Daños en el oído: El arco eléctrico puede emitir hasta 160 dB (un arma de fuego emite 130 dB), por lo cual el tímpano puede sufrir lesiones a tal punto de provocar sordera. La emisión sonora se genera debido al choque constante de iones y de partículas.

Inhalación de gases tóxicos: Dentro del arco se pueden generar temperaturas de hasta 8000 °C, esto provoca que los materiales circundantes al arco se evaporicen y generen humos tóxicos.

Lesiones debido a la metralla sobrecalentada: El constante paso de energía en el momento en el que se produce el arco genera un aumento en la temperatura y otros fenómenos físicos como la convección y radiación, esto provoca un sobrecalentamiento en el aire y una gran presión. Cuando ocurre esto, distintos materiales metálicos se sobrecalientan por la temperatura y adicionalmente se desprenden y salen disparados a causa de la presión, a esto se le conoce como metralla sobrecalentada, que al golpear a una persona puede provocarle lesiones e inclusive una conmoción cerebral.

Quemaduras: El constante choque de iones negativos y positivos en un arco eléctrico genera un gran aumento en la temperatura, lo cual ocasiona el sobrecalentamiento del aire que, al ser expulsado por la presión interna, puede tener contacto con la piel humana y provocar fácilmente quemaduras de segundo e inclusive tercer grado.

El arco eléctrico genera riesgos altos, ya que puede generar heridas graves e inclusive la muerte, por lo tanto, es necesario, además de reducir la probabilidad de ocurrencia del evento de arco

eléctrico y de mitigar la energía incidente que produce este, establecer que equipo de protección personal (EPP) debe llevar la persona que podría estar cerca a la ocurrencia de este fenómeno, normalmente en celdas de subestaciones o en trabajos de línea viva.

Para saber que EPP se debe usar, es necesario calcular la energía incidente que provoca el arco eléctrico, la cual está dada convencionalmente en unidades de cal/cm². Debido a que el arco eléctrico tiene tantos fenómenos físicos implicados, su cálculo no está establecido por una ecuación o metodología específica, por lo tanto, para hacer este cálculo en términos ideales, se deben hacer pruebas de laboratorio para saber específicamente la energía incidente que se puede generar en la celda. No obstante, los costos de estas pruebas y la complejidad para realizarlas representan un gran inconveniente, así que, mientras se encuentran formas de subsanar esta falencia en pruebas de laboratorio, es permitido hacer el cálculo de este parámetro por medio de ecuaciones empíricas.

2.2 Metodologías de Cálculo de Energía Incidente

En el Anexo D de la NFPA 70E [6] se encuentran diferentes metodologías de análisis tradicionales recomendadas para la estimación de energía incidente, las cuales se utilizan como base para la selección del EPP necesario para mitigar los riesgos ante choque y arco eléctrico. Entre las metodologías mencionadas, se encuentra la establecida en la IEEE Std. 1584 – 2018 [7], aplicada en niveles de tensión comprendidos entre 0.208 kV y 15 kV. A su vez, se menciona el método propuesto por Ralph Lee, el cual es un modelo basado en principios físicos para estimar la energía calorífica que puede ser usado de manera general ante cualquier nivel de tensión, pero que se vuelve cada vez más conservativo frente al aumento de tensión.

Por otro lado, las metodologías de cálculo alternativas han sido desarrolladas a partir de estudios independientes y han sido incorporadas en los análisis de energía incidente para tensiones superiores a 15 kV debido a que ofrecen estimaciones más realistas de la energía calorífica generada durante un evento de arco eléctrico.

En este apartado se describen tanto las metodologías tradicionales previamente mencionadas como las metodologías alternativas propuestas en estudios realizados por EPRI y por Terzija/Koglin, las cuales sirven de base para el análisis comparativo que se plantea en esta tesis.

2.2.1 IEEE 1584 – 2002 y 2018

La IEEE 1584 es una guía para establecer los riesgos de arco eléctrico a partir de ecuaciones empíricas sustentadas en pruebas realizadas a equipos variados. Dichas ecuaciones son usadas ampliamente en la industria en el cálculo de la energía incidente, brindando valores aceptados por el sector, para el dimensionamiento del EPP correspondiente. La versión del 2018 [7] se basa en más de 1800 pruebas, mientras que la IEEE Std. 1584–2002 [14] se fundamenta en aproximadamente 300 pruebas, es decir, la IEEE Std. 1584–2018 amplió en gran medida la cantidad de pruebas realizadas e incluyó factores de corrección asociados a las diversas configuraciones de electrodos en los equipos, que la metodología del 2002 no tenía en cuenta. Estos factores

impactaron sustancialmente los resultados para configuraciones horizontales, garantizando mayor seguridad para equipos que tuviesen electrodos con dicha orientación.

Metodología de cálculo

Las metodologías de cálculo de ambas versiones de la guía IEEE 1584 contienen un conjunto de ecuaciones empíricas basadas en diversas pruebas realizadas. La versión del 2018 considera una mayor cantidad de variables y ecuaciones en comparación con la versión del 2002. Entre las variables adicionales consideradas se incluyen las dimensiones del encerramiento, la corriente de arco mínima (que impacta en el tiempo de actuación de la protección sobrecorriente de característica de tiempo inverso), así como configuraciones de electrodos que no eran contempladas en metodologías anteriores. Esta última variable resulta especialmente relevante, ya que las configuraciones incluidas en las normas IEEE Std. 1584-2002 [14] y NFPA 70E [6] se limitan a VCB (electrodos verticales en encerramiento) y VOA (electrodos verticales al aire libre). En cambio, la metodología propuesta en la IEEE Std. 1584-2018 [7] amplía el alcance del análisis, permitiendo simular la energía incidente en configuraciones adicionales como VCBB (electrodos verticales terminados en una barrera aislante en encerramiento), HCB (electrodos horizontales en encerramiento) y HOA (electrodos horizontales en aire libre). Dichas configuraciones afectan en el resultado final de la energía incidente debido a la forma en que se genera la nube de plasma del arco eléctrico y, por ende, en la incidencia que tendría sobre una persona que está cerca al fenómeno. Particularmente, las configuraciones horizontales son las más peligrosas.

En la IEEE Std. 1584-2018 [7] también se incluye un factor de corrección para el cálculo de la corriente de arco mínima, con el fin de evaluar el impacto de una protección con característica de tiempo inverso. Al ser analizada esta corriente en una gráfica de tiempo inverso de la protección asociada, se obtiene un tiempo de despeje más prolongado con respecto al tiempo de despeje que se produce con la corriente de arco inicial. En algunas ocasiones el aumento de este tiempo eleva la energía incidente final con una corriente de arco menor a la inicialmente estimada. Por tal motivo, es importante realizar ambos cálculos y validar cuál es la situación con mayor riesgo.

Para llevar a cabo estos métodos de cálculo, se debe seguir los procedimientos descritos en las guías de cada versión [14] [7], en las cuales se presentan todos los insumos necesarios para calcular la energía incidente.

Rangos de aplicación

Los rangos de aplicación de cada metodología se establecen a partir de los estándares IEEE Std. 1584-2002 e IEEE Std. 1584-2018 [14] [7].

- **IEEE Std. 1584-2002:**
 - Tensión nominal L-L entre 208 V y 15 000 V.
 - Frecuencia de 50 Hz a 60 Hz.
 - Corriente de cortocircuito (rms simétrica) de 700 A a 106 000 A.
 - Distancia entre conductores (gaps) de 13 mm a 152 mm.
- **IEEE Std. 1584-2018:**
 - Tensión nominal L-L entre 208 V y 15 000 V.

- Frecuencia de 50 Hz a 60 Hz.
- Corriente de cortocircuito (rms simétrica) de:
 - 208 V a 600 V: 500 A a 106 000 A.
 - 601 V a 15 000 V: 200 A a 65 000 A.
- Distancia entre conductores (gaps) de:
 - 208 V a 600 V: 6.35 mm a 76.2 mm.
 - 601 V a 15 000 V: 19.05 mm a 254 mm.

2.2.2 Método de Cálculo de Ralph Lee

Este método de cálculo estima la energía incidente basado en un valor teórico máximo disipado por fallas de arco [1]. Sus ecuaciones solo tienen en cuenta el arco en aire libre, es decir, no posee parámetros de cálculo que aborden encerramientos. Aun así, este método puede llegar a arrojar valores poco realistas más allá de 4.16 kV, considerándose un modelo demasiado conservador en toda la industria [15].

Metodología de cálculo

Mediante las siguientes ecuaciones se calcula la energía incidente y la distancia de quemadura de acuerdo con el método de Ralph Lee.

$$E = \frac{793 \times F \times V \times t_A}{D^2} \quad (1)$$

$$D_c = (53 \times MVA \times t)^{\frac{1}{2}} \quad (2)$$

Donde,

- E Energía incidente estimada en el punto máximo de potencia [cal/cm²];
- F Corriente de cortocircuito de falla atornillada [kA];
- V Tensión del sistema [kV];
- t_A Tiempo de duración del arco [s];
- D Distancia de trabajo desde el arco [pulgadas];
- D_c distancia de la persona a la fuente de arco para justo una quemadura curable (es decir, la temperatura de la piel se mantiene a menos de 80°C.) [pies];
- MVA Falla sólida MVA en el punto involucrado;
- t Tiempo de exposición al arco [s].

Rangos de aplicación

Calcula la energía incidente y la frontera de relámpago de arco abierto al aire libre; conservador sobre los 600 V y se hace más conservador al incrementar la tensión [6].

2.2.3 Método a partir de ecuaciones Terzija/Koglin

Este método utiliza como bases teóricas los estudios desarrollados por Vladimir V. Terzija y Hans-Jürgen Koglin, en las cuales, a partir de sus artículos [16] [17] [18], se presentan ecuaciones para el cálculo de resistencia de arco y tensión de arco.

Marroquin, Rehman y Madani presentan las ecuaciones para el cálculo de la energía incidente en [19], estableciendo una ecuación para calcular la energía incidente en aire libre, y otra ecuación para calcular este valor en un encerramiento.

Metodología de cálculo

De acuerdo con los estudios de Terzija/Koglin en [17] y [18], la forma de onda de la tensión de arco tiende a comportarse como una onda cuadrada. Como resultado del desarrollo matemático presentado en *New Approach to Arc Resistance Calculation* [16], la resistencia de arco se calcula a partir de la ecuación (3).

$$R_{arc} = \frac{2 \times \sqrt{2} \times U_a}{\pi \times I_{arc}} \quad (3)$$

Donde,

R_{arc} Resistencia de arco [Ω];
 U_a Magnitud de la tensión de arco [V];
 I_{arc} Corriente de arco [A].

La magnitud de la tensión de arco U_a se puede calcular así:

$$U_a = E_a \times G \quad (4)$$

Donde,

E_a Gradiente de tensión de arco [V/m];
 G Gap entre electrodos [m].

Para gaps largos el gradiente de arco se puede calcular con la ecuación (5).

$$E_a = B + \frac{5000}{I_{arc}} \quad (5)$$

Donde,

B Gradiente de tensión [V/m];
 I_{arc} Corriente de arco [A].

Una de las limitaciones que trae este conjunto de ecuaciones es el cálculo de la corriente de arco, puesto que su valor está sujeto a la resistencia de arco que se presente en el momento de falla por arco eléctrico. Es por ello por lo que Marroquin, Rehman y Madani realizan el cálculo de estos valores de forma iterativa [19].

La tensión de arco rms se calcula a partir de la ecuación (6) partiendo de que la tensión en un evento de falla de arco tiende a generar un comportamiento de onda cuadrada.

$$V_{arc} = \frac{U_a \times 2 \times \sqrt{2}}{\pi} \quad (6)$$

Donde,

V_{arc} Tensión de arco rms [V].

Una vez se tiene el valor de la corriente de arco y la tensión de arco, se pueden calcular resultados de interés con las siguientes ecuaciones.

$$R_{arc} = \frac{2 \times \sqrt{2} \times U_a}{\pi \times I_{arc}} = \frac{V_{arc}}{I_{arc}} \quad (7)$$

$$P_{arc} = V_{arc} \times I_{arc} \quad (8)$$

$$E_{arc} = V_{arc} \times I_{arc} \times T_{arc} \quad (9)$$

Donde,

- R_{arc} Resistencia de arco [Ω];
- P_{arc} Potencia de arco [W];
- E_{arc} Energía de arco [J];
- T_{arc} Tiempo de duración de arco [s].

Para calcular la energía incidente en aire libre se debe usar la ecuación (10) tomada de [19], considerando que el calor generado por el arco eléctrico tiende a propagarse de una forma esférica o cilíndrica cuando la distancia entre electrodos es corta o larga, respectivamente. Para tener en cuenta ambos casos, se utiliza un coeficiente de distancia x que variará dependiendo de la distancia entre electrodos (gap). El valor de x podría considerarse de 2 asumiendo que el medio plasmático del arco tiene forma esférica. Sin embargo, ante la incertidumbre en el resultado con gaps largos, en los que el calor del arco eléctrico se distribuya con una forma más semejante a la de un cilindro, es necesario utilizar datos conservativos de x .

$$E = \frac{E_{arc}}{4 \times \pi \times d^x} \quad (10)$$

Donde,

- E Energía incidente [J/mm²];
- E_{arc} Energía de arco [J];
- d Distancia de trabajo [mm];
- x Coeficiente de exponente de distancia.

Por su parte, en el artículo *High Voltage Arc Flash Assessment and Applications - Part 1* [19] se plantea el uso de la ecuación (11) para calcular la energía incidente en encerramientos. Para ello, es necesario considerar los coeficientes de reflectividad de Wilkins [20] presentados en [2] y que dependen de las dimensiones físicas (altura, anchura y profundidad) del encerramiento. Adicionalmente, se tiene en cuenta el coeficiente de exponente de distancia x .

$$E = k \times \frac{E_{arc}}{a^2 + d^x} \quad (11)$$

Donde,

- E Energía incidente [J/mm²];
- E_{arc} Energía de arco [J];
- d Distancia de trabajo [mm];

x	Coeficiente de exponente de distancia;
k	Coeficiente “k” de reflectividad de Wilkins;
a	Coeficiente “a” de reflectividad de Wilkins [mm].

Los valores de “a” y “K” se extraen de [2].

Por último, la energía incidente se expresa convencionalmente en calorías por centímetro al cuadrado [cal/cm²], unidad sobre la cual se basan las categorías de riesgo por arco eléctrico establecidas en diversas normativas. Cuando los resultados se obtienen en julios por milímetro cuadrado [J/mm²], es necesario realizar una conversión para expresar los valores en [cal/cm²], utilizando la siguiente equivalencia:

$$E_{cal} = E \times 23.9006 \quad (12)$$

Donde,

E_{cal}	Energía incidente [cal/cm ²];
E	Energía incidente [J/mm ²].

Rangos de aplicación

Los rangos de aplicación se establecen a partir de las limitaciones y la información referenciada de la sección 41 y 44 de la NESC C2 2023 [21].

- Tensión nominal L-L de 1 kV a 800 kV.
- Distancia de trabajo mayor a 381 mm.
- Distancia entre electrodos (gaps) mayores a 50.8 mm.

2.2.4 Método a partir de ecuaciones de EPRI

Este método utiliza como bases teóricas los estudios presentados en el Technical Report de EPRI [15], en el cual se presentan resultados variados de testeos con arcos largos, además de presentar ecuaciones empíricas fundamentadas en los resultados obtenidos.

Metodología de cálculo

En el capítulo 8 de [15] se presenta un resumen de ecuaciones derivadas a partir de pruebas que se realizaron para arcos con espacios largos que varían en longitud de 1 pie (304.8 mm) a 5 pies (1524 mm). Estas ecuaciones se basan de algunos supuestos que simplifican su desarrollo para el uso en el cálculo del flujo y la energía térmica incidente.

El supuesto más importante utilizado en la simplificación de las ecuaciones, de acuerdo con lo especificado en el capítulo 8 del Technical Report de EPRI [15], es considerar que un gradiente de tensión promedio en la región del arco es suficiente para desarrollar modelos de arco razonables. Teniendo en cuenta la consideración anterior, en la ecuación (13) se establece el cálculo del gradiente de tensión promedio.

$$E_{ave} = 0.0000112G^{-8} + 1.19 + (0.0069G^{-1.239} - 0.0126)I_{arc} \quad (13)$$

Donde,

E_{ave} Gradiente de arco promedio [kV_{pico}/m];
 G Longitud del gap entre electrodos [m];
 I_{arc} Corriente de arco [kA_{rms}].

A partir del cálculo de E_{ave} , se hace uso de la ecuación (14) para calcular la tensión de arco en valores rms.

$$V_{arc} = \frac{E_{ave} \times G}{1.1255} = 0.8885 \times E_{ave} \times G \quad (14)$$

Donde,

V_{arc} Tensión de arco [kV_{rms}].

La potencia de arco se puede calcular así:

$$P_{arc} = V_{arc} \times I_{arc} \quad (15)$$

Donde,

P_{arc} Potencia de arco [MW].

La ecuación simplificada para el cálculo del flujo de energía térmica incidente es:

$$\Phi = 6.7 \times E_{ave} \times I_{arc} \times G^{0.58} \times D^{-1.58G^{-0.152}} \quad (16)$$

Donde,

Φ Flujo de energía térmica incidente [cal/(s*cm²)];
 D Distancia de trabajo [pies].

Para D en metros, la ecuación sería:

$$\Phi = 6.7 \times E_{ave} \times I_{arc} \times G^{0.58} \times (3.2810 \times D)^{-1.58G^{-0.152}} \quad (17)$$

Donde,

Φ Flujo de energía térmica incidente [cal/(s*cm²)];
 D Distancia de trabajo [m].

Multiplicando el flujo de energía térmica incidente con el tiempo de duración del arco se calcula la energía térmica incidente, tal como se muestra en la ecuación (18).

$$W = \Phi \times T \quad (18)$$

Donde,

W Energía incidente térmica [cal/cm²];
 T Duración de la exposición al arco [s].

Como las ecuaciones derivadas se basan en resultados promedios, en el capítulo 8 de [15], se plantea un factor de ajuste estadístico con el fin de calcular un valor de energía térmica incidente conservativo. Con la ecuación (19) se logra calcular este valor con el ajuste propuesto.

$$W = \Phi \times T \times k = \Phi \times T \times (1 + n \times \sigma) \quad (19)$$

Donde,

W Energía incidente térmica [cal/cm²];
 T Duración de la exposición al arco [s];

k	Factor de ajuste estadístico;
n	Factor de multiplicación estadístico;
σ	Desviación estándar.

La desviación estándar de la población de resultados en las pruebas es de $\sigma = 0.196$. El valor del multiplicador n se deja a criterio del usuario. Valores de $n = 2$ y $n = 3$ son mencionados en el documento como buenas opciones para su uso, ya que, con dichos valores, el número de datos que exceden el valor de energía térmica incidente calculado son de 10 y 4 respectivamente, frente a un total de 578 puntos de datos medidos [15].

Al igual que en la metodología de cálculo de Terzija/Koglin, una de las limitaciones que trae este conjunto de ecuaciones es el cálculo de la corriente de arco, puesto que su valor está sujeto al gradiente de tensión promedio. Es por ello por lo que Marroquin, Rehman y Madani realizan el cálculo de estos valores de forma iterativa [19].

Rangos de aplicación

Los rangos de aplicación se establecen a partir de las limitaciones y la información referenciada del Technical Report de EPRI y de la sección 41 y 44 de la NESC C2 2023 [15] [21].

- Tensión nominal L-L de 1 kV a 800 kV.
- Corriente de cortocircuito de 8 kA a 40 kA.
- Distancia de trabajo mayor a 381 mm.
- Distancia entre electrodos (gaps) entre 304.8 mm y 1524 mm.

2.3 Modelos de Machine Learning

Los modelos de Machine Learning ofrecen una capacidad avanzada para analizar patrones, relaciones y correlaciones en los datos, superando las limitaciones de los enfoques estadísticos tradicionales, que a menudo no pueden captar la complejidad de los conjuntos de datos de manera intuitiva o directa. Gracias a su capacidad de aprendizaje a partir de grandes volúmenes de información, estos modelos permiten descubrir estructuras subyacentes y tendencias ocultas que podrían pasar desapercibidas en análisis convencionales [22].

Además de facilitar una comprensión más profunda de los datos, el Machine Learning permite generar predicciones altamente precisas sobre nuevos datos o una clasificación adecuada de los mismos, incluso cuando estos no han sido previamente observados en el proceso de entrenamiento. Esto lo convierte en una herramienta eficaz para la toma de decisiones basada en evidencia.

Cada modelo de Machine Learning posee características específicas que influyen en su capacidad para aprender y representar patrones en los datos. Esto significa que diferentes modelos pueden interpretar un mismo conjunto de datos de maneras distintas, logrando captar relaciones con distintos niveles de precisión. Mientras que algunos modelos pueden no ser lo suficientemente flexibles para identificar estructuras complejas, otros pueden detectar tendencias subyacentes con mayor eficacia. Por esta razón, es fundamental evaluar múltiples modelos y comparar su

desempeño para identificar aquellos que mejor se adapten a la naturaleza del problema. Este proceso implica la aplicación de técnicas como la validación cruzada y el ajuste de hiperparámetros para mejorar el rendimiento de los modelos seleccionados. El objetivo es fortalecerlos hasta minimizar los errores en la predicción o clasificación, logrando así resultados más precisos y confiables.

Algunas de las estrategias más comunes incluyen [22] [23]:

- **Regresión Lineal y Logística:** Son herramientas clásicas para modelar relaciones entre variables numéricas y predecir resultados categóricos. Métodos como Ridge y Lasso ayudan a mejorar la estabilidad y precisión del modelo al regularizar coeficientes y evitar el sobreajuste.
- **Series Temporales:** Modelos como ARIMA, Prophet o técnicas de suavizamiento exponencial permiten analizar datos secuenciales y hacer proyecciones futuras basadas en patrones históricos.
- **Árboles de Decisión y Modelos de Ensemble:** Algoritmos como Random Forest y XGBoost combinan múltiples árboles de decisión para mejorar la precisión y robustez en tareas de clasificación y regresión.
- **Redes Neuronales:** Cuando el problema requiere modelar relaciones altamente complejas y no lineales, frameworks como TensorFlow y PyTorch permiten construir arquitecturas avanzadas para el aprendizaje profundo.

2.3.1 Modelos de regresión aplicados

Regresión Lineal Regularizada mediante Elastic Net

La regresión lineal es uno de los modelos fundamentales en el análisis estadístico aplicado al Machine Learning. Su principal uso radica en el modelamiento de relaciones lineales, asignando a cada variable independiente un coeficiente (β_i) con el fin de estimar una variable de salida. El objetivo es determinar un valor óptimo para cada coeficiente a partir del conjunto de datos de entrenamiento, de modo que la salida pueda ser estimada con un bajo error [22][24].

Para mejorar la capacidad generalizadora del modelo y evitar el sobreajuste, se introducen técnicas de regularización, que añaden penalizaciones al proceso de ajuste de los coeficientes. Las dos técnicas más comunes son:

- **Lasso Regression (L1):** Penaliza la suma de los valores absolutos de los coeficientes. Estabiliza el modelo frente a variables correlacionadas y evita coeficientes grandes.
- **Ridge Regression (L2):** Penaliza la suma de los cuadrados de los coeficientes. Favorece la selección de variables al poder llevar coeficientes exactamente a cero.

La regresión Elastic Net es una extensión de la regresión lineal regularizada que combina las penalizaciones de Lasso y Ridge, donde cada una de estas tiene asignado un peso, lo que permite beneficiarse de las propiedades de ambas técnicas. Cuando se aplica Elastic Net en un modelo de

regresión lineal se debe ajustar la intensidad de la regularización y la proporción de penalización L1 y L2 [25][26][27].

Regresión Polinómica de Grado 2

La regresión polinómica es una extensión de la regresión lineal que permite modelar relaciones no lineales al incluir términos elevados a diferentes potencias de las variables predictoras. Al aumentar el grado del polinomio, el modelo puede capturar patrones más complejos en los datos. Sin embargo, elegir un grado demasiado alto puede llevar a un sobreajuste (overfitting), donde el modelo se ajusta demasiado a los datos de entrenamiento y pierde capacidad de generalización. Para que un modelo de regresión lineal pueda aprender propiedades polinómicas, se requiere que el conjunto de datos se le generen nuevas características al elevar las variables a la potencia especificada [22][24][25][28].

Redes Neuronales

Las redes neuronales son modelos inspirados en el funcionamiento de las neuronas biológicas. Estas redes utilizan variables de entrada para generar salidas a través de múltiples capas interconectadas constituidas de perceptrones (neuronas artificiales).

Según Kanal en [29] el componente básico de un perceptrón es un elemento que acepta un número de entradas x_i , $i = 1, \dots, N$, y calcula una suma ponderada de estas entradas donde, para cada entrada, su peso fijo ω puede ser solo +1 o -1. La suma se compara entonces con un umbral θ , y se produce una salida "y" que es 0 o 1, dependiendo de si la suma supera o no el umbral, en otras palabras:

$$y = \begin{cases} 1, & \left(\sum_{i=1}^n \omega_i \cdot x_i \right) \geq \theta \\ 0, & \left(\sum_{i=1}^n \omega_i \cdot x_i \right) < \theta \end{cases}$$

De acuerdo con Matich en [30], hay tres tipos de capas que representan un modelo de redes neuronales:

- **Capa de entrada:** En esta capa se recibe toda la información de entrada.
- **Capas ocultas:** Son capas internas a la red y su número de niveles ocultos se puede seleccionar de acuerdo con la complejidad de los datos. Asimismo, la interconexión entre capas puede ser definida de diferentes formas para establecer la topología de la red neuronal.
- **Capa de salida:** Son aquellas encargadas de transmitir los resultados obtenidos en las capas internas.

En la figura siguiente se presenta un esquema de una red neuronal:

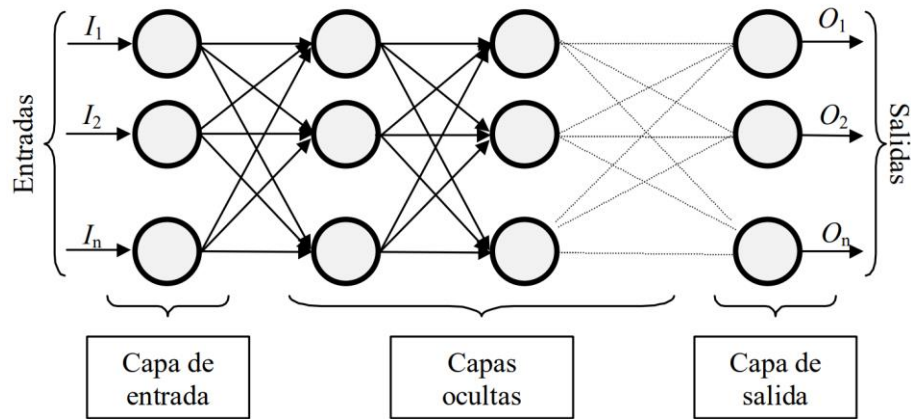


Figura 1. Ejemplo de red neuronal tomada de [30]

Cada neurona recibe las entradas y les asigna pesos y sesgos, realiza una suma ponderada y aplica una función no lineal, lo que permite modelar relaciones complejas en los datos. Las salidas generadas por cada neurona se transmiten a la siguiente capa como nuevas entradas, y este proceso continúa hasta la capa final, que produce el número de salidas requeridas. Para mejorar la precisión del modelo, el proceso de cálculo de las salidas se repite durante múltiples épocas e iteraciones, ajustando continuamente los pesos y sesgos mediante retroalimentación (retropropagación). Este proceso iterativo permite que el modelo aprenda de los datos y mejore su desempeño a lo largo del tiempo [22][24][30][31].

2.3.2 Escalado de datos

El escalado de datos es una etapa fundamental en la preparación de los datos para algoritmos de Machine Learning, especialmente cuando se utilizan modelos sensibles a las magnitudes relativas de las variables. El objetivo principal es normalizar el rango de las variables de entrada, de manera que todas contribuyan de forma equilibrada al proceso de aprendizaje del modelo. Esto previene que variables con valores numéricamente mayores dominen sobre otras de menor escala, lo cual podría sesgar el resultado del modelo o afectar su capacidad de generalización [32].

Una técnica comúnmente empleada para este propósito es el escalado estándar, también conocido como normalización por puntuaciones Z (Z-score). Este método transforma cada valor de una variable restando la media y dividiendo por la desviación estándar de dicha variable. La fórmula utilizada es la siguiente:

$$Z = \frac{x - \mu}{\sigma}$$

Donde \$x\$ es el valor original, \$\mu\$ representa la media y \$\sigma\$ la desviación estándar de la variable. Como resultado, la variable transformada tiene una media de 0 y una desviación estándar de 1, lo que garantiza que todos los atributos se comporten de manera homogénea ante el modelo [32][33].

2.3.3 Evaluación de los modelos de Machine Learning

La evaluación de los modelos de Machine Learning corresponde a un paso fundamental para determinar su desempeño y validar si cumplen con el propósito para el cual fueron desarrollados. Este proceso generalmente se realiza a partir de técnicas estadísticas, las cuales no solo permiten medir la precisión y la capacidad de generalización del modelo, sino que también facilita la identificación de posibles deficiencias, como el sobreajuste (overfitting) o el subajuste (underfitting). Además, la evaluación proporciona información valiosa para optimizar el modelo mediante ajustes en los parámetros, la selección de hiperparámetros o incluso modificaciones en su arquitectura. Esto permite mejorar su rendimiento frente a nuevos datos y aumentar su robustez ante variaciones en el conjunto de entrenamiento.

Métricas de evaluación para modelos de Machine Learning

Existen múltiples métricas para evaluar el desempeño de un modelo de Machine Learning, por lo que es fundamental seleccionar aquellas que mejor se ajusten al objetivo del modelo y a las características del problema. La elección de métricas adecuadas permite no solo medir la precisión del modelo, sino también identificar posibles sesgos, sobreajustes o deficiencias en la generalización.

Las métricas más comunes para la evaluación del desempeño de modelos de predicción son [23]:

1. **Error Absoluto Medio (MAE):** El MAE es una métrica que mide la diferencia promedio entre los valores predichos y los valores reales, sin considerar la dirección del error, que permite verificar que tan alejados están los valores entre sí.
2. **Raíz del Error Cuadrático Medio (RMSE):** Similar al MAE, el RMSE cuantifica la diferencia entre predicciones y valores reales, pero penaliza más los errores grandes al elevarlos al cuadrado antes de promediarlos. Esto hace que sea especialmente útil en casos donde se desea minimizar el impacto de grandes desviaciones en las predicciones.
3. **Coefficiente de Determinación (R^2):** El R^2 mide qué tan bien explica el modelo la variabilidad de los datos en comparación con la media de los valores reales. Un valor cercano a 1 indica que el modelo explica perfectamente los valores reales, mientras que valores cercanos a 0 o negativos sugieren un modelo ineficaz.

El uso conjunto de estas tres métricas permite evaluar adecuadamente el desempeño del modelo, ya que se valida tanto la magnitud de los errores como la capacidad del modelo para capturar la variabilidad de los datos en cada conjunto de evaluación.

3. METODOLOGÍA PROPUESTA

Para calcular la energía incidente en sistemas con niveles de tensión superiores a 15 kV en encerramientos, mediante metodologías alternativas, es necesario aplicar factores de ajuste. Estos factores permiten escalar los resultados obtenidos con metodologías basadas en arcos monofásicos al aire libre, hacia valores más conservadores que representen adecuadamente arcos trifásicos en encerramientos.

Con el fin de estimar dichos factores de ajuste, se propone una metodología para desarrollar un Modelo de Ajuste de Energía Incidente (MAEI). Este modelo, basado en técnicas de Machine Learning, permite predecir los factores de ajuste a partir de nuevos datos. Entre varios modelos evaluados, se selecciona aquel con el mejor desempeño predictivo. La metodología para obtener el MAEI se resume en la figura siguiente.

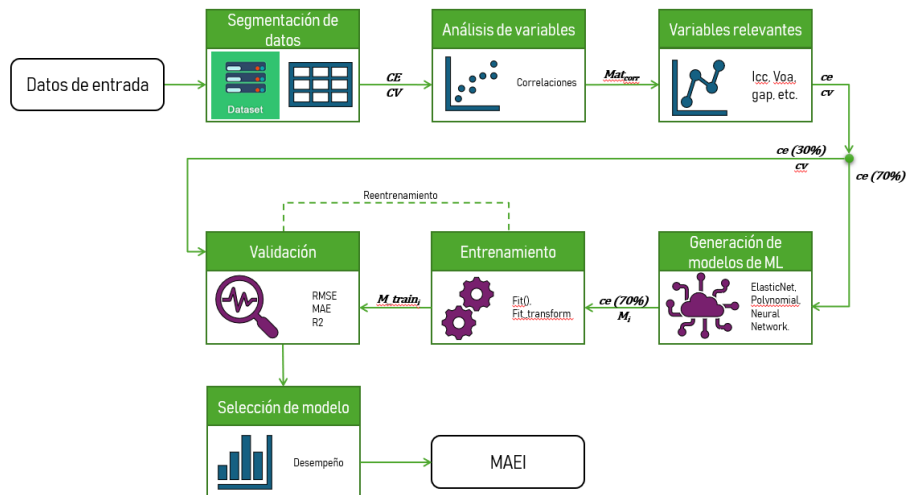


Figura 2. Proceso metodológico para obtener MAEI

Datos de entrada: Corresponde al conjunto de datos utilizado para la construcción de los modelos y se construye a partir del planteamiento de escenarios de análisis en subestaciones que operan con celdas tipo AIS, específicamente en niveles de tensión inferiores a 15 kV. Para ello, se generan casos con sensibilidades en las variables de entrada dependientes de la configuración de la celda o de labores de mantenimiento, tales como la distancia entre electrodos (gap), configuración de electrodos y distancia de trabajo.

Los casos son diseñados de forma que cubran un rango amplio de condiciones del sistema y de la configuración de las celdas, garantizando así la diversidad y representatividad del conjunto de datos. Esta estrategia permite que los modelos de Machine Learning sean entrenados y validados con información robusta que permita mejorar la capacidad de generalización del modelo resultante.

Segmentación de datos: En esta etapa, el conjunto de datos de entrada es organizado y segmentado con el propósito de estructurar dos grupos diferenciados: casos de entrenamiento (CE) y casos de validación (CV). Esta división permite alimentar adecuadamente los modelos de Machine Learning, asegurando que el proceso de entrenamiento se realice sobre un conjunto de datos suficientemente amplio y diverso, mientras que la validación se efectúa sobre un conjunto independiente, no utilizado previamente por el modelo.

Para ello, se diseñan dos tipos de escenarios:

- Escenarios orientados al entrenamiento y prueba, que permiten al modelo aprender patrones, relaciones y tendencias entre las variables de entrada y los factores de ajuste.
- Escenarios específicos de validación, destinados a evaluar la capacidad de generalización del modelo frente a nuevos datos, lo que resulta fundamental para garantizar su aplicación confiable en situaciones reales.

Los escenarios son definidos aplicando sensibilidades sistemáticas a las variables de entrada, lo cual permite explorar un amplio rango de combinaciones posibles. Como resultado de este proceso, se generan 13.920 casos para el entrenamiento y prueba del modelo, y 5.568 casos adicionales destinados exclusivamente para la validación.

Análisis de variables: En este bloque se busca identificar las variables de entrada que ejercen mayor influencia sobre los resultados de energía incidente y los factores de ajuste, realizando un análisis detallado utilizando herramientas estadísticas. Para ello, se emplean funciones de cálculo de correlaciones, como el coeficiente de Pearson, junto con su representación gráfica en un mapa de calor.

El resultado central de este análisis es la construcción de una matriz de correlación (denominada Mat_{corr}), la cual resume las relaciones lineales entre todas las variables involucradas. Esta matriz permite observar qué variables presentan una mayor asociación con la variable objetivo, así como detectar posibles redundancias entre variables predictoras.

El análisis de correlación no solo cumple un papel exploratorio, sino que también orienta la selección de características para los modelos de Machine Learning. Al conocer qué variables aportan información significativa y cuáles pueden inducir ruido o multicolinealidad, se optimiza el desempeño del modelo y se mejora su interpretabilidad.

Variables relevantes: En esta etapa se seleccionan las variables que presentan mayor influencia sobre los resultados de energía incidente y los factores de ajuste a partir del análisis de correlación y exploración estadística realizado en el bloque anterior. Esta selección se basa en criterios cuantitativos, como los coeficientes de correlación, y en criterios técnicos asociados al conocimiento del fenómeno físico subyacente.

Con las variables seleccionadas se conforman conjuntos de datos depurados que se destinan al entrenamiento (ce) y validación (cv) de los modelos de Machine Learning. Esta depuración contribuye a reducir la dimensionalidad del problema y mejora la eficiencia computacional.

Generación de modelos de ML: Para este bloque se realiza la selección e implementación de modelos de Machine Learning adecuados para la predicción de factores de ajuste en el cálculo de energía incidente en encerramientos. Dada la complejidad y variabilidad en las relaciones entre las variables de entrada y la variable objetivo, se exploran múltiples enfoques de modelado que permiten capturar tanto patrones lineales como no lineales, así como interacciones multivariadas.

La estrategia adoptada contempla la aplicación de diferentes algoritmos de regresión supervisada, tales como regresión lineal, regresión polinómica y redes neuronales artificiales. Esta diversidad de métodos tiene como objetivo identificar el modelo que mejor se acople a la estructura y distribución de los datos disponibles, maximizando la precisión y capacidad de generalización del modelo.

Adicionalmente, se lleva a cabo un ajuste inicial de parámetros e hiperparámetros para cada uno de los modelos seleccionados. Este ajuste permite optimizar el comportamiento del algoritmo, mejorando su rendimiento en tareas de predicción.

Entrenamiento: En esta etapa del proceso metodológico se realiza el entrenamiento de los modelos de Machine Learning utilizando el 70 % del conjunto de datos *ce* (casos de entrenamiento depurado). Durante el entrenamiento, cada modelo aplica su lógica interna para ajustar los parámetros necesarios que le permitan aprender los patrones y relaciones existentes entre las variables de entrada y la variable objetivo: el factor de ajuste.

Cada modelo ajusta sus parámetros internos de acuerdo con su estructura particular. Por ejemplo, los modelos de regresión estiman coeficientes que minimizan una función de error, aplicando regularización en el caso de ciertos enfoques para evitar sobreajuste. En modelos más complejos, como redes neuronales, los parámetros se ajustan de manera iterativa mediante algoritmos de optimización como el descenso de gradiente, utilizando funciones de activación y retropropagación para mejorar progresivamente la precisión de las predicciones.

Validación: En este bloque se evalúa el desempeño general de los modelos entrenados utilizando dos conjuntos de datos: el 30 % restante del conjunto de datos *ce*, reservado para prueba, y un conjunto adicional de casos de validación (*cv*), diseñados para representar condiciones distintas a las vistas durante el entrenamiento.

La validación permite determinar la capacidad de generalización de los modelos, es decir, su habilidad para realizar predicciones precisas sobre datos nuevos. Para cuantificar su desempeño, se emplean métricas estadísticas comúnmente utilizadas en regresión, tales como:

- **MAE (Mean Absolute Error):** Mide el promedio de los errores absolutos entre las predicciones y los valores reales.
- **RMSE (Root Mean Squared Error):** Penaliza los errores grandes al elevarlos al cuadrado antes de promediarlos, ofreciendo una visión más sensible a desviaciones importantes.
- **R² (Coeficiente de determinación):** Indica qué proporción de la variabilidad de la variable dependiente puede ser explicada por el modelo.

Si el modelo no alcanza un rendimiento satisfactorio según estas métricas, ya sea por errores elevados o baja capacidad explicativa, se procede a un proceso iterativo de ajuste de parámetros o hiperparámetros, seguido de un nuevo ciclo de entrenamiento. Este ajuste puede implicar la modificación de configuraciones internas del modelo, la reconsideración de ciertas variables o la aplicación de técnicas de regularización adicionales.

El objetivo final de esta etapa es identificar el modelo que mejor represente el comportamiento del factor de ajuste en distintos escenarios, con un equilibrio entre precisión, robustez y generalización.

Selección de modelo: Como último paso, se selecciona el modelo que muestra el menor error de predicción y el mayor grado de ajuste a los datos, sin evidenciar síntomas de sobreajuste (overfitting) ni pérdida de robustez ante condiciones distintas a las de entrenamiento. Este modelo será adoptado formalmente como el Modelo de Ajuste de Energía Incidente (MAEI), y se usará como herramienta para la predicción de factores de ajuste en el cálculo de energía incidente en celdas tipo AIS. Su implementación permitirá complementar o mejorar las metodologías alternativas, brindando una solución adaptable y basada en datos para distintos niveles de tensión y configuraciones de equipos.

3.1 Datos de entrada

En esta sección se presentan las variables seleccionadas como datos de entrada para la generación de los distintos casos de estudio. Estas variables han sido definidas en función de su influencia sobre el cálculo de la energía incidente considerando la aplicación de todas las metodologías de interés. Las variables planteadas se presentan a continuación:

1. Escenario.
2. Tensión nominal (V).
3. Corriente de cortocircuito (I_{bf}).
4. Tipo de falla.
5. Distancia de Trabajo (D).
6. Gap (G).
7. Configuración de electrodos (EC).
8. Coeficiente de encerramiento (Cenc).
9. Gradiente de tensión (B).
10. IEEE 1584 – 2018 Iarc.
11. IEEE 1584 – 2018 Earc.
12. EPRI Iarc.
13. EPRI Earc.
14. Terzija/Koglin Iarc.
15. Terzija/Koglin Earc.
16. Relación IEEE – EPRI Earc Aire Libre.

17. Relación IEEE – Terzija/Koglin Earc Aire Libre.

Cada una de estas variables ha sido definida con el propósito de capturar información esencial para el análisis, garantizando que los datos recopilados sean coherentes, representativos y adecuados para su procesamiento.

NOTA: Considerando que el tiempo de despeje de falla o de exposición al arco eléctrico tiene un comportamiento lineal en todas las metodologías de cálculo de energía incidente, se establece con un valor constante de 200 ms para efectos de cálculo.

En los siguientes subnumerales se establecen los escenarios implementados para el entrenamiento y validación del modelo, así como las sensibilidades que se realizaron en las variables correspondientes a la distancia de trabajo, distancia entre electrodos y configuración de electrodos. Además, se presentan definiciones sobre algunas variables de entrada y se plantean los métodos iterativos para el uso de las metodologías Terzija/Koglin.

3.1.1 Escenarios de entrenamiento y validación implementados

La red eléctrica tomada para la implementación de los escenarios de entrenamiento y validación corresponde a cuatro subestaciones eléctricas ubicadas en la zona CQR (Caldas – Quindío – Risaralda) del sistema eléctrico colombiano a nivel de 33 kV y 13.2 kV. Estas subestaciones son denominadas Subestación A, Subestación B, Subestación C y Subestación D.

En particular, la subestación A cuenta con un transformador bidevanado de 33/13.2 kV, mientras que las subestaciones B, C y D disponen de transformadores tridevanados de 33/13.2/4.16 kV. La presencia de un terciario a 4.16 kV en estas últimas subestaciones introduce un nuevo nivel de tensión que es incorporado como un escenario adicional de análisis, lo que permite ampliar el rango de condiciones de operación evaluadas, para observar cómo varía la energía incidente en función de la tensión nominal del sistema. El diagrama unifilar, los parámetros eléctricos y los datos de la Red Equivalente de cada subestación se muestran en el Anexo A.

La corriente de cortocircuito de falla atornillada para los casos se toma desde el modelo de red esquematizado en PowerFactory. Este valor será el mayor entre la falla trifásica, monofásica y bifásica a tierra. En la siguiente tabla se presentan los escenarios con sus respectivas corrientes de cortocircuito máximas respecto a cada subestación y nivel de tensión a evaluar.

Tabla 1. Escenarios implementados para el análisis

Escenario	Subestación	Nivel de tensión [kV]	Corriente de cortocircuito [kA]	Tipo de falla
Escenario 1	Subestación A	13.2	5.65	Bifásica a tierra
Escenario 2	Subestación B	13.2	20.98	Bifásica a tierra
Escenario 3	Subestación C	13.2	11.80	Bifásica a tierra
Escenario 4	Subestación B	4.16	24.75	Trifásica
Escenario 5	Subestación C	4.16	14.51	Trifásica
Escenario 6	Subestación D	13.2	20.91	Bifásica a tierra
Escenario 7	Subestación D	4.16	24.49	Trifásica

3.1.2 Sensibilidades de variables

En el cálculo de la energía incidente, variables como la configuración de electrodos, la distancia entre electrodos (gap) y la distancia de trabajo tienen una influencia significativa y están directamente relacionadas con el diseño físico de la celda y las condiciones de operación durante tareas de mantenimiento.

La configuración de electrodos depende de cómo se disponen los conductores y superficies conductoras dentro del encerramiento, lo cual está directamente determinado por el diseño de la celda AIS. Por este motivo, se incluyen todas las configuraciones de electrodos en encerramientos especificadas por la IEEE Std. 1584-2018 (VCB, VCBB y HCB), con el fin de representar fielmente las condiciones reales que podrían encontrarse en distintas subestaciones.

La distancia entre electrodos (gap) se refiere al espacio libre entre conductores que podrían originar un arco eléctrico, y varía según el diseño interno de los equipos, la disposición de los componentes y las condiciones de mantenimiento. Por su parte, la distancia de trabajo puede cambiar según las posturas, herramientas o posiciones adoptadas por el personal durante las maniobras de intervención, lo cual introduce variabilidad adicional que debe ser considerada en los análisis.

Con el fin de analizar la sensibilidad de estas variables en los resultados, se establece un rango amplio y representativo para el gap y la distancia de trabajo: de 50 mm a 1600 mm para el gap, y de 200 mm a 1600 mm para la distancia de trabajo, en ambos casos con incrementos de 50 mm por intervalo. Esta estrategia permite observar el impacto progresivo de estas variables tanto en los valores de energía incidente estimados por las distintas metodologías (IEEE, EPRI, Terzija/Koglin) como en el desarrollo de los modelos de Machine Learning para predecir factores de ajuste.

A partir de la variación propuesta en las variables mencionadas, se construye un conjunto amplio de combinaciones posibles. Para cada escenario planteado, esta combinación de variables da lugar a un total de 2.784 casos únicos.

3.1.3 Coeficiente de encerramiento y gradiente de tensión B

Coeficiente de encerramiento

Como se había mencionado en el numeral 2.2.3, en [19] se plantea el uso de la ecuación (11) para calcular la energía incidente en encerramientos. Esta ecuación ayuda a establecer un factor de multiplicación en encerramientos, que se mencionará como Coeficiente de encerramiento (C_{enc}) para objeto del análisis, el cual se calcula a partir de la ecuación (20) y corresponde a la relación entre la ecuación (11) y ecuación (10). Este coeficiente se puede usar fácilmente para calcular el valor de la energía incidente en encerramientos a partir del calculado en aire libre para las metodologías alternativas Terzija/Koglin y EPRI.

$$C_{enc} = \frac{4 \times \pi \times k}{1 + \frac{a^2}{d^x}} \quad (20)$$

Donde,

C_{enc}	Coeficiente de encerramiento;
d	Distancia de trabajo [mm];
x	Coeficiente de exponente de distancia;
k	Coeficiente “k” de reflectividad de Wilkins;
a	Coeficiente “a” de reflectividad de Wilkins [mm].

Gradiente de tensión B

Para el desarrollo del análisis de esta tesis se plantea definir el valor de B en 1300 V/m para corrientes de arco menores a 10 kA y 1500 V/m para corrientes de arco iguales o mayores a 10 kA. Estos valores se establecen a partir de los resultados presentados en [34].

3.1.4 Variables calculadas

Las variables calculadas que complementan el conjunto de datos de entrada corresponden a:

- IEEE1584 – 2018 Iarc: Se calcula usando los métodos expuestos en la IEEE Std. 1584 – 2018.
- IEEE1584 – 2018 Earc: Se calcula usando los métodos expuestos en la IEEE Std. 1584 – 2018.
- EPRI Iarc: Se debe emplear un método iterativo para calcularla a partir de la corriente de falla.
- EPRI Earc: Se usa la ecuación (19) para su cálculo.
- Terzija/Koglin Iarc: Se debe emplear un método iterativo para calcularla a partir de la corriente de falla.
- Terzija/Koglin Earc: Se usa la ecuación (12) para su cálculo y se debe emplear un método iterativo.
- Relación IEEE – EPRI Earc Aire Libre (Factor de ajuste EPRI): Cociente entre el valor de energía incidente calculado según IEEE Std. 1584-2018 y el valor obtenido mediante la metodología EPRI.
- Relación IEEE – Terzija/Koglin Earc Aire Libre (Factor de ajuste Terzija/Koglin): Cociente entre el valor de energía incidente calculado según IEEE Std. 1584-2018 y el valor obtenido mediante la metodología Terzija/Koglin.

3.1.5 Métodos iterativos propuestos

En este apartado se proponen los siguientes métodos iterativos para hacer uso de las metodologías de cálculo de energía incidente de Terzija/Koglin y EPRI.

Terzija/Koglin

En el método de Terzija/Koglin es necesario usar iteraciones a partir de la resistencia de falla y corriente de falla, utilizando estos valores en las ecuaciones (4), (5) y (6) con el fin de encontrar el valor de tensión de arco, corriente de arco y resistencia de arco. Los pasos establecidos para el método iterativo son los siguientes:

1. Utilizando una herramienta o programa especializado para cálculo de corrientes de cortocircuito, se extraen los valores de impedancias de secuencia asociadas al nodo o barra de interés.

2. Se define el tipo de falla (Trifásica, monofásica o bifásica a tierra) que se debe calcular durante el proceso iterativo según se considere para el análisis.
3. Se inicializa el proceso iterativo con una resistencia de falla igual a cero ($R_f = 0$), y luego se calcula la corriente de falla (I_f) siguiendo el método IEC 60909 [35].
4. Se multiplica la resistencia de falla (R_f) por la corriente de falla (I_f) para encontrar el valor de la tensión de falla (V_f).
5. A partir de la ecuación (5) se calcula el gradiente de arco (E_a) usando el valor de la corriente de falla (I_f) y el gradiente de tensión (B).
6. Haciendo uso de la ecuación (4) se calcula la magnitud de la tensión de arco (U_a) a partir del gradiente de arco (E_a) y el gap (G).
7. Con la ecuación (6) se calcula la tensión de arco rms (V_{arc}) a partir de la magnitud de la tensión de arco (U_a).
8. Se compara la tensión de falla (V_f) obtenida en el **paso 4** con la tensión de arco rms (V_{arc}) del **paso 7** para determinar si el valor calculado de la corriente de falla tiene un comportamiento teórico semejante al de la corriente de arco.
9. En cada iteración se debe cambiar el valor de la resistencia de falla (R_f) para obtener un nuevo valor calculado de corriente de falla (I_f), con el fin de encontrar una magnitud de corriente de arco que satisfaga las ecuaciones establecidas en la metodología de Terzija/Koglin.

Criterio de parada

Teniendo en cuenta que el objetivo es encontrar un valor de corriente de falla que genere valores de tensión de falla y de tensión de arco similares, un criterio de parada adecuado es establecer un delta entre ambos valores, que se calcula con el valor absoluto de la diferencia de las tensiones calculadas. Es decir:

$$\Delta = |V_f - V_{arc}| \quad (21)$$

Una vez el valor de Delta sea lo suficientemente bajo, es posible parar el proceso iterativo, especificando al final los siguientes valores:

Corriente de arco (I_{arc}): Se toma la última corriente de falla del proceso iterativo.

Gradiente de arco (E_a): Se toma el último gradiente de arco E_a del proceso iterativo.

Tensión de arco (V_{arc}): Se toma la última tensión de arco rms (V_{arc}) del proceso iterativo.

Con las ecuaciones (7), (8) y (9) se calcula el valor de resistencia de arco, potencia de arco y energía de arco. Por último, se calcularía la energía incidente al aire libre a partir de las ecuaciones (10) y (12).

EPRI

En el método de EPRI es necesario usar iteraciones a partir de la resistencia de falla y corriente de falla, utilizando estos valores en las ecuaciones (13) y (14) con el fin de encontrar el valor de tensión de arco y corriente de arco. Los pasos establecidos para el método iterativo son los siguientes:

1. Utilizando una herramienta o programa especializado para cálculo de corrientes de cortocircuito, se extraen los valores de impedancias de secuencia asociadas al nodo o barra de interés.
2. Se define el tipo de falla (Trifásica, monofásica o bifásica a tierra) que se debe calcular durante el proceso iterativo según se considere para el análisis.
3. Se inicializa el proceso iterativo con una resistencia de falla igual a cero ($R_f = 0$), y luego se calcula la corriente de falla (I_f) siguiendo el método IEC 60909 [35].
4. Se multiplica la resistencia de falla (R_f) por la corriente de falla (I_f) para encontrar el valor de la tensión de falla (V_f).
5. A partir de la ecuación (13) se calcula el gradiente de arco promedio (E_{ave}) usando el valor de la corriente de falla (I_f) y el gap (G).
6. Haciendo uso de la ecuación (14) se calcula la magnitud de la tensión de arco rms (V_{arc}) a partir del gradiente de arco promedio (E_{ave}) y el gap (G).
7. Se compara la tensión de falla (V_f) obtenida en el **paso 4** con la tensión de arco rms (V_{arc}) del **paso 6** para determinar si el valor calculado de la corriente de falla tiene un comportamiento teórico semejante al de la corriente de arco.
8. En cada iteración se debe cambiar el valor de la resistencia de falla (R_f) para obtener un nuevo valor calculado de corriente de falla (I_f), con el fin de encontrar una magnitud de corriente de arco que satisfaga las ecuaciones establecidas en la metodología de EPRI.

Criterio de parada

Al igual que en la metodología de Terzija/Koglin, en EPRI el objetivo también consiste en encontrar un valor de corriente de falla que genere valores de tensión de falla y de tensión de arco similares, por lo cual, el criterio de parada se establece con el delta de las tensiones usando la ecuación (21). Una vez el valor de Delta sea lo suficientemente bajo, es posible parar el proceso iterativo, especificando al final los siguientes valores:

Corriente de arco (I_{arc}): Se toma la última corriente de falla del proceso iterativo.

Gradiente de arco promedio (E_{ave}): Se toma el último gradiente de arco E_{ave} del proceso iterativo.

Tensión de arco (V_{arc}): Se toma la última tensión de arco rms (V_{arc}) del proceso iterativo.

Con la ecuación (16) o (17) se halla el valor del flujo de energía térmico (Φ) y luego se calcularía la energía incidente al aire libre a partir de la ecuación (19).

3.1.6 Definiciones adicionales

En este numeral se definen los valores y consideraciones de las variables que son usadas en el cálculo de los datos principales para el análisis.

Impedancias de secuencia

Para la aplicación del cálculo de corrientes de cortocircuito, es necesario hacer uso de las impedancias de secuencia del caso al que se le requiera realizar el análisis de energía incidente. Por lo tanto, en la siguiente tabla se consignan los valores de impedancias de secuencia de PowerFactory de los escenarios planteados en la Tabla 1.

Tabla 2. Impedancias de secuencia de cada escenario

Escenario	Subestación	Nivel de tensión [kV]	Z0 [Ω]	Z1 [Ω]	Z2 [Ω]
Escenario 1	Subestación A	13.2	0 + 1.1412j	0.5267 + 2.1407j	0.5267 + 2.1407j
Escenario 2	Subestación B	13.2	0.0057 + 0.2384j	0.0759 + 0.7188j	0.0759 + 0.7188j
Escenario 3	Subestación C	13.2	0.0222 + 0.4800j	0.1608 + 1.1617j	0.1608 + 1.1617j
Escenario 4	Subestación B	4.16	50749559.5706 + 3.6093j	0.0044 + 0.1066j	0.0044 + 0.1066j
Escenario 5	Subestación C	4.16	74916013.4476 - 0.1357j	0.0115 + 0.1817j	0.0115 + 0.1817j
Escenario 6	Subestación D	13.2	0.0026 + 0.2349j	0.0692 + 0.7305j	0.0692 + 0.7305j
Escenario 7	Subestación D	4.16	50749556.9700 + 7.2875j	0.0037 + 0.1078j	0.0037 + 0.1078j

Factor de ajuste estadístico k

El factor de ajuste estadístico “k” es mencionado en el Technical Report de EPRI [15] como un factor multiplicativo igual a $1 + (n \times \sigma)$ con el cual se puede calcular la energía térmica de manera conservativa, buscando alcanzar los mayores valores encontrados en los tests realizados por EPRI. Considerando la necesidad de obtener valores conservativos durante el análisis, se plantea usar una desviación estándar $\sigma = 0.196$ y un valor de $n = 3$, partiendo de lo mencionado en el numeral 2.2.4.

Dimensiones de encerramiento para metodología IEEE Std. 1584 – 2018

La metodología presentada en la IEEE Std.1584-2018 realiza cálculos de energía incidente para encerramientos, en donde las dimensiones de estos tienen un impacto en el resultado final. Con el fin de mantener un análisis estándar, se establecen los valores de Height (H), Width (W) y Depth (Dp), correspondientes al alto, ancho y profundidad del encerramiento respectivamente, de acuerdo con [7]. Los valores seleccionados estarán sujetos al nivel de tensión en el cual se realice el cálculo. El tipo de equipo se seleccionará así:

- Si la tensión es menor o igual a 1 kV el Tipo de Equipo o Equipment Type (ET) será “Low-voltage switchgear”.
- Si la tensión es mayor que 1 kV y menor o igual a 5 kV el ET será igual a “5 kV switchgear”.
- Si la tensión es mayor que 5 kV y menor o igual a 15 kV el ET será “15 kV switchgear”.

Coeficiente de exponente de distancia x

El coeficiente de exponente de distancia varía respecto a la forma en que se desarrolle el arco eléctrico. Para gaps largos la forma tiende a ser cada vez más cilíndrica, como es el caso de fallas en alta tensión en donde la separación entre posibles electrodos es considerable [19]. Para celdas de media tensión en donde el entorno es cerrado y las distancias entre los equipos eléctricos son próximas, el gap se puede considerar pequeño, por ende, para la aplicación de las ecuaciones (10) y (11) se plantea hacer uso de un coeficiente de exponente de distancia igual 2 ($x=2$), puesto que se espera que la nube de plasma del arco adopte una forma más esférica.

3.2 Segmentación de casos

La segmentación de datos se realiza a través de la división de los datos de entrada, buscando generar conjuntos estructurados enfocados al entrenamiento, prueba y validación de los modelos de Machine Learning. Para ello, se crean de dos conjuntos de datos correspondientes a los casos que se usarán para el entrenamiento y validación de los modelos:

- **Conjunto de datos de entrenamiento (CE):** Este conjunto de datos estará compuesto por los primeros cinco escenarios establecidos en la Tabla 1, correspondientes a los Escenarios de análisis asociados a la Subestación A, Subestación B y Subestación C.
- **Conjunto de datos de validación (CV):** En este segundo conjunto de datos se usarán los dos Escenarios restantes expuestos en la Tabla 1, es decir, el Escenario 6 y Escenario 7 asociados a la Subestación D.

Tanto el conjunto de datos de entrenamiento como el de validación han sido organizados bajo una estructura tabular, en la cual cada variable identificada como dato de entrada se representa como una columna, obteniéndose un total de 17 columnas por conjunto de datos.

El conjunto de datos de entrenamiento, derivado de la combinación de escenarios y sensibilidades de variables, cuenta con un total de 13.920 registros (filas). Por su parte, el conjunto de datos de validación contiene 5.568 registros, generados a partir de los escenarios faltantes. Cada fila representa un caso único y completo que será utilizado para el aprendizaje o evaluación de los modelos de Machine Learning.

3.3 Análisis de variables

Durante esta fase se lleva a cabo un análisis general de los conjuntos de datos para visualizar sus comportamientos y obtener estadísticas descriptivas asociadas a cada conjunto de datos, así como, identificar valores nulos, los cuales se eliminarán para evitar errores de precisión en el modelo. Posteriormente, se profundiza en el análisis de las correlaciones entre variables con el objetivo de evaluar su influencia en los factores de ajuste.

3.3.1 Análisis general de los conjuntos de datos

Con el objetivo de garantizar la integridad, consistencia y adecuada estructuración de los datos generados, se realiza un análisis exploratorio mediante herramientas estadísticas implementadas en Python. Este proceso permite validar la conformación de los conjuntos de datos, evaluar la distribución de las variables, y detectar posibles inconsistencias o valores atípicos que puedan afectar el desarrollo del modelo. La estadísticas descriptivas son:

- **Encabezado:** Muestra las primeras filas del conjunto de datos, en donde se logra apreciar la estructura de los conjuntos de datos, verificando que estén contenidas todas las variables.
- **Descripción:** Muestra estadísticas descriptivas del conjunto de datos, tales como:
 - count: Conteo de valores no nulos.
 - mean: Media aritmética.
 - Std: Desviación estándar.
 - Min: Valor mínimo.
 - 25%: Primer cuartil.
 - 50%: Mediana.
 - 75%: Tercer cuartil.
 - Max: Valor máximo.
- **Información:** Muestra información general sobre el conjunto de datos, incluyendo la cantidad de valores no nulos y el tipo de dato (int64, float64, object, etc.).
- **Valores Nulos:** Muestra la cantidad de valores nulo por columna.
- **Distribución de atributos:** Genera histogramas de los atributos numéricos contenidos en el DataFrame.

Con estas medidas se puede observar el comportamiento general de los conjuntos de datos, identificando valores nulos, dispersión de los datos, valor medio del conjunto de datos, separación de los datos y valores mínimos y máximos.

Las variables de mayor interés corresponden a el Coeficiente de encerramiento (Cenc), la Relación IEEE – EPRI Earc Aire Libre (Factor de ajuste EPRI) y la Relación IEEE – Terzija Earc Aire Libre (Factor de ajuste Terzija/Koglin), las cuales tienen las siguientes estadísticas descriptivas:

Tabla 3. Estadísticas descriptivas conjunto de datos entrenamiento

Descripción	Coeficiente de encerramiento (Cenc)	Relación IEEE - EPRI Earc Aire Libre	Relación IEEE - Terzija Earc Aire Libre
count	13920	12789	13920
mean	2.54	1.28	1.63
std	1.08	0.53	2.00
min	0.22	0.05	0.16

Descripción	Coficiente de encerramiento (Cenc)	Relación IEEE - EPRI Earc Aire Libre	Relación IEEE - Terzija Earc Aire Libre
25%	1.70	0.81	0.68
50%	2.87	1.26	1.01
75%	3.50	1.63	1.64
max	3.87	3.00	19.13

Tabla 4. Estadísticas descriptivas conjunto de datos validación

Descripción	Coficiente de encerramiento (Cenc)	Relación IEEE - EPRI Earc Aire Libre	Relación IEEE - Terzija Earc Aire Libre
count	5568	5133	5568
mean	2.61	1.30	1.60
std	1.05	0.54	1.97
min	0.22	0.05	0.16
25%	1.84	0.79	0.67
50%	2.97	1.33	1.01
75%	3.50	1.68	1.61
max	3.87	2.82	18.56

De las estadísticas descriptivas presentadas en la Tabla 3 y Tabla 4, los parámetros más relevantes son count, mean, std, min y max.

En la variable Coeficiente de Encerramiento (Cenc), se registra un valor mínimo de 0.22, lo cual resulta inconsistente con el propósito de este coeficiente, ya que se espera que la energía incidente en un entorno confinado sea mayor en comparación con la obtenida en condiciones de aire libre. Al analizar los conjuntos de datos, se identifican múltiples valores inferiores a 1, lo que se atribuye al impacto del nivel de tensión en la selección de los coeficientes de Wilkins y, principalmente, a la influencia de la distancia de trabajo. En particular, se observa que el coeficiente de encerramiento toma valores por debajo de 1 en las siguientes condiciones:

- Niveles de tensión inferiores a 5 kV con distancias de trabajo menores o iguales a 200 mm.
- Niveles de tensión superiores a 5 kV con distancias de trabajo menores o iguales a 450 mm.

Para la variable Relación IEEE - EPRI Earc Aire Libre, se observa una menor cantidad de valores registrados en comparación con la variable Relación IEEE - Terzija Earc Aire Libre. Esto sugiere que, en algunas iteraciones, no se obtuvo un resultado dentro del número máximo de iteraciones establecidas, lo que impidió el cálculo de la corriente de arco y energía incidente. Además, destaca el valor mínimo de 0.05, lo que indica que, en ciertos casos, la energía incidente calculada mediante la metodología EPRI es significativamente mayor que la obtenida con la IEEE Std. 1584 - 2018.

Por otro lado, la variable Relación IEEE - Terzija Earc Aire Libre presenta un valor mínimo de 0.16, evidenciando un comportamiento similar al observado con la metodología EPRI. Sin embargo, el valor máximo obtenido con la metodología de Terzija/Koglin es 19.13, lo que, al compararse con la media aritmética, el tercer cuartil y la desviación estándar, sugiere que existen uno o varios valores atípicos. Estos valores extremos indican que, en ciertos escenarios, la energía incidente calculada con la IEEE Std. 1584 - 2018 supera considerablemente la obtenida con Terzija/Koglin.

Al analizar las tablas en detalle, se observa que, para gaps reducidos de 50 mm, 100 mm y 150 mm, la variable Relación IEEE - Terzija Earc Aire Libre alcanza valores significativamente más altos. Esto evidencia una limitación importante de la metodología Terzija/Koglin, ya que su precisión se ve afectada cuando la distancia entre electrodos es pequeña.

Información y valores nulos

A través de los módulos de información y valores nulos, se puede observar que únicamente las variables EPRI Iarc [kA], EPRI Earc Aire Libre [cal/cm²] y Relación IEEE - EPRI Earc Aire Libre contienen valores nulos. Esto refleja la ineficacia del proceso iterativo para obtener soluciones en todas las variaciones planteadas dentro de la metodología de análisis.

Sin embargo, al analizar en detalle los conjuntos de datos de entrenamiento y validación, se detecta que los valores nulos se presentan exclusivamente en gaps pequeños de 50 mm, 100 mm y 150 mm. Esta situación puede deberse a que la metodología EPRI fue desarrollada en base a pruebas realizadas para gaps largos, mayores a 1 pie (304.8 mm).

Distribución de atributos

En la Figura 3 y Figura 4 se presenta la distribución de atributos de los conjuntos de datos de entrenamiento y validación respectivamente. Para cada variable del conjunto de datos se refleja la cantidad de veces que se repite un valor haciendo usos de histogramas, de esta forma es fácil observar cómo es la distribución de los valores que fueron calculados.

En general, se observa que la mayor densidad de atributos para las variables de interés se distribuye en los siguientes rangos:

- Coeficiente de Encerramiento (Cenc): entre 3 y 4.
- Relación IEEE - EPRI Earc Aire Libre: entre 0.5 y 1.5.
- Relación IEEE - Terzija Earc Aire Libre: entre 0.5 y 2.

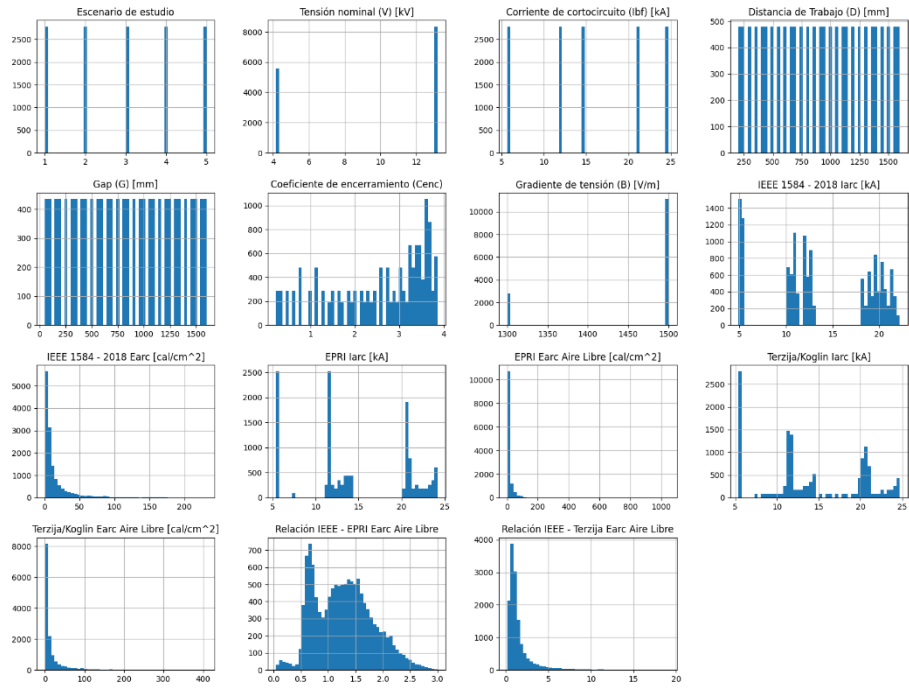


Figura 3. Distribución de atributos del conjunto de datos entrenamiento

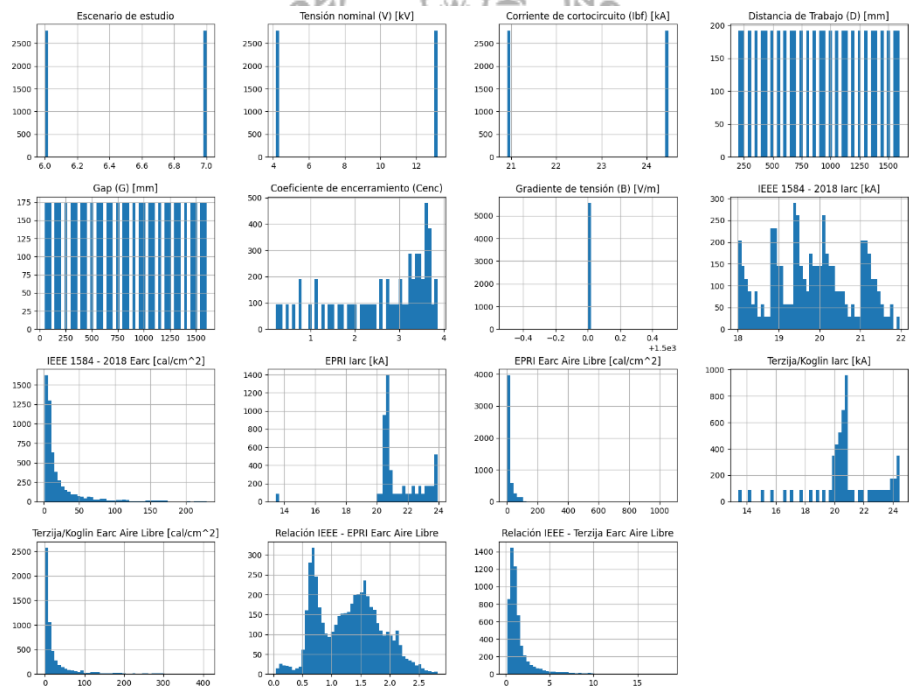


Figura 4. Distribución de atributos del conjunto de datos validación

Limpieza de datos

Considerando la distribución de atributos y los valores nulos encontrados, se decide eliminar las filas correspondientes a gaps menores a 200 mm. Esta decisión responde a dos aspectos clave:

1. **Eliminación de valores atípicos:** En el cálculo de variables mediante la metodología Terzija/Koglin, se identificaron valores elevados que afectan la estabilidad del análisis y podrían generar ruido a los modelos de Machine Learning y a las conclusiones que se extraigan del análisis.
2. **Manejo de valores nulos:** En la metodología EPRI, la presencia de valores nulos en estos rangos de gaps sugiere que el proceso iterativo no logró converger en estas condiciones, lo que compromete la integridad del conjunto de datos para el análisis comparativo y el entrenamiento de los modelos.

Al aplicar este filtro, se garantiza un conjunto de datos más representativo y coherente para el análisis, evitando sesgos y mejorando la calidad del entrenamiento de los modelos de Machine Learning. La cantidad de casos resultantes después de eliminar los valores atípicos y nulos es de 12.615 para el conjunto de datos de entrenamiento y 5.046 para el conjunto de datos de validación.

3.3.2 Análisis de las correlaciones entre variables

Con el fin de identificar las variables con mayor influencia en el cálculo de la relación entre la energía incidente estimada mediante la metodología IEEE Std. 1584–2018 y las metodologías alternativas (EPRI y Terzija/Koglin), se construye una matriz de correlaciones a partir del cálculo del coeficiente de correlación de Pearson. Este coeficiente permite cuantificar el grado de asociación lineal entre pares de variables numéricas, proporcionando una medida normalizada de su covariación [36].

El análisis se enfoca exclusivamente en aquellas variables numéricas directamente involucradas en las ecuaciones de cálculo de energía incidente de cada metodología. Por tanto, se excluyen del análisis variables categóricas como Escenario y Tipo de falla, ya que su función es únicamente descriptiva o clasificatoria, sin aportar información directa para la relación matemática entre variables continuas. De igual forma, el Coeficiente de encerramiento (Cenc) es omitido para los casos en aire libre (Aire Libre), ya que no se aplica en ese tipo de entorno. Por su parte, el Gradiente de tensión (B) se mantuvo constante en los escenarios simulados, dependiendo exclusivamente del nivel de corriente de cortocircuito. Dado que no varía entre casos, no aporta al análisis de correlación y también se descarta.

Una vez seleccionadas las variables relevantes, se procede al cálculo del coeficiente de Pearson para cada par de variables. Este coeficiente toma valores entre -1 y 1, donde:

- +1 indica una correlación lineal positiva perfecta.
- -1 indica una correlación lineal negativa perfecta.
- 0 sugiere la ausencia de una relación lineal significativa.

La matriz de correlaciones resultante permite una visualización integral del comportamiento lineal entre las variables, contribuyendo a:

- Detectar colinealidades potenciales,
- Priorizar variables para el entrenamiento de modelos de Machine Learning,
- Comprender el impacto relativo de cada variable en las metodologías de cálculo evaluadas.

Los resultados de este análisis se resumen en las tablas siguientes, donde se reportan tanto los valores de r (coeficiente de correlación de Pearson) como sus valores absolutos para facilitar la interpretación.

Tabla 5. Matriz de correlación conjunto de datos entrenamiento

Variable 1	Variable 2	r	abs_r
Relación IEEE - Terzija Earc Aire Libre	Gap (G) [mm]	-6.61E-01	6.61E-01
Gap (G) [mm]	Relación IEEE - Terzija Earc Aire Libre	-6.61E-01	6.61E-01
Relación IEEE - Terzija Earc Aire Libre	Relación IEEE - EPRI Earc Aire Libre	5.62E-01	5.62E-01
Relación IEEE - EPRI Earc Aire Libre	Relación IEEE - Terzija Earc Aire Libre	5.62E-01	5.62E-01
Corriente de cortocircuito (Ibf) [kA]	Tensión nominal (V) [kV]	-4.96E-01	4.96E-01
Tensión nominal (V) [kV]	Corriente de cortocircuito (Ibf) [kA]	-4.96E-01	4.96E-01
Relación IEEE - EPRI Earc Aire Libre	Gap (G) [mm]	-3.72E-01	3.72E-01
Gap (G) [mm]	Relación IEEE - EPRI Earc Aire Libre	-3.72E-01	3.72E-01
Distancia de Trabajo (D) [mm]	Relación IEEE - Terzija Earc Aire Libre	2.67E-01	2.67E-01
Relación IEEE - Terzija Earc Aire Libre	Distancia de Trabajo (D) [mm]	2.67E-01	2.67E-01
Relación IEEE - EPRI Earc Aire Libre	Configuración de electrodos (EC)	-2.38E-01	2.38E-01
Configuración de electrodos (EC)	Relación IEEE - EPRI Earc Aire Libre	-2.38E-01	2.38E-01
Configuración de electrodos (EC)	Relación IEEE - Terzija Earc Aire Libre	-1.60E-01	1.60E-01
Relación IEEE - Terzija Earc Aire Libre	Configuración de electrodos (EC)	-1.60E-01	1.60E-01
Relación IEEE - EPRI Earc Aire Libre	Corriente de cortocircuito (Ibf) [kA]	4.94E-02	4.94E-02
Corriente de cortocircuito (Ibf) [kA]	Relación IEEE - EPRI Earc Aire Libre	4.94E-02	4.94E-02
Relación IEEE - Terzija Earc Aire Libre	Corriente de cortocircuito (Ibf) [kA]	-2.15E-02	2.15E-02
Corriente de cortocircuito (Ibf) [kA]	Relación IEEE - Terzija Earc Aire Libre	-2.15E-02	2.15E-02
Relación IEEE - EPRI Earc Aire Libre	Tensión nominal (V) [kV]	-1.57E-02	1.57E-02
Tensión nominal (V) [kV]	Relación IEEE - EPRI Earc Aire Libre	-1.57E-02	1.57E-02
Relación IEEE - EPRI Earc Aire Libre	Distancia de Trabajo (D) [mm]	-9.16E-03	9.16E-03
Distancia de Trabajo (D) [mm]	Relación IEEE - EPRI Earc Aire Libre	-9.16E-03	9.16E-03
Tensión nominal (V) [kV]	Relación IEEE - Terzija Earc Aire Libre	-8.03E-03	8.03E-03
Relación IEEE - Terzija Earc Aire Libre	Tensión nominal (V) [kV]	-8.03E-03	8.03E-03
Tensión nominal (V) [kV]	Gap (G) [mm]	1.70E-15	1.70E-15
Gap (G) [mm]	Tensión nominal (V) [kV]	1.70E-15	1.70E-15
Corriente de cortocircuito (Ibf) [kA]	Gap (G) [mm]	-1.34E-15	1.34E-15
Gap (G) [mm]	Corriente de cortocircuito (Ibf) [kA]	-1.34E-15	1.34E-15
Tensión nominal (V) [kV]	Configuración de electrodos (EC)	-6.28E-16	6.28E-16

Variable 1	Variable 2	r	abs_r
Configuración de electrodos (EC)	Tensión nominal (V) [kV]	-6.28E-16	6.28E-16
Gap (G) [mm]	Distancia de Trabajo (D) [mm]	-5.71E-16	5.71E-16
Distancia de Trabajo (D) [mm]	Gap (G) [mm]	-5.71E-16	5.71E-16
Corriente de cortocircuito (Ibf) [kA]	Distancia de Trabajo (D) [mm]	4.809E-16	4.809E-16
Distancia de Trabajo (D) [mm]	Corriente de cortocircuito (Ibf) [kA]	4.809E-16	4.809E-16
Configuración de electrodos (EC)	Distancia de Trabajo (D) [mm]	4.48E-16	4.48E-16
Distancia de Trabajo (D) [mm]	Configuración de electrodos (EC)	4.48E-16	4.48E-16
Configuración de electrodos (EC)	Corriente de cortocircuito (Ibf) [kA]	2.085E-16	2.085E-16
Corriente de cortocircuito (Ibf) [kA]	Configuración de electrodos (EC)	2.085E-16	2.085E-16
Tensión nominal (V) [kV]	Distancia de Trabajo (D) [mm]	-1.97E-17	1.967E-17
Distancia de Trabajo (D) [mm]	Tensión nominal (V) [kV]	-1.97E-17	1.967E-17
Configuración de electrodos (EC)	Gap (G) [mm]	1.77E-17	1.77E-17
Gap (G) [mm]	Configuración de electrodos (EC)	1.77E-17	1.77E-17

Tabla 6. Matriz de correlación conjunto de datos validación

Variable 1	Variable 2	r	abs_r
Tensión nominal (V) [kV]	Corriente de cortocircuito (Ibf) [kA]	-1.00E+00	1.00E+00
Corriente de cortocircuito (Ibf) [kA]	Tensión nominal (V) [kV]	-1.00E+00	1.00E+00
Relación IEEE - Terzija Earc Aire Libre	Gap (G) [mm]	-6.41E-01	6.41E-01
Gap (G) [mm]	Relación IEEE - Terzija Earc Aire Libre	-6.41E-01	6.41E-01
Relación IEEE - EPRI Earc Aire Libre	Relación IEEE - Terzija Earc Aire Libre	5.27E-01	5.27E-01
Relación IEEE - Terzija Earc Aire Libre	Relación IEEE - EPRI Earc Aire Libre	5.27E-01	5.27E-01
Relación IEEE - EPRI Earc Aire Libre	Gap (G) [mm]	-2.83E-01	2.83E-01
Gap (G) [mm]	Relación IEEE - EPRI Earc Aire Libre	-2.83E-01	2.83E-01
Relación IEEE - EPRI Earc Aire Libre	Configuración de electrodos (EC)	-2.83E-01	2.83E-01
Configuración de electrodos (EC)	Relación IEEE - EPRI Earc Aire Libre	-2.83E-01	2.83E-01
Distancia de Trabajo (D) [mm]	Relación IEEE - Terzija Earc Aire Libre	2.63E-01	2.63E-01
Relación IEEE - Terzija Earc Aire Libre	Distancia de Trabajo (D) [mm]	2.63E-01	2.63E-01
Configuración de electrodos (EC)	Relación IEEE - Terzija Earc Aire Libre	-1.88E-01	1.88E-01
Relación IEEE - Terzija Earc Aire Libre	Configuración de electrodos (EC)	-1.88E-01	1.88E-01
Tensión nominal (V) [kV]	Relación IEEE - Terzija Earc Aire Libre	-2.90E-02	2.90E-02
Relación IEEE - Terzija Earc Aire Libre	Tensión nominal (V) [kV]	-2.90E-02	2.90E-02
Relación IEEE - Terzija Earc Aire Libre	Corriente de cortocircuito (Ibf) [kA]	2.90E-02	2.90E-02
Corriente de cortocircuito (Ibf) [kA]	Relación IEEE - Terzija Earc Aire Libre	2.90E-02	2.90E-02
Distancia de Trabajo (D) [mm]	Relación IEEE - EPRI Earc Aire Libre	-2.37E-02	2.37E-02
Relación IEEE - EPRI Earc Aire Libre	Distancia de Trabajo (D) [mm]	-2.37E-02	2.37E-02
Relación IEEE - EPRI Earc Aire Libre	Corriente de cortocircuito (Ibf) [kA]	-1.14E-02	1.14E-02
Corriente de cortocircuito (Ibf) [kA]	Relación IEEE - EPRI Earc Aire Libre	-1.14E-02	1.14E-02
Relación IEEE - EPRI Earc Aire Libre	Tensión nominal (V) [kV]	1.14E-02	1.14E-02
Tensión nominal (V) [kV]	Relación IEEE - EPRI Earc Aire Libre	1.14E-02	1.14E-02

Variable 1	Variable 2	r	abs_r
Distancia de Trabajo (D) [mm]	Corriente de cortocircuito (Ibf) [kA]	2.06E-15	2.06E-15
Corriente de cortocircuito (Ibf) [kA]	Distancia de Trabajo (D) [mm]	2.06E-15	2.06E-15
Distancia de Trabajo (D) [mm]	Tensión nominal (V) [kV]	-1.67E-15	1.67E-15
Tensión nominal (V) [kV]	Distancia de Trabajo (D) [mm]	-1.67E-15	1.67E-15
Distancia de Trabajo (D) [mm]	Gap (G) [mm]	-1.08E-15	1.08E-15
Gap (G) [mm]	Distancia de Trabajo (D) [mm]	-1.08E-15	1.08E-15
Gap (G) [mm]	Tensión nominal (V) [kV]	7.841E-16	7.841E-16
Tensión nominal (V) [kV]	Gap (G) [mm]	7.841E-16	7.841E-16
Gap (G) [mm]	Corriente de cortocircuito (Ibf) [kA]	-6.63E-16	6.633E-16
Corriente de cortocircuito (Ibf) [kA]	Gap (G) [mm]	-6.63E-16	6.633E-16
Configuración de electrodos (EC)	Corriente de cortocircuito (Ibf) [kA]	3.794E-16	3.794E-16
Corriente de cortocircuito (Ibf) [kA]	Configuración de electrodos (EC)	3.794E-16	3.794E-16
Configuración de electrodos (EC)	Tensión nominal (V) [kV]	-3.76E-16	3.763E-16
Tensión nominal (V) [kV]	Configuración de electrodos (EC)	-3.76E-16	3.763E-16
Configuración de electrodos (EC)	Distancia de Trabajo (D) [mm]	1.513E-16	1.513E-16
Distancia de Trabajo (D) [mm]	Configuración de electrodos (EC)	1.513E-16	1.513E-16
Configuración de electrodos (EC)	Gap (G) [mm]	1.108E-17	1.108E-17
Gap (G) [mm]	Configuración de electrodos (EC)	1.108E-17	1.108E-17

Con el objetivo de facilitar la interpretación de las correlaciones entre las variables de entrada y las variables objetivo, se recurre a la generación de un mapa de calor. Esta herramienta gráfica permite visualizar, de forma intuitiva y detallada, el grado de asociación lineal entre variables, a través de una escala cromática que representa la intensidad y dirección de los coeficientes de correlación de Pearson.

En particular, el interés del análisis se centra en identificar la influencia de las variables de entrada sobre las variables de salida Relación IEEE – EPRI Earc Aire Libre (Factor de ajuste EPRI) y Relación IEEE – Terzija Aire Libre (Factor de ajuste Terzija/Koglin).



Figura 5. Mapa de calor – Matriz de correlaciones conjunto de datos entrenamiento



Figura 6. Mapa de calor – Matriz de correlaciones conjunto de datos validación

Los mapas de calor mostrados en la Figura 5 y Figura 6 reflejan que las variables de salida Relación IEEE – EPRI Earc Aire Libre (Factor de ajuste EPRI) y Relación IEEE – Terzija Aire Libre (Factor de ajuste Terzija/Koglin), tienen una mayor dependencia a la distancia de trabajo, distancia entre electrodos (gap) y configuración de electrodos. Principalmente el gap tiene una mayor impacto en el cálculo del factor de ajuste Terzija/Koglin, alcanzando un coeficiente absoluto de 0.66.

Por su parte, la tensión nominal y la corriente de cortocircuito no tienen un impacto lineal representativo en los factores de ajuste, lo que podría indicar que el impacto de estas variables en el cálculo de la energía incidente bajo ambas metodologías es similar, sin generar discrepancias significativas entre los valores obtenidos con IEEE Std. 1584 – 2018 y las metodologías alternativas.

3.4 Selección de variables relevantes

Considerando los análisis realizados previamente, se considera relevante incluir las siguientes variables de entrada para el entrenamiento del modelo:

- **Distancia de trabajo:** Esta variable mostró una alta correlación con la relación calculada entre las metodologías IEEE y Terzija/Koglin. Su inclusión como variable de entrada es crucial para reflejar adecuadamente la interacción entre estos dos enfoques y mejorar la precisión del modelo.
- **Distancia entre electrodos (gap):** La distancia entre electrodos se presentó como la variable con mayor valor de correlación en ambas relaciones, tanto en la Relación IEEE – EPRI Earc Aire Libre como en la Relación IEEE – Terzija Earc Aire Libre. Su alta correlación y relevancia en el cálculo de energía incidente la convierten en una variable notable para el modelo.
- **Configuración de electrodos:** Además de su relevancia en el cálculo de la energía incidente con la metodología IEEE, la variabilidad que genera en los resultados destaca la importancia de considerar esta variable. Su inclusión mejora la capacidad del modelo para generalizar correctamente entre diferentes configuraciones y metodologías.

El uso de estas variables de entrada se basa en su importancia para el análisis y en su capacidad para influir directamente en la precisión del modelo, permitiendo una mejor predicción y adaptación a diferentes escenarios de tensión y configuración.

3.5 Generación de modelos de Machine Learning

La selección de uno o varios modelos de Machine Learning debe partir de una clara definición del objetivo del análisis y la razón de su aplicación al conjunto de datos en cuestión. En este caso, dado que los conjuntos de datos creados no requieren una clasificación de datos, el enfoque debe

centrarse en modelos de regresión, cuya capacidad predictiva puede ser clave para estimar factores de ajuste.

Estos factores permitirán implementar metodologías alternativas, como las propuestas por EPRI y Terzija/Koglin, para el cálculo de la energía incidente generada por arcos trifásicos en encerramiento. El uso de Machine Learning en este contexto contribuiría a generar estimaciones más precisas y conservativas de energía incidente, facilitando la selección adecuada de Equipos de Protección Personal (EPP) para la realización de actividades eléctricas cerca de partes energizadas a niveles de tensión superiores a 15 kV. De esta manera, se fortalecería la seguridad de los trabajadores, minimizando el riesgo de lesiones graves en caso de un evento de arco eléctrico.

Cuando no se dispone de un conocimiento preciso sobre las relaciones entre variables dentro de un conjunto de datos, es recomendable emplear distintos enfoques de modelado en Machine Learning. Cada tipo de modelo presenta ventajas y limitaciones específicas que pueden influir en su capacidad para capturar los patrones y tendencias subyacentes de los datos. Por esta razón, se opta por evaluar diversos modelos de Machine Learning para la predicción de factores de ajuste en el cálculo de energía incidente, contrastando tanto su precisión como el costo computacional asociado al entrenamiento y validación.

A partir del análisis de correlaciones desarrollado en el numeral 3.3, se identificó que múltiples variables presentan relaciones lineales significativas con los factores de ajuste derivados de cada metodología. Esto justifica la implementación de modelos de regresión lineal y regresión polinómica de segundo grado, los cuales, además de ser modelos interpretables, requieren bajo procesamiento computacional y permiten establecer líneas base sólidas para comparar el desempeño de modelos más complejos.

Complementariamente, dada la dimensión de los conjuntos de datos analizados, que alcanzan hasta 20.000 filas, el uso de modelos basados en redes neuronales resulta una alternativa viable y prometedora. En este contexto, las dependencias lineales detectadas entre variables podrían facilitar el proceso de convergencia de las redes, reduciendo la necesidad de realizar un ajuste intensivo de hiperparámetros y acortando los tiempos de entrenamiento sin comprometer la precisión.

En total se entrenan y evalúan los siguientes tres modelos de Machine Learning:

1. Regresión Lineal.
2. Regresión polinómica de grado 2.
3. Redes neuronales.

Para cada estrategia de modelado, se generan dos modelos independientes: uno orientado a capturar los patrones asociados a la Relación IEEE – EPRI Earc Aire Libre (Factor de ajuste EPRI), y otro enfocado en la Relación IEEE – Terzija Aire Libre (Factor de ajuste Terzija/Koglin). Esta duplicación por estrategia permite especializar cada modelo en la predicción de una relación específica, optimizando así la precisión del ajuste en función del tipo de factor a estimar.

Los modelos de Machine Learning mencionados se establecen en Python y se ajustan tal como se indica a continuación:

Modelo ElasticNet: Se construye utilizando la función ElasticNet. Se establecen los hiperparámetros en $\alpha = 0.5$, asegurando un grado de penalización moderado, y $l1_ratio = 0.1$, favoreciendo la regularización Ridge (L2) para evitar la eliminación de variables importantes. Se seleccionan estos valores para mantener un equilibrio entre la penalización Lasso (L1) y Ridge (L2), evitando que alguna de las variables clave sea descartada durante el entrenamiento.

Modelo de Regresión Polinómica de Grado 2: Se emplea la función PolynomialFeatures para transformar los datos de entrenamiento y validación al grado polinómico deseado, en este caso, grado 2 ($degree = 2$). Posteriormente, se crea un modelo de Regresión Lineal con la función LinearRegression, el cual será entrenado con los datos transformados.

Modelo Red Neural: Se implementa con la función Sequential, permitiendo definir múltiples capas de neuronas mediante la función Dense. Para este modelo, se utilizarán cuatro capas configuradas para generar una única salida así: $64 \rightarrow 32 \rightarrow 16 \rightarrow 1$ neuronas. Además, se establece el número de épocas (epochs) igual a 100 y el tamaño de lote (batch_size) de 32, siendo valores adecuados en problemas de baja complejidad. Las épocas determinan cuántas veces el modelo recorrerá el conjunto de datos durante el entrenamiento, mientras que el tamaño del lote define cuántas muestras se procesan antes de actualizar los pesos del modelo. Estos parámetros influyen en la convergencia y el rendimiento de la red, por lo que su ajuste adecuado es esencial para obtener buenos resultados.

A todos los modelos desarrollados se les aplicó una normalización mediante escalado estándar (puntuaciones Z) con el objetivo de transformar cada una de las variables de entrada para que presenten una media de cero y una desviación estándar unitaria, lo cual resulta fundamental en modelos sensibles a la escala de los datos, como las regresiones regularizadas y las redes neuronales. Esta transformación ayuda a nivelar el impacto de cada variable, evitando que aquellas con valores mucho mayores condicionen el resultado del modelo.

Se debe mencionar que el establecimiento de los hiperparámetros no suele tener criterios específicos, es por ello por lo que se suelen hacer múltiples validaciones hasta obtener un desempeño adecuado del modelo evaluado.

3.6 Entrenamiento de modelos

El entrenamiento de modelos de Machine Learning es fundamental para que el modelo aprenda de los datos, detecte patrones y pueda hacer predicciones precisas y confiables para su aplicación práctica. Asimismo, un entrenamiento adecuado permite que el modelo sea capaz de generalizar acorde a un conjunto de datos nuevos, evitando posibles sobreajustes.

Durante esta fase, el conjunto de datos de entrenamiento es empleado para ajustar los parámetros internos de los modelos. Para ello, es necesario separar claramente las variables de entrada (atributos) y las variables de salida (factores de ajuste), ya que el proceso de aprendizaje

consiste en establecer una relación funcional que permita, a partir de los atributos, predecir adecuadamente el valor de salida esperado.

Dado que por cada estrategia de modelado se han generado dos modelos específicos, es fundamental establecer cuáles se utilizarán para la predicción del factor de ajuste correspondiente a la metodología EPRI y cuáles para la metodología Terzija/Koglin, asegurando así que cada modelo sea entrenado de acuerdo con su objetivo predictivo. Para asegurar que el modelo se entrene y pruebe de manera adecuada, el conjunto de datos de entrenamiento se divide en los siguientes dos subconjuntos:

- **Conjunto de entrenamiento (70%):** El modelo aprende de estos datos, ajustando sus parámetros internos para capturar patrones y relaciones.
- **Conjunto de prueba (30%):** Permite poner a prueba la capacidad de generalización del modelo, evitando el sobreajuste a los datos de entrenamiento. Este conjunto se usará en la etapa de validación de los modelos.

El entrenamiento de los modelos se realiza mediante la función `.fit()`, la cual está disponible en diversas librerías de Python según el tipo de modelo utilizado. Esta función ajusta los parámetros internos del modelo a partir de los datos de entrenamiento previamente procesados. Como resultado, se obtiene el modelo entrenado, el cual pasará a la etapa de validación.

Es importante destacar que, durante la fase de validación de los modelos, es posible obtener un rendimiento bajo, lo que indica la necesidad de reajustar los parámetros e hiperparámetros, con el fin de repetir el proceso de entrenamiento hasta alcanzar un desempeño aceptable para la predicción de factores de ajuste.

3.7 Validación de modelos y selección de MAEI

En la etapa de validación se evalúa el desempeño de los modelos generados mediante el análisis de las siguientes métricas estadísticas:

- Error Absoluto Medio (MAE)
- Raíz del Error Cuadrático Medio (RMSE)
- Coeficiente de Determinación (R^2)

Estas métricas se calculan para cada uno de los conjuntos de datos: entrenamiento, prueba y validación. Los resultados obtenidos en los conjuntos de entrenamiento y prueba permiten identificar posibles casos de sobreajuste o subajuste en los modelos. Por su parte, el análisis sobre el conjunto de validación permite evaluar la capacidad de generalización de cada modelo.

El proceso de evaluación se dará por terminado cuando al menos uno de los modelos alcance un desempeño satisfactorio, determinado a partir de los resultados obtenidos en las métricas de evaluación para cada conjunto de datos. El objetivo principal es minimizar los valores del MAE y RMSE, utilizando la media de los valores reales de cada conjunto como referencia para evaluar la magnitud del error. Por otro lado, el coeficiente de determinación (R^2) deberá alcanzar un valor

igual o superior a 0.9, lo que garantiza que el modelo explica adecuadamente la variabilidad de los datos.

A continuación, se presentan los resultados obtenidos de la validación con métricas estadísticas para cada uno de los modelos entrenados.

Resultados modelo ElasticNet

En la evaluación del modelo ElasticNet, se observó un desempeño ineficaz en la identificación de tendencias y patrones entre las variables. Como se muestra en la Figura 7, Figura 8 y Figura 9, las métricas de evaluación reflejan valores poco satisfactorios para la predicción de relaciones, tanto en el caso de Terzija/Koglin como en EPRI. El Error Absoluto Medio (MAE) presentó valores entre 0.3644 y 0.4302, lo que, en comparación con la media de los datos de salida, representa aproximadamente un tercio de su valor. Esto indica que las predicciones del modelo presentan un error significativo respecto a los valores reales. La Raíz del Error Cuadrático Medio (RMSE) se situó entre 0.4812 y 0.5592, con los mayores errores registrados en el modelo ElasticNet para la relación IEEE – Terzija Earc Aire Libre. Este resultado refleja la sensibilidad del RMSE a desviaciones elevadas en los datos, evidenciando una dispersión de casi la mitad de la media de los valores reales. Respecto al coeficiente de determinación (R^2), el modelo ElasticNet IEEE – EPRI Earc Aire Libre mostró valores inferiores a 0.2, lo que indica una baja capacidad para explicar la variabilidad de los datos. Para el caso de Terzija/Koglin, el valor de R^2 se aproximó a 0.5, sugiriendo una capacidad moderada para capturar la tendencia general, aunque aún insuficiente para una predicción precisa.

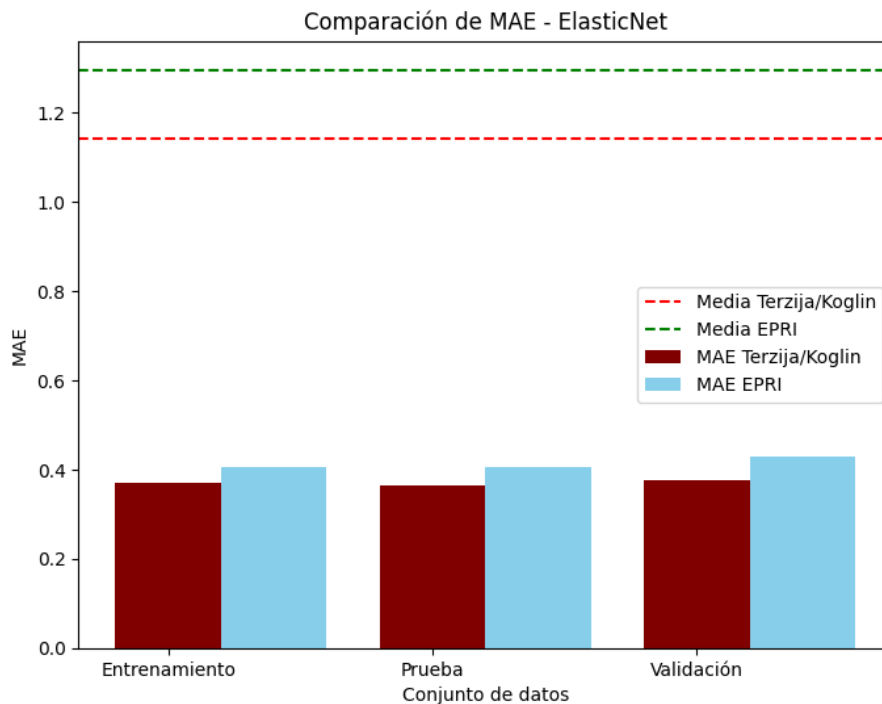


Figura 7. Evaluación de conjuntos de datos - Métrica MAE - ElasticNet

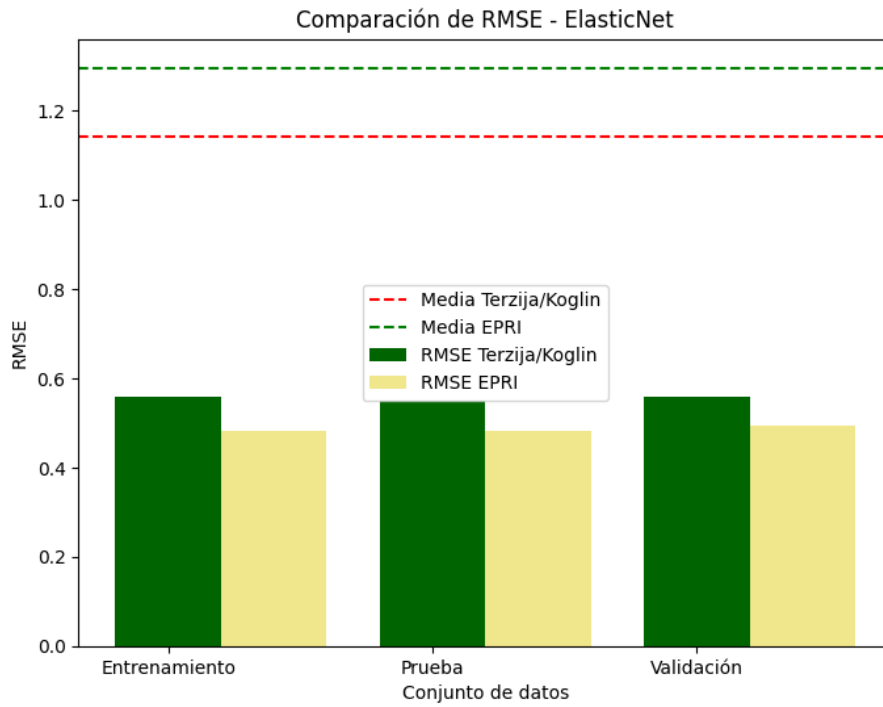


Figura 8. Evaluación de conjuntos de datos - Métrica RMSE - ElasticNet

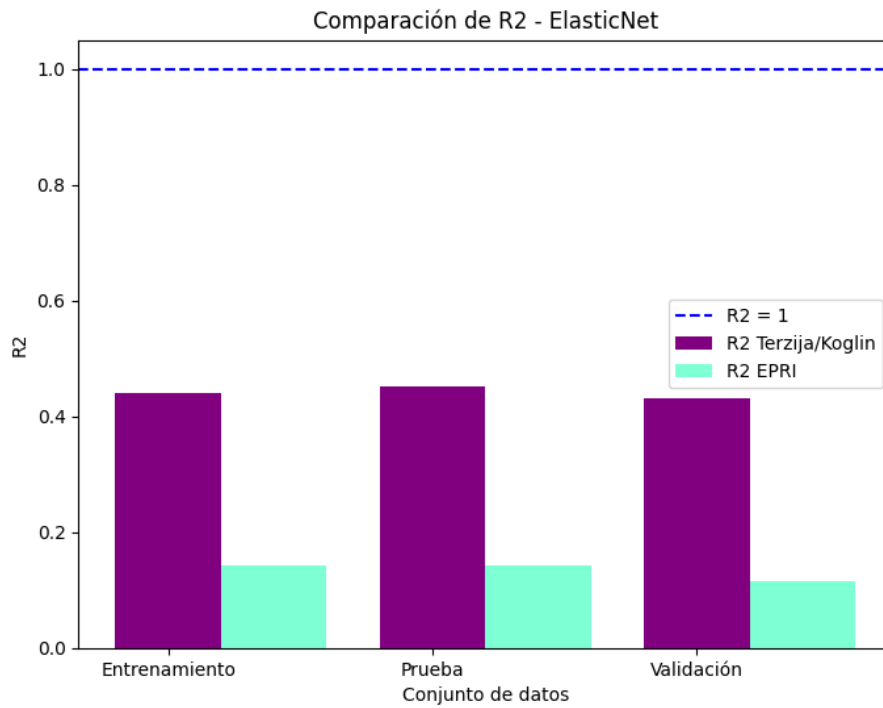


Figura 9. Evaluación de conjuntos de datos - Métrica R2 - ElasticNet

Los valores obtenidos en las métricas de evaluación indican que el modelo ElasticNet no es adecuado para la predicción de datos en las relaciones entre metodologías de cálculo de energía incidente, ya que las relaciones entre las variables no se pueden explicar solamente con relaciones lineales.

Resultados modelo Regresión Polinomial de Grado 2

El modelo de Regresión Polinomial de Grado 2 mostró un mejor desempeño en las métricas de evaluación utilizadas durante el entrenamiento. Como se observa en la Figura 10 y Figura 11, el Error Absoluto Medio (MAE) oscila entre 0.1494 y 0.1851, mientras que la Raíz del Error Cuadrático Medio (RMSE) varía entre 0.2342 y 0.2573. Estos valores reflejan una mejora significativa en comparación con el modelo ElasticNet, aunque aún sin alcanzar resultados completamente satisfactorios para la predicción de los factores de ajuste. Por otro lado, el coeficiente de determinación (R^2), presentado en la Figura 12, evidencia una mejora considerable, alcanzando valores cercanos a 0.8 para las relaciones con EPRI y aproximadamente 0.9 en el caso de Terzija/Koglin, lo que indica una mayor capacidad del modelo para captar la variabilidad de los datos.

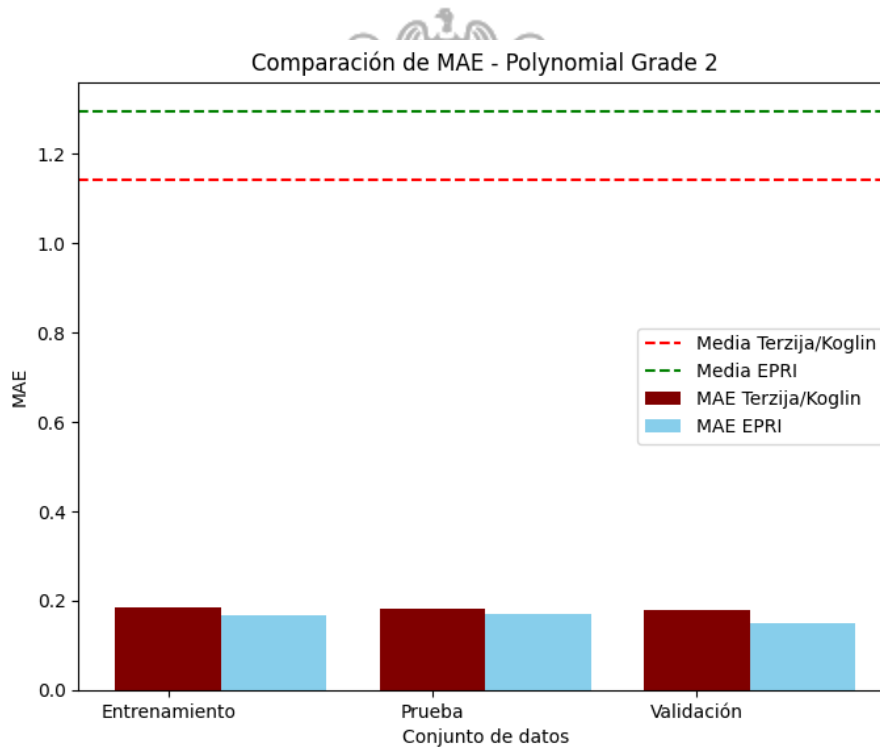


Figura 10. Evaluación de conjuntos de datos - Métrica MAE - Polynomial Grade 2

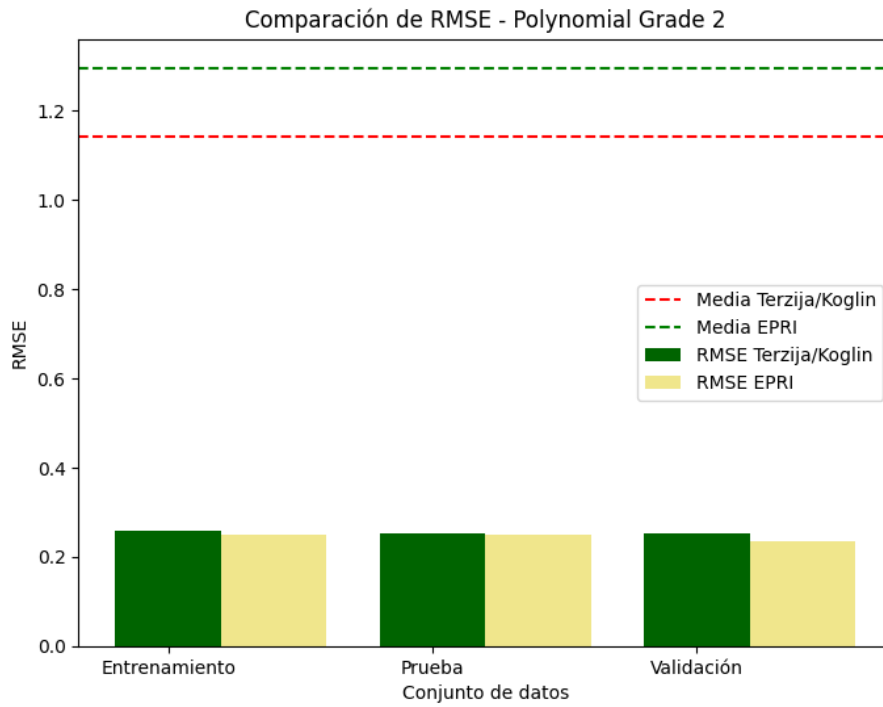


Figura 11. Evaluación de conjuntos de datos - Métrica RMSE - Polynomial Grade 2

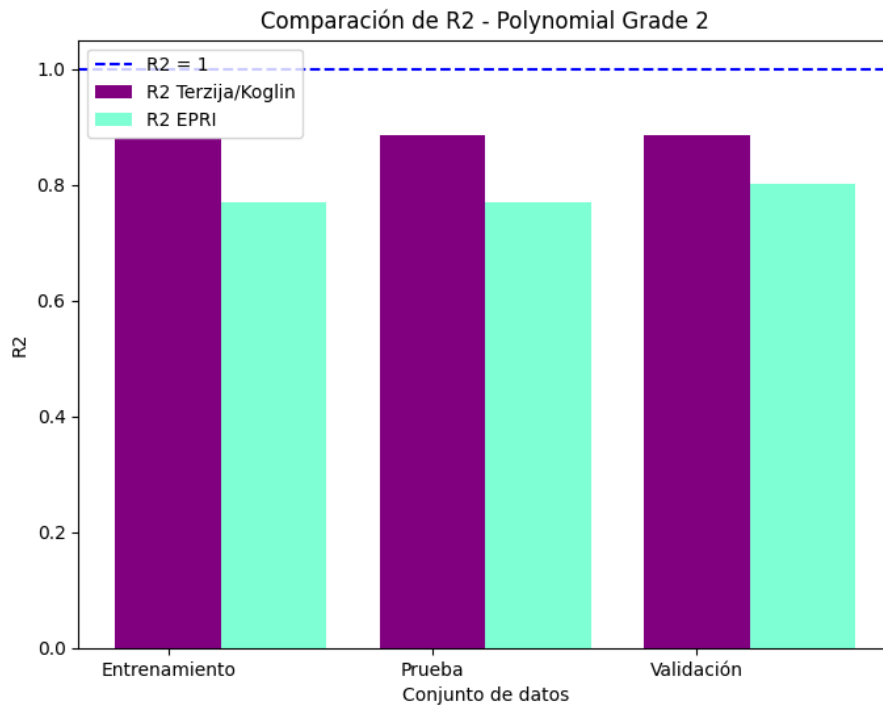


Figura 12. Evaluación de conjuntos de datos - Métrica R2 - Polynomial Grade 2

Los resultados obtenidos en la evaluación del modelo de Regresión Polinomial de Grado 2 indican que, incluso con modelos relativamente simples, es posible capturar en cierta medida las tendencias y patrones en la relación entre metodologías de cálculo de energía incidente. La eficacia de este modelo podría estar asociada a la relación exponencial de algunas variables dentro de las ecuaciones analizadas. Sin embargo, factores como la configuración de electrodos podrían influir en los valores de las métricas MAE y RMSE, afectando la precisión del modelo. Esto se debe a la forma en que dicha variable impacta las ecuaciones de la metodología IEEE Std. 1584 – 2018, lo que sugiere la necesidad de explorar modelos más complejos o enfoques adicionales para mejorar la capacidad predictiva.

Resultados modelo Red Neuronal

En la Figura 13, Figura 14 y Figura 15 se presentan los resultados de la evaluación del modelo de Red Neuronal a partir de las métricas MAE, RMSE y R^2 . Los valores obtenidos evidencian un desempeño elevado para la predicción de factores de ajuste, ya que se obtuvo un MAE entre 0.0796 y 0.0907, un RMSE entre 0.1087 y 0.1239, y un coeficiente de determinación (R^2) entre 0.9507 y 0.9785.

Este rendimiento no solo se manifiesta en el conjunto de entrenamiento, sino que también se mantiene en los conjuntos de prueba y validación, lo que indica una excelente capacidad de generalización del modelo. Además, al comparar los valores de MAE y RMSE con la media de los datos de salida, se observa que representan menos del 10% de su valor, lo que confirma una alta precisión en la predicción y una desviación aceptable respecto a los valores reales.

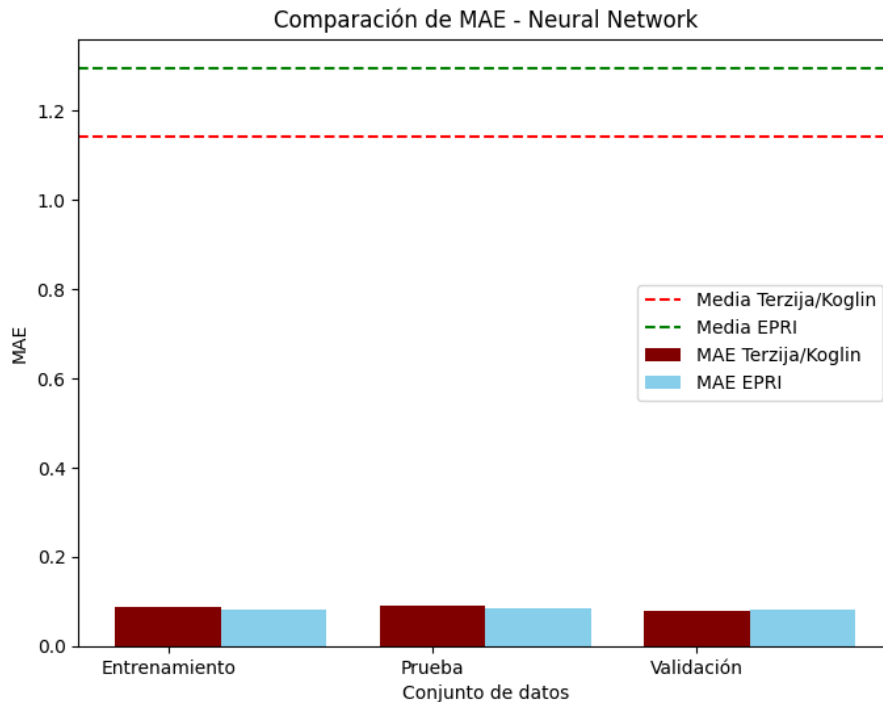


Figura 13. Evaluación de conjuntos de datos - Métrica MAE - Red Neuronal

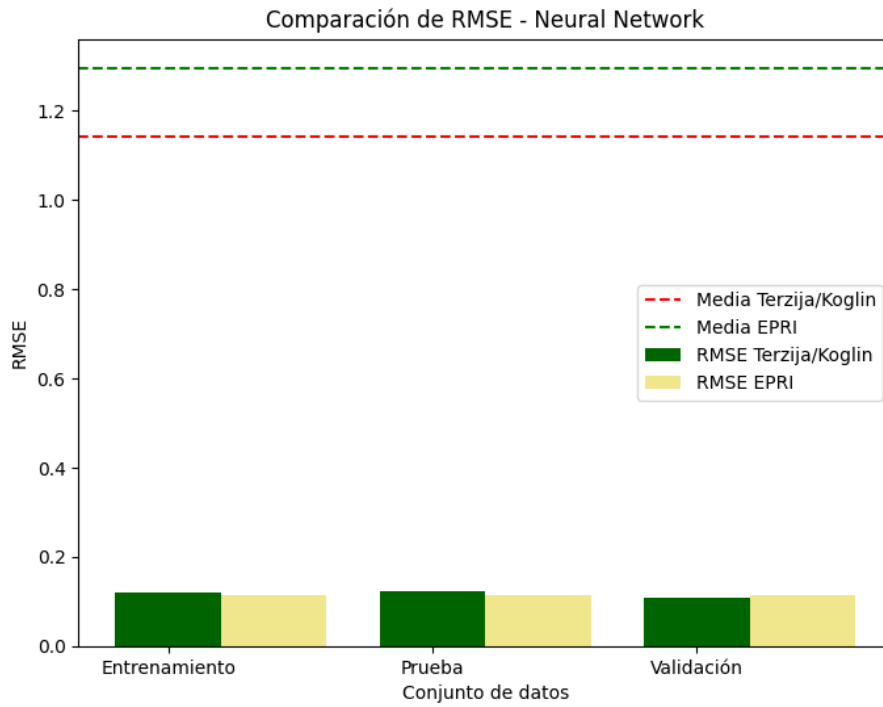


Figura 14. Evaluación de conjuntos de datos - Métrica RMSE - Red Neuronal

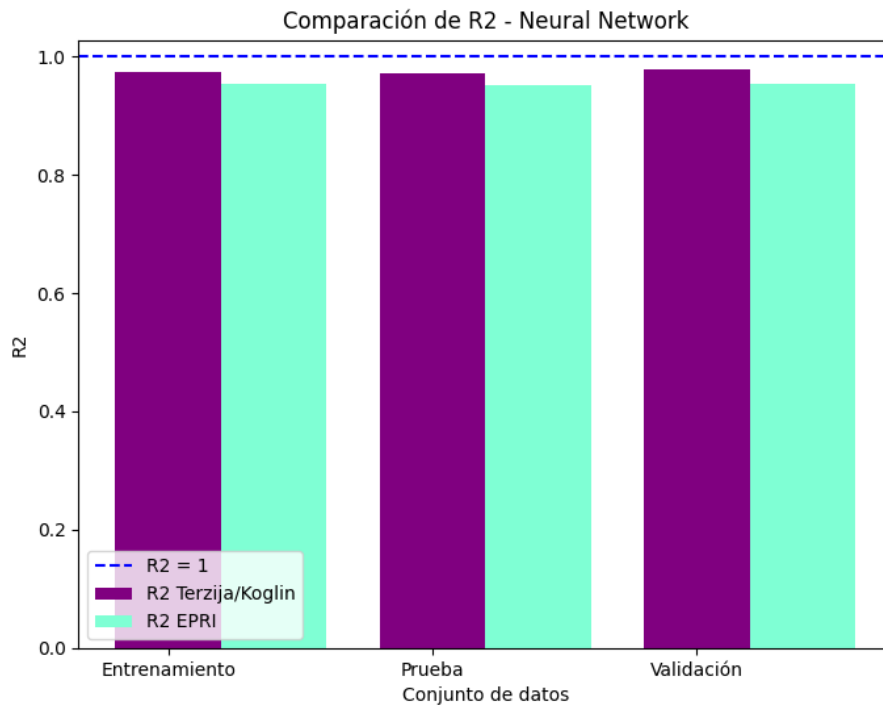


Figura 15. Evaluación de conjuntos de datos - Métrica R2 - Red Neuronal

Dado que el modelo de Red Neuronal alcanzó un desempeño deseado en la predicción de los factores de ajuste, se decide no realizar ajustes adicionales en los parámetros ni hiperparámetros de los modelos previamente analizados. La precisión obtenida con la red neuronal sugiere que este enfoque es el más adecuado para capturar la relación entre las variables de las diferentes metodologías de cálculo de energía incidente, lo que lo convierte en una alternativa viable para la predicción de estos factores.

El éxito de la Red Neuronal en comparación con los modelos de regresión se debe a su capacidad para modelar relaciones no lineales complejas y capturar patrones subyacentes en los datos que los modelos tradicionales no pueden identificar fácilmente. A diferencia de la Regresión Polinomial de Grado 2, que solo puede ajustar relaciones de tipo cuadrático, la red neuronal puede aprender y representar interacciones no lineales de orden superior, lo que le permite adaptarse mejor a las ecuaciones de las metodologías analizadas.

Otro factor clave en su desempeño es su capacidad para manejar grandes volúmenes de datos y descubrir relaciones ocultas mediante el ajuste de pesos en sus capas internas. La red neuronal puede asignar diferentes niveles de importancia a cada variable de entrada, lo que es especialmente relevante en el caso de metodologías como la IEEE Std. 1584 – 2018, donde variables como la configuración de electrodos del encerramiento afectan significativamente los resultados.

Teniendo en cuenta que no se realiza un reentrenamiento de los modelos, se procede a realizar la selección del Modelo de Ajuste de Energía Incidente (MAEI).

Selección de MAEI

Considerando los resultados obtenidos, el modelo que tiene un mejor desempeño para ser usado como Modelo de Ajuste de Energía Incidente (MAEI) es el de Red Neuronal, ya que demostró una mejor capacidad de generalización para la predicción de factores de ajuste. Es por ello por lo que se utilizará este modelo, tanto para la metodología Terzija/Koglin como para EPRI, para complementar los cálculos de energía incidente en encerramientos con metodologías alternativas.

Los modelos de Machine Learning están disponibles en el repositorio público de GitHub 15_kV_Arc_Flash [37].

4. IMPLEMENTACIÓN Y EVALUACIÓN DEL MAEI

En este capítulo se presenta la propuesta de implementación del MAEI empleando metodologías alternativas, para el cálculo de energía incidente en encerramientos. Asimismo, se realizan validaciones con metodologías tradicionales que permiten comparar el desempeño de los resultados obtenidos con la propuesta, evaluando su rendimiento frente a la metodología IEEE Std. 1584 – 2018 en tensiones menores a 15 kV y frente a cálculos basados en metodologías comúnmente utilizadas en niveles de tensión superiores a 15 kV.

4.1 Propuesta de implementación

Con el fin de hacer un uso adecuado de las metodologías alternativas de cálculo de energía incidente en encerramientos y ajustar los resultados a valores más conservadores aplicables a estudios de arco eléctrico en celdas AIS, se plantea la siguiente propuesta:

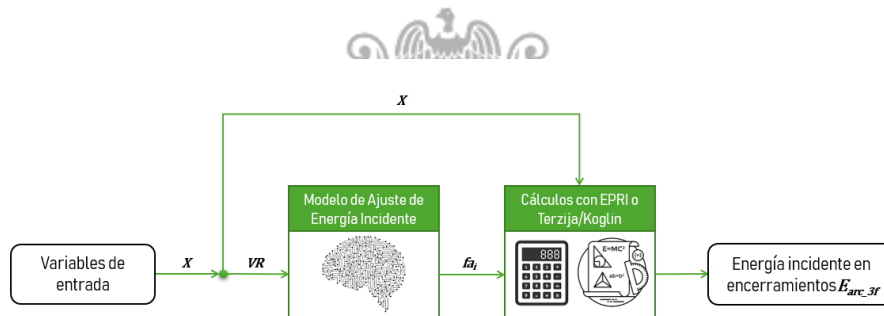


Figura 16. Propuesta de implementación de MAEI

Variables de entrada: Corresponden a los datos requeridos por las metodologías alternativas para el cálculo de la energía incidente generada por arcos eléctricos monofásicos en aire libre. Estas variables se establecen en función del Escenario de estudio específico que se pretenda analizar, e incluyen tanto parámetros eléctricos generales como variables particulares de cada metodología.

Entre las variables eléctricas generales se encuentran la tensión nominal del sistema, la corriente de cortocircuito, la distancia de trabajo, y el gap. Estas variables son fundamentales en cualquier evaluación de energía incidente.

Adicionalmente, cada metodología contempla variables propias que responden a sus fundamentos teóricos y empíricos. Por ejemplo, la metodología Terzija/Koglin incorpora el gradiente de tensión B , el cual establece una relación entre la corriente de cortocircuito y la tensión del sistema para caracterizar el arco. Por otro lado, la metodología EPRI introduce el factor estadístico de ajuste k que se define a partir de los análisis realizados en el Technical Report de EPRI [15].

Asimismo, es importante considerar la configuración de electrodos, la cual, si bien no interviene directamente en los cálculos de energía incidente realizados mediante las metodologías alternativas, sí se incorpora como un atributo de entrada en el modelo MAEI para la predicción de factores de ajuste.

Modelo de Ajuste de Energía Incidente (MAEI): En este bloque de la propuesta de implementación se ejecuta la predicción de los factores de ajuste (f_a) mediante el Modelo de Ajuste de Energía Incidente (MAEI), el cual ha sido previamente entrenado con las Variables Relevantes (VR) seleccionadas en el numeral 3.4, que corresponden a la distancia de trabajo, distancia entre electrodos (gap) y configuración de electrodos.

La aplicación del modelo es directa, ya que con el conjunto de valores relevantes, el MAEI entrega como salida un conjunto de factores de ajuste que permiten aumentar el valor de energía incidente en función del escenario evaluado.

Cálculos con EPRI o Terzija/Koglin: En esta etapa se emplean las ecuaciones propuestas por alguna de las metodologías alternativas seleccionadas (EPRI o Terzija/Koglin) para calcular la energía incidente generada por arcos monofásicos al aire libre. El procedimiento inicia con la recepción de las variables de entrada requeridas, tales como la corriente de cortocircuito, el gap entre electrodos, la distancia de trabajo, el nivel de tensión, entre otros parámetros característicos de la metodología utilizada.

Adicionalmente, se incorpora el factor de ajuste previamente estimado mediante el modelo MAEI, el cual representa el incremento esperado de la energía incidente al pasar de un arco monofásico al aire libre a un arco trifásico confinado en un tipo específico de encerramiento. De esta manera, se obtendría el valor de energía incidente en encerramientos ($E_{arc,3f}$).

A continuación, se lleva a cabo una evaluación del desempeño para niveles de tensión tanto inferiores como superiores a 15 kV. En esta evaluación, la energía incidente se calcula considerando las etapas definidas en la propuesta de implementación del MAEI empleando metodologías alternativas, cuyos resultados son equivalentes a arcos trifásicos en encerramientos. Finalmente, estos valores se comparan con los obtenidos a partir de metodologías tradicionales y referencias de la literatura técnica.

4.2 Evaluación de desempeño del MAEI en niveles de tensión inferiores a 15 kV

En esta sección se evalúa el desempeño del MAEI en escenarios con niveles de tensión inferiores a 15 kV, con el objetivo de contrastar sus resultados frente a los obtenidos mediante las metodologías tradicionales IEEE Std. 1584-2018 y el enfoque propuesto por Ralph Lee. Para ello, se plantean los siguientes escenarios de análisis:

Tabla 7. Escenarios de estudio - MAEI en niveles de tensión inferiores a 15 kV

Parámetros generales				
Escenario de estudio	Subestación	Nivel de tensión [kV]	Corriente de cortocircuito [kA]	Tipo de falla
Escenario 8	Subestación E	13.2	13.7	Bifásica a tierra
Escenario 9	Subestación E	4.16	15.06	Trifásica
Impedancias de secuencia				
Escenario de estudio	Subestación	Z0 [Ω]	Z1 [Ω]	Z2 [Ω]
Escenario 8	Subestación E	0.0002 + 0.3741j	0.0689 + 1.0854j	0.0678 + 1.0844j
Escenario 9	Subestación E	74916010.2352 + 1.8939j	0.0021 + 0.1755j	0.0020 + 0.1754j

Para ambos escenarios de validación, se mantienen constantes tanto la separación entre electrodos (Gap) como la distancia de trabajo, con valores de 304.8 mm y 914.4 mm, respectivamente. Asimismo, el tiempo de despeje del arco se fija en 0.2 segundos, garantizando condiciones homogéneas para la comparación de resultados. El diagrama unifilar, los parámetros eléctricos y los datos de la Red Equivalente la Subestación E se muestran en el Anexo A.

A partir de los datos presentados en la tabla anterior, y considerando las mismas definiciones utilizadas para la generación de datos descritas en el numeral 3.1, se llevó a cabo el cálculo de la energía incidente en configuraciones con encerramiento empleando las metodologías IEEE Std. 1584 – 2018, EPRI y Terzija/Koglin. El objetivo fue comparar los resultados obtenidos bajo estas metodologías, considerando las siguientes configuraciones de electrodos: HCB, VCB y VCBB. Adicionalmente, se incorporó el cálculo utilizando la metodología de Ralph Lee, con el propósito de contrastar su comportamiento respecto a las metodologías alternativas mencionadas. A continuación, se detallan los resultados obtenidos a partir de este análisis comparativo:

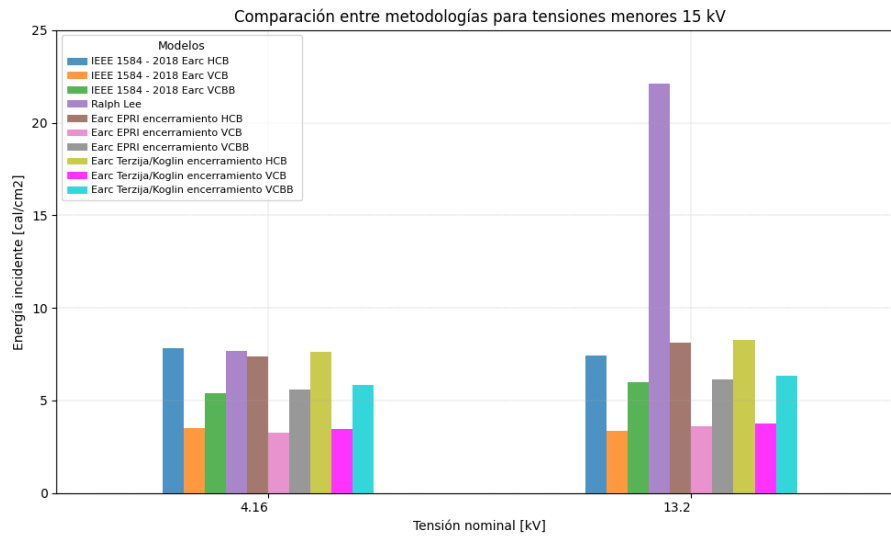


Figura 17. Comparación entre metodologías para tensiones menores a 15 kV

En la Figura 17 se observa que las metodologías alternativas empleadas logran estimaciones de energía incidente con una gran similitud con respecto a los valores obtenidos mediante la metodología IEEE Std. 1584 – 2018, para cada una de las configuraciones de electrodos analizadas. Estos resultados evidencian la capacidad del MAEI para predecir adecuadamente los factores de ajuste en niveles de tensión inferiores a 15 kV. Asimismo, se destaca la sobreestimación generada por la metodología de Ralph Lee en el Escenario 8, correspondiente a una tensión de 13.2 kV.

A continuación, se presenta una tabla con los resultados de energía incidente al aire libre calculados mediante las metodologías alternativas, junto con los factores de ajuste estimados por el MAEI para cada una de ellas:

Tabla 8. Resultados – MAEI en niveles de tensión inferiores a 15 kV

Escenario de estudio	Nivel de tensión [kV]	Configuración de electrodos	Earc EPRI aire libre [cal/cm²]	k _{3f} EPRI	Earc Terzija/Koglin aire libre [cal/cm²]	k _{3f} Terzija
Escenario 8	13.2	HCB	2.88	2.56	2.56	2.98
Escenario 8	13.2	VCB	2.88	1.13	2.56	1.35
Escenario 8	13.2	VCBB	2.88	1.94	2.56	2.28
Escenario 9	4.16	HCB	3.16	2.56	2.78	2.98
Escenario 9	4.16	VCB	3.16	1.13	2.78	1.35
Escenario 9	4.16	VCBB	3.16	1.94	2.78	2.28

4.3 Evaluación de desempeño del MAEI en niveles de tensión superiores a 15 kV

Con el propósito de evaluar el desempeño del MAEI en niveles de tensión superiores a 15 kV, se plantean dos instancias de validación. La primera consiste en comparar los resultados obtenidos mediante las metodologías alternativas con los del método de Ralph Lee, aplicados a escenarios representativos de un nivel de tensión de 33 kV, utilizando como referencia las características eléctricas de las subestaciones A, B, C, D y E previamente descritas en esta tesis. La segunda validación se basa en el contraste con resultados reportados en la literatura técnica, lo que permite evaluar el comportamiento del modelo frente a datos publicados.

4.3.1 Validación con subestaciones de 33 kV

En la siguiente tabla se presentan las características eléctricas de las subestaciones A, B, C, D y E en un nivel de tensión de 33 kV:

Tabla 9. Escenarios de estudio – MAEI en niveles de tensión superiores a 15 kV

Parámetros generales				
Escenario de estudio	Subestación	Nivel de tensión [kV]	Corriente de cortocircuito [kA]	Tipo de falla
Escenario 10	Subestación A	33	3.03	Trifásica
Escenario 11	Subestación B	33	16.23	Bifásica a tierra
Escenario 12	Subestación C	33	7.87	Trifásica
Escenario 13	Subestación D	33	17.30	Bifásica a tierra
Escenario 14	Subestación E	33	21.26	Bifásica a tierra
Impedancias de secuencia				
Escenario de estudio	Subestación	Z0 [Ω]	Z1 [Ω]	Z2 [Ω]
Escenario 10	Subestación A	4.7843 + 15.2222j	2.9397 + 6.2560j	2.9397 + 6.2560j
Escenario 11	Subestación B	0.2179 + 1.2565j	0.2786 + 1.2934j	0.2786 + 1.2934j
Escenario 12	Subestación C	0.6142 + 3.2625j	0.7262 + 2.5635j	0.7262 + 2.5635j
Escenario 13	Subestación D	0.0997 + 1.1206j	0.2359 + 1.3662j	0.2359 + 1.3662j
Escenario 14	Subestación E	0.0066 + 0.4117j	0.1344 + 2.1289j	0.1278 + 2.1231j

Para todos escenarios de validación, se mantienen constantes tanto la separación entre electrodos (Gap) como la distancia de trabajo, con valores de 304.8 mm y 914.4 mm, respectivamente. Asimismo, el tiempo de despeje del arco se fija en 0.2 segundos, garantizando condiciones homogéneas para la comparación de resultados.

La siguiente figura presenta los resultados obtenidos a partir de las metodologías de Ralph Lee, EPRI y Terzija/Koglin, considerando las configuraciones de electrodos HCB, VCB y VCBB.

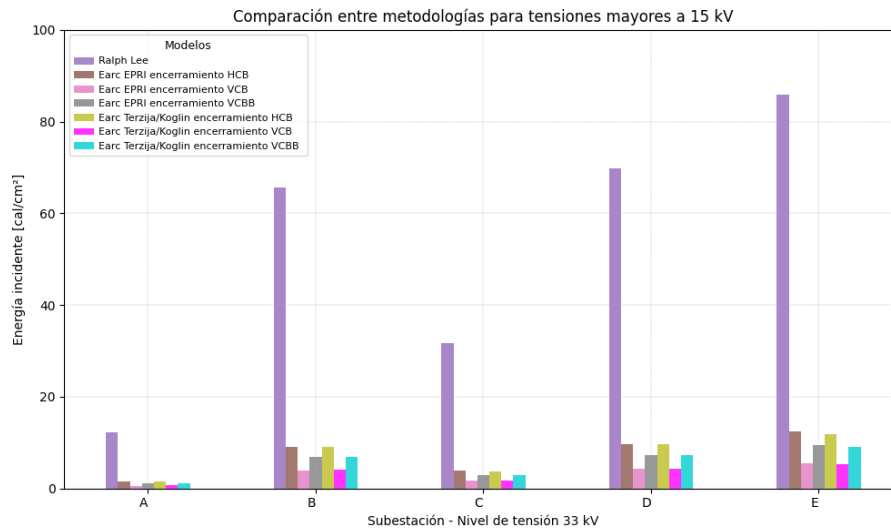


Figura 18. Comparación entre metodologías para tensiones mayores a 15 kV

Para niveles de tensión superiores a 15 kV, la metodología de Ralph Lee presenta estimaciones de energía incidente considerablemente más elevadas en comparación con las metodologías alternativas de EPRI y Terzija/Koglin. Este comportamiento evidencia un enfoque excesivamente conservador que podría derivar en una selección inadecuada de Equipos de Protección Personal (EPP). En contraste, los resultados obtenidos mediante EPRI y Terzija/Koglin muestran una alta concordancia entre sí, lo que respalda la precisión del MAEI en la predicción de factores de ajuste. Además, los valores estimados por estas metodologías tienden a representar de forma más realista el fenómeno del arco eléctrico en escenarios con altos niveles de tensión.

Los factores de ajuste estimados mediante el MAEI basado en la metodología EPRI, para los escenarios analizados, se encontraron en un rango de 1.13 a 2.56. En el caso del MAEI correspondiente a la metodología de Terzija/Koglin, los factores obtenidos oscilaron entre 1.35 y 2.98.

4.3.2 Validación con resultados de literatura técnica

En el artículo "High Voltage Arc Flash Assessment and Applications", desarrollado por Albert Marroquin, Abdur Rehman y Ali Madani [19], se presenta una comparación de los valores de energía incidente en configuraciones con encerramiento para niveles de tensión de 5 kV, 15 kV, 25 kV y 35 kV. Los resultados, ilustrados en la Fig. 6 del artículo mencionado [19], han sido tomados como referencia para el análisis comparativo realizado en este trabajo.

Para cada nivel de tensión se evalúa un valor distinto de separación entre electrodos (gap), asignado de la siguiente manera:

5 kV → 101.6 mm; 15 kV → 152.4 mm; 25 kV → 228.6 mm; 35 kV → 304.8 mm.

Las condiciones comunes consideradas para todos los casos son:

- Corriente de cortocircuito: 10 kA.

- Distancia de trabajo: 914.4 mm.
- Tiempo de despeje del arco: 0.2 segundos.

Para el cálculo de la energía incidente utilizando las metodologías EPRI y Terzija/Koglin, se adoptan las siguientes consideraciones:

- La corriente de arco (I_{arc}) se asume igual a la corriente de cortocircuito.
- Para la metodología EPRI, el valor del gap se unifica como sigue:
5 kV → 228.6 mm; 15 kV → 228.6 mm; 25 kV → 228.6 mm; 35 kV → 304.8 mm. Esta decisión se fundamenta en las inconsistencias observadas en los resultados de EPRI cuando el gap es inferior a 200 mm.
- El factor de ajuste estadístico (k) utilizado en las ecuaciones de EPRI se define según lo expuesto en el numeral 3.1.6.
- El gradiente de tensión (B), necesario para las ecuaciones de Terzija/Koglin, se establece conforme a lo indicado en el numeral 3.1.3.
- El coeficiente de exponente de distancia (x) en la metodología de Terzija/Koglin se adopta de acuerdo con lo establecido en el numeral 3.1.6.

La siguiente figura muestra la comparación de los resultados obtenidos mediante las metodologías alternativas, incluyendo la propuesta de implementación del MAEI, frente a valores de referencia extraídos de la literatura técnica. Las barras etiquetadas con "Fig. 6" corresponden a datos publicados en la Figura 6 del artículo [19], y representan las metodologías IEEE 2002, ArcPro, EPRI y Terzija Koglin. Por otro lado, los resultados asociados al MAEI se identifican con las configuraciones de encerramiento HCB, VCB y VCBB, las cuales hacen referencia al tipo de electrodo considerado en el cálculo de la energía incidente para cada metodología alternativa.

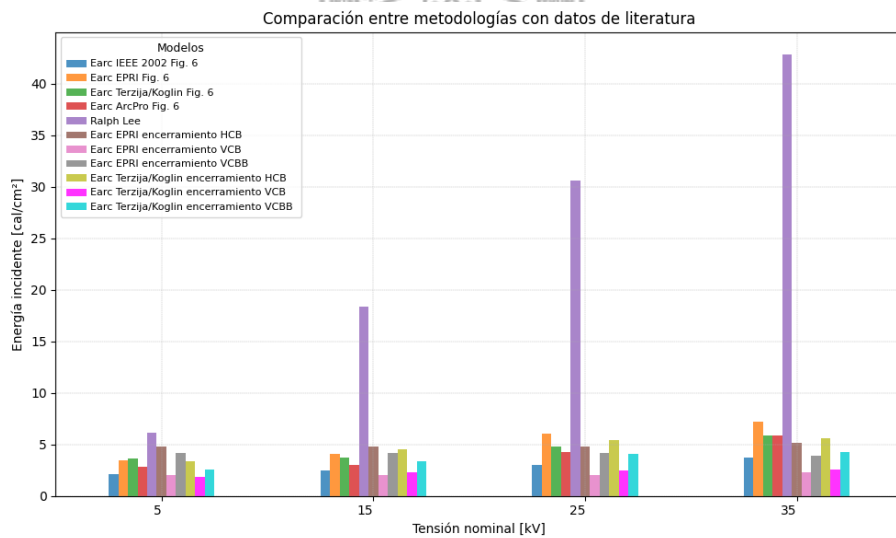


Figura 19. Comparación entre metodologías con datos de literatura técnica

A partir del análisis de la Figura 19, se evidencia una alta concordancia entre los resultados obtenidos a partir de la literatura y aquellos calculados mediante las metodologías alternativas, utilizando los factores de ajuste generados por el MAEI. En particular, los valores de energía incidente estimados con las metodologías de EPRI y Terzija/Koglin bajo configuración HCB presentan la mayor cercanía con los resultados reportados en la Figura 6 del artículo [19], los cuales incluyen los métodos de EPRI, Terzija/Koglin y ArcPro.

Este comportamiento sugiere que los factores de ajuste utilizados en [19] fueron seleccionados con un enfoque conservador, buscando asegurar una protección eficaz frente a eventos de arco eléctrico. No obstante, dicha selección evitó incurrir en sobreestimaciones, como las comúnmente observadas en la metodología de Ralph Lee.

A partir de los análisis realizados, tanto en las subestaciones de 33 kV como en los resultados extraídos de la literatura, se concluye que las metodologías EPRI y Terzija/Koglin pueden ser aplicadas eficazmente para el cálculo de la energía incidente en encerramientos, utilizando factores de ajuste estimados mediante el modelo MAEI. No obstante, dado que el MAEI no garantiza predicciones completamente precisas en todos los casos, se recomienda incorporar una sobreestimación adicional del 10 % como margen de seguridad. Esta recomendación se fundamenta en los resultados obtenidos durante la validación del modelo de redes neuronales, donde se evidenció una desviación cercana al 10 % al comparar las métricas MAE y RMSE respecto a la media de los datos analizados.



5. CONCLUSIONES Y FUTUROS DESARROLLOS

A continuación, se presentan las conclusiones generales derivadas del análisis realizado en esta tesis.

5.1 Conclusiones generales

- Se describieron las ecuaciones y variables de entrada necesarias para aplicar las metodologías IEEE Std. 1584 – 2018, Ralph Lee, EPRI y Terzija/Koglin en la estimación de energía incidente.
- Se definió e implementó una metodología orientada al desarrollo de un Modelo de Ajuste de Energía Incidente (MAEI) basada en técnicas de Machine Learning que permitiera realizar la predicción de factores de ajuste.
- Se determinaron escenarios para el entrenamiento y validación de los modelos de Machine Learning a partir de subestaciones con celdas tipo AIS a niveles de tensión inferiores a 15 kV.
- Se analizó el comportamiento entre las variables de interés planteadas en los escenarios de entrenamiento y validación mediante evaluación de correlaciones. En este análisis se logró evidenciar que las variables con mayor relevancia en la estimación de factores de ajuste son la distancia de trabajo, la configuración de electrodos y la distancia entre electrodos (gap).
- Se identificó que las metodologías alternativas de ERPI y Terzija/Koglin arrojan resultados incoherentes con una distancia entre electrodos (gap) menor a 200 mm.
- El Coeficiente de encerramiento (Cenc) definido en el numeral 3.1.3 a partir de las ecuaciones de cálculo energía incidente al aire libre y en encerramientos, mostró resultados por debajo de 1 para algunas configuraciones de valores entre la distancia de trabajo y los coeficientes de reflectividad de Wilkins. Este tipo de resultados demuestran incoherencia, ya que la energía incidente en encerramientos es mayor que en al aire libre a causa de la reflexión de energía causada en el entorno cerrado. Por tal motivo, no se recomienda su uso sin un previo análisis que determine un resultado de energía incidente confiable en encerramientos.

- Se entrenaron y evaluaron distintos modelos de Machine Learning desarrollados para la predicción de factores de ajuste, correspondientes a modelos basados en ElasticNet, regresión polinómica de grado 2 y redes neuronales. Este último fue seleccionado como Modelo de Ajuste de Energía Incidente (MAEI), debido al desempeño alcanzado durante la etapa de validación.
- Se validó el desempeño del MAEI en niveles de tensión inferiores a 15 kV a partir de la comparación de resultados entre las metodologías IEEE Std. 1584 – 2018, Ralph Lee, EPRI y Terzija/Koglin. Los resultados obtenidos evidenciaron un desempeño adecuado del MAEI en la predicción de factores de ajuste.
- Se validó el desempeño del MAEI en niveles de tensión superiores a 15 kV, contrastando resultados con la metodología de Ralph Lee y resultados reportados en la literatura. En este análisis se logró evidenciar un desempeño aceptable en la predicción de valores con el MAEI, alcanzado valores similares a los establecidos en la literatura técnica. Sin embargo, se recomienda aplicar una sobreestimación adicional del 10 %, dado que durante la etapa de validación se observó una desviación de este orden al comparar las métricas MAE y RMSE con respecto al valor medio de los datos analizados.
- Se realizó un análisis comparativo entre las metodologías de cálculo de energía incidente IEEE 1584 – 2018, EPRI y Terzija/Koglin en celdas AIS, en el cual, a través del uso de herramientas estadísticas, fue posible identificar las variables más relevantes para el desarrollo de modelos de Machine Learning orientados a la predicción de factores de ajuste. Estos factores permiten escalar los valores de energía incidente obtenidos para condiciones de arcos monofásicos al aire libre hacia escenarios de arcos trifásicos en ambientes encerrados.

5.2 Aportes

- Se obtuvieron deficiencias en el uso de la ecuación (11) para el cálculo de energía incidente en encerramientos a partir de la aplicación de las metodologías alternativas EPRI y Terzija/Koglin.
- Se integró el uso de modelos de Machine Learning en la predicción de factores de ajuste, abordando una de las principales debilidades de las metodologías alternativas, que corresponde a la falta de un criterio claro para la selección de dichos factores.

- Se desarrollaron scripts en Python integrados en cuadernos de Google Colab, lo que facilita el análisis automatizado de las relaciones entre metodologías, la detección de valores atípicos y la comparación de resultados de energía incidente. Esta implementación optimiza el procesamiento de datos y proporciona una herramienta accesible y reproducible para fundamentar los valores calculados en estudios de arco eléctrico, mejorando la trazabilidad y la validación de los resultados.
- Se validó el uso de metodologías alternativas para el cálculo de energía incidente aplicables a niveles de tensión superiores a 15 kV, lo que permite una selección más adecuada de los equipos de protección personal (EPP), representando una ventaja significativa en comparación al uso del método de Ralph Lee, el cual tiende a sobreestimar los valores de energía incidente al punto de obligar una desenergización de las celdas en tareas de mantenimiento. Por lo tanto, realizar los cálculos con las metodologías alternativas reduce pérdidas económicas asociadas a interrupciones operativas innecesarias sin sacrificar la seguridad para el operario.

5.3 Futuros desarrollos

Considerando la capacidad de mejora de los análisis realizados y las necesidades encontradas durante el mismo, se plantean los siguientes desarrollos futuros para complemento del estado del arte en el estudio del arco eléctrico:

- **Ampliación y optimización del modelo de redes neuronales:** Desarrollar un modelo basado en redes neuronales utilizando un conjunto de datos más extenso y representativo, abarcando una mayor variabilidad amplio de niveles de tensión. Esto permitirá una evaluación más precisa de la influencia de la tensión nominal en el cálculo de energía incidente y en la determinación de factores de ajuste para metodologías alternativas.
- **Incorporación de criterios más específicos en normativas internacionales:** Proponer y establecer criterios en normativas internacionales de seguridad eléctrica que proporcionen mayor claridad y precisión en el cálculo de energía incidente en celdas y tableros con niveles de tensión superiores a 15 kV.
- **Validación experimental en celdas y encerramientos de alta tensión:** Realizar pruebas experimentales en celdas y encerramientos con niveles de tensión superiores a 15 kV que permitan definir metodologías más confiables, similares a la establecida en la IEEE Std. 1584-2018.

- **Reducción de energía incidente mediante estudios de coordinación de protecciones:** Implementar estrategias de mitigación de energía incidente a través de la reducción del tiempo de despeje de fallas en estudios de coordinación de protecciones. El objetivo es minimizar la duración del arco eléctrico y, con ello, reducir significativamente la energía incidente, mejorando la seguridad del personal y optimizando la selección de equipos de protección personal (EPP).
- **Mejora de equipos de protección personal (EPP):** Investigar y mejorar los materiales y diseños de los EPP utilizados en labores de mantenimiento con riesgo de arco eléctrico, que permitan tener mayor movilidad y soportar una energía incidente alta a través de una mayor capacidad de disipación térmica.



6. REFERENCIAS BIBLIOGRÁFICAS

- [1] R. H. Lee, "The Other Electrical Hazard: Electric Arc Blast Burns," in *IEEE Transactions on Industry Applications*, vol. IA-18, no. 3, pp. 246-251, May 1982, doi: 10.1109/TIA.1982.4504068.
- [2] R. F. Ammerman, T. Gammon, P. K. Sen and J. P. Nelson, "Comparative study of arc modeling and arc flash incident energy exposures," *2008 55th IEEE Petroleum and Chemical Industry Technical Conference*, Cincinnati, OH, USA, 2008, pp. 1-12, doi: 10.1109/PCICON.2008.4664001.
- [3] A. D. Stokes and D. K. Sweeting, "Electric arcing burn hazards," in *IEEE Transactions on Industry Applications*, vol. 42, no. 1, pp. 134-141, Jan.-Feb. 2006, doi: 10.1109/TIA.2005.861911.
- [4] A. D. Stokes and D. K. Sweeting, "Closure to discussions of "Electric arcing burn hazards"," in *IEEE Transactions on Industry Applications*, vol. 42, no. 1, pp. 146-147, Jan.-Feb. 2006, doi: 10.1109/TIA.2005.861908.
- [5] ABB, "Protección contra arcos eléctricos Integración entre el dispositivo Monitor de arco interno (TVOC-2) y los interruptores de bastidor abierto Emax". ABB, Tech. Publication ID. 1TXA007446G0701 REV. A, 2012. Acceso: Feb. 15, 2025. [En línea]. Disponible: <https://search.abb.com/library/Download.aspx?DocumentID=1TXA007446G0701&LanguageCode=es&DocumentPartId=&Action=Launch>
- [6] NFPA 70E, *Standard for Electrical Safety in the Workplace*, 2024.
- [7] "IEEE Guide for Performing Arc-Flash Hazard Calculations," in *IEEE Std 1584-2018 (Revision of IEEE Std 1584-2002)*, vol., no., pp.1-134, 30 Nov. 2018, doi: 10.1109/IEEESTD.2018.8563139.
- [8] Data Carpentry. "Data Analysis and Visualization in Python for Ecologists". *Datacarpentry.org*. Acceso: Feb. 15, 2025. [En línea]. Disponible: <https://datacarpentry.github.io/python-ecology-lesson/index.html>
- [9] S. Navarro. "¿Qué son los datasets? [4 sitios donde encontrarlos]". *Keepcoding.io*. Acceso: Feb. 15, 2025. [En línea]. Disponible: <https://keepcoding.io/blog/que-son-datasets/>
- [10] IEC 62271-200, *High-voltage switchgear and controlgear - Part 200: AC metal-enclosed switchgear and controlgear for rated voltages above 1 kV and up to and including 52 kV*, 2021.
- [11] Química.es. "Enciclopedia de química". *Quimica.es*. Acceso: Feb. 15, 2025. [En línea]. Disponible: <https://www.quimica.es/enciclopedia/>
- [12] RETIE, *Reglamento Técnico de Instalaciones Eléctricas*, 2024.
- [13] A. Roldán. "Qué es un script: definición, uso y guía para crearlo". *Aplicaciones-ai*. Acceso: Feb. 15, 2025. [En línea]. Disponible: <https://aplicaciones-ai.com/que-es-un-script/>

- [14] "IEEE Guide for Performing Arc Flash Hazard Calculations," in *IEEE Std 1584-2002*, vol., no., pp.1-113, 23 Sept. 2002, doi: 10.1109/IEEESTD.2002.94138.
- [15] EPRI, "Arc Flash Issues in Transmission and Substation Environments: Results from Tests with Longs Arcs". Electric Power Research Institute, TR-1022632, Palo Alto, CA, 2011.
- [16] V. V. Terzija and H. . -J. Koglin, "New approach to arc resistance calculation," *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194)*, Columbus, OH, USA, 2001, pp. 781-787 vol.2, doi: 10.1109/PESW.2001.916958.
- [17] V. V. Terzija and H. . -J. Koglin, "Long arc in free air: testing, modelling and parameter estimation. I," *Ninth International Conference on Harmonics and Quality of Power. Proceedings (Cat. No.00EX441)*, Orlando, FL, USA, 2000, pp. 404-409 vol.2, doi: 10.1109/ICHQP.2000.897714.
- [18] V. V. Terzija and H. . -J. Koglin, "Long arc in free air: testing, modelling and parameter estimation. II," *Ninth International Conference on Harmonics and Quality of Power. Proceedings (Cat. No.00EX441)*, Orlando, FL, USA, 2000, pp. 481-486 vol.2, doi: 10.1109/ICHQP.2000.897726.
- [19] A. Marroquin, A. Rehman and A. Madani, "High Voltage Arc Flash Assessment and Applications – Part 1," *The InterNational Electrical Testing Association Journal*, June 01, 2019. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://netaworldjournal.org/high-voltage-arc-flash-assessment-and-applications-part-1/>
- [20] R. Wilkins, "Simple Improved Equations for Arc Flash Hazard Analysis," *IEEE Electrical Safety Forum*, Aug. 30, 2004.
- [21] NESC C2, *National Electrical Safety Code*, 2023.
- [22] M. A. Acevedo y K. E. Vargas, "Machine Learning: algoritmos de clasificación y sus aplicaciones en el análisis de datos", Tesis de licenciatura, FES Acatlán, UNAM, Naucalpan de Juárez, México, 2017. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://ru.dgb.unam.mx/bitstream/20.500.14330/TES01000767007/3/0767007.pdf>
- [23] UN-GCPDS, *Curso de Inteligencia Artificial CHEC*, GitHub, 2024. Acceso: Feb. 16, 2025. [En línea]. Disponible: https://github.com/UN-GCPDS/curso_IA_CHEC
- [24] InteractiveChaos. "Tutorial de Machine Learning". Interactivechaos.com. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://interactivechaos.com/es/course/tutorial-de-machine-learning>
- [25] Scikit-learn. "Linear Models". *Scikit-learn*. Acceso: Feb. 16, 2025. [En línea]. Disponible: https://scikit-learn.org/stable/modules/linear_model.html
- [26] IBM. "Regresión lineal de Elastic Net". *IBM Corporation*, Sept. 30, 2024. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://www.ibm.com/docs/es/spss-statistics/saas?topic=features-linear-elastic-net-regression>

- [27] J. I. von Lücken, "Métodos de Regularización Lasso, Ridge y Elastic Net: Una aplicación a los seguros de no vida", Trabajo final de máster, Facultad de Economía y Empresa, UB, Barcelona, España, 2021. Acceso: Abr. 14, 2025. [En línea]. Disponible: https://diposit.ub.edu/dspace/bitstream/2445/179200/1/TFM-CAF_L%C3%BCckenGim%C3%A9nez_2021.pdf
- [28] El Laberinto de Falken, "Regresión Polinómica (Polinomial)," *El Laberinto de Falken*, Dec. 05, 2019. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://www.ellaberintodefalken.com/2019/12/regresion-polinomica-polinomial.html>
- [29] L. N. Kanal, "Perceptron," in *Encyclopedia of Computer Science*, 2003, pp.1383–1385. [En línea]. Disponible: <https://dl.acm.org/doi/pdf/10.5555/1074100.1074686>
- [30] D. J. Matich, "Redes Neuronales: Conceptos Básicos y Aplicaciones", Cátedra: Informática Aplicada a la Ingeniería de Procesos – Orientación I, Dept. de Ingeniería Química, UTN, Rosario, Argentina, 2001. Acceso: Abr. 14, 2025. [En línea]. Disponible: https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monografias/matich-redesneuronales.pdf
- [31] TensorFlow. "Sequential model". *TensorFlow*, Jan. 01, 2022. Acceso: Feb. 16, 2025. [En línea]. Disponible: https://www.tensorflow.org/guide/keras/sequential_model?hl=es-419
- [32] I. Goodfellow, Y. Bengio y A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016. Acceso: Abr. 14, 2025. [En línea]. Disponible: <https://www.deeplearningbook.org/>
- [33] E. D. Angulo, "Entrenamiento de modelos de clasificación con aprendizaje federado preservando la privacidad de los datos", Monografía de maestría, Dept. de Ingeniería de Sistemas y Computación, UniNorte, Barranquilla, Colombia, 2022. Acceso: Abr. 14, 2025. [En línea]. Disponible: <https://manglar.uninorte.edu.co/bitstream/handle/10584/11704/1140897206.pdf>
- [34] A. P. Strom, "Long 60-Cycle Arcs in Air," in *Transactions of the American Institute of Electrical Engineers*, vol. 65, no. 3, pp. 113–118, March 1946, doi: 10.1109/T-AIEE.1946.5059312.
- [35] IEC 60909-0, *Short-circuit currents in three-phase a.c. systems – Part 0: Calculation of currents*, 2016.
- [36] C. C. M. Vara de Rey, *Coeficiente de correlación lineal de Pearson*, material de la asignatura Análisis de Datos en Psicología II, Universidad de Sevilla. [En línea]. Disponible: <https://personal.us.es/vararey/adatos2/correlacion.pdf>
- [37] A. Villegas, *15 kV Arc Flash*, GitHub, 2025. Acceso: Feb. 16, 2025. [En línea]. Disponible: https://github.com/alvillegasru/15_kV_Arc_Flash
- [38] C. Salazar y S. Del Castillo, *Fundamentos Básicos de Estadística*, 1ª ed. Portoviejo, Ecuador: Universidad San Gregorio de Portoviejo, 2018. Acceso: Feb. 16, 2025. [En línea]. Disponible: <http://librodigital.sangregorio.edu.ec/librosusgp/B0009.pdf>

- [39] D. R. Doan, "Arc Flash Calculations for Exposures to DC Systems," in *IEEE Transactions on Industry Applications*, vol. 46, no. 6, pp. 2299-2302, Nov.-Dec. 2010, doi: 10.1109/TIA.2010.2070480.
- [40] Brady. "La guía para seguridad de arco eléctrico en el área de trabajo". *BradyID*. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://www.bradyid.com.mx/formularios-de-descarga/guia-de-arco-electrico>
- [41] Y. Goda, M. Iwata, K. Ikeda and S. Tanaka, "Arc voltage characteristics of high current fault arcs in long gaps," in *IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 791-795, April 2000, doi: 10.1109/61.853021.
- [42] B. Underwood y K. Mara, "Distribution System Arc-Flash Calculation Case Studies," Informe CRN08-01, Cooperative Research Network, National Rural Electric Cooperative Association, Arlington, VA, USA, 2008. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://www.cooperative.com/programs-services/bts/Documents/Reports/ArcFlashCalculationCaseStudies.pdf>
- [43] "IEEE Guide for Testing Switchgear Rated up to 52 kV for Internal Arcing Faults—Corrigendum 1," in *IEEE Std C37.20.7-2017/Cor 1-2021 (Corrigendum to IEEE Std C37.20.7-2017)*, vol., no., pp.1-14, 2 April 2021, doi: 10.1109/IEEESTD.2021.9393650.
- [44] R. Wilkins, M. Allison and M. Lang, "Calculating hazards," in *IEEE Industry Applications Magazine*, vol. 11, no. 3, pp. 40-48, May-June 2005, doi: 10.1109/MIA.2005.1423901.
- [45] A. Marroquin, A. Rehman and A. Madani, "High Voltage Arc Flash Assessment and Applications – Part 2," *The InterNational Electrical Testing Association Journal*, Sept. 01, 2019. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://netaworldjournal.org/high-voltage-arc-flash-assessment-and-applications-part-1/>
- [46] E. A. Cano-Plata, A. J. Ustariz-Farfan and O. J. Soto-Marin, "Electric Arc Furnace Model in Distribution Systems," in *IEEE Transactions on Industry Applications*, vol. 51, no. 5, pp. 4313-4320, Sept.-Oct. 2015, doi: 10.1109/TIA.2015.2429638.
- [47] E. A. Cano-Plata *et al.*, "Hybrid Simulation of Power Quality Assessment: An Application for Power Ground Grid in Arc Furnace Systems," in *IEEE Transactions on Industry Applications*, vol. 52, no. 3, pp. 2648-2653, May-June 2016, doi: 10.1109/TIA.2016.2516501.
- [48] P. A. Loaiza, "Marco comparativo entre la norma IEEE 1584 y la norma NFPA 70E", Tesis de maestría, Dept. de Ingeniería Eléctrica, Electrónica y Computación, UNAL, Manizales, Caldas, 2018. Acceso: Feb. 16, 2025. [En línea]. Disponible: <https://repositorio.unal.edu.co/bitstream/handle/unal/69568/1053812049.2018.pdf>
- [49] R. Wilkins, M. Allison and M. Lang, "Improved method for arc flash hazard analysis," *Conference, 2004 IEEE Industrial and Commercial Power Systems Technical*, Clearwater Beach, FL, USA, 2004, pp. 55-62, doi: 10.1109/ICPS.2004.1314982.
- [50] V. V. Terzija and H. . -J. Koglin, "Testing, modeling and simulation of long arc in still air," *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat.*

- No.01CH37194), Columbus, OH, USA, 2001, pp. 1140-1145 vol.3, doi: 10.1109/PESW.2001.917233.
- [51] M. L. Eblen, T. A. Short and W. -J. Lee, "Medium-Voltage Arc Flash in Switchgear and Live-Front Transformers," in *IEEE Transactions on Industry Applications*, vol. 52, no. 6, pp. 5280-5288, Nov.-Dec. 2016, doi: 10.1109/TIA.2016.2598300.
- [52] M. D. Fontaine and P. E. P. Walsh, "DC arc flash calculations — Arc-in-open-air & arc-in-a-box — Using a simplified approach (Multiplication factor method)," *2012 IEEE IAS Electrical Safety Workshop*, Daytona Beach, FL, 2012, pp. 1-8, doi: 10.1109/ESW.2012.6165544.
- [53] V. Terzija, G. Preston, M. Popov and N. Terzija, "New Static "AirArc" EMTP Model of Long Arc in Free Air," in *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1344-1353, July 2011, doi: 10.1109/TPWRD.2010.2086082.
- [54] R. L. Doughty, T. E. Neal and H. L. Floyd, "Predicting incident energy to better manage the electric arc hazard on 600-V power distribution systems," in *IEEE Transactions on Industry Applications*, vol. 36, no. 1, pp. 257-269, Jan.-Feb. 2000, doi: 10.1109/28.821823.
- [55] E. Hoagland, M. Golovkov, C. Maurice and P. Clark, "Clothing in Arc Flash: Four Types of Arc Exposure and the Effect of Moisture on Garment Appearance After an Arc Event," in *IEEE Transactions on Industry Applications*, vol. 50, no. 4, pp. 2370-2374, July-Aug. 2014, doi: 10.1109/TIA.2014.2300414.
- [56] M. Golovkov and E. H. Hoagland, "Arc flash testing update: Effect of arc electrode geometry and distance on cotton shirt ignition," *2014 IEEE IAS Electrical Safety Workshop*, San Diego, CA, USA, 2014, pp. 1-8, doi: 10.1109/ESW.2014.6766905.
- [57] R. F. Ammerman, T. Gammon, P. K. Sen and J. P. Nelson, "DC-Arc Models and Incident-Energy Calculations," in *IEEE Transactions on Industry Applications*, vol. 46, no. 5, pp. 1810-1819, Sept.-Oct. 2010, doi: 10.1109/TIA.2010.2057497.