



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Juan Camilo Cardona Posada

Universidad Nacional de Colombia
Facultad de Minas, Departamento de Energía Eléctrica y Automática
Medellín, Colombia
2021

Detección y clasificación de fallas eléctricas en sistemas de distribución de energía eléctrica mediante el uso de la Transformada Wavelet Continua y Funciones Madre de Soporte Infinito

Juan Camilo Cardona Posada

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de:

Magister en Ingeniería Eléctrica

Director (a):

Ph.D Freddy Bolaños Martínez

Codirector (a):

Ph.D Ernesto Pérez Gonzales

Línea de Investigación:

Protección de Sistemas Eléctricos de Potencia

Grupo de Investigación:

GAUNAL

Universidad Nacional de Colombia

Facultad de Minas, Departamento de Energía Eléctrica y Automática

Medellín, Colombia

2021

A mis padres, cuyo apoyo incondicional me permitió perseverar en la escritura de esta disertación. A mí hermana, cuya crianza me dio el carácter que poseo y el deseo de crecimiento personal que me llevó a iniciar esta maestría.

Al docente Freddy Bolaños Martínez, cuya dirección académica demostrada en sus enseñanzas y calidad humana, exaltada en sus diarias publicaciones de humor, moldearon la persona que soy hoy y permitieron la entrega de este trabajo.

La presión es un privilegio.

Christopher Bumstead.

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

Juan Camilo Cardona Posada.

Fecha 18/02/2022.

Agradecimientos

Quisiera formalmente agradecer al docente Ernesto Pérez González (Ph.D) por su apoyo como codirector de esta disertación y su ayuda en aspectos específicos relacionados a simulaciones y conceptos de sistemas de protección eléctricas y a David Alejandro Sánchez Muñoz, M.Sc. por su colaboración en la automatización de las simulaciones de fallas eléctricas.

Por otro lado, al docente Freddy Bolaños Martínez (Ph.D) por su labor como director de tesis la cual se reflejó en un apoyo incondicional durante todo el transcurso de la investigación y por haberme formado en el método científico con su experiencia y carisma. También le agradezco por su ayuda en el desarrollo de códigos de MATLAB, revisión de literatura y la definición de la metodología que se presenta en esta disertación.

También quisiera agradecer al Proyecto "Estrategia de transformación del sector energético colombiano en el horizonte de 2030" financiado en la convocatoria 778 de Colciencias Ecosistema Científico mediante el Contrato FP44842-210-2018 por darme la oportunidad de involucrarme en la transformación energética del país, desarrollando mi sentido investigativo y apoyándome económicamente.

Finalmente, a mi Alma Mater, la Universidad Nacional de Colombia, la cual me dio la oportunidad de cursar esta maestría por medio de una beca de excelencia de pregrado. Sin esta oportunidad, probablemente no habría tomado este reto académico y personal del cual he hallado tanta satisfacción y crecimiento personal.

Resumen

Detección y clasificación de fallas eléctricas en sistemas de distribución de energía eléctrica mediante el uso de la Transformada Wavelet Continua y Funciones Madre de Soporte Infinito

***Descripción:** El análisis de fallas eléctricas en Sistemas Eléctricos de Potencia ha visto un avance significativo en la implementación de metodologías fundamentadas en el Procesamiento Digital de Señales, como el análisis en tiempo y frecuencia. En este trabajo se propone una metodología para utilizar filtros con Respuesta Infinita al Impulso como Funciones Madre para computar la Transformada Wavelet Continua mediante la evaluación del concepto del remuestreo digital de señales. Esta metodología es evaluada al compararla con la metodología más comúnmente reportada en la literatura (Transformada Wavelet Discreta con Función Madre tipo Daubechies 4) y se investigan los efectos de los parámetros más relevantes inherentes al método y a las señales.

La implementación del código se realiza en el software de MATLAB y los resultados se validan comparando la energía espectral asociada a los coeficientes Wavelet obtenidos. Por otro lado, se propone evaluar las etapas de detección y clasificación de fallas eléctricas en sistemas de distribución de energía eléctrica utilizando una serie de señales simuladas en el Software ATP/EMTP y redes neuronales artificiales. Los resultados obtenidos validan la metodología propuesta además de presentar una mejora notoria en cuanto a la tasa de detección y clasificación de eventos al compararla con el método tradicional. Finalmente, se resumen las limitaciones de la investigación y se propone una serie de recomendaciones a modo de trabajo futuro con el objetivo de continuar la evaluación de la metodología propuesta.

Palabras clave: Transformada Wavelet, Análisis de Fallas Eléctricas, Sistemas de Distribución de Energía Eléctrica, Funciones Madre, Respuesta al Impulso de Filtros Digitales.

Abstract

Electric Fault detection and Classification in Power Distribution Systems using the Continuous Wavelet Transform and Infinite Support Mother Functions

***Description:** Electric fault analysis in Power Systems has seen significant progress in the adoption of methodologies based on Digital Signal Processing, such as time and frequency analysis. In this work, a novel method is proposed to use filters with Infinite Impulse Response as Mother Functions to compute the Continuous Wavelet Transform by evaluating the concept of digital signal resampling. This methodology is evaluated by comparing it with the most common method in the literature (Discrete Wavelet Transform with Daubechies 4-type Mother Function) and the effects of the most relevant parameters inherent to the method and to the signals are investigated.

The implementation of the code is carried out in the MATLAB software and the results are validated by comparing the spectral energy associated with the Wavelet Coefficients obtained. On the other hand, an evaluation of the stages of detection and classification of electrical faults in electrical power distribution systems is proposed by using a series of simulated signals in the ATP/EMTP Software and artificial neural networks. The results obtained validate the proposed methodology in addition to presenting a notable improvement in the rate of detection and classification of events when compared to the traditional method. Finally, the limitations of the research are summarized, and a series of recommendations are proposed for future work with the aim of continuing the evaluation of the proposed method.

Keywords: Wavelet Transform, Electric Fault Analysis, Power Distribution Systems, Mother Functions, Digital Filter Impulse Response.

Contenido

	Pág.
1. Introducción	1
2. Estado del arte y antecedentes	5
2.1 Detección y clasificación de fallas eléctricas en sistemas de distribución de energía eléctrica	5
2.2 La Transformada Wavelet	12
2.3 Funciones Madre: Observaciones desde la teoría de filtros digitales.....	22
2.4 Extracción de características.....	30
2.4.1 Densidad de Potencia Espectral (PSD)	30
2.4.2 Entropía Wavelet	31
2.4.3 Parámetros estadísticos	33
2.4.4 Máximos locales y Módulo Máximo Wavelet.....	33
2.4.5 Modos de correlación Wavelet.....	36
2.5 Sistemas de detección y clasificación de eventos	39
2.5.1 Umbrales de detección y clasificación	39
2.5.2 Algoritmos de clasificación basados en Machine Learning e Inteligencia Artificial	41
3. Propuesta metodológica.....	45
3.1 Preprocesamiento digital de la señal.....	46
3.2 Análisis Multirresolución utilizando filtros digitales con respuesta infinita al impulso	50
3.2.1 Enventanado de la señal a analizar	56
3.2.2 Definición del filtro o función Madre	56
3.2.3 Definición de la resolución en tiempo y frecuencia del método	57
3.2.4 Remuestreo de las ventanas	58
3.2.5 Computación y normalización de coeficientes	60
3.3 Extracción de características de los coeficientes Wavelet.....	61
3.4 Detección y clasificación de eventos.....	63
4. Resultados y discusión	74
4.1 Simulación de eventos a detectar y clasificar	74
4.2 Validación de la metodología	77
4.3 Sistemas de detección y clasificación.....	102
5. Conclusiones y recomendaciones.....	113
5.1 Conclusiones.....	113
5.2 Recomendaciones.....	114

XIV Detección y clasificación de fallas eléctricas en sistemas de distribución de
energía eléctrica mediante el uso de la Transformada Wavelet Continua y
Funciones Madre de Soporte Infinito

Lista de figuras

	Pág.
Figura 2-1: Resolución en tiempo y frecuencia para las series de tiempo (izquierda) y la Transformada de Fourier (derecha).....	9
Figura 2-2: Resolución en tiempo y frecuencia para la Transformada de Fourier de Tiempo Corto.	11
Figura 2-3: Resolución en tiempo y frecuencia para la Transformada Wavelet.....	16
Figura 2-4: Algoritmo Piramidal para MRA de señales utilizando la DWT.....	18
Figura 2-5: Escalograma para una senoide con frecuencias de 100Hz y 50Hz y una componente de ruido blanco.	20
Figura 2-6: Wavelets de Daubechies de ordenes 2, 3 y 4.....	23
Figura 2-7: Coiflets de ordenes 1, 2 y 3.....	24
Figura 2-8: Wavelet de Morlet real e imaginaria.	25
Figura 2-9: Representación de una señal en el tiempo (arriba) mediante la WT en un escalograma en escala de grises (medio) y los esqueletos correspondientes al WMM (abajo).....	35
Figura 2-10: Ejemplo de árboles de decisión para detección (izquierda) y clasificación (derecha) de fallas.....	40
Figura 2-11: Ejemplo de Random Forest con N árboles de decisiones.....	42
Figura 2-12: Hiperplano óptimo para clasificación utilizando SVM.....	43
Figura 3-1: Diagrama de flujo de la propuesta metodológica. Elaboración propia.....	45
Figura 3-2: Ventanas de longitud fija con traslapo del 0%.	46
Figura 3-3: Ventanas de longitud fija con traslapo del 50%.	46
Figura 3-4: Ventanas de longitud fija con traslapo del 100%.	47
Figura 3-5: Representación en el dominio del tiempo (arriba) y de la frecuencia (abajo) para las funciones de enventanado rectangular, triangular y de Hann, de izquierda a derecha.	49
Figura 3-6: Respuesta a una secuencia impulsiva de entrada para un filtro FIR (arriba) y un filtro IIR (abajo).....	52
Figura 3-7: Respuesta en fase para un filtro FIR de orden 6 (izquierda) y un filtro IIR de orden 16 (derecha).....	53
Figura 3-8: Respuesta en magnitud para filtros IIR resonantes con diferentes radios de convergencia.....	55
Figura 3-9: Representación de ANN tipo Feedforward.	64
Figura 3-10: Representación gráfica del concepto de Bias en las neuronas de una ANN.	67
Figura 3-11: Vista general del modo de operación de una neurona en una ANN.....	68
Figura 3-12: Funciones de activación Sigmoidal, Gaussiana, ReLu y SoftMax.....	69
Figura 3-13: Esquema de ANN superficiales (izquierda) y profundas (derecha).	71
Figura 4-1: Diagrama unifilar de sistema de prueba IEEE 9 barras.	74

Figura 4-2: Modelo de simulación del sistema de prueba IEEE 9 barras implementado en ATP/EMTP.....	75
Figura 4-3: Señal sinusoidal con armónico (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 4800Hz.....	78
Figura 4-4: Señal sinusoidal con armónico (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 48000Hz.....	78
Figura 4-5: Señal de falla simulada (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 4800Hz.	79
Figura 4-6: Señal de falla simulada (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 48000Hz.	80
Figura 4-7: Escalograma obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 4800 Hz.....	81
Figura 4-8: Escalograma obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 48000 Hz.....	81
Figura 4-9: Escalograma obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 4800 Hz.....	82
Figura 4-10: Escalograma obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 48000 Hz.....	82
Figura 4-11: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 4800 Hz.	84
Figura 4-12: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 48000 Hz.	85
Figura 4-13: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 4800 Hz.	86
Figura 4-14: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 48000 Hz.	87
Figura 4-15: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado de Taylor. ...	89
Figura 4-16: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado de Hann.....	89
Figura 4-17: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado Rectangulares.	90
Figura 4-18: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.8.	91
Figura 4-19: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.95.	91
Figura 4-20: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.999.	92
Figura 4-21: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.99999.....	92
Figura 4-22: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 60Hz.	93

Figura 4-23: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 1200Hz..... 94

Figura 4-24: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 2400Hz..... 94

Figura 4-25: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 15Hz$ 95

Figura 4-26: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 30Hz$ 96

Figura 4-27: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 60Hz$ 96

Figura 4-28: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes wavelet sin normalización. 98

Figura 4-29: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes con normalización por remuestreo. 98

Figura 4-30: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes con normalización por remuestreo y normalización L1..... 99

Figura 4-31: Diagrama de flujo asociado a la generación de los conjuntos de entrada y salida para los algoritmos de detección y clasificación de eventos y la evaluación y clasificación de éstos.....102

Figura 4-32: Ventanas pre-falla, en falla y post-falla propuestas para la extracción de características.103

Figura 4-33: Matriz de características obtenidas con los coeficientes wavelet hallados usando la descomposición mediante DWT.....104

Figura 4-34: Matriz de características obtenidas con los coeficientes wavelet hallados usando la metodología propuesta, por ventana.105

Figura 4-35: Matriz de características obtenidas con los coeficientes wavelet hallados usando la metodología propuesta, por escala de frecuencia.105

Figura 4-36: Esquema de ANN propuesto para la detección y clasificación de eventos.107

Figura 4-37: Matriz de confusión para ANN usando DWT e información por nivel de descomposición.....109

Figura 4-38: Matriz de confusión para ANN usando CWT e información por ventana....110

Figura 4-39: Matriz de confusión para ANN usando CWT e información por escala de frecuencia.....111

Lista de tablas

Pág.

Tabla 1-1: Cifras de participación de los sistemas de distribución en los Sistemas Eléctricos de Potencia. Tomado de [1]	2
Tabla 2-1: Síntesis de la revisión de la literatura relacionada al análisis de fallas mediante el uso de la WT en SEP. Elaboración personal.	29
Tabla 4-1: Eventos simulados y características de falla (Impedancia, ángulo de inyección y tipo de falla). Elaboración propia.	76
Tabla 4-2: Tiempos promedio de simulación para el algoritmo propuesto en las señales sinusoidal y de falla simulada en función de la frecuencia de muestreo, frecuencia de resonancia, radio de convergencia, función de enventanado y paso de frecuencia. Elaboración propia.	100
Tabla 4-3: Codificación y categorización de eventos simulados. Elaboración propia. ...	106
Tabla 4-4: Parámetros de las redes neuronales utilizadas para los conjuntos de información obtenidos. Elaboración propia.	107

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Δf	Paso de frecuencia
Δt	Paso de tiempo
$C(a, b)$	Coeficiente Wavelet para factor de escala a y factor de desplazamiento b
F_S	Frecuencia de Muestreo
H	Entropía de Shannon
K	Curtosis
\bar{M}	Matriz Wavelet
p / q	Factores de remuestreo
r	Radio de convergencia del filtro resonante
s	Factor de desplazamiento de las Wavelet Hijas
T_S	Período de muestreo
u	Factor de escala de las Wavelet Hijas
W	Pesos neuronales
δt	Resolución en tiempo
$\delta \omega$	Resolución en frecuencia
$\bar{\lambda}$	Matriz diagonal de valores singulares
μ	Coeficiente de regularización
φ	Función de activación neuronal
ω_o	Frecuencia angular de resonancia
ψ	Wavelet Madre
$\Psi_{(u,s)}$	Wavelet Hija escalada por u y desplazada por s

Abreviaturas

ANN	Redes Neuronales Artificiales
CNN	Red Neuronal Convolutiva
Col	Cono de Influencia
CWT	Transformada Wavelet Continua
DC	Corriente directa
DFT	Transformada de Fourier en tiempo Discreto
DT	Árbol de decisión
DT-CWT	Dual Tree Complex Wavelet Transform

DWT	Transformada Wavelet Discreta
FFT	Transformada Rápida de Fourier
FIR	Respuesta Finita al Impulso
FT	Transformada de Fourier
IA	Inteligencia Artificial
IEEE	Instituto de Ingenieros Eléctricos y Electrónicos
IIR	Respuesta Infinita al Impulso
IWT	Transformada Wavelet Inversa
LTI	Lineal e Invariante en el Tiempo
ML	Machine Learning
MRA	Análisis Multirresolución
PNN	Red Neuronal Probabilística
PSD	Densidad de Potencia Espectral
QMF	Banco de Filtros Espejo en Cuadratura
RF	Bosques Aleatorios
RMS	Valor Cuadrático Medio
SEP	Sistema Eléctrico de Potencia
STFT	Transformada de Fourier de Tiempo Corto
SV	Sampled Values
SVD	Descomposición en Valores Singulares
SVM	Máquinas de Soporte Vectorial
WAE	Entropía Wavelet Promedio
WDE	Entropía Wavelet de Distancia
WE	Entropía Wavelet
WEE	Entropía de Energía Wavelet
WM	Wavelet Madre
WMM	Módulo Máximo Wavelet
WPT	Wavelet Packet Transform
WSE	Entropía Singular Wavelet
WT	Transformada Wavelet
WTE	Entropía de Tiempo Wavelet
WTFE	Entropía Tiempo-Frecuencia Wavelet
WTMM	Transformada Wavelet de Módulo Máximo

1. Introducción

Los *Sistemas Eléctricos de Potencia* (SEP) se han convertido en un pilar fundamental para el desarrollo de la civilización y estilo de vida humano modernos. Idealmente, un SEP se diseña para cumplir con el objetivo de generar, transmitir y distribuir energía eléctrica cumpliendo con unos estándares de confiabilidad, seguridad, estabilidad y eficiencia mientras se garantizan ciertos factores económicos relevantes para los entes operadores del sistema.

La confiabilidad en los SEP se refiere, en términos generales, a la garantía de que en el largo plazo existirá una capacidad instalada suficiente para cubrir la demanda proyectada, al igual que el suministro de energía eléctrica tendrá las mínimas interrupciones, cumpliendo con unos estándares de calidad de la energía. Por otro lado, la seguridad se entiende como la capacidad del SEP para minimizar situaciones que impliquen riesgos para las personas, el medio ambiente, el sistema mismo o los equipos de un usuario [1].

En cuanto al concepto de estabilidad, Kundur [1] la define como la habilidad de un SEP, dadas ciertas condiciones operativas iniciales, de recuperar un estado operativo de equilibrio luego de ser sujeto a una perturbación física, en donde la mayoría de las variables del sistema estén limitadas de forma que el sistema se encuentre, prácticamente, intacto [2]. En el ámbito económico, el ciclo de planeamiento busca cumplir con el objetivo técnico planteado con el desarrollo del SEP de menor costo posible maximizando el retorno a la inversión bajo las restricciones estipuladas en el plan de financiación, las políticas de cobro y las instituciones involucradas en el mercado de energía eléctrica.

En este sentido, las fallas de carácter eléctrico que afectan a los equipos y dispositivos que conforman un SEP se convierten en el centro de atención de la mayoría de las investigaciones y proyectos que buscan mantener o mejorar alguno de los indicadores mencionados anteriormente [3]. Dado que los SEP se componen, a grandes rasgos, por tres sistemas diferenciados, siendo estos Generación, Transmisión y Distribución, el

análisis de las perturbaciones eléctricas que se presentan en cada uno puede variar considerablemente, al igual que la definición misma de falla eléctrica.

Con el fin de focalizar el objetivo de esta disertación, se puede asignar prioridad a los sistemas de Distribución de energía eléctrica dado el elevado porcentaje de participación en múltiples instancias relacionadas con la operación confiable, segura, estable y eficiente técnica y económicamente de un SEP como se observa en la Tabla 1-1 [1].

Tabla 1-1: Cifras de participación de los sistemas de distribución en los Sistemas Eléctricos de Potencia. Tomado de [1] .

Ítem	Descripción	Participación [%]
1	Pérdidas de energía	70
2	Número de fallas	90
3	Indisponibilidad	99
4	Energía no suministrada	75
5	Costo de entrega de energía	40
6	Inversión anual para transporte	70
7	Costo de operación	20

Se puede observar que en la mayoría de los ítems listados en la tabla 1-1, los sistemas de Distribución de energía eléctrica tienen una gran influencia, al igual que un elevado potencial para mejorar las características de operación deseables del sistema.

Considerando las consecuencias que acarrearán las fallas eléctricas sobre la operación correcta de los SEP y en particular de los sistemas de Distribución de energía eléctrica, los sistemas de protecciones eléctricas y los esquemas para el análisis de fallas eléctricas se han vuelto necesarios para cualquier sistema moderno. En términos generales, el análisis de fallas eléctricas puede dividirse en tres etapas marcadas: la detección, clasificación y localización de fallas eléctricas.

La etapa de detección se centra en el reconocimiento de un estado operativo de falla en algún punto del sistema, sin preocuparse por su naturaleza o ubicación específica; su variable de estudio es el momento en el que ocurre el evento que conlleva al sistema a un estado de falla. Por otro lado, la clasificación se fundamenta en la extracción de características de los eventos que permitan diferenciarlas unas de otras, es decir, catalogarlos bajo algún esquema de clasificación. Finalmente, la localización se refiere a la determinación precisa de la ubicación en la que ocurre un evento que lleva al sistema al

estado operativo de falla y según la metodología aplicada, el tiempo de incidencia del evento puede o no ser de relevancia. Según la revisión de la literatura, abarcar los tres aspectos fundamentales del análisis de fallas extiende considerablemente el alcance de una investigación, por lo que se propone hacer énfasis en las etapas de detección y clasificación.

Tradicionalmente, el análisis de fallas eléctricas utiliza métodos basados en la medición de la impedancia observada desde un punto del sistema a frecuencia industrial. De esta manera se usan algoritmos que descartan la información proveniente de los componentes transitorios de alta frecuencia observados luego de una falla eléctrica [4]. En años recientes se ha considerado usar las componentes transitorias de alta frecuencia, las cuales contienen información adicional a la que aportan las componentes a frecuencia industrial.

En particular, se destacan dos enfoques: los esquemas de onda viajera [5]–[8] y los esquemas que utilizan el espectro en frecuencia de las señales de tensión y corriente [9]–[13]. Los esquemas de onda viajera utilizan la información de la onda electromagnética que se genera cuando ocurre un evento transitorio en el sistema, la cual se propaga a través de las líneas de transmisión y distribución e interactúa con ciertos componentes como cargas, transformadores en circuito abierto o centros de generación reflejándose. Dado que los sistemas de Distribución de energía eléctrica suelen ser enmallados o radiales y están sujetos a cambios topológicos constantes debido a la conexión o desconexión de cargas en los centros de consumo, el uso de ondas viajeras pierde eficiencia y exactitud, obligando a efectuar implementaciones más robustas y de mayor costo computacional [9], [14].

Según estas características el uso de esquemas que utilizan el espectro en frecuencia de las señales cobra relevancia. Estos enfoques utilizan métodos de procesamiento digital de señales para sintetizar la información contenida en las señales de falla y por lo general se acompaña de una etapa de extracción de características relevantes y algoritmos de *Machine Learning* (ML) e *Inteligencia Artificial* (IA) [15].

El fundamento del esquema es el del *Análisis Multirresolución* (MRA), el cual permite obtener información de las señales en el dominio del tiempo y de la frecuencia para diferentes escalas de resolución, entendiéndose resolución en este punto como una medida de la precisión del espectro en frecuencia de la señal [16]. El eje central de esta investigación estriba sobre la *Transformada Wavelet* (WT) y sus funciones fundamentales,

denominadas *Funciones Madre* o *Wavelets Madre* (WM) de la WT, las cuales definen un conjunto o familia de filtros. Específicamente, se buscará validar una metodología para utilizar una familia de filtros que no ha sido reportada en la literatura analizada y los resultados en términos de detección y clasificación de eventos simulados utilizando esta propuesta.

El resto de este escrito está organizado como sigue: En el Capítulo 2 se indaga en aspectos más específicos de los esquemas de detección y clasificación de fallas mediante el uso del espectro en frecuencia de las señales de tensión y corriente, con énfasis en el MRA y las observaciones relevantes desde la teoría de filtros digitales, además de la revisión de las metodologías reportadas para la extracción de características y la detección y clasificación de fallas; este capítulo también contiene la síntesis de la información reportada en la literatura consultada relacionada a la detección y clasificación de fallas mediante el uso de la Transformada Wavelet, al igual que expone las limitaciones y carencias detectadas en esta revisión literaria. El Capítulo 3 presenta una metodología para la implementación de filtros resonantes tipo IIR (Infinite Impulse Response) en conjunto con la Transformada Wavelet Continua y evalúa las implicaciones que conllevan el uso de estos filtros en la respuesta del sistema y detalla el esquema de extracción de características y los algoritmos de detección y clasificación de eventos propuestos. Por otro lado, en el Capítulo 4 se propone el uso del sistema IEEE de 9 barras para la simulación de los eventos de interés, se demuestran los resultados correspondientes a la validación de la metodología propuesta en función de las características del sistema y filtros utilizados; además, se comparan los resultados en términos de detección y clasificación utilizando la metodología propuesta con respecto a una metodología tradicional reportada en la literatura. Finalmente, en el Capítulo 5 se reportan las conclusiones halladas en la investigación y las contribuciones realizadas en el campo de estudio, al igual que demarcar el trabajo futuro.

2.Estado del arte y antecedentes

2.1 Detección y clasificación de fallas eléctricas en sistemas de distribución de energía eléctrica

El análisis de fallas eléctricas en líneas de distribución se opera, en gran mayoría, bajo esquemas tradicionales de medición de sobrecorriente, las cuales permiten detectar fallas eléctricas, pero presenta dificultades para ubicarlas de forma precisa. Funciones de protección más robustas, como la función de impedancia, podría aportar mejores resultados en líneas de transmisión para las cuales no se esperan cambios topológicos significativos entre sus dos terminaciones. La medida o estimación de la impedancia efectiva observada en el punto de falla permite concluir con cierta precisión si ocurre o no un evento que pueda catalogarse como falla eléctrica, al igual que el tipo de evento analizando las componentes de secuencia positiva y cero de ésta [13], [17], [18].

En términos generales, para la medición o estimación de la impedancia efectiva se pueden utilizar métodos de una única terminación en los que sólo se consideran los parámetros observados desde un lado de la línea para calcular la impedancia efectiva y métodos de doble terminación, en los que se considera la información observada en ambos extremos de la línea para calcular el valor de la impedancia efectiva. En ambos casos es posible estimar la ubicación del evento utilizando metodologías similares, basadas en el valor de impedancia por unidad de longitud de la línea y el valor calculado de impedancia efectiva.

Si bien estos enfoques han presentado resultados prometedores en su aplicación clásica, se han identificado limitaciones que han llevado al desarrollo de nuevas técnicas de análisis de fallas eléctricas en SEP [15]. En particular, dado que luego de un evento de falla se generan componentes transitorios, el cálculo de los componentes fundamentales o a frecuencia industrial de la tensión y corriente se vuelve complejo y requiere el uso de algoritmos robustos y de gran costo computacional [5].

Por otro lado, esta clase de esquemas de análisis de fallas es particularmente lenta con respecto a su interacción con los Sistemas de Protección del SEP, además de ser altamente susceptibles a factores como la impedancia y ángulo de inyección de las fallas [19], [20], [16]. Inclusive en el caso de líneas de transmisión de sistemas de alta o extra-alta tensión, la corriente de desplazamiento capacitiva y corrientes DC pseudo-amortiguadas causadas por la alta relación de reactancia y resistencia de estos sistemas han llevado a la disminución progresiva de la efectividad y precisión de estos esquemas [5].

Una de las primeras alternativas a los esquemas basados en componentes fundamentales o a frecuencia industrial es el de las Ondas Viajeras, el cual fue inicialmente propuesto por Bewley en la década de los 30 [8] aunque su implementación en aplicaciones reales no fue posible hasta la década de los 60 en la que se desarrollaron los primeros modelos computacionales digitales del efecto de los eventos transitorios electromagnéticos en los SEP, siendo Barthold y Carter los encargados de proponer su primera solución e implementación digital [21]. Este tipo de esquemas de análisis de fallas se centran en la extracción de los componentes transitorios de las señales de tensión y corriente utilizando la naturaleza ondulatoria de los eventos electromagnéticos que se presentan en los SEP.

Cuando se presenta un evento transitorio (no necesariamente una falla eléctrica) se generan ondas electromagnéticas que se propagan a una velocidad cercana a la de la luz a través de la línea de transmisión o distribución e interactúan con ciertos componentes como cargas, generadores o transformadores en circuito abierto, reflectándose y generando frentes de onda que regresan al punto de origen de la onda original (punto donde ocurrió la falla). La información contenida en estos frentes de onda que se propagan y reflejan permite establecer si ocurrió un evento y si este corresponde a una falla eléctrica o a una condición de operación normal del sistema.

Los métodos para la localización de fallas utilizando el modelo de onda viajera se pueden categorizar en tres grandes grupos: única terminación, doble terminación y múltiples terminaciones. En los esquemas de terminación única sólo se tiene disponibilidad de información de uno de los extremos de la línea de interés y no se requiere sincronización en tiempo o un esquema de comunicación de información dado que sólo se requiere del

registro de los tiempos de llegada de los frentes de onda al extremo de medición. Por otro lado, en los esquemas de terminación doble se necesitan mediciones sincronizadas mediante GPS a ambos extremos de la línea en cuestión o bien el uso de métodos de procesamiento digital de señales avanzados para medir o estimar la diferencia en el tiempo de llegada de los frentes de onda a cada uno de los extremos de la línea. Estos esquemas utilizan las diferencias en las medidas de tiempo de llegada de las ondas a los extremos de la línea y la velocidad de propagación de éstas para estimar el punto de falla [22]. De forma similar, los métodos de múltiples terminaciones requieren mediciones sincronizadas en todos los extremos de la línea, lo que ha llevado a que estos esquemas se eviten en aplicaciones prácticas. [23]

En general, los esquemas basados en el método de las Ondas Viajeras no se ven afectados significativamente por la impedancia o el tipo de falla como se ha reportado en la literatura [8], [21], [22]. Sin embargo, su principal aplicación en el análisis de fallas eléctricas se centra en la localización en líneas de transmisión, mientras que su aplicación en líneas de distribución es limitada debido a la complejidad elevada de estos sistemas, los cuales presentan cambios topológicos constantes al igual que múltiples terminaciones debido a su naturaleza radial [24], [25].

Por otro lado, la detección y clasificación de fallas eléctricas basadas en el método de las Ondas Viajeras se apoya en aplicaciones del procesamiento digital de las señales de tensión y corriente como el uso de la función de correlación [26]. Esta metodología no está ligada únicamente a los sistemas de análisis de fallas basados en el método de las Ondas Viajeras, dado que no requieren específicamente la información contenida en los frentes de onda, sino que también pueden fundamentarse en la información del espectro en frecuencia de las señales de tensión y corriente de falla.

Un primer acercamiento a esta metodología es el uso de la Transformada de Fourier (FT) de las señales de falla, la cual permite representar el contenido del espectro de frecuencia de las señales de tensión y corriente mediante su magnitud y fase. La FT de una función en el tiempo es una función compleja en el dominio de la frecuencia cuya magnitud representa cuán presente está una componente de frecuencia en la función original, mientras que su argumento o ángulo corresponde al desfase de la senoide en dicha frecuencia.

Originalmente, la FT de una señal es una función continua de una función integrable, definida como se observa a continuación:

Sea $g: \mathbb{R} \rightarrow \mathbb{C}$

$$G(\omega) = \int_{-\infty}^{\infty} g(t)e^{-j\omega t} dt \quad (2.1)$$

Donde G corresponde a la Transformada de Fourier de g y ω denota la frecuencia angular relacionada a cierto componente en frecuencia. Sin embargo, esta definición de la FT no es utilizada comúnmente en aplicaciones de procesamiento digital de señales dado que las señales que se tratan son de naturaleza discreta, además que comúnmente los sistemas computacionales encargados de calcular la FT de las señales operan de forma discreta. Así, se puede extender la definición de la Ecuación (2.1) en el dominio discreto para una secuencia finita de muestras $\{x_n\} = \{x_0, x_1, \dots, x_{N-1}\}$ como se observa a continuación:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j\frac{2\pi}{N}kn} \quad (2.2)$$

En este caso, la Transformada de Fourier Discreta (DFT) convierte una secuencia finita de muestras equi-espaciadas de una función en una secuencia de igual longitud y separación que componen una función compleja en el dominio de la frecuencia. En las referencias estudiadas que aplican la DFT (o una versión más eficiente computacionalmente conocida como la Transformada de Fourier Rápida o FFT) se observan aplicaciones relativamente sencillas; en el caso de las publicaciones actuales, su objetivo fundamental es el de comparar los resultados obtenidos mediante el uso de la DFT con respecto a otros métodos para el análisis de fallas en el dominio de la frecuencia [27], [28], [29], [30].

Para comprender las razones que limitan la aplicación de la FT en el análisis de fallas en SEP se debe introducir el concepto de resolución. En términos generales, la resolución de una medida está relacionada con su nivel de granularidad con el que es posible discernir detalles. En este sentido, se definen dos tipos de resoluciones: la resolución en tiempo y la resolución en frecuencia. La convención utilizada es la siguiente: cuando la resolución

mejora es posible discernir detalles cada vez más finos, por lo que el valor numérico de la resolución se reduce, mientras que cuando la resolución empeora, su valor aumenta.

Estos conceptos están relacionados bajo el principio de incertidumbre de Heisenberg demostrado en la Ecuación (2.3), el cual demuestra que si la resolución en tiempo de una señal aumenta su resolución en el dominio de la frecuencia disminuye.

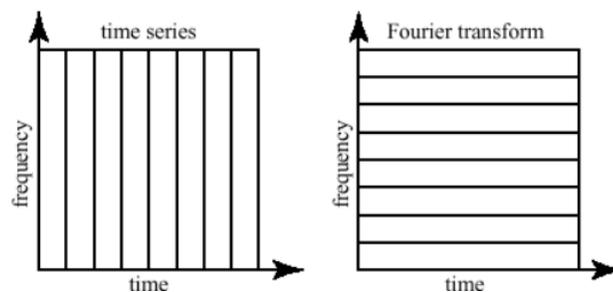
$$\delta t \cdot \delta \omega \geq 4\pi \quad (2.3)$$

Aquí, δt corresponde a la resolución en tiempo y $\delta \omega$ corresponde a la resolución en frecuencia (angular).

En este punto es útil comparar la FT con las series de tiempo o de Shannon, debido a su naturaleza opuesta. Por un lado, las series de tiempo representan una función en el dominio del tiempo con la mayor resolución posible, mientras que no existe información alguna del contenido espectral de frecuencia; esto permite localizar fenómenos y eventos temporales de forma exacta, aunque el contenido en frecuencia es desconocido.

De forma similar, la FT es la representación en fase y amplitud en el dominio de la frecuencia de una señal, por lo que se tiene la máxima resolución posible en frecuencia, aunque es imposible tener información en el tiempo sobre estos componentes en la frecuencia; así, es posible localizar fenómenos y eventos en el dominio de la frecuencia con exactitud, aunque es imposible asignarles información temporal. Una representación gráfica de este fenómeno se presenta en la Figura 2-1 [31].

Figura 2-1: Resolución en tiempo y frecuencia para las series de tiempo (izquierda) y la Transformada de Fourier (derecha).



Nombre de la fuente: Heart rate variability: linear and nonlinear analysis with applications in human physiology [31].

Considerando el problema de capturar información localizada tanto en el dominio del tiempo como en el dominio de la frecuencia, una evolución natural para la FT consiste en dividir la secuencia o señal original en ventanas de longitud constante y luego computar la FT para cada uno de estos segmentos de forma separada. Este es el fundamento de la Transformada de Fourier de Tiempo Corto (STFT), la cual puede calcularse utilizando una función de enventanado $w(t - \tau, m)$ para restringir el dominio en el que se toman muestras o valores de la señal o secuencia original. En este sentido, la STFT en tiempo continuo para una señal se computa como se observa a continuación:

$$X(\tau, \omega) = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-j\omega t} dt \quad (2.4)$$

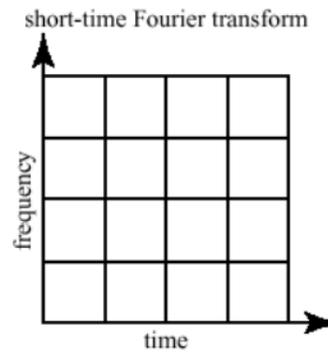
Mientras que la STFT de una secuencia discreta de valores corresponde a:

$$X_k(m) = \sum_{n=0}^{N-1} x_n w(n - m) e^{-j\frac{2\pi}{N}kn} \quad (2.5)$$

Donde $\tau \in R$ y $m \in Z$ corresponden a los índices temporales. En el caso de la función de enventanado $w(t - \tau, m)$ es común utilizar una ventana Gaussiana centrada en cero o bien una ventana de Hann [32].

En cuanto a la resolución en tiempo y frecuencia de la STFT, dado que se definen ventanas en todo el dominio de la señal o secuencia analizada es posible localizar eventos en tiempo para los componentes en frecuencia de alguno de los segmentos definidos. En este caso la resolución permanece constante durante todo el dominio de la señal dado el ancho fijo de las ventanas definidas por la función de enventanado.

Figura 2-2: Resolución en tiempo y frecuencia para la Transformada de Fourier de Tiempo Corto.



Nombre de la fuente: Heart rate variability: linear and nonlinear analysis with applications in human physiology [31].

Algunas implementaciones relacionadas al análisis de fallas eléctricas en SEP utilizando STFT se reportan en [33], [34]. En general estas metodologías funcionan, aunque presentan una precisión menor en comparación con otros métodos reportados en la literatura por lo que se requieren enfoques más robustos en su aplicación.

Con respecto a lo anterior, si bien es posible localizar eventos en el tiempo y frecuencia utilizando la STFT existe una relación de compromiso con respecto a los eventos analizados dado que no es posible tener una resolución en tiempo arbitrariamente buena simultáneamente con buena resolución en frecuencia, por lo que se deben redefinir las ventanas de la transformada para cada caso. En otros casos se acostumbra a realizar dos o más análisis con la STFT para cada señal: uno de ventana corta y otro de ventana larga, con el fin de cubrir ambos extremos del espectro en frecuencias. Esta característica de la STFT hace que sea poco práctica en aplicaciones en las que se busca analizar eventos en diferentes rangos de frecuencia y escalas de tiempo.

Es claro que es necesario introducir métodos de análisis en tiempo y frecuencia que permitan implementar una resolución variable que se ajuste a las características de la señal o secuencia a ser analizada; este es el fundamento del Análisis Multi-Resolución (MRA).

2.2 La Transformada Wavelet

Con el fin de introducir el concepto teórico de la Transformada Wavelet (WT) es pertinente introducir algunas definiciones básicas utilizadas en el procesamiento digital de señales. En primera instancia, es relevante definir la representación dispersa de una señal como aquella que contiene la información de relevancia de ésta en un número finito de coeficientes. En general, el procesamiento de las señales es más rápido y sencillo cuando se realiza de forma dispersa.

Las representaciones dispersas pueden construirse al descomponer las señales en formas de onda elementales escogidas de alguna familia de funciones elementales. Esta familia es denominada, en el contexto matemático, como un diccionario. La búsqueda de diccionarios capaces de obtener la información de relevancia de una señal es el punto de partida del concepto de la WT.

Ahora bien, uno de los objetivos principales del uso de diccionarios en la representación de señales es la búsqueda de bases ortogonales, las cuales son diccionarios del mínimo tamaño posible capaces de entregar una representación dispersa cuando se diseñan para concentrar la energía de la señal en un conjunto de vectores [35]. En otras palabras, una base ortogonal, en el contexto de la representación de señales digitales, contiene un número mínimo de coeficientes que concentran la energía de la señal y son capaces de representarla de forma dispersa.

Por otro lado, el soporte de una función en un dominio arbitrario puede definirse como el intervalo en el que la función es diferente de cero y es un concepto relevante en cuanto a la convolución de dos señales y la recuperación de información localizada.

En cuanto a la progresión histórica de la WT, podría decirse que el primer acercamiento al concepto de Wavelets se dio en 1910 cuando el matemático Alfréd Haar construyó la siguiente función a trozos:

$$\psi(t) = \begin{cases} 1, & \text{sí } 0 \leq t < 1/2 \\ -1, & \text{sí } \frac{1}{2} \leq t < 1 \\ 0, & \text{en otro caso} \end{cases} \quad (2.6)$$

Al analizar esta función, Haar descubrió que se puede generar una base ortonormal, para cierto espacio de Hilbert, ejecutando translaciones y dilataciones de la función:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-2^j n}{2^j}\right) \quad (2.7)$$

Con $(j, n) \in Z$

Cada función $\psi_{j,n}(t)$ corresponde a una Wavelet de Haar, que en conjunto introducen las nociones fundamentales necesarias para definir las condiciones de admisibilidad de las bases ortonormales de una familia de Wavelets o diccionario las cuales serán evaluadas con mayor detalle en el capítulo 2.3. En particular, cada Wavelet de Haar tiene soporte en el intervalo $[2^j n, 2^j(n+1)]$, por lo que se dice que su soporte es compacto o finito, y el valor medio de la función en su intervalo de soporte es cero [35]. En la práctica, esta función primitiva puede identificar transiciones fuertes en una señal, lo cual marcó el inicio del MRA.

Luego, en 1980 el matemático Jan-Olov Strömberg propuso otra función lineal a trozos ψ que también genera una base ortonormal y presentaba mejores aproximaciones en funciones suaves [36].

El nombre Wavelet -que traduce literalmente ondícula o pequeña onda- fue originalmente acuñado por el geofísico Jean Morlet y el físico Alex Grossman en 1984 [37]. Morlet había trabajado en sismología con reflexión de ondas, en la que se envían formas de onda de baja duración a alta frecuencia a través del subsuelo con el fin de separar las capas geofísicas que se encuentran espaciadas de forma fina. Dichas formas de onda se conocen como Wavelets en el campo de la geofísica.

En este sentido, la primera aplicación real de la WT fue aplicada por Morlet mediante el envío de señales cada vez más cortas en duración a frecuencias más altas. Por su parte, Grossman utilizó un enfoque similar aplicado a estados cuánticos coherentes. En 1984, Grossman y Morlet proponen la descomposición de funciones relacionadas al campo de la física cuántica en Wavelets cuadradas e integrables de forma constante [38].

Por otro lado, el matemático Yves Meyer, sin conocer el resultado de Strömberg, intentó probar que no existían wavelets regulares ψ que generaran una base ortonormal. Su

investigación concluyó con la construcción una nueva familia de bases ortonormales con funciones ψ infinita y continuamente diferenciables que luego sería conocida como la Wavelet de Meyer [39]. Esto finalmente llevaría a Ingrid Daubechies a desarrollar la familia de Wavelets Daubechies [40], [41], la cual es una de las más utilizadas en el contexto de análisis de fallas eléctricas en SEP.

Estos aportes crearon el fundamento teórico de la Transformada Wavelet, la cual se basa en el uso de las Wavelet, las cuales son funciones que deben cumplir, inicialmente con las siguientes propiedades:

- Media cero en su dominio $\int_{-\infty}^{\infty} \psi(t) dt = 0$.
- Deben estar normalizadas $|\psi| = 1$ y centradas en cero.

Cuando una función cumple con estas propiedades, es posible crear un diccionario de la representación en el dominio del tiempo y la frecuencia al escalar ψ por un factor s y trasladándola por un factor u (siendo ésta una generalización al procedimiento desarrollado por Haar). Así se define el diccionario D como:

$$D = \left\{ \psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \right\} \quad (2.8)$$

Donde $u \in R$ y $s \in Z$ corresponden a los parámetros de traslación y escala, respectivamente. $\psi_{u,s}$ corresponde a la Wavelet ψ modificada por los parámetros de traslación y escala. Para efectos prácticos, la Wavelet ψ es denominada la función Madre o Wavelet Madre (MW), dado que a partir de ésta se generan las funciones $\psi_{u,s}$ conocidas como funciones hijas o Wavelets hijas. Cada una de las funciones hijas mantiene su magnitud unitaria a través del escalamiento y traslación ($|\psi_{u,s}| = 1$) [42].

Se puede extender el concepto de Wavelets al aplicarlo a una función continua e integrable $f(t)$ para un instante de tiempo u y escala s como se observa a continuación:

$$CWT\{f(t)\} = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt \quad (2.9)$$

La Ecuación (2.9) define la WT en tiempo Continuo o CWT de la función $f(t)$. En ésta, el operador $*$ denota la conjugación compleja.

Esta definición puede extenderse al tiempo discreto de la siguiente forma: dada una señal $x(t)$ con representación discreta $x[n]$ obtenida en intervalos de muestreo uniformes N^{-1} , se puede computar una Transformada Wavelet Discreta o DWT para las escalas $s \in Z \vee N^{-1} < s < 1$. En este caso, los parámetros de escala y traslación quedan definidos de forma discreta como $s = a_0^j$ y $u = nb_0s$, respectivamente, con $j \in Z$. De esta forma, la DWT de $x[n]$ queda definida por:

$$DWT\{x[n]\} = \frac{1}{\sqrt{a_0^m}} \sum_n x[n] \psi\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \quad (2.10)$$

En esta representación los parámetros de escala y traslación son funciones de valores enteros $j \in Z$, por lo que se obtiene un factor de escala s que varía de forma geométrica como $\{1, \frac{1}{a}, \frac{1}{a^2}, \dots\}$ y un factor de traslación u que varía de forma lineal como $\{0, n, 2n, \dots\}$ [43]. La representación de la DWT presentada en la Ecuación (2.10) puede reescribirse como:

$$DWT\{x[n]\} = \frac{1}{\sqrt{a_0^m}} \sum_k x[k] \psi(a_0^{-m}n - b_0k) \quad (2.11)$$

En esta ecuación es posible observar una similitud con respecto a la Respuesta Finita al Impulso (FIR) de filtros digitales, relacionada con el teorema de convolución como se observa a continuación:

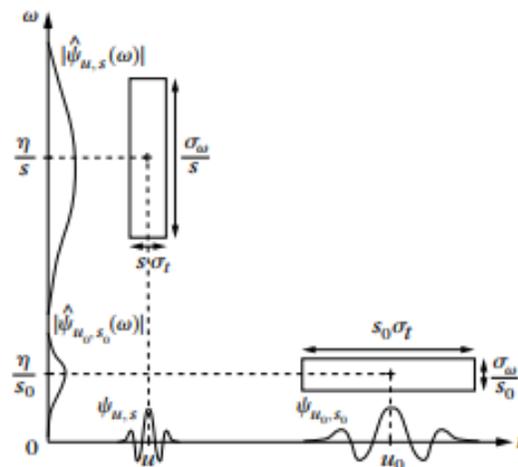
$$y[n] = \frac{1}{L} \sum_k x[k] h(n - k) = x[n] * h(n) \quad (2.12)$$

Donde $h(n)$ corresponde a la respuesta a una entrada tipo impulso del sistema. Según la Ecuación (2.12), $h(n) = \psi(a_0^{-m}n - b_0k)$. Bajo este enfoque es posible simplificar la escogencia de escalas y fijar $a_0 = 2$ para un $b_0 > 1$ arbitrario, lo cual crea una escala diádica para las divisiones en el eje de la frecuencia. De esta forma, el parámetro $s = a_0^j$ se duplica en cada nivel de frecuencia subsecuente con respecto al anterior. De forma

análoga, el número de puntos en los que la Wavelet está definida se reduce a la mitad lo cual reduce la resolución en frecuencia e introduce el MRA en la teoría de la WT. El fenómeno descrito anteriormente es formalmente conocido como Downsampling.

Por consiguiente, tanto en su versión continua como discreta, las Wavelets que componen la WT tienen una resolución que varía dependiendo del factor de escala. En el caso de la CWT, la Wavelet definida por $\psi_{u,s}$ tiene un soporte en el tiempo que es proporcional a s y está centrado en $t = u$, mientras que, en el dominio de la frecuencia, se encuentra localizada en un rango de frecuencias positivas de tamaño proporcional a $\frac{1}{s}$. De aquí se puede concluir que cuando el parámetro de escala cambia, el ancho de la resolución en tiempo y frecuencia también se modifican, pero el área que éstos definen permanece constante.

Figura 2-3: Resolución en tiempo y frecuencia para la Transformada Wavelet.



Nombre de la fuente: A Wavelet Tour of Signal Processing The Sparse Way [35].

La figura 2-3 [35] muestra la invarianza del área producida por el ancho en la resolución en tiempo y frecuencia para dos escalas diferentes en una Wavelet. Por otro lado, se observa la característica más atractiva de las Wavelets en el MRA, la cual se resume en que si se obtiene un coeficiente de la WT lo suficientemente grande es posible detectar y medir variaciones cortas de alta frecuencia dado que tienen localización angosta en el dominio del tiempo para altas frecuencias; Mientras que a bajas frecuencias la resolución

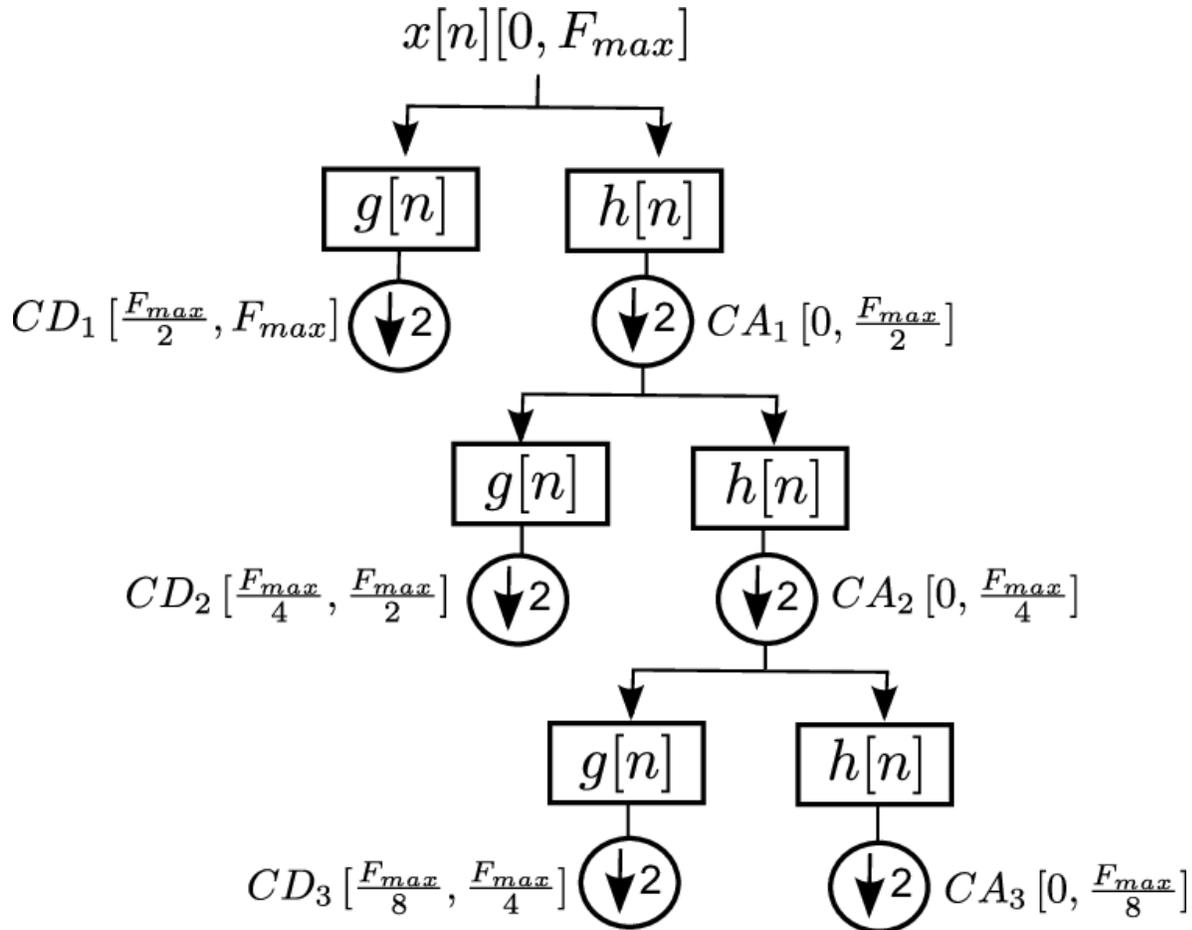
en tiempo es menor, aunque tiene mejor resolución en frecuencia, permitiendo localizar eventos en frecuencia con mayor precisión. Dado que los fenómenos transitorios que caracterizan las fallas en los SEP son, en general, de alta frecuencia y corta duración, el uso de Wavelets se ajusta de forma natural a éstas características.

Con respecto a la correspondencia entre la DWT y la Respuesta Finita al Impulso de los filtros digitales, dado que es posible establecer una escala diádica para las divisiones en el eje de la frecuencia y utilizando el hecho de que la operación de convolución en el dominio del tiempo representa una multiplicación en el dominio de la frecuencia, se puede definir la DWT de una secuencia digital como una serie finita de filtros en cascada. Cabe destacar que, si se considera la MW ψ como un filtro cuya frecuencia central es f_o , cada una de las funciones hija $\psi_{u,s}$ también será un filtro con frecuencia central f_o/s .

En términos prácticos, las Wavelets hijas se definen mediante un par de filtros digitales: un filtro pasa-altas $g[n]$ y un filtro pasa-bajas $h[n]$. Estos filtros pueden aplicarse a una secuencia digital $x[n]$ cuya frecuencia de muestreo es F_s mediante la convolución en el tiempo (Ecuación (2.12)), resultando en dos nuevas señales discretas; la señal resultante de aplicar el filtro pasa-altas es llamada componente de detalle, dado que recoge los detalles de alta frecuencia de la señal original, y dada la naturaleza diádica de esta descomposición, contiene la información de la banda de frecuencia $[F_s/2, F_s]$, mientras que la convolución con el filtro pasa-bajas produce la componente de aproximación, la cual contiene la información de baja frecuencia de la señal original comprendida en la banda de frecuencia $[0, F_s/2]$. Debido al fenómeno de Downsampling, cada una de las componentes resultantes tiene la mitad de las muestras de la secuencia digital original $x[n]$.

El procedimiento descrito genera el primer nivel de descomposición y puede extenderse aplicándose sobre el componente de aproximación generado, creando un nuevo par de componentes de aproximación y detalle, que definen dos nuevas bandas de frecuencia $[0, F_s/4]$ y $[F_s/4, F_s/2]$, cada una con la mitad de las muestras de la componente de aproximación de la cual fueron generadas. Cada vez que se computan un nuevo par de componentes de detalle y aproximación se crea un nuevo nivel de descomposición N , el cual va a contener la información de las bandas de frecuencia $[0, F_s/2^N]$ y $[F_s/2^N, F_s/2^{N-1}]$. Este procedimiento es comúnmente denominado Algoritmo Piramidal, debido a la forma en la que se generan nuevos coeficientes.

Figura 2-4: Algoritmo Piramidal para MRA de señales utilizando la DWT.



Nombre de la fuente: Discrete wavelet transform optimal parameters estimation for arc fault detection in low-voltage residential power networks [44].

El algoritmo piramidal ha sido ampliamente utilizado en el contexto de análisis de fallas en SEP, desde redes residenciales en baja tensión [44], hasta Smart Grids [45] y sistemas de distribución [12], [46] y transmisión [27], [47]. Debido a su amplia base de investigación, y a su rápida implementación computacional, la DWT es preferida. Efectivamente, es notorio que la investigación del uso de la CWT al problema de análisis de fallas eléctricas en SEP es mucho más reducida que su contraparte en tiempo discreto, aunque se resaltan algunas características de interés en este sentido:

- Las limitaciones computacionales observadas tanto en los Software de simulación de transitorios electromagnéticos como en los equipos de protección y medida de SEP

se han reducido considerablemente en la última década, por lo que la implementación de la CWT no presenta inconvenientes mayores en cuanto a carga computacional ni tiempo de procesado. Por otro lado, en la práctica la CWT debe implementarse de forma discretizada en los procesadores digitales que se encargan de computarla, lo que permite utilizar una versión modificada del algoritmo piramidal para alivianar la carga computacional asociada a ésta [48], [49], [50].

- En comparación con la DWT, la CWT permite desarrollar análisis más detallados del espectro de energía de los fenómenos transitorios electromagnéticos [49], [50].
- En presencia de transductores de potencia (hallados en equipos de medida de los SEP), la CWT ha demostrado tener una mayor precisión general con respecto a la DWT [45], [51].
- En el caso de aplicaciones con técnicas avanzadas para el análisis en el dominio de la frecuencia, como el uso de imágenes en escalas de grises, se ha demostrado que los detalles adicionales entregados por la CWT mejoran la precisión del método de detección de fallas con respecto a la DWT [48].

Como se observa, existe un potencial que no ha sido explorado por completo con respecto a la investigación en el uso de la CWT para el análisis de fallas en SEP. Bajo este argumento y los presentados con referencia en la literatura, el foco de la presente investigación será la CWT.

Ahora bien, tal como se aclaró anteriormente, es común discretizar la CWT para implementarla numéricamente en el Software de procesamiento encargado de computarla (o bien el Software se encarga de discretizar el sistema y la WT de forma interna). En este sentido, si se considera una señal en el tiempo $x(t)$ cuya representación discreta es la secuencia $x[n]$, la implementación numérica de la CWT sobre la señal $x(t)$ está denotada por los coeficientes de la matriz $C(a, b)$, donde:

$$C(a, b) = C(a, iT_s) = T_s \frac{1}{\sqrt{|a|}} \sum_{n=0}^{N-1} \psi^* \left(\frac{(n-i)T_s}{a} \right) x[nT_s] \quad (2.13)$$

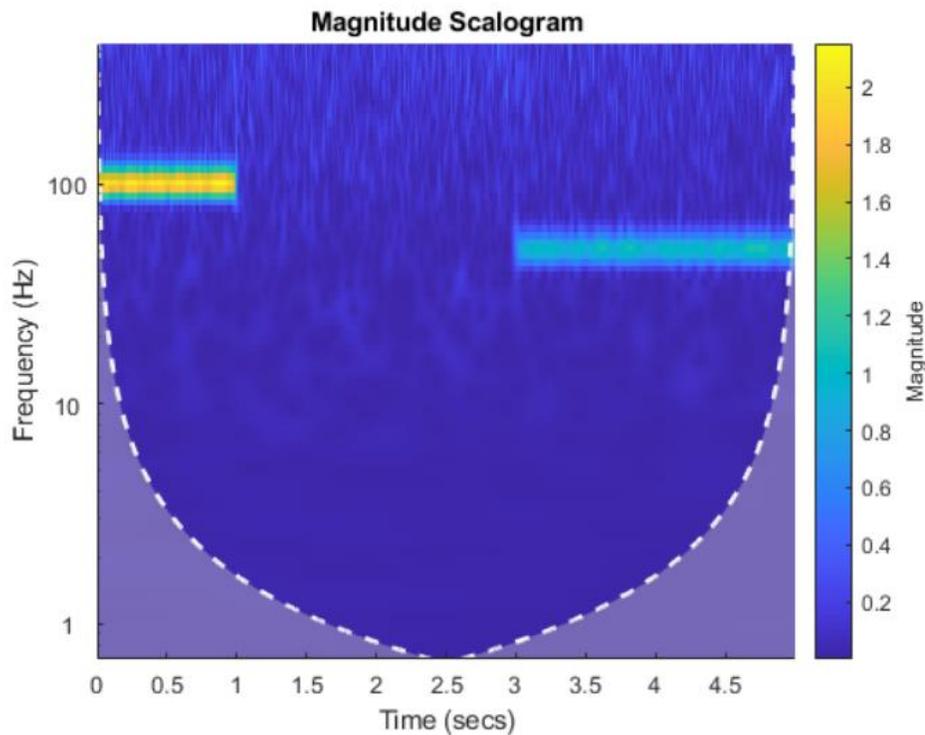
Donde:

- a corresponde al coeficiente de escala.
- $b = iT_s$ corresponde al factor de traslación en el tiempo.

- N corresponde al total de registros de la secuencia $x[n]$.
- T_s corresponde al período de muestreo asociado a la secuencia $x[n]$.

Por lo general estos coeficientes se utilizan de forma netamente numérica para obtener características de las señales, aunque es posible observar los resultados de la WT de forma gráfica utilizando un escalograma. Los escalogramas son los análogos a los espectrogramas en el caso de la FT dado que presentan la energía contenida en la señal para una escala de frecuencia y momento del tiempo en particular. Como se verá a fondo más adelante, la energía espectral se define como el cuadrado de los WC y puede ser ajustada a un mapa de calor cuyo eje vertical corresponde a las escalas de frecuencia y su eje horizontal al tiempo. Un ejemplo de un espectrograma se observa en la figura 2 – 5; en este caso la señal utilizada es una senoide con frecuencias de 100Hz y 50Hz y una componente de ruido blanco [52].

Figura 2-5: Escalograma para una senoide con frecuencias de 100Hz y 50Hz y una componente de ruido blanco.



Nombre de la fuente: <https://la.mathworks.com/help/wavelet/gs/continuous-wavelet-analysis.html>

[52].

Aquí podemos observar que el valor de la energía se relaciona a una escala de colores cuya magnitud se representa a la derecha del escalograma; además, se logran identificar las componentes de 100Hz y 50Hz en los espacios de tiempo durante los que están presentes en la señal analizada. Sin embargo, se puede observar una zona gris debajo de una línea punteada la cual es conocida como el Cono de Influencia (CoI), el cual muestra las zonas en las que la información presentada por el escalograma puede verse afectada por los efectos de frontera intrínsecos a la WT [53].

Finalmente, se destacan algunos casos particulares de la WT que han sido implementados en esquemas de análisis de fallas en SEP. En primer lugar, se encuentra la Wavelet Packet Transform (WPT), la cual es una extensión de la DWT implementada bajo el algoritmo piramidal. En esta versión de la WT el proceso de filtrado asociado a un nuevo nivel de descomposición se aplica tanto para la señal resultante del filtro pasa-altas $g[n]$ y el filtro pasa-bajas $h[n]$, resultando en una representación más detallada en el espectro de frecuencia de la señal en el tiempo.

Algunas implementaciones destacadas de esta variación de la DWT se presentan en [4], [54]–[56]. En particular se resalta que estas implementaciones no se ven afectadas por el ángulo de inyección de falla o impedancia de falla; en comparación con la DWT, la WPT presenta una mayor precisión en la detección y clasificación de eventos de falla.

Por otro lado, se ha propuesto el uso de filtros de valores complejos para implementar el algoritmo piramidal de la DWT. Este método es conocido como Dual Tree Complex Wavelet Transform (DT-CWT) y opera de forma similar a la WPT, aunque utiliza filtros complejos en lugar de filtros reales y los fasores de tensión y corriente en lugar de las señales en el tiempo. En general, la precisión de los algoritmos de detección y clasificación de fallas es mayor con la DT-CWT en comparación con la DWT [57], [58], aunque el requerimiento de utilizar filtros complejos y los fasores de las señales representa un reto en cuanto a la implementación física de este tipo de metodologías.

2.3 Funciones Madre: Observaciones desde la teoría de filtros digitales

En las operaciones clásicas del procesamiento digital de señales es común utilizar sistemas u operadores Lineales e Invariantes en el Tiempo (LTI). De forma sucinta, un sistema es lineal si satisface el principio de superposición [59], mientras que la invarianza en el tiempo requiere que si una entrada se retrasa por un factor τ , su salida se retrase por el mismo factor.

Ahora bien, al analizar la Ecuación (2.11), se estableció que existe una similitud entre la forma en la que se computa la DWT con la respuesta de un sistema Lineal e Invariante en el Tiempo ante una entrada impulso de Dirac. Por otro lado, en la Ecuación (2.12) se relaciona la respuesta de un sistema LTI ante una entrada tipo impulso con el teorema de convolución. En este sentido es posible considerar las funciones madre o MW $\psi_{u,s}$ como filtros digitales pasabandas escalados por un factor s y retrasados en el tiempo por un factor u . La CWT también puede observarse desde este enfoque de implementación, aunque los filtros (funciones madre e hijas) deben estar descritos en tiempo continuo [35].

La consecuencia del escalamiento por un factor s de las funciones hijas $\psi_{u,s}$ se resume en que si la MW ψ es un filtro pasabandas con frecuencia central f_o , la función hija $\psi_{u,s}$ tendrá una frecuencia central f_o/s . Esta propiedad permite extraer la información de diferentes escalas de frecuencia de la señal original para un instante de tiempo, mientras que la traslación introducida por el factor u posibilita desplazarse en el tiempo y evaluar diferentes instantes de tiempo.

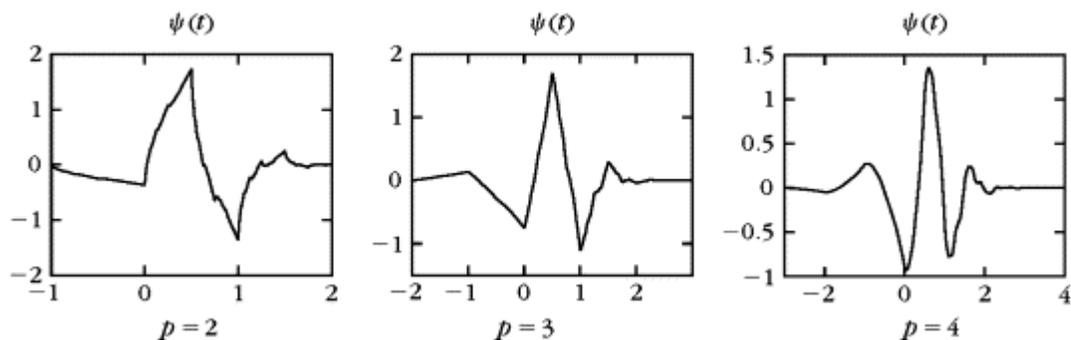
En cuanto al soporte de las funciones (tanto hijas como madre), al analizarlas en el dominio de la frecuencia como filtros pasabandas, se puede encontrar una similitud con el ancho de banda del filtro. De esta forma, un soporte compacto en el tiempo indica una localización más precisa en frecuencia y un menor ancho de banda.

Por otro lado, existen múltiples familias de MW o filtros que se han desarrollado a través de la historia de las Wavelet; el uso de alguna familia particular dependerá en primera instancia de la aplicación en la que se desea implementar la WT, por lo que su selección

en muchos casos queda a criterio del investigador. En el caso del análisis de fallas en SEP se destacan las siguientes MW:

- a) Daubechies:** Fueron introducidas en 1990 por Ingrid Daubechies. Corresponden a una familia de Wavelets ortogonales que definen DWT por lo que normalmente están expresadas como filtros digitales. Matemáticamente se caracterizan por tener un número máximo de *momentos de desvanecimiento*, los cuales son una medida de la regularidad de una función y son definidos como momentos estadísticos, los cuales definen su localización en la frecuencia [43]. De forma similar, los momentos de desvanecimiento definen el soporte de la función o el orden del filtro pasabandas asociado. En la figura 2-6 se pueden observar algunas funciones madre de la familia Daubechies para diferentes momentos de desvanecimiento P (para efectos prácticos, el orden del filtro corresponde al número de momentos de desvanecimiento).

Figura 2-6: Wavelets de Daubechies de ordenes 2, 3 y 4.



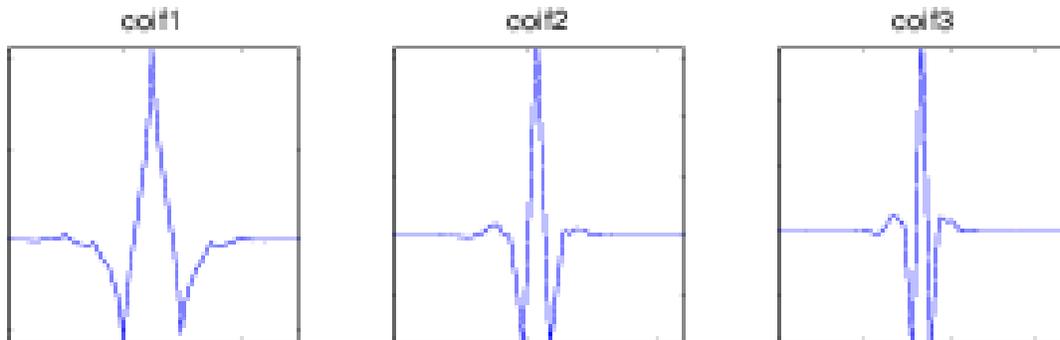
Nombre de la fuente: A Wavelet Tour of Signal Processing The Sparse Way [35].

En la literatura consultada hasta el momento, la familia de Wavelets de Daubechies es la más utilizada en el contexto de análisis de fallas eléctricas en SEP. En particular, se suelen utilizar filtros de cuarto orden, denotados por dB4, los cuales han sido extensamente utilizados [22], [44], [48] y comparados con respecto a otras familias [60], [61], [62]. En instancias reducidas se ha observado el uso de filtros de orden 2 [63] y 10 [54].

- b) Coiflet:** Esta familia también fue desarrollada por Ingrid Daubechies bajo una petición de su colega Ronald Coifman (el nombre Coiflet está dado por la unión de las palabras “Coifman” y “Wavelet”) [43]. Esta familia de funciones es

aproximadamente simétrica alrededor de su frecuencia central por lo que tiene mayor aplicabilidad en análisis armónico de señales. La figura 2-7 muestra los filtros Coiflet de primer, segundo y tercer orden.

Figura 2-7: Coiflets de ordenes 1, 2 y 3.



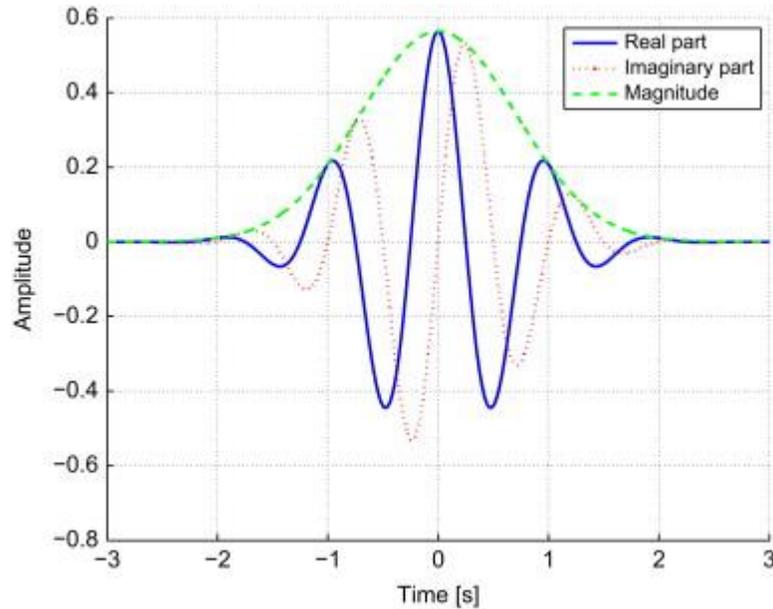
Nombre de la fuente: A Wavelet Tour of Signal Processing The Sparse Way [35].

Si bien esta familia de funciones no ha sido tan utilizada en análisis de fallas eléctricas en SEP se destacan algunas aplicaciones que demuestran resultados comparables con el filtro Daubechies de orden 4 [55], [60], [61].

- c) Morlet:** Reciben su nombre de Jean Morlet, el cual propuso formalmente la CWT en 1984 en conjunto con Grossman. Las Morlet surgieron a partir de un trabajo del físico Dennis Gabor, el cual trabajó la descomposición en tiempo y frecuencia de señales con sinusoides enventanadas con funciones Gaussianas; Morlet modificó estas funciones utilizando un factor exponencial complejo con lo cual introdujo la capacidad de realizar MRA con Wavelets.

En este caso la localización en frecuencia está dada por una función Gaussiana cuya frecuencia central corresponde al máximo global de ésta; dado que se utiliza un factor exponencial complejo, es común utilizar solo la parte real del filtro (aunque el filtro complejo ha sido ampliamente utilizado en aplicaciones de la física cuántica) [35]. Una representación de una Wavelet de Morlet en su representación compleja (filtro real e imaginario) se presenta en la figura 2-8. Nótese cómo la magnitud del filtro corresponde a una envolvente alrededor de las componentes real e imaginaria.

Figura 2-8: Wavelet de Morlet real e imaginaria.



Nombre de la fuente: Signal reconstruction, modeling and simulation of a vehicle full-scale crash test based on Morlet wavelets [64].

En cuanto al análisis de fallas eléctricas en SEP, las Wavelets de Morlet han visto un incremento en cuanto al número de publicaciones [49], [61] debido a su simetría, aunque en sus aplicaciones iniciales se consideraron innecesariamente complicadas por tratarse de filtros complejos, además de no detectar todos los componentes asociados a la reflexión de ondas en el método de las Ondas viajeras [65].

Todas estas familias de funciones cumplen con ciertas condiciones requeridas según la teoría de Wavelets. En particular, para una función hija $\psi_{u,s}$ cuya función madre es ψ se debe garantizar que:

- El valor medio de la función madre e hija sea cero en su dominio $\int_{-\infty}^{\infty} \psi(t) dt = 0$ y $\int_{-\infty}^{\infty} \psi_{u,s}(t) dt = 0$.
- La función madre e hija deben estar normalizadas $|\psi| \vee 1$ y $|\psi_{u,s}| = 1$.
- La función madre debe estar centrada en $t = 0$.
- La función hija debe estar centrada en $t = u$.

Estas condiciones son suficientes para garantizar que una señal puede descomponerse en las escalas de frecuencia definidas por el parámetro s . Ahora bien, existe otra condición denominada *condición de admisibilidad* (Ecuación (2.14)), la cual es necesaria para garantizar que la WT es matemáticamente completa y cumple con el principio de conservación de la energía. Esta condición fue desarrollada y propuesta por Calderón en 1964 [66], aunque también fue demostrada por Morlet y Grossman, los cuales desconocían sobre el trabajo de Calderón dado que éste trabajó sobre el teorema de convolución y no sobre las Wavelet, las cuales eran relativamente nuevas [38]. Es relevante resaltar que en este caso $\psi(\omega)$ representa la FT de $\psi(t)$.

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty \quad (2.14)$$

Por otro lado, existen condiciones de suficiencia que garantizan que una familia de funciones cumple con la condición de admisibilidad; en particular:

- a) El valor medio de las funciones debe ser cero.
- b) $\psi(\omega)$ es continuamente diferenciable.

En particular, la condición de suficiencia b se cumple si se garantiza que las funciones $\psi(t)$ (y por ende $\psi_{u,s}$) decrece arbitrariamente rápido a cero cuando el tiempo tiende a infinito, lo cual puede expresarse de forma matemática bajo la siguiente expresión:

$$\int_{-\infty}^{\infty} (1 + |t|)\psi(t)dt < \infty \quad (2.15)$$

En última instancia, existe otra condición que cumplen las familias tradicionales de Wavelets, la cual es conocida como la *condición de ortogonalidad*; esta condición es particularmente útil cuando se requiere utilizar la descomposición obtenida en tiempo y frecuencia de una señal luego de aplicar la WT para reconstruir la señal original. Este proceso es formalmente conocido como la Transformada Wavelet Inversa (IWT), cuya definición formal se presenta a continuación.

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C(a, b)\psi_{a,b}(t) \frac{da db}{a^2} \quad (2.16)$$

Donde C_ψ corresponde al coeficiente de admisibilidad de la WT y $C(a, b)$ corresponde a los coeficientes wavelet de $f(t)$ para los parámetros $s = a$ y $u = b$.

En el caso en se requiera realizar una reconstrucción de la señal original es estrictamente necesario que las funciones madre construidas cumplan con la condición de ortogonalidad. Algunas aplicaciones de análisis de fallas en SEP que utilizan este tipo enfoques se pueden observar en [12], [46].

Por otro lado, en [50] Borghetti demuestra que no es necesario cumplir la condición de ortogonalidad en las funciones madre construidas cuando no se pretende hacer una reconstrucción de la señal original. Las funciones propuestas por Borghetti utilizan la información de la señal del evento a detectar para crear una función similar a la señal de falla. Con el fin de cumplir con las condiciones de admisibilidad, estas funciones se normalizan y centran, asegurando que su valor medio sea cero, además de introducir un factor exponencial decreciente que garantiza que la Ecuación (2.15) se cumple.

En cuanto a los resultados observados en [50], las funciones construidas presentan, en general, una mejor tasa de detección y clasificación de fallas eléctricas en relación con una Wavelet de una familia tradicional (Morlet) y son capaces de discriminar correctamente las frecuencias relacionadas a las reflexiones de onda esperados (en este caso se trabajó con el método de las Ondas Viajeras). En el apartado de conclusiones se destaca que un trabajo futuro consiste en garantizar la ortogonalidad de las funciones construidas con el fin de reconstruir la señal de falla original para cada uno de los componentes en frecuencia detectados. Esta metodología fue replicada en [24], llegando a conclusiones similares, aunque se detalló que el algoritmo propuesto parece tener una menor tasa de detección y clasificación de eventos al inicio de la línea debido a la relación inversa entre la frecuencia característica de las ondas generadas y la distancia a la que se generan.

Por otra parte, en [67] se generan funciones madre asociadas con la familia paraunitaria de bancos de filtros espejo en cuadratura (QMF), los cuales son los filtros digitales comúnmente utilizados en el algoritmo piramidal en la implementación de la DWT. En este caso se trabajó el método de las ondas viajeras con esquema de terminación doble para una línea de transmisión. La caracterización de las funciones madre se basa en la selección óptima de los coeficientes de Lattice QMF, los cuales determinan la función de transferencia de los filtros pasabajas y pasaaltas. En este caso las funciones propuestas

cumplen con los principios de admisibilidad y ortogonalidad por lo que es posible reconstruir la señal original a partir de la descomposición obtenida. En particular, se encontró que a comparación de las funciones madre de la familia Daubechies, los filtros propuestos presentan un menor retraso en la medida de los tiempos de llegada a los extremos de la línea de transmisión; también se observó una mejora en la característica de identificación de eventos de falla debido a la forma de los picos de las funciones madre propuestas y su baja tasa de oscilación en comparación con la familia Daubechies.

A modo de síntesis de la revisión del Estado del Arte para el tópico de análisis de fallas mediante el uso de la WT en SEP (tanto distribución como transmisión) se presenta la tabla 2-1. En esta es posible observar que la gran mayoría de aplicaciones se centran en el uso de la DWT debido a su implementación sencilla, mientras que el uso de la CWT u otros esquemas como la DT-CWT y WPT comprenden un campo reducido de investigación. Por otro lado, está claro que el uso de familias de funciones madre convencionales como Daubechies, Morlet y Coiflets es la preferencia de los investigadores, mientras que la generación de función adaptadas a las características de la señal de falla presenta una amplia posibilidad en investigación. Finalmente, es relevante resaltar que en múltiples ocasiones no se menciona la función o funciones madre utilizadas para evaluar la WT, aunque es pertinente suponer que en estos casos se utilizaron funciones de familiar convencionales debido a la poca relevancia que se le da a este aspecto.

Tabla 2-1: Síntesis de la revisión de la literatura relacionada al análisis de fallas mediante el uso de la WT en SEP. Elaboración personal.

Referencias	Esquema de implementación de Transformada Wavelet				Funciones Madre		
	DWT	CWT	CWT y DWT	Otro	Familia convencional	Funciones generadas	No especificada
[10], [35], [50], [51], [56], [59], [60], [61], [63], [64], [67], [68], [69], [70], [71], [73], [76], [77], [78], [80], [81], [82], [83]	x				x		
[55]	x					x	
[9], [10], [18], [52], [66], [72], [74]	x						x
[13], [16], [40], [41], [58], [62], [75], [79]		x			x		
[15], [42]		x				x	
[7], [65]		x					x
[43]			x				x
[6], [39], [44], [45],				x	x		
[47], [48]				x			x

2.4 Extracción de características

Regresando al contexto de detección y clasificación de eventos en señales digitales de sistemas de potencia mediante el uso de la WT es relevante resaltar que, en general, la información que se obtiene de forma directa al aplicar ésta no es la más adecuada para discriminar eventos de interés. Si bien los coeficientes (en el caso de la CWT) o las descomposiciones (en el caso de la DWT) wavelet obtenidos contienen información en tiempo y frecuencia de las señales analizadas, es común utilizar métodos que sinteticen esta información, bien sea para aumentar la capacidad de detección y clasificación del sistema o para reducir la dimensionalidad de la información obtenida cuando se trabajan entornos en los que la velocidad de procesamiento tiene una mayor relevancia.

La mayoría de trabajos analizados utiliza los coeficientes wavelet obtenidos de la descomposición discreta de forma directa ([16], [68]–[73]) dada su conveniencia, reportando resultados aceptables, aunque se han propuesto metodologías en los que se computan tanto los coeficientes DWT como los CWT como en [51] en el que la información discreta se usa para detectar si se presenta o no un evento y para localizar de forma aproximada el punto de falla, mientras que la CWT se utiliza para mejorar el grado de localización de ésta.

Según la literatura, existen múltiples acercamientos para la extracción de características de los coeficientes o descomposiciones wavelet en el contexto de interés. A continuación, se presentan los métodos más comúnmente referenciados y de mayor aplicabilidad para el objetivo de esta disertación:

2.4.1 Densidad de Potencia Espectral (PSD)

La Densidad de Potencia Espectral o PSD por sus siglas en inglés es una métrica comúnmente usada en el campo del procesamiento digital de señales dado que describe la potencia existente en una señal como función de la frecuencia, por unidad de frecuencia. En este contexto, la potencia de la señal puede corresponder a la potencia física de ésta, aunque existen casos abstractos en los que se extiende una definición de la potencia de una señal como el cuadrado del valor de ésta. En el caso particular de la WT la PSD se

define para cada uno de los niveles o escalas de frecuencia como se observa en la Ecuación (2.17) en función del tipo de coeficiente (detalle y aproximación).

$$\sum_t |f(t)|^2 = E_A + E_D \quad (2.17)$$

Donde E_A corresponde a la PSD contenida en los componentes de aproximación del último nivel de descomposición y E_D corresponde a la PSD contenida en los componentes de detalle, para todos sus niveles. Las fórmulas correspondientes a la PSD para los niveles de aproximación y detalle se presentan en las Ecuaciones (2.18) y (2.19), respectivamente.

$$E_A = \sum_k |a_j(k)|^2 \quad (2.18)$$

$$E_D = \sum_{j=1}^J \left(\sum_k |d_j(k)|^2 \right) \quad (2.19)$$

Donde:

- $f(t)$ corresponde a la señal original en el dominio del tiempo.
- $a_j(k)$ corresponde al coeficiente de aproximación en la escala de frecuencia j e instante k .
- $d_j(k)$ corresponde al coeficiente de detalle en la escala de frecuencia j e instante k .

En la literatura analizada se utiliza sólo la energía correspondiente a los coeficientes de detalle, dado que éstos contienen la información relacionada a eventos transitorios que se desean detectar y clasificar, en lugar de la información de la onda estacionaria contenida en los coeficientes de aproximación ([20], [61], [74]–[77]). En otros casos se ha optado por mapear la información obtenida de los coeficientes wavelet en imágenes en escala de grises ([48]) generando un escalograma aproximado, debido a las estructuras que se utilizaron para la clasificación de eventos.

2.4.2 Entropía Wavelet

La entropía es un concepto relacionado a la teoría de la información que mide la incertidumbre en estructuras y patrones de datos. Para cualquier evento, su incertidumbre

está relacionada con sus estados y las probabilidades respectivas de éstos [78]; para un espacio muestral $\{x_1, x_2, \dots, x_n\}$, cada porción de información tiene una probabilidad $P(x_i) = P_i$, la entropía clásica (o de Shannon) H se define como:

$$H = -\sum_i P_i \log_a(P_i) \quad (2.20)$$

En el caso de la WT, existen múltiples tipos de Entropía Wavelet (WE) [79], entre los que se destacan:

- Entropía de Energía Wavelet (WEE): En este caso las probabilidades de cada estado están relacionadas a la razón entre la energía de una escala particular de descomposición y la energía total de la señal.
- Entropía de Tiempo Wavelet (WTE): Esta clase de entropía es comúnmente usada en sistemas con ventanas móviles, dado que las probabilidades de los estados están relacionadas a la probabilidad de que un coeficiente para una escala particular esté contenido en cierta ventana de tiempo.
- Entropía Singular Wavelet (WSE): Las probabilidades de los estados se computan utilizando los valores singulares o propios que se generan al crear la Matriz Wavelet (referirse a la sección 2.4.5) y hallando la razón en entre el valor propio de cada escala de frecuencia y la sumatoria de éstos para todas las escalas.
- Entropía Tiempo-Frecuencia Wavelet (WTFE): Corresponde a un mapeo discreto de las WEE y WTE considerando las escalas diádicas de frecuencia y los coeficientes wavelet para cada uno de los puntos del tiempo discreto contenidos en una de las ventanas móviles.
- Entropía Wavelet Promedio (WAE): Es una extensión de la WEE, en la que las energías de las ventanas móviles se normalizan con respecto a la energía total de la señal antes de computar el valor de Entropía. Este método indica de forma satisfactoria la complejidad promedio de la señal analizada, aunque existe una pérdida considerable en cuanto a la detección de variaciones en tiempo y frecuencia.
- Entropía Wavelet de Distancia (WDE): Este tipo de entropía utiliza la distancia entre cada uno de coeficientes wavelet para cada uno de los puntos del tiempo discreto

contenidos en una de las ventanas móviles en conjunto con el valor mismo del coeficiente. Es particularmente útil en series de tiempo.

Si bien el concepto de WE tiene un desarrollo significativo, su aplicación para la detección y clasificación de fallas en sistemas eléctricos de potencia no ha sido muy investigada; sin embargo, se destaca que en [80] se realizaron estudios para demostrar la validez de la WSE en este contexto, obteniendo mejoras notables en comparación con respecto a la PSD.

2.4.3 Parámetros estadísticos

El uso de parámetros o indicadores estadísticos es uno de los métodos de extracción de características más utilizados en la literatura analizada ([17], [18], [80]–[84]), dado que comprende varios métodos que se resumen, de forma sucinta, en utilizar algún indicador estadístico para sintetizar la información o resaltar alguna característica deseada. Siendo rigurosos, tanto la PSD como la WE son extensiones de esta categoría, aunque se han destinado a otras categorizaciones bien por su amplio uso como en el caso de la PSD (la cual ha demostrado ser capaz de extraer las características de las descomposiciones Wavelet satisfactoriamente) o por su particularidad y potencial investigativo futuro como en el caso de la WE.

Según la literatura consultada, los parámetros estadísticos más utilizados son la media, valor cuadrático medio (RMS), los momentos estadísticos de N-ésimo orden, la desviación estándar y la curtosis y asimetría. Las fórmulas de la estadística descriptiva con las que tradicionalmente se computan estos parámetros para un conjunto de valores se extiende al contexto de la WT al aplicarlas sobre los coeficientes wavelet de cada una de las escalas de frecuencia, en las ventanas de tiempo predefinidas.

Otros trabajos han propuesto utilizar otros indicadores poco convencionales como el coeficiente de alienación, propio de las series de tiempo, para extraer la información de los CW, reportando resultados satisfactorios [85].

2.4.4 Máximos locales y Módulo Máximo Wavelet

La teoría detrás del uso de los máximos locales para extraer características de los coeficientes wavelet se soporta por el hecho de que las singularidades que normalmente

se asocian a los eventos transitorios que se buscan detectar o clasificar, generan formas particulares en las descomposiciones discretas wavelet (DWT) y escalogramas (CWT). Se ha observado que para los eventos transitorios electromagnéticos en los sistemas eléctricos de potencia (y otros casos externos a la aplicación de esta disertación), la descomposición de las señales en sus niveles o coeficientes Wavelet presenta picos afilados los cuales pueden ser localizados en tiempo y escala. Algunos autores han utilizado estos máximos locales de forma directa para alimentar sus sistemas de detección y clasificación de eventos ([7], [62], [86]) reportando resultados positivos.

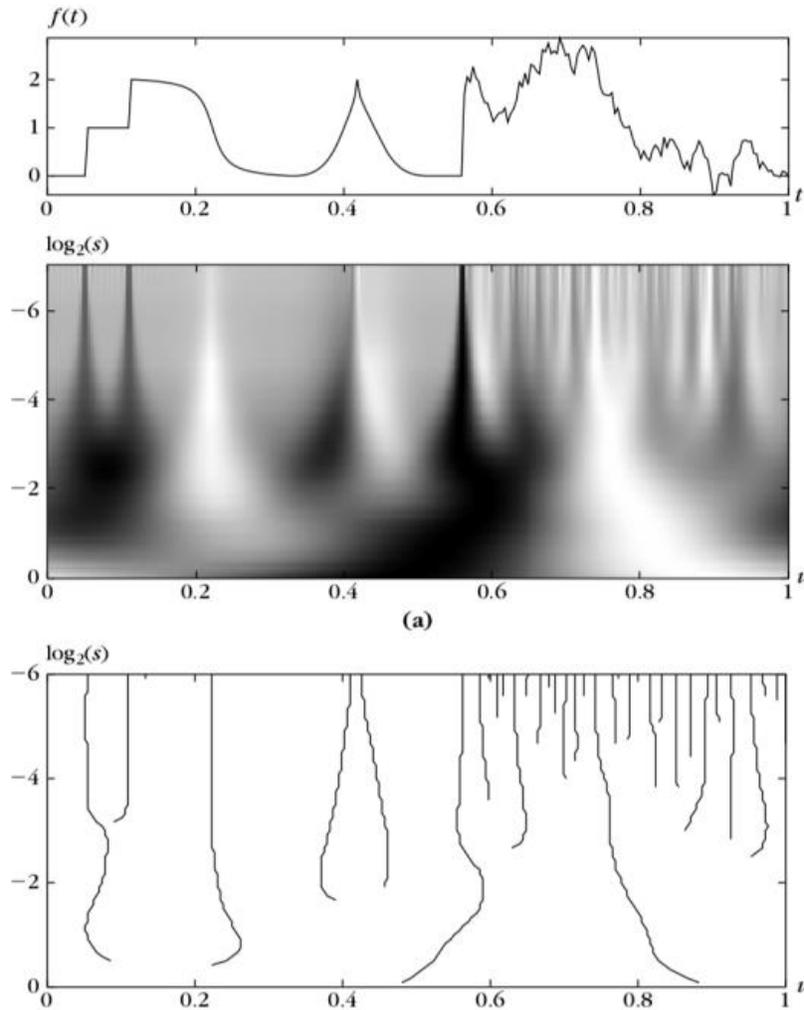
La detección de los máximos locales en el dominio del tiempo y frecuencia se fundamenta en la siguiente definición:

$$|W_f(s_0, x)| \leq |W_f(s_0, x_0)| \quad (2.21)$$

Donde $W_f(s_0, x)$ corresponde al coeficiente resultante de la WT en la escala s_0 e instante x .

Otro acercamiento que ha sido usado en aplicaciones que utilizan la CWT para hallar los coeficientes wavelet de las señales ([87], [88]) es del Módulo Máximo Wavelet (WMM) y está ligado a una variante particular de la WT llamada Transformada Wavelet de Módulo Máximo (WTMM). En términos sucintos, se busca crear “esqueletos” en el espectrograma asociado a la descomposición en escalas wavelet hallando los máximos locales en una escala o nivel de descomposición de referencia (normalmente la primera escala); este proceso se repite para las escalas más finas y se determina si los máximos locales se “alinean” [35]. Una demostración gráfica del uso del WMM se observa en la figura 2-9, en la que se presenta una señal de referencia $f(t)$ (arriba), su escalograma en escala de grises (medio) y los esqueletos resultantes (abajo); en este caso, se puede observar cómo los cambios abruptos en la señal generan WC de mayor magnitud (representados por valores más oscuros), los cuales se traducen en esqueletos más largos.

Figura 2-9: Representación de una señal en el tiempo (arriba) mediante la WT en un escalograma en escala de grises (medio) y los esqueletos correspondientes al WMM (abajo).



Nombre de la fuente: Digital Signal Filtering, Analysis and Restoration [95].

Esta alineación de las escalas está inmersa en la incertidumbre propia de las resoluciones en tiempo y frecuencia de la WT, por lo que no existe un método único para detectarla, aunque se han propuesto las siguientes alternativas:

- Utilizar un delta de tiempo (hacia adelante y atrás) desde el punto de localización de los máximos locales, con el fin de alivianar la rigurosidad en el proceso de alineación. Este delta de tiempo dependerá de las características del sistema como la frecuencia de muestreo, además de la incertidumbre en tiempo y frecuencia asociada a la WT.
- Declarar un umbral de profundidad de los esqueletos, medido en escalas o niveles, a partir del cual los eventos se pueden considerar como singularidades de interés.

Esto permite detectar eventos de menor magnitud, aunque está sujeto a la escogencia de un umbral, lo que requiere experimentación previa con la profundidad de los esqueletos para diferentes eventos.

2.4.5 Modos de correlación Wavelet

Este método de extracción de características es propio de la DWT dado que utiliza los coeficientes de detalle de los N niveles de descomposición y el último coeficiente de aproximación discreta de las señales [90]. Con estos se crea una matriz llamada Matriz Wavelet (\bar{M}) como se observa a continuación

$$\bar{M}_{N+1,n} = \begin{bmatrix} cD_1(1) & cD_1(2) & \cdots & cD_1(n) \\ cD_2(1) & cD_2(2) & \cdots & cD_2(n) \\ \vdots & \vdots & \ddots & \vdots \\ cD_N(1) & cD_N(2) & \cdots & cD_N(n) \\ cA_N(1) & cA_N(2) & \cdots & cA_N(n) \end{bmatrix} \quad (2.22)$$

A continuación, se utiliza la teoría del álgebra lineal y se aplica la descomposición en valores singulares (SVD) en la matriz \bar{M} con el fin de analizar la información contenida en ésta identificando factores subyacentes difíciles de identificar en la matriz original. Según la teoría de SVD, la matriz M se puede descomponer como:

$$\bar{M} = U \cdot \bar{\lambda} \cdot V^T \quad (2.23)$$

Donde:

- U es una matriz N+1 x r, cuyas columnas son ortogonales.
- $\bar{\lambda}$ es una matriz diagonal r x r.
- V es una matriz r x n, cuyas filas son ortogonales.

$$\bar{\lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \lambda_r \end{bmatrix} \quad (2.24)$$

Los Modos de Correlación Wavelet (WCM) están relacionados a los valores propios o singulares contenidos en la principal diagonal de la matriz λ dado que éstos son una medida directa del contenido en frecuencia de cada uno de los niveles de descomposición. De igual forma, cuando ocurre un cambio significativo o abrupto en el estado del sistema analizado, las altas frecuencias generadas se ven reflejadas en los valores singulares, particularmente en los primeros valores (λ_1 es el valor más significativo, seguido de λ_2 y los demás).

2.5 Sistemas de detección y clasificación de eventos

En términos generales, la información obtenida mediante la extracción de características puede utilizarse para validación de metodologías, detección de eventos, clasificación de eventos y localización de eventos. En la literatura analizada, la finalidad de la aplicación de la WT y la extracción de características depende de los objetivos particulares de los autores o patrocinadores de los trabajos.

Considerando el objetivo principal de esta disertación, nos centraremos en la detección y clasificación de eventos utilizando la información extraída al aplicar la WT y los métodos de extracción de características. Así, se destacan dos categorías amplias que concentran la mayoría de los sistemas reportados en la literatura.

2.5.1 Umbrales de detección y clasificación

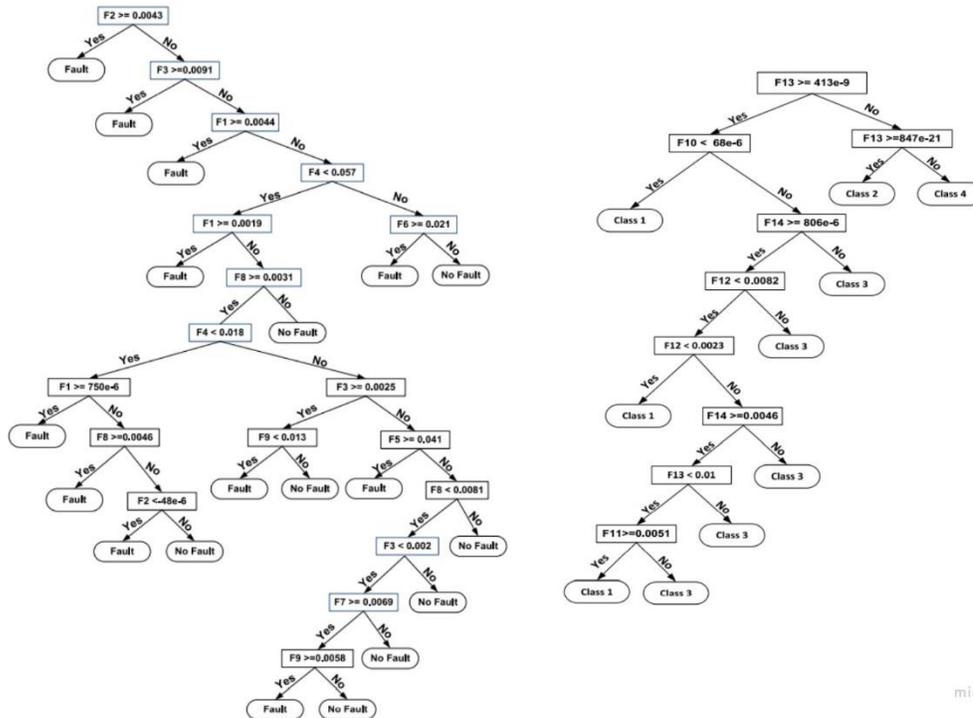
El uso de umbrales para la detección y/o clasificación de eventos transitorios comprende múltiples acercamientos, los cuales requieren de cierto conocimiento previo del comportamiento del sistema o los valores obtenidos en la etapa de extracción de características y su relación con los eventos de interés.

Algunos trabajos han utilizado directamente los coeficientes obtenidos en los niveles o escalas de descomposición de ciertas variables como las corrientes de secuencia en conjunto con umbrales conocidos del sistema para determinar si se presenta o no un evento, además de clasificarlo. En [55] por ejemplo, se utilizan las corrientes y tensiones de secuencia cero para determinar la polaridad de los eventos simulados con el fin de determinar en qué ramal del sistema de distribución utilizado se presenta el evento.

En otros casos como [77], [80], [91], [92] los valores extraídos de la etapa de extracción de características para un conjunto de señales de muestras se analizan con el fin de determinar umbrales de decisión. Estos umbrales están directamente relacionados a los métodos de extracción de características y varían sustancialmente entre los autores analizados, aunque se comparte un objetivo común de comparar los valores correspondientes a una señal de interés con los umbrales predeterminados para decidir si el evento es o no una falla eléctrica y determinar el tipo de ésta, en caso afirmativo.

Generalmente, las metodologías utilizadas en la detección y clasificación basada en umbrales utilizan estructuras de decisión anidadas para mejorar su desempeño y robustez al incluir más variables de decisión. Este tipo de estructuras llamado árboles de decisión (DT), es análogo a un árbol biológico, en el que existe un tronco con una decisión y umbral iniciales, del cual se derivan ramas con pares de decisión y umbrales propios, repitiéndose hasta llegar una decisión final. En cuanto a la clasificación y detección de eventos utilizando las características extraídas luego de procesar las señales de interés, es común separar estas categorías en árboles de decisión diferentes; un ejemplo de esta metodología se presenta en la figura 2-10, en la que los autores decidieron utilizar múltiples parámetros estadísticos como entropía, PSD y desviación estándar y umbrales obtenidos de la observación del comportamiento de estos parámetros ante diferentes eventos simulados, para crear dos árboles de decisión, el primero (izquierda) para la detección de fallas y el segundo (derecha) para la clasificación de éstas.

Figura 2-10: Ejemplo de árboles de decisión para detección (izquierda) y clasificación (derecha) de fallas.



Nombre de la fuente: A combined wavelet and data mining based intelligent protection scheme for microgrid [93].

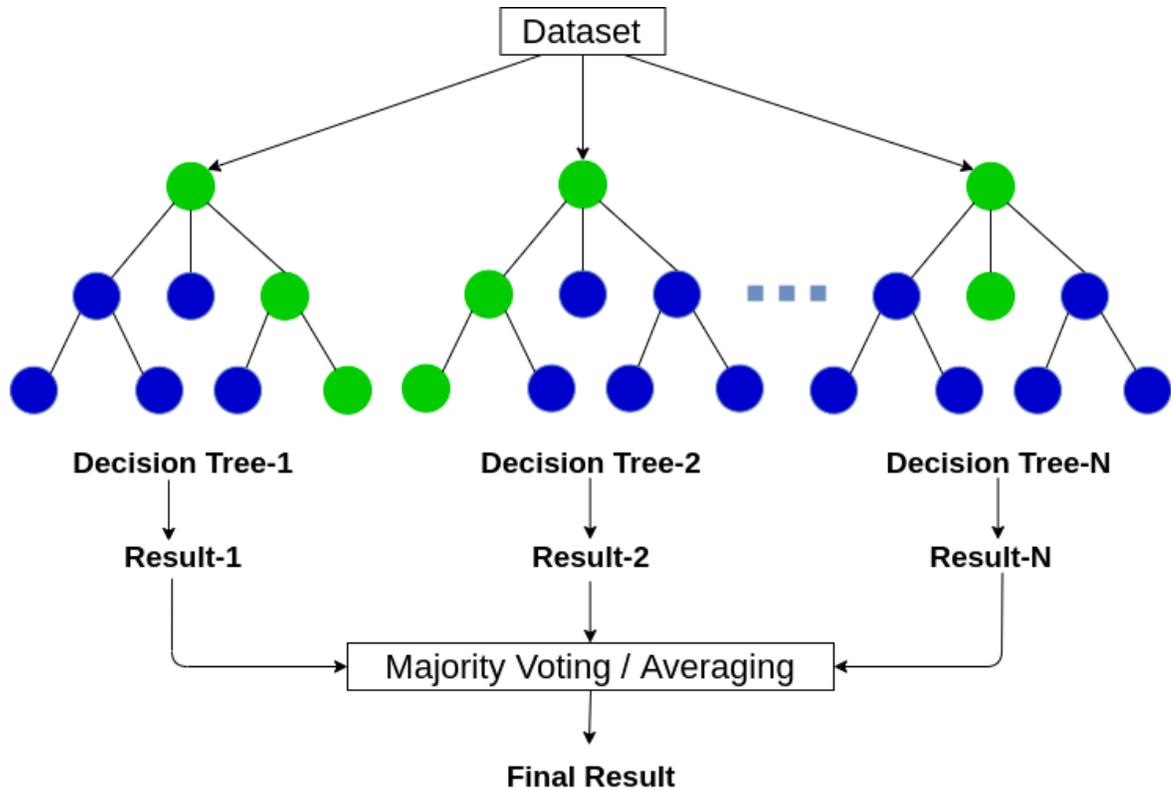
Es relevante resaltar que la complejidad de los esquemas de decisión basados en umbrales crece sustancialmente con la introducción de más parámetros de falla como el ángulo de carga, impedancia de falla, ángulo de inyección de la falla, tipo de falla (y fases falladas), localización de la falla, condiciones de pre-falla y magnitud de la carga, entre otras. En estos casos, aunque es posible implementar esquemas basados en umbrales, es recomendable analizar el sistema y los valores asociados a algunos eventos para obtener familiaridad con los mismos, aunque la complejidad del problema conlleva a la mayoría de los autores a recurrir a otro tipo de alternativas.

2.5.2 Algoritmos de clasificación basados en Machine Learning e Inteligencia Artificial

Dentro de esta categoría de sistemas para la detección y/o clasificación de eventos en el contexto de esta disertación, existen múltiples metodologías reportadas en la literatura entre las que se destacan:

- **Bosques Aleatorios (RF):** Se definen como una colección de estructuras de clasificación basadas en árboles de decisión, en los que cada estructura arroja la salida más probable según sus parámetros y entrenamiento, y se decide sobre ésta utilizando un voto basado en pesos [94] como se observa en la figura 2-11. Este tipo de estructuras suele ser robusta ante el sobreajuste siempre y cuando se utilice un número suficiente de nodos o estructuras. En [5], [63] se utilizaron estos algoritmos, reportando buenos resultados en comparación con los árboles de decisiones convencionales.

Figura 2-11: Ejemplo de Random Forest con N árboles de decisiones.

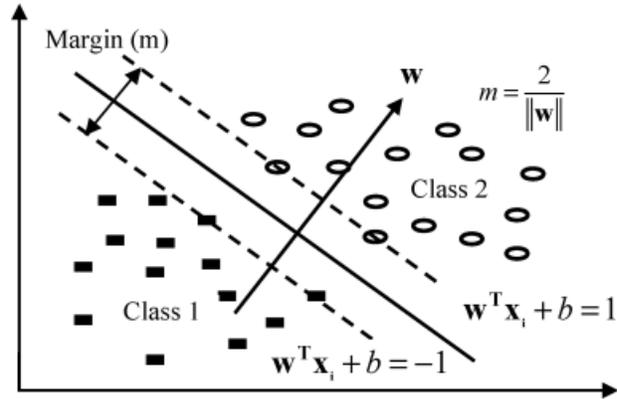


Nombre de la fuente: Random Forests Understanding [95].

- Máquinas de Soporte Vectorial (SVM):** En problemas de clasificación, las SVM buscan un hiperplano para separar la información de acuerdo con la clase a la que pertenecen, de modo que esta separación sea máxima; dado que esta separación máxima no puede ser garantizada, por lo que este hiperplano ideal se conoce como el hiperplano óptimo (figura 2-12). Utilizando una función de costo que busca maximizar la distancia entre clases al hiperplano modificando los pesos y bias que definen la curva del hiperplano [96]. Esto se logra buscando vectores que satisfacen una serie de restricciones que minimizan el número de clases clasificadas incorrectamente, conocidos como vectores de soporte, los cuales determinan la superficie de decisión o el hiperplano de separación de las clases. Este tipo de sistemas de clasificación se utilizó en [68], [82] para comparar su eficiencia con respecto a otros métodos como Redes Neuronales Artificiales resultando en tasas

de detección similares, aunque el porcentaje de clasificación exitoso fue menor al reportado por otros métodos.

Figura 2-12: Hiperplano óptimo para clasificación utilizando SVM.



Nombre de la fuente: Combined wavelet-SVM technique for fault zone detection in a series compensated transmission line [68].

- **Redes Neuronales Artificiales (ANN):** Las redes neuronales artificiales han sido ampliamente utilizadas en contextos investigativos, comerciales e industriales en las últimas décadas en problemas de clasificación y ajuste o regresión de datos. Este tipo de algoritmo se fundamenta en las redes neuronales biológicas, en las que un conjunto de neuronas se conecta de cierto modo y se alimenta con información conocida con el fin de entrenarse y poder, eventualmente, generalizar el conocimiento obtenido en su entrenamiento para clasificar información nueva. En la literatura analizada, las ANN son uno de los algoritmos de aprendizaje y clasificación más utilizados ([20], [73], [75], [82]–[84], [97], [98]) debido en gran parte al extenso trasfondo investigativo que existe en torno a éstas, además de su efectividad para clasificar eventos.

Existen múltiples tipos de ANN, aunque el más utilizado es el perceptrón multicapa, el cual se adhiere más fielmente a su analogía con las redes biológicas y se destaca por su facilidad de ajuste. Por otro lado, otro tipo de redes neuronales como las redes neuronales probabilísticas (PNN) y las redes neuronales convolucionales (CNN), han visto aplicaciones limitadas debido a la naturaleza de la información obtenida en la etapa de extracción de características, la cual corresponde a una

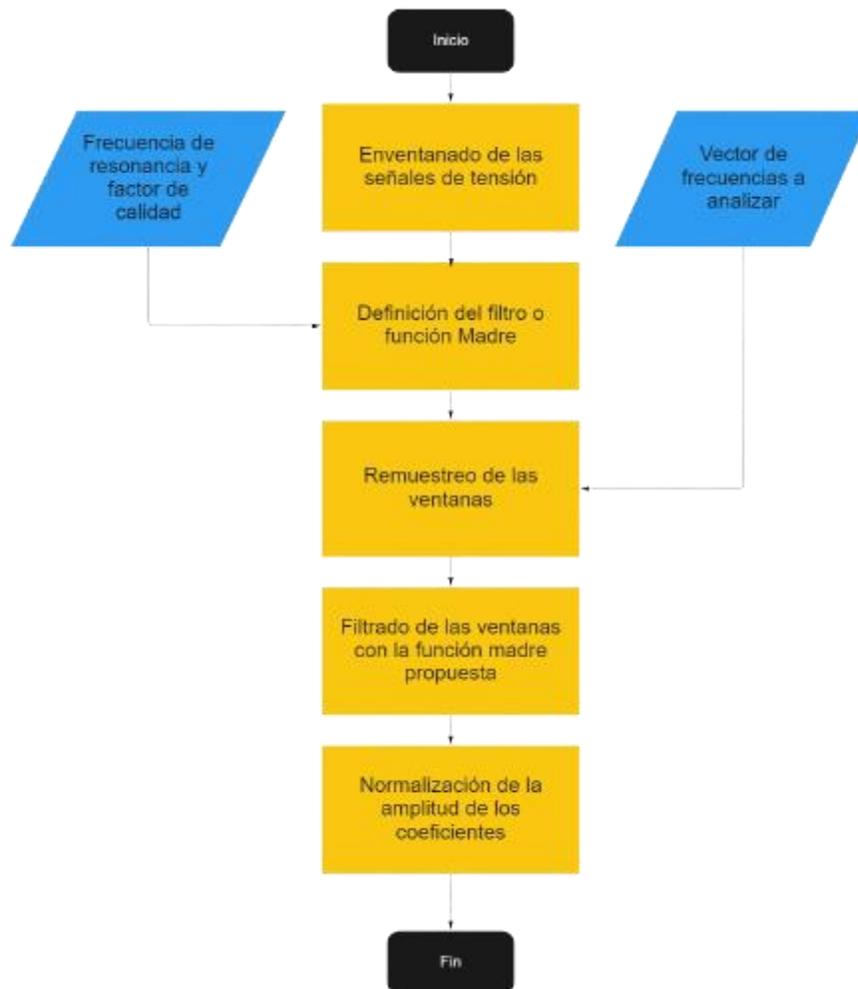
serie de características que definen a los eventos de interés. Sin embargo, algunos autores han utilizado este tipo de algoritmos [18], [81] y [99] en el caso de las PNN y CNN, respectivamente, reportando resultados positivos.

Finalmente, otros autores han propuesto el uso de otros algoritmos de clasificación como el k vecinos más cercanos (k-NN) [13], lógica difusa [100] y clasificadores bayesianos [17] aunque no se han observado publicaciones recientes que soporten el uso de éstos.

3.Propuesta metodológica

A modo de introducción de la propuesta metodológica relacionada a la computación de la Transformada Wavelet Continúa utilizando funciones madre de soporte infinito se presenta el diagrama de flujo de la figura 3-1. Cada una de las subsecciones de este capítulo está relacionada a un conjunto de bloques y su fundamento teórico.

Figura 3-1: Diagrama de flujo de la propuesta metodológica. Elaboración propia.

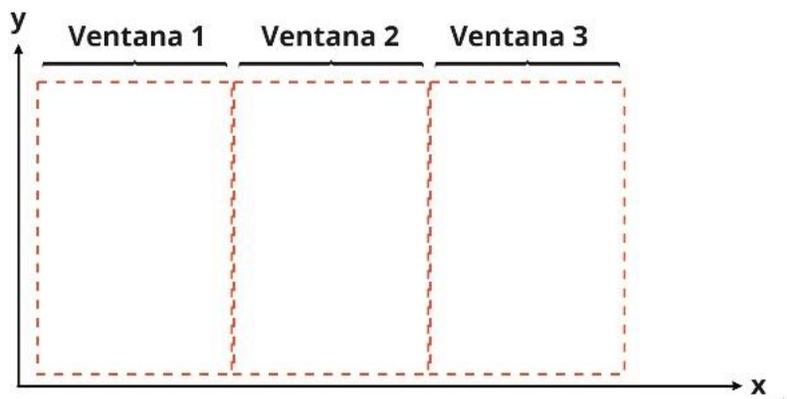


Nombre de la fuente: Elaboración propia.

3.1 Preprocesamiento digital de la señal

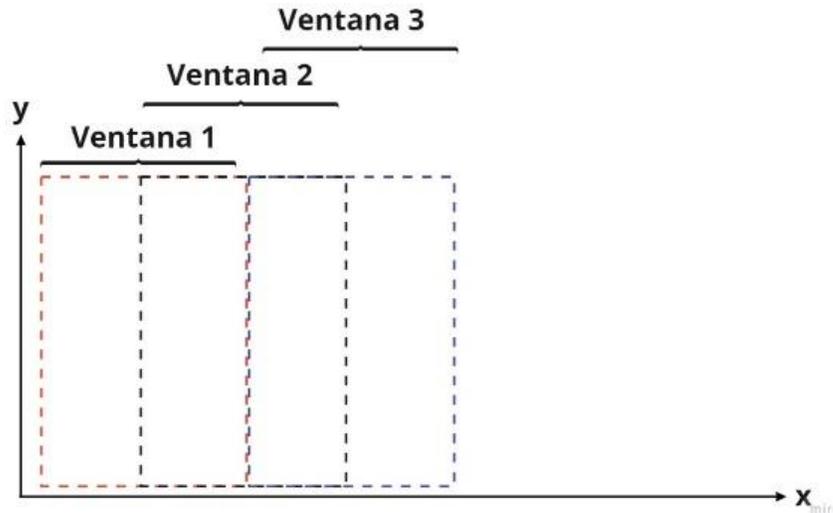
Con el fin de representar un sistema de transmisión de información en una subestación moderna en la que los dispositivos de medida y canales de información contemplan un modelo de suscripción/ subscriptor mediante el uso de tramas que contienen la información de los valores medidos de forma digital, las señales simuladas son descompuestas en segmentos de longitud constante M . Ahora bien, estos segmentos de la señal podrían tomarse de forma consecutiva o considerando un traslapo sobre el segmento previo. El concepto de traslapo puede definirse como el porcentaje de un segmento i -ésimo de la señal original que se repite en el segmento $i+1$ -ésimo; en este sentido, un traslapo de 0% significa que no hay repetición alguna entre segmentos consecutivos (figura 3-2), mientras que un traslapo del 100% se define como la repetición de todos datos del segmento anterior exceptuando el primero dado que el nuevo segmento debe moverse en el tiempo para tomar una nueva muestra (figura 3-4). El traslapo de ventanas al 50% se ejemplifica de forma gráfica en la figura 3-3.

Figura 3-2: Ventanas de longitud fija con traslapo del 0%.



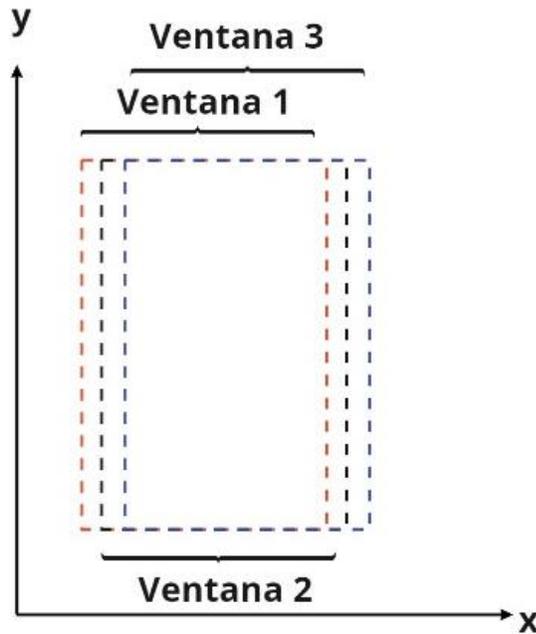
Nombre de la fuente: Elaboración propia.

Figura 3-3: Ventanas de longitud fija con traslapo del 50%.



Nombre de la fuente: Elaboración propia.

Figura 3-4: Ventanas de longitud fija con traslapo del 100%.



Nombre de la fuente: Elaboración propia.

Es relevante resaltar que utilizar transformaciones o métodos de procesamiento digital de señales sobre los segmentos generados de la señal no es equivalente a hacerlo sobre la señal completa debido a los efectos de frontera que se generan en los límites de los

segmentos, similar al fenómeno de Gibbs que se presenta al computar una reconstrucción mediante coeficientes de Fourier [101].

Así, para minimizar las distorsiones ocasionadas por los efectos de frontera introducidos al segmentar la señal es común utilizar señales de enventanado o Windowing functions. Estas funciones deben cumplir con una serie de propiedades (en el dominio del tiempo o de la frecuencia), que pueden resumirse, sin pérdida de generalidad [101] para el alcance de la presente investigación en:

- La función de enventanado $f(t)$ debería ser real y no-negativa.
- La función de enventanado $f(t)$ debería ser par. Esto implica que la transformada de Fourier de la función de enventanado $F(j\omega)$ es real.
- La función de enventanado $f(t)$ debería alcanzar su máximo global en $t = \tau$, donde τ corresponde al punto en el que se centraliza la ventana.
- La transformada de Fourier de la función de enventanado $F(j\omega)$ debería tener un lóbulo principal en el origen y lóbulos menores a cada lado.
- El lóbulo principal debería ser lo más angosto posible y contener la mayor energía posible de la señal.

Dadas estas propiedades, múltiples funciones de enventanado han sido propuestas, tales como:

- A. **Función rectangular o Box Car:** Es la función de enventanado más simple (y antigua, propuesta formalmente en 1958 por Blackman y Tuckey) dado que consiste en una ventana uniforme [102] definida como:

$$f(t) = \begin{cases} 1, & |t| \leq \tau \\ 0, & |t| > \tau \end{cases} \quad (3.1)$$

$$F(j\omega) = \frac{2\tau \cdot \sin(\omega\tau)}{\omega\tau} \quad (3.2)$$

B. Función Triangular o de Barlett: Esta función de enventanado se obtiene al realizar una convolución lineal de dos ventanas rectangulares de media duración [102]. Su definición formal está dada por:

$$f(t) = \begin{cases} 1 - \frac{|t|}{\tau}, & |t| \leq \tau \\ 0, & |t| > \tau \end{cases} \quad (3.3)$$

$$F(j\omega) = \tau \left[\frac{\sin\left(\frac{\omega\tau}{2}\right)}{\frac{\omega\tau}{2}} \right]^2 \quad (3.4)$$

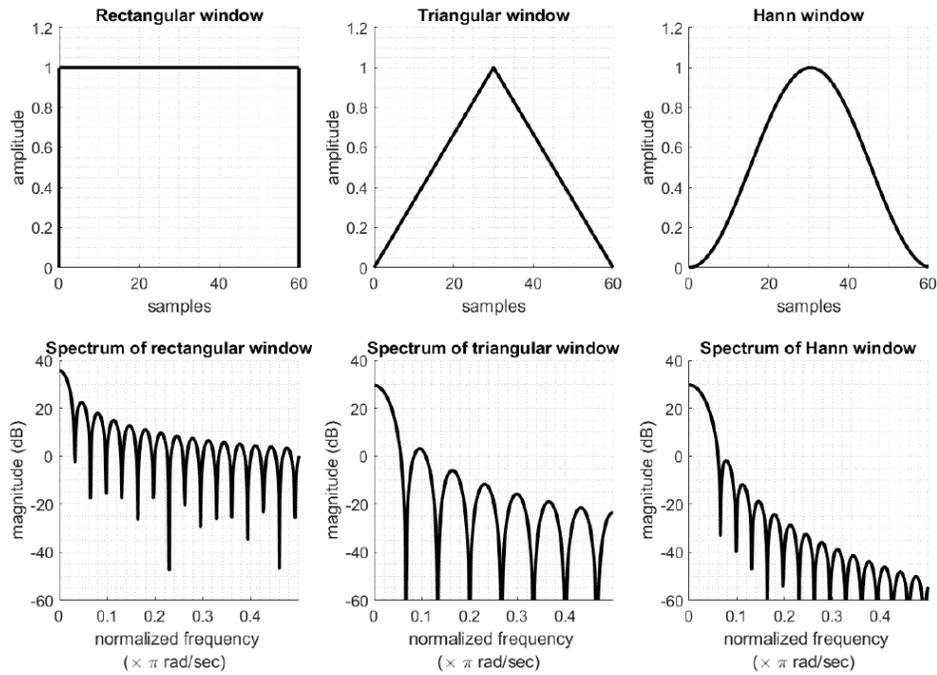
C. Función de Han o Coseno Elevado: Esta función fue propuesta por Julius von Hann, un meteorólogo Austriaco, y es una de las ventanas más comúnmente aplicadas en la literatura [103]. Su definición en el tiempo y frecuencia está dada por:

$$f(t) = \begin{cases} \frac{1}{2} \left(1 + \cos\left(\frac{\pi t}{\tau}\right) \right), & |t| \leq \tau \\ 0, & |t| > \tau \end{cases} \quad (3.5)$$

$$F(j\omega) = \frac{\sin(\omega\tau)}{\omega\tau} + \frac{1}{2} \frac{\sin\left(\left(\omega + \frac{\pi}{\tau}\right)\tau\right)}{\left(\omega + \frac{\pi}{\tau}\right)} + \frac{\sin\left(\left(\omega - \frac{\pi}{\tau}\right)\tau\right)}{\left(\omega - \frac{\pi}{\tau}\right)} \quad (3.6)$$

La gráfica de estas funciones, tanto en el dominio del tiempo como en el de la frecuencia, puede observarse en la figura 3-5.

Figura 3-5: Representación en el dominio del tiempo (arriba) y de la frecuencia (abajo) para las funciones de enventanado rectangular, triangular y de Hann, de izquierda a derecha.



Nombre de la fuente: On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform [97].

3.2 Análisis Multirresolución utilizando filtros digitales con respuesta infinita al impulso

Los filtros digitales son sistemas que operan sobre una señal muestreada de forma discreta para modificar (mejorar, disminuir o eliminar) ciertas características de ésta [105]. Como su nombre lo indica, son la versión en tiempo discreto de los filtros análogos comúnmente utilizados en aplicaciones electrónicas como sistemas de sonido y radios; por lo general, muchos de los conceptos que se utilizan en el contexto de filtros analógicos pueden ser extrapolados a su versión digital como su clasificación basada en su aplicación de atenuación de frecuencias (filtros pasabandas, rechazabandas, pasaaltas y pasabajas).

Una de las formas más comunes de representar un filtro digital es mediante su ecuación en diferencias finitas, la cual está compuesta por los coeficientes digitales del filtro y expresa la relación de la salida actual del filtro con respecto a la entrada actual, además

de las entradas y salidas pasadas, según la memoria del filtro o sistema. La forma canónica de la ecuación en diferencias [106] se presenta en la Ecuación (3.7):

$$y(n) = \sum_{k=0}^M b_k x(n-k) - \sum_{k=1}^N a_k y(n-k) \quad (3.7)$$

Donde:

- $x(n)$ y $y(n)$ corresponden a las secuencias de entrada y salida del filtro digital, respectivamente.
- Los coeficientes b_k son denominados coeficientes feedforward del filtro.
- Los coeficientes a_k son denominados coeficientes de realimentación del filtro o feedback coefficients.
- Los enteros M y N hacen referencia al número de coeficientes del filtro y se utilizan para determinar el orden del mismo (el orden del filtro se define como el mayor entre estos dos enteros). Un filtro de mayor orden tendrá una mejor respuesta en frecuencia (anchos de banda más angostos, o filtros más selectivos), aunque requiere una mayor cantidad de recursos (físicos en el caso de filtros analógicos) computacionales para su implementación.

Por otro lado, al manipular la ecuación en diferencias es posible obtener una representación en términos de la función de transferencia del sistema (Ecuación (3.8)), la cual es particularmente útil en aplicaciones de control.

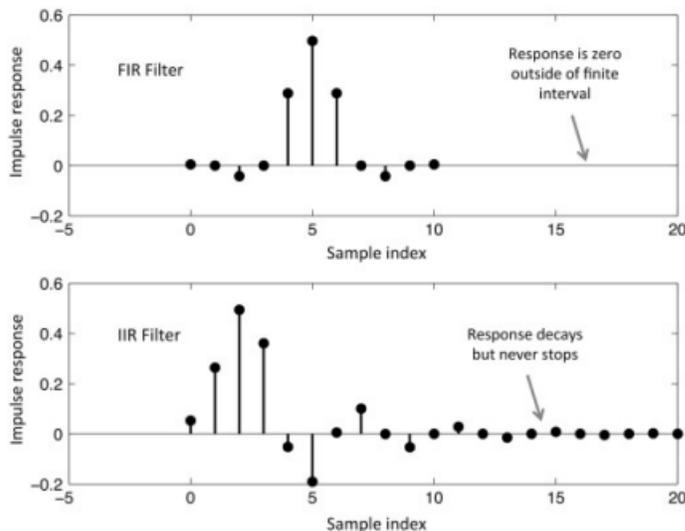
$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^M b_i z^{-i}}{1 + \sum_{k=1}^N a_k z^{-k}} \quad (3.8)$$

Cuando en un filtro digital todos los coeficientes de realimentación son nulos (en otras palabras, no existen polos), se dice que es un filtro con Respuesta Finita al Impulso o FIR. La respuesta al impulso de un filtro digital se determina utilizando una entrada tipo escalón correspondiente al Delta de Kronecker $\delta(n)$ y comúnmente se denota como $h[k]$ o h_k [105]. Claramente, la respuesta al impulso de un filtro digital cobra relevancia cuando se pretende identificar las características de éste, dado que si h_k es exactamente igual a la secuencia de coeficientes del filtro se puede afirmar que el sistema es LTI y FIR [107].

Por otro lado, si el filtro tiene alguna realimentación en su topología, el denominador en su ecuación de diferencias finitas no es unitario, lo que implica que el sistema es recursivo y, por ende, una salida del sistema en un tiempo determinado dependerá tanto de las entradas actuales y pasadas, como de salidas pasadas. Este último tipo de filtros se conocen como filtros con respuesta infinita al impulso o IIR [108]. En éstos, cada uno de los términos de realimentación en la ecuación de transferencia constituyen un polo en el dominio de la frecuencia y producen un pico de resonancia que puede observarse en la respuesta en amplitud.

La diferencia de los filtros FIR e IIR en el dominio del tiempo puede observarse claramente en la figura 3-6, en la que se observa como los filtros FIR (arriba) tienen una respuesta de soporte finito, que decae a cero por fuera de un intervalo finito, mientras que los filtros IIR tienen una respuesta que, también, decae, aunque de forma asintótica.

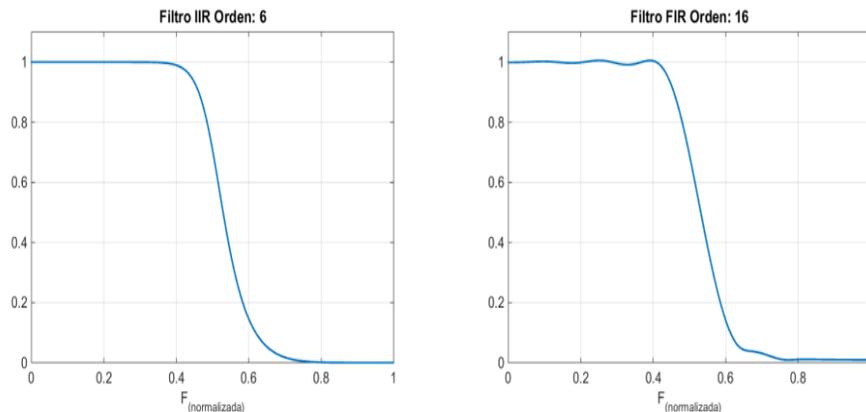
Figura 3-6: Respuesta a una secuencia impulsiva de entrada para un filtro FIR (arriba) y un filtro IIR (abajo).



Nombre de la fuente: Passive, active, and digital filters [101].

En general, los filtros IIR requieren un orden mucho menor que un equivalente FIR para obtener respuestas a la frecuencia en fase y amplitud similares, como puede observarse en la figura 3-7; sin embargo, los filtros FIR son siempre estables y se puede garantizar que su respuesta en fase es lineal [105].

Figura 3-7: Respuesta en fase para un filtro FIR de orden 6 (izquierda) y un filtro IIR de orden 16 (derecha).



Nombre de la fuente: Elaboración propia.

Regresando a lo discutido en la sección 2.2, la Ecuación (2.12) expresa la relación entre la respuesta de un sistema LTI ante una entrada tipo impulso con el teorema de convolución. De igual modo, se mencionó que un filtro FIR puede aplicarse a un sistema utilizando la convolución en el tiempo (equivalente a una multiplicación en el dominio de la frecuencia), aunque este acercamiento no puede ser extendido de forma trivial a los filtros IIR. En particular, debido al soporte finito de los filtros FIR, la convolución resulta en un número finito de coeficientes correspondientes a la respuesta del sistema ante el filtro, mientras que los filtros IIR tienen respuestas de longitud infinita ante entradas de tipo impulso por lo que la convolución no puede calcularse de forma trivial.

Un primer acercamiento para evaluar el comportamiento de los filtros IIR es utilizar una aproximación con coeficientes finitos; esto es, conociendo la longitud de entrada de la señal del sistema (número de muestras discretizadas), es posible tomar un número de coeficientes del filtro IIR tal que la longitud de la señal del sistema sea despreciable, o bien

que la cantidad de coeficientes del filtro aparente ser infinita. Siendo estrictos, esto corresponde a la aplicación de un filtro tipo FIR de muy alta dimensión, por lo que no es muy conveniente.

Por otro lado, podemos recurrir a la teoría de la WT para extender el concepto de los filtros IIR y proponer una metodología particular para utilizarlos; como se observó en capítulo 2, sección 3 (Funciones madre: Observaciones desde la teoría de filtros digitales), las funciones madre que se utilizan para computar la Transformada Wavelet $\psi_{u,s}$ pueden ser consideradas como un par de filtros digitales, uno pasabajas y otro pasaltas en el caso de la Transformada Wavelet Discreta, mientras que, en su contraparte Continua, las funciones madre corresponden a filtros pasabandas; estos filtros son escalados por un factor s y retrasados en el tiempo por un factor u .

Las familias wavelet clásicas como las mencionadas en el en capítulo 2, sección 3 (Funciones madre: Observaciones desde la teoría de filtros digitales), corresponden en su totalidad a filtros FIR debido a la necesidad de utilizar una convolución discreta entre la señal y el filtro tal y como se observa en las Ecuaciones (2.11) y (2.13). En este sentido, utilizar filtros tipo IIR para computar la WT no es posible mediante métodos tradicionales (a menos que se utilicen aproximaciones como la mencionada anteriormente).

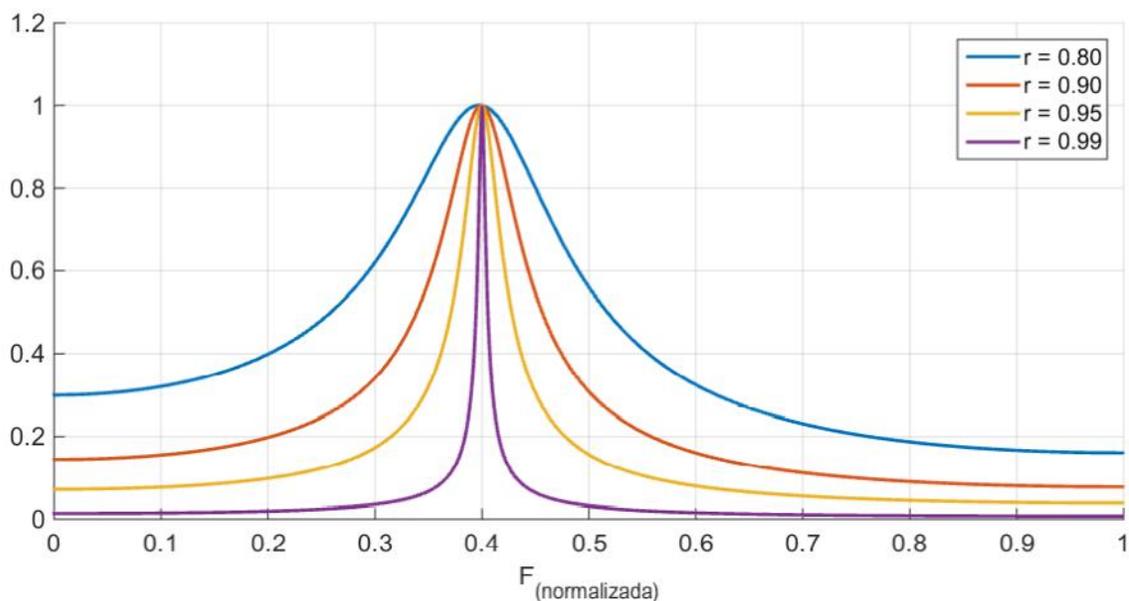
Así, se propone utilizar una metodología similar a la empleada para computar la CWT de forma tradicional incluyendo modificaciones tales que se puedan utilizar filtros tipo IIR como funciones madre o sus análogas a la metodología propuesta. En primera instancia, es pertinente proponer un filtro IIR particular que cumpla con características deseables para la detección de eventos por cambios en la frecuencia; en particular, es de interés un filtro pasabandas, cuyo ancho de banda sea ajustable con el fin de poder detectar las componentes en frecuencia de alta resolución en caso de ser necesario. Para este efecto se propone un filtro IIR de tipo resonante o tipo pico, el cual se representa mediante la siguiente función de transferencia:

$$H(z) = \frac{|(1-r)(1-re^{-2j\omega_0})|}{1-2r \cos(\omega_0)z^{-1}+r^2z^{-2}} \quad (3.9)$$

Donde; ω_o es la frecuencia angular de resonancia del filtro y r corresponde al radio de convergencia de la función de transferencia del filtro. El parámetro r está directamente relacionado con el factor de calidad del filtro, dado que a medida que éste (r) aumenta, también aumenta la selectividad de éste. Este tipo de filtros es estable siempre y cuando $r < 1$ [107].

En la siguiente figura se observa el comportamiento de un filtro resonante para diferentes valores de r . En todos los casos, el filtro está sintonizado para un valor de frecuencia de resonancia de 0.4 (en la escala de frecuencia normalizada).

Figura 3-8: Respuesta en magnitud para filtros IIR resonantes con diferentes radios de convergencia.



Nombre de la fuente: Elaboración propia.

Ahora bien, regresando a la teoría original de la WT desde la perspectiva de filtros digitales, los coeficientes wavelet resultantes se computan, de forma sucinta, realizando la convolución entre la señal de interés y una wavelet hija, la cual corresponde a una versión desplazada en el tiempo y modificada en amplitud o escala de la función madre. En el contexto de filtros digitales, la operación anterior corresponde a remuestrear la función madre para obtener un banco de filtros compuesto por las funciones hijas y aplicar estos filtros (recordando que la convolución en el dominio del tiempo corresponde a una multiplicación en el dominio de la frecuencia) a la señal original. Cada vez que se filtra la

señal original por uno de los filtros del banco generado se obtiene un coeficiente $C(a, b)$ como se demuestra en la Ecuación (2.13).

Si bien es posible remuestrear un filtro IIR para generar el análogo a una función wavelet hija, como se detalló anteriormente, su implementación a un sistema digital no es trivial. En este orden de ideas, se propone una metodología alterna, la cual corresponde a remuestrear la señal original, creando efectivamente un banco de señales las cuales pueden ser filtradas con un único filtro resonante y obteniendo los coeficientes wavelet deseados (o su equivalente bajo la metodología propuesta).

3.2.1 Enventanado de la señal a analizar

En esta etapa se propone utilizar ventanas de longitud constante M y traslapo del 100% mientras que la función de enventanado será sujeto de análisis en el apartado de resultados de esta disertación. El valor de M está relacionado con la resolución en tiempo y frecuencia de la metodología, además de las características físicas de las señales de interés; en particular, se suelen utilizar ventanas cuya longitud corresponde a una fracción de un ciclo a la frecuencia industrial del sistema de potencia (60Hz en Colombia), destacándose ventanas de un cuarto de ciclo, medio ciclo y un ciclo completo. En teoría, el uso de ventanas más pequeñas puede garantizar mayor precisión temporal en la detección de eventos, aunque, como se discutirá más adelante puede afectar negativamente la resolución en frecuencia de los resultados.

3.2.2 Definición del filtro o función Madre

La extensión a la función Madre propuesta bajo esta metodología corresponde a un filtro IIR resonante, cuya frecuencia de resonancia permanece constante. En este caso, al estar operando en un sistema digital muestreado a una frecuencia F_S , sólo se puede extraer información veraz en el intervalo $[0, F_S/2]$. Este límite está definido por la frecuencia de Nyquist, la cual recibe su nombre en honor al ingeniero Harry Nyquist, el cual trabajó sobre las restricciones en el ancho de banda en sistemas digitales en el siglo 20.

Este límite puede generar inconvenientes cuando se trabaja sobre sistemas que están limitados de forma física, como por ejemplo los equipos en las subestaciones digitales que operan bajo el estándar IEC 61850 en el que se estandariza una frecuencia de muestreo de 4800Hz para SV, por lo que en teoría, sólo se obtendría información hasta 2400Hz, el cual es un valor relativamente pequeño en comparación de las frecuencias características observadas en los fenómenos transitorios electromagnéticos en sistemas de potencia, las cuales pueden alcanzar los 60kHz para fallas en sistemas de distribución [25], [110] o hasta 200kHz para eventos de conmutación en sistemas de transmisión [10].

En estos casos, es relevante recordar que la frecuencia de Nyquist, la cual define el límite teórico hasta el cual se pueden medir frecuencias en los sistemas digitales, también recibe el nombre de frecuencia de plegado, dado que cualquier frecuencia mayor a la frecuencia de Nyquist presente en una señal continua se verá solapada o plegada en las frecuencias menores de ésta cuando la señal se muestrea de forma discreta (este fenómeno es formalmente conocido como Aliasing). En nuestro caso, las altas frecuencias están presentes de forma plegada en los componentes de baja frecuencia, por lo que la información contenida en éstas no se pierde por completo.

3.2.3 Definición de la resolución en tiempo y frecuencia del método

Para definir las resoluciones en tiempo y frecuencia de un sistema es útil recurrir al principio de incertidumbre de Heisenberg para tiempo y frecuencia (Ecuación (2.3)). En términos de frecuencia temporal, este principio se describe por la Ecuación (3.10).

$$2 \cdot \delta f \cdot \delta t \geq \frac{1}{4} \quad (3.10)$$

Donde:

- δf corresponde a la resolución en frecuencia del sistema medida en Hz.
- δt corresponde a la resolución en tiempo del sistema medida en segundos.

Claramente, obtener una representación tiempo y frecuencia en la que se alcance el límite propuesto por la desigualdad presente en el principio de incertidumbre es complejo, por lo que estos valores de resolución en tiempo y frecuencia deberían ser considerados como

los umbrales mínimos que pueden obtenerse. Así, para la metodología propuesta se proponen Δf y Δt tales que:

$$\Delta f > \delta f \quad (3.11)$$

$$\Delta t > \delta t \quad (3.12)$$

En este caso, Δf corresponde al paso discreto entre las escalas de frecuencia que se analizarán con la metodología propuesta, mientras que Δt corresponde al paso discreto de tiempo que depende de la frecuencia de muestreo de éste. Considerando que se propone operar sobre ventanas de longitud constante M y no sobre la señal completa, se puede definir Δt como:

$$\Delta t = \frac{M}{F_S} \quad (3.13)$$

Y utilizando 3.10, se puede obtener un límite para δf :

$$\delta f \geq \frac{F_S}{8M} \quad (3.14)$$

Considerando la desigualdad en 3.14:

$$\Delta f = \frac{F_S}{8M} \quad (3.15)$$

3.2.4 Remuestreo de las ventanas

Teniendo definido el paso discreto en frecuencias Δf , es posible generar un conjunto de frecuencias de interés cómo:

$$F = \{f_i | f_i = F_C \pm \Delta f \wedge f_i \geq f_{min}\} \quad (3.16)$$

Donde:

- f_i corresponde una de las frecuencias válidas para el MRA.
- F_C corresponde a la frecuencia de resonancia del filtro IIR.
- f_{min} corresponde a la frecuencia mínima de interés o la mínima frecuencia esperada en el sistema.

Por motivos que se harán evidentes en la sección de resultados, la frecuencia de resonancia del filtro IIR es escogida como $F_C = F_S/2$, aunque una explicación sucinta de esta decisión radica en el hecho de que el remuestreo propuesto de las señales enventanadas es más eficiente en términos computacionales cuando F_C es un múltiplo o submúltiplo de F_S . Por otro lado, f_{min} puede ser tomada como la frecuencia mínima esperada del sistema, correspondiente a la frecuencia industrial de 60Hz, aunque se puede intuir que una parte de la información contenida en la banda de frecuencia cercana a f_{min} se perderá debido a que es el límite inferior de cómputo de la metodología.

La relevancia del conjunto de frecuencias generado estriba en el concepto fundamental de la metodología propuesta, el cual es el remuestreo de las señales enventanadas. Cómo se explicó de forma breve, se busca obtener un banco de filtros compuesto por el análogo de las funciones hija remuestreando cada una de las ventanas de la señal original por una serie de factores. Estos factores se pueden obtener de la razón f_i/F_C , la cual genera un número racional de la forma p/q , con los cuales se computan las señales remuestreadas.

Si bien existen múltiples métodos para computar el remuestreo de señales digitales [111], el método utilizado corresponde al utilizado por el software de MATLAB en el que se usa un proceso de tres etapas, las cuales consisten en:

1. Insertar ceros a la secuencia digital $x[n]$ para sobremuestrear por el factor p .
2. Aplicar un filtro antialiasing tipo FIR a la señal sobremuestreada.
3. Descartar muestras hasta submuestrear la señal obtenida en la segunda etapa por un factor q .

La relevancia del filtro antialiasing utilizado en la segunda etapa radica en el filtrado de componentes indeseadas de frecuencia que se introducen en las fronteras del filtro convencional utilizado para implementar el remuestreo.

3.2.5 Computación y normalización de coeficientes

A continuación, el banco de filtros obtenido al remuestrear las señales enventanadas se filtra utilizando el filtro IIR resonante propuesto al inicio de la metodología. Este filtrado se computa utilizando la función `filter` de MATLAB, la cual utiliza la ecuación en diferencias del filtro obtenida a partir de su ecuación de transferencias para computar la respuesta en cada una de las muestras de la secuencia [112]. El resultado de la aplicación del filtro es una nueva señal discreta $Y[n]$.

$$Y(p, q)[n] = \text{resample}(x[n], p, q) \quad (3.17)$$

Con el objetivo de obtener resultados comparables a los coeficientes wavelets tradicionales, es necesario introducir dos factores de normalización sobre esta secuencia; el primer factor de normalización está relacionado a la modificación de amplitudes introducida al remuestrear las ventanas de la señal original por los factores p/q . En este caso, el factor de corrección introducido corresponde a q/p , esto es, el factor recíproco al utilizado para computar el remuestreo.

$$Y^{corr}(p, q)[n] = \frac{q}{p} Y_n(p, q)[n] \quad (3.18)$$

Por otro lado, en la Ecuación (2.13) referente a la CWT discretizada, se introduce una normalización mediante el factor de escala (a) de la forma $\frac{1}{\sqrt{|a|}}$. Este factor proviene de las características físicas de la CWT, la cual en su forma integral conserva la energía; sin embargo, al implementar la CWT de forma numérica la energía no es conservada, por lo que se introduce una normalización de tipo $L1$ [53].

La normalización de tipo $L1$ asigna diferentes pesos a cada escala de frecuencia, de forma que los picos de alta frecuencia se atenúan más que los picos en baja frecuencia. En cuanto a su implementación en MATLAB, la normalización $L1$ puede computarse dividiendo la señal por la suma de los valores absolutos de los valores de la secuencia filtrada. Finalmente, para obtener los coeficientes análogos a los coeficientes wavelets

basta con computar la suma de los valores filtrados y normalizados como se observa en la Ecuación (3.19).

$$C(p, q) = \frac{\sum_{n=1}^N Y^{corr}(p, q)[n]}{\sum_{n=1}^N |Y^{corr}(p, q)[n]|} \quad (3.19)$$

3.3 Extracción de características de los coeficientes

Wavelet

Considerando la metodología utilizada para computar los coeficientes wavelet se consideraron múltiples acercamientos para la extracción de características relevantes de éstos. Finalmente, se concluyó que para el conjunto de datos generados un enfoque en parámetros estadísticos permite obtener información suficiente sin convertirse en el foco de la disertación (recordando que el objetivo principal de ésta consiste en comprobar la validez de la metodología propuesta anteriormente).

En este sentido, se propuso utilizar los parámetros de la media estadística, el valor cuadrático medio, la densidad de potencia espectral y la curtosis de los conjuntos de información obtenidos al computar la CWT mediante la metodología propuesta sobre las ventanas de las señales analizadas. A continuación, se da una breve definición de cada uno de estos parámetros y se presenta su fórmula, en función de una secuencia de puntos $x[n] = x_1, x_2, \dots, x_n$.

- **Media aritmética:** A veces llamada promedio, es una medida de la tendencia central de una variable la cual es susceptible a la variabilidad de las muestras. Es útil en el contexto de la detección y clasificación de eventos dado que contiene, en cierto grado, la información relacionada a cambios repentinos en el contenido en frecuencia de las señales. Su fórmula se observa en la Ecuación (3.20).

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.20)$$

- **Valor cuadrático medio (RMS):** Se define como la raíz cuadrada de la media cuadrada, la cual es la media aritmética de los cuadrados de los valores de una variable aleatoria (Ecuación (3.21)). Es particularmente útil en aplicaciones de

señales eléctricas debido a que para una corriente alterna, el valor RMS equivale a la corriente directa que produciría una disipación de energía idéntica para una carga resistiva.

$$X_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (3.21)$$

- Densidad de Potencia Espectral (PSD): Como fue mencionado en el apartado del Estado del Arte, es una medida de la energía total contenida en la señal de interés para ciertas escalas de frecuencia. En este caso se propone utilizar únicamente la PSD de los coeficientes de detalle (Ecuación (2.19)).
- Curtosis: La Curtosis K (por su traducción al inglés Kurtosis) es una medida de la forma de una distribución de probabilidad y en su forma clásica está directamente relacionada con el momento central de cuarto orden de una variable aleatoria. Dado que está relacionada con la “forma” de la distribución de probabilidad, es una variable cuya definición tiende a ser vaga y está sujeta a ser formalizada de múltiples formas [113].

Para esta disertación se utilizó la función integrada en el toolbox estadístico de MATLAB [114], en el cual la curtosis de una secuencia de entradas $x[n]$ se computa como:

$$k = \frac{E(x[n]-\mu)^4}{\sigma^4} \quad (3.22)$$

Donde:

- μ corresponde a la media aritmética de $x[n]$.
- σ corresponde a la desviación estándar de $x[n]$.
- $E(t)$ corresponde al valor esperado para la variable aleatoria t.

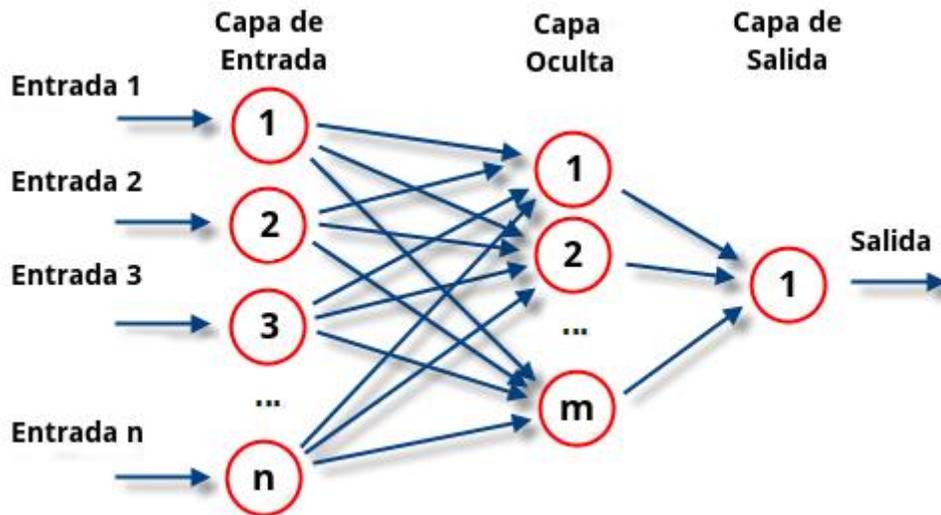
3.4 Detección y clasificación de eventos

Según el análisis de la literatura reportado en el capítulo 2, sección 4, existen dos opciones generales para desarrollar un sistema de detección y clasificación de eventos: el uso de umbrales y DT o el uso de herramientas basadas en ML e IA. Considerando que el objetivo principal de la presente disertación se centra en la propuesta metodológica y la complejidad de las características extraídas de los coeficientes wavelet, se propone utilizar un esquema de detección y clasificación basado en ANN debido a la familiaridad con estos algoritmos y la revisión y aplicación extensa que se ha propuesto en las referencias de interés.

Como se mencionó brevemente en el Estado del Arte, las ANN operan de forma similar a sus análogas biológicas, reemplazando las neuronas por nodos que cumplen las funciones de recibir entradas de información, procesarla y transmitirla hacia otras neuronas; el objetivo de este procedimiento es aprender dado un conjunto de datos de forma que el algoritmo tenga la capacidad de clasificar correctamente salidas dadas unas características de entrada. En particular, el beneficio de las ANN está en que, dada suficiente información y con el ajuste correcto, pueden realizar predicciones lógicas con información nueva en función del aprendizaje que han experimentado.

Así, dados los conjuntos de entrada y salida obtenidos (los cuales consisten en una serie de características de entrada y categorías de salida) se propone el uso de una ANN de tipo Feedforward. El término Feedforward hace referencia a la forma en la que la ANN transmite la información y computa las salidas: un grupo de neuronas (denominado capa) interactúa únicamente otro grupo en un nivel posterior, por lo que la información se propaga de forma unidireccional, desde la entrada a la salida de la ANN. Esta representación se puede observar en la figura 3-9.

Figura 3-9: Representación de ANN tipo Feedforward.



Nombre de la fuente: Neural networks: A comprehensive foundation [116].

Con respecto al concepto de las capas en las ANN, pueden existir tantas como sean necesarias, aunque por lo general se pueden agrupar en algunas de las siguientes categorías:

- Capas de entrada
- Capas de salida
- Capas ocultas

Las capas de entrada y salida, como su nombre lo indican, son aquellas donde la información entra y sale, respectivamente, mientras que las capas ocultas se encuentran en medio de éstas (entrada y salida); las características de estas capas se verán con más detalle más adelante en esta sección. Las capas ocultas reciben este nombre debido a que normalmente su estructura completa no es directamente visible, aunque es posible abstraer el concepto de éstas para manipularlas.

Con respecto al aprendizaje, este puede definirse como un problema de regresión o ajuste de curva [116], [117] de una función continua en varias variables. De este modo, el modo en el que una ANN realiza predicciones está basado en la interpolación de la información

de entrada para obtener la salida más probable. En este orden de ideas, es fundamental introducir algunos conceptos utilizados en el proceso de ajuste y aprendizaje de las ANN, los cuales son los pesos, las funciones de activación y el bias.

Los pesos (W) son parámetros exclusivos de cada neurona que modifican la información que ingresa a éstas; el valor de los pesos se determina a través de procesos iterativos como un método de relajación (como el descenso de gradiente) [121] aunque cuando hay una gran cantidad de datos o un elemento que introduce una alta no linealidad, se recomiendan otros métodos [118].

La Ecuación (3.23) ejemplifica el método de descenso de gradiente para determinar los pesos de una neurona dada, ya que establece que el peso de la i -ésima neurona debido al efecto de la j -ésima neurona (en el caso de un feed-forward esta neurona pertenecería a una capa directamente anterior a la capa que contiene la i -ésima neurona) en la iteración $m+1$ está dada por el peso de la iteración anterior más un gradiente que se modifica hasta satisfacer un cierto error máximo.

$$W_{ij}^{m+1} = W_{ij}^m + \Delta W_{ij}^m \quad (3.23)$$

El valor por el que se actualiza el peso ΔW_{ij}^m es comúnmente conocido como la tasa de aprendizaje (o learning rate) y corresponde a un parámetro ajustable, usualmente en el rango (0,1). Un learning rate muy pequeño requiere que el algoritmo de entrenamiento se compute más veces dado que el aprendizaje es lento y los pasos que se introducen a las actualizaciones de los pesos son pequeñas; por otro lado, un learning rate grande agiliza el tiempo de entrenamiento al incrementar los pasos en la actualización de los pesos, aunque es aumenta la probabilidad de que el modelo llegue rápidamente a una solución (recordando que el entrenamiento es, fundamentalmente un problema de ajuste de curvas multivariadas) que no es óptima [119].

La forma en la que las neuronas procesan la información de entrada para obtener una salida es mediante el uso de las funciones base, dentro de las que se destacan dos, las cuales se asemejan a superficies en las que la información de entrada pasa por una transformación que genera un valor utilizado para crear dicha superficie. Cuando los pesos son muy lineales, una función base hiperplana da resulta en una buena aproximación, pero

cuando existe alta no-linealidad, es recomendable utilizar una superficie hiperesférica [121].

- **Función base hiperplana o lineal:** En este caso la salida es una combinación lineal de las entradas.

$$y_i(\omega, x) = \sum_{j=1}^n \omega_{ij} x_j = \sum_{j=1}^n u_j \quad (3.24)$$

- **Función base hiperesférica o radial:** En este caso los pesos se utilizan para medir una distancia al centro de un conjunto de esferas. Esta función base se usa principalmente en aplicaciones altamente no lineales o casos con una gran cantidad de datos.

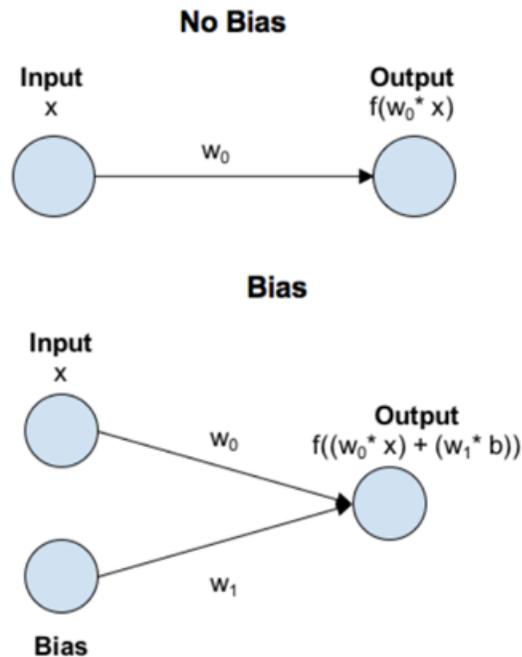
$$y_i(\omega, x) = \sqrt{\sum_{j=1}^n (x_j - \omega_{ij})^2} \quad (3.25)$$

Donde:

- W_{ij} es el peso de la i-ésima neurona debido al efecto de la j-ésima neurona.
- x_j es la entrada de la j-ésima neurona.
- y_j es la salida de la i-ésima neurona.

Por otro lado, el concepto de Bias está directamente relacionado con su significado gramático (de parcialidad) dado que modifica la salida de las neuronas dada por la función base al introducir una constante; esto puede relacionarse al equivalente de una constante en una función lineal, la cual desplaza la función en todos los puntos por el valor de ésta. En el escenario con Bias, la salida de la neurona se modifica por el valor del Bias b multiplicado por el peso asignado al Bias w_b , desplazando el resultado por $(b \cdot w_b)$. La diferencia generada por la inclusión del Bias se presenta en la figura 3-10.

Figura 3-10: Representación gráfica del concepto de Bias en las neuronas de una ANN.

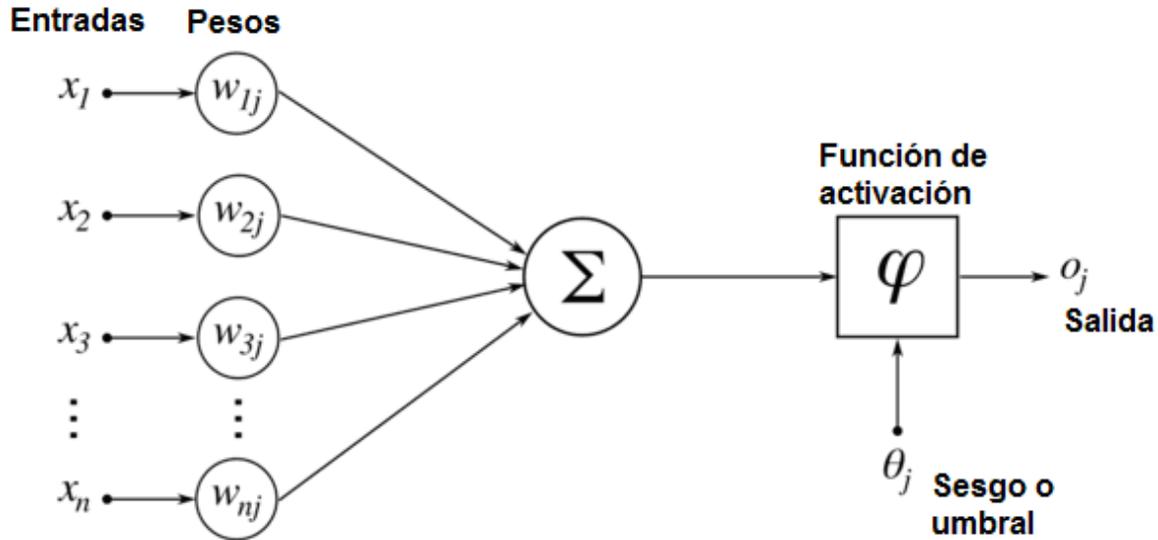


Nombre de la fuente: <https://www.pico.net/kb/the-role-of-bias-in-neural-networks/> [120].

El último concepto relevante para comprender el método de aprendizaje de las ANN es la función de activación ϕ ; esta función se aplica a la salida de una neurona para determinar si dicha salida tiene o no relevancia en la decisión o salida final.

La operación fundamental de una neurona dentro de una ANN se presenta en la figura 3-11; como se observa, una neurona recoge los valores ponderados propagados por las neuronas en una capa anterior mediante la función base y suma el bias; finalmente, se utiliza la función de activación para determinar la relevancia de la salida de la neurona [121].

Figura 3-11: Vista general del modo de operación de una neurona en una ANN.



Nombre de la fuente: *Artificial neural networks: an introduction to ANN theory and practice* [114].

A continuación, se presentan algunas de las funciones de activación más utilizadas y sus aplicaciones convencionales.

Las funciones de activación sigmoideal y gaussiana son comúnmente usadas en las capas ocultas de las ANN. Las ecuaciones de estas funciones se presentan en las Ecuaciones (3.26) y (3.27), respectivamente.

$$\varphi_{\text{sigmoideal}}(y_i) = \frac{1}{1+e^{-y_i}} \quad (3.26)$$

$$\varphi_{\text{gaussiana}}(y_i) = ce^{-y_i^2} \quad (3.27)$$

En el caso de las funciones Unidad Lineal Rectificada (ReLU) y SoftMax, su aplicación común es en la capa de salida. Las ecuaciones de estas funciones de activación se presentan en las Ecuaciones (3.28) y (3.29), respectivamente.

$$\varphi_{\text{ReLU}}(y_i) = \begin{cases} 0, & t < 0 \\ y_i, & t \geq 0 \end{cases} \quad (3.28)$$

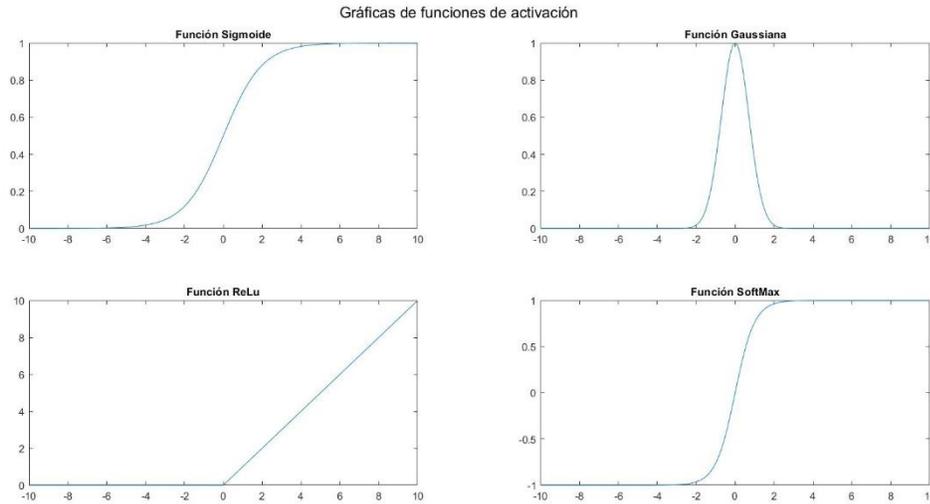
$$\varphi_{SoftMax}(y_i) = \frac{e^{y_i}}{\sum_{j=1}^m e^{y_j}} \tag{3.29}$$

Donde:

- y_i corresponde a la salida de la neurona i -ésima.
- m corresponde al número de neuronas en la capa.

La representación gráfica de estas funciones se observa en la figura 3-12. Es relevante destacar que la función Gaussiana es mucho más selectiva en cuanto al peso que se asigna a las salidas en comparación con el resto de las funciones de activación.

Figura 3-12: Funciones de activación Sigmoideal, Gaussiana, ReLu y SoftMax.



Nombre de la fuente: Elaboración propia.

Con estos conceptos es posible introducir una explicación simplificada de cómo una ANN reconoce con éxito un patrón (en nuestro caso, un evento en el SEP) es la siguiente: supongamos que los pesos individuales de cada neurona (y su valor de bias en caso de tenerlo) ya estén determinados por un método cualquiera y que la ANN es de tipo Feedforward con una sola capa oculta; en este caso, la información de cada evento es ingresada a la ANN en la capa de entrada mediante un vector que contiene las

características extraídas de las señales en conjunto con la categoría (tipo de evento) al que describen éstas.

Cuando se alimenta una nueva entrada a la red, esta información se modifica por el peso de la entrada y luego se propaga a la capa oculta, donde las neuronas de ésta operan sobre el vector de entrada, modificándolo con los pesos individuales (y bias) de cada una y se aplica la función base para recopilar la información de salida de las neuronas; a continuación, se aplica la función de activación con la que se decide si la salida de una neurona es significativa o no en la clasificación del vector de características de entrada. Finalmente, la información de la capa oculta se propaga a la capa de salida donde se asigna la categoría que más se ajusta a la salida obtenida.

En cuanto al proceso de aprendizaje de la ANN y el ajuste de los pesos de las neuronas en su capa oculta, es útil considerar las siguientes etapas:

- **Entrenamiento:** En esta etapa se utilizan algoritmos como el descenso del gradiente para ajustar los pesos de las neuronas utilizando un conjunto de entradas y salidas, y parámetros definidos como el learning rate.
- **Validación o generalización:** En esta etapa, la ANN usa la información y los pesos del proceso de aprendizaje para generalizar su conocimiento, con lo que es posible obtener salidas (siempre y cuando éstas correspondan a alguna de las salidas usadas para entrenar la ANN) para un conjunto de características de entrada nuevo.

Si entendemos el proceso de entrenamiento y aprendizaje de la ANN como una reconstrucción de una hipersuperficie como se ha dicho anteriormente, la generalización podría entenderse como un método estadístico que se centra en completar los huecos que pueda tener la hipersuperficie asumiendo una superficie lisa que minimiza el error y es estadísticamente aceptable [118].

El proceso de generalización es particularmente importante en el contexto del ajuste de la red, dado que permite diagnosticar si existe sobreajuste o subajuste. El concepto de sobreajuste se utiliza cuando una red neuronal se entrena hasta el punto en que aprende sobre los detalles y el ruido de las entradas, lo que afecta su

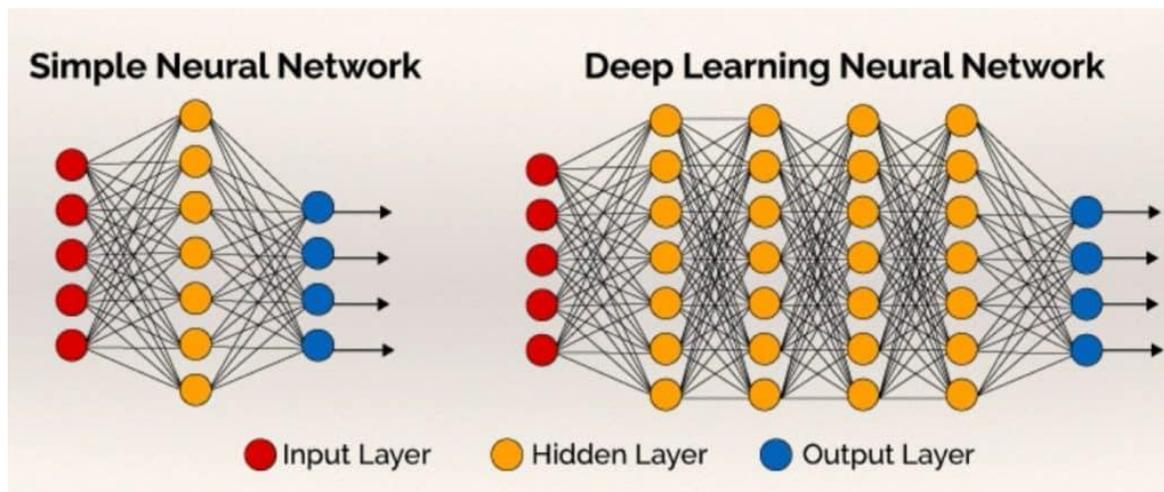
capacidad de predecir salidas para nueva información. Por otro lado, el subajuste se presenta cuando el algoritmo no modela de forma adecuada la información de entrenamiento (y por ende tampoco puede generalizar de forma correcta).

- **Prueba:** Se utiliza un conjunto de entradas y salidas que se extrae del conjunto de datos original, los cuales se usan cuando la red ha sido entrenada para comparar las salidas reales con respecto a las predichas por el algoritmo y se determina la eficiencia de clasificación y aprendizaje de la red neuronal.

La proporción de información que se le asigna a los conjuntos de entrenamiento, validación y prueba varía con respecto al autor y aplicación, aunque una partición de 70% de datos para entrenamiento, 15% de datos para validación y 15% de datos para prueba es un acercamiento convencional.

Finalmente, es relevante introducir el concepto de la Redes Neuronales Artificiales Profundas, en las que se utilizan múltiples capas ocultas entre las capas de entrada y salida. Por lo general, una única capa profunda suele arrojar resultados aceptables, aunque para problemas complejos o para diseñadores con experiencia, el uso de múltiples capas profundas puede mejorar significativamente la eficiencia de éstas.

Figura 3-13: Esquema de ANN superficiales (izquierda) y profundas (derecha).



Nombre de la fuente: Elements of Artificial Neural Networks [112].

Por otro lado, se pueden implementar capas ocultas que no contienen un arreglo de neuronas sino funciones o estructuras con objetivos particulares que buscan mejorar el rendimiento del algoritmo. En este sentido, es relevante introducir los conceptos de capa de normalización y regularización.

Las capas de regularización buscan prevenir los efectos del sobreajuste y subajuste actuando sobre los pesos individuales de las neuronas en la etapa de entrenamiento de la red neuronal. Por lo general, cuando existe sobreajuste o subajuste, los pesos individuales de algunas neuronas crecen (o decrecen) desproporcionadamente, impactando negativamente la capacidad de la ANN de aprender y generalizar. La regularización actúa penalizando los pesos individuales de las neuronas de forma que no crezcan o disminuyan desproporcionadamente en la etapa de entrenamiento en relación con el resto de las neuronas [117].

Los algoritmos de regularización más comunes en aplicaciones de ML son la regularización L1 y L2. Éstos añaden un término adicional en el proceso de actualización de los pesos de las neuronas (Ecuación (3.23)), lo cual regula la tasa de actualización de las neuronas, mejorando la capacidad de generalización de la ANN [118]. La diferencia entre los algoritmos L1 y L2 radica en el término utilizado.

$$L2 = \mu \sum \|w\|^2 \quad (3.30)$$

$$L1 = \mu \sum \|w\| \quad (3.31)$$

Donde:

- μ corresponde al parámetro de regularización.
- w corresponde al peso de una neurona cualquiera.

Como puede observarse, la regularización L1 penaliza el valor absoluto de los pesos lo que limita directamente el tamaño de los pesos y su actualización; por otro lado, la regularización L2 penaliza el cuadrado de la magnitud de los pesos, lo que fuerza a los pesos a decaer rápidamente a un valor cercano a cero.

Finalmente, las capas de normalización se utilizan para obtener un conjunto de información de entrenamiento con media cercana a cero (normalización estadística) lo cual, por lo general, aumenta la velocidad de la etapa de aprendizaje; esto es relevante cuando se tienen múltiples características de entrada con rangos diferentes, dado que es posible que la magnitud de ciertas entradas afecte desproporcionadamente las salidas.

Otra aplicación relevante de la normalización está relacionada con las funciones de activación, las cuales son, en su mayoría, no lineales y sus salidas están en rangos predeterminados (-1, 0 y 1). En estos casos, se debería garantizar que los objetivos o salidas estén localizados en el rango de las funciones de activación, lo cual puede ser solucionado con la normalización de las capas ocultas o de salida de la ANN.

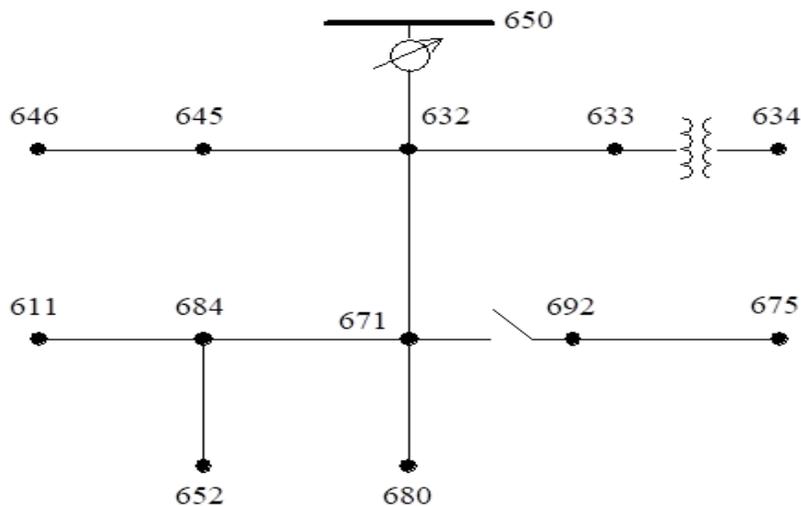
4. Resultados y discusión

4.1 Simulación de eventos a detectar y clasificar

Dado el acercamiento metodológico propuesto, se consideraron varios métodos y sistemas de distribución para obtener las señales de falla de interés. Finalmente, se optó por utilizar un sistema con múltiples referencias en la literatura revisada, que permitiese obtener un banco de eventos lo suficientemente grande como para validar la metodología propuesta cuya implementación no fuese muy compleja, dado que el objetivo de esta disertación se centra en la metodología y no en el sistema.

Para esto se utilizó el Sistema de distribución de la IEEE de 13 barras, el cual opera a una tensión de 4,16kV y se caracteriza por ser corto, tener líneas aéreas y cables subterráneos, dos capacitores en derivación, un transformador de subestación y otro de reducción de tensión, además de tener una carga relativamente alta en comparación con la potencia nominal del sistema con un total de 9 cargas desbalanceadas [124].

Figura 4-1: Diagrama unifilar de sistema de prueba IEEE 13 barras.



Nombre de la fuente: IEEE 13 Node Test Feeder [125].

La tabla con los eventos simulados se muestra a continuación:

Tabla 4-1: Eventos simulados y características de falla (Impedancia, ángulo de inyección y tipo de falla). Elaboración propia.

Nodo	Resistencias de falla [Ω]	Ángulos de inyección [$^{\circ}$]	Tipos de Falla
632	1, 20, 100	0°, 45°	1 Φ - A, 1 Φ - B, 1 Φ - C
			2 Φ - AB, 2 Φ - AC, 2 Φ - BC
			3 Φ
675	1, 20, 100	0°, 45°	1 Φ - A, 1 Φ - B, 1 Φ - C
			2 Φ - AB, 2 Φ - AC, 2 Φ - BC
			3 Φ
645	1, 20, 100	0°, 45°	1 Φ - B, 1 Φ - C
			2 Φ - BC
632	1, 20, 100	0°, 45°	1 Φ - A

Así, se obtienen 108 eventos simulados, para los que se registraron los valores de tensión en cada fase durante dos ciclos a 60Hz (equivalentes a 33.3 ms). En cuanto a la generación de los eventos de falla, se adaptó un script de Python provisto por David Alejandro Sánchez Muñoz, M.Sc. [126], el cual utiliza la herramienta de cómputo de ATP con archivos de texto plano generados a partir de la herramienta gráfica ATP Draw, para crear y ejecutar archivos temporales con las modificaciones necesarias para la simulación de los eventos propuestos, como la impedancia y tipo de falla.

Por otro lado, es necesario proveer múltiples situaciones de operación no nominal que no representan una falla eléctrica, con el fin de mejorar la capacidad de detección y clasificación del algoritmo. Para ello, se simularon los siguientes escenarios de operación:

- Operación nominal del sistema.
- Elevación de tensión del 10% en el nodo de generación del sistema.
- Reducción de tensión del 10% en el nodo de generación del sistema.
- Conexión de carga en los nodos de interés. Para ello se aumentó la magnitud de la impedancia de carga equivalente por 20%.
- Desconexión de carga en los nodos de interés. Para ello se redujo la magnitud de la impedancia de carga equivalente por 20%.

- Conexión de bancos de condensadores en las líneas entre los nodos 692 y 675 y los nodos 632 y 671. La reactancia de estos bancos de condensadores se calculó de modo que el factor de potencia de la carga al final de la línea se llevara a 0.95. Así, se obtuvieron 22 casos de operación normal del sistema los cuales se simularon de forma manual. Los cálculos utilizados para determinar las nuevas impedancias de carga y la reactancia de los bancos de condensadores se detallan en el anexo A.

4.2 Validación de la metodología

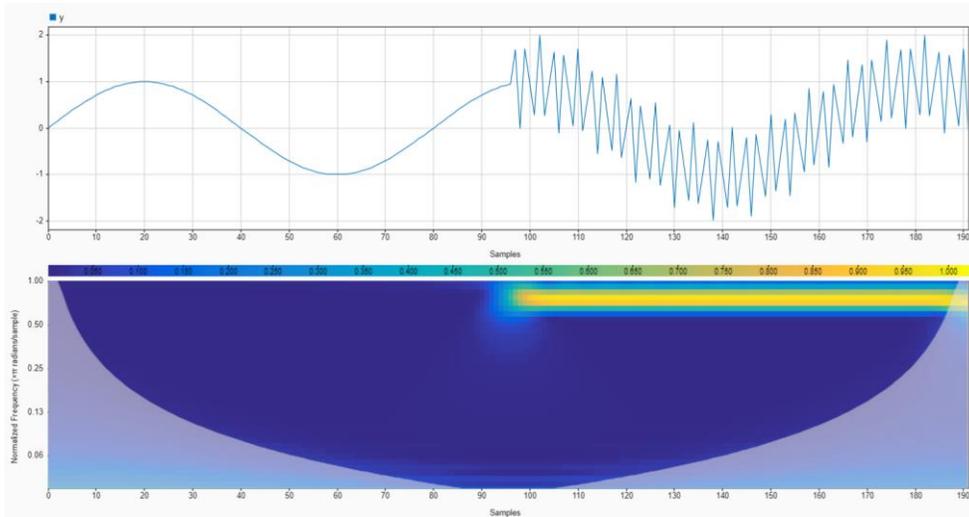
Las señales generadas en ATP/EMTP se simularon a una tasa de muestreo de 480 kHz por razones de estabilidad y convergencia en los modelos de las líneas de distribución y cables subterráneos utilizados. Sin embargo, al exportar las señales se realiza un submuestreo por un factor de 10, obteniendo señales con frecuencia de muestreo de 48kHz. En este punto es relevante analizar el impacto que tiene el valor de la frecuencia de muestreo en los resultados de la computación de los coeficientes mediante el algoritmo propuesto; para ello se utilizaron dos señales de prueba: una señal sinusoidal con duración de 40ms a 60Hz con un armónico de 1800Hz introducido a los 20ms y una señal simulada en ATP/EMTP correspondiente a una falla monofásica a tierra con impedancia de falla de 10Ω en el sistema IEEE 39 barras propuesto, en el nodo 675.

Con el fin de validar la metodología propuesta se propuso sintetizar los resultados obtenidos en un escalograma y comparar éste con el escalograma arrojado por el analizador de señales de MATLAB, el cual computa el escalograma utilizando la DWT y una función madre tipo db4. Todos los escalogramas que se presenten de aquí en adelante tienen un mapa de colores idéntico (en particular es el mapa *parula*, de MATLAB), en el que las zonas con baja energía tienen un color azul, mientras que las zonas con mayor concentración de energía tienen asociado un color amarillo. El efecto de la frecuencia de muestreo sobre los escalogramas generados por el analizador de señales de MATLAB puede observarse en las figuras 4-3, 4-4, 4-5 y 4-6.

En las figuras 4-3 y 4-4 se observan los escalogramas para la función sinusoidal con componente armónico de alta frecuencia muestreada a 4800Hz y 48000Hz respectivamente. En este caso, ambos escalogramas representan adecuadamente la componente de alta frecuencia (1800Hz) aunque se observa una mayor dispersión de la

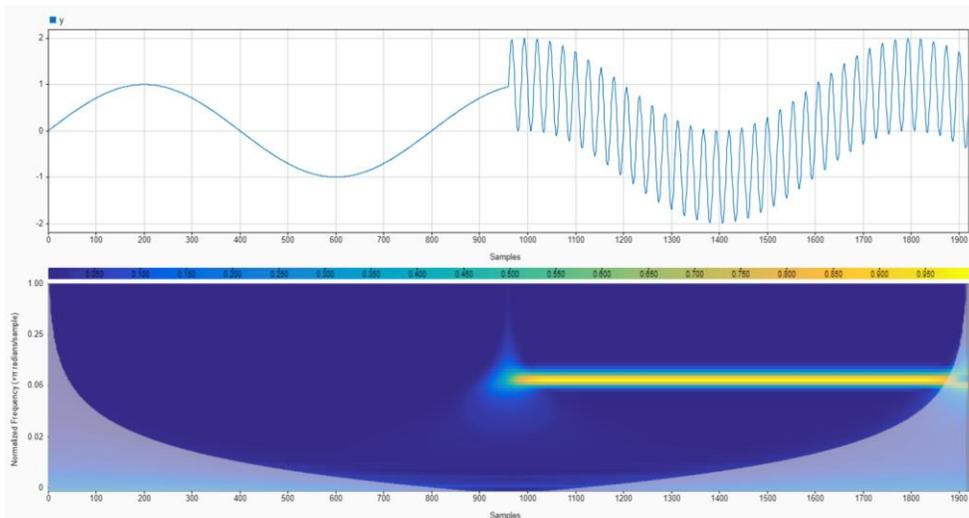
energía en el caso de 4800Hz. Es relevante destacar que en ningún caso se detecta adecuadamente el componente de 60Hz dado que éste se encuentra, en su mayoría, por fuera del cono de influencia del escalograma.

Figura 4-3: Señal sinusoidal con armónico (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 4800Hz.



Nombre de la fuente: Elaboración propia.

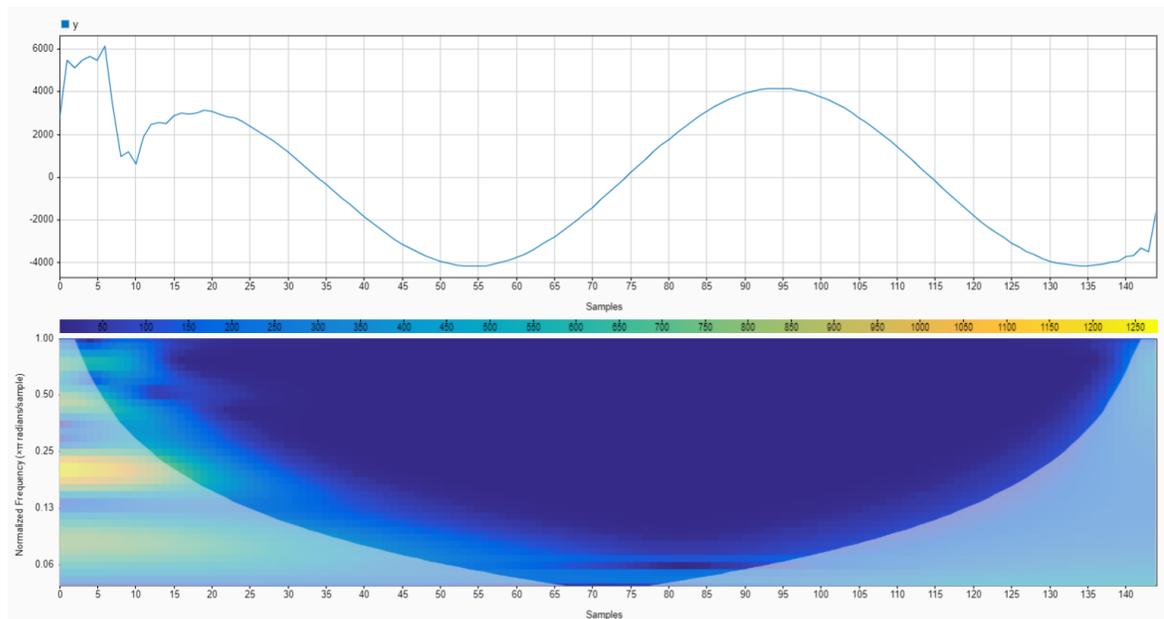
Figura 4-4: Señal sinusoidal con armónico (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 48000Hz.



Nombre de la fuente: Elaboración propia.

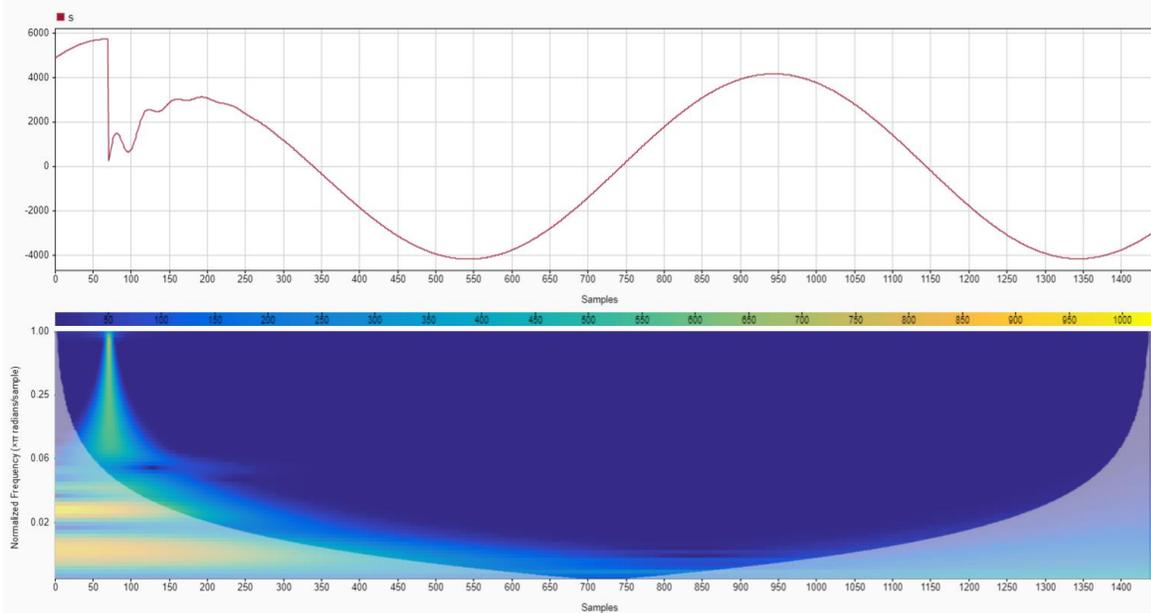
Por otro lado, cuando comparamos los escalogramas asociados a la señal de falla simulada (figuras 4-5 y 4-6) podemos observar una diferencia significativa en la dispersión del contenido de alta frecuencia que se presenta al inicio de la señal (momento en el que ocurre el evento transitorio). En este caso la localización del componente de alta frecuencia y su contenido de energía es claramente superior cuando se utiliza la frecuencia de muestreo de 48000Hz.

Figura 4-5: Señal de falla simulada (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 4800Hz.



Nombre de la fuente: Elaboración propia.

Figura 4-6: Señal de falla simulada (arriba) y su escalograma asociado (abajo), generado por MATLAB utilizando DWT y función madre db4 para una frecuencia de muestreo de 48000Hz.

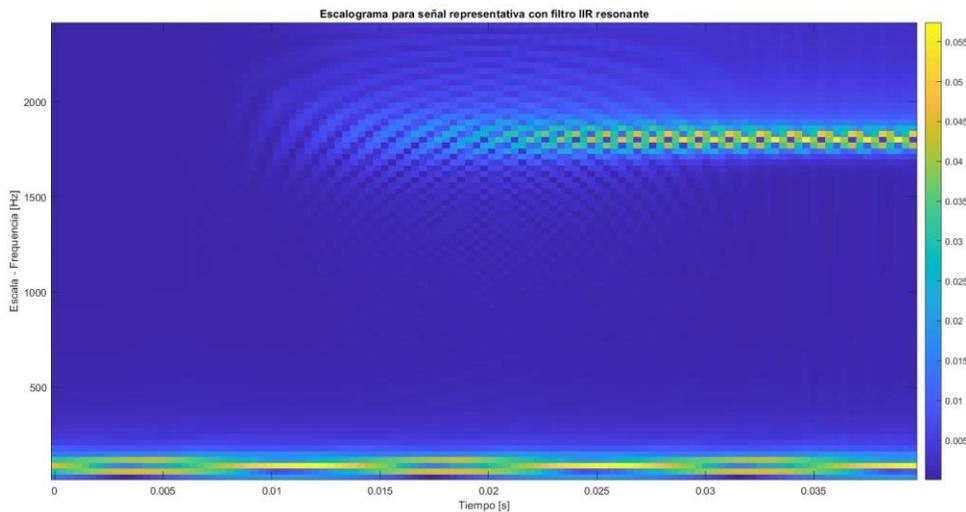


Nombre de la fuente: Elaboración propia.

Los resultados obtenidos al implementar la metodología propuesta con CWT y funciones madre IIR se pueden observar en las figuras 4-7, 4-8, 4-9 y 4-10. Las figuras 4-7 y 4-8 muestran el escalograma obtenido para la señal sinusoidal utilizando frecuencias de muestreo de 4800 Hz y 48000 Hz, respectivamente; en este caso la dispersión de energía que introduce la menor frecuencia de muestreo resulta evidente, en particular alrededor de los 0.2 segundos, donde se inyecta la componente de alta frecuencia. Por otro lado, en la figura 4-7 se observan puntos fríos (color azul) en las bandas de frecuencia que contienen las componentes de 60 Hz y 1800 Hz. Este comportamiento se analizará en conjunto con la influencia de la longitud de las ventanas, más adelante. Finalmente, cabe resaltar que en ambos casos el componente de 60 Hz se detecta de forma satisfactoria, aunque la representación dada para el caso con frecuencia de muestreo a 48000 Hz parece ser mucho más estable en tanto el color asignado no varía significativamente, en contraste a la representación para 4800 Hz en donde se observan de nuevo los puntos fríos mencionados previamente.

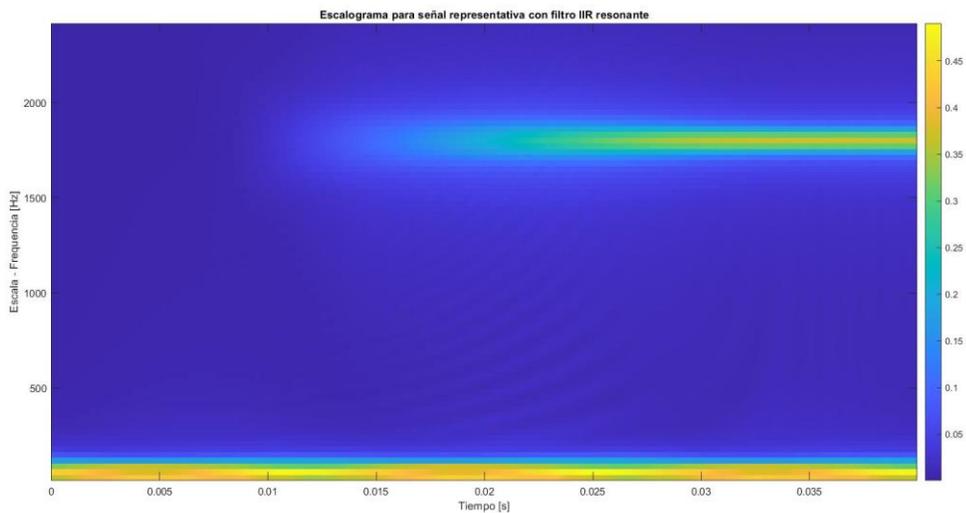
Por otro lado, las figuras 4-9 y 4-10 muestran el escalograma obtenido para la señal de falla simulada utilizando frecuencias de muestreo de 4800 Hz y 48000 Hz, respectivamente. En este caso, la diferencia en el contenido de energía en las bandas de alta frecuencia relacionadas al evento transitorio no es tan significativa dado que contiene poca energía en relación con el componente a frecuencia industrial.

Figura 4-7: Escalograma obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 4800 Hz.



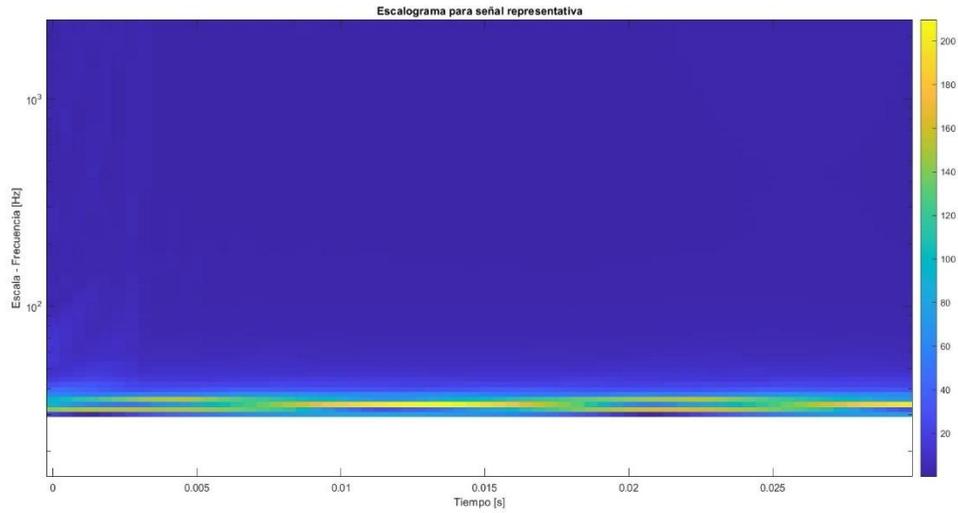
Nombre de la fuente: Elaboración propia.

Figura 4-8: Escalograma obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 48000 Hz.



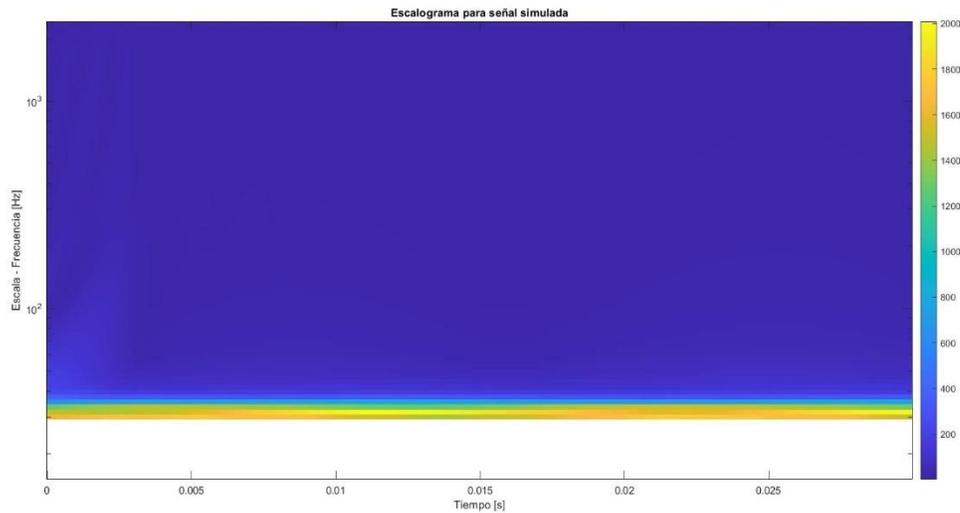
Nombre de la fuente: Elaboración propia.

Figura 4-9: Escalograma obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 4800 Hz.



Nombre de la fuente: Elaboración propia.

Figura 4-10: Escalograma obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 48000 Hz.

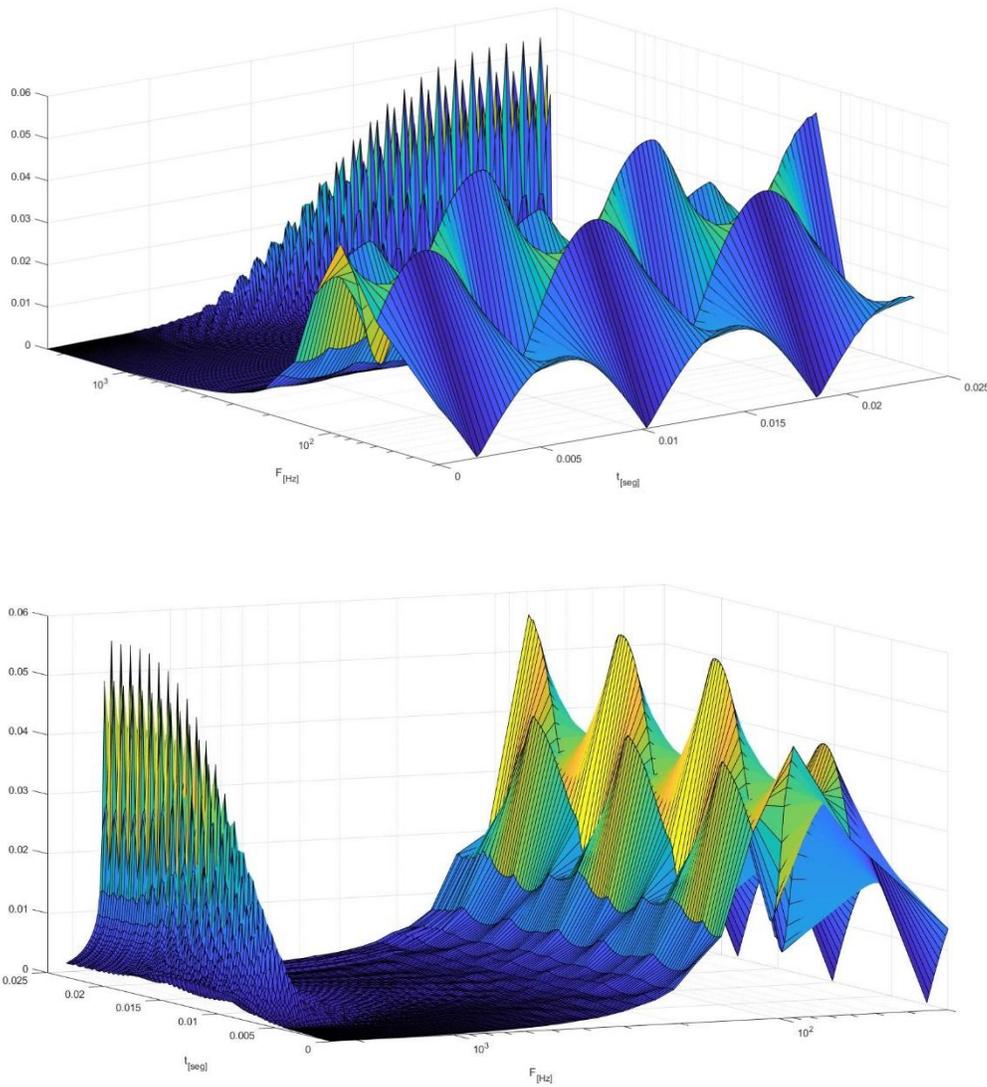


Nombre de la fuente: Elaboración propia.

También se propuso utilizar una vista en tres dimensiones del escalograma, en donde la energía constituye el eje Z y se crea una superficie dinámica. Cada figura contiene dos vistas diferentes del escalograma generado con el fin de observar diferentes dinámicas que podrían estar ocultas en una única vista. En las figuras 4-11 y 4-12 se presentan los escalogramas tridimensionales para la señal sinusoidal utilizando frecuencias de muestreo de 4800 Hz y 48000 Hz, respectivamente; en éstas es posible observar las variaciones en la magnitud de la energía con mayor facilidad dado que las vistas generadas para la frecuencia de muestreo de 4800 Hz presenta picos muy marcados en la componente de alta frecuencia, mientras que en el caso a 48000Hz, este comportamiento se suaviza y los puntos fríos desaparecen.

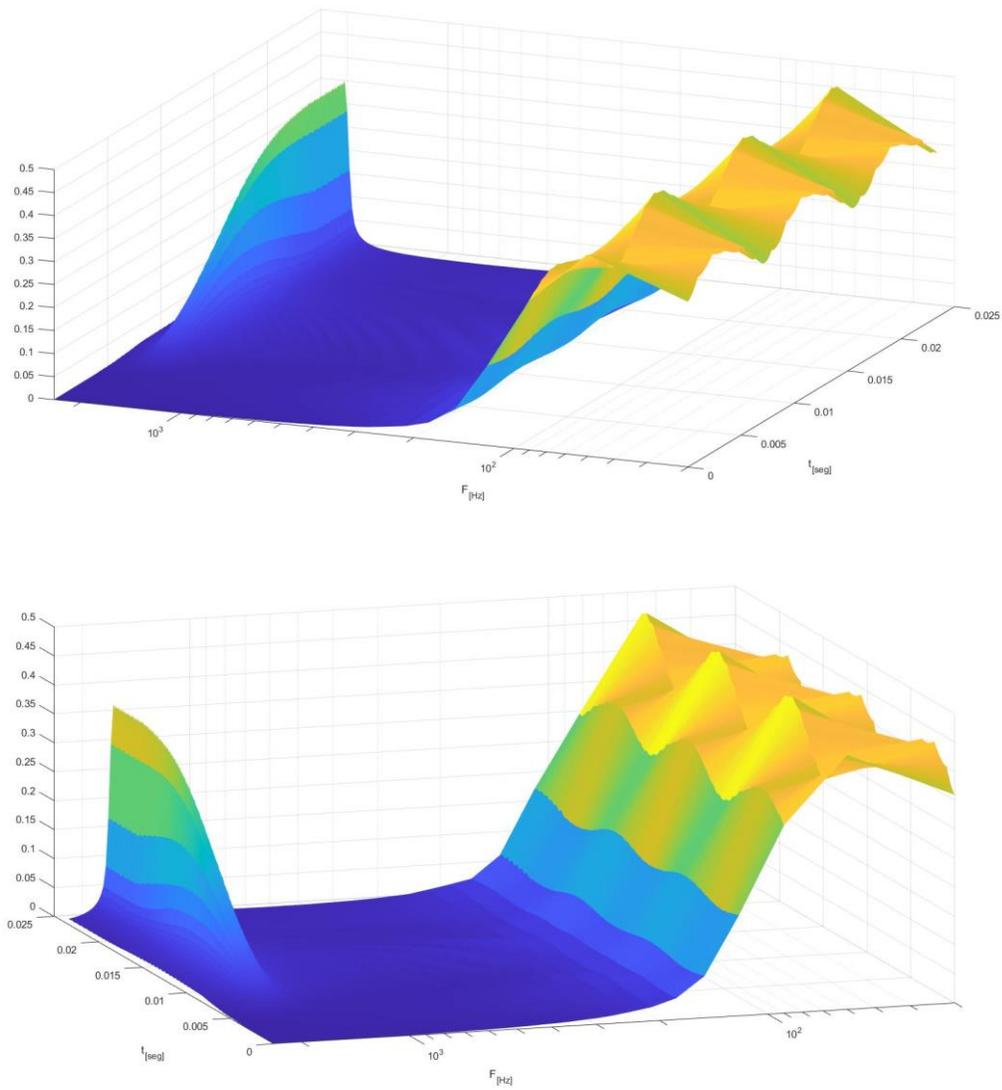
Por otro lado, las figuras 4-13 y 4-14 muestran los escalogramas tridimensionales para la señal de falla simulada utilizando frecuencias de muestreo de 4800 Hz y 48000 Hz, respectivamente. En éstas la energía asociada a la componente de alta frecuencia sigue teniendo una magnitud relativamente baja con respecto a la energía asociada a la componente a frecuencia industrial, aunque, de nuevo, puede observarse cómo la variabilidad presente a 4800 Hz se suaviza significativamente utilizando una frecuencia de muestreo mayor.

Figura 4-11: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 4800 Hz.



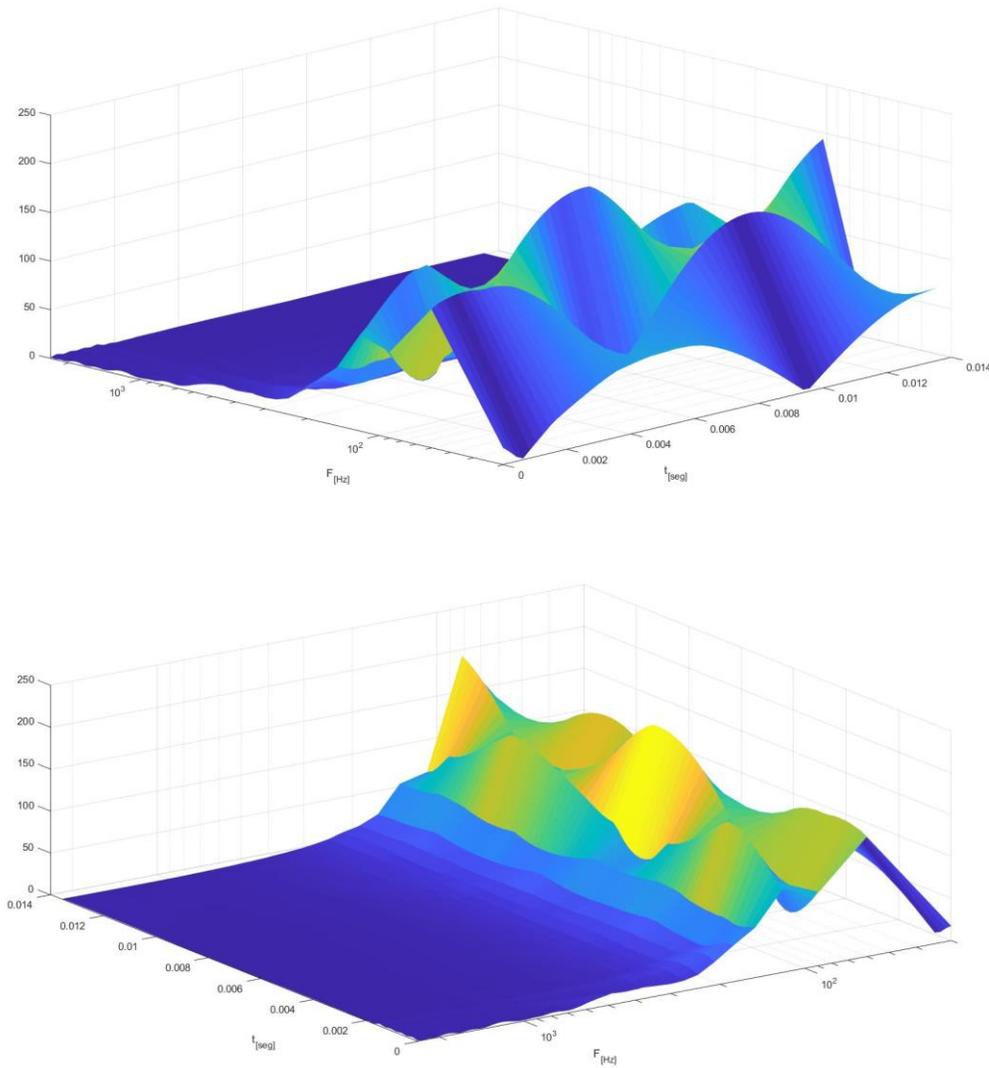
Nombre de la fuente: Elaboración propia.

Figura 4-12: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal sinusoidal y frecuencia de muestreo de 48000 Hz.



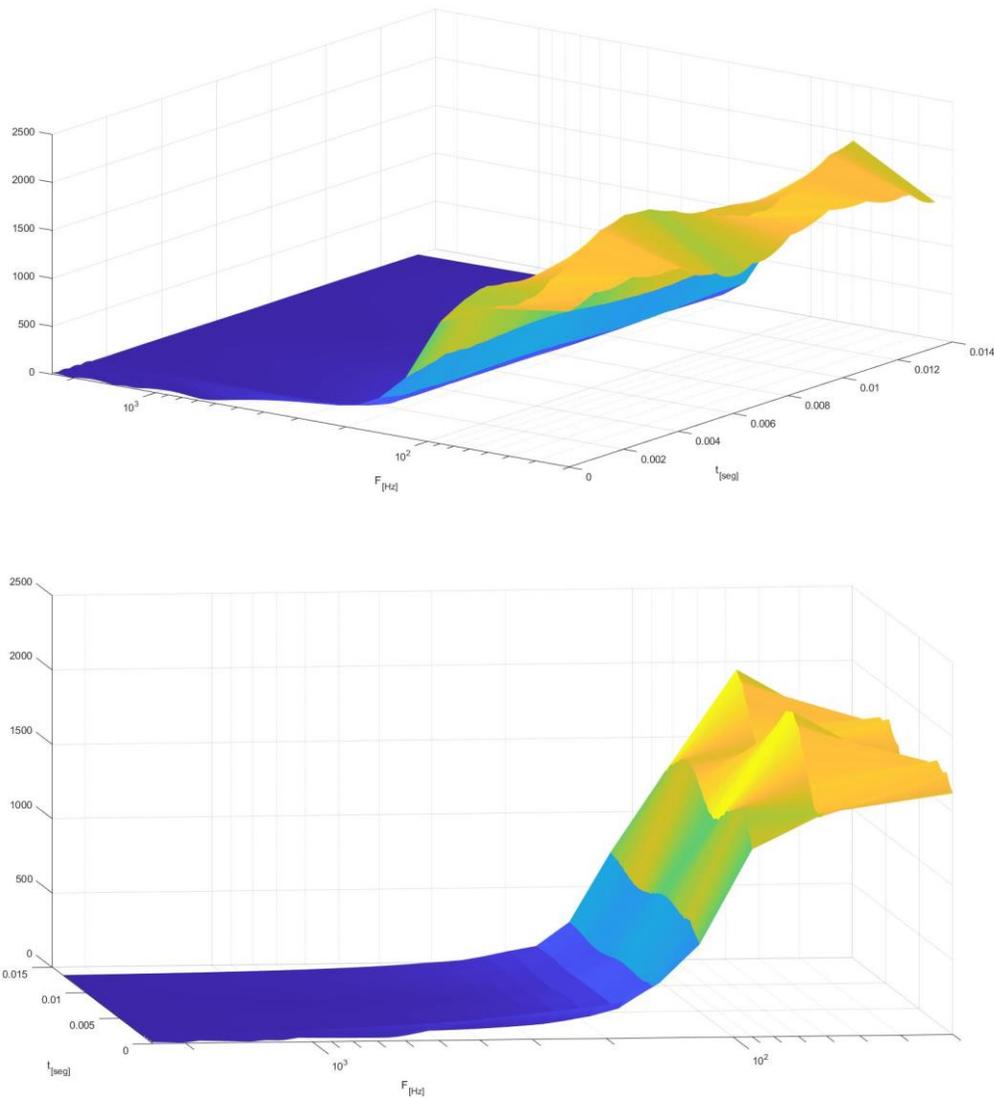
Nombre de la fuente: Elaboración propia.

Figura 4-13: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 4800 Hz.



Nombre de la fuente: Elaboración propia.

Figura 4-14: Escalograma en tres dimensiones obtenido con la metodología propuesta para la señal de falla simulada y frecuencia de muestreo de 48000 Hz.



Nombre de la fuente: Elaboración propia.

Con respecto al comportamiento variable de las componentes en frecuencia en las simulaciones obtenidas a 4800 Hz, es probable que esto se deba a la cantidad de componentes resultantes del proceso de remuestreo propuesto en la metodología dado que, cada vez que se computa el remuestreo para un factor p/q donde p es más pequeño, se tendrán menos muestras en la secuencia resultante. Este comportamiento es esperado, aunque cuando se utilizan frecuencias de muestreo menores, la cantidad de puntos de las

secuencias que componen las ventanas tomadas de la señal original, disminuye, lo que conlleva a que las secuencias resultantes de los remuestreos tengan una menor cantidad de coeficientes, lo que puede influenciar negativamente la información contenida en ésta.

Aludiendo a las Ecuaciones ((3.13) y (3.15)), en las que se relacionan los parámetros de F_s y M (longitud de las ventanas) con las resoluciones en tiempo y frecuencia del sistema, podemos observar una característica relevante del comportamiento del sistema en función de la frecuencia de muestreo escogida dado que afecta directamente la resolución en frecuencia e inversamente la resolución en tiempo. Si consideramos la frecuencia de muestreo propuesta de 48kHz y una ventana de un ciclo a frecuencia industrial (60Hz), obtenemos:

$$M = \frac{48000 \text{ Hz}}{60 \text{ Hz}} = 800 \text{ muestras}$$

$$\Delta t = \frac{1}{60} \text{ segundos} = 0.0167 \text{ segundos}$$

$$\Delta F = 7.5 \text{ Hz}$$

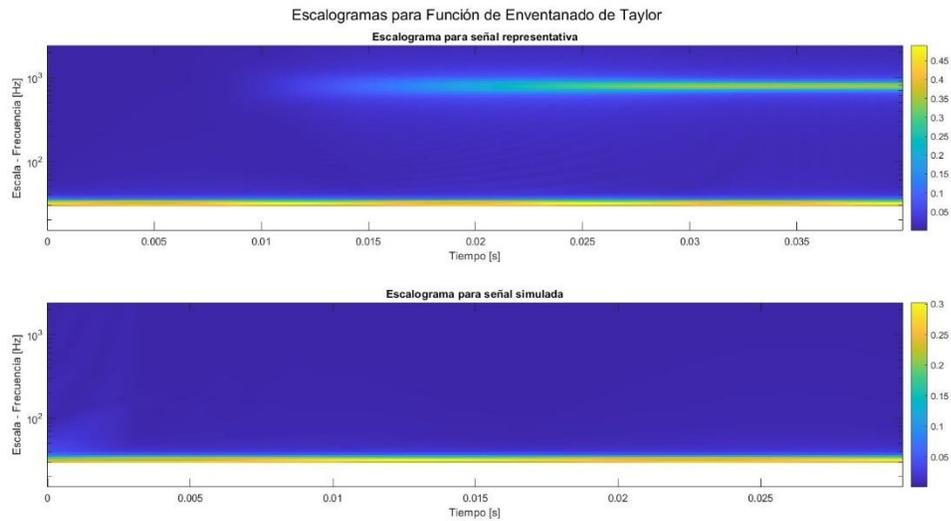
En este caso óptimo, en el que tomamos las ventanas de mayor longitud en relación con lo reportado en la literatura (en la que es común usar ventanas de un ciclo, medio ciclo o un cuarto de ciclo) y una frecuencia de muestreo de 48000 Hz, cada ventana conforma una secuencia digital de 800 puntos. Por otro lado, si utilizamos una tasa de muestreo de 4800 Hz, las ventanas tendrán una longitud máxima de 80 puntos.

Dado que se realizan remuestreos en función de los parámetros p/q (relacionados a f_i/F_c), podemos analizar un caso extremo en el que se pretenda computar la longitud de la secuencia remuestreada para un factor 1/40 (el cual se obtiene de considerar una banda de frecuencia a 60 Hz para una frecuencia de resonancia de 2400 Hz); en este caso, la longitud de la señal resultante será 39 muestras más corta que la original, lo cual no resulta significativo para 800 muestras en el caso a 48000 Hz, pero elimina casi el 50% de la longitud de las señales en el caso a 4800Hz.

Por otro lado, para la función de enventanado, se probaron las funciones de Taylor, de Hann y Rectangular, como se observa en las figuras 4-15, 4-16 y 4-17, respectivamente.

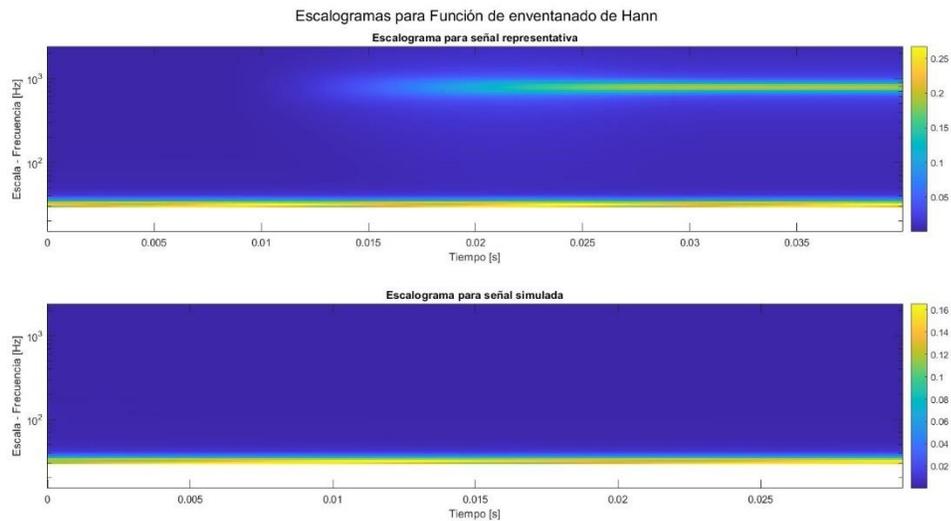
Al comparar los resultados obtenidos con las funciones de Taylor y Gibbs no se observó una diferencia significativa, aunque en el caso de los resultados obtenidos con las funciones rectangulares se observa una dispersión de energía en todo el espectro de frecuencia, probablemente relacionado al fenómeno de Gibbs introducido al inventanar de forma brusca la señal analizada.

Figura 4-15: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado de Taylor.



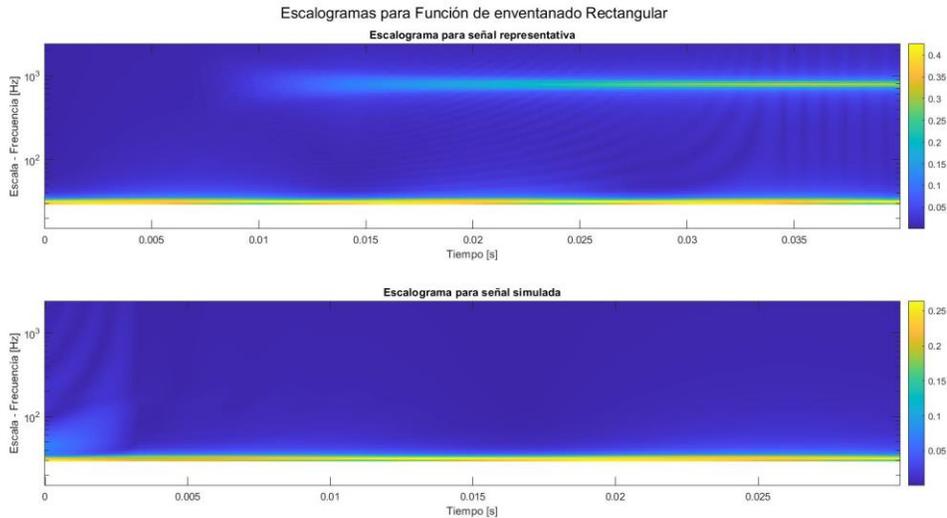
Nombre de la fuente: Elaboración propia.

Figura 4-16: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado de Hann.



Nombre de la fuente: Elaboración propia.

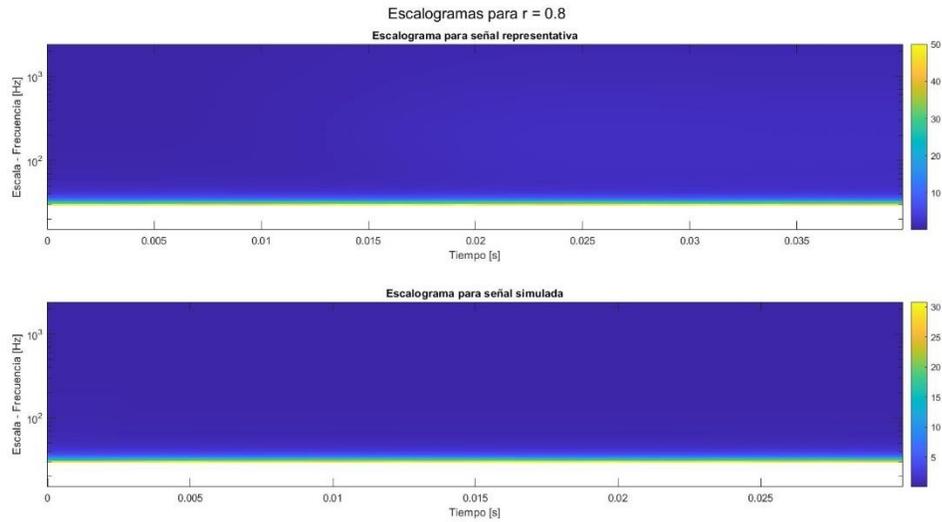
Figura 4-17: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con funciones de enventanado Rectangulares.



Nombre de la fuente: Elaboración propia.

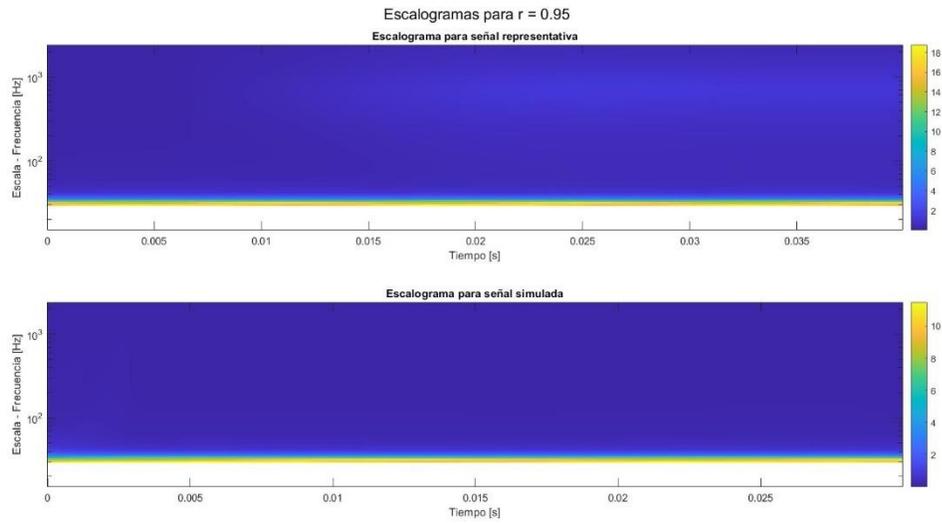
En cuanto al radio de convergencia del filtro, se probaron múltiples valores en el intervalo $(0.8, 0.99999)$, entre los que se destacan los valores de $r = 0.8$, $r = 0.95$, $r = 0.999$ y $r = 0.99999$, cuyos escalogramas, para ambas señales de prueba, se presentan en las figuras 4-18, 4-19, 4-20 y 4-21, respectivamente. Como puede observarse, para $r = 0.8$ y $r = 0.95$ la energía contenida en las escalas de alta frecuencia no se representa de forma satisfactoria dado que el filtro no es lo suficientemente selectivo; por otro lado, entre los valores de $r = 0.999$ y $r = 0.99999$ no existe una diferencia significativa en la representación de energía en los escalogramas, aunque, como se observará más adelante, existe una diferencia en los tiempos de computación del algoritmo.

Figura 4-18: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.8.



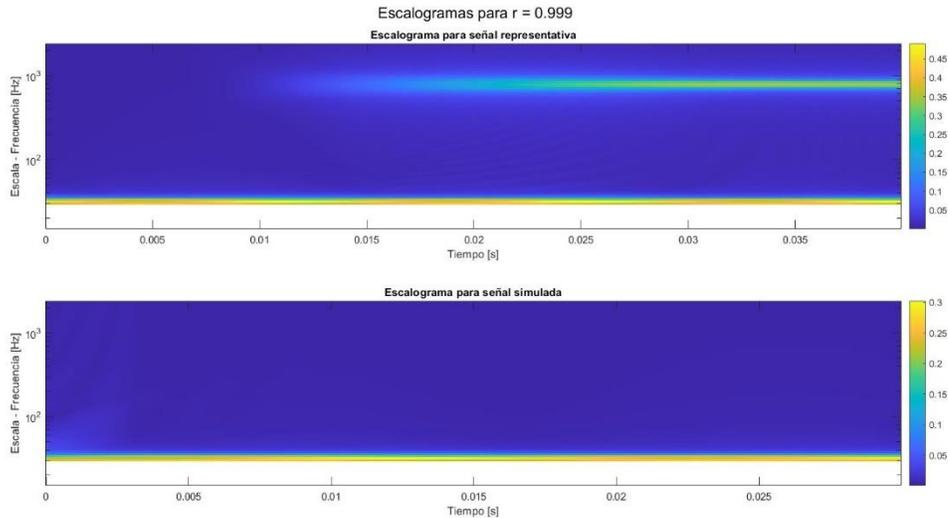
Nombre de la fuente: Elaboración propia.

Figura 4-19: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.95.



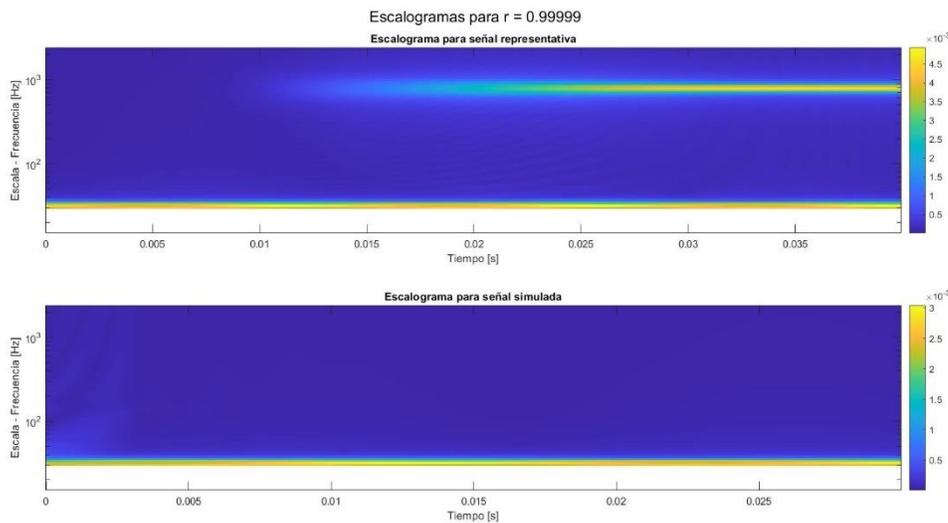
Nombre de la fuente: Elaboración propia.

Figura 4-20: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.999.



Nombre de la fuente: Elaboración propia.

Figura 4-21: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para radio de convergencia de 0.99999.



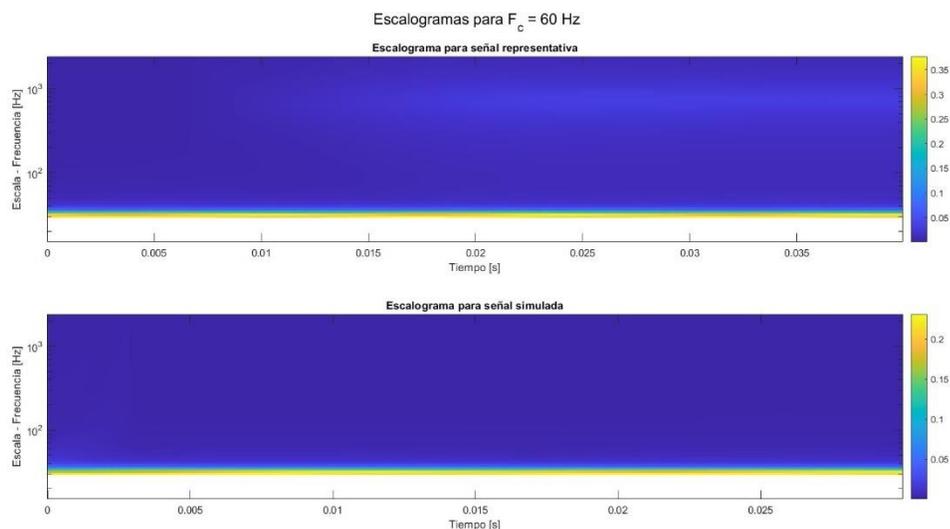
Nombre de la fuente: Elaboración propia.

Analicemos ahora la frecuencia de resonancia F_C del filtro IIR resonante; para ello se propone analizar su impacto en los tiempos de computación del algoritmo y la capacidad de generar una escala de frecuencias F , de forma tal que las componentes de este conjunto de frecuencia sean múltiplos o submúltiplos de F_C .

Esto último es relevante dado que se requiere computar la fracción p/q a partir de f_i/F_C para realizar el remuestreo de las señales inventanadas extraídas de la señal original; en términos de recursos y tiempos computacionales es mucho más eficiente trabajar con fracciones del tipo $3/7$ que $299/701$, por ejemplo, aunque ambas representen una fracción similar en términos absolutos. Las figuras 4-22, 4-23 y 4-24 muestran los escalogramas obtenidos para ambas señales de prueba para frecuencias de resonancia de 60 Hz, 1200 Hz y 2400 Hz, respectivamente.

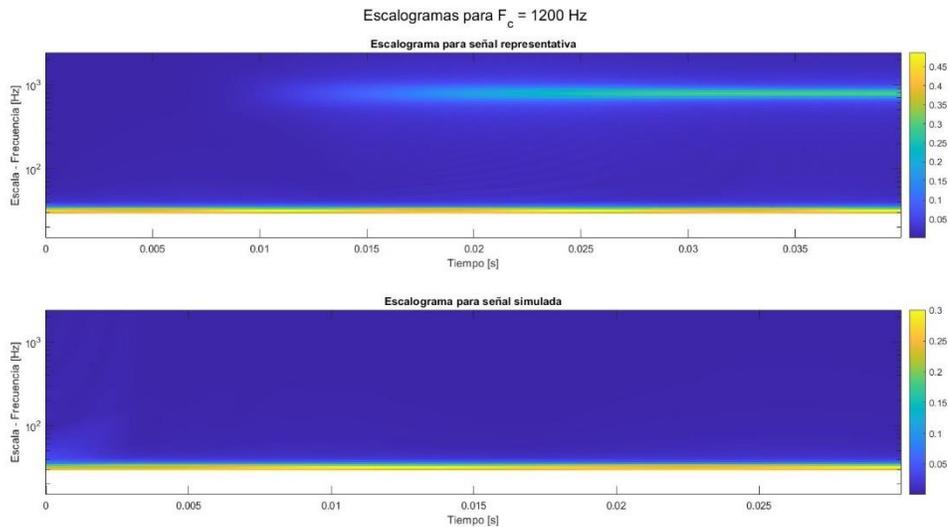
También cabe resaltar que se probaron otras metodologías para hallar los coeficientes p/q , computando el primer remuestreo de la ventana utilizando un submuestreo por una muestra en el numerador, de forma que se generasen factores de la forma $M/M - 1$ y continuar generando remuestreos a partir de la señal previa y eliminando una muestra. Si bien los resultados obtenidos son similares a los de la metodología propuesta, el tiempo de computación aumenta significativamente dado que se computaban M remuestreos de cada una de las ventanas de interés.

Figura 4-22: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 60Hz.



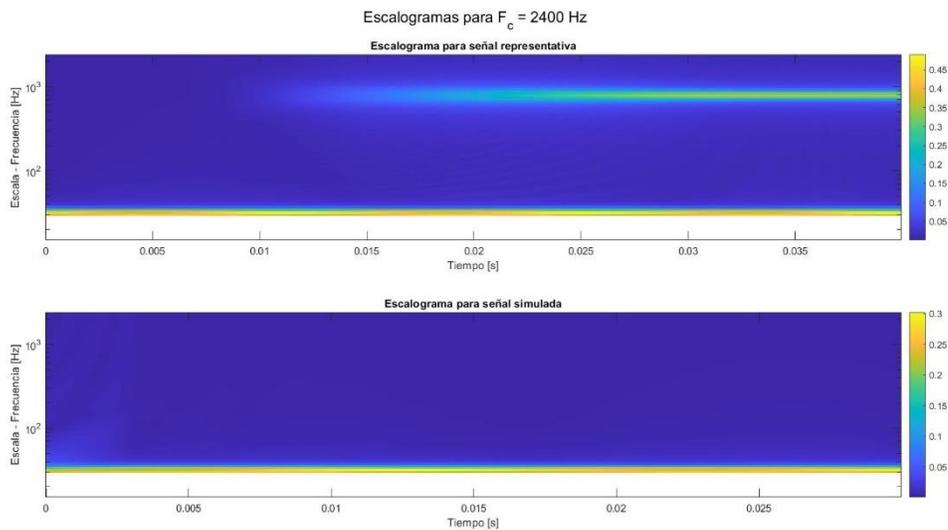
Nombre de la fuente: Elaboración propia.

Figura 4-23: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 1200Hz.



Nombre de la fuente: Elaboración propia.

Figura 4-24: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para frecuencia de resonancia del filtro IIR de 2400Hz.



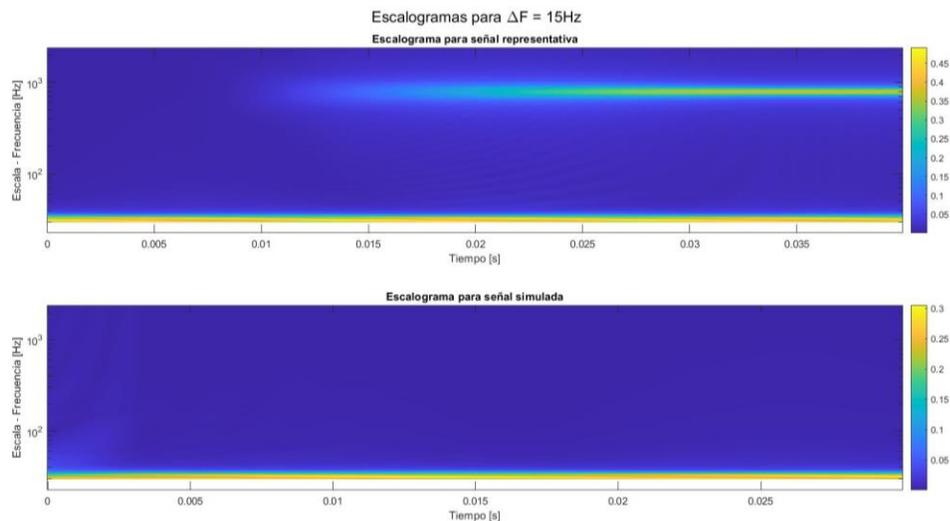
Nombre de la fuente: Elaboración propia.

En términos de los resultados obtenidos en los escalogramas, no existe una diferencia significativa entre el uso de una frecuencia de resonancia de 2400 Hz o de 1200 Hz, mientras que, para una frecuencia de resonancia de 60 Hz, la energía asociada a las componentes de alta frecuencia se reduce hasta casi desaparecer. El análisis de los tiempos de computación se expondrá al final de esta subsección del capítulo, en el que se presenta de forma integral esta característica para los parámetros analizados.

En relación con la discusión anterior, el método propuesto para generar las escalas de frecuencia F utiliza separaciones uniformes ΔF a partir de un valor de frecuencia mínima; cómo se observó en el análisis la frecuencia de muestreo, utilizando ventanas de longitud de un ciclo y una frecuencia de muestreo de 48000Hz, la resolución de frecuencia obtenida es de 7.5Hz. Así, se probaron escalas de frecuencia con separaciones de 15Hz, 30Hz y 60Hz cuyos resultados se presentan en las figuras 4-25, 4-26 y 4-27, respectivamente.

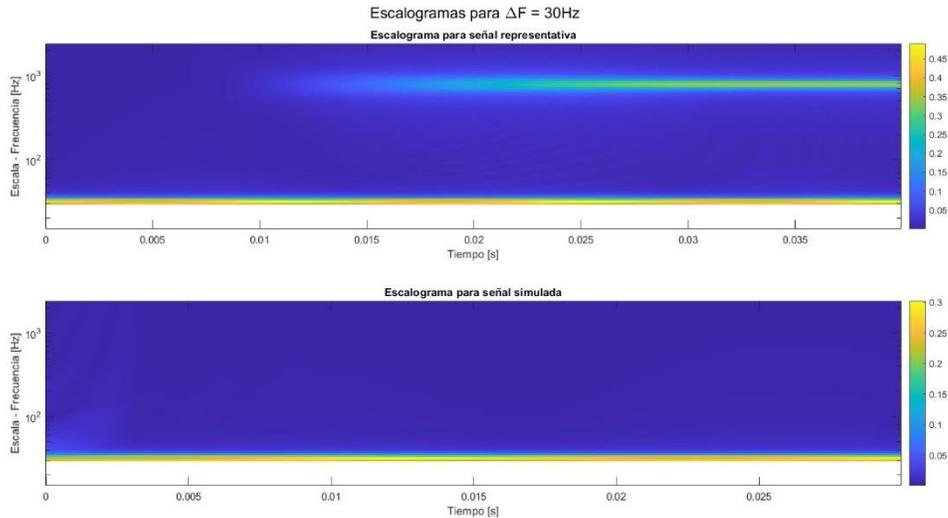
En términos de los coeficientes obtenidos, para valores lejanos al límite de resolución de frecuencia, como 60Hz, se puede observar (figura 4-27) la dispersión de energía que se obtiene al utilizar escalas de frecuencia más separadas, mientras que, para valores más cercanos, como 30Hz y 15Hz, los resultados obtenidos son similares, aunque el escalograma obtenido al utilizar pasos de frecuencia de 15Hz (figura 4-25) presenta divisiones más finas y suaves en comparación con el obtenido para pasos de frecuencia de 30Hz (figura 4-26).

Figura 4-25: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 15\text{Hz}$..



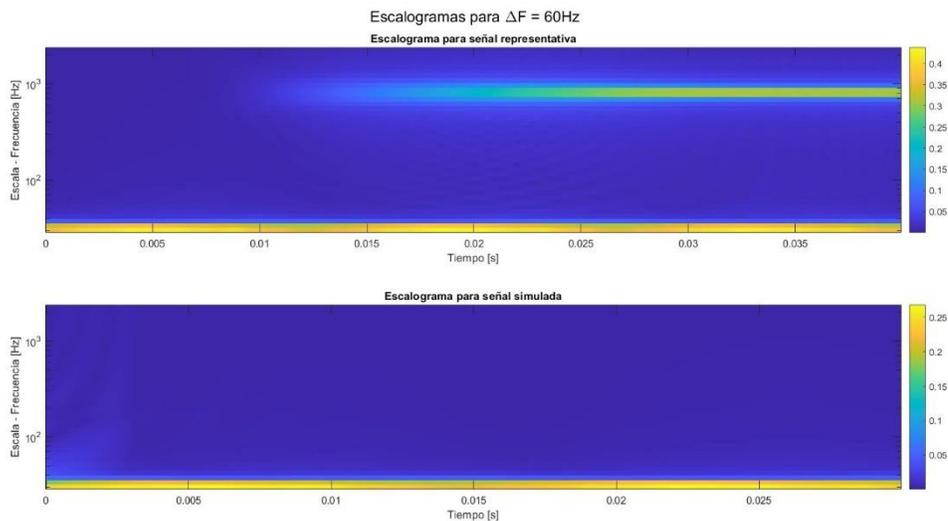
Nombre de la fuente: Elaboración propia.

Figura 4-26: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 30\text{Hz}$.



Nombre de la fuente: Elaboración propia.

Figura 4-27: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) para $\Delta F = 60\text{Hz}$.



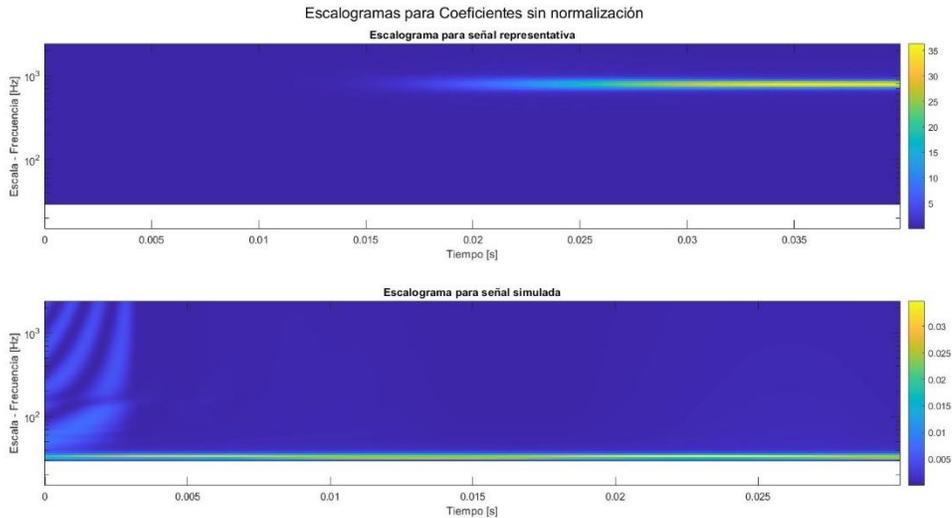
Nombre de la fuente: Elaboración propia.

Por otro lado, se probó el algoritmo utilizando una escala logarítmica con el fin de concentrar los remuestreos alrededor de F_C ; sin embargo, no se puede garantizar la multiplicidad de las frecuencias generadas y F_C lo que alarga los tiempos de computación y se incrementa la dispersión en la densidad de potencia espectral de forma significativa en las frecuencias alejadas a F_C . Se concluyó que utilizar esta metodología podría introducir errores en las frecuencias de interés que no estén centradas en F_C .

Finalmente, se observó el impacto de las normalizaciones propuestas para los coeficientes Wavelet obtenidos por la metodología propuesta en los escalogramas generados. En las figuras 4-28, 4-29 y 4-30 se presentan los resultados de los escalogramas para los coeficientes computados sin normalización, con normalización debido al remuestreo de las ventanas y con normalización debido al remuestreo de las ventanas con normalización L1, respectivamente.

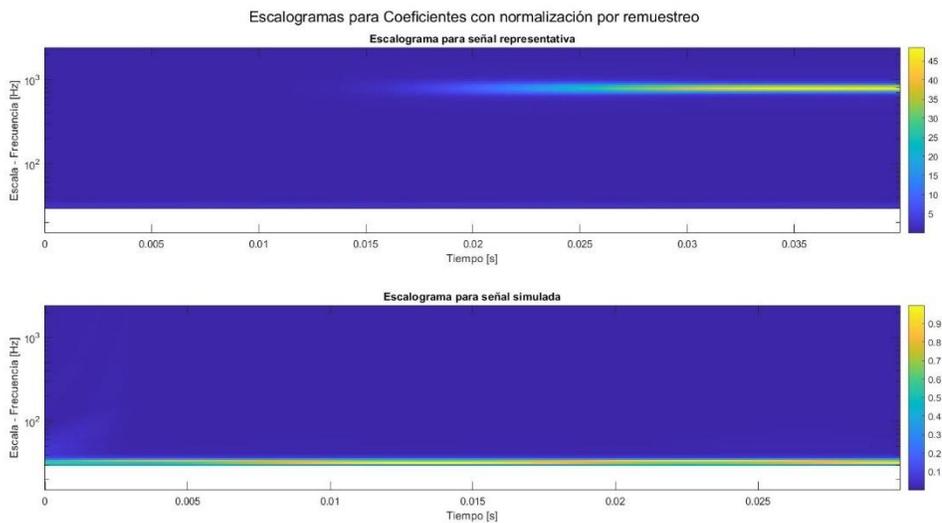
Como puede observarse en la figura 4-28 cuando no se utiliza normalización alguna no se detecta el componente de baja frecuencia en la señal sinusoidal mientras que para la señal de falla simulada las frecuencias asociadas al evento transitorio tienen una magnitud desproporcionada. Por otro lado, cuando sólo se corrigen los coeficientes por el efecto del remuestreo de las ventanas (figura 4-29), se logra adecuar las altas frecuencias observadas en la señal de falla simulada y aparece un coeficiente en baja frecuencia para la señal sinusoidal, aunque su magnitud no es la adecuada. Sin embargo, al introducir la normalización L1 (figura 4-30) se logran detectar tanto los coeficientes de alta y baja frecuencia en ambas señales con magnitudes similares a las reportadas por el método de referencia (DWT con db4), aunque se puede observar un aumento en la dispersión de energía cerca de los puntos donde ocurren los eventos transitorios.

Figura 4-28: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes wavelet sin normalización.



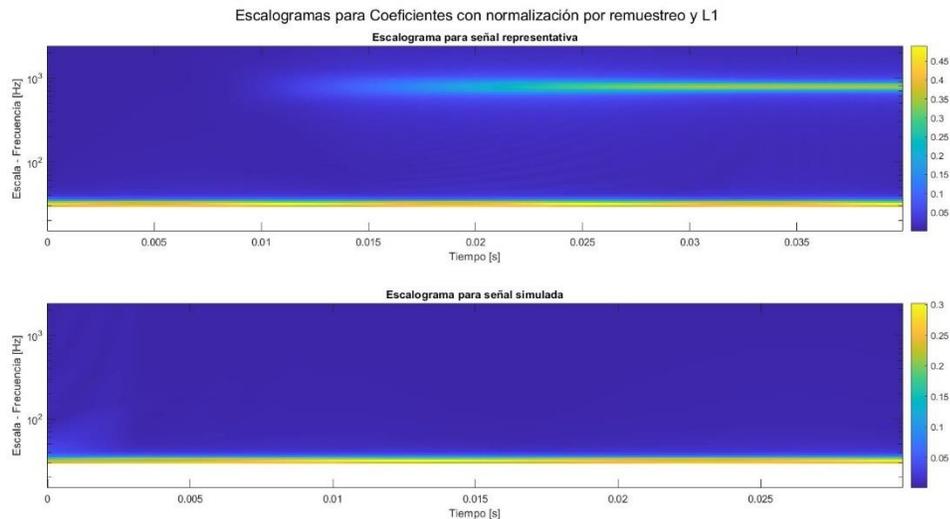
Nombre de la fuente: Elaboración propia.

Figura 4-29: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes con normalización por remuestreo.



Nombre de la fuente: Elaboración propia.

Figura 4-30: Escalograma obtenido con la metodología propuesta para la señal sinusoidal (arriba) y la señal de falla simulada (abajo) con coeficientes con normalización por remuestreo y normalización L1.



Nombre de la fuente: Elaboración propia.

Aquí es relevante resaltar que no existe una diferencia significativa en cuanto al tiempo de computación con o sin las normalizaciones propuestas dado que MATLAB realiza estas operaciones de forma matricial con un consumo mínimo de recursos, por lo que se optó por no reportar estos resultados en análisis de tiempos computacionales que se presenta a continuación. En la tabla 4-2 se presentan los tiempos de simulación en MATLAB utilizando un equipo personal con un procesador Intel Core i7 de décima generación y 16 Gb de memoria RAM. Los resultados reportados corresponden sólo al remuestreo de las ventanas, la aplicación del filtro y el cómputo de los coeficientes dado que éstos son los procesos que deberían realizarse en tiempo real si se deseara implementar esta metodología. Para reportar estos tiempos se realizó cada simulación 10 veces y tomó el promedio de tiempos, asegurando condiciones similares en todos los casos con respecto al entorno de simulación (no había otros procesos consumiendo recursos); sin embargo, debido al número reducido de muestras no se puede asegurar la significancia estadística de la diferencia de la media de los tiempos reportados.

Tabla 4-2: Tiempos promedio de simulación para el algoritmo propuesto en las señales sinusoidal y de falla simulada en función de la frecuencia de muestreo, frecuencia de resonancia, radio de convergencia, función de enventanado y paso de frecuencia. Elaboración propia.

Parámetro	Valor/Tipo	Tiempo de simulación [segundos]	
		Señal sinusoidal	Señal de falla simulada
Frecuencia de Muestreo [Hz]	4800	2.94	2.25
	48000	32.57	26.11
Frecuencia de Resonancia [Hz]	60	65.48	36.35
	1200	41.49	17.55
	2400	30.66	16.14
Radio de Convergencia	0.8	32.11	18.43
	0.95	33.89	19.41
	0.999	33.97	19.71
	0.99999	34.61	19.97
Función de enventanado	Taylor	31.21	23.03
	Han	31.92	23.72
	Rectangular	30.35	22.95
ΔF [Hz]	15	73.56	43.06
	30	31.43	24.23
	60	18.47	11.82

Los resultados se sintetizan en los siguientes puntos de interés:

- La frecuencia de muestreo es uno de los parámetros con mayor influencia en los tiempos de computación del algoritmo debido a que ésta impacta directamente el número de puntos que contiene cada ventana. Está claro que, controlando las demás variables, el procesar N ventanas de longitud M_1 con respecto a N ventanas de longitud M_2 donde $M_1 > M_2$ consumirá más recursos y tardará más tiempo al considerar que se debe remuestrear cada ventana una cantidad de veces dependiendo de la escala de frecuencias y aplicar el filtro a cada una de éstas.
- En cuanto a la frecuencia de resonancia del filtro podemos observar que ésta afecta de forma inversa al tiempo de cómputo. En este caso, es relevante recordar que se intentó utilizar valores de frecuencia de resonancia que fuesen múltiplos o submúltiplos de las componentes de la escala de frecuencias, por lo que se optó por utilizar múltiplos de 60 Hz. El comportamiento observado puede explicarse recordando que el factor de muestreo computado proviene de la fracción p/q a partir de f_i/F_C y que la escala de frecuencias es un vector de frecuencias de la

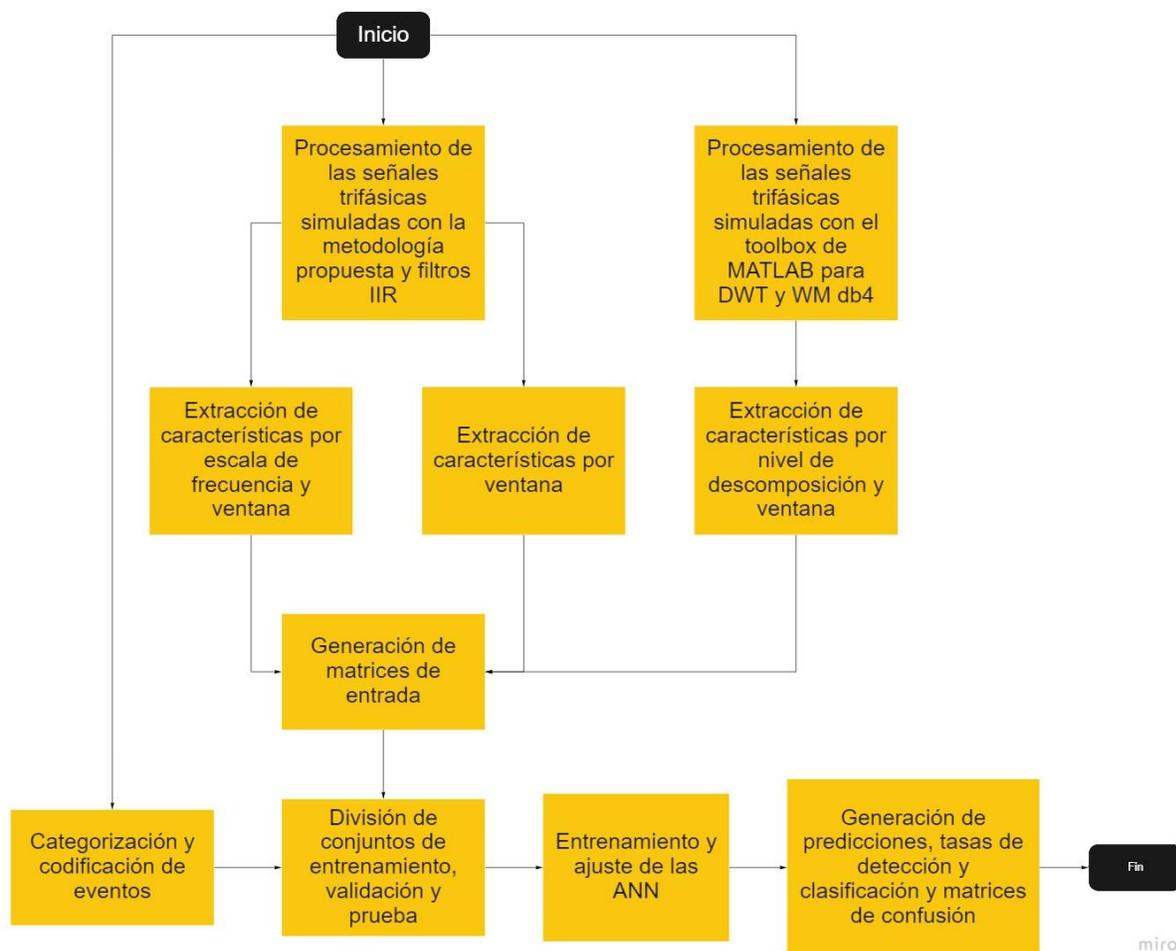
forma $\{f_{min}, f_{min} + \Delta F, f_{min} + 2\Delta F, \dots, F_c\}$ (ecuación 40); cuando se utiliza $F_c = 60\text{Hz}$, por ejemplo, los factores p/q serán, en su mayoría, mayores a 1, por lo que se deben ejecutar más sobremuestras que submuestras, mientras que para $F_c = 2400\text{Hz}$, los factores serán menores a 1, por lo que se deben ejecutar más submuestras que sobremuestras. Según la forma en que MATLAB computa el remuestreo de las señales digitales es evidente que los sobremuestras tienen una mayor carga computacional que los submuestras dado que éstos alargan la señal original antes de aplicar el filtro de aliasing.

- Para el radio de convergencia del filtro IIR resonante se puede observar una relación directa entre la proximidad de este valor a la unidad y el tiempo de simulación del algoritmo. En particular, hay una diferencia significativa entre 0.8 y 0.95, aunque a partir de este valor los cambios no son tan marcados por lo que el rendimiento con respecto a la validez de los coeficientes toma mayor relevancia.
- Con respecto a la función de enventanado, aunque se observan ciertas diferencias en los tiempos, es difícil atribuirlos a las funciones en sí dado que éstas se aplican a las ventanas antes de ejecutar el algoritmo y no entran en la medida de los tiempos reportados. Dado que el algoritmo recibe ventanas de igual longitud luego de ser enventanadas no deberían existir diferencias apreciables en cuanto a los tiempos de ejecución, por lo que se éstas se pueden atribuir a aleatoriedad y al número reducido de muestras.
- Finalmente, el paso de frecuencia utilizado para crear la escala de frecuencias juega un papel fundamental en el consumo de recursos para el algoritmo dado que está directamente relacionado al número de remuestras que se deben computar (para 60Hz se computan 40 remuestras, mientras que para 30 Hz son 80 y a 15 Hz se deben ejecutar 160) para cada una de las ventanas. Reducir el paso de frecuencia en un 50% aparentemente dobla el tiempo de ejecución del algoritmo, aunque produce resultados más precisos como se observó en el análisis de coeficientes, por lo que su selección se vuelve un asunto de balance entre consumo de recursos e incertidumbre en la precisión de los resultados.

4.3 Sistemas de detección y clasificación

Luego de validar la metodología propuesta se utilizó el banco de eventos simulados para obtener los coeficientes wavelet asociados a cada uno de éstos. Es relevante aclarar que en cada evento se registraron las tensiones trifásicas, por lo que se generaron coeficientes trifásicos y, por ende, las características estadísticas extraídas también son trifásicas. A continuación, se presentan un diagrama de flujo del procedimiento realizado para generar los conjuntos de información de los eventos, el entrenamiento de los algoritmos de detección y clasificación y la evaluación de su rendimiento:

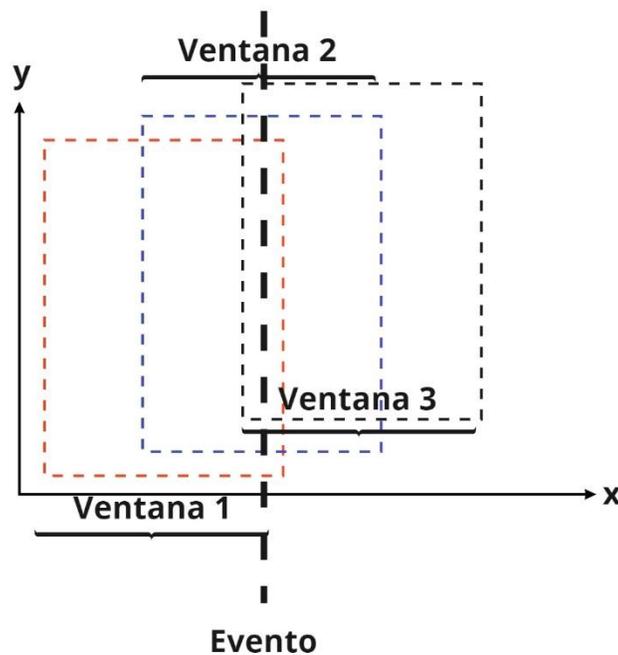
Figura 4-31: Diagrama de flujo asociado a la generación de los conjuntos de entrada y salida para los algoritmos de detección y clasificación de eventos y la evaluación y clasificación de éstos.



Nombre de la fuente: Elaboración propia.

Aquí es relevante resaltar que para cada evento que se pretendía detectar y clasificar se consideraron 3 ventanas alrededor del momento en el que ocurre éste: una ventana antes del momento de falla (que incluye el 10% de la información inicial del evento) otra ventana que centraliza el momento de la falla y la última luego del momento de falla (que incluye el 10% de la información final del evento). Esta metodología se presenta en la figura 4-32.

Figura 4-32: Ventanas pre-falla, en falla y post-falla propuestas para la extracción de características.



Nombre de la fuente: Elaboración propia.

La utilización de 3 ventanas por evento tiene dos objetivos fundamentales:

- Considerar las dinámicas previas y posteriores al momento de los eventos (pre y post-falla).
- Ampliar el conjunto de información de aprendizaje de los algoritmos de detección y clasificación.

Ahora analicemos los conjuntos de información obtenidos luego de obtener los coeficientes wavelet y realizar la extracción de características mediante parámetros estadísticos. Como se observa en la figura 4-31, luego del procesamiento de las señales, bien sea por la metodología propuesta o utilizando el ToolBox de MATLAB, se computan los parámetros

Figura 4-34: Matriz de características obtenidas con los coeficientes wavelet hallados usando la metodología propuesta, por ventana.

12 características

		FASE A				FASE B				FASE C			
		AVG	Kurtosis	PSD	RMS	AVG	Kurtosis	PSD	RMS	AVG	Kurtosis	PSD	RMS
Evento 1	Ventana 1												
	Ventana 2												
	Ventana 3												
...	...												
Evento N	Ventana 1												
	Ventana 2												
	Ventana 3												

miro

Nombre de la fuente: Elaboración propia.

Figura 4-35: Matriz de características obtenidas con los coeficientes wavelet hallados usando la metodología propuesta, por escala de frecuencia.

960 características

		FASE A												...	FASE C											
		AVG			Kurtosis			PSD			RMS			...	AVG			Kurtosis			PSD			RMS		
		SF 1	...	SF 80	SF 1	...	SF 80	SF 1	...	SF 80	SF 1	...	SF 80	...	SF 1	...	SF 80	SF 1	...	SF 80	SF 1	...	SF 80	SF 1	...	SF 80
Evento 1	Ventana 1																									
	Ventana 2																									
	Ventana 3																									
...	...																									
Evento N	Ventana 1																									
	Ventana 2																									
	Ventana 3																									

miro

Nombre de la fuente: Elaboración propia.

Por otro lado, la codificación y categorización de los eventos de interés se presenta en la tabla 4-3. Estas categorías corresponden las salidas asignadas cada uno de los eventos y ventanas con los que se generaron las matrices de entrada presentadas anteriormente.

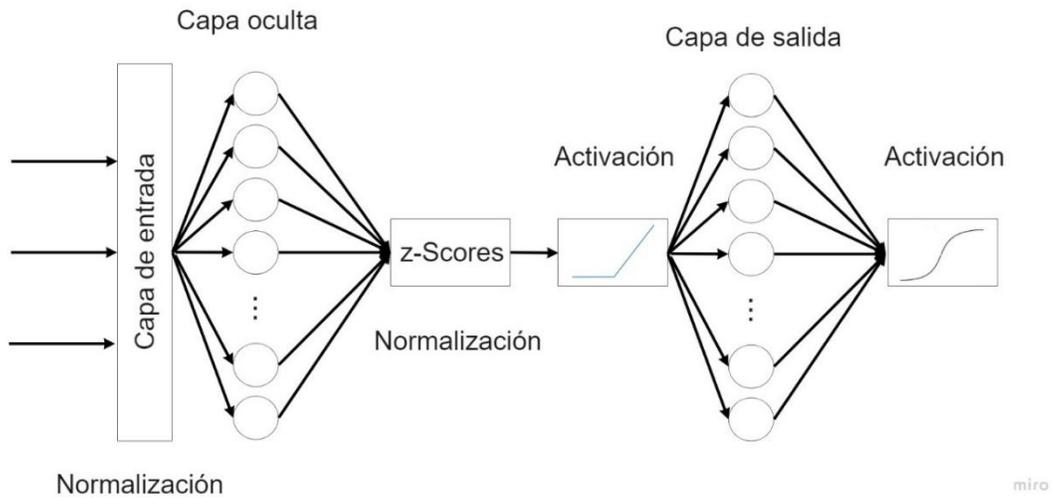
Tabla 4-3: Codificación y categorización de eventos simulados. Elaboración propia.

Salida			Categorización
0	0	0	Operación normal
0	0	1	Falla trifásica
0	1	0	Falla bifásica - Fases AB
0	1	1	Falla bifásica - Fases BC
1	0	0	Falla bifásica - Fases AC
1	0	1	Falla monofásica - Fase A
1	1	0	Falla monofásica - Fase B
1	1	1	Falla monofásica - Fase C

A continuación, para todos los conjuntos de información se realizó una división en los conjuntos de prueba (15%) y entrenamiento (85%). Sobre este conjunto de entrenamiento se realizó otra división para obtener el conjunto de validación (30%). Utilizando los conjuntos de entrenamiento obtenidos de cada una de las matrices presentadas en las figuras 4-33, 4-34 y 4-35 se realizó un proceso de ensayo y error hasta obtener resultados aceptables, reconocidos, en primera instancia, por su precisión de validación.

Luego de probar múltiples esquemas de redes neuronales profundas, se observó que el modelo más adecuado en términos de precisión en los resultados obtenidos es el presentado en la figura 4-36. En este modelo se tiene una capa de entrada (y normalización), una única capa oculta cuyo número de neuronas depende del conjunto de datos a utilizar, una capa de normalización que utiliza z-Scores, una capa de activación de la capa oculta con función ReLu, una capa de salida con 8 neuronas (una para cada categoría de clasificación) y su respectiva capa de activación con función SoftMax.

Figura 4-36: Esquema de ANN propuesto para la detección y clasificación de eventos.



Nombre de la fuente: Elaboración propia.

Los parámetros utilizados para cada una de las ANN propuestas, según el esquema de la Figura 4-36 se presentan en la Tabla 4-4. Estos parámetros fueron determinados mediante ensayo y error evaluando la precisión de validación durante la etapa de entrenamiento.

Tabla 4-4: Parámetros de las redes neuronales utilizadas para los conjuntos de información obtenidos. Elaboración propia.

	Conjunto de información		
	DWT	CWT por ventana	CWT por escala de frecuencia
Neuronas en la capa oculta	60	100	150
Épocas	1000	1000	600
Tamaño de conjunto de entrenamiento	8	8	10
Learning Rate	0.003	0.001	0.004
Coeficiente L2	0.008	0.0025	0.001
Precisión [%]	78.9	71.9	80.2

La precisión reportada en la tabla anterior hace referencia al porcentaje de aciertos totales obtenidos al comparar las salidas predichas por la ANN con el conjunto de entrada de prueba y su respectiva salida. Esta información está contenida, a su vez, en la matriz de confusión, la cual es el método de presentación de resultados que se propone en este apartado dado que muestra el desempeño de la red para cada una de las clases y su desempeño general.

Las filas de la matriz de confusión corresponden a las clases predichas por la ANN, mientras que las columnas corresponden a la clase verdadera. Las celdas observadas en la diagonal principal representan las observaciones clasificadas correctamente, mientras que las celdas fuera de la diagonal muestra las clasificaciones erróneas y cada celda muestra el total de observaciones y su porcentaje respecto al total de observaciones. La fila al final de la matriz muestra los porcentajes de todos los ejemplos (del conjunto de prueba) que pertenecen a cada clase y si éstos son correcta o incorrectamente clasificados; estas métricas se conocen como tasa de verdaderos positivos y tasa de falsos positivos, respectivamente. Finalmente, la columna a la derecha de la matriz muestra el porcentaje de todos los casos que se predice que pertenecen a cada clase y que son correcta o incorrectamente clasificados; estas métricas son conocidas como precisión y tasa de descubrimiento falso, respectivamente.

A continuación, en las Figuras 4-37, 4-38 y 4-39 se presentan las matrices de confusión obtenidas para las ANN entrenadas con los conjuntos de información generados mediante la DWT, la metodología propuesta (CWT) por ventana y la metodología propuesta por escala de frecuencia, respectivamente.

Figura 4-37: Matriz de confusión para ANN usando DWT e información por nivel de descomposición.

Confusion Matrix

000	11 8.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
001	0 0.0%	9 7.0%	1 0.8%	0 0.0%	1 0.8%	3 2.3%	0 0.0%	0 0.0%	64.3% 35.7%
010	0 0.0%	5 3.9%	14 10.9%	0 0.0%	0 0.0%	4 3.1%	1 0.8%	1 0.8%	56.0% 44.0%
011	0 0.0%	1 0.8%	0 0.0%	10 7.8%	0 0.0%	0 0.0%	0 0.0%	2 1.6%	76.9% 23.1%
100	0 0.0%	1 0.8%	0 0.0%	0 0.0%	13 10.2%	4 3.1%	0 0.0%	0 0.0%	72.2% 27.8%
101	1 0.8%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	14 10.9%	0 0.0%	0 0.0%	87.5% 12.5%
110	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	17 13.3%	1 0.8%	94.4% 5.6%
111	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	13 10.2%	100% 0.0%
	91.7% 8.3%	56.2% 43.8%	93.3% 6.7%	100% 0.0%	86.7% 13.3%	56.0% 44.0%	94.4% 5.6%	76.5% 23.5%	78.9% 21.1%
	000	001	010	011	100	101	110	111	
	Target Class								

Nombre de la fuente: Elaboración propia.

Figura 4-38: Matriz de confusión para ANN usando CWT e información por ventana.

Confusion Matrix

Output Class	000	001	010	011	100	101	110	111	
000	7 5.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	0 0.0%	0 0.0%	87.5% 12.5%
001	0 0.0%	9 7.0%	2 1.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	81.8% 18.2%
010	0 0.0%	7 5.5%	10 7.8%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	1 0.8%	52.6% 47.4%
011	0 0.0%	1 0.8%	0 0.0%	12 9.4%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	85.7% 14.3%
100	0 0.0%	3 2.3%	0 0.0%	0 0.0%	6 4.7%	4 3.1%	0 0.0%	2 1.6%	40.0% 60.0%
101	0 0.0%	1 0.8%	2 1.6%	0 0.0%	0 0.0%	14 10.9%	0 0.0%	1 0.8%	77.8% 22.2%
110	0 0.0%	1 0.8%	0 0.0%	3 2.3%	0 0.0%	0 0.0%	12 9.4%	3 2.3%	63.2% 36.8%
111	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.8%	1 0.8%	22 17.2%	91.7% 8.3%
	100% 0.0%	40.9% 59.1%	71.4% 28.6%	80.0% 20.0%	100% 0.0%	70.0% 30.0%	85.7% 14.3%	73.3% 26.7%	71.9% 28.1%
	000	001	010	011	100	101	110	111	
	Target Class								

Nombre de la fuente: Elaboración propia.

Figura 4-39: Matriz de confusión para ANN usando CWT e información por escala de frecuencia.

Confusion Matrix

Output Class	000	001	010	011	100	101	110	111	
000	5 6.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
001	0 0.0%	7 8.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
010	0 0.0%	2 2.5%	5 6.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	71.4% 28.6%
011	0 0.0%	0 0.0%	0 0.0%	10 12.3%	0 0.0%	0 0.0%	4 4.9%	0 0.0%	71.4% 28.6%
100	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 8.6%	1 1.2%	0 0.0%	0 0.0%	87.5% 12.5%
101	0 0.0%	2 2.5%	0 0.0%	0 0.0%	1 1.2%	7 8.6%	0 0.0%	0 0.0%	70.0% 30.0%
110	0 0.0%	0 0.0%	3 3.7%	1 1.2%	0 0.0%	0 0.0%	14 17.3%	0 0.0%	77.8% 22.2%
111	0 0.0%	0 0.0%	1 1.2%	0 0.0%	1 1.2%	0 0.0%	0 0.0%	10 12.3%	83.3% 16.7%
	100% 0.0%	63.6% 36.4%	55.6% 44.4%	90.9% 9.1%	77.8% 22.2%	87.5% 12.5%	77.8% 22.2%	100% 0.0%	80.2% 19.8%
	000	001	010	011	100	101	110	111	

Target Class

Nombre de la fuente: Elaboración propia.

Observando la Figura 4.37 podemos afirmar que el método convencional de utilizar la DWT con funciones madre db4 es efectivo para detectar eventos dado que la precisión de clasificación para la clase '000' correspondiente a la operación normal del sistema es del 100%; por otro lado, la mayoría de clases tiene una precisión de clasificación cercana al 90% aunque las clases '001' y '010', correspondientes a fallas trifásicas y fallas bifásicas

en fases A y B, respectivamente, tienen precisiones cercanas al 60% lo que afecta negativamente la precisión total del algoritmo de clasificación.

Por otro lado, para la metodología propuesta utilizando la CWT (Figura 4.38) y extrayendo la información por ventana hay una tasa de detección del 87.5%, mientras que el resto de las clases tiene una precisión cercana al 80%, entre las que se destaca la clase '100' correspondiente a fallas bifásicas entre las fases A y C, la cual tiene una precisión del 40%, por lo que la precisión total del algoritmo de clasificación es significativamente menor en comparación con el método convencional.

Finalmente, al observar los resultados de la metodología propuesta y extrayendo la información por escalas de frecuencia (Figura 4.39), podemos determinar que el algoritmo de clasificación asociado tiene una tasa de detección de eventos del 100%, al igual que en el caso de la DWT; sin embargo, la menor precisión observada a través de las clases es del 70%, lo que mejora su precisión total en relación con la reportada para la DWT en 1.3%. También es relevante destacar que este conjunto de información implica una menor utilización de recursos y tiempos en la etapa de entrenamiento del algoritmo de clasificación tal y como se observa en la cantidad de épocas de entrenamiento y el tamaño del conjunto de entrenamiento, reportados en la Tabla 4 – 4; esto implica que el conjunto de información obtenido por la metodología propuesta es superior al generado en la metodología convencional dado que existe una mayor capacidad de aprendizaje y explicación de las características.

5. Conclusiones y recomendaciones

5.1 Conclusiones

Observando los resultados de la sección 4.2, es pertinente afirmar que se cumplió con el objetivo principal planteado para el desarrollo de la tesis dado que la metodología para la implementación de la Transformada Wavelet Continua con filtros madre tipo IIR fue validada satisfactoriamente al compararla con los resultados obtenidos al utilizar la Transformada Wavelet Discreta con funciones madre convencionales en dos señales de prueba.

En cuanto a los efectos de los parámetros evaluados en la metodología propuesta, se destaca la influencia de la frecuencia de muestreo y el tamaño de ventana en las resoluciones en tiempo y frecuencia del sistema, lo cual afecta directamente la precisión de los resultados observados en los escalogramas obtenidos. Para efectos prácticos, la frecuencia de muestreo propuesta en las normas internacionales como la IEC 61850 para subestaciones digitales, de 4800 Hz, arroja resultados imprecisos en comparación con las metodologías convencionales.

Por otro lado, se encontró que valores del radio de convergencia del filtro IIR más cercanos a uno, pasos de frecuencia más cercanos a la resolución en frecuencia del sistema y una frecuencia de resonancia igual a la frecuencia de Nyquist ($F_s/2$) mejoran significativamente la precisión de los resultados obtenidos con la metodología propuesta. Además, se comprobó que la normalización de los Coeficientes Wavelet en función del factor de remuestreo y por la longitud del filtro (normalización tipo L1) es necesaria para simular el efecto de conservación de la energía de la Transformada Wavelet en forma integral cuando ésta se implementa de forma discretizada.

En cuanto al consumo de recursos computacionales evidenciados en el tiempo de cálculo de los Coeficientes Wavelet, se observó que tanto el paso en frecuencia como la frecuencia

de muestreo utilizadas afectan significativamente el tiempo de computación. Este resultado es consistente con la cantidad de escalas de frecuencia (determinadas por el paso en frecuencia) o el número de puntos en cada ventana (determinadas por la frecuencia de muestreo) dado que esto implica computar una mayor cantidad de remuestreos o manejar señales de longitud más larga.

El algoritmo de detección de eventos basado en ANN tuvo una eficiencia del 100% tanto al utilizar el conjunto de información hallado mediante la metodología propuesta como la tradicional, aunque se obtuvo una leve mejora en la precisión de clasificación de eventos utilizando el primero. Por otro lado, considerando el número de épocas y tamaño de conjunto de entrenamiento, se puede afirmar que el tiempo entrenamiento para la ANN alimentada utilizando la información obtenida mediante la metodología propuesta es menor en comparación con respecto a utilizar la información obtenida por la metodología tradicional; de esta forma es pertinente concluir que la metodología propuesta resulta en mejoras observables con respecto al uso de la metodología tradicional.

Si bien la precisión de los algoritmos de detección y clasificación fueron aceptables, es relevante destacar que se utilizó un conjunto de información relativamente reducido (aproximadamente 130 señales), el cual fue aumentado artificialmente utilizando múltiples mediciones por ventana. En la literatura analizada se ha observado que los problemas que utilizan Redes Neuronales Artificiales como medio de clasificación utilizan conjuntos de información mucho más grandes (del orden de miles de eventos), por lo que es posible que los resultados varíen significativamente al aumentar el conjunto real de información utilizada.

5.2 Recomendaciones

- En el desarrollo de la metodología propuesta para computar los Coeficientes Wavelet, no se tuvo en cuenta los efectos de incertidumbre introducidos por el cono de influencia de las representaciones en tiempo y frecuencia. Dado que se optó por centralizar las ventanas analizadas los efectos del cono de influencia deberían ser nulos o insignificantes, aunque para una comparación más rigurosa se sugiere evaluar el efecto de este sobre la metodología usada para una señal completa.

-
- Dado que el objetivo principal de esta disertación consistía en evaluar la validez de la metodología propuesta para computar los Coeficientes Wavelet, algunos aspectos secundarios se proponen como trabajo futuro para esta tesis. En particular, el sistema de distribución utilizado para la simulación de eventos (IEEE de 13 Nodos) no es necesariamente representativo de las dinámicas más complejas de los sistemas de distribución reales por lo que se propone evaluar la metodología validada con un conjunto de información simulado en sistemas más complejos o bien registros históricos de sistemas reales.
 - La localización de eventos no fue considerada dentro de la validación de resultados de esta disertación debido al alcance de esta; sin embargo, este aspecto es uno de los más relevantes para los Operadores de Red y podría ser decisivo para llevar la metodología evaluada a una prueba de concepto en tiempo real. En este sentido, se propone a modo de trabajo futuro utilizar alguna de las metodologías reportadas en la literatura, como el uso de Ondas Viajeras, en conjunto con los Coeficientes Wavelet resultantes para evaluar la validez de la metodología en el contexto de localización de fallas.
 - En cuanto al problema del análisis en tiempo real existen múltiples problemas que deben ser investigados para poder llevar la metodología propuesta a un modelo o prueba de concepto. En particular, se destacan las restricciones de la frecuencia de muestreo y los tiempos de computación reportados. El problema de la frecuencia de muestreo debe ser sopesado de forma que el uso de las frecuencias de muestreo propuestas por las normas internacionales conlleve a resultados aceptables según la resolución de frecuencia del sistema. Por otro lado, los tiempos de computación podrían disminuir significativamente utilizando computación paralela dado que solo se destinó una CPU al cálculo de los coeficientes, o bien utilizar herramientas de procesamiento gráfico.

A. Anexo: Características del modelo IEEE de 9 Barras

Los siguientes datos son traídos directamente del documento oficial del subcomité de análisis de sistemas de distribución de la IEEE [129].

Overhead Line Configuration Data:

Config.	Phasing	Phase	Neutral	Spacing
		ACSR	ACSR	ID
601	B A C N	556,500 26/7	4/0 6/1	500
602	C A B N	4/0 6/1	4/0 6/1	500
603	C B N	1/0	1/0	505
604	A C N	1/0	1/0	505
605	C N	1/0	1/0	510

Underground Line Configuration Data:

Config.	Phasing	Cable	Neutral	Space ID
606	A B C N	250,000 AA, CN	None	515
607	A N	1/0 AA, TS	1/0 Cu	520

Line Segment Data:

Node A	Node B	Length(ft.)	Config.
632	645	500	603
632	633	500	602
633	634	0	XFM-1
645	646	300	603
650	632	2000	601
684	652	800	607
632	671	2000	601

671	684	300	604
671	680	1000	601
671	692	0	Switch
684	611	300	605
692	675	500	606

Transformer Data:

	kVA	kV-high	kV-low	R - %	X - %
Substation:	5,000	115 - D	4.16 Gr. Y	1	8
XFM -1	500	4.16 – Gr.W	0.48 – Gr.W	1.1	2

Capacitor Data:

Node	Ph-A	Ph-B	Ph-C
	kVAr	kVAr	kVAr
675	200	200	200
611			100
Total	200	200	300

Spot Load Data:

Node	Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
	Model	kW	kVAr	kW	kVAr	kW	kVAr
634	Y-PQ	160	110	120	90	120	90
645	Y-PQ	0	0	170	125	0	0
646	D-Z	0	0	230	132	0	0
652	Y-Z	128	86	0	0	0	0
671	D-PQ	385	220	385	220	385	220
675	Y-PQ	485	190	68	60	290	212
692	D-I	0	0	0	0	170	151
611	Y-I	0	0	0	0	170	80
	TOTAL	1158	606	973	627	1135	753

Distributed Load Data:

Node A	Node B	Load	Ph-1	Ph-1	Ph-2	Ph-2	Ph-3	Ph-3
		Model	kW	kVAr	kW	kVAr	kW	kVAr
632	671	Y-PQ	17	10	66	38	117	68

Impedances

Configuration 601:

Z (R +jX) in ohms per mile

0.3465	1.0179	0.1560	0.5017	0.1580	0.4236
		0.3375	1.0478	0.1535	0.3849
				0.3414	1.0348

B in micro Siemens per mile

6.2998	-1.9958	-1.2595
	5.9597	-0.7417
		5.6386

Configuration 602:

Z (R +jX) in ohms per mile

0.7526	1.1814	0.1580	0.4236	0.1560	0.5017
		0.7475	1.1983	0.1535	0.3849
				0.7436	1.2112

B in micro Siemens per mile

5.6990	-1.0817	-1.6905
	5.1795	-0.6588
		5.4246

Configuration 603:

Z (R +jX) in ohms per mile

0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		1.3294	1.3471	0.2066	0.4591
				1.3238	1.3569

B in micro Siemens per mile

0.0000	0.0000	0.0000
--------	--------	--------

4.7097 -0.8999
 4.6658

Configuration 604:

Z (R +jX) in ohms per mile
 1.3238 1.3569 0.0000 0.0000 0.2066 0.4591
 0.0000 0.0000 0.0000 0.0000
 1.3294 1.3471

B in micro Siemens per mile
 4.6658 0.0000 -0.8999
 0.0000 0.0000
 4.7097

Configuration 605:

Z (R +jX) in ohms per mile
 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
 0.0000 0.0000 0.0000 0.0000
 1.3292 1.3475

B in micro Siemens per mile
 0.0000 0.0000 0.0000
 0.0000 0.0000
 4.5193

Configuration 606:

Z (R +jX) in ohms per mile
 0.7982 0.4463 0.3192 0.0328 0.2849 -0.0143
 0.7891 0.4041 0.3192 0.0328
 0.7982 0.4463

B in micro Siemens per mile
 96.8897 0.0000 0.0000
 96.8897 0.0000
 96.8897

Configuration 607:

```

      Z (R +jX) in ohms per mile
1.3425  0.5124   0.0000  0.0000   0.0000  0.0000
          0.0000  0.0000   0.0000  0.0000
                          0.0000  0.0000

      B in micro Siemens per mile
      88.9912   0.0000   0.0000
              0.0000   0.0000
                          0.0000

```

B. Anexo: Código en MATLAB para la validación de la metodología propuesta

```

Fs = 800*60; %Sampling frequency
Fc = 2400; %Peak filter resonance frequency
M = 800; %Window length (equivalent to 1 cycle at 60 hz)

% polos = r .* exp(j.*wo.*[-1 1]);
% den = poly(polos);
% num = abs(exp(j.*wo) - polos(1)) * abs(exp(j.*wo) - polos(2))

t1 = 0:1/Fs:0.04-(1/Fs);
% dT1 = t1(2)-t1(1);
y1 = sin(2*pi*60*t1);

for k = 1:length(t1)
    if t1(k) > 0.02
        y1(k) = y1(k) + sin(2*pi*1800*t1(k));
    end
end

%Base Voltage
BaseV = 4.16e3*sqrt(2);

```

```

%File name of the simulated faults
fname = 'N632_T1.json';
fid = fopen(fname);
raw = fread(fid,inf);
str = char(raw);
mystruct = jsondecode(str);
ts2 = mystruct.N632F.SVabc.A_1.SVmeas_1.time';
dummy = strcat('mystruct.N632F.SVabc.A_1.SVmeas_1.sv_A');

%Transform the signals to per unit
s2 = eval(dummy)/BaseV;

%Resample the signal to Fs
dT2 = ts2(2);

[a,b] = rat(Fs*dT2);
y2 = resample(s2,a,b);
t2 = (0:(length(y2)-1))*b/(a/dT2);

% Plot the signal
figure(1)
subplot(2,1,1)
plot(t1,y1)
title ('Señal Representativa')
xlabel ('Tiempo [s]')
ylabel ('Amplitud [p.u]')
subplot(2,1,2)
plot(t2,y2)
title ('Señal Simulada')
xlabel ('Tiempo [s]')
ylabel ('Amplitud [p.u]')

% Resonant IIR Filter
wo = 2*pi*Fc/Fs; %Resonant Frequency
r = 0.999; %Pole radius
% num1 = abs((1-r)*(1-r*exp(-2*1i*wo)));
a1 = -2*r*cos(wo);
a2 = r^2;
b0 = (1-r^2)/2;
b1=0;
b2=-b0;
num = [b0 b1 b2];
den = [1 a1 a2];

% 100% overlapping rectangular windows
W1 = buffer(y1,M,M-1);
W1 = W1(:,M:end);

```

```

W2 = buffer(y2,M,M-1);
W2 = W2(:,M:end);

% Taylor window definition
win = taylorwin(M);

% Taylor window application to rectangular windows
TW1 = zeros(size(W1));
TW2 = zeros(size(W2));

for j = 1:size(W1,2)
    for i = 1:size(W1,1)
        TW1(i,j) = W1(i,j)*win(i);
    end
end

for j = 1:size(W2,2)
    for i = 1:size(W2,1)
        TW2(i,j) = W2(i,j)*win(i);
    end
end

F = 30: 30 : 2400; %Frequency vector
[p, q] = rat(F./ Fc); %Resample coefficients
E1 = zeros(length(F),size(W1,2));
C1 = zeros(length(F),size(W1,2));
C1pq = zeros(length(F),size(W1,2));
C1L1 = zeros(length(F),size(W1,2));

%Signal resampling and filtering
tic
for i = 1:size(W1,2)
    for j = 1:length(F)
        rs1 = resample(TW1(:,i),p(j),q(j))*(q(j)/p(j));
        filt1 = filter(num,den,rs1);
        E1(j,i) = sum(filt1.^2)/sum(abs(filt1));
        C1(j,i) = sum(filt1.^2)*(p(j)/q(j));
        C1pq(j,i) = sum(filt1.^2);
        C1L1(j,i) = sum(filt1.^2)/sum(abs(filt1));
    end
end
toc

E2 = zeros(length(F),size(W2,2));
C2 = zeros(length(F),size(W2,2));
C2pq = zeros(length(F),size(W2,2));
C2L1 = zeros(length(F),size(W2,2));

tic
for i = 1:size(W2,2)
    for j = 1:length(F)

```

```

rs2 = resample(TW2(:,i),p(j),q(j))*(q(j)/p(j));
filt2 = filter(num,den,rs2);
E2(j,i) = sum(filt2.^2)/sum(abs(filt2));
C2(j,i) = sum(filt2.^2)*(p(j)/q(j));
C2pq(j,i) = sum(filt2.^2);
C2L1(j,i) = sum(filt2.^2)/sum(abs(filt2));
end
end
toc

% %New time vector considering overlapping windows
dT1 = t1(end)/length(W1);
tw1 = 0 : dT1 : t1(end)-dT1;

dT2 = t2(end)/length(W2);
tw2 = 0 : dT2 : t2(end)-dT2;

% %Scalogram with IIR Filter Bank

figure(2)
subplot(2,1,1)
imagesc(tw1,F,E1);
colormap(parula)
colorbar
axis xy
set(gca, 'YScale', 'log')
xlabel('Tiempo [s]')
ylabel('Escala - Frecuencia [Hz]')
title ('Escalograma para señal representativa')
subplot(2,1,2)
imagesc(tw2,F,E2);
colormap(parula)
colorbar
axis xy
set(gca, 'YScale', 'log')
xlabel('Tiempo [s]')
ylabel('Escala - Frecuencia [Hz]')
title ('Escalograma para señal simulada')
sgtitle('Escalogramas para F_c = 2400 Hz')

%3D Surfplot
figure(5)
h = subplot(1,1,1);
z = surf((0:length(E)-1)./Fs,F,E);
xlabel('t_{[seg]}'); ylabel('F_{[Hz]}');
set(h, 'YScale', 'log');
set(z, 'LineStyle', 'none')

```

Bibliografía

- [1] C. Zapata, *Confiabilidad de Sistemas Eléctricos de Potencia*. Pereira: Universidad Tecnológica de Pereira, 2011. doi: 10.2307/j.ctt1zk0m77.5.
- [2] P. Kundur *et al.*, "Definition and classification of power system stability," *IEEE Transactions on Power Systems*, 2004, doi: 10.1109/TPWRS.2004.825981.
- [3] M. Ventosa, P. Linares, and I. J. Pérez-Arriaga, "Power System Economics," *Power Systems*, vol. 61, no. October 1982, pp. 47–123, 2013, doi: 10.1007/978-1-4471-5034-3_2.
- [4] S. Silva, P. Costa, M. Santana, and D. Leite, "Evolving neuro-fuzzy network for real-time high impedance fault detection and classification," *Neural Computing and Applications*, vol. 32, no. 12, pp. 7597–7610, 2020, doi: 10.1007/s00521-018-3789-2.
- [5] P. Jafarian and M. Sanaye-Pasand, "A traveling-wave-based protection technique using wavelet/pca analysis," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 588–599, 2010, doi: 10.1109/TPWRD.2009.2037819.
- [6] D. Spoor and J. G. Zhu, "Improved single-ended traveling-wave fault- location algorithm based on experience with conventional substation transducers," *IEEE Transactions on Power Delivery*, 2006, doi: 10.1109/TPWRD.2006.878091.
- [7] M. Parsi, P. Crossley, P. L. Dragotti, and D. Cole, "Wavelet based fault location on power transmission lines using real-world travelling wave data," *Electric Power Systems Research*, vol. 186, no. February, p. 106261, 2020, doi: 10.1016/j.epsr.2020.106261.
- [8] L. Y. Bewley, "Traveling Waves on Transmission Systems," *Transactions of the American Institute of Electrical Engineers*, 1931, doi: 10.1109/T-AIEE.1931.5055827.
- [9] M. Pourahmadi-Nakhli and A. A. Safavi, "Path characteristic frequency-based fault locating in radial distribution systems using wavelets and neural networks," *IEEE Transactions on Power Delivery*, 2011, doi: 10.1109/TPWRD.2010.2050218.

- [10] S. M. Korobeynikov, N. Y. Ilyushov, Y. A. Lavrov, S. S. Shevchenko, and V. A. Loman, "High-Frequency Transients Suppression at Substation," 2019. doi: 10.1109/ICHVE.2018.8641972.
- [11] Y. Gu and M. H. J. Bollen, "Time-frequency and time-scale domain analysis of voltage disturbances," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1279–1284, 2000, doi: 10.1109/61.891515.
- [12] F. H. Magnago and A. Abur, "A new fault location technique for radial distribution systems based on high frequency signals," 1999. doi: 10.1109/PSS.1999.784386.
- [13] T. M. Lai, L. A. Snider, E. Lo, and D. Sutanto, "High-impedance fault detection using discrete wavelet transform and frequency range and RMS conversion," *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 397–407, 2005, doi: 10.1109/TPWRD.2004.837836.
- [14] M. Goudarzi, B. Vahidi, R. A. Naghizadeh, and S. H. Hosseini, "Improved fault location algorithm for radial distribution systems with discrete and continuous wavelet analysis," *International Journal of Electrical Power and Energy Systems*, vol. 67, no. May, pp. 423–430, 2015, doi: 10.1016/j.ijepes.2014.12.014.
- [15] S. Khokhar, A. A. B. Mohd Zin, A. S. B. Mokhtar, and M. Pesaran, "A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances," *Renewable and Sustainable Energy Reviews*. 2015. doi: 10.1016/j.rser.2015.07.068.
- [16] A. M. Gaouda, M. M. A. Salama, M. R. Sultan, and A. Y. Chikhani, "Power quality detection and classification using wavelet-multiresolution signal decomposition," *IEEE Transactions on Power Delivery*, vol. 14, no. 4, pp. 1469–1476, 1999, doi: 10.1109/61.796242.
- [17] A. R. Sedighi, M. R. Haghifam, O. P. Malik, and M. H. Ghassemian, "High impedance fault detection based on wavelet transform and statistical pattern recognition," *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2414–2421, 2005, doi: 10.1109/TPWRD.2005.852367.
- [18] Z. Moravej, S. H. Mortazavi, and S. M. Shahrtash, "DT-CWT based event feature extraction for high impedance faults detection in distribution system," *International Transactions on Electrical Energy Systems*, 2015, doi: 10.1002/etep.2035.
- [19] C. L. Tu, W. L. Hwang, and J. Ho, "Analysis of singularities from modulus maxima of complex wavelets," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1049–1062, 2005, doi: 10.1109/TIT.2004.842706.
- [20] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines based on wavelet transform and ANN," *IEEE Transactions on*

- Power Delivery*, vol. 21, no. 4, pp. 2058–2063, 2006, doi: 10.1109/TPWRD.2006.876659.
- [21] L. O. Barthold and G. K. Carter, “Digital Traveling-Wave Solutions: I—Single-Phase Equivalents,” *Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, 1961, doi: 10.1109/AIEEPAS.1961.4501145.
- [22] F. Namdari and M. Salehi, “Fault classification and location in transmission lines using traveling waves modal components and continuous wavelet transform (CWT),” *Journal of Electrical Systems*, vol. 12, no. 2, pp. 373–386, 2016.
- [23] M. Vrhel, C. Lee, and M. Unser, “THE CONTINUOUS WAVELET TRANSFORM: A TOOL FOR SIGNAL INVESTIGATION AND UNDERSTANDING,” *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 1995.
- [24] L. U. Iurinic, A. S. Bretas, and E. S. Guimaraes, “Continuous-wavelet transform fault location algorithm inferred from faulty signal,” 2012. doi: 10.1109/PESGM.2012.6344853.
- [25] A. Borghetti, M. Bosetti, C. A. Nucci, M. Paolone, and A. Abur, “Integrated use of time-frequency wavelet decompositions for fault location in distribution networks: Theory and experimental validation,” *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 3139–3146, 2010, doi: 10.1109/TPWRD.2010.2046655.
- [26] H. w. Dommel and J. M. Michels, “HIGH SPEED RELAYING USING TRAVELING WAVE TRANSIENT ANALYSIS.,” 1978.
- [27] M. T. Hagh, K. Razi, and H. Taghizadeh, “Fault classification and location of power transmission lines using artificial neural network,” 2007.
- [28] A. Abdollahi and S. Seyedtabaai, “Comparison of fourier & wavelet transform methods for transmission line fault classification,” *PEOCO 2010 - 4th International Power Engineering and Optimization Conference, Program and Abstracts*, no. June, pp. 579–584, 2010, doi: 10.1109/PEOCO.2010.5559232.
- [29] L. J. Collazo and E. O’Neill-Carrillo, “Comparison of windowed Fourier transform and dynamic phasors for power quality analysis,” *2009 IEEE/PES Power Systems Conference and Exposition, PSCE 2009*, pp. 1–5, 2009, doi: 10.1109/PSCE.2009.4840147.
- [30] D. Donnelly, “The Fast Fourier and Hilbert-Huang Transforms: A Comparison,” in *IMACS Multiconference on “Computational Engineering in Systems Applications” (CESA)*, 2006, vol. 12, no. 3, pp. 84–88.

- [31] S. Vandepu, *Heart rate variability: linear and nonlinear analysis with applications in human physiology*. 2010.
- [32] J. B. Allen and L. R. Rabiner, "A Unified Approach to Short-Time Fourier Analysis and Synthesis," *Proceedings of the IEEE*, 1977, doi: 10.1109/PROC.1977.10770.
- [33] R. Subhashree, C. S. Preethi, and P. Supriya, "Fault distance identification in transmission line using STFT algorithm," 2016. doi: 10.1109/ICCCI.2016.7480036.
- [34] R. Bhattacharjee and A. De, "A robust STFT based algorithm for detection of symmetrical fault during power swing," 2017. doi: 10.1109/POWERI.2016.8077364.
- [35] S. Mallat, *A Wavelet Tour of Signal Processing The Sparse Way*. 2009.
- [36] J. O. Strömberg, "A modified franklin system and higher-order spline systems on R_n as unconditional bases for hardy spaces," in *Fundamental Papers in Wavelet Theory*, 2009. doi: 10.1515/9781400827268.197.
- [37] J. D. Kropotov, *Functional Neuromarkers for Psychiatry: Applications for Diagnosis and Treatment*. 2016. doi: 10.1016/C2012-0-07144-X.
- [38] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape," *SIAM Journal on Mathematical Analysis*, vol. 15, no. 4, 1984.
- [39] Y. Meyer, *Wavelets and Operators*. 1993. doi: 10.1017/cbo9780511623820.
- [40] A. Cohen, I. Daubechies, and J. -C Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, 1992, doi: 10.1002/cpa.3160450502.
- [41] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, 1988, doi: 10.1002/cpa.3160410705.
- [42] I. Daubechies, "2. The Continuous Wavelet Transform," in *Ten Lectures on Wavelets*, 1992. doi: 10.1137/1.9781611970104.ch2.
- [43] I. Daubechies, *Ten Lectures on Wavelets*. 1992. doi: 10.1137/1.9781611970104.
- [44] P. Qi, S. Jovanovic, J. Lezama, and P. Schweitzer, "Discrete wavelet transform optimal parameters estimation for arc fault detection in low-voltage residential power networks," *Electric Power Systems Research*, vol. 143, no. June 2018, pp. 130–139, 2017, doi: 10.1016/j.epsr.2016.10.008.

- [45] D. Spoor and J. G. Zhu, "Improved single-ended traveling-wave fault- location algorithm based on experience with conventional substation transducers," *IEEE Transactions on Power Delivery*, 2006, doi: 10.1109/TPWRD.2006.878091.
- [46] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Transactions on Power Delivery*, 1998, doi: 10.1109/61.714808.
- [47] F. B. Costa, B. A. Souza, N. S. D. Brito, J. A. C. B. Silva, and W. C. Santos, "Real-time detection of transients induced by high-impedance faults based on the boundary wavelet transform," *IEEE Transactions on Industry Applications*, 2015, doi: 10.1109/TIA.2015.2434993.
- [48] M. F. Guo, X. D. Zeng, D. Y. Chen, and N. C. Yang, "Deep-Learning-Based Earth Fault Detection Using Continuous Wavelet Transform and Convolutional Neural Network in Resonant Grounding Distribution Systems," *IEEE Sensors Journal*, 2018, doi: 10.1109/JSEN.2017.2776238.
- [49] A. Borghetti, S. Corsi, C. A. Nucci, M. Paolone, L. Peretto, and R. Tinarelli, "On the use of continuous-wavelet transform for fault location in distribution power systems," *International Journal of Electrical Power and Energy Systems*, 2006, doi: 10.1016/j.ijepes.2006.03.001.
- [50] A. Borghetti, M. Bosetti, M. Di Silvestro, C. A. Nucci, and M. Paolone, "Continuous-wavelet transform for fault location in distribution power networks: Definition of mother wavelets inferred from fault originated transients," *IEEE Transactions on Power Systems*, 2008, doi: 10.1109/TPWRS.2008.919249.
- [51] K. Nanayakkara, A. D. Rajapakse, and R. Wachal, "Fault Location in Extra Long HVDC Transmission Lines using Continuous Wavelet Transform," 2011.
- [52] MATHWORKS, "Continuous Wavelet Analysis," <https://la.mathworks.com/help/wavelet/gs/continuous-wavelet-analysis.html>, 2021.
- [53] J. Lilly, "A data analysis package for Matlab," 2016.
- [54] A. R. Adly, S. H. E. Abdel Aleem, M. A. Elsadd, and Z. M. Ali, "Wavelet packet transform applied to a series-compensated line: A novel scheme for fault identification," *Measurement: Journal of the International Measurement Confederation*, vol. 151, p. 107156, 2020, doi: 10.1016/j.measurement.2019.107156.
- [55] G. Jing and L. Ru, "A new wavelet packet method of single-phase earth fault line selection in distribution network based on the maximum difference comparison," 2009. doi: 10.1109/ICEMS.2009.5382649.

- [56] E. Y. Hamid and Z. I. Kawasaki, "Wavelet-based data compression of power system disturbances using the minimum description length criterion," *IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 460–466, 2002, doi: 10.1109/61.997918.
- [57] L. Jin, M. Jiang, and G. Yang, "Fault analysis of microgrid and adaptive distance protection based on complex wavelet transform," 2014. doi: 10.1109/PEAC.2014.7037882.
- [58] Z. Moravej, S. H. Mortazavi, and S. M. Shahrtash, "DT-CWT based event feature extraction for high impedance faults detection in distribution system," *International Transactions on Electrical Energy Systems*, 2015, doi: 10.1002/etep.2035.
- [59] J. G. Proakis and D. G. Monolakis, *Digital signal processing: principles, algorithms, and applications*. 1996.
- [60] J. Klomjit, A. Ngaopitakkul, and B. Sreewirote, "Comparison of mother wavelet for classification fault on hybrid transmission line systems," 2017. doi: 10.1109/ICAwST.2017.8256514.
- [61] C. Pothisarn, J. Klomjit, A. Ngaopitakkul, C. Jettanasen, D. A. Asfani, and I. M. Y. Negara, "Comparison of various mother wavelets for fault classification in electrical systems," *Applied Sciences (Switzerland)*, vol. 10, no. 4, pp. 1–16, 2020, doi: 10.3390/app10041203.
- [62] W. Fluty and Y. Liao, "Electric Transmission Fault Location Techniques Using Traveling Wave Method and Discrete Wavelet Transform," *Clemson University Power Systems Conference, PSC 2020*, 2020, doi: 10.1109/PSC50246.2020.9131271.
- [63] B. Rathore and A. G. Shaik, "Wavelet-Alienation Based Transmission Line Protection Scheme," pp. 1–15.
- [64] H. R. Karimi, W. Pawlus, and K. G. Robbersmyr, "Signal reconstruction, modeling and simulation of a vehicle full-scale crash test based on Morlet wavelets," *Neurocomputing*, 2012, doi: 10.1016/j.neucom.2012.04.010.
- [65] A. Borghetti, M. Bosetti, M. Di Silvestro, C. A. Nucci, and M. Paolone, "Continuous-wavelet transform for fault location in distribution power networks: Definition of mother wavelets inferred from fault originated transients," *IEEE Transactions on Power Systems*, 2008, doi: 10.1109/TPWRS.2008.919249.
- [66] A. Calderón, "Intermediate spaces and interpolation, the complex method," *Studia Mathematica*, 1964, doi: 10.4064/sm-24-2-113-190.

- [67] P. E. Argyropoulos and H. Lev-Ari, "Wavelet Customization for Improved Fault-Location Quality in Power Networks," *IEEE Transactions on Power Delivery*, 2015, doi: 10.1109/TPWRD.2015.2429590.
- [68] U. B. Parikh, B. Das, and R. P. Maheshwari, "Combined wavelet-SVM technique for fault zone detection in a series compensated transmission line," *IEEE Transactions on Power Delivery*, vol. 23, no. 4, pp. 1789–1794, 2008, doi: 10.1109/TPWRD.2008.919395.
- [69] R. Kumar, B. Singh, D. T. Shahani, A. Chandra, and K. Al-Haddad, "Recognition of Power-Quality Disturbances Using S-Transform-Based ANN Classifier and Rule-Based Decision Tree," *IEEE Transactions on Industry Applications*, vol. 51, no. 2, pp. 1249–1258, 2015, doi: 10.1109/TIA.2014.2356639.
- [70] N. Zhang and M. Kezunovic, "Transmission line boundary protection using wavelet transform and neural network," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 859–869, 2007, doi: 10.1109/TPWRD.2007.893596.
- [71] N. Tailor, S. Joshi, and O. P. Mahela, "Transmission line protection schemes based on wigner distribution function and discrete wavelet transform," *PIICON 2020 - 9th IEEE Power India International Conference*, 2020, doi: 10.1109/PIICON49524.2020.9113011.
- [72] A. K. Pradhan, A. Routray, S. Pati, and D. K. Pradhan, "Wavelet fuzzy combined approach for fault classification of a series-compensated transmission line," *IEEE Transactions on Power Delivery*, 2004, doi: 10.1109/TPWRD.2003.822535.
- [73] Z. L. Gaing, "Wavelet-based neural network for power disturbance recognition and classification," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1560–1568, 2004, doi: 10.1109/TPWRD.2004.835281.
- [74] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2295–2304, 2016, doi: 10.1109/TSG.2015.2487501.
- [75] M. Dashtdar, M. Esmailbeig, M. Najafi, and M. E. N. Bushehri, "Fault Location in the Transmission Network Using Artificial Neural Network," *Automatic Control and Computer Sciences*, vol. 54, no. 1, pp. 39–51, 2020, doi: 10.3103/S0146411620010022.
- [76] A. A. P. Biscoaro, R. A. F. Pereira, M. Kezunovic, and J. R. S. Mantovani, "Integrated Fault Location and Power-Quality Analysis in Electric Power Distribution Systems," *IEEE Transactions on Power Delivery*, 2016, doi: 10.1109/TPWRD.2015.2464098.

- [77] A. H. Osman and O. P. Malik, "Transmission line distance protection based on wavelet transform," *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 515–523, 2004, doi: 10.1109/TPWRD.2003.822531.
- [78] A. Bermúdez, E. Spinelli, and C. Muravch, "Detección de eventos en señales de EEG mediante Entropía," 2011. Accessed: Oct. 20, 2021. [Online]. Available: http://www.sabi2011.fi.mdp.edu.ar/proceedings/SABI/Pdf/SABI2011_75.pdf
- [79] H. Zheng-you, C. Xiaoqing, and L. Guoming, "Wavelet Entropy Measure Definition and Its Application for Transmission Line Fault Detection and Identification; (Part I: Definition and Methodology)," 2007. doi: 10.1109/icpst.2006.321939.
- [80] Z. He, L. Fu, S. Lin, and Z. Bo, "Fault detection and classification in EHV transmission line based on wavelet singular entropy," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2156–2163, 2010, doi: 10.1109/TPWRD.2010.2042624.
- [81] S. Khokhar, A. A. Mohd Zin, A. P. Memon, and A. S. Mokhtar, "A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network," *Measurement: Journal of the International Measurement Confederation*, vol. 95, pp. 246–259, 2017, doi: 10.1016/j.measurement.2016.10.013.
- [82] H. H. Chang and N. V. Linh, "Statistical feature extraction for fault locations in nonintrusive fault detection of low voltage distribution systems," *Energies*, vol. 10, no. 5, pp. 1–19, 2017, doi: 10.3390/en10050611.
- [83] H. Lala and S. Karmakar, "Continuous wavelet transform and artificial neural network based fault diagnosis in 52 bus hybrid distributed generation system," *4th Students Conference on Engineering and Systems, SCES 2015*, 2016, doi: 10.1109/SCES.2015.7506463.
- [84] Y. Y. Hong and M. T. A. M. Cabatac, "Fault Detection, Classification, and Location by Static Switch in Microgrids Using Wavelet Transform and Taguchi-Based Artificial Neural Network," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2725–2735, 2020, doi: 10.1109/JSYST.2019.2925594.
- [85] B. Rathore and A. G. Shaik, "Wavelet-alienation based transmission line protection scheme," *IET Generation, Transmission and Distribution*, vol. 11, no. 4, pp. 995–1003, Mar. 2017, doi: 10.1049/iet-gtd.2016.1022.
- [86] P. E. Argyropoulos and H. Lev-Ari, "Wavelet Customization for Improved Fault-Location Quality in Power Networks," *IEEE Transactions on Power Delivery*, 2015, doi: 10.1109/TPWRD.2015.2429590.

- [87] F. Namdari and M. Salehi, "Fault classification and location in transmission lines using traveling waves modal components and continuous wavelet transform (CWT)," *Journal of Electrical Systems*, vol. 12, no. 2, pp. 373–386, 2016.
- [88] G. Jing and L. Ru, "A new wavelet packet method of single-phase earth fault line selection in distribution network based on the maximum difference comparison," 2009. doi: 10.1109/ICEMS.2009.5382649.
- [89] M. Hirn, "Lecture 16: Wavelet Modulus Maxima," 2020. Accessed: Sep. 14, 2021. [Online]. Available: https://matthewhirn.files.wordpress.com/2020/03/math994_spring2020_lecture16-1.pdf
- [90] D. Guillen, M. R. A. Paternina, A. Zamora, J. M. Ramirez, and G. Idarraga, "Detection and classification of faults in transmission lines using the maximum wavelet singular value and Euclidean norm," *IET Generation, Transmission and Distribution*, 2015, doi: 10.1049/iet-gtd.2014.1064.
- [91] C. Pothisarn, J. Klomjit, A. Ngaopitakkul, C. Jettanasen, D. A. Asfani, and I. M. Y. Negara, "Comparison of various mother wavelets for fault classification in electrical systems," *Applied Sciences (Switzerland)*, 2020, doi: 10.3390/app10041203.
- [92] P. Lertwanitrot and A. Ngaopitakkul, "Discriminating between Capacitor Bank Faults and External Faults for an Unbalanced Current Protection Relay Using DWT," *IEEE Access*, vol. 8, pp. 180022–180044, 2020, doi: 10.1109/ACCESS.2020.3026744.
- [93] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," *IEEE Transactions on Smart Grid*, 2016, doi: 10.1109/TSG.2015.2487501.
- [94] L. Breiman, "Random Forests," Berkeley, CA: University of California, 2001.
- [95] H. Ampadu, "Random Forests Understanding," <https://ai-pool.com/a/s/random-forests-understanding>, May 01, 2021.
- [96] M. Awad and R. Khanna, "Support Vector Machines for Classification," in *Efficient Learning Machines*, 2015. doi: 10.1007/978-1-4302-5990-9_3.
- [97] A. A. P. Biscaro, R. A. F. Pereira, M. Kezunovic, and J. R. S. Mantovani, "Integrated Fault Location and Power-Quality Analysis in Electric Power Distribution Systems," *IEEE Transactions on Power Delivery*, vol. 31, no. 2, pp. 428–436, 2016, doi: 10.1109/TPWRD.2015.2464098.
- [98] S. Kaewarsa, K. Attakitmongcol, and T. Kulworawanichpong, "Recognition of power quality events by using multiwavelet-based neural network," *Proceedings -*

- 6th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2007; 1st IEEE/ACIS International Workshop on e-Activity, IWEA 2007*, no. Icis, pp. 993–998, 2007, doi: 10.1109/ICIS.2007.153.
- [99] M. F. Guo, X. D. Zeng, D. Y. Chen, and N. C. Yang, “Deep-Learning-Based Earth Fault Detection Using Continuous Wavelet Transform and Convolutional Neural Network in Resonant Grounding Distribution Systems,” *IEEE Sensors Journal*, 2018, doi: 10.1109/JSEN.2017.2776238.
- [100] O. A. S. Yousef, “Combined fuzzy-logic wavelet-based fault classification technique for power system relaying,” *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 582–589, 2004, doi: 10.1109/TPWRD.2004.826386.
- [101] J. Jan, *Digital Signal Filtering, Analysis and Restoration*. 2000. doi: 10.1049/pbte044e.
- [102] R. B. Blackman and J. W. Tukey, “The Measurement of Power Spectra from the Point of View of Communications Engineering — Part II,” *Bell System Technical Journal*, vol. 37, no. 2, 1958, doi: 10.1002/j.1538-7305.1958.tb01530.x.
- [103] F. J. Harris, “On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform,” *Proceedings of the IEEE*, vol. 66, no. 1, 1978, doi: 10.1109/PROC.1978.10837.
- [104] Chegg, “Problem 8 (Window Functions),” <https://www.chegg.com/homework-help/questions-and-answers/problem-8-window-functions-figure-71-shows-three-window-functions-corresponding-amplitude--q51974895>.
- [105] S. W. Smith, *Properties of Convolution: The Scientist and Engineer’s Guide to Digital Signal Processing*. 1999.
- [106] S. Bon, M. Vigny, and J. Massoulie, “Order A Journal On The Theory Of Ordered Sets And Its Applications,” *Biochemistry*, vol. 76, no. 6, 1979.
- [107] W. K. Chen, *Passive, active, and digital filters*. 2009. doi: 10.1201/9781315219141.
- [108] R. B. Randall, J. Antoni, and P. Borghesani, “Applied Digital Signal Processing,” in *Handbook of Experimental Structural Dynamics*, 2020. doi: 10.1007/978-1-4939-6503-8_6-1.
- [109] R. Holmes, “Applied Underwater Acoustics,” *Physics Bulletin*, vol. 18, no. 1, 1967, doi: 10.1088/0031-9112/18/1/021.
- [110] A. Borghetti, M. Bosetti, M. di Silvestro, C. A. Nucci, and M. Paolone, “Continuous-wavelet transform for fault location in distribution power networks: Definition of

- mother wavelets inferred from fault originated transients,” *IEEE Transactions on Power Systems*, 2008, doi: 10.1109/TPWRS.2008.919249.
- [111] T. Carsey and J. Harden, *Monte Carlo Simulation and Resampling Methods for Social Science*. 2014. doi: 10.4135/9781483319605.
- [112] MATHWORKS, “FILTER REFERENCE,”
<https://la.mathworks.com/help/matlab/ref/filter.html>.
- [113] K. P. Balanda and H. L. Macgillivray, “Kurtosis: A critical review,” *American Statistician*, vol. 42, no. 2, 1988, doi: 10.1080/00031305.1988.10475539.
- [114] MATHWORKS, “KURTOSIS REFERENCE,”
<https://la.mathworks.com/help/stats/kurtosis.html>.
- [115] ATRIA Innovation, “Qué son las redes neuronales y sus funciones,”
<https://www.atriainnovation.com/que-son-las-redes-neuronales-y-sus-funciones/>,
Oct. 19, 2019.
- [116] B. Yegnanarayana, “Artificial neural networks for pattern recognition,” *Sadhana*, vol. 19, no. 2, 1994, doi: 10.1007/BF02811896.
- [117] K. Mehrotra, C. Mohan, and S. Ranka, *Elements of Artificial Neural Networks*. 2019. doi: 10.7551/mitpress/2687.001.0001.
- [118] I. N. da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, *Artificial neural networks: A practical course*. 2016. doi: 10.1007/978-3-319-43162-8.
- [119] P. J. Braspenning, F. Thuijsman, and A. Weijters, *Artificial neural networks: an introduction to ANN theory and practice*, vol. 931. 1995.
- [120] PICO, “<https://www.pico.net/kb/the-role-of-bias-in-neural-networks/>,”
<https://www.pico.net/kb/the-role-of-bias-in-neural-networks/>.
- [121] R. Tadeusiewicz, “Neural networks: A comprehensive foundation,” *Control Engineering Practice*, vol. 3, no. 5, 1995, doi: 10.1016/0967-0661(95)90080-2.
- [122] I. S. Korovin and M. v. Khisamutdinov, “Hybrid method of dynamograms wavelet analysis for oil-production equipment state identification,” in *Advanced Materials Research*, 2014, vol. 909. doi: 10.4028/www.scientific.net/AMR.909.252.
- [123] J. Villanueva, “Redes neuronales desde cero (I) – Introducción,”
<https://www.iartificial.net/redes-neuronales-desde-cero-i-introduccion/>.
- [124] IEEE, “IEEE PES Test Feeder,” <https://cmte.ieee.org/pes-testfeeders/resources/>.

- [125] Distribution System Analysis Subcommittee IEEE, "IEEE 13 Node Test Feeder."
- [126] D. A. Sánchez Muñoz, "Estrategia de detección y localización de fallas para el esquema de protección distancia en redes con alta penetración de energía renovable de tipo eólica," Medellín, 2020.