

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA ELÉCTRICA



**ANÁLISIS ESTADÍSTICO DE LA CALIDAD DE
FRECUENCIA EN LA RED ELÉCTRICA DE EL SALVADOR**

PRESENTADO POR:

JOSUE GEOVANNI LÚE GONZÁLEZ

CARLOS ANTONIO ARCE ARAUZ

PARA OPTAR POR EL TÍTULO DE:

INGENIERO ELECTRICISTA

CUIDAD UNIVERSITARIA, SEPTIEMBRE DE 2024

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSC. JUAN ROSA QUINTANILLA

SECRETARIO GENERAL:

LIC. PEDRO ROSALÍO ESCOBAR CASTANEDA

FACULTAD DE INGENIERÍA Y ARQUITECTURA

DECANO:

ING. LUIS SALVADOR BARRERA MANCÍA

SECRETARIO:

ARQ. RAÚL ALEXANDER FABIÁN ORELLANA

ESCUELA DE INGENIERÍA ELÉCTRICA

DIRECTOR INTERINO:

ING. WERNER DAVID MELÉNDEZ VALLE

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA ELÉCTRICA

TRABAJO DE GRADUACION PREVIO A LA OPCION AL GRADO DE:
INGENIERO ELECTRICISTA

TITULO:

**ANÁLISIS ESTADÍSTICO DE LA CALIDAD DE
FRECUENCIA EN LA RED ELÉCTRICA DE EL SALVADOR**

PRESENTADO POR:

JOSUÉ GEOVANNI LÚE GONZÁLEZ

CARLOS ANTONIO ARCE ARAUZ

TRABAJO DE GRADUACION APROBADO POR:

DOCENTE ASESOR:

PhD. CARLOS OSMIN POCASANGRE JÍMENEZ

SAN SALVADOR, SEPTIEMBRE DE 2024

TRABAJO DE GRADUACION APROBADO POR:

DOCENTE ASESOR:

PhD. CARLOS OSMIN POCASANGRE JÍMENEZ

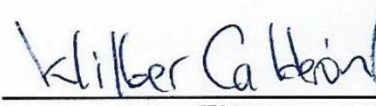
NOTA Y DEFENSA FINAL


En esta fecha, jueves 29 de agosto de 2024, en la Sala de Lectura de la Escuela de Ingeniería Eléctrica, a las 10:00 a.m. horas, en presencia de las siguientes autoridades de la Escuela de Ingeniería Eléctrica de la Universidad de El Salvador:

1. Ing. Werner David Meléndez Valle
Director Interino


Firma

2. MSc. José Wilber Calderón Urrutia
Secretario


Firma



Y, con el Honorable Jurado de Evaluación integrado por las personas siguientes:

- DR. CARLOS OSMIN POCASANGRE JIMENEZ
(Docente Asesor)


Firma

- MSC. JORGE ALBERTO ZETINO CHICAS


Firma

- ING. JOSÉ ROBERTO RAMOS LÓPEZ


Firma

Se efectuó la defensa final reglamentaria del Trabajo de Graduación:

ANÁLISIS ESTADÍSTICO DE LA CALIDAD DE FRECUENCIA EN LA RED ELÉCTRICA DE EL SALVADOR

A cargo de los Bachilleres:

ARCE ARAUZ CARLOS ANTONIO

LÚE GONZÁLEZ JOSUÉ GEOVANNI

Habiendo obtenido en el presente Trabajo una nota promedio de la defensa final: 9.1

(Nueve punto uno)

AGRADECIMIENTOS

Quiero expresar mi más profundo agradecimiento a Dios, por haberme dado la fortaleza y la oportunidad de llegar hasta este momento tan importante en mi vida. Es una inmensa alegría haber culminado este proyecto, y sé que no habría sido posible sin su guía.

A mi madre, Virginia Arauz, le debo todo mi agradecimiento. Su apoyo incondicional en cada etapa de mi vida ha sido el pilar que me ha sostenido, y su amor y dedicación me han dado la fuerza para seguir adelante, incluso en los momentos más difíciles.

Mi gratitud eterna también va para mi abuela, Angélica Aguilar, quien desde mi infancia me enseñó el valor del bien y me guio con sabiduría en el camino de la vida. Sus enseñanzas han sido una brújula que me ha orientado a lo largo de este recorrido.

A Don Carlos Pineda, le agradezco profundamente por haber sido un gran apoyo en esta etapa de mi vida. Su ayuda y su presencia han sido fundamentales para que pudiera superar los desafíos que encontré en el camino.

Finalmente, quiero agradecer a mis tíos: Lilian Arauz, Roxana Arauz y Ever Arauz, quienes, con su apoyo moral y sus palabras de aliento, me sostuvieron cuando las dudas me hicieron pensar en abandonar. Gracias a todos por estar a mi lado y ayudarme a llegar hasta aquí.

¡Per Aspera Ad Astra!

Antonio Arauz

AGRADECIMIENTOS

En primer lugar, agradezco a Dios por permitirme llegar hasta este punto de mi vida y superar los retos que he enfrentado en el camino. También agradezco a mi madre, Dora González, quien ha estado a mi lado brindándome su apoyo incondicional y consejo, junto a mi padrastro, Dionicio Lúe. A mis abuelos maternos, Concepción Palacios y Félix González, les agradezco profundamente por su compañía y respaldo. Agradezco también por el apoyo brindado por mi hermano, Raúl González, durante la carrera. También agradezco el apoyo recibido por los miembros de mi familia materna y paterna primos.

Agradezco mi amiga Grecia García, cuya presencia a lo largo de mi vida universitaria, especialmente en pandemia, ha sido invaluable, además de darme apoyo para la elaboración de este trabajo de graduación. Agradezco a todos mis compañeros y amigos que he conocido en la EIE-UES y en la FIA-UES, por todo el apoyo brindado durante todos estos años, a pesar de la modalidad extraordinaria en la que hemos convivido. Estoy agradecido enormemente con todos aquellos que se quedaron conmigo hasta el final, dándome apoyo durante todo este proceso.

Agradezco al personal docente y administrativo de la Escuela de Ingeniería Eléctrica por el apoyo brindado durante este proyecto. Al PhD. Carlos Pocasangre, por asesorarnos durante el proyecto, facilitar contactos y cooperar para la elaboración de dos artículos científicos que han sido enviados a los congresos CONESCAPAN Panamá 2024 y CONCAPAN Costa Rica 2024. También agradecemos a las autoridades de la Facultad de Ingeniería y Arquitectura, por brindar financiamiento para la asistencia al congreso CONESCAPAN 2024.

Josué Lúe

AGRADECIMIENTOS GENERALES

Agradecemos a las instituciones externas y sus miembros, con los cuales se logró la realización de este trabajo de graduación:

- A la Universidad Nacional Autónoma de México, por almacenar brindar los datos de frecuencia recolectados en El Salvador y asesorar sobre la configuración de equipos.
- A la Unidad de Transacciones, por interesarse en el proyecto y brindar las mediciones de frecuencia del sistema SCADA, los cuales han permitido complementar este trabajo.
- Y a la Universidad de Tennessee, Knoxville, quienes gestionan el proyecto FNET/GridEye y han confiado en la Universidad de El Salvador para la participación del proyecto.
- A las autoridades de la FIA-UES, por dar apoyo al proyecto FNET/GridEye.

Además de profesionales de otras instituciones, a nivel nacional y extranjero, que nos brindaron asesoramiento para la realización de este proyecto.

INTRODUCCIÓN

La frecuencia eléctrica es uno de los parámetros más fundamentales dentro de la red eléctrica de una región, ya que su estabilidad es crucial para el correcto funcionamiento de motores eléctricos y otros dispositivos dependientes de esta. Es necesario que la frecuencia sea controlada, monitoreada y analizada continuamente para garantizar la estabilidad y confiabilidad de la red eléctrica. Las perturbaciones en la frecuencia pueden ocasionar consecuencias graves, afectando tanto la infraestructura eléctrica como los dispositivos conectados a ella.

Este proyecto se enfoca en realizar un análisis estadístico de las perturbaciones de frecuencia eléctrica monitoreadas en El Salvador. Se describe estadísticamente el comportamiento de la frecuencia eléctrica tanto en condiciones normales como durante eventos de perturbación, además de identificar posibles patrones y tendencias en los datos recopilados. Para este análisis, se han utilizado dos fuentes de datos principales: los datos proporcionados por la Unidad de Transacciones (UT) de su sistema de medición SCADA, que abarcan el periodo completo de 2022 y 2023, y los datos medidos por la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador (UES) mediante un «Medidor de Perturbaciones de Frecuencia» (FDR) donado por la Universidad de Tennessee en Knoxville (UTK), los cuales comprenden el periodo de septiembre de 2022 a junio de 2023.

Además, dado que la frecuencia eléctrica está intrínsecamente relacionada con el equilibrio entre la potencia generada y la potencia demandada en una región, se ha realizado un análisis correlativo preliminar entre los flujos de potencia regionales y las variaciones de frecuencia observadas. Este análisis ha sido posible gracias a los reportes de eventos proporcionados por el Ente Operador Regional (EOR), los cuales están disponibles públicamente en su sitio web. Debido a las limitaciones en el acceso a información más detallada, este estudio se considera preliminar, sirviendo como base para futuras investigaciones más exhaustivas a nivel centroamericano.

Finalmente, se presenta una introducción al Machine Learning, abarcando los distintos tipos de modelos predictivos. Además, se detalla la implementación de un modelo predictivo utilizando redes neuronales LSTM para evaluar su precisión, empleando los datos del FDR-UES.

ÍNDICE

Capítulo 1. Perfil del proyecto

1.1.	Planteamiento del proyecto	1
1.2.	Objetivos	2
1.2.1.	Objetivo general.....	2
1.2.2.	Objetivos específicos	2
1.3.	Alcances.	3
1.4.	Antecedentes	4
1.5.	Planteamiento del problema	6
1.6.	Justificación.....	7
1.7.	Resultados esperados	8

Capítulo 2. Conceptos básicos acerca de la frecuencia electrica de la red salvadoreña

2.1.	Definición del concepto de frecuencia, instituciones y mercados.	10
2.2.	Estándares, recomendaciones y normativas en la frecuencia eléctrica.....	12
2.3.	Acerca del control de la frecuencia eléctrica.....	14
2.3.1.	Equilibrio entre la generación y carga en el sistema eléctrico.	14
2.3.2.	Acerca de la estabilidad del sistema eléctrico de potencia	15
2.3.3.	Respuesta en frecuencia ante una gran pérdida de generación	16
2.3.4.	Ejemplificación de evento de perdida de generación.	20

Capítulo 3. Obtención y tratamiento de datos

3.1. Acerca de FNET/GridEye y la obtención de datos registrados durante el periodo sept. 2022 – jun. 2023 por la unidad FDR-UES 1628.	24
3.2. Almacenamiento de mediciones del FDR-UES #1703 desde marzo 2024.....	27
3.3. Solicitud de datos a la Unidad de Transacciones.....	28
3.4. Utilización de reportes de eventos de anomalías en el SER (EOR).....	29
3.5. Acerca del manejo de los datos.....	30
3.5.1. Creación de set de datos.....	31
3.5.2. Validación de datos del FDR-1628 y el SCADA.....	42

Capítulo 4. Análisis descriptivo del valor de la frecuencia eléctrica y su tendencia en el tiempo

4.1. Descripción del comportamiento del valor de la frecuencia eléctrica.	49
4.1.1. Medidas de tendencia central y variabilidad.....	49
4.1.2. Tabla de distribución de frecuencias	50
4.1.3. Gráficos de cajas y bigotes.....	51
4.1.4. Histograma y curva de densidad.....	53
4.1.5. Prueba de aproximación a normalidad: Asimetría y apuntamiento.	55
4.2. Identificación de la distribución del valor de la frecuencia en el tiempo.....	57
4.2.1. Tendencias anuales (2022 y 2023): Datos del SCADA-UT.....	57
4.2.2. Tendencias mensuales.....	58
4.2.3. Tendencias para cada día de la semana	63
4.2.4. Tendencias para cada hora del día.	66

Capítulo 5. Acerca del registro de eventos de perturbación de frecuencia y su análisis

5.1.	Creación de lista de eventos de perturbación de frecuencia.	74
5.1.1.	Propuesta de clasificación de eventos de perturbación según su magnitud.	74
5.1.2.	Identificación de eventos de perturbación.....	75
5.1.3.	Registro de información del evento	78
5.2.	Descripción del comportamiento de los registros de eventos de perturbación.....	79
5.2.1.	Contabilización de eventos de perturbación en el periodo de estudio.	79
5.2.2.	Diagrama de dispersión de los eventos de perturbación.....	80
5.2.3.	Comportamiento de eventos de perturbación según su clasificación.	83
5.3.	Tendencias en el tiempo.....	89
5.3.1.	Tendencias anuales.....	89
5.3.2.	Tendencias mensuales.....	91
5.3.3.	Tendencias para cada día de la semana	94
5.3.4.	Tendencias para cada hora del día.	95

Capítulo 6. Análisis preliminar de la frecuencia eléctrica y los flujos de potencia regionales

6.1.	Contabilización de eventos reportados en el SER por el EOR.	98
6.2.	Interconexión Sistema Eléctrico Regional – Sistema Eléctrico Mexicano.	102
6.3.	Contabilización de eventos reportados en El Salvador por el EOR: Causas.	105

Capítulo 7. Implementación de Modelo Predictivo con Redes Neuronales LSTM para la Predicción de Frecuencia Eléctrica usando Datos del FDR-1628

7.1.	Definición de modelo predictivo.....	110
7.2.	Breve introducción sobre las Redes Neuronales.....	110
7.2.1.	Red neuronal básica.....	110
7.2.2.	Red neuronal Recurrente.....	114
7.2.3.	Red neuronal de corto y largo plazo (LSTM).....	114
7.3.	Preparación del entorno.....	117
7.4.	Diseño e implementación de una red LSTM.....	118
7.4.1.	Configuración de la Red Neuronal.....	118
7.4.2.	Construcción del modelo.....	119
7.5.	Simulación de modelo.....	122
7.5.1.	Resultados de simulación.....	122
7.5.2.	Observaciones Clave.....	132
Conclusiones		134
Recomendaciones		137
Referencias		139

ÍNDICE DE FIGURAS

FIGURAS DEL CAPITULO I

Figura 1.1. Medidor de Perturbaciones de Frecuencia donado a la Universidad de El Salvador (FDR-UES), por el Laboratorio de Tecnología de la Información de Energía de la Universidad de Tennessee.	1
Figura 1.2. Primeras unidades PMU experimentales desarrolladas en Virginia Tech en 1980s, extraído de [1].	4

FIGURAS DEL CAPITULO II

Figura 2.1. Representación de una frecuencia de 60 Hz durante un segundo, para una onda senoidal.	10
Figura 2.2. Analogía del balance entre la generación y la carga eléctrica. Extraído de [17]	14
Figura 2.3. Clasificación CRIGE expandida de la estabilidad de los sistemas eléctricos de potencia, simplificado.....	15
Figura 2.4. Incidente de caída de frecuencia por pérdida de generación, extraído de IEEE C.37.106 [23].	16
Figura 2.5. Respuesta en la frecuencia típica del sistema de potencia ante una perturbación. Extraído de [25].	17
Figura 2.6. Identificación de puntos A, B y C. Extraído de [33].	19
Figura 2.7. Respuesta en frecuencia ausente (Withdrawal) durante evento de perturbación. Extraído de [35].	20
Figura 2.8. Frecuencia del Sistema eléctrico Regional (SER) durante evento de perturbación de frecuencia. Extraído de [36].	21
Figura 2.9. Pérdida de generación reportada en El Salvador, causante del evento de perturbación de frecuencia. Extraído de [36].	22

FIGURAS DEL CAPITULO III

Figura 3.1. Medidor de perturbaciones de frecuencia (FDR), segunda generación, desarrollado por el Laboratorio de Tecnología de la Información de Energía del departamento de Ingeniería Eléctrica y Ciencias de la Computación de la Universidad de Tennessee. Extraído de [38].....	24
Figura 3.2. Ilustración de la instalación típica de un FDR. Extraído de [38]	25
Figura 3.3. Captura de pantalla tomada en marzo del 2024, de las variaciones de frecuencia registradas por el FDR-UES. Disponible en [39].	25
Figura 3.4. Muestra de datos descargados del medidor FDR-UES	26
Figura 3.5. CPU marca DELL, tipo OptiPlex 7040SFF, utilizado para el almacenamiento de registros.....	27
Figura 3.6. Reportes de eventos en el SER, disponibles en su sitio web para consulta/descarga.	29
Figura 3.7. Resumen de información relevante extraído para cada evento de pérdida de generación/carga ocurrido en el SER.....	29
Figura 3.8. Repositorio de GitHub del proyecto de graduación. Captura tomada el 25 de abril del 2024.....	30
Figura 3.9. Set de datos creados del SCADA UT, del año 2022.....	31
Figura 3.10. Datos fuente de mediciones realizadas por el FDR-1628 desde el 02 de septiembre del 2022 hasta el 22 de junio del 2023.....	32
Figura 3.11. Archivo csv de datos proporcionados, correspondiente a 01feb-23feb2023.csv.....	32
Figura 3.12. Limpieza de ruido de la frecuencia eléctrica aplicada la Transformada de Fourier, Transformada de Wavelet y Promedio Móvil. Fecha: 08 de febrero del 2023.	35
Figura 3.13. Limpieza de ruido de la frecuencia eléctrica aplicada la Transformada de Fourier, Transformada de Wavelet y Promedio Móvil. Fecha: 14 de febrero del 2023.	36
Figura 3.14. Carpetas de datos proporcionados por la Unidad de Transacciones S.A de C.V	38
Figura 3.15. Carpetas anuales de «Datos Demanda, Intercambios, Frecuencia» proporcionados por la Unidad de Transacciones.	38

Figura 3.16. Carpetas mensuales, correspondientes al año 2022, de «Datos Demanda, Intercambios, Frecuencia», proporcionados por la Unidad de Transacciones.....	39
Figura 3.17. Archivos csv para cada uno de los días del mes de mayo del 2022, de «Datos Generación, Intercambios, Frecuencia» proporcionados por la Unidad de Transacciones.	39
Figura 3.18. Carpetas de «Datos Generación» proporcionados por la Unidad de Transacciones.	40
Figura 3.19. Archivos csv de generación, de «Datos Generación» proporcionados por la Unidad de Transacciones.	41
Figura 3.20. Set de datos del SCADA-UT, mensualizados, limpiados y unificados.....	41
Figura 3.21. Comparación de mediciones: FDR-UES 1628 (datos brutos) vs SCADA UT, en condición normal, entre las 12:16 hasta las 12:20 del 19 de febrero del 2023 (UTC-6).....	42
Figura 3.22. Comparación de mediciones: FDR-UES 1628 (remuestreado a cuatro segundos) vs SCADA UT, en condición normal, entre las 12:16 hasta las 12:20 del 19 de febrero del 2023 (UTC-6).	43
Figura 3.23. Comparación de mediciones: FDR-UES 1628 (datos brutos) vs SCADA UT, en condición crítica, entre las 09:16 hasta las 09:20 del 14 de febrero del 2023 (UTC-6).	44
Figura 3.24. Comparación de mediciones: FDR-UES 1628 (remuestreado a cuatro segundos) vs SCADA UT, en condición crítica, entre las 09:16 hasta las 09:20 del 14 de febrero del 2023 (UTC-6).	44

FIGURAS DEL CAPITULO IV

Figura 4.1. Partes de un diagrama de cajas y bigotes.	51
Figura 4.2. Diagrama de cajas y bigotes: Distribución de la frecuencia eléctrica.	52
Figura 4.3. Diagrama de cajas y bigotes (sin valores atípicos): Distribución de la frecuencia eléctrica.	53
Figura 4.4. Histograma y curva de densidad: Valores de la frecuencia en El Salvador (FDR-UES 1628).....	54

Figura 4.5. Histograma y curva de densidad: Valores de la frecuencia en El Salvador (SCADA-UT).....	54
Figura 4.6. Distribuciones de datos: (a) Asimetrías, (b) Apuntamiento. Extraídos de la web.	55
Figura 4.7. Diagrama de cajas y bigotes (sin valores atípicos): Distribución de la frecuencia eléctrica en el año 2022 y en el año 2023, según mediciones del SCADA-UT.....	57
Figura 4.8. Diagrama de cajas y bigotes (con valores atípicos): Distribución de la frecuencia eléctrica en el año 2022 y en el año 2023, según mediciones del SCADA-UT.....	58
Figura 4.9. Tendencia de la desviación estándar para cada mes del año: SCADA-UT	60
Figura 4.10. Tendencia de la desviación estándar para cada mes del año: FDR-UES 1628.....	60
Figura 4.11. Distribución de los valores de la frecuencia para cada uno de los meses (2022 y 2023): SCADA-UT.....	63
Figura 4.12. Tendencia de la desviación estándar para cada día de la semana: SCADA-UT y FDR-UES 1628.....	63
Figura 4.13. Distribución de los valores de la frecuencia para cada día de la semana: SCADA-UT.....	64
Figura 4.14. Distribución de los valores de la frecuencia para cada día de la semana: FDR-UES 1628.....	65
Figura 4.15. Tendencia de la desviación estándar para hora del día: SCADA-UT y FDR-UES 1628.....	66
Figura 4.16. Tendencia de la desviación estándar para hora del día, considerando únicamente el periodo del día 02 de septiembre del 2022 hasta el 22 junio del 2023: SCADA-UT y FDR-UES 1628.....	66
Figura 4.17. Distribución de los valores de la frecuencia para cada hora del día: FDR-UES 1628.....	69
Figura 4.18. Distribución de los valores de la frecuencia para cada hora del día: SCADA-UT ..	69
Figura 4.19. Tendencia de la desviación estándar para hora del día, para cada cuatrimestre: SCADA-UT.....	72

Figura 4.20. Tendencia de la desviación estándar para hora del día, para días laboral y no laboral: SCADA-UT.....	72
---	----

FIGURAS DEL CAPITULO V

Figura 5.1. Ilustración del funcionamiento del algoritmo del promedio móvil y la desviación estándar móvil.....	76
Figura 5.2. Detección de puntos de cambio: (a) Por un cambio en el valor de la media, (b) Por un cambio en el valor de la desviación estándar. Extraído de [65].....	77
Figura 5.3. Gráfico de dispersión de los eventos de perturbaciones de frecuencia registradas, de la magnitud máxima reportada y la duración de los eventos: Escala semilogarítmica (eje vertical).	80
Figura 5.4. Gráfico del comportamiento de la frecuencia durante eventos de perturbación.	83
Figura 5.5. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación menores.....	83
Figura 5.6. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación moderados.....	85
Figura 5.7. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación críticos.....	86
Figura 5.8. Comportamiento de dos eventos de perturbación críticos: Casos atípicos.....	88
Figura 5.9. Contabilización de eventos de perturbación de frecuencia, por año.	89
Figura 5.10. Promedio de desviación máxima de eventos de perturbación de frecuencia reportados, por año.	90
Figura 5.11. Contabilización de eventos de perturbación de frecuencia, por mes.....	91
Figura 5.12. Contabilización de eventos de perturbación de frecuencia, por día de la semana....	94
Figura 5.13. Contabilización de eventos de perturbación de frecuencia, por cada intervalo de horas del día.....	96

FIGURAS DEL CAPITULO VI

Figura 6.1. Esquema del Sistema Eléctrico Regional de América Central (SER). Extraído de [68].	98
Figura 6.2. Diagrama de cajas y bigotes: Desviaciones máximas que ha presentado la frecuencia eléctrica en eventos de pérdida de generación/carga, según el origen. Realizado con reportes del EOR de 2022 y 2023.....	100
Figura 6.3. Visualización de eventos registrado y no registrado por el EOR, desde la perspectiva salvadoreña, utilizando datos brindados por la Unidad de Transacciones S.A de C.V.	101
Figura 6.4. Distribución de la frecuencia eléctrica (medida en El Salvador) debido a la interconexión SEM-SER.....	104
Figura 6.5. Distribución de pérdidas de generación/cargas reportadas por tipo de tecnología, según reportes del Ente Operador Regional, durante 2022 - 2023.....	108

FIGURAS DEL CAPITULO VII

Figura 7.1. Estructura de una red neuronal. Extraído de [74].....	111
Figura 7.2. Composición de capas de una red neuronal. Extraído de [77].	112
Figura 7.3. Arquitectura de una Red Neuronal tipo LSTM. Extraído de [82].....	115
Figura 7.4. Datos reales de frecuencia eléctrica, con un resampleo de seis segundos.....	123
Figura 7.5. Datos reales de frecuencia eléctrica, con un resampleo de una hora.....	123
Figura 7.6. Datos reales con datos predichos de frecuencia eléctrica con un resampleo de seis segundos.....	124
Figura 7.7. Datos reales con datos predichos de frecuencia eléctrica con resampleo de una hora.	125
Figura 7.8. Pérdidas durante el entrenamiento con remuestro de seis segundos.	126
Figura 7.9. Pérdidas durante el entrenamiento con remuestro de una hora.	127

Figura 7.10. Representación del periodo de entrenamiento de datos (remuestreo de seis segundos).	127
Figura 7.11. Representación del periodo de entrenamiento de datos (remuestreo de una hora). 128	
Figura 7.12. Forecasting de frecuencia eléctrica con resamplero de seis segundos. Pronostico realizado para los próximos 120 segundos.	129
Figura 7.13. Forecasting de valores de frecuencia con resamplero de una hora. Pronóstico realizado para las próximas 120 horas.....	130

ÍNDICE DE TABLAS

TABLA I. Ajustes del EDCBF actual en El Salvador (revisado en dic. 2023).....	18
Tabla II. Medidas de tendencia central y variabilidad del valor de la frecuencia eléctrica: SCADA-UT y FDR-UES.....	50
Tabla III. Tabla de distribución de los valores de la frecuencia eléctrica: FDR-UES y SCADA-UT.....	50
Tabla IV. Parámetros de los diagramas de caja y bigotes.	53
Tabla V. Interpretación de resultados de asimetría (Skewness) y apuntamiento (Kurtosis)	55
Tabla VI. Medidas de tendencia central y variabilidad del valor de la frecuencia eléctrica: SCADA-UT	57
Tabla VII. Parámetros de los diagramas de caja y bigotes anuales: SCADA-UT	58
Tabla VIII. Medidas de tendencia central y variabilidad para cada uno de los meses y años: SCADA-UT.....	59
Tabla IX. Medidas de tendencia central y variabilidad para cada uno de los meses y años: FDR-UES 1628	60
Tabla X. Medidas de tendencia central y variabilidad de cada mes: SCADA-UT	61
Tabla XI. Parámetros de los diagramas de caja y bigotes mensuales: SCADA-UT	62
Tabla XII. Parámetros de los diagramas de caja y bigotes mensuales: FDR-UES 1628	62
Tabla XIII. Medidas de tendencia central y variabilidad para cada uno de los días de la semana: SCADA-UT.....	64
Tabla XIV. Medidas de tendencia central y variabilidad para cada uno de los días de la semana: FDR-UES 1628.....	64
Tabla XV. Parámetros de los diagramas de caja y bigotes para cada día de la semana: SCADA-UT	65
Tabla XVI. Parámetros de los diagramas de caja y bigotes para cada día de la semana: FDR-UES 1628.....	65

Tabla XVII. Medidas de tendencia central y variabilidad para cada hora del día (general): SCADA-UT.	67
Tabla XVIII. Medidas de tendencia central y variabilidad para cada hora del día (general): FDR-UES 1628.	68
Tabla XIX. Parámetros de los diagramas de caja y bigotes para hora del día: SCADA-UT	70
Tabla XX. Parámetros de los diagramas de caja y bigotes para hora del día: FDR-UES 1628....	71
Tabla XXI. Contabilización de eventos de perturbación, para el periodo de estudio, según clasificación y tipo. Realizado en base a mediciones del SCADA-UT y el FDR-UES 1628.....	79
Tabla XXII. Contabilización de valores máximos de la frecuencia durante eventos.	81
Tabla XXIII. Interpretación de la magnitud del coeficiente de correlación, según sugerencias de Cohen. Extraído de [66], [67].	82
Tabla XXIV. Contabilización de valores máximos de la frecuencia durante eventos menores....	84
Tabla XXV. Parámetros de los diagramas de caja y bigotes para eventos menores.	84
Tabla XXVI. Contabilización de valores máximos de la frecuencia durante eventos moderados.	85
Tabla XXVII. Parámetros de los diagramas de caja y bigotes para eventos moderados.	86
Tabla XXVIII. Contabilización de valores máximos de la frecuencia durante eventos críticos. .	87
Tabla XXIX. Parámetros de los diagramas de caja y bigotes para eventos críticos.	88
Tabla XXX. Contabilización de eventos registrados, según su clasificación, para año	89
Tabla XXXI. Promedio de duraciones reportadas para eventos de perturbación de frecuencia, según su clasificación, para cada año del estudio.	90
Tabla XXXII. Contabilización de eventos registrados, según su clasificación, para cada mes y año.	92
Tabla XXXIII. Contabilización de eventos registrados, según su clasificación, para cada mes (sumados).....	93

Tabla XXXIV. Contabilización de eventos registrados, según su clasificación, para cada día de la semana.....	94
Tabla XXXV. Contabilización de eventos registrados, según su clasificación, para cada día de la semana.....	96
Tabla XXXVI. Contabilización de eventos reportados por el Ente Operador Regional, durante el periodo 2022 a 2023. Realizado en base a la contabilización de reportes preliminares, los cuales son públicos en su sitio web.....	99
Tabla XXXVII. Contabilización de eventos reportados por el Ente Operador Regional, durante el periodo 2022 a 2023, teniendo en consideración solo los eventos en que se observaron perturbaciones de frecuencia.....	99
Tabla XXXVIII. Parámetros de los diagramas de caja y bigotes.....	100
Tabla XXXIX. Contabilización de aperturas en la interconexión SER-SEM, para el periodo 2022 - 2023. Elaborado mediante reportes públicos del Ente Operador Regional.	103
Tabla XL. Contabilización de aperturas en la interconexión SER-SEM, para el periodo 2022 – 2023, y su impacto en la frecuencia eléctrica.	103
Tabla XLI. Parámetros de los diagramas de caja y bigotes.	104
Tabla XLII. Resumen de pérdida de generación/carga iniciados en El Salvador, durante el periodo 2022 – 2023, reportados por el Ente Operador Regional.....	105
Tabla XLIII. Tipo de falla ocurrido para las pérdidas de generación/cargas reportadas en El Salvador por el Ente Operador Regional, para el periodo 2022-2023.....	107
Tabla XLIV. Parámetros de los diagramas de caja y bigotes.	108
Tabla XLV. Métricas de evaluación de modelo elaborado.....	131

CAPITULO 1. PERFIL DEL PROYECTO

1.1. Planteamiento del proyecto.

La Facultad de Ingeniería y Arquitectura (FIA) de la Universidad de El Salvador (UES) ha tenido participación en el proyecto FNET/GridEye, que proporciona una red de medición de frecuencia sincronizada con GPS, permitiendo un monitoreo rápida y rentable y en tiempo real de parámetros eléctricos como lo sería la frecuencia. La Universidad de Tennessee en Knoxville (UTK) brindó a la UES un «Medidor de Perturbaciones de Frecuencia» (FDR), mediante el cual se pueden visualizar y transmitir los valores de la voltaje y frecuencia en tiempo real hacia la nube de internet. Mediante la cooperación de ingenieros de la Universidad Nacional Autónoma de México (UNAM), se realizará la configuración básica a un ordenador perteneciente a la Escuela de Ingeniería Eléctrica (EIE) para tener un almacenamiento de las mediciones de frecuencia realizadas localmente por el FDR-UES.



Figura 1.1. Medidor de Perturbaciones de Frecuencia donado a la Universidad de El Salvador (FDR-UES), por el Laboratorio de Tecnología de la Información de Energía de la Universidad de Tennessee.

La UNAM ha facilitado los datos recopilados por el medidor FDR-UES durante el periodo de septiembre 2022 hasta junio 2023. Además de ello, la Unidad de Transacciones (UT) ha facilitado los datos de frecuencia existentes para el periodo de enero 2022 a diciembre 2023. Por medio de técnicas estadísticas, se describirá el comportamiento de la frecuencia eléctrica, la realización de un análisis correlativo con respecto a los flujos de potencia, mediante los datos públicos existentes y los que puedan ser facilitados, y además de realizar un modelo predictivo.

1.2. Objetivos

1.2.1. Objetivo general

1. Analizar estadísticamente los registros de las mediciones de frecuencia obtenidos por el uso del medidor FDR-UES, realizando una descripción del comportamiento del valor de la frecuencia mediante la identificación y contabilización de los eventos de perturbación en el tiempo y el establecimiento de correlaciones con los flujos de potencia.

1.2.2. Objetivos específicos

1. Establecer una entrada de flujo de datos y su almacenamiento para las mediciones de frecuencia obtenidas por la unidad FDR-UES de correlativo #1703 mediante un ordenador de la EIE con el software openPDC desde la UNAM.
2. Crear un criterio de identificación de eventos de perturbación de frecuencia, teniendo en consideración recomendaciones brindadas por instituciones.
3. Aplicar técnicas estadísticas para poder realizar una descripción de los datos recopilados por el FDR-UES, para los datos recopilados desde septiembre 2022 hasta junio 2023, de los valores de la frecuencia y los eventos identificados.
4. Identificar tendencias existentes en el valor de la frecuencia eléctrica y los eventos de perturbación de frecuencia para distintos periodos de tiempo y épocas, para los datos recopilados desde septiembre 2022 hasta junio 2023.
5. Establecer correlaciones entre variaciones de la frecuencia y los flujos de potencia.
6. Evaluar la precisión de un modelo predictivo que permita determinar momentos más probables en que pueda ocurrir un evento de perturbación.

1.3. Alcances.

1. Realizar las adecuaciones necesarias, tanto en software como hardware, al ordenador perteneciente a la EIE para la correcta recepción y almacenamiento de flujo de datos.
2. Definir valores límites de variación de la frecuencia eléctrica, así como tiempos, para la identificación del inicio y finalización de un evento de perturbación, considerando en lo mayor posible las normativas, estándares y recomendaciones establecidos por instituciones nacionales e internacionales.
3. Utilizar técnicas estadísticas descriptivas para analizar los datos de frecuencia de la red eléctrica recopilados por la unidad FDR-UES #1628, obtenidos gracias a la UNAM durante el período especificado, como lo serían medidas de tendencia central, medidas de dispersión, y la creación de gráficos de distintos tipos y visualizaciones para representar la distribución de los datos y los eventos registrados, incluso para distintas divisiones de tiempo.
4. Analizar los datos recopilados por la unidad FDR-UES #1628 para identificar tendencias en el valor de la frecuencia eléctrica y registros de eventos de perturbación, según cada mes, día de la semana y hora del día.
5. Utilizar técnicas estadísticas adecuadas para determinar la relación existente entre el comportamiento de la frecuencia eléctrica y el flujo de potencia.
6. Evaluar el uso y precisión de modelos de regresión para determinar los momentos más probables en que puedan ocurrir eventos de perturbación de frecuencia en la red eléctrica, dejando planteado posibles limitantes, mejoras o ajustes necesarios.

1.4. Antecedentes.

Las Unidades de Medición Fasorial (PMU) son dispositivos desarrollados por Virginia Tech durante la década de 1980. Durante esa década, el despliegue de satélites GPS permitió el desarrollo de un sistema de medición que utilizaba señales horarias GPS y podía proporcionar una imagen instantánea del estado del sistema eléctrico. Debido al despliegue incompleto del sistema satelital, los primeros receptores GPS eran costosos, y la instalación de todo el equipo requerido era costoso y llevaba mucho tiempo, lo cual limitó su amplio uso en la industria eléctrica [1].

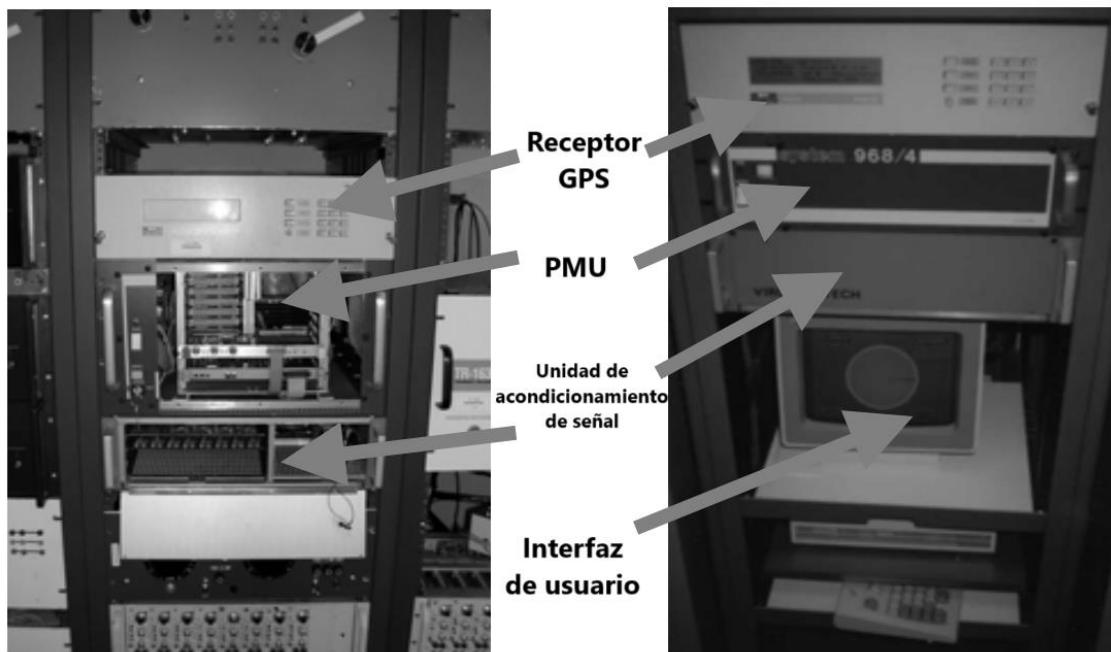


Figura 1.2. Primeras unidades PMU experimentales desarrolladas en Virginia Tech en 1980s, extraído de [1].

En el año 2000, un equipo de Virginia Tech liderado por Yilu Liu comenzó a desarrollar el proyecto de la red FNET, que consiste en la medición de fasores de bajo costo, y de rápida implementación con alta precisión dinámica, con una instalación tan simple desde un tomacorriente de pared. El sistema FNET entró en funcionamiento en noviembre del 2003 [2].

Desde 2010 se ha convertido en una red de monitoreo de red de área amplia que cubre las tres principales redes eléctricas de América del Norte y 16 de las redes más grandes del mundo.

Desde hace unos pocos años, la Escuela de Ingeniería Eléctrica inició su participación en el proyecto, gracias al donativo de uno de estos dispositivos por parte de la Universidad de Tennessee y la cooperación de la Universidad Nacional Autónoma de México.

Durante el año 2023 se realizó un proyecto de ingeniería, en el que se estableció contacto con los encargados de la UNAM para la obtención de las mediciones recopiladas por el medidor FDR de correlativo #1628, al cual nos referimos como FDR-UES #1628, o unidad 1628. Durante ese periodo de tiempo, se estableció un primer flujo de entrada de datos a un ordenador personal de forma exitosa. Sin embargo, debido a factores externos, el equipo 1628 sufrió daños irreparables, siendo necesario un reemplazo que no se efectuó hasta el mes de noviembre del 2023, por una nueva unidad de numero correlativo #1703, al cual nos referiremos normalmente como FDR-UES #1703, FDR-UES 2024, o unidad 1703. Los datos recolectados por la unidad 1628 comprendían desde septiembre del 2022 hasta junio del 2023, los cuales estaban almacenados en el equipo de la UNAM y nos fueron facilitados.

A su vez, durante el año 2024 se estableció un primer contacto con la Unidad de Transacciones para solicitar los datos que ellos poseían en su sistema SCADA. El ente accedió a brindar los datos, y gracias a ello se podrá realizar la comparación entre los datos del FDR-UES y el SCADA, además de ampliar el periodo de estudio de la investigación realizada. El sistema Supervisory Control and Data Acquisition (SCADA) surgió durante los años 1970, producto de la Tercera Revolución Industrial que surgió por esos mismos años siendo esto también útil para el monitoreo y recopilación de datos en tiempo real [3].

1.5. Planteamiento del problema

La Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador cuenta con un medidor FDR mediante el cual se puede monitorear y transmitir en tiempo real el valor de la frecuencia eléctrica, pero los datos monitoreados no han sido almacenados en un equipo propio en el que se puedan obtener con facilidad; siendo necesario solicitarlos a la Universidad Nacional Autónoma de México.

A su vez, no existen estudios públicos que describan el comportamiento de la frecuencia, y por ende del sistema eléctrico de potencia regional; lo que dificulta la simple realización de estudios más completos que puedan identificar factores y proponer medidas preventivas.

1.6. Justificación

Las variaciones de frecuencia en el sistema eléctrico pueden impactar significativamente la calidad de la energía eléctrica, generando consecuencias en el rendimiento y la integridad de los equipos eléctricos como los motores y su velocidad, los transformadores y su relación de transformación, la precisión y vida útil de dispositivos sensibles, entre otras consecuencias.

La EIE tiene la necesidad de disponer de datos confiables y accesibles para la realización de los análisis y estudios pertinentes sobre el comportamiento de la frecuencia, teniendo en consideración la relación existente de este parámetro con la estabilidad del sistema eléctrico.

Un análisis estadístico que describa el comportamiento de la frecuencia, aunque sea básico, permite comprender mejor al sistema eléctrico de potencia regional, logrando identificar fortalezas y debilidades existentes dentro del sistema, identificando los momentos en que el sistema se encuentra más vulnerable.

1.7. Resultados esperados

1. La recepción y el almacenamiento de mediciones de frecuencia de manera eficiente y confiable, mediante un flujo de entrada de datos facilitado por la UNAM.
2. El desarrollo de un criterio sólido y fundamentado para la identificación de eventos de perturbación de frecuencia.
3. Un análisis estadístico de los datos de frecuencia recopilados que dé una comprensión de su comportamiento en el período especificado.
4. La identificación de tendencias en el comportamiento de la frecuencia eléctrica y los eventos de perturbación a lo largo del tiempo, permitiendo entender mejor los patrones y variaciones de la frecuencia en el sistema eléctrico.
5. El identificar la correlación existente entre las variaciones en la frecuencia eléctrica y los flujos de potencia en el sistema eléctrico.
6. La determinación de la efectividad del uso de un modelo predictivo regresivo en la anticipación de los eventos de perturbación de frecuencia en el sistema eléctrico.

CAPITULO 2. CONCEPTOS BÁSICOS ACERCA DE LA FRECUENCIA ELECTRICA DE LA RED SALVADOREÑA

2.1. Definición del concepto de frecuencia, instituciones y mercados.

Presente en los sistemas de corriente alterna, cuya forma es senoidal para sistemas de distribución, la frecuencia eléctrica es el número de veces que una onda sinusoidal completa un ciclo de positivo a negativo, siendo esto medido en la unidad Hertz (Hz) [4]. Para El Salvador, el valor nominal de la frecuencia de la red eléctrica es de 60 Hz, lo que quiere significar que la onda se deberá de repetir 60 veces durante cada segundo.

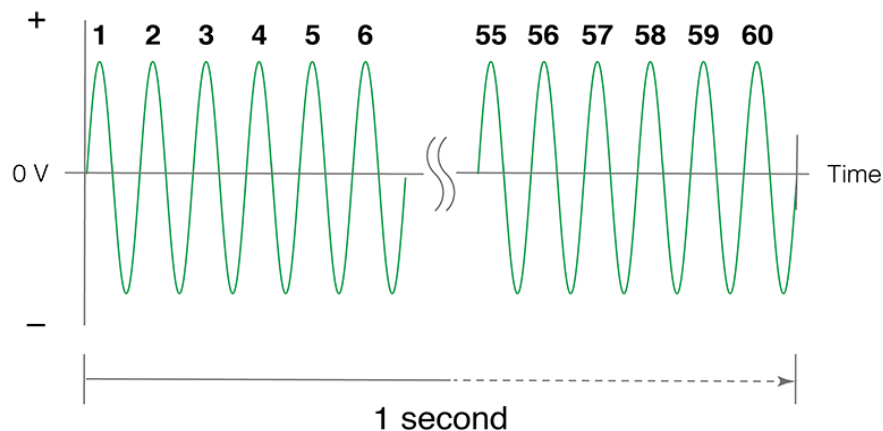


Figura 2.1. Representación de una frecuencia de 60 Hz durante un segundo, para una onda senoidal.

En muchas ocasiones haremos referencia de estándares y recomendaciones de instituciones internacionales como el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) y la Comisión Electrotécnica Internacional (IEC), o instituciones regionales.

- **IEEE:** El Instituto de Ingenieros Eléctricos y Electrónicos es la organización profesional técnica sin fines de lucro más grande del mundo, quienes establecen estándares (normas técnicas, recomendaciones prácticas, guías, etc.) para la industria eléctrica [5].
- **IEC:** La Comisión Electrotécnica Internacional (IEC) es una organización global de membresía sin fines de lucro que prepara y publica estándares internacionales para todas las tecnologías eléctricas, electrónicas y relacionadas [6].

En El Salvador, la Superintendencia General de Electricidad y Telecomunicaciones (SIGET) y la Unidad de Transacciones (UT) son entidades que tienen roles diferentes pero complementarios en el sector eléctrico.

- **SIGET:** Es la entidad competente para aplicar las normas contenidas en tratados internacionales sobre electricidad vigentes en El Salvador, en las leyes que rigen dicho sector y sus reglamentos, así como para conocer del incumplimiento de estas [7].
- **UT:** La UT es una entidad nacional que se dedica a la generación, transmisión y distribución de energía eléctrica en el mercado mayorista de El Salvador [8], el cual se rige bajo el reglamento aprobado por SIGET.

Aunque es cierto que el enfoque del trabajo realizado es en El Salvador, no hay que ignorar el hecho que nosotros como país pertenecemos a un mercado eléctrico más grande, el cual es el Mercado Eléctrico Regional (MER). Entiéndase el mercado eléctrico como un espacio financiero intangible en el que se compra y se vende la energía eléctrica de acuerdo con las condiciones de la oferta y la demanda existentes en la región de América Central.

El Mercado Eléctrico Regional (MER) está conformado por seis países, los cuales conforman al Sistema Eléctrico Regional (SER): Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica y Panamá. Su diseño general conceptualiza al MER como un séptimo mercado, superpuesto con los seis mercados o sistemas nacionales existentes [9], [10].

Dos organismos regionales mencionados en este trabajo son los siguientes:

- **Ente Operador Regional (EOR):** Es la entidad encargada de responsable del despacho e intercambios de energía entre países, en su calidad de administrador del mercado [10], [11].
- **Comisión Regional de Interconexión Eléctrica (CRIE):** responsable de regular las relaciones comerciales entre las instituciones públicas y privadas que se conectan al sistema y de fijar las tarifas [10].

Pueden existir otras instituciones o mercados que sean mencionados en este trabajo, sin embargo, consideramos que los aportes son apenas puntuales y podemos omitir la explicación de estas. Con únicamente lo mencionado hasta el momento creemos que el lector ya puede estar mejor orientado.

2.2. Estándares, recomendaciones y normativas en la frecuencia eléctrica.

Una perturbación de frecuencia eléctrica significa que la frecuencia fundamental del sistema de potencia ha tenido una variación con respecto a su valor nominal especificado. Debido a varias cuestiones que se mencionarán más adelante, se tendrá que el valor de la frecuencia eléctrica no siempre será igual a su valor nominal (establecida en 60 Hz para el caso de El Salvador).

Todo evento de perturbación se caracteriza por tener cierta magnitud y duración. Se puede intuir que existe cierto rango de variación que puede ser aceptable para el valor de la frecuencia en la red eléctrica. Dos criterios bien establecidos los brindan las instituciones IEEE e IEC, en su estándar 1159 y su reporte técnico 61000-2-5, respectivamente.

Para el caso de IEEE std.1159, se establecen que deben cumplirse las siguientes características para hablar de una operación normal [12]:

- La magnitud de las variaciones de frecuencia debe ser inferior a 0.10 Hz.
- La duración de las variaciones de frecuencias debe ser inferior a los diez segundos.

Para el caso de IEC TR 61000-2-5, se establece que [13]:

- La frecuencia fundamental de una red de potencia es muy estable y no varía en más de un 0.2% (equivalente a 0.12 Hz para El Salvador).

El Reglamento de Operación del Sistema de Transmisión y del Mercado Mayorista Basado en Costos de Producción (ROBCP), en su anexo «Normas de calidad y seguridad operativas», establece varios criterios acerca de la regulación de la frecuencia, destacando los siguientes puntos aplicables en la red eléctrica nacional [14]:

- La frecuencia nominal de la red de transmisión es de 60.00 Hz.
- En condición normal, la UT deberá mantener la frecuencia dentro de un rango entre 60.12 Hz y 59.88 Hz ($\pm 0.2\%$ de la frecuencia nominal).
- En condición de emergencia operativa, la UT deberá mantener la frecuencia dentro de un rango entre 60.60 Hz y 59.40 Hz ($\pm 1\%$ de la frecuencia nominal).
- Entre otros.

Además de ello, se establece como Criterios de Calidad, Seguridad y Desempeño del Sistema Eléctrico Regional (SER) en el Reglamento del Mercado Eléctrico Regional (RMER) lo siguiente para la frecuencia eléctrica, en los apartados 16.2.5.5 y 16.2.5.6 [15]:

- La frecuencia nominal del SER es de 60 Hz
- Durante la operación normal, el 90% de las variaciones de la frecuencia promedio en períodos de 10 minutos, deberán estar comprendidas dentro del rango de $60 \pm 1.65\sigma$ Hz, donde σ es la desviación estándar de la frecuencia promedio en períodos de 10 minutos. El valor de σ será de 0.03 Hz (agregado: el rango de valores sería de 60 ± 0.05 Hz).

2.3. Acerca del control de la frecuencia eléctrica.

2.3.1. Equilibrio entre la generación y carga en el sistema eléctrico.

El concepto de control de frecuencia en los sistemas de energía está estrechamente relacionado con el equilibrio entre la generación de energía y el consumo de energía. Por lo tanto, una potencia generada excedente conduce a la aceleración en generadores síncronos, velocidad de rotación y, por lo tanto, desviación positiva de la frecuencia de potencia. Por otro lado, un aumento en la demanda eléctrica o, de manera equivalente, una pérdida repentina de una unidad de generación da como resultado una caída en la frecuencia del sistema, afectando a toda carga motora [16].

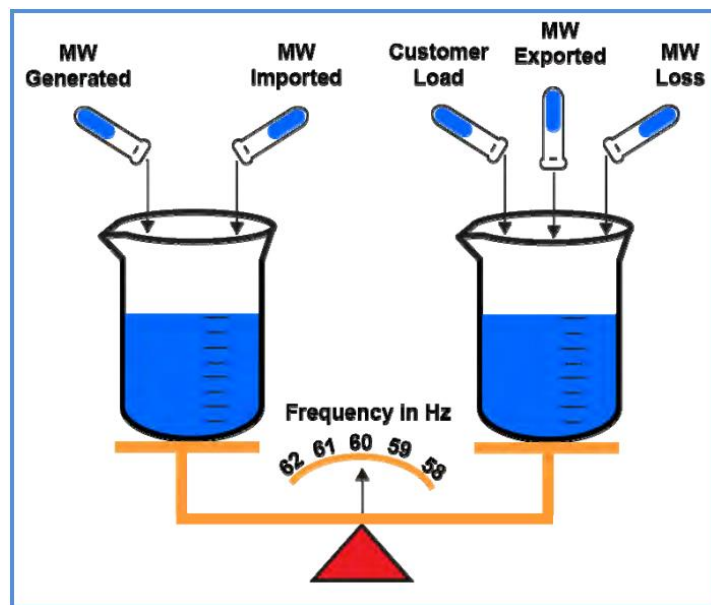


Figura 2.2. Analogía del balance entre la generación y la carga eléctrica. Extraído de [17]

En la práctica es imposible poder tener una igualdad entre la generación y demanda, siendo algo que se logra en pocos momentos del día [17]. La demanda tiene cambios a lo largo del día, y ello se ve reflejado en los Boletines Estadísticos de 2022 y 2023 de la UT [18], [19]. Sin embargo, mediante automatización se puede lograr que la diferencia existente sea la mínima posible.

Las perturbaciones en la frecuencia pueden afectar negativamente a equipos electrónicos y electromecánicos sensibles; existiendo un mayor riesgo en máquinas síncronas, incluidos generadores y motores síncronos, ya que algunos de sus devanados pueden experimentar flujos de corriente irregulares y sobrecargarse [20].

2.3.2. Acerca de la estabilidad del sistema eléctrico de potencia

La estabilidad de un sistema eléctrico de potencia es la capacidad de recuperar un estado de equilibrio operativo después de haber sido sometido a una perturbación física [21], [22]. Dependiendo de la topología de la red, condición de operación del sistema y la forma de la perturbación, diferentes conjuntos de fuerzas opuestas pueden experimentar un desbalance sostenido que conlleve a diferentes formas de inestabilidad.

Debido a las variadas formas de inestabilidad que un sistema de potencia puede alcanzar, se considera que una clasificación es esencial para realizar un análisis práctico significativo que pueda lograr solucionar los problemas de estabilidad presentados. Las clasificaciones realizadas se basan en las siguientes consideraciones [21]:

- La naturaleza física del modo resultante de inestabilidad como lo indica la principal variable en la cual la inestabilidad puede ser observada.
- El tamaño de la perturbación considerada, el cual influye en el método de cálculo y predicción de la estabilidad.
- Los dispositivos, procesos, y el tiempo que debe ser tomado en consideración en orden a restablecer la estabilidad.

Desde 2021, IEEE/CIGRE tiene en consideración las clasificaciones que se muestran [22]:

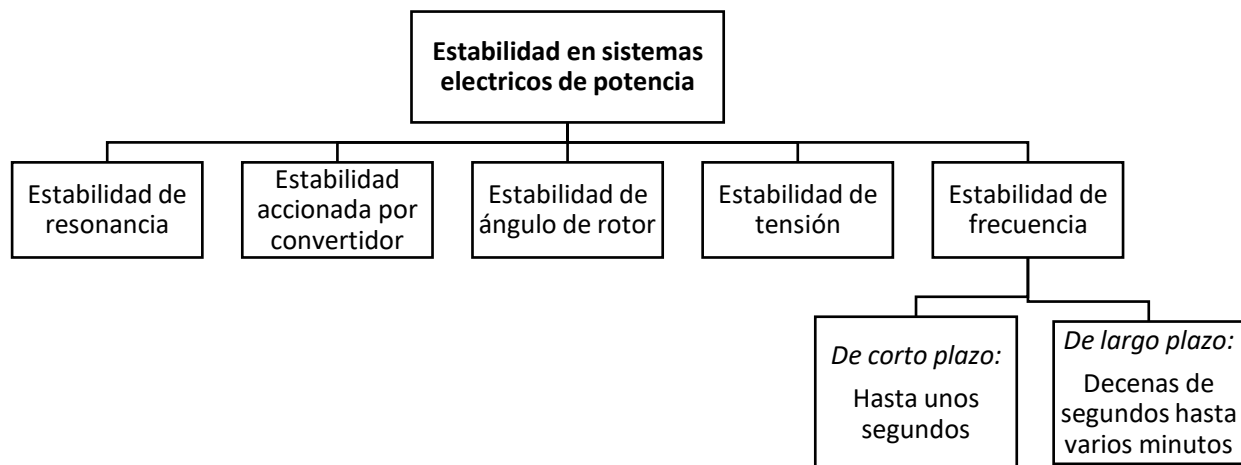


Figura 2.3. Clasificación CRIGE expandida de la estabilidad de los sistemas eléctricos de potencia, simplificado.

Dentro de lo que sería la estabilidad de la frecuencia, se refiere a la capacidad de un sistema de potencia para mantener la frecuencia constante después de una severa alteración del sistema que resulta en un desequilibrio significativo entre la generación y la carga [21], [22].

Un ejemplo de un incidente típico es el que se observa en la siguiente figura, extraído de la normativa IEEE C.37.106-2003, en la que se muestra la pérdida de una interconexión en la región este de los Estados Unidos el día 13 de mayo del 2002.

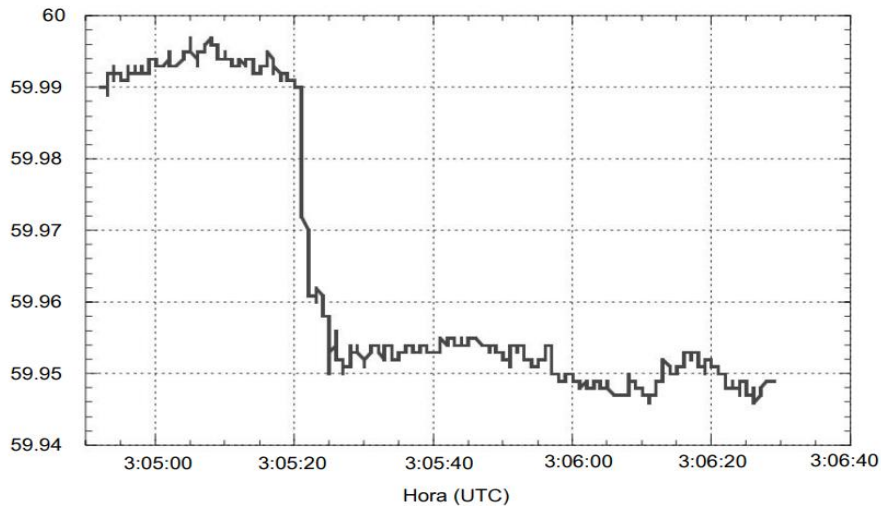


Figura 2.4. Incidente de caída de frecuencia por pérdida de generación, extraído de IEEE C.37.106 [23]

El comportamiento de la frecuencia podrá variar debido a factores como lo podrían ser las condiciones meteorológicas, la generación y consumo de potencia, la configuración adecuada de protecciones, entre otros factores. Según IEEE C37.106-2003, se puede considerar que un nivel de frecuencia inferior a 59.95 Hz es motivo de preocupación, y en algunos lugares ya desencadena la declaración de un estado de alarma. Cada región y país tiene sus criterios establecidos [23].

2.3.3. Respuesta en frecuencia ante una gran pérdida de generación

La respuesta en frecuencia es la capacidad de un sistema o elementos del sistema para reaccionar o responder a un cambio en la frecuencia del sistema [24]. En la figura muestra una respuesta genérica de un sistema de energía típico a una gran pérdida de generación. Se puede apreciar que hay tres períodos de respuesta [25]:

- i. El período inicial de respuesta hasta el punto más bajo de la desviación de frecuencia.
- ii. La recuperación inicial, determinada por la respuesta primaria y el comportar de la carga.
- iii. El retorno final de la frecuencia del sistema a la frecuencia nominal como se logra mediante el control automático de generación (AGC).

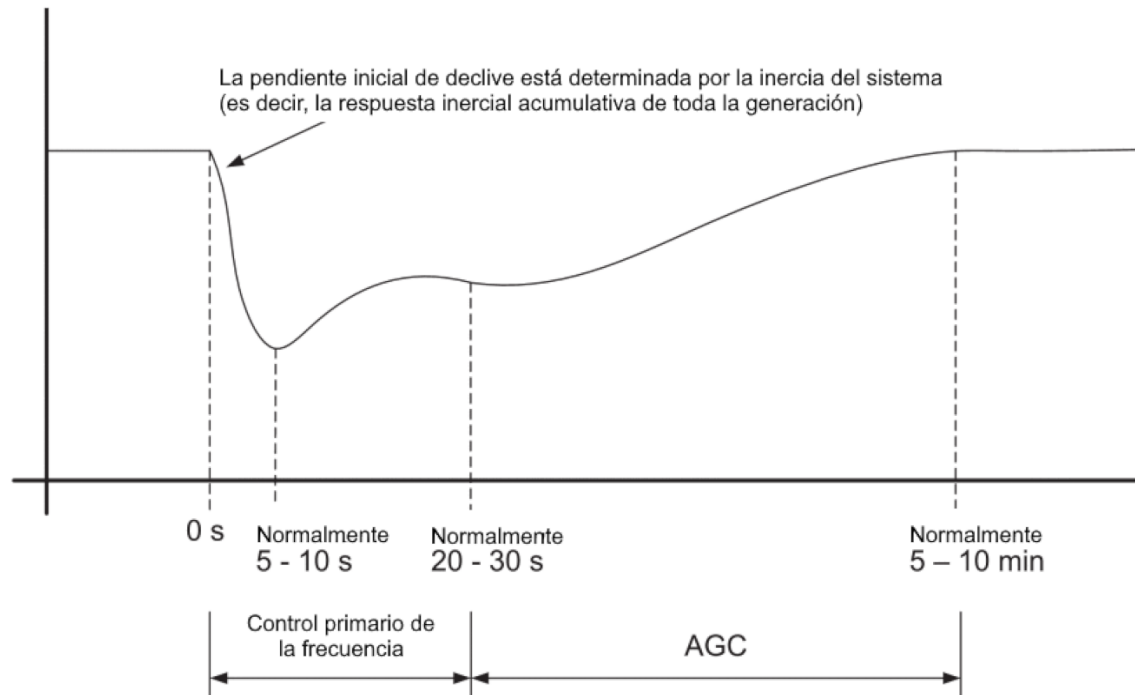


Figura 2.5. Respuesta en la frecuencia típica del sistema de potencia ante una perturbación. Extraído de [25].

Control primario de la frecuencia:

Dentro de lo que sería el control primario de la frecuencia, actúa lo que conoceremos como «Inercia del Sistema», el cual representa la sumatoria de toda la inercia existente por cada generador. La inercia es resistencia del generador ante un cambio de su velocidad, y mientras mayor se requerirá de un mayor cambio en la potencia para lograr cambiar el valor de la frecuencia, siendo por ello que la frecuencia no cambia inmediatamente ocurrido un evento de cambio en la potencia eléctrica generada y demandada [4], [17], [23], [25]. Por naturaleza, las generadoras que aportan mayor inercia al sistema son las térmicas e hidroeléctricas [26].

Dentro del control primario de la frecuencia, también se deben de hacer uso de las reservas existente en cada uno de los generadores, con el fin de que el valor de la frecuencia eléctrica deje de caer, dado que la respuesta inercial lo único que se encarga es de disminuir el ROCOF [26],

[27]. Para el caso de las infrafrecuencias, en muchas ocasiones este puede ser algo difícil, por no decir imposible, por lo que la otra alternativa es realizar la desconexión de carga. Para mitigar los efectos de las perturbaciones de frecuencia, se hace uso de relés de frecuencia, los cuales supervisan constantemente la frecuencia y generar un disparo cuando es necesario [28].

La Unidad de Transacciones ha realizado un esquema de desconexión de carga por baja frecuencia, en base a la realidad energética existente en El Salvador y Centroamérica. El esquema consta de seis etapas, especificando el porcentaje de demanda a desconectar. Las primeras tres etapas están sincronizadas a nivel centroamericano. Estos esquemas son revisados cada año, típicamente durante el mes de diciembre [29].

TABLA I. Ajustes del EDCBF actual en El Salvador (revisado en dic. 2023)

Etapa	f [Hz]	Retraso de tiempo (ciclos)	Carga a desconectar
I*	59.3	6	3%
II*	59.1	6	4%
III*	58.9	6	5%
IV	58.5	0	19%
V	58.2	0	7%
VI	57.9	0	5%

Nota: La apertura de interconexiones se realiza a partir de 58.8 Hz, de acuerdo con lo establecido por el Ente Operador Regional
 *Valores coordinados a nivel centroamericano por el Ente Operador Regional

Además de ello, el Ente Operador Regional ha establecido la apertura de interconexiones dentro de la región para cuando se presenten valores de frecuencia inferiores a 58.80 Hz. Esta desconexión es escalonada, existiendo los siguientes tiempos de retraso para las interconexiones [30]:

- Honduras – Nicaragua: 0.2 segundos.
- Nicaragua – Costa Rica: 0.6 segundos.
- Resto de interconexiones: 1.0 segundo.

Estas acciones ayudan a que la diferencia entre la potencia suministrada (P_G) y la potencia demandada (P_L) se minimice, reduciendo esta última y permitiendo que el valor de la frecuencia pueda ir a un valor más seguro, que evite el colapso del sistema regional.

Control secundario de la frecuencia

El Control Primario de la Frecuencia es efectivo para limitar la caída de la frecuencia, pero no puede eliminar el error de estado estacionario en la señal de frecuencia industrial. Por lo tanto, se requiere una regulación de frecuencia secundaria para corregir este error y restaurar el sistema a su estado previo a la contingencia.

El control secundario es un algoritmo centralizado (conocido como Control Automático de Generación o AGC) ejercido automática o manualmente por una entidad de control superior en la jerarquía de control de carga (generalmente el operador del sistema de transmisión) [16]. El objetivo principal de AGC es mantener o recuperar la frecuencia del sistema y el intercambio de potencia de la línea de enlace en sus valores programados, ajustando salida de las unidades generadoras según los cambios en la demanda de carga y las condiciones del sistema [17].

Durante el evento de perturbación, se identifican algunos puntos relevantes que pueden ser calculados para estimar el valor de la respuesta en frecuencia [31], [32], [33]:

1. El punto A es el valor promedio de la frecuencia existente antes de una perturbación
2. El punto B es el valor promedio de la frecuencia en su etapa de recuperación
3. El punto C es el valor mínimo existente de la frecuencia durante el evento.

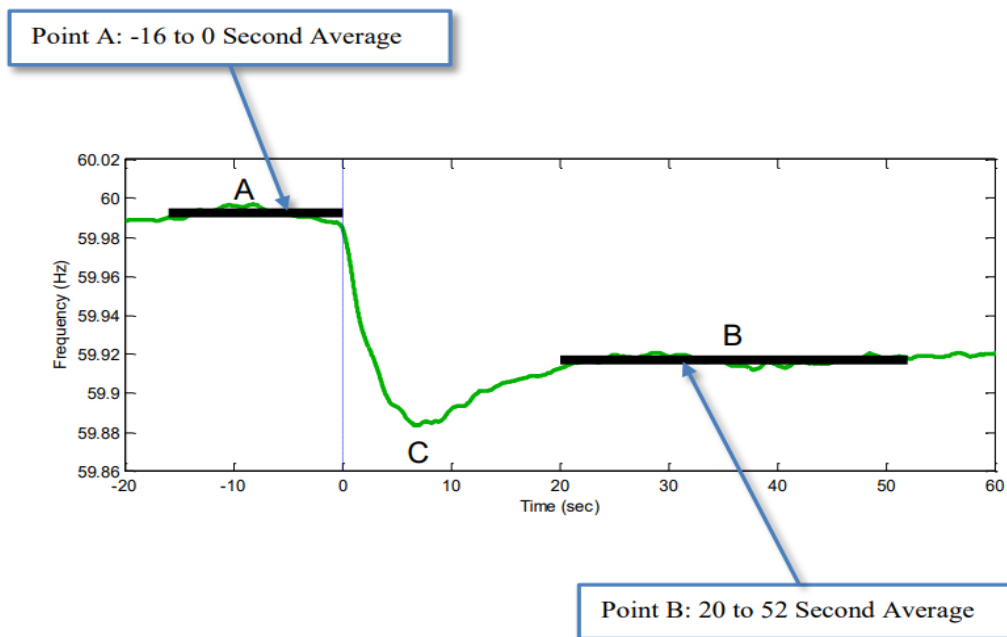


Figura 2.6. Identificación de puntos A, B y C. Extraído de [33].

Cabe aclarar que puede ocurrir que no se dé una respuesta típica de la frecuencia en todos los casos de pérdida de generación. En ocasiones, puede darse que exista una «Respuesta Ausente» (en inglés, el termino es Response Withdrawal), el cual se puede identificar porque el punto mínimo de la frecuencia no fue encontrado durante los primeros 20 segundos del evento [34], [35].

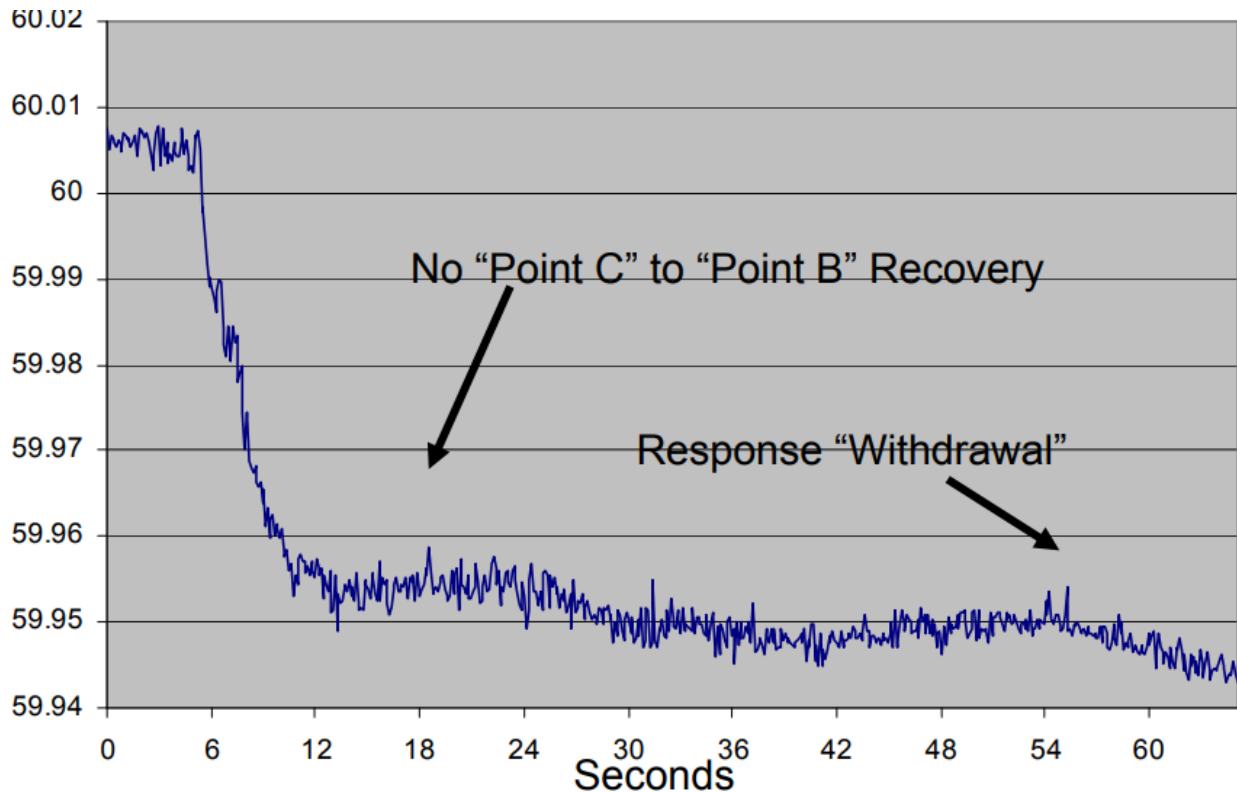


Figura 2.7. Respuesta en frecuencia ausente (Withdrawal) durante evento de perturbación. Extraído de [35].

2.3.4. Ejemplificación de evento de pérdida de generación.

Un caso para ejemplificar es el ocurrido el 07 de mayo del 2023, a las 10:54 hora local salvadoreña, el cual, según el informe realizado por el Ente Operador Regional, ocurrió debido a una pérdida en la generación existente en la Central Energía del Pacífico, en El Salvador.

Para el ejemplo mostrado en las figuras, se puede observar que el evento de falla ocurrido a las 10:54 horas provocó que posteriormente descendiera a 59.096 Hz [36].

Citando lo descrito por el Ente Operador Regional para el evento mencionado:

“Por el disparo de Central Energía del Pacífico se registra en el SEP El Salvador, la pérdida de 340 MW de generación, ocasionando que el flujo por la línea de interconexión de 400 kV de México a Guatemala alcance un valor de 420 MW, provocando la desconexión de esta línea por actuación del Esquema de Disparo Transferido por Bajo Voltaje (EDALTBV). Situación que origina un déficit de generación significativo en el SER, con la consecuente operación del EDACBF hasta la segunda etapa, desconectando 399.77 MW de carga. Operadores del sistema utilizaron su reserva rodante para contribuir con la estabilización de la frecuencia del SER.” – EOR, informe preliminar 91-05-2023 [36].

Utilizando este ejemplo, retómese algunos conceptos recién mencionados. Se puede observar que la pérdida de generación inicial que ocurrió en El Salvador dio a que iniciara una inestabilidad en el Sistema Eléctrico Regional de América Central. Debido a este déficit inicial ocurrido en El Salvador, la interconexión existente con México tuvo una apertura, surgiendo otro déficit de generación adicional. Debido al gran déficit existente en el sistema, inició la operación del EDCBF para disminuir la carga, y por ende reducir el déficit actualmente existente.

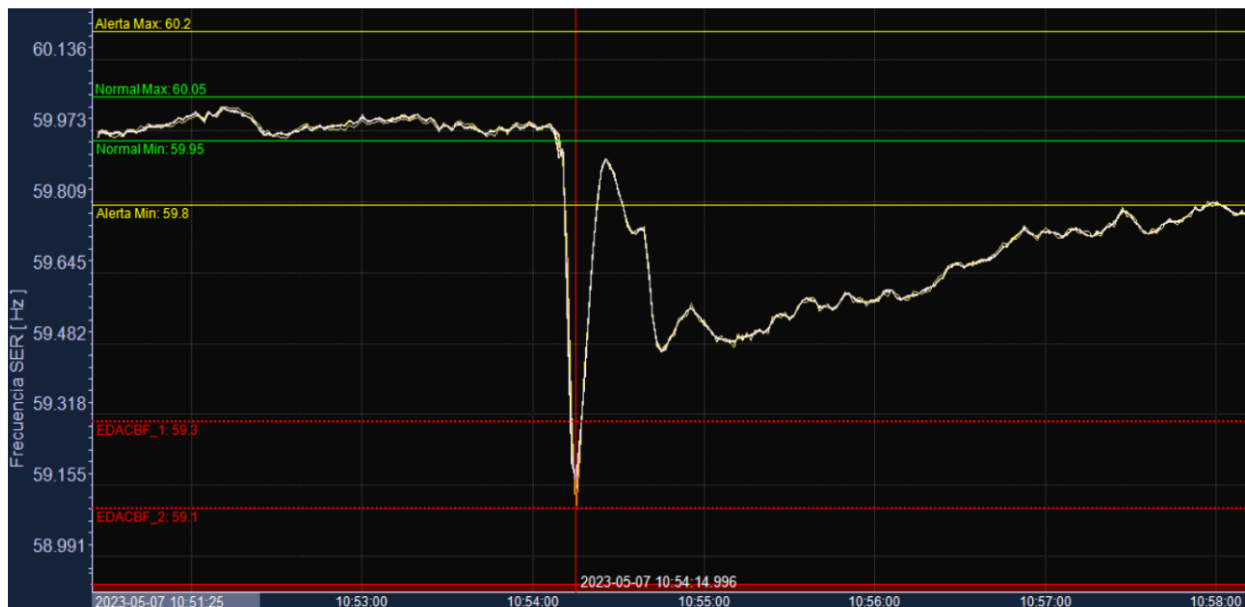


Figura 2.8. Frecuencia del Sistema eléctrico Regional (SER) durante evento de perturbación de frecuencia. Extraído de [36]

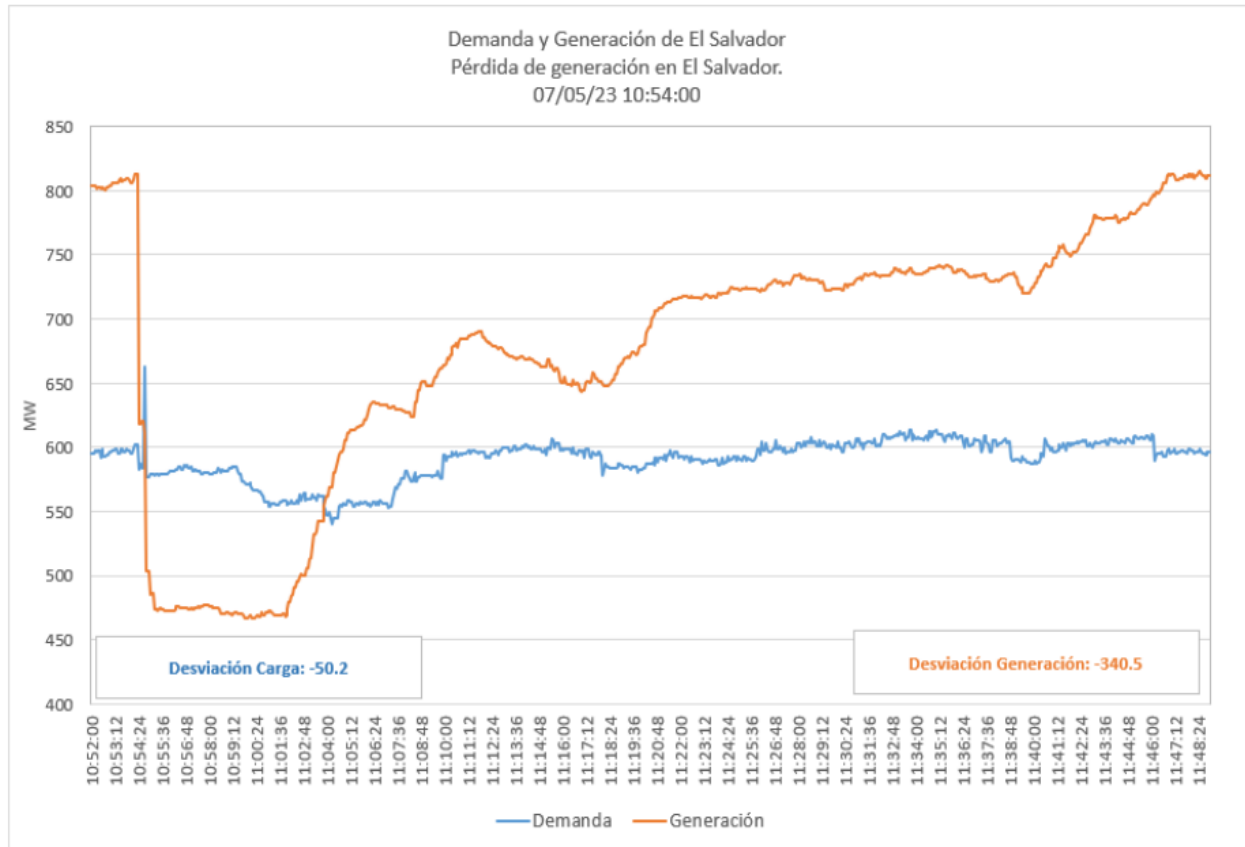


Figura 2.9. Pérdida de generación reportada en El Salvador; causante del evento de perturbación de frecuencia. Extraído de [36]

CAPITULO 3. OBTENCIÓN Y TRATAMIENTO DE DATOS

3.1. Acerca de FNET/GridEye y la obtención de datos registrados durante el periodo sept. 2022 – jun. 2023 por la unidad FDR-UES 1628.

En el año 2000, un equipo de Virginia Tech liderado por Yilu Liu comenzó a desarrollar el proyecto de la red FNET, que consiste en la medición de fasores de bajo costo, y de rápida implementación con alta precisión dinámica, con una instalación tan simple desde un tomacorriente de pared. El sistema FNET entró en funcionamiento en noviembre del 2003 [2].

Operado por el Laboratorio de Tecnología de la Información de Energía de la Universidad de Tennessee, FNET/GridEye es una red de medición de frecuencia de área amplia sincronizada por GPS de bajo costo y despliegue rápido [37], del cual la Facultad de Ingeniería y Arquitectura de la Universidad de El Salvador hace participación. Los FDR de alta precisión desarrollados se utilizan para medir y registrar principalmente los valores de la frecuencia de la señal de potencia que se encuentra en los enchufes eléctricos ordinarios de 120 V o 240 V.



Figura 3.1. Medidor de perturbaciones de frecuencia (FDR), segunda generación, desarrollado por el Laboratorio de Tecnología de la Información de Energía del departamento de Ingeniería Eléctrica y Ciencias de la Computación de la Universidad de Tennessee. Extraído de [38]

Además del FDR, se tendrá que hacer uso de complementos como lo sería una antena GPS, un cable de alimentación y un cable de Ethernet. Típicamente, la instalación tendría que resultar de la forma en que se muestra en la figura. Existen más especificaciones sobre la instalación y funcionamiento de un FDR, los cuales pueden ser consultados fácilmente en [38].

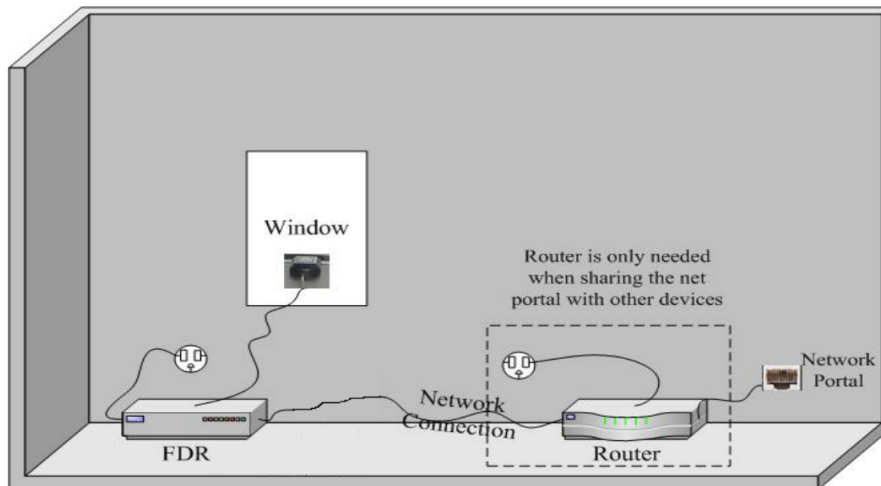


Figura 3.2. Ilustración de la instalación típica de un FDR. Extraído de [38]

Los datos de medición se transmiten continuamente a través de Internet al servidor FNET/GridEye alojado en la Universidad de Tennessee, habiendo un total de 80 de ellos conectados en Estados Unidos y el mundo. Se pueden observar los valores de las desviaciones de la frecuencia en tiempo real en el sitio web oficial del proyecto FNET/GridEye. Cabe aclarar que los FDR, incluido el FDR-UES, registra los datos en hora universal UTC.

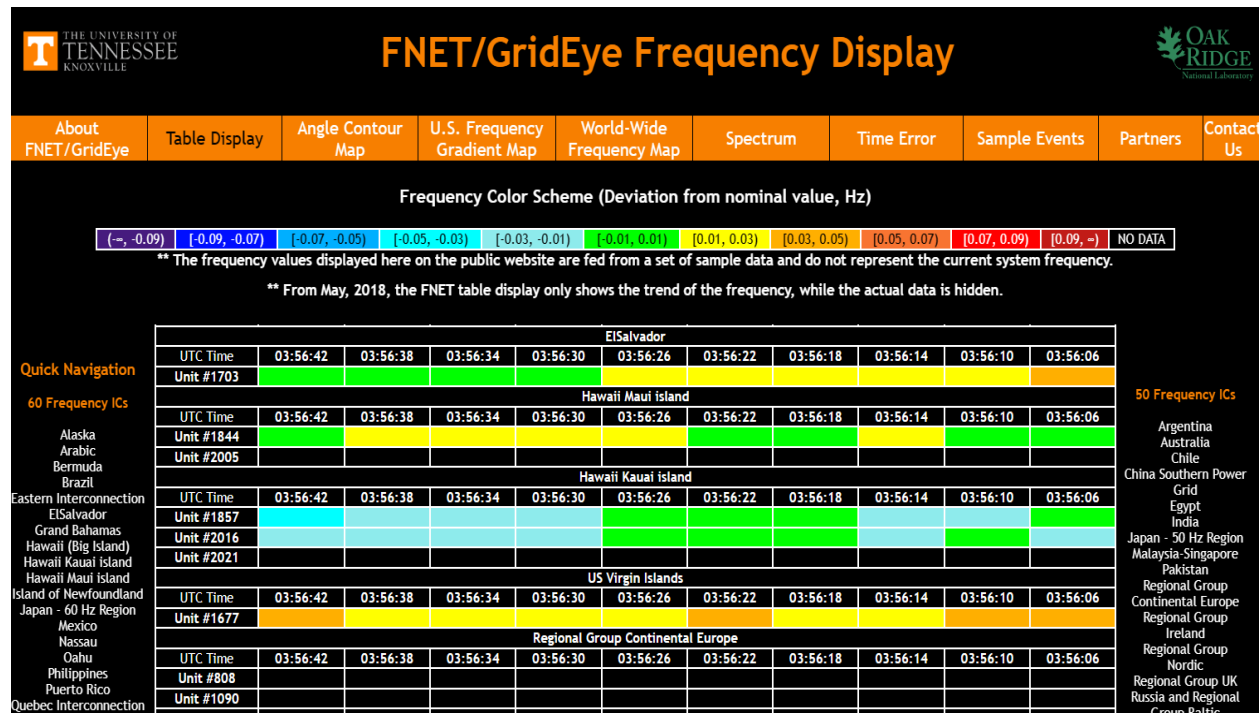


Figura 3.3. Captura de pantalla tomada en marzo del 2024, de las variaciones de frecuencia registradas por el FDR-UES. Disponible en [39].

Desde hace unos pocos años, la Escuela de Ingeniería Eléctrica inició su participación en el proyecto, gracias al donativo de uno de estos dispositivos por parte de la Universidad de Tennessee y la cooperación de la Universidad Nacional Autónoma de México. Primero se inició mediante una unidad FDR #1628, el cual tenía datos desde septiembre 2022 hasta junio 2023. Para obtener algunos de los datos registrados de los valores de la frecuencia, se realizó la solicitud para el acceso al servidor de la Universidad Nacional Autónoma de México para la respectiva descarga de datos registrados por la unidad.

Los datos han sido obtenidos en formato de valores separados por coma (.csv), y se muestran en la Figura 3.4. Agradecemos la ayuda del PhD. Mario Roberto Arrieta Paternina y el Ing. José Luis Zárate Gómez con la facilitación de estos.

```

1 Historian Data Viewer Export: 11/01/2022 00:00:00.000 to 11/30/2022 23:59:59.900
2
3 Average,NaN,60.0025452732045,NaN
4 Maximum,NaN,67.3886032104492,NaN
5 Minimum,NaN,57.1935997009277,NaN
6
7 Time,GPA_F-NET_UNIT1385:F F-Net Unit Frequency,GPA_F-NET_UNIT1628:F F-Net Unit 1628 Frequency,GPA_D-PMU_JUAREZ
8 01-nov.-2022 00:00:00.000,NaN,59.9943,NaN
9 01-nov.-2022 00:00:00.100,NaN,60.0028,NaN
10 01-nov.-2022 00:00:00.200,NaN,59.9975,NaN
11 01-nov.-2022 00:00:00.300,NaN,59.9981,NaN
12 01-nov.-2022 00:00:00.400,NaN,60.0014,NaN
13 01-nov.-2022 00:00:00.500,NaN,59.9977,NaN
14 01-nov.-2022 00:00:00.600,NaN,60.0061,NaN
15 01-nov.-2022 00:00:00.700,NaN,59.9976,NaN
16 01-nov.-2022 00:00:00.800,NaN,60.0045,NaN
17 01-nov.-2022 00:00:00.900,NaN,60,NaN
18 01-nov.-2022 00:00:01.000,NaN,60.0008,NaN
19 01-nov.-2022 00:00:01.100,NaN,60.0041,NaN
20 01-nov.-2022 00:00:01.200,NaN,59.9983,NaN
21 01-nov.-2022 00:00:01.300,NaN,60.0071,NaN
22 01-nov.-2022 00:00:01.400,NaN,59.9983,NaN
23 01-nov.-2022 00:00:01.500,NaN,60.0044,NaN
24 01-nov.-2022 00:00:01.600,NaN,60.001,NaN
25 01-nov.-2022 00:00:01.700,NaN,60.0003,NaN

```

Figura 3.4. Muestra de datos descargados del medidor FDR-UES

3.2. Almacenamiento de mediciones del FDR-UES #1703 desde marzo 2024.

El equipo utilizado para la recolección de datos es un CPU marca DELL tipo OptiPlex 7040SFF modelo D11S, al cual se le ha incorporado una antena Wifi tipo DELL 0GJRH7 A00. El equipo brindado pertenece a la Escuela de Ingeniería Eléctrica, y además de ser usado para almacenamiento de los datos recopilados por la unidad #1703 u otras unidades, también será útil a la hora de ejecutar códigos para la realización del análisis estadístico.



Figura 3.5. CPU marca DELL, tipo OptiPlex 7040SFF, utilizado para el almacenamiento de registros

Haciendo uso de varios programas y configuraciones al equipo, se realizará el almacenamiento de los datos. Todas las instrucciones para seguir quedarán escritas en un **manual**, el cual servirá como documentación en caso de que tenga que hacer cambios al equipo o volverlo a configurar.

Las instrucciones y la documentación realizada han sido basada en base a documentación compartida por la Universidad Nacional Autónoma de México [40], [41], realizando las adecuaciones correspondientes para el equipo de la Escuela de Ingeniería Eléctrica.

A fecha de septiembre del 2024, el docente PhD. Carlos Pocasangre (carlos.pocasangre@ues.edu.sv) de la EIE está gestionando el proyecto y es el que posee la documentación respectiva. Para evitar cualquier problema, esta documentación, se ha optado por dejar de forma interna.

3.3. Solicitud de datos a la Unidad de Transacciones

El primer día de marzo del 2024, se procedió a hacer contacto a la Unidad de Transacciones con el objetivo de poder solicitar datos de frecuencia eléctrica y adicionales, con tal de:

- Poder complementar información vacía existente del FDR-1628, debido a mantenimientos o fallas en la transmisión de datos.
- Poder ampliar el periodo de estudio, para cubrir todo el año 2022 y 2023, y de esa forma poder identificar de mejor forma algunas tendencias existentes en la red eléctrica.
- De ser posible, poder realizar una correlación con los flujos de potencia.

El lunes 4 de marzo del 2024 se recibió respuesta por parte de un analista de sistemas de potencia en la Gerencia de Operación y Estudios de la Unidad de Transacciones, y posteriormente el viernes 8 de marzo se realizó una reunión para profundizar acerca de que consistía la investigación, y los alcances planteados. Finalmente, la solicitud fue aprobada, siendo proporcionados los datos desde el día 18 de marzo del 2024 hasta el 17 de abril del 2024.

Los datos solicitados y proporcionados fueron los siguientes:

- Datos de demanda, intercambios y frecuencia, los cuales fueron proporcionados en archivos csv y XML de forma diaria, ordenadas en carpetas de forma mensual y anual. De estos datos, optamos por hacer uso únicamente de los de formato csv.
- Datos de generación, los cuales fueron proporcionados en archivos csv según el tipo de generación, ordenados en carpetas en distintos periodos de tiempo.

Finalmente, solo se hizo uso de los datos de frecuencia, y se optó por hacer uso de reportes públicos del EOR para la correlación con los flujos de potencia. Para más información, ver el Apartado 3.4 y el Capítulo 6 de este reporte.

3.4. Utilización de reportes de eventos de anomalías en el SER (EOR).

El Ente Operador Regional (EOR) tiene en su portal web todos los reportes de eventos de anomalías ocurridos a nivel del Sistema Eléctrico Regional. En total, para el periodo de estudio, son 558 los reportes, los cuales fueron revisados uno por uno para extraer la información relevante.

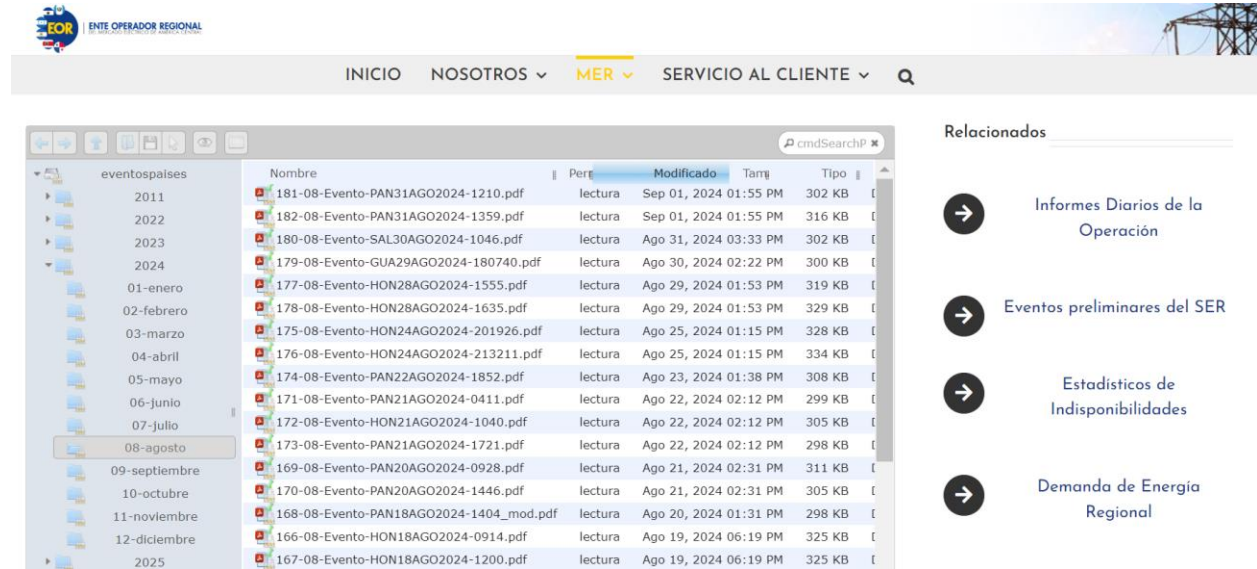


Figura 3.6. Reportes de eventos en el SER, disponibles en su sitio web para consulta/descarga.

ID	Time	Country	MW load	MW gen	Δf	SEM-SER
01-01-2022	01-01-2022 / 12:55	HON	0	62	-0.185	CLOSED
02-01-2022	04-01-2022 / 19:42	SLV	0.00	78.00	-0.02	CLOSED
03-01-2022	05-01-2022 / 15:03	HON	0	325.08	-0.849	OPEN
04-01-2022	08-01-2022 / 22:37	SLV	70.00	0.00	-0.017	CLOSED
05-01-2022	09-01-2022 / 11:31	HON	0	70	-0.09	CLOSED
06-01-2022	11-01-2022 / 12:33	PAN	0	72	-0.016	CLOSED
07-01-2022	12-01-2022 / 01:41	PAN	0	296	-0.85	OPEN
08-01-2022	16-01-2022 / 13:30	PAN	149.6	0	0.158	Maintenance
09-01-2022	16-01-2022 / 16:08	HON	460	483	0.135	Maintenance
10-01-2022	18-01-2022 / 18:17	SLV	0	80	-0.06	CLOSED
11-01-2022	18-01-2022 / 21:23	SLV	0	90	-0.122	CLOSED
12-01-2022	20-01-2022 / 20:28	PAN	200	0	0.092	CLOSED
13-01-2022	24-01-2022 / 04:40	CRI	0	60	-0.062	CLOSED
14-02-2022	01-02-2022 / 19:17	HON	0	122.7	-0.042	CLOSED
15-02-2022	01-02-2022 / 09:41	SLV	74.10	52.70	-0.053	CLOSED
16-02-2022	02-02-2022 / 05:50	PAN	0	137	-0.101	CLOSED
17-02-2022	02-02-2022 / 07:37	PAN	105	0	0.067	CLOSED
18-02-2022	05-02-2022 / 12:43	SER:SLV-HON			-0.6635	OPEN

Figura 3.7. Resumen preliminar de información relevante extraído para cada evento de pérdida de generación/carga ocurrido en el SER, realizado en hoja de Excel.

3.5. Acerca del manejo de los datos.

Para procesar todos los datos existentes, se hará uso de lenguajes de programación de uso libre y gratuito, destacando el uso de Python y los cuadernos de Jupyter que permitan organizar de mejor forma la información. Para algunas necesidades específicas, se podría recurrir a otros lenguajes como lo es R, o también se podría hacer uso de otras herramientas, pero su uso será muy menor a comparación de los cuadernos de Jupyter y Python. Los scripts que sean realizados se almacenarán en un repositorio de GitHub.

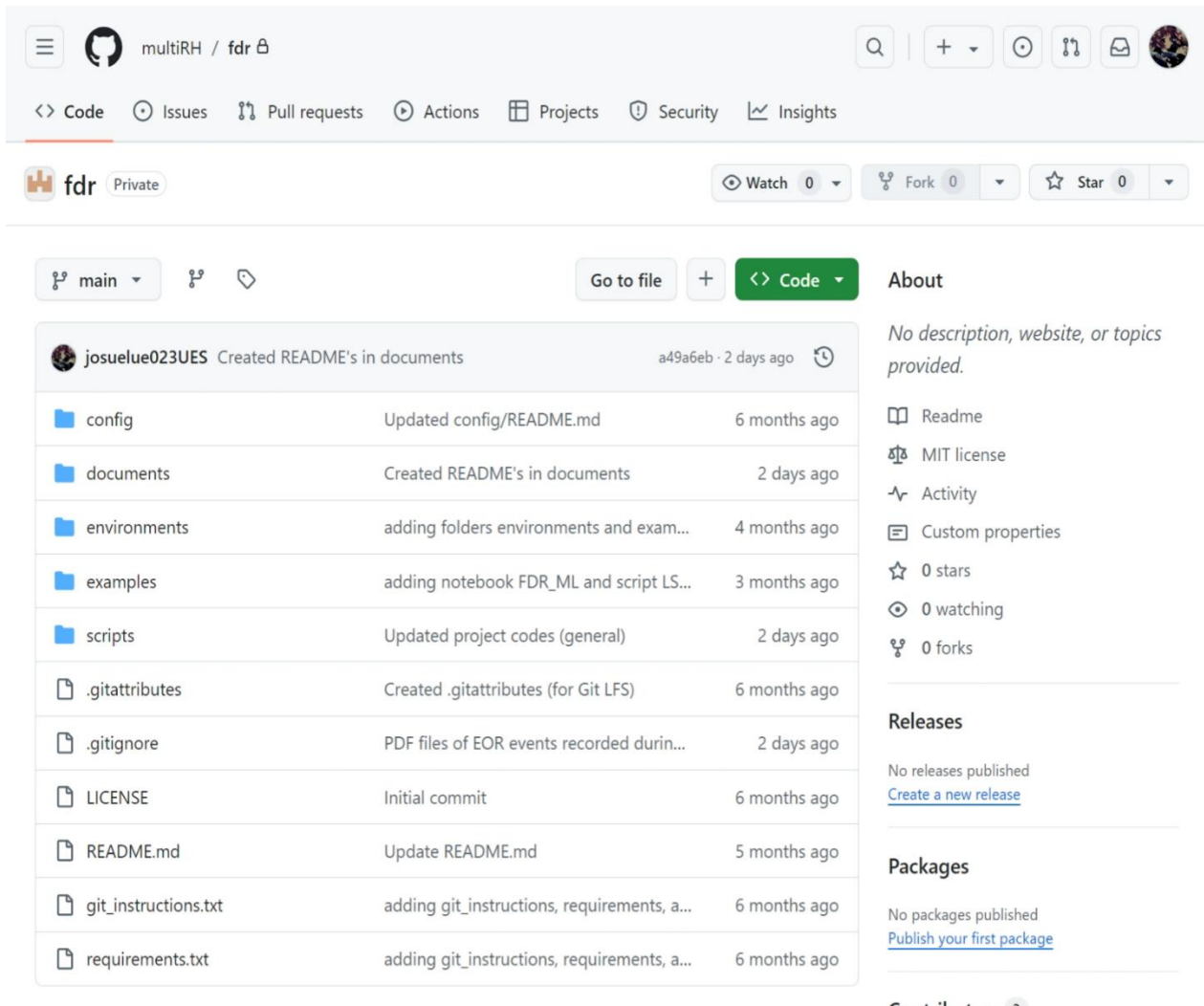


Figura 3.8. Repositorio de GitHub del proyecto.

Para la realización de futuros estudios, estos scripts realizados pueden ser reutilizados y mejorados. Inclusive, si se desea, estos scripts pueden hacerse públicos sin ningún inconveniente.

Los scripts realizados se pueden dividir en varios bloques, entre los que podemos enlistar:

1. Creación de set de datos (limpieza y organización de datos).
2. Realización de análisis descriptivo y de tendencias.
3. Realización de análisis correlativo.
4. Realización de modelo predictivo.
5. Otras tareas generales.

3.5.1. Creación de set de datos.

Se ha realizado dos procesos distintos para la creación de los sets de datos: el primer proceso es con los datos de los medidores #1628 y #1703 de la Escuela de Ingeniería Eléctrica, mientras que el segundo proceso es con los datos proporcionados por la Unidad de Transacciones del SCADA.

Para ambos casos, se ha optado por la creación de archivos mensuales, con el fin de poder manejar de forma más óptima los datos, y para su identificación se ha seguido el formato:

[Measurer ID] - Month Year.csv

A su vez, se han dividido los archivos en carpetas que corresponden a cada año.













Nombre	Fecha de modificación	Tipo	Tamaño
 [SCADA UT] - April 2022.csv	18/6/2024 20:42	Archivo CSV	81,970 KB
 [SCADA UT] - August 2022.csv	18/6/2024 21:31	Archivo CSV	84,879 KB
 [SCADA UT] - December 2022.csv	18/6/2024 22:20	Archivo CSV	84,440 KB
 [SCADA UT] - February 2022.csv	18/6/2024 20:17	Archivo CSV	77,040 KB
 [SCADA UT] - January 2022.csv	18/6/2024 20:05	Archivo CSV	85,229 KB
 [SCADA UT] - July 2022.csv	18/6/2024 21:18	Archivo CSV	84,663 KB
 [SCADA UT] - June 2022.csv	18/6/2024 21:06	Archivo CSV	82,097 KB
 [SCADA UT] - March 2022.csv	18/6/2024 20:30	Archivo CSV	85,553 KB
 [SCADA UT] - May 2022.csv	18/6/2024 20:54	Archivo CSV	84,242 KB
 [SCADA UT] - November 2022.csv	18/6/2024 22:07	Archivo CSV	81,782 KB
 [SCADA UT] - October 2022.csv	18/6/2024 21:55	Archivo CSV	83,563 KB
 [SCADA UT] - September 2022.csv	18/6/2024 21:43	Archivo CSV	81,189 KB

Figura 3.9. Set de datos creados del SCADA UT, del año 2022.

A. Proceso de creación de set de datos para los medidores FDR-1628 y FDR-1703.

Los datos puros, extraídos del medidor FDR #1628, están de la siguiente forma, en archivos csv.

Nombre	Fecha de modificación	Tipo	Tamaño
01ene-31ene2023.csv	13/6/2023 11:52	Archivo CSV	1,116,300 ...
01nov-30nov2022.csv	12/6/2023 17:33	Archivo CSV	1,079,274 ...
23feb-25marzo2023.csv	14/6/2023 04:38	Archivo CSV	1,077,980 ...
02oct-31oct2022.csv	12/6/2023 07:13	Archivo CSV	1,048,884 ...
junio2023.csv	2/7/2023 20:43	Archivo CSV	1,038,594 ...
25marzo - 30abril2023.csv	14/6/2023 12:03	Archivo CSV	923,856 KB
20abril-15mayo2023.csv	14/6/2023 19:24	Archivo CSV	874,528 KB
01dic-24dic2022.csv	13/6/2023 00:16	Archivo CSV	832,416 KB
01feb-23feb2023.csv	13/6/2023 19:13	Archivo CSV	803,580 KB
15may-31may2022.csv	15/6/2023 03:55	Archivo CSV	540,872 KB

Figura 3.10. Datos fuente de mediciones realizadas por el FDR-1628 desde el 02 de septiembre del 2022 hasta el 22 de junio del 2023.

Cada uno de esos archivos tiene una estructura similar a la de la siguiente figura.

```

01feb-23feb2023.csv
1 Historian Data Viewer Export: 02/01/2023 00:00:00.000 to 02/28/2023 23:59:59.900
2
3 Average,NaN,60.0026289753721,NaN
4 Maximum,NaN,66.3295974731445,NaN
5 Minimum,NaN,55.5793991088867,NaN
6
7 Time,GPA_F-NET_UNIT1385:F F-Net Unit Frequency,GPA_F-NET_UNIT1628:F F-Net Unit 1628 Frequency,GPA_D-PMU
8 01-feb.-2023 00:00:00.000,NaN,NaN,NaN
9 01-feb.-2023 00:00:00.100,NaN,NaN,NaN
10 01-feb.-2023 00:00:00.200,NaN,NaN,NaN
11 01-feb.-2023 00:00:00.300,NaN,NaN,NaN
12 01-feb.-2023 00:00:00.400,NaN,NaN,NaN
13 01-feb.-2023 00:00:00.500,NaN,NaN,NaN
14 01-feb.-2023 00:00:00.600,NaN,NaN,NaN
15 01-feb.-2023 00:00:00.700,NaN,NaN,NaN
16 01-feb.-2023 00:00:00.800,NaN,NaN,NaN
17 01-feb.-2023 00:00:00.900,NaN,NaN,NaN
18 01-feb.-2023 00:00:01.000,NaN,NaN,NaN
19 01-feb.-2023 00:00:01.100,NaN,NaN,NaN
20 01-feb.-2023 00:00:01.200,NaN,NaN,NaN
21 01-feb.-2023 00:00:01.300,NaN,NaN,NaN
22 01-feb.-2023 00:00:01.400,NaN,NaN,NaN
23 01-feb.-2023 00:00:01.500,NaN,NaN,NaN
24 01-feb.-2023 00:00:01.600,NaN,NaN,NaN
25 01-feb.-2023 00:00:01.700,NaN,NaN,NaN

```

Figura 3.11. Archivo csv de datos proporcionados, correspondiente a 01feb-23feb2023.csv

Existen algunos detalles que inicialmente se pueden destacar:

- Existen columnas que no corresponden al medidor de interés, y todas están llenas de valores NaN. Esto se debió a un error que existió con la generación de los archivos csv.
- A primera vista, es difícil identificar cuáles son los datos de interés, y se utilizan nombres que podrían ser extensos y pueden llegar a confundir.
- La fecha está en un formato no estándar, y algunas están duplicadas.
- Existen errores de medición de frecuencia.

Además de ello, agreguemos que los datos a primera vista están totalmente desordenados, pues no existe una periodización de todos los datos. Para ello, se ha mencionado que se ordenarán de forma mensual por cuestión de rendimiento y practicidad.

La idea es simple: Se enlistarán los archivos csv existentes en la carpeta donde se almacenan los registros del FDR-UES, y se leerán uno por uno. Para cada archivo que se lea, se observarán las fechas, detectando los meses y escribiendo los datos de cada mes en archivos separados.

Por ejemplo, para el archivo identificado de 20abril-15mayo2023:

1. Se detecta que existen datos de dos meses: abril y mayo del 2023.
2. Los datos existentes del 20 al 30 de abril del 2023 se adjuntan a un archivo csv que almacena los datos del mes de abril 2023
3. Los datos existentes del 1 al 15 de mayo se adjuntan a un archivo csv que almacena los meses de mayo del 2023.

Y ese mismo proceso se realizará para todos los archivos. Cabe aclarar que los datos están originalmente en horario UTC, por lo que se tienen que cambiar de una vez a la hora local salvadoreña (UTC-6), o de la región respectiva.

Sin embargo, este proceso no termina de completar, pues los archivos recién generados tendrán que pasar por un filtro el cual ordene las fechas, elimine datos duplicados e incoherentes. Mucho de esto se realizará antes de escribir los nuevos archivos csv.

Con respecto a los errores de medición, se aplicará un algoritmo básico para verificar si un valor puede considerarse válido: se verificarán cinco muestras antes y cinco muestras posteriores si

existen desviaciones de frecuencia de la misma magnitud. Este algoritmo aplicará a partir de desviaciones de 0.06 Hz en adelante, y consiste en que:

$$f_{tol} = (1 - tolerance) \times \Delta f_{i,60}$$

Donde, para una muestra i , se tendrá que $\Delta f_{i,60} = f_i - 60$ Hz.

El valor del parámetro ‘tolerancia’ se ha establecido en 0.50, después de la realización de varias pruebas. Para conservar un valor atípico f_k , deberá de cumplirse con que:

- Si $f_k < 60$ Hz, deberá existir al menos un valor entre f_{k-5} y f_{k+5} que sea menor a f_{tol} .
- Si $f_k > 60$ Hz, deberá existir al menos un valor entre f_{k-5} y f_{k+5} que sea mayor a f_{tol} .

Mientras mayor sea el valor del parámetro de tolerancia, más permisivo llega a ser el algoritmo al permitir el conservar valores de medición (por el término « $1 - tolerancia$ »). Si tenemos una tolerancia de 0.25, la desviación de la otra medición deberá de ser por lo menos un 75% de la referencia dentro de la ventana establecida.

Para considerar que $f_k = 59$ Hz es un valor valido y no un error de medición, se requiere que dentro de la ventana exista un valor igual o menor a 59.50 Hz para que este no sea descartado, considerando una tolerancia de 0.50.

Finalmente, si visualizamos la frecuencia eléctrica, se observa que tiene un montón de variaciones, que podemos interpretar como ruido, el cual puede llegar a afectar algunos parámetros.

Para minimizar el ruido podemos plantear hacer uso de las siguientes tres herramientas: Transformada Rápida de Fourier, Transformada de Wavelet Discreta, y Promedio Móvil. Se optó por graficar los siguientes dos momentos:

1. Entre las 13:15:30 hasta las 13:16:00 del 08 de febrero del 2023 (UTC-6).
2. Entre las 09:16:49 hasta las 09:16:54 del 14 de febrero del 2023 (UTC-6).

Comparison of Noise Reduction Methods for Signal
Date: February 08, 2023

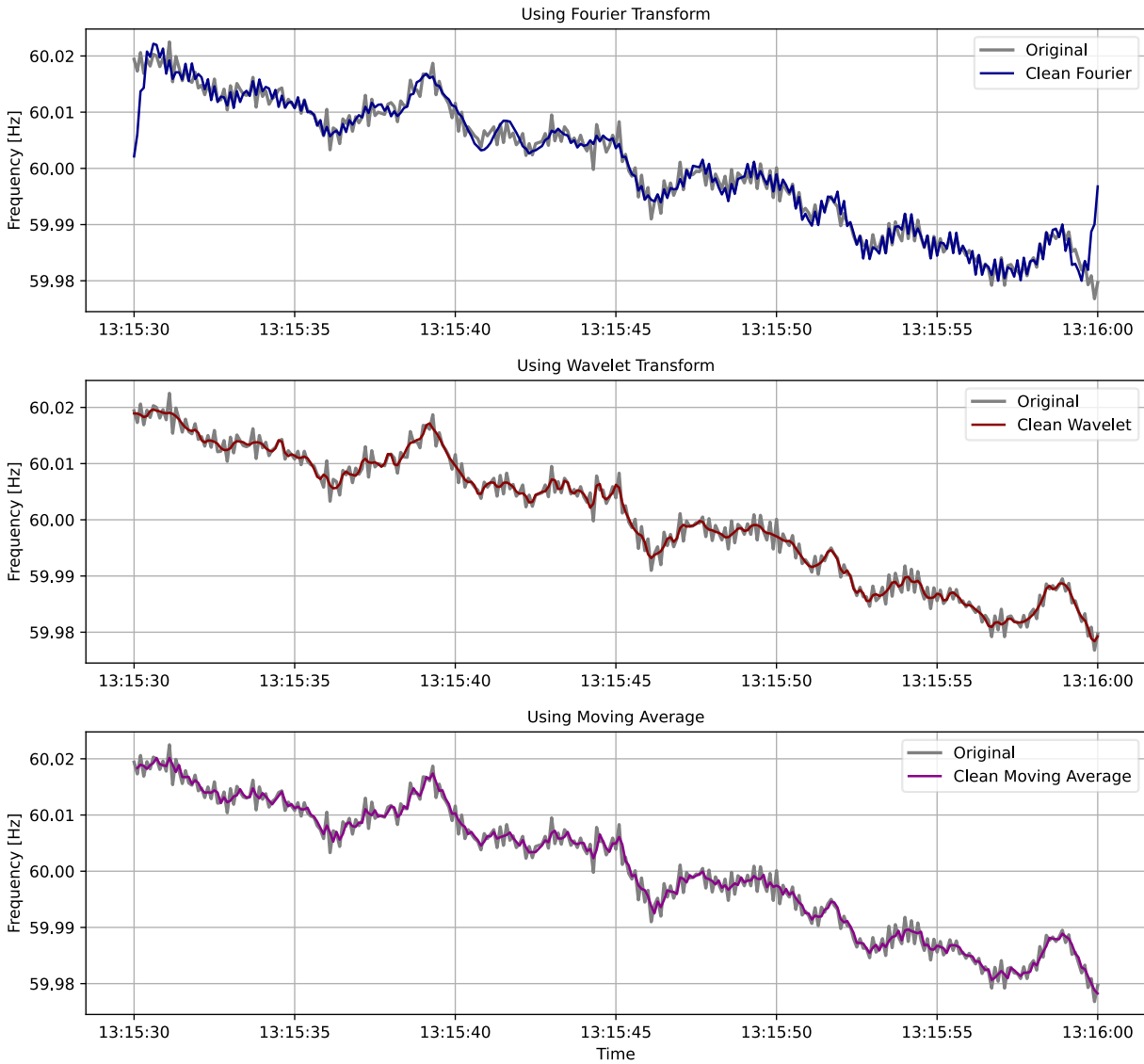


Figura 3.12. Limpieza de ruido de la frecuencia eléctrica aplicada la Transformada de Fourier, Transformada de Wavelet y Promedio Móvil. Fecha: 08 de febrero del 2023.

Comparison of Noise Reduction Methods for Signal
Date: February 14, 2023

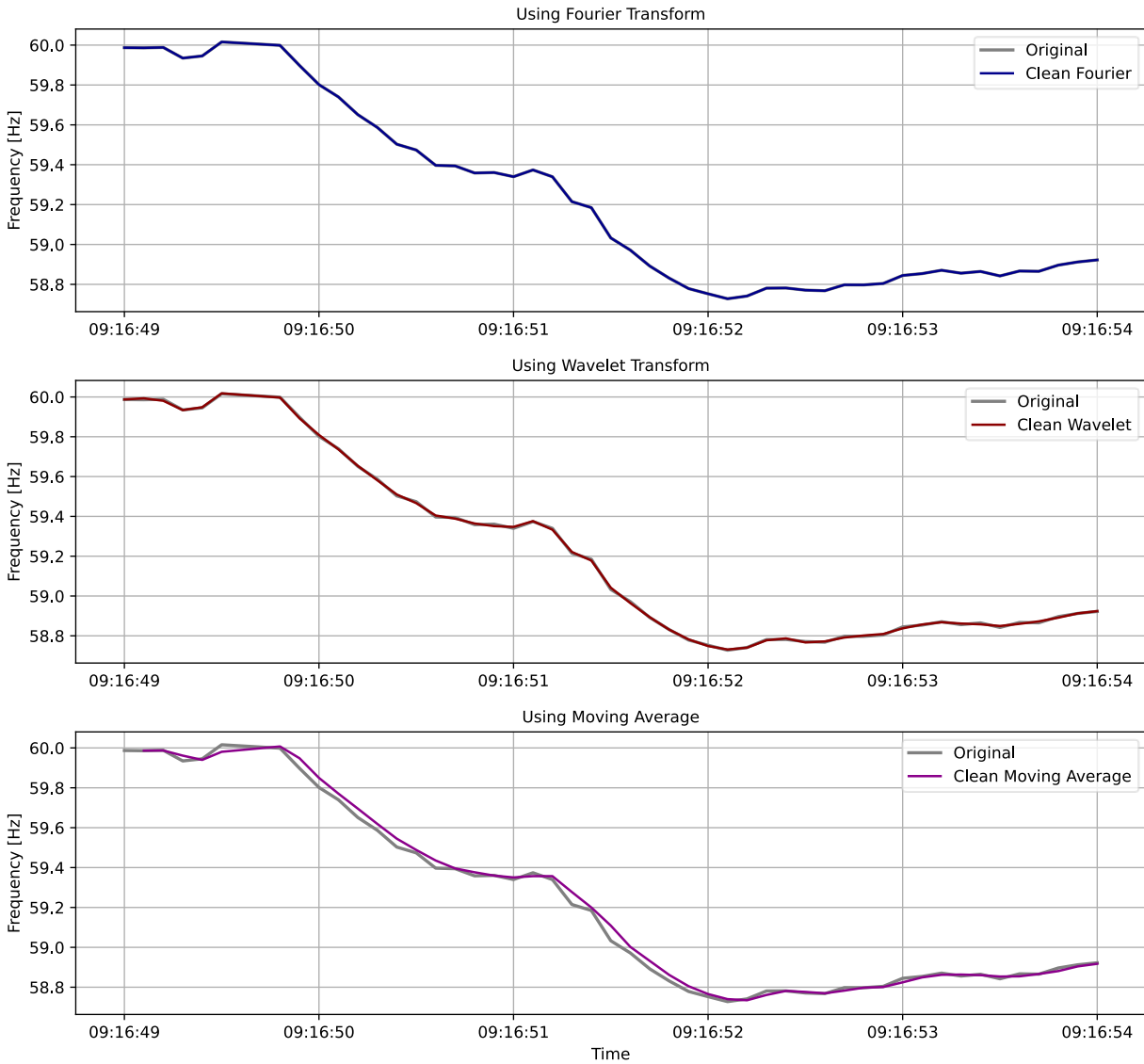


Figura 3.13. Limpieza de ruido de la frecuencia eléctrica aplicada la Transformada de Fourier, Transformada de Wavelet y Promedio Móvil. Fecha: 14 de febrero del 2023.

De las tres opciones planteadas, la mejor es la transformada de Wavelet, y por ende **es la que se procederá a aplicar**. No profundizaremos mucho sobre el funcionamiento de las transformadas en general, pero daremos las bases generales. La Transformada de Fourier proporciona información solo en el dominio de la frecuencia, lo que significa que no puede capturar cambios en la señal a lo largo del tiempo; mientras que la Transformada de Wavelet proporciona información tanto en el dominio de la frecuencia como en el dominio del tiempo [42].

Podemos interpretar que la Transformada de Wavelet se comportaría como una “Transformada de Fourier Móvil”. Mientras la Transformada de Fourier hace un análisis de la frecuencia **en general**, la Transformada de Wavelet hace este análisis en una ventana [42]. Si observamos la Figura 3.12, se puede observar que para la Transformada de Fourier hay momentos en que existe más ruido, mientras que en otros momentos se nota que ha sido hasta suavizada. Mientras tanto, la Transformada de Wavelet no sufre de ello. Para el caso del promedio móvil, cuando se está en una condición normal puede llegar a ser una alternativa válida siempre y cuando se seleccione un buen tamaño de ventana. Sin embargo, y tal como se ve en la Figura 3.13, en eventos que existen grandes perturbaciones puede llegar a presentar problemas, mientras que la Transformada de Wavelet se puede observar que se ajusta mucho mejor a la señal.

Los procedimientos para eliminar el ruido emplean la transformada para alterar el dominio de la señal. En este nuevo dominio, se realizan diversas operaciones de coeficientes, que suelen implicar una umbralización y finalmente se aplica la transformada inversa para retornar la señal modificada al dominio del tiempo [42], [43], [44]. Hay varias fuentes en la web que pueden explicar toda esta teoría con más detalle. En Python, todo lo dicho se implementa así:

```
import pywt
import pandas as pd

def wavelet_denoise(df: pd.DataFrame, threshold = 0.01, wavelet = 'db4'):

    # Se aplica la transformada a una serie de tiempo, al valor de la
    # frecuencia. Se obtendrán coeficientes de aproximación (cA) y detalle (cD).
    (cA, cD) = pywt.dwt(df['Frequency'].values, wavelet)

    # Se aplica un umbral (0.01) a los coeficientes de detalle, siendo que si
    # son menores que el umbral seleccionado se considerarán ruido.
    cD_thresh = pywt.threshold(cD, threshold)

    # Se reconstruye la señal a partir de cA y los modificados cD.
    reconstructed_signal_wavelet = pywt.waverec([cA, cD_thresh], wavelet)

    # Por naturaleza de la función, se retomaría una lista con dimensión par.
    # De esta forma se garantiza que sea del mismo tamaño en caso de que sea impar.
    reconstructed_signal_wavelet = reconstructed_signal_wavelet[:len(df['Frequency'])]

    # Se crea un nuevo DataFrame con la señal con el ruido removido
    df_denoised = df.copy()
    df_denoised['Frequency'] = reconstructed_signal_wavelet

    # Se devuelve el nuevo dataframe.
    return df_denoised
```

B. Proceso de creación de set de datos para el SCADA UT

Los datos proporcionados por parte de la Unidad de Transacciones corresponden al periodo del 01 de enero del 2022 hasta el 31 de diciembre de 2023, y contiene lo siguiente:

- Valores de la frecuencia eléctrica.
- Valores de la demanda existente.
- Intercambios de flujos de potencia con Guatemala y Honduras, desde las subestaciones de Ahuachapán y 15 de septiembre.
- Datos de generación eléctrica por tecnología: eólica, hidroeléctrica, térmica, etc.
- Entre otros datos.

Los archivos, una vez descomprimidos, se tienen en dos carpetas bien identificadas.

 Datos Demanda, Intercambios, Frecuencia	Fecha de modificación: 7/4/2024 12:59
 Datos Generacion	Fecha de modificación: 17/4/2024 18:37

Figura 3.14. Carpetas de datos proporcionados por la Unidad de Transacciones S.A de C.V

B.1. Datos de demanda, intercambios y frecuencia.

Dentro de la carpeta «Datos Demanda, Intercambios, Frecuencia» se encuentran la información clasificada por años, incluyendo algunos datos de inicios del 2024.




 2022	Fecha de modificación: 7/4/2024 13:00
 2023	Fecha de modificación: 7/4/2024 13:01
 2024	Fecha de modificación: 7/4/2024 13:02

Figura 3.15. Carpetas anuales de «Datos Demanda, Intercambios, Frecuencia» proporcionados por la Unidad de Transacciones.

Si accedemos a una de las carpetas, nos encontraremos con que hay más datos proporcionados de forma mensual, para cada uno de los meses del año correspondiente. Finalmente, si accedemos a una de esas carpetas, nos encontraremos con un archivo csv por cada uno de los días del mes (además, existen archivos XML, los cuales no fueron utilizados).













Nombre	Fecha de modificación	Tipo	Tamaño
 SAL-ACEabr2022	3/5/2022 00:01	Carpeta de archivos	
 SAL-ACEago2022	1/9/2022 00:02	Carpeta de archivos	
 SAL-ACEdic2022	2/1/2023 00:21	Carpeta de archivos	
 SAL-ACEene2022	1/2/2022 00:06	Carpeta de archivos	
 SAL-ACEfeb2022	1/3/2022 00:05	Carpeta de archivos	
 SAL-ACEjul2022	2/8/2022 00:34	Carpeta de archivos	
 SAL-ACEjun2022	2/7/2022 00:02	Carpeta de archivos	
 SAL-ACEmar2022	3/4/2022 00:10	Carpeta de archivos	
 SAL-ACEmay2022	2/6/2022 00:03	Carpeta de archivos	
 SAL-ACEnov2022	2/12/2022 00:05	Carpeta de archivos	
 SAL-ACEoct2022	4/11/2022 00:13	Carpeta de archivos	
 SAL-ACEsep2022	1/10/2022 00:03	Carpeta de archivos	

Figura 3.16. Carpetas mensuales, correspondientes al año 2022, de «Datos Demanda, Intercambios, Frecuencia», proporcionados por la Unidad de Transacciones.











Nombre	Fecha de modificación	Tipo	Tamaño
 SAL-ACE01Abr22.csv	2/4/2022 00:05	Archivo CSV	1,384 KB
 SAL-ACE02Abr22.csv	3/4/2022 00:08	Archivo CSV	1,374 KB
 SAL-ACE03Abr22.csv	4/4/2022 00:03	Archivo CSV	1,408 KB
 SAL-ACE04Abr22.csv	5/4/2022 00:02	Archivo CSV	1,375 KB
 SAL-ACE05Abr22.csv	6/4/2022 00:04	Archivo CSV	1,381 KB
 SAL-ACE06Abr22.csv	7/4/2022 00:03	Archivo CSV	1,372 KB
 SAL-ACE07Abr22.csv	8/4/2022 00:02	Archivo CSV	1,407 KB
 SAL-ACE08Abr22.csv	9/4/2022 00:08	Archivo CSV	1,410 KB
 SAL-ACE09Abr22.csv	10/4/2022 00:02	Archivo CSV	1,420 KB
 SAL-ACE10Abr22.csv	11/4/2022 00:02	Archivo CSV	1,415 KB

Figura 3.17. Archivos csv para cada uno de los días del mes de mayo del 2022, de «Datos Generación, Intercambios, Frecuencia» proporcionados por la Unidad de Transacciones.

Teniendo en consideración todo lo anterior dicho, se procedió a realizar lo siguiente:

1. Se verifica si hay ausencia de archivos csv de cada día correspondiente.
2. Considerando que los datos tienen una tasa de muestreo de cuatro segundos, se tendrán en total 21 600 mediciones diarias. Se verificará si existe una cantidad mayor o menor, y se tomarán acciones:
 - a. Se eliminarán las mediciones duplicadas, conservando únicamente la original. En caso de que la fecha sea igual, pero la medición sea distinta, se reportará.
 - b. Se contabilizarán la cantidad de mediciones faltantes, con tal de saber si estos pueden afectar los resultados o ser despreciados.
 - c. Se verificará que las columnas y encabezados sean coherentes.

Mediante un script realizado en Jupyter Notebook, se verificaron todos los puntos anteriores, encontrando que los datos proporcionados de demanda, intercambios y frecuencia estaban ya limpios de origen.

B.2. Datos de generación.

Dentro de la carpeta «Datos Generación» se encuentran las carpetas de la siguiente forma. ‘ene22’, ‘ene23’, ‘abr23’ y ‘ene24’ representa el mes desde el cual inician los datos (no estrictamente inician en el día 01 de mes).

Nombre	Fecha de modificación	Tipo	Tamaño
abr23	17/4/2024 18:22	Carpeta de archivos	
ene22	17/4/2024 18:23	Carpeta de archivos	
ene23	17/4/2024 18:25	Carpeta de archivos	
ene24	17/4/2024 18:25	Carpeta de archivos	

Figura 3.18. Carpetas de «Datos Generación» proporcionados por la Unidad de Transacciones.

Dentro de cada una de las carpetas se encuentra un archivo csv por cada tipo de tecnología de generación. Sin embargo, ocurre que no están totalmente estandarizados todos los nombres. Por ejemplo, puede darse el caso de que exista un archivo ‘hidroelectrico.csv’ en una carpeta, mientras que en otra puede llamarse ‘hidro.csv’, por lo que se opta por estandarizar el nombre de todos los archivos csv.







Nombre	Fecha de modificación	Tipo	Tamaño
 eolico.csv	14/3/2024 11:37	Archivo CSV	177,664 KB
 geotermico.csv	14/3/2024 11:40	Archivo CSV	172,456 KB
 hidroelectrico.csv	17/4/2024 18:40	Archivo CSV	172,051 KB
 ingenios.csv	16/4/2024 15:09	Archivo CSV	179,513 KB
 solar.csv	14/3/2024 11:49	Archivo CSV	180,239 KB
 termico.csv	16/4/2024 16:17	Archivo CSV	173,210 KB

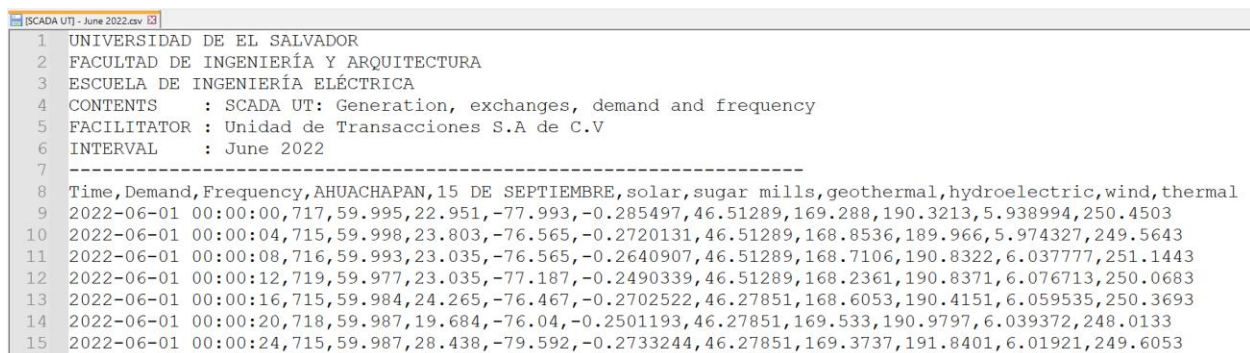
Figura 3.19. Archivos csv de generación, de «Datos Generación» proporcionados por la Unidad de Transacciones.

Teniendo en consideración todo lo anterior dicho, se procedió a realizar lo siguiente:

1. Se verifica si hay ausencia de archivos csv de cada generación correspondiente
2. Se procedió a verificar si existían datos duplicados, conservando únicamente el original.
3. Haciendo uso de un muestreo con los datos de generación existentes en <https://estadistico.ut.com.sv/OperacionDiaria.aspx>, se realizó una verificación manual para verificar que los datos de cada generación fueran correctos.

Mediante un script realizado en Jupyter Notebook, se verificaron todos los puntos anteriores, encontrando algunas incoherencias. Estas incoherencias fueron comentadas a la Unidad de Transacciones, y fueron minimizadas con éxito.

Finalmente, estos datos se ordenan de forma mensual y se unifican en un mismo archivo csv.



```

1 UNIVERSIDAD DE EL SALVADOR
2 FACULTAD DE INGENIERÍA Y ARQUITECTURA
3 ESCUELA DE INGENIERÍA ELÉCTRICA
4 CONTENTS : SCADA UT: Generation, exchanges, demand and frequency
5 FACILITATOR : Unidad de Transacciones S.A de C.V
6 INTERVAL : June 2022
7 -----
8 Time, Demand, Frequency, AHUACHAPAN, 15 DE SEPTIEMBRE, solar, sugar mills, geothermal, hydroelectric, wind, thermal
9 2022-06-01 00:00:00, 717, 59.995, 22.951, -77.993, -0.285497, 46.51289, 169.288, 190.3213, 5.938994, 250.4503
10 2022-06-01 00:00:04, 715, 59.998, 23.803, -76.565, -0.2720131, 46.51289, 168.8536, 189.966, 5.974327, 249.5643
11 2022-06-01 00:00:08, 716, 59.993, 23.035, -76.565, -0.2640907, 46.51289, 168.7106, 190.8322, 6.037777, 251.1443
12 2022-06-01 00:00:12, 719, 59.977, 23.035, -77.187, -0.2490339, 46.51289, 168.2361, 190.8371, 6.076713, 250.0683
13 2022-06-01 00:00:16, 715, 59.984, 24.265, -76.467, -0.2702522, 46.27851, 168.6053, 190.4151, 6.059535, 250.3693
14 2022-06-01 00:00:20, 718, 59.987, 19.684, -76.04, -0.2501193, 46.27851, 169.533, 190.9797, 6.039372, 248.0133
15 2022-06-01 00:00:24, 715, 59.987, 28.438, -79.592, -0.2733244, 46.27851, 169.3737, 191.8401, 6.01921, 249.6053

```

Figura 3.20. Set de datos del SCADA-UT, mensualizados, limpiados y unificados.

3.5.2. Validación de datos del FDR-1628 y el SCADA.

Una de las mayores inquietudes que puede existir es si los datos de la unidad FDR-UES coinciden con los registros que se tiene por parte de la Unidad de Transacciones. Para ello, se han realizado:

1. Una comprobación grafica.
2. El cálculo de la diferencia existente entre las mediciones del SCADA y el FDR-UES.
3. Una prueba estadística de hipótesis.

Para la realización de estas pruebas, se hará uso únicamente de una pequeña muestra (un mes), la cual consideramos que sería suficiente para corroborar la igualdad de los datos.

A. Mediante comprobaciones graficas.

Se optó por graficar los siguientes dos momentos:

1. En condición normal, entre las 12:16 hasta las 12:20 del 19 de febrero del 2023 (UTC-6).
2. En condición crítica, entre las 09:16 hasta las 09:20 del 14 de febrero del 2023 (UTC-6).

Se procede a realizar el grafico para la condición normal, tanto con los datos del FDR-UES en bruto como remuestreados a cuatro segundos.

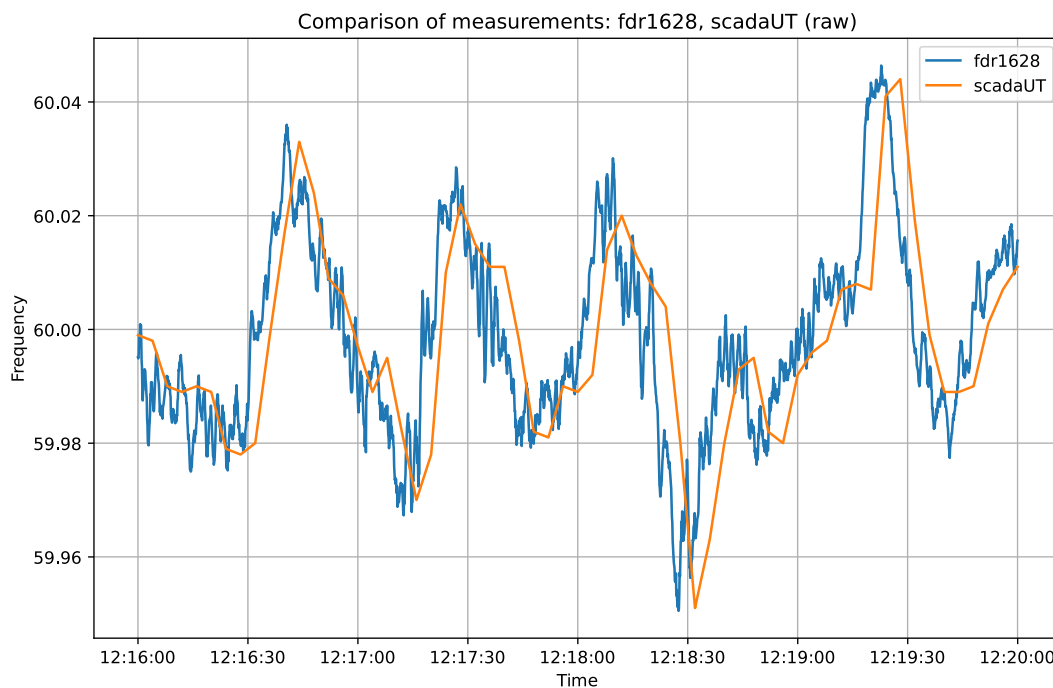


Figura 3.21. Comparación de mediciones: FDR-UES 1628 (datos brutos) vs SCADA UT, en condición normal, entre las 12:16 hasta las 12:20 del 19 de febrero del 2023 (UTC-6).

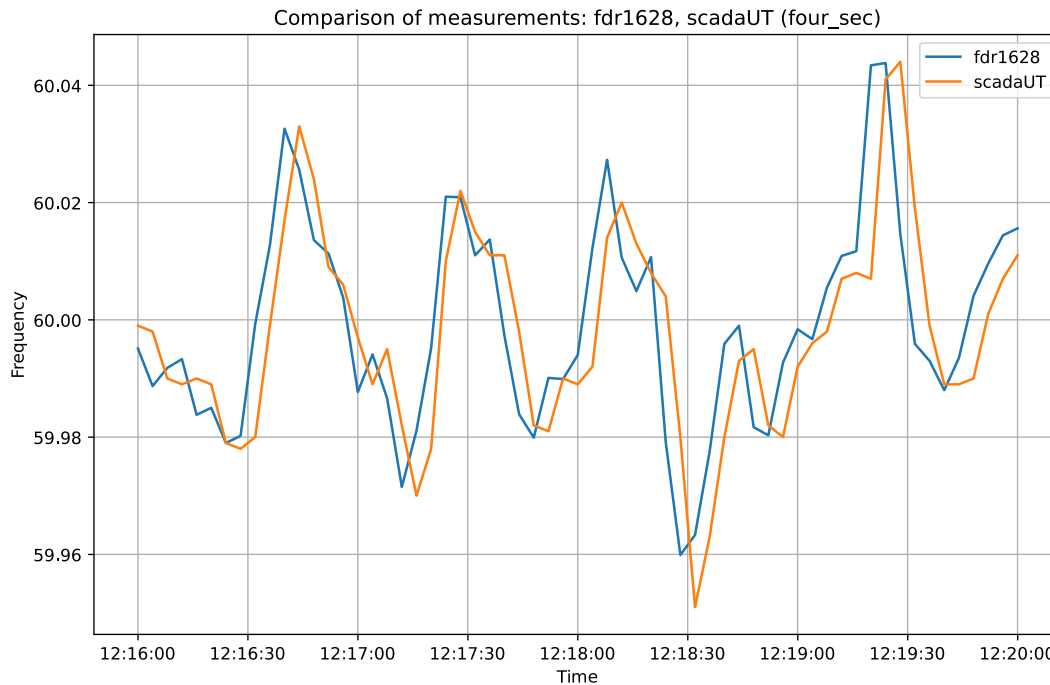


Figura 3.22. Comparación de mediciones: FDR-UES 1628 (remuestreado a cuatro segundos) vs SCADA UT, en condición normal, entre las 12:16 hasta las 12:20 del 19 de febrero del 2023 (UTC-6).

Es evidente que las mediciones del FDR-UES coinciden con la tendencia que marca el sistema SCADA de la Unidad de Transacciones. Si optamos por conservar únicamente las mediciones múltiplos de cuatro segundos del FDR-UES, se puede apreciar de mejor forma que existe un cierto atraso en las mediciones del SCADA.

Se ha optado por estimar, aproximadamente, el tiempo de atraso. Para ello, mediante Python, calculamos el tiempo en que el valor del SCADA y de FDR-1628 (remuestreado) es máximo y mínimo, y se saca la diferencia. El resultado finalmente **fue de cuatro segundos**, por lo que podemos considerar que no será una diferencia que llegue a dar grandes variaciones a los resultados.

Si observamos los gráficos durante un evento de perturbación crítico, podemos observar que (a pesar del atraso del SCADA), las tendencias coinciden.

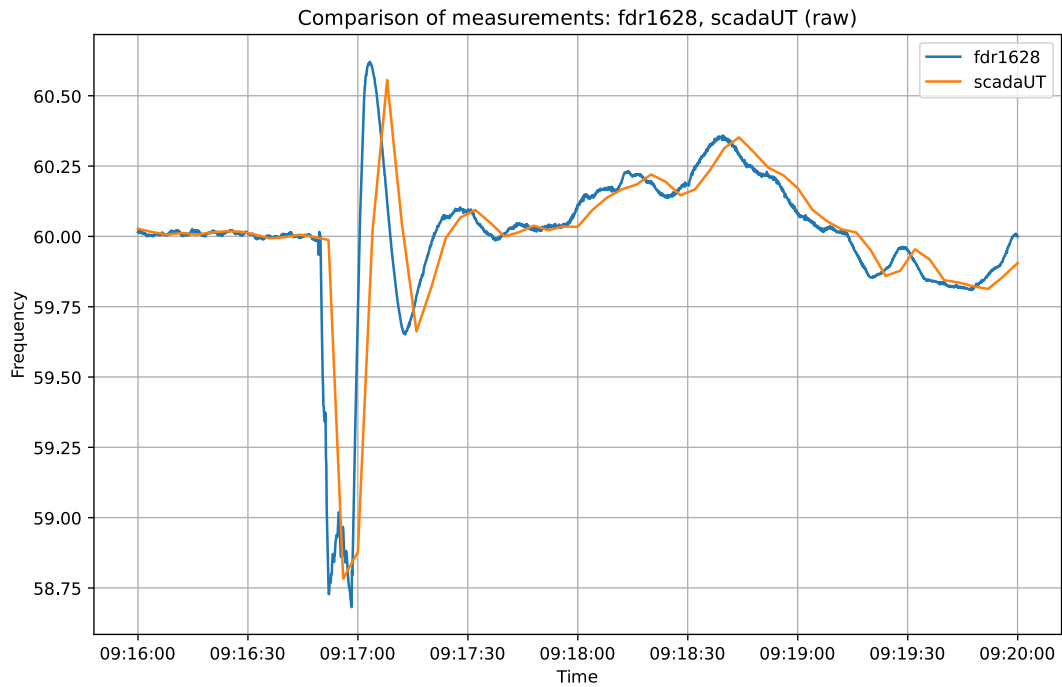


Figura 3.23. Comparación de mediciones: FDR-UES 1628 (datos brutos) vs SCADA UT, en condición crítica, entre las 09:16 hasta las 09:20 del 14 de febrero del 2023 (UTC-6).

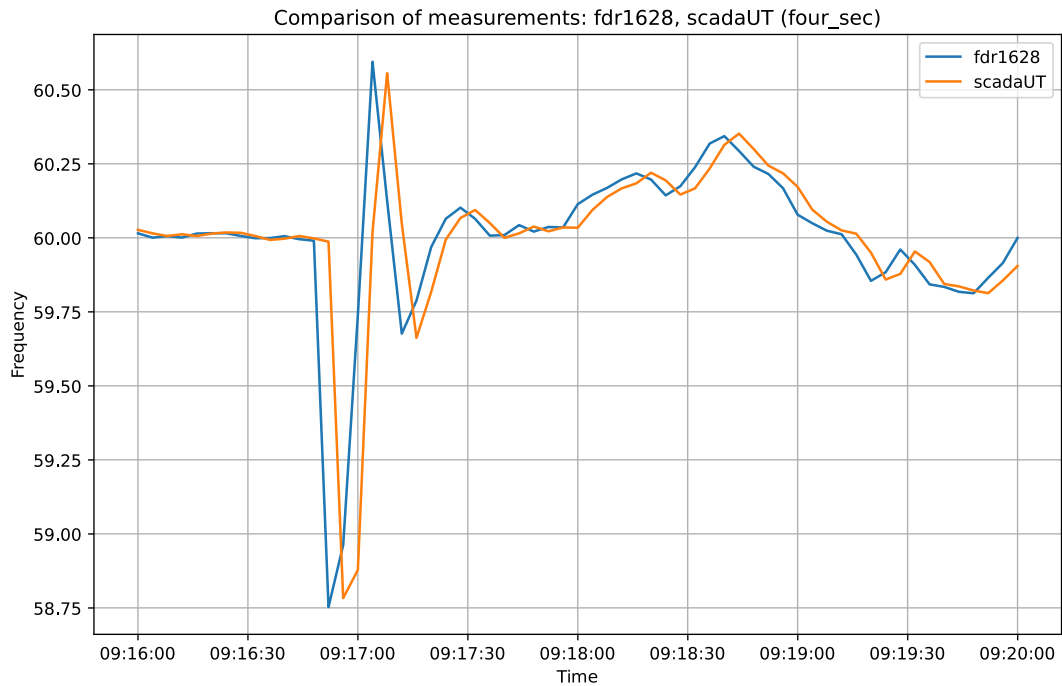


Figura 3.24. Comparación de mediciones: FDR-UES 1628 (remuestreado a cuatro segundos) vs SCADA UT, en condición crítica, entre las 09:16 hasta las 09:20 del 14 de febrero del 2023 (UTC-6).

B. Cálculo de la diferencia existente entre las mediciones

Primero, definamos los dos parámetros a utilizar:

1. *Root Mean Square Error (RMSE)*: Métrica que nos indica la raíz cuadrada de la diferencia cuadrática media entre las mediciones [45]. El RMSE son sensibles a grandes errores debido a la cuadratura de los residuos [46]. Este parámetro se calcula de la siguiente forma:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (f_{SCADA_i} - f_{FDR_i})^2}{n}}$$

2. *Mean Absolute Error (MAE)*: Métrica que nos indica la diferencia media absoluta entre las mediciones [45]. El MAE no es sensible a los valores atípicos, por lo que puede dar una imagen más clara de las diferencias comúnmente existentes, ya que es mucho más sencillo de poder entender [46]. Este parámetro se calcula de la siguiente forma:

$$MAE = \frac{\sum_{i=1}^n |f_{SCADA_i} - f_{FDR_i}|}{n}$$

Se procedió a calcular ambos parámetros para la muestra seleccionada. Cabe aclarar que se optó por remuestrear los datos del FDR-UES a una tasa de cuatro segundos (misma del SCADA). Finalmente, los resultados son los siguientes:

- RMSE = 0.0125 Hz
- MAE = 0.0093 Hz

Los valores obtenidos pueden ser considerados aceptables, en especial si tenemos en consideración que los datos del SCADA están ligeramente atrasados con respecto a los del FDR-UES. Pero si quitamos el atraso de los datos de SCADA, entonces se obtienen mejores valores:

- RMSE = 0.0054 Hz
- MAE = 0.0038 Hz

Y al ser RMSE bastante cercano al MAE, se puede considerar que no existen grandes diferencias atípicas entre un medidor y otro. Téngase en consideración que estamos hablando de distintos tipos de medidores, un PMU (al cual se remuestreo para esta prueba) y un SCADA, por lo que las diferencias existentes en las mediciones siempre serán algo normal.

C. Prueba de hipótesis estadística y prueba T para muestras independientes.

Para este subapartado, tenemos la necesidad de adelantar algunos conceptos y resultados que se obtendrán durante el siguiente capítulo. Por “orden narrativo” hemos optado por hacer la prueba de hipótesis estadística en este punto.

Para poder hacer uso de la prueba T, necesitamos suponer que la distribución existente es normal, o puede ser aproximada a una. En el siguiente capítulo llegamos a la conclusión de que la distribución no es totalmente normal, pero sí nos podemos tomar la libertad de aproximarla.

Las pruebas T se utilizan para determinar si existe una diferencia significativa entre las medias de dos variables y nos permite saber si pertenecen a la misma distribución [47], [48], [49]. La más común es la prueba t-student, sin embargo, tiene la desventaja de asumir que las desviaciones estándar son iguales [47], aspecto que en este punto no conocemos. Entonces se optará por hacer uso de una adaptación, que estaría siendo la prueba t-Welch’s, el cual no tiene este inconveniente y para estos casos es más fiable [50]. Mediante Python, la prueba se puede realizar de forma sencilla mediante la siguiente línea de la librería SciPy [48]:

```
# Realizando la prueba t
from scipy import stats
t_stat, p_val = stats.ttest_ind(data1, data2, equal_var=False)
```

Teniendo en consideración lo expuesto hasta ahora, podemos plantear una hipótesis estadística. Una hipótesis estadística es una aseveración o conjetura respecto a una o varias poblaciones o conjunto de datos. La verdad o falsedad de una hipótesis estadística nunca se sabe con absoluta certeza, a menos que se examine toda la población [47]. La estructura de la prueba de hipótesis se establece usando el término hipótesis nula, el cual se refiere a cualquier hipótesis que se desea probar y se denota con H_0 , y el rechazo de H_0 conduce a la aceptación de una hipótesis alternativa, que se denota con H_1 [47]. Para nuestro, nosotros estableceremos como hipótesis nula que existe igualdad en las medias de las mediciones reportados con el SCADA y los reportados por el FDR. La hipótesis alternativa sería lo contrario, negar la igualdad.

$$H_0: \bar{x}_{SCADA} = \bar{x}_{FDR}$$

$$H_1: \bar{x}_{SCADA} \neq \bar{x}_{FDR}$$

Finalmente, queda definir el nivel de significancia α con el que vamos a trabajar. Este parámetro es la probabilidad de rechazar la hipótesis nula cuando en realidad es verdadera [47].

El rechazar la hipótesis nula, siendo que esta es verdadera, sería bastante crítico para la investigación. Un nivel de significación de 0.05 es el más común, sin embargo, es responsabilidad del analista determinar cuánta evidencia requerir para concluir que se rechace la hipótesis nula [51]. En nuestro caso, estableceremos nuestro nivel de significancia a 0.01 (1.0%), ya que se desea tener una seguridad absoluta de los resultados. El escoger un valor de 0.01 es una práctica muy común en varios campos, como lo sería en la medicina o ingeniería.

Finalmente, se realiza la prueba t-Welch's, dando como resultado 0.0238. Dado que es superior a nuestro valor establecido de $\alpha = 0.01$, entonces conservamos la hipótesis nula. Dada también las otras pruebas presentadas en los apartados anteriores, podemos considerar entonces que las mediciones pueden ser tratados como iguales en el mundo estadístico. Para el periodo de tiempo en que estos coinciden, sus valores estadísticos y su distribución estarían siendo prácticamente iguales, con diferencias que podrían despreciarse para el estudio.

CAPITULO 4. ANÁLISIS DESCRIPTIVO DEL VALOR DE LA FRECUENCIA ELÉCTRICA Y SU TENDENCIA EN EL TIEMPO

4.1. Descripción del comportamiento del valor de la frecuencia eléctrica.

El análisis descriptivo puede mostrar las tendencias claves en los datos existentes y observar las situaciones que conduzcan a nuevos hechos que permitan organizar, resumir y analizar la información existente [47]. El análisis descriptivo es una fase inicial en el proceso de manipulación de datos, donde se resumen los datos pasados para extraer información relevante o para prepararlos para análisis más detallados, como lo podría ser el registro de eventos de perturbación.

Durante el Capítulo 3. se explicó que se poseen dos datos existentes de mediciones de frecuencia: Los obtenidos mediante la unidad FDR-UES 1628 durante el periodo de septiembre del 2022 hasta junio del 2023, y los proporcionados del sistema SCADA de la Unidad de Transacciones del periodo de 2022 a 2023.

4.1.1. Medidas de tendencia central y variabilidad.

Las medidas de tendencia central, también conocidas como medidas de localización, están diseñadas para brindar al analista algunos valores cuantitativos de la ubicación central o de otro tipo de los datos en una muestra [47].

Dentro de las medidas de tendencia central podemos destacar los siguientes:

- *La media:* es el promedio de todos los valores en un conjunto de datos, calculado sumando todos los valores y dividiendo entre el número total de valores. Se denotará como \bar{x} .
- *La mediana:* es el valor que se encuentra en el medio de un conjunto de datos ordenados, tiene como objetivo representar el punto medio de la muestra sin ser afectada por valores atípicos o extremos. Se denotará en este documento como \tilde{x} .
- *La moda:* es el valor que se repite en más ocasiones en un conjunto de datos. Se denotará en este documento como Mo .
- *La desviación estándar:* es una medida de dispersión que indica cuánto se desvían los valores de un conjunto de datos de su media. La desviación estándar se denotará en este documento como σ .

Para cada uno de los medidores, se obtuvieron los siguientes resultados:

Tabla II. Medidas de tendencia central y variabilidad del valor de la frecuencia eléctrica: SCADA-UT y FDR-UES

	Mean	Std. Dev.	Mode	Median	Min	Max
SCADA 2024 (UT)	60.002	0.024	60.010	60.002	58.195	62.115
FDR-UES 1628 (UES)	60.0021	0.0244	60.0132	60.0026	58.6816	60.6206

Es posible el observar que existen ciertas diferencias con lo que reporta con el SCADA UT con el FDR-UES 1628, pero estas diferencias son bien mínimas (ignorando mínimos y máximos). Observando los datos en general, se puede apreciar que existe un pequeño sesgo a la hora de observar la media con la mediana y la moda, siendo estos dos valores ligeramente superiores a la media. A su vez, el valor de la desviación estándar es inferior al dicho (0.03 Hz) en el Reglamento del Mercado Eléctrico Regional [10].

4.1.2. Tabla de distribución de frecuencias

Una tabla de distribución es una herramienta estadística que organiza y muestra la cantidad de veces que se repiten los datos o intervalos de datos. Estas tablas pueden presentarse en términos de porcentajes, lo que simplifica aún más la interpretación de los datos, ya que el porcentaje se calcula en relación con el total de datos. [47]. La tabla de distribución se dividió en dos columnas: Una para las mediciones del SCADA y otra para las mediciones del FDR-UES.

Tabla III. Tabla de distribución de los valores de la frecuencia eléctrica: FDR-UES y SCADA-UT.

Intervalos [Hz]	%scadaUT	%fdr1628	Intervalos [Hz]	%scadaUT	%fdr1628
< 59.88	0.03%	0.04%	[60.0 - 60.012)	16.53%	16.92%
[59.88 - 59.892)	0.01%	0.01%	[60.012 - 60.024)	17.21%	16.97%
[59.892 - 59.904)	0.01%	0.01%	[60.024 - 60.036)	12.39%	12.16%
[59.904 - 59.916)	0.02%	0.02%	[60.036 - 60.048)	5.49%	5.49%
[59.916 - 59.928)	0.05%	0.06%	[60.048 - 60.06)	1.48%	1.59%
[59.928 - 59.94)	0.17%	0.20%	[60.06 - 60.072)	0.31%	0.36%
[59.94 - 59.952)	0.75%	0.86%	[60.072 - 60.084)	0.06%	0.07%
[59.952 - 59.964)	3.36%	3.64%	[60.084 - 60.096)	0.02%	0.02%
[59.964 - 59.976)	9.88%	10.05%	[60.096 - 60.108)	0.01%	0.01%
[59.976 - 59.988)	15.55%	15.39%	[60.108 - 60.12)	0.00%	0.00%
[59.988 - 60.0)	16.66%	16.11%	≥ 60.12	0.01%	0.01%

Para los valores de la frecuencia, se ha realizado una tabla de distribución, en intervalos de 0.012Hz, desde 59.88 Hz hasta 60.12 Hz. El tamaño de cada intervalo se seleccionó considerando que $\sigma \approx 0.024$ Hz, es decir, la mitad de la desviación estándar; y se consideró un número de intervalos suficientes para cubrir la condición normal de operación que establece la Unidad de Transacciones para El Salvador [14].

Podemos observar que el intervalo de 60 a 60.024 Hz tiene una mayor cantidad de datos agrupados que su equivalente de 59.976 a 60 Hz. Para los valores extremos la tendencia es otra, por lo que se puede llegar a la fácil conclusión de que, fuera de la condición normal de operación, existe una mayor abundancia de valores de infrafrecuencia que de sobrefrecuencia.

4.1.3. Gráficos de cajas y bigotes.

Una gráfica de caja y bigotes es una herramienta visual utilizada para representar la distribución de un conjunto de datos. La gráfica de caja y bigote encierra el rango intercuartil de los datos en una caja que contiene la mediana representada

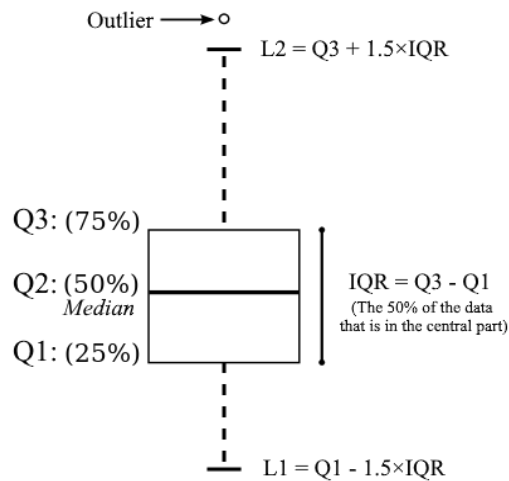


Figura 4.1. Partes de un diagrama de cajas y bigotes.

En general, podemos identificar varios elementos de un gráfico de cajas:

- Q_1 es el primer cuartil, siendo este el valor que separa el 25% de los valores inferiores.
- Q_2 es el segundo cuartil, también conocido como la mediana. Es el valor que separa los datos en dos mitades iguales.
- Q_3 es el tercer cuartil, siendo este el valor que separa el 25% de los valores superiores.

El tamaño de la caja está determinado por el rango intercuartílico ($IQR = Q3 - Q1$). Este tamaño es importante para poder calcular los límites superiores e inferiores de los bigotes. Cualquier valor que esté fuera de los límites superiores e inferiores anteriormente mencionados estarían siendo considerado valores atípicos (outliers).

- L_1 , también conocido como L_i , es el límite inferior de los datos. Normalmente se calcula como $L_1 = Q_1 - 1.5 \times IQR$.
- L_2 , también conocido como L_s , es el límite superior de los datos. Normalmente se calcula como $L_2 = Q_3 + 1.5 \times IQR$.

Las gráficas de caja y bigotes son útiles para comparar la distribución de varios conjuntos de datos e identificar valores atípicos, en especial para cuando se tiene una gran cantidad de datos, en especial para reconfirmar la existencia de ciertas asimetrías en la distribución de los datos [47]. En nuestro caso, podemos hacer uso de ellos para observar la distribución de los valores de la frecuencia monitoreada por ambos registros. Primero optaremos por observar con los valores atípicos (outliers).

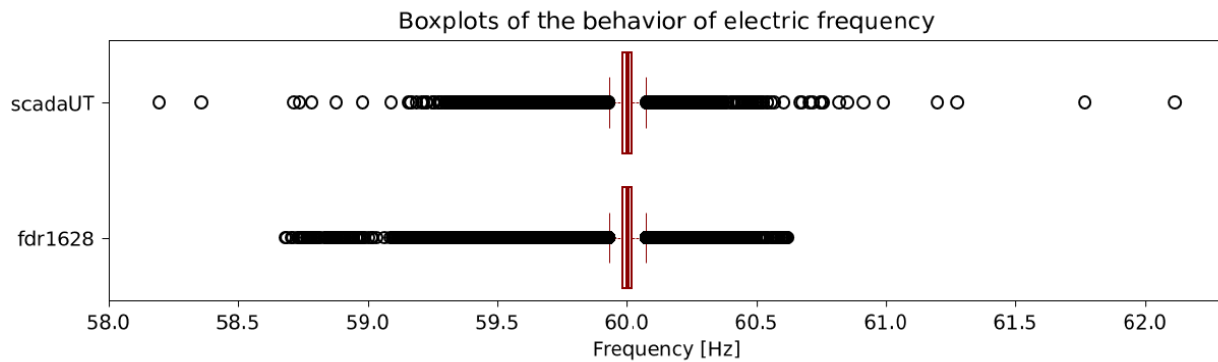


Figura 4.2. Diagrama de cajas y bigotes: Distribución de la frecuencia eléctrica.

Se puede hacer evidente que, salvo excepciones que pueden ser fácilmente contabilizadas, las desviaciones de frecuencia más grandes tienden a darse para los eventos de infrafrecuencia (frecuencia menor a la nominal), mientras que la sobrefrecuencia (frecuencia superior a la nominal) no es tan común que alcance valores atípicos tan altos, aunque existe la posibilidad de ello en base a los propios datos proporcionados por la Unidad de Transacciones.

Si procedemos a observar los gráficos de caja, sin los valores atípicos, entonces se puede apreciar que en condiciones normales se espera que el rango de la frecuencia varíe inclusive de mayor manera en valores superiores de 60 Hz que en valores inferiores de los 60 Hz. Bajo condiciones estadísticamente normales, es más esperable que existan sobrefrecuencias de mayor magnitud que las infrafrecuencias.

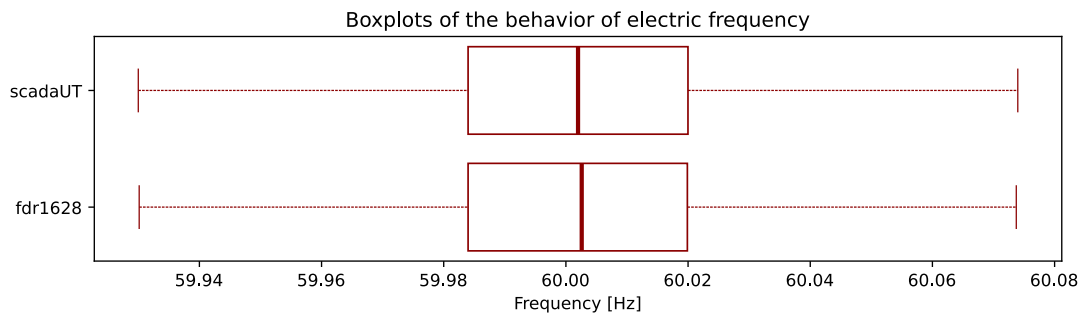


Figura 4.3. Diagrama de cajas y bigotes (sin valores atípicos): Distribución de la frecuencia eléctrica.

Los valores numéricos de lo anterior dicho se pueden visualizar en la siguiente tabla:

Tabla IV. Parámetros de los diagramas de caja y bigotes.

	L1	Q1	Q2	Q3	L2
SCADA 2024 (UT)	59.93	59.984	60.002	60.02	60.074
FDR-UES 1628 (UES)	59.93015	59.984	60.0026	60.0199	60.07375

También, mediante los gráficos de caja, se puede observar que el 50% de los datos se concentrarán en el intervalo de 59.984 hasta 60.020 Hz, un intervalo que podría ser aproximado a ± 0.020 Hz, pero que en la realidad se puede apreciar que inclusive acá se ve el pequeño sesgo anteriormente mencionado en la distribución de los datos. Con respecto a ello, se profundizará más haciendo uso de histogramas y la visualización de la curva de densidad.

4.1.4. Histograma y curva de densidad.

En base a los datos existentes de una tabla de distribución, se puede realizar un histograma. La utilidad de estos es que permite ver la distribución de los datos de forma clara, de tal forma que incluso se puedan identificar fácilmente distribuciones. Además de la tabla de distribución, esta

puede ser acompañado por un histograma y por una curva de densidad, mediante los cuales se pueda identificar de mejor manera los posibles sesgos existentes en la distribución.

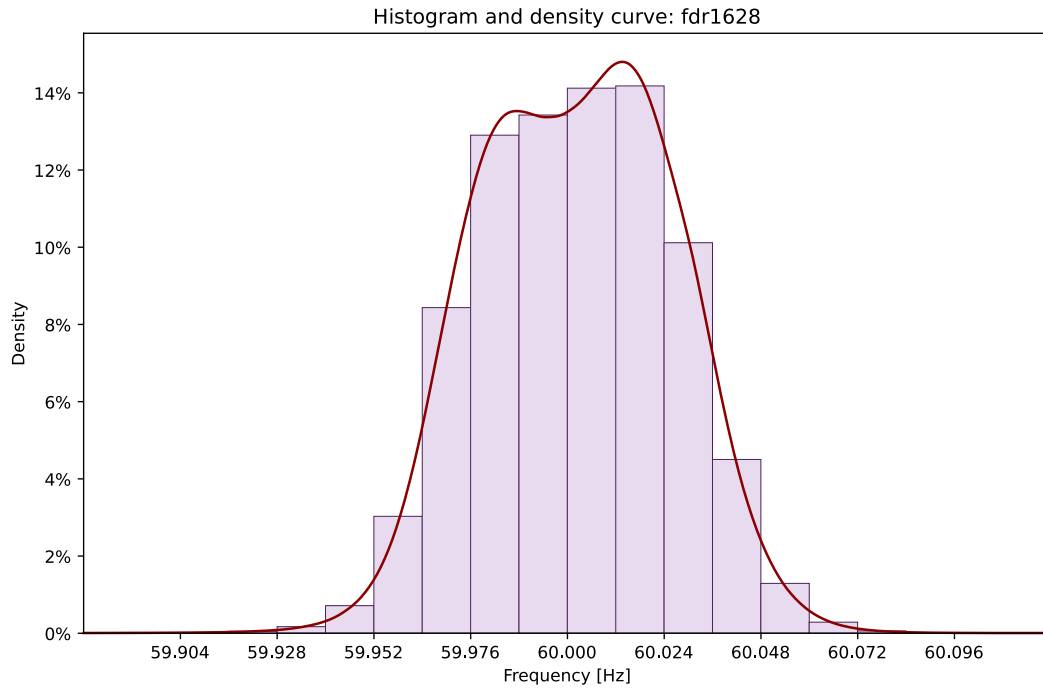


Figura 4.4. Histograma y curva de densidad: Valores de la frecuencia en El Salvador (FDR-UES 1628)

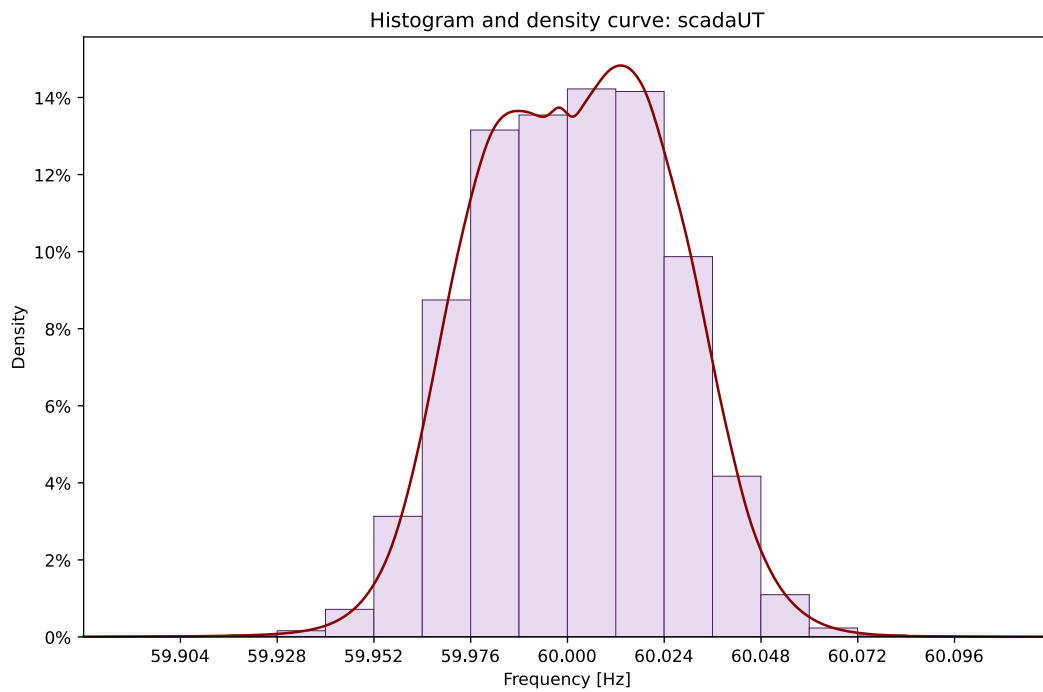


Figura 4.5. Histograma y curva de densidad: Valores de la frecuencia en El Salvador (SCADA-UT)

4.1.5. Prueba de aproximación a normalidad: Asimetría y apuntamiento.

Hasta ahora hemos llegado a la conclusión que la distribución de los valores de la frecuencia no presenta una simetría perfecta, existiendo una agrupación de datos en el lado izquierdo. Sin embargo, es necesario el calcular la asimetría y el apuntamiento de los datos para poder decir que tanto se distancian de una distribución normal. Los parámetros por calcular se definen como [52]:

$$\text{Skewness} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3 / n}{\sigma^3}$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^n (x_i - \bar{x})^4 / n}{\sigma^4} - 3$$

Finalmente se obtienen los siguientes resultados, según las mediciones realizadas:

- Para el FDR-UES 1628, se obtuvo una asimetría de -0.02 y un apuntamiento de -0.447.
- Para el SCADA-UT, se obtuvo una asimetría de -0.01 y un apuntamiento de -0.492.

Los resultados obtenidos se interpretarían en base al contenido de la siguiente tabla

Tabla V. Interpretación de resultados de asimetría (Skewness) y apuntamiento (Kurtosis)

Skewness		Kurtosis
< 0	Distribución asimétrica a la izquierda	Distribución platicúrtica (muy dispersos a la media)
= 0	Distribución simétrica	Distribución mesocúrtica (similar a la distribución normal)
> 0	Distribución asimétrica a la derecha	Distribución leptocúrtica (muy cercanos a la media)

Para mejor representación visual, también se puede consultar las siguientes figuras:

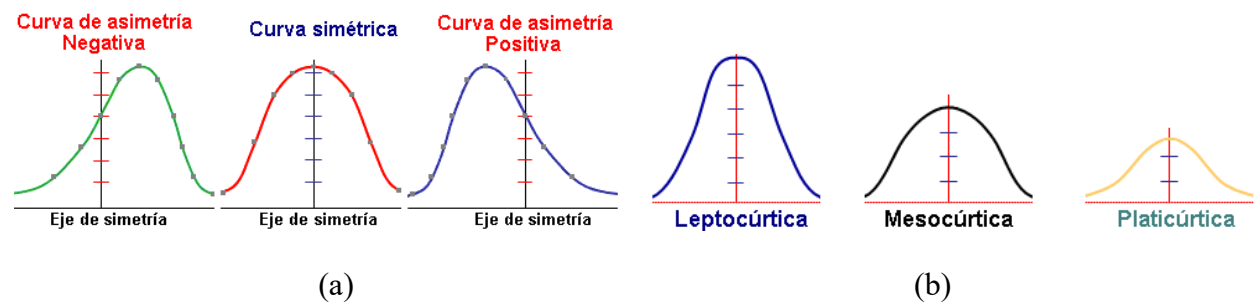


Figura 4.6. Distribuciones de datos: (a) Asimetrías, (b) Apuntamiento. Extraídos de la web.

El tipo de distribución que se tiene es una distribución asimétrica a la izquierda (negativa) con un apuntamiento platicurtica. Naturalmente los resultados no darán a existir una normalidad exacta, en ningún caso, debido a que exista cierta 'aleatoriedad' con los datos medidos de la frecuencia eléctrica. Una distribución se denomina "normal aproximada" si la asimetría o la curtosis de los datos están entre -1 y $+1$ [53], otros consideran que puede ser flexibilizado al rango -2 y $+2$ [54]. Este es un método menos confiable en el tamaño de muestra pequeño a moderado ($n < 300$) porque no puede ajustar el error estándar, pero para cantidad muestras superiores a 300, se puede guiar de valores absolutos de asimetría y el apuntamiento, además del auxilio de curvas de densidad e histogramas para considerar o no la aproximación a la normalidad [55].

Aunque somos conscientes que la distribución no es perfectamente normal, podemos llegar a aproximarla a una; dado que los valores de asimetría y apuntamiento no se encuentran fuera de los intervalos definidos anteriormente en [53], [54].

4.2. Identificación de la distribución del valor de la frecuencia en el tiempo.

Cuando hablamos de análisis de tendencias, nos referimos a la evaluación e identificación de patrones, cambios o tendencias en los datos a medida que transcurre el tiempo. Es importante tener en cuenta que el valor nominal de la frecuencia (60 Hz) generalmente se mantiene constante, pero, dado que no es perfecto, se busca observar su distribución y las variaciones que experimenta a lo largo del tiempo, para los siguientes periodos de tiempo: anual, mensual, días de la semana y horario.

4.2.1. Tendencias anuales (2022 y 2023): Datos del SCADA-UT

Para las tendencias anuales lo mejor es utilizar los datos del SCADA-UT para poder sacar conclusiones confiables, debido a que el FDR-UES 1628 no cubrió completamente ningún año. Del 2022 al 2023 se puede apreciar que no existe un gran cambio en lo que sería la distribución de los datos, si hacemos uso de las medidas de tendencia central y variabilidad. En las siguientes tablas se puede observar cómo los parámetros permanecieron sin grandes cambios.

Tabla VI. Medidas de tendencia central y variabilidad del valor de la frecuencia eléctrica: SCADA-UT

	Mean	Std. Dev.	Mode	Median	Min	Max
SCADA-UT: year 2022	60.0015	0.0239	60.01	60.002	58.195	62.115
SCADA-UT: year 2023	60.0023	0.0242	60.01	60.003	58.783	60.556

A la hora de observar la distribución de los valores, se puede apreciar que del 2022 al 2023 la distribución de los valores normales se mantuvo prácticamente igual, existiendo un pequeño incremento en 2023 que, a la hora de observar los datos, se puede considerar que puede ser despreciado.

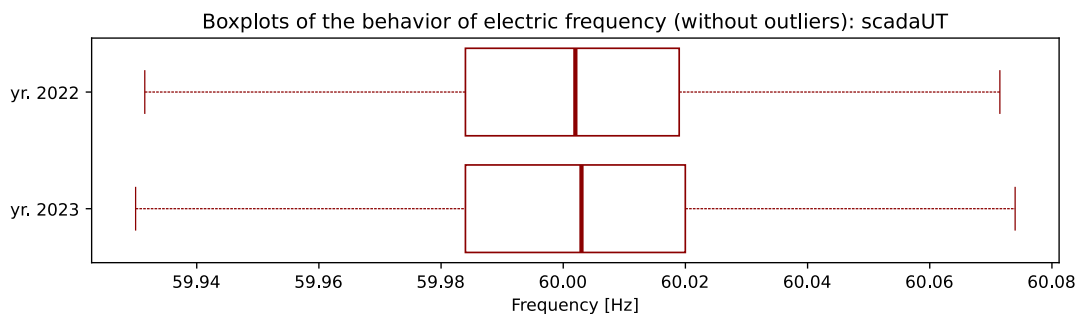


Figura 4.7. Diagrama de cajas y bigotes (sin valores atípicos): Distribución de la frecuencia eléctrica en el año 2022 y en el año 2023, según mediciones del SCADA-UT.

Los valores numéricos de lo anterior dicho se pueden visualizar en la siguiente tabla:

Tabla VII. Parámetros de los diagramas de caja y bigotes anuales: SCADA-UT

	L1	Q1	Q2	Q3	L2
Year 2022	59.9315	59.984	60.002	60.019	60.0715
Year 2023	59.93	59.984	60.003	60.02	60.074

Para el caso de los valores atípicos, se puede apreciar que fue durante el año 2022 que se dieron valores atípicos más extremos, ocurriendo que inclusive se dio una sobrefrecuencia más grande que todas las infrafrecuencias reportadas en ese año, lo cual es una situación atípica y podría considerarse una excepción a la regla [23]. Más adelante, en el siguiente capítulo, se profundizará acerca de ello.

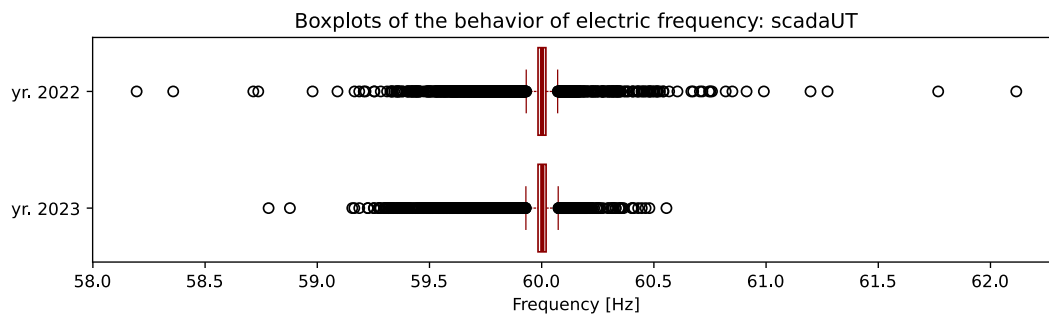


Figura 4.8. Diagrama de cajas y bigotes (con valores atípicos): Distribución de la frecuencia eléctrica en el año 2022 y en el año 2023, según mediciones del SCADA-UT.

4.2.2. Tendencias mensuales.

Cuando hacemos referencia al análisis de tendencias mensuales del valor de la frecuencia y su distribución, puede considerarse también que es el equivalente a realizar el análisis por épocas. Debido a varios factores, como lo podría ser la propia temporada de lluvias, se espera que la situación del sistema eléctrico de potencia sea distinta según la época (y por ende esto se vea reflejado en la distribución de la frecuencia). Este aspecto también se espera que se pueda visualizar en la cantidad de eventos de perturbación de frecuencia que se reporten, de lo cual se profundizará acerca de todo ello en el siguiente capítulo.

Para cada uno de los meses del estudio, se tienen las siguientes medidas de tendencia central y variabilidad, para cada uno de los conjuntos de datos:

Tabla VIII. Medidas de tendencia central y variabilidad para cada uno de los meses y años: SCADA-UT.

	Mean	Std. Dev.	Mode	Median	Min	Max
January 2022	60.0018	0.0257	60.01	60.002	59.166	60.161
February 2022	60.0017	0.0263	60.011	60.002	59.334	60.126
March 2022	60.0018	0.026	60.017	60.002	59.283	60.157
April 2022	60.0013	0.0238	60.013	60.002	59.326	60.188
May 2022	60.0003	0.0242	60.012	60	58.195	62.115
June 2022	60.001	0.0214	60.013	60.001	59.703	60.193
July 2022	60.001	0.0219	60.013	60.001	59.331	60.151
August 2022	60.001	0.0215	60.011	60.001	59.187	60.134
September 2022	60.0013	0.0217	60.012	60.001	59.838	60.193
October 2022	60.002	0.0241	60.014	60.003	59.207	60.142
November 2022	60.0026	0.0248	60.013	60.003	59.331	60.164
December 2022	60.0027	0.0242	60.016	60.003	59.862	60.167
January 2023	60.0026	0.0252	60.016	60.003	59.226	60.189
February 2023	60.0027	0.0257	60.015	60.003	58.783	60.556
March 2023	60.0026	0.0245	60.010	60.003	59.187	60.199
April 2023	60.0023	0.0252	60.014	60.003	59.272	60.226
May 2023	60.0015	0.0239	60.011	60.002	59.156	60.147
June 2023	60.0018	0.0221	60.012	60.002	59.165	60.246
July 2023	60.002	0.0222	60.014	60.003	59.282	60.175
August 2023	60.0018	0.0217	60.013	60.002	59.225	60.158
September 2023	60.0026	0.0218	60.012	60.003	59.31	60.192
October 2023	60.0025	0.0244	60.014	60.003	59.275	60.147
November 2023	60.0028	0.0262	60.018	60.004	59.253	60.185
December 2023	60.0027	0.0261	60.018	60.004	59.304	60.144

Tabla IX. Medidas de tendencia central y variabilidad para cada uno de los meses y años: FDR-UES 1628

	Mean	Std. Dev.	Mode	Median	Min	Max
September 2022	60.0012	0.0218	60.0129	60.0013	59.822	60.2142
October 2022	60.0016	0.0236	60.0130	60.002	59.0899	60.1317
November 2022	60.0025	0.0248	60.0141	60.0031	59.2914	60.1666
December 2022	60.0026	0.0239	60.0150	60.003	59.8416	60.1682
January 2023	60.0025	0.0252	60.0143	60.0031	59.1977	60.1989
February 2023	60.0026	0.0257	60.0157	60.0033	58.6797	60.6168
March 2023	60.0026	0.0247	60.0143	60.0031	59.1433	60.2065
April 2023	60.0018	0.0252	60.0143	60.0025	59.1363	60.2309
May 2023	60.0011	0.0244	60.0138	60.0018	59.126	60.1541
June 2023	60.0017	0.0226	60.0127	60.0024	59.0973	60.2812

Considerando que la media es prácticamente 60 Hz en cualquier época del año, el parámetro de mayor interés para observar su variación es la desviación estándar, dado que indica cuánto se desvían los valores de un conjunto de datos de su media.

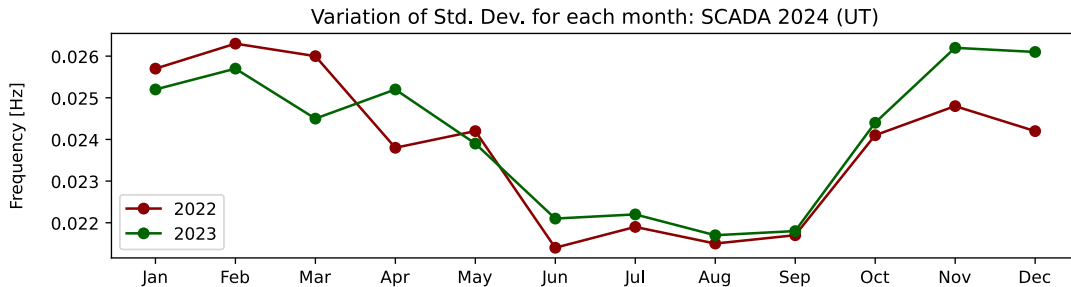


Figura 4.9. Tendencia de la desviación estándar para cada mes del año: SCADA-UT

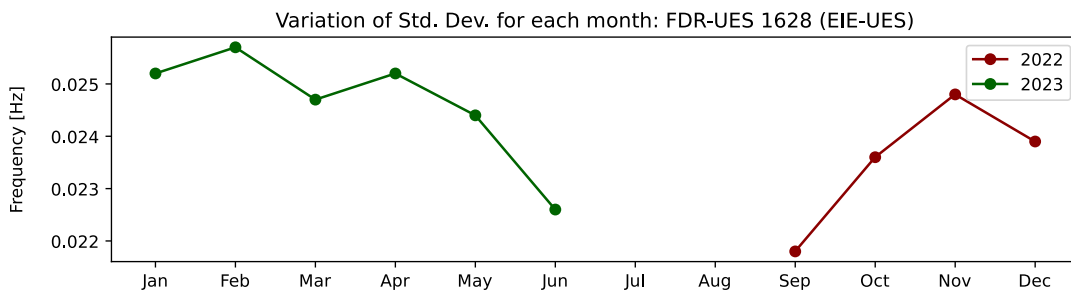


Figura 4.10. Tendencia de la desviación estándar para cada mes del año: FDR-UES 1628.

Para ambas figuras se puede apreciar la tendencia de que entre junio hasta septiembre de tiene una desviación estándar mucho más baja de lo que sería durante el resto del año. Para el caso del FDR, debido a que solo recolectó diez meses, no se podía confirmar esta teoría. Pero con los datos proporcionados con la Unidad de Transacciones del sistema SCADA se pudo confirmar esta observación, dado que son dos años completos de datos.

Las estadísticas, considerando únicamente los meses y sin dividirlos en años, se presentan a continuación para el caso del SCADA-UT. En ellas se puede observar de mejor forma la tendencia anteriormente mencionada con la desviación estándar.

Tabla X. Medidas de tendencia central y variabilidad de cada mes: SCADA-UT

	Mean	Std. Dev.	Mode	Median	Min	Max
January	60.0022	0.0255	60.010	60.003	59.166	60.189
February	60.0022	0.026	60.015	60.003	58.783	60.556
March	60.0022	0.0253	60.010	60.003	59.187	60.199
April	60.0018	0.0246	60.012	60.002	59.272	60.226
May	60.0009	0.0241	60.017	60.001	58.195	62.115
June	60.0014	0.0217	60.012	60.002	59.165	60.246
July	60.0015	0.0221	60.014	60.002	59.282	60.175
August	60.0014	0.0216	60.011	60.002	59.187	60.158
September	60.0019	0.0218	60.012	60.002	59.31	60.193
October	60.0023	0.0243	60.014	60.003	59.207	60.147
November	60.0027	0.0255	60.015	60.004	59.253	60.185
December	60.0027	0.0252	60.016	60.003	59.304	60.167

También, a la hora de observar la distribución de los valores de la frecuencia mediante diagramas de caja y bigotes, podemos confirmar que durante la temporada de junio – septiembre de cada año es que los bigotes de los diagramas tenían el menor tamaño, llegando a presentarse variación de 0.01 Hz en la longitud de estos. Inclusive se puede apreciar, aunque en mucha menor medida, que el tamaño de las cajas de los diagramas tienden a ser menor durante la temporada junio – septiembre.

Los datos numéricos, para el SCADA-UT y el FDR-UES 1628, se presentan en las tablas:

Tabla XI. Parámetros de los diagramas de caja y bigotes mensuales: SCADA-UT

	L1	Q1	Q2	Q3	L2
January	59.9260	59.983	60.003	60.021	60.078
February	59.9245	59.983	60.003	60.022	60.0805
March	59.9260	59.983	60.003	60.021	60.078
April	59.9300	59.984	60.002	60.02	60.074
May	59.9330	59.984	60.001	60.018	60.069
June	59.9355	59.985	60.002	60.018	60.0675
July	59.9355	59.985	60.002	60.018	60.0675
August	59.9355	59.985	60.002	60.018	60.0675
September	59.9355	59.985	60.002	60.018	60.0675
October	59.9260	59.983	60.003	60.021	60.078
November	59.9245	59.983	60.004	60.022	60.0805
December	59.9245	59.983	60.003	60.022	60.0805

Tabla XII. Parámetros de los diagramas de caja y bigotes mensuales: FDR-UES 1628

	L1	Q1	Q2	Q3	L2
January	59.9272	59.9836	60.0031	60.0212	60.0776
February	59.9265	59.9835	60.0033	60.0215	60.0785
March	59.9289	59.9841	60.0031	60.0209	60.0761
April	59.9300	59.9839	60.0025	60.0198	60.0737
May	59.9309	59.9837	60.0018	60.0189	60.0717
June	59.9360	59.9853	60.0024	60.0182	60.0675
July & August	-	-	-	-	-
September	59.9358	59.9847	60.0013	60.0173	60.0662
October	59.9307	59.9838	60.002	60.0192	60.0723
November	59.9285	59.9839	60.0031	60.0208	60.0762
December	59.9299	59.9842	60.003	60.0204	60.0747

Los diagramas de caja y bigotes, para el caso del SCADA-UT se muestran en la siguiente figura. Cabe aclarar que, para observar la distribución de los datos, se ha hecho uso principalmente de los datos del SCADA-UT debido a que este cubre dos años completos.

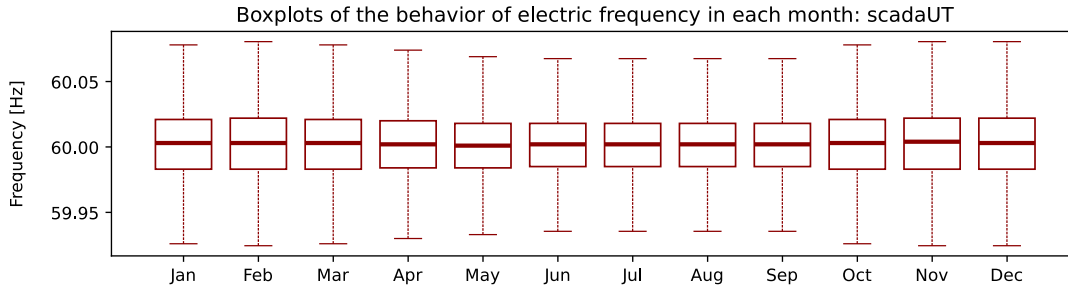


Figura 4.11. Distribución de los valores de la frecuencia para cada uno de los meses (2022 y 2023): SCADA-UT.

4.2.3. Tendencias para cada día de la semana

Se observará la distribución del valor de la frecuencia para cada uno de los días de la semana, desde el lunes hasta el domingo. Para ello, se hará de forma similar a como se realizó para ver las tendencias mensuales. Primero, observemos el comportamiento de la desviación estándar para cada uno de los días de la semana (ver figura).

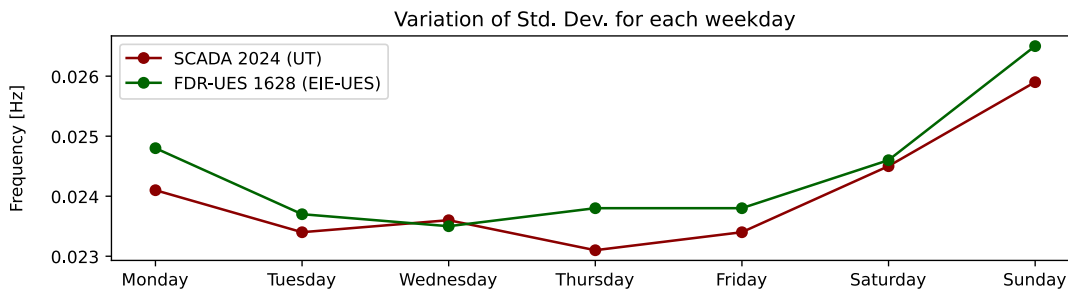


Figura 4.12. Tendencia de la desviación estándar para cada día de la semana: SCADA-UT y FDR-UES 1628.

En general, es evidente que los mayores valores de desviación estándar se dan durante el domingo, con valores de 0.026 Hz (0.0265 Hz, según el FDR-UES 1628). Los otros días que podemos llegar a destacar sería el lunes y sábado, en los cuales se presentan mayores valores, pero que todavía son inferiores al domingo. Para el resto de los días (martes a viernes), se presentan valores similares en lo que sería la desviación estándar.

Para cada uno de los días de la semana, se tienen las siguientes medidas de tendencia central y variabilidad, para cada uno de los conjuntos de datos:

Tabla XIII. Medidas de tendencia central y variabilidad para cada uno de los días de la semana: SCADA-UT.

	Mean	Std. Dev.	Mode	Median	Min	Max
Monday	60.0018	0.0241	60.012	60.002	59.331	60.144
Tuesday	60.0018	0.0234	60.014	60.002	58.783	60.556
Wednesday	60.0016	0.0236	60.013	60.002	58.195	61.274
Thursday	60.0017	0.0231	60.010	60.002	59.165	60.246
Friday	60.0019	0.0234	60.010	60.002	59.09	62.115
Saturday	60.0021	0.0245	60.016	60.003	59.225	60.192
Sunday	60.0026	0.0259	60.015	60.003	59.156	60.204

Tabla XIV. Medidas de tendencia central y variabilidad para cada uno de los días de la semana: FDR-UES 1628

	Mean	Std. Dev.	Mode	Median	Min	Max
Monday	60.0022	0.0248	60.0131	60.0028	59.2299	60.1441
Tuesday	60.0018	0.0237	60.0113	60.0023	58.6797	60.6168
Wednesday	60.0019	0.0235	60.0141	60.0024	59.158	60.1617
Thursday	60.002	0.0238	60.0130	60.0026	59.0973	60.2812
Friday	60.002	0.0238	60.0123	60.0025	59.1745	60.1652
Saturday	60.0022	0.0246	60.0147	60.0026	59.1977	60.1989
Sunday	60.0027	0.0265	60.0158	60.0033	59.126	60.2142

Mediante diagramas de caja podemos observar la distribución de los valores de la frecuencia para cada uno de los días de la semana:

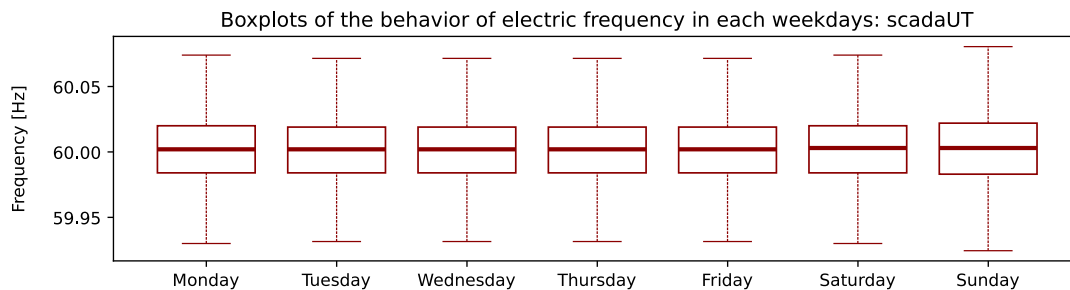


Figura 4.13. Distribución de los valores de la frecuencia para cada día de la semana: SCADA-UT.

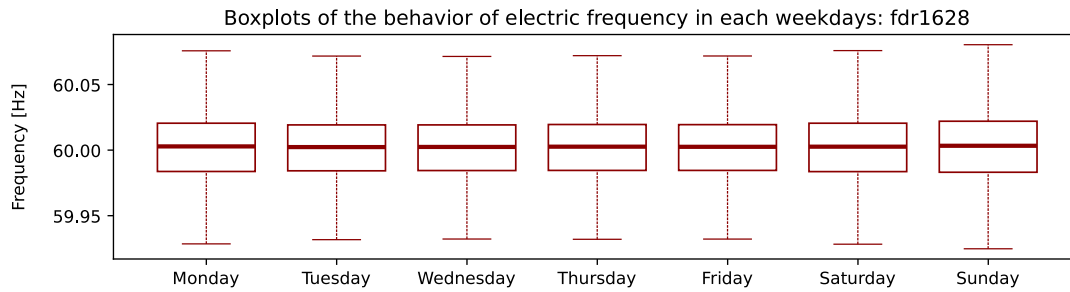


Figura 4.14. Distribución de los valores de la frecuencia para cada día de la semana: FDR-UES 1628

Se puede apreciar que el tamaño de los bigotes e inclusive la caja es mayor durante el domingo, mientras que el menor tamaño de ambos parámetros se da durante los martes a viernes. Los datos numéricos se presentan en las siguientes tablas.

Tabla XV. Parámetros de los diagramas de caja y bigotes para cada día de la semana: SCADA-UT

	L1	Q1	Q2	Q3	L2
Monday	59.93	59.984	60.002	60.02	60.074
Tuesday	59.9315	59.984	60.002	60.019	60.0715
Wednesday	59.9315	59.984	60.002	60.019	60.0715
Thursday	59.9315	59.984	60.002	60.019	60.0715
Friday	59.9315	59.984	60.002	60.019	60.0715
Saturday	59.93	59.984	60.003	60.02	60.074
Sunday	59.9245	59.983	60.003	60.022	60.0805

Tabla XVI. Parámetros de los diagramas de caja y bigotes para cada día de la semana: FDR-UES 1628

	L1	Q1	Q2	Q3	L2
Monday	59.9285	59.9837	60.0028	60.0205	60.0757
Tuesday	59.9317	59.9842	60.0023	60.0192	60.0717
Wednesday	59.9322	59.9844	60.0024	60.0192	60.0714
Thursday	59.932	59.9845	60.0026	60.0195	60.072
Friday	59.93215	59.9845	60.0025	60.0194	60.07175
Saturday	59.92825	59.9836	60.0026	60.0205	60.07585
Sunday	59.92475	59.9831	60.0033	60.022	60.08035

4.2.4. Tendencias para cada hora del día.

Se observará la distribución del valor de la frecuencia para cada una de las horas de los días. Para ello, se hará de forma similar a como se realizó para ver las tendencias mensuales y para cada día de la semana, observando el comportamiento de la desviación estándar para cada hora (ver hora).

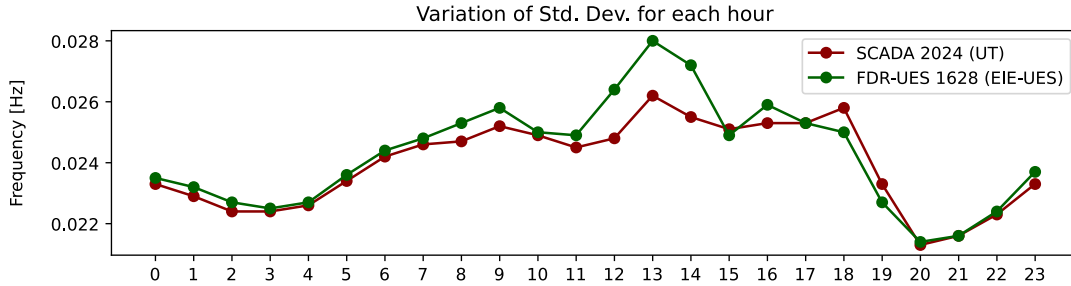


Figura 4.15. Tendencia de la desviación estándar para hora del día: SCADA-UT y FDR-UES 1628.

Existen algunas diferencias a notar en lo que serían las mediciones del FDR-UES y del registro por el SCADA-UT, y de hecho al inicio puede llegar a ser algo preocupante. Sin embargo, consideremos que las mediciones del SCADA-UT cubren desde enero 2022 hasta diciembre 2023, mientras que los del FDR-UES cubren desde septiembre 2022 hasta junio 2023. Para ello, obsérvese el siguiente gráfico, en el que los datos de SCADA-UT están restringidos al mismo periodo que el FDR-UES.

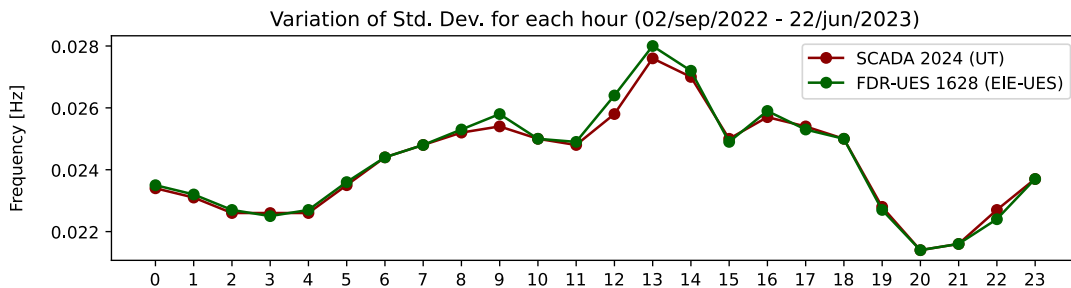


Figura 4.16. Tendencia de la desviación estándar para hora del día, considerando únicamente el periodo del día 02 de septiembre del 2022 hasta el 22 junio del 2023: SCADA-UT y FDR-UES 1628.

Retomando lo visto para las tendencias mensuales, en el periodo de septiembre a junio se dan valores de la desviación estándar más altos, por lo que ello puede explicar la diferencia de las mediciones del SCADA-UT y del FDR-UES. Indirectamente, esto ha servido para poder observar que a lo largo del año puede cambiar la tendencia de la distribución del valor de la frecuencia, el siendo este un detalle que profundizaremos más adelante.

Las medidas de tendencia y variabilidad para el caso de los registros del sistema SCADA-UT, y también para el FDR-UES 1628. Téngase en consideración los tiempos que registraron ambos medidores, y por ende las diferencias existentes entre ambos.

Tabla XVII. Medidas de tendencia central y variabilidad para cada hora del día (general): SCADA-UT.

Hour	Mean	Std. Dev.	Mode	Median	Min	Max
00:00	60.0019	0.0233	60.012	60.002	59.858	60.185
01:00	60.0018	0.0229	60.017	60.002	59.311	60.14
02:00	60.0014	0.0224	60.017	60.001	59.862	60.147
03:00	60.0009	0.0224	60.014	60.001	59.654	60.149
04:00	60.0006	0.0226	59.981	60	59.883	60.141
05:00	60.0015	0.0234	59.984	60.001	59.441	60.15
06:00	60.0018	0.0242	60.016	60.002	59.275	60.152
07:00	60.0033	0.0246	60.010	60.004	59.451	60.164
08:00	60.0032	0.0247	60.016	60.004	59.165	60.213
09:00	60.0027	0.0252	60.014	60.003	58.783	62.115
10:00	60.0021	0.0249	60.014	60.003	59.156	60.367
11:00	60.0012	0.0245	60.015	60.002	59.317	60.188
12:00	60.0017	0.0248	60.013	60.003	59.187	60.158
13:00	60.0012	0.0262	60.015	60.002	59.215	60.193
14:00	60.0012	0.0255	60.010	60.002	59.187	60.246
15:00	60.0013	0.0251	60.015	60.002	59.166	60.193
16:00	60.0019	0.0253	60.016	60.003	59.375	60.161
17:00	60.0018	0.0253	60.010	60.002	59.253	60.19
18:00	60.0002	0.0258	59.986	60	58.195	61.274
19:00	60.0022	0.0233	60.010	60.002	59.355	60.568
20:00	60.0028	0.0213	60.010	60.003	59.207	60.121
21:00	60.0034	0.0216	60.010	60.004	59.872	60.151
22:00	60.0028	0.0223	60.010	60.003	59.488	60.13
23:00	60.0036	0.0233	60.010	60.004	59.872	60.136

Tabla XVIII. Medidas de tendencia central y variabilidad para cada hora del día (general): FDR-UES 1628.

Hour	Mean	Std. Dev.	Mode	Median	Min	Max
00:00	60.0019	0.0235	60.0163	60.0023	59.8545	60.0991
01:00	60.0019	0.0232	60.014	60.002	59.8902	60.1061
02:00	60.0011	0.0227	60.0149	60.0012	59.8418	60.1052
03:00	60.001	0.0225	60.0137	60.0009	59.8911	60.098
04:00	60.0008	0.0227	59.9854	60.0006	59.8799	60.1288
05:00	60.0015	0.0236	59.9844	60.0012	59.7517	60.163
06:00	60.0021	0.0244	60.0161	60.0019	59.8494	60.161
07:00	60.0032	0.0248	60.0144	60.0038	59.822	60.1666
08:00	60.0046	0.0253	60.0165	60.0057	59.0973	60.2257
09:00	60.003	0.0258	60.0152	60.0038	58.6797	60.6168
10:00	60.0021	0.025	60.0136	60.0028	59.126	60.14
11:00	60.0012	0.0249	60.0134	60.0019	59.3188	60.1782
12:00	60.0016	0.0264	60.0148	60.0027	59.0899	60.1663
13:00	60.0011	0.028	60.0144	60.0024	59.1977	60.1989
14:00	60.0012	0.0272	60.0118	60.0023	59.1433	60.2812
15:00	60.0015	0.0249	60.0103	60.0021	59.8416	60.1468
16:00	60.0017	0.0259	60.015	60.0025	59.2145	60.17
17:00	60.0017	0.0253	60.0129	60.0022	59.8606	60.2142
18:00	60	0.025	59.9844	59.9998	59.8702	60.1127
19:00	60.0027	0.0227	60.0101	60.003	59.2738	60.1556
20:00	60.003	0.0214	60.0111	60.0036	59.158	60.1241
21:00	60.0043	0.0216	60.012	60.0052	59.8679	60.1123
22:00	60.0034	0.0224	60.0144	60.004	59.8851	60.1037
23:00	60.0039	0.0237	60.0122	60.0045	59.8693	60.1292

Por primera vez se hará un comentario que no involucra a la desviación estándar, y es que puede observarse que durante las 04:00 y las 05:00 horas se tiene una moda de 59.9854 y 59.9844 Hz respectivamente. Esto es algo que llega a destacar bastante, pues si verificamos todas las tablas

anteriores, durante este capítulo, podemos ver que es primera vez que se presenta un valor de moda inferior a los 60 Hz. Salvo ese detalle, no hay más que se pudiera agregar observando otras medidas de tendencia central, y la desviación estándar está dentro de lo normal.

Para complementar lo visto anteriormente, se muestran los gráficos de caja y bigotes para cada una de las horas del día. Para ambos medidores se puede apreciar que existe una distribución bastante similar en lo que serían los valores de la frecuencia, aunque evidentemente durante las 20:00 y 21:00 horas se puede observar que se presenta la menor distribución, mientras que la mayor se presenta en horas como las 08:00 o 16:00 horas.

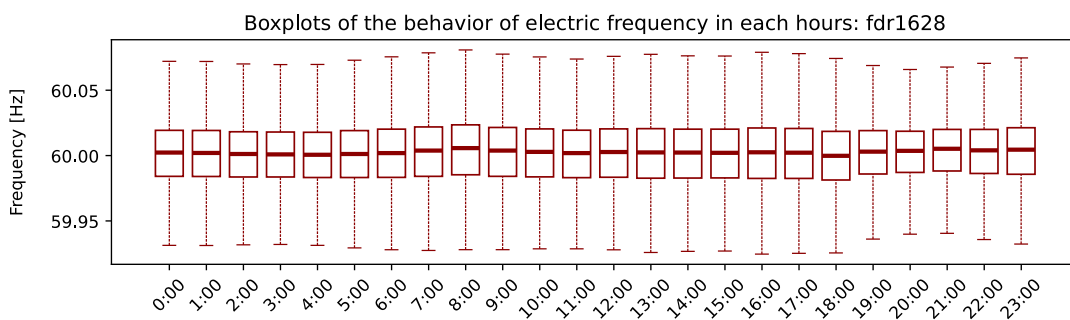


Figura 4.17. Distribución de los valores de la frecuencia para cada hora del día: FDR-JES 1628.

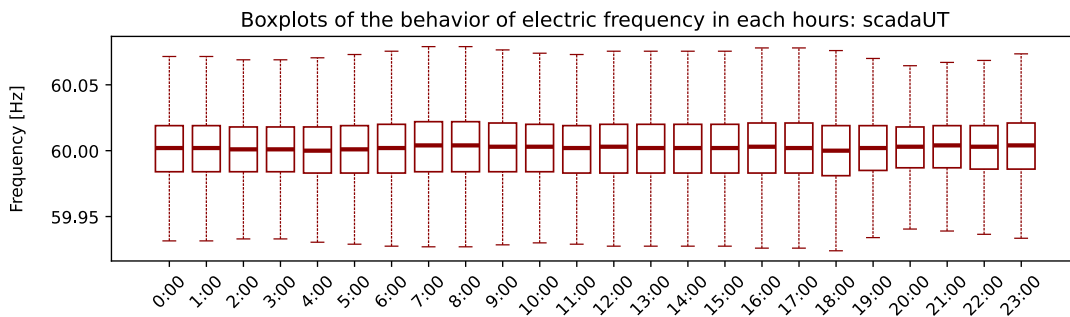


Figura 4.18. Distribución de los valores de la frecuencia para cada hora del día: SCADA-UT

En las correspondientes tablas se encuentra con las distribuciones mencionadas:

Tabla XIX. Parámetros de los diagramas de caja y bigotes para hora del día: SCADA-UT

	L1	Q1	Q2	Q3	L2
0	59.9315	59.984	60.002	60.019	60.0715
1	59.9315	59.984	60.002	60.019	60.0715
2	59.933	59.984	60.001	60.018	60.069
3	59.933	59.984	60.001	60.018	60.069
4	59.9305	59.983	60	60.018	60.0705
5	59.929	59.983	60.001	60.019	60.073
6	59.9275	59.983	60.002	60.02	60.0755
7	59.927	59.984	60.004	60.022	60.079
8	59.927	59.984	60.004	60.022	60.079
9	59.9285	59.984	60.003	60.021	60.0765
10	59.93	59.984	60.003	60.02	60.074
11	59.929	59.983	60.002	60.019	60.073
12	59.9275	59.983	60.003	60.02	60.0755
13	59.9275	59.983	60.002	60.02	60.0755
14	59.9275	59.983	60.002	60.02	60.0755
15	59.9275	59.983	60.002	60.02	60.0755
16	59.926	59.983	60.003	60.021	60.078
17	59.926	59.983	60.002	60.021	60.078
18	59.924	59.981	60	60.019	60.076
19	59.934	59.985	60.002	60.019	60.07
20	59.9405	59.987	60.003	60.018	60.0645
21	59.939	59.987	60.004	60.019	60.067
22	59.9365	59.986	60.003	60.019	60.0685
23	59.9335	59.986	60.004	60.021	60.0735

Tabla XX. Parámetros de los diagramas de caja y bigotes para hora del día: FDR-UES 1628

	L1	Q1	Q2	Q3	L2
0	59.9315	59.984	60.002	60.019	60.0715
1	59.9315	59.984	60.002	60.019	60.0715
2	59.933	59.984	60.001	60.018	60.069
3	59.933	59.984	60.001	60.018	60.069
4	59.9305	59.983	60	60.018	60.0705
5	59.929	59.983	60.001	60.019	60.073
6	59.9275	59.983	60.002	60.02	60.0755
7	59.927	59.984	60.004	60.022	60.079
8	59.927	59.984	60.004	60.022	60.079
9	59.9285	59.984	60.003	60.021	60.0765
10	59.93	59.984	60.003	60.02	60.074
11	59.929	59.983	60.002	60.019	60.073
12	59.9275	59.983	60.003	60.02	60.0755
13	59.9275	59.983	60.002	60.02	60.0755
14	59.9275	59.983	60.002	60.02	60.0755
15	59.9275	59.983	60.002	60.02	60.0755
16	59.926	59.983	60.003	60.021	60.078
17	59.926	59.983	60.002	60.021	60.078
18	59.924	59.981	60	60.019	60.076
19	59.934	59.985	60.002	60.019	60.07
20	59.9405	59.987	60.003	60.018	60.0645
21	59.939	59.987	60.004	60.019	60.067
22	59.9365	59.986	60.003	60.019	60.0685
23	59.9335	59.986	60.004	60.021	60.0735

Se procede a observar las curvas horarias para cada uno de los meses del año, y también para cada día de la semana, haciendo uso de los datos del SCADA-UT (dado que cubren un mayor periodo de tiempo). Para una mejor visualización, se observará el comportamiento de la desviación estándar de forma agrupada: cuatrimestres, días laborales y no laborales.

Para el caso de cada cuatrimestre, puede observarse que los menores valores de desviación estándar ocurren durante el tercer cuatrimestre (julio, agosto, septiembre), mientras que los valores más altos de desviación estándar se dan en el cuatrimestre I y IV. La forma de las curvas puede presentar ligeras variaciones de un cuatrimestre a otro, como lo podría ser la hora en que se den los picos máximos de desviación estándar (y por ende dispersión en el valor de la frecuencia), pero la mayoría coinciden en la tendencia.

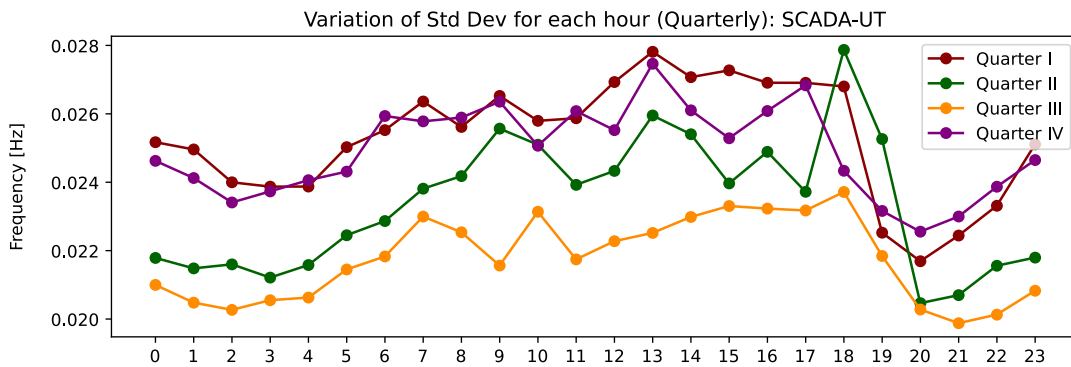


Figura 4.19. Tendencia de la desviación estándar para hora del día, para cada cuatrimestre: SCADA-UT.

Si observamos las curvas para días laborales y no laborales, se puede observar que los valores de la desviación estándar (Y por ende la distribución de la frecuencia) es mayor durante los fines de semana, si se compara con la curva de los días laborales.

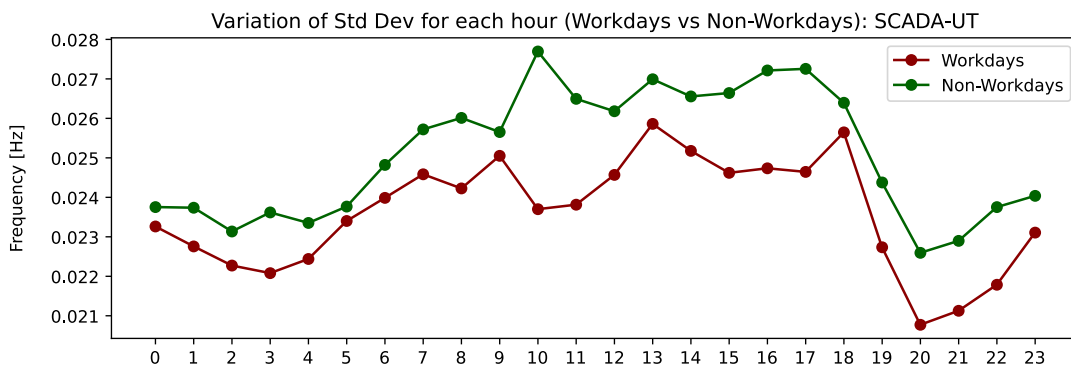


Figura 4.20. Tendencia de la desviación estándar para hora del día, para días laboral y no laboral: SCADA-UT.

CAPITULO 5. ACERCA DEL REGISTRO DE EVENTOS DE PERTURBACIÓN DE FRECUENCIA Y SU ANÁLISIS.

5.1. Creación de lista de eventos de perturbación de frecuencia.

Para llevar a cabo un análisis completo, es necesario identificar, siquiera de forma básica, los eventos de perturbación que han ocurrido durante el periodo de tiempo en cuestión. Para la identificación y clasificación de eventos de desviación de la frecuencia eléctrica nos enfocaremos únicamente en la magnitud de la desviación máxima reportada durante el evento.

Lo propuesto en este documento es un criterio básico pero funcional para la realización de un análisis estadístico. Para el control de la frecuencia pueden encontrarse varias formas de realizar el registro de eventos e inclusive identificar el tipo de perturbación, como el hacer uso de métodos basados en el ROCOF o el uso de Machine Learning [56].

5.1.1. Propuesta de clasificación de eventos de perturbación según su magnitud.

Hemos decidimos realizar una clasificación propia para mayores facilidades en el estudio, tanto para los eventos mayores como para poder contabilizar pequeñas fluctuaciones pequeñas consideradas cotidianas. Para ello, se ha tenido todo lo siguiente:

- El estándar IEEE 1159-2019 recomienda que, en condición normal, la frecuencia eléctrica no deberá de superar los 0.10 Hz de variación, y que el periodo de tiempo del evento de la variación no deberá ser mayor a los 10 segundos [12].
- Según IEC 61000-2-5, "la frecuencia fundamental de una red de suministro de energía es generalmente muy estable y no varía más del 0.2%" [13].
- Según el Reglamento de Operación del Sistema de Transmisión y del Mercado Mayorista, la frecuencia podría variar hasta 0.12 Hz y seguir en condición operativa. A su vez, el rango de operación de emergencia sería entre 59.40 hasta 59.88 Hz [14].
- Según el Reglamento del Mercado Eléctrico Regional, las variaciones de frecuencia promedio deben de estar comprendidas en un rango de $60 \pm 1.65\sigma$. Establecido $\sigma = 0.03$ Hz, entonces el rango estaría siendo de variaciones de hasta 0.05 Hz [15].

También se desea tener en consideración los eventos más pequeños de frecuencia que se dan cotidianamente en la red. Las fluctuaciones en la frecuencia es ya algo característico del sistema, pues en raras ocasiones se dará que la generación será exactamente igual a la demanda. Sin embargo, de forma cotidiana se pueden identificar fluctuaciones de frecuencia más grandes que otras, siendo nuestro interés el identificar las más altas que existan.

Para esta investigación, tendremos en consideración para registro los eventos que presenten valores de frecuencia a partir de 0.05 Hz; tomando como referencia recomendaciones y guías técnicas usadas para la red de América del Norte que establecen que toda variación menor a ± 0.05 Hz no generan preocupación y puede ser considerado variaciones normales [23], [17], [57], [58], [59], además de que el intervalo de ± 0.05 Hz es lo que sugiere el CRIE para el 90% del tiempo [15].

Con respecto a la magnitud, tendríamos las siguientes clasificaciones:

1. *Leve*: La desviación máxima en la frecuencia ronda entre 0.05 Hz hasta 0.12 Hz.
2. *Moderado*: La desviación máxima está en el rango de 0.12 Hz hasta 0.60 Hz.
3. *Critico*: La desviación máxima es mayor o igual a 0.60 Hz.

Con respecto a su duración, basado en lo recomendado en [12], únicamente consideraremos eventos de perturbación cuya duración sea igual o mayor a 10 segundos.

5.1.2. Identificación de eventos de perturbación.

Se implementará un algoritmo simple para la detección de los eventos de perturbación, teniendo en cuenta que los datos utilizados provienen del FDR-UES y el SCADA-UT. Se descartarán los métodos y técnicas basados en la estimación del ROCOF como lo realizado en [56], [60], debido a que es necesario utilizar para todo el tiempo mediciones sincrofásicas que cumplan con lo establecido en [61], como el poseer tasas de muestreo de 10, 12, 10, 12, 15, 20, 30, 60 o 120 muestras por segundo. Por coherencia para los resultados del análisis estadístico, se desarrollará un algoritmo que se aplique tanto para los datos del SCADA como del FDR-UES.

A. Detección del inicio de un evento de perturbación.

De [56], [60] hemos retomado la idea de hacer uso de una ventana móvil para la detección de eventos de perturbación. Para un punto de tiempo t_k existente, se tomará una ventana de datos de tamaño n , tal que cubra un total de 12 segundos (debido a la tasa de muestreo del SCADA, el tiempo debe ser un múltiplo de 4). Se calculará su promedio y desviación estándar, y se asignarán esos valores al punto de tiempo t_k actual. Luego la ventana se desplaza al siguiente punto de tiempo t_{k+1} , repitiendo el proceso. Al primer parámetro mencionado es comúnmente conocido en la literatura como promedio móvil [62], mientras que el segundo podemos llamarle, por analogía, una desviación estándar móvil.

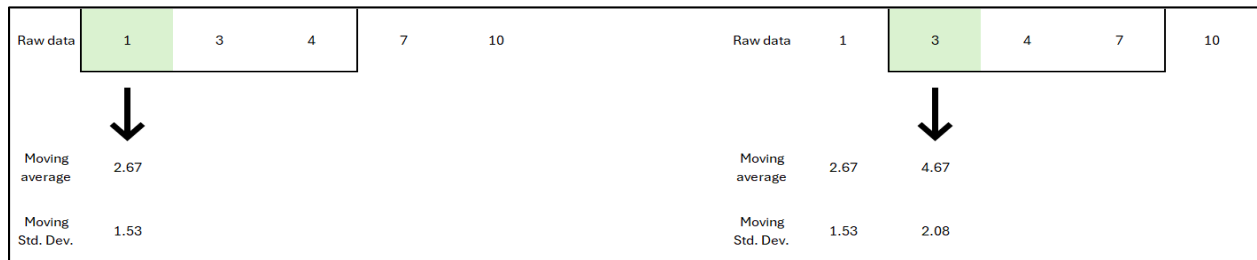


Figura 5.1. Ilustración del funcionamiento del algoritmo del promedio móvil y la desviación estándar móvil.

Si se da que alguno de estos parámetros supere un umbral respectivo, entonces se detectará que ha iniciado un evento de perturbación.

- El promedio móvil deberá de tener un valor superior a 0.05 Hz.
- La desviación estándar móvil deberá de superar el valor de la desviación estándar general, que podemos aproximar a 0.025 Hz en base a los resultados del Capítulo 4.

Nota: Para que el evento de perturbación quede registrado, tuvo que existir una desviación máxima de la frecuencia de ± 0.05 Hz, es decir, entrar en clasificación de evento menor.

Todo lo mencionado sirve para identificar los puntos donde la distribución de probabilidad de los datos cambia dentro de una serie de datos [63], [64]. En este rol, la media y la desviación estándar juegan un papel importante en la teoría de la detección de puntos de cambio, ya que un cambio en estos parámetros puede indicar la existencia de un cambio en la señal.

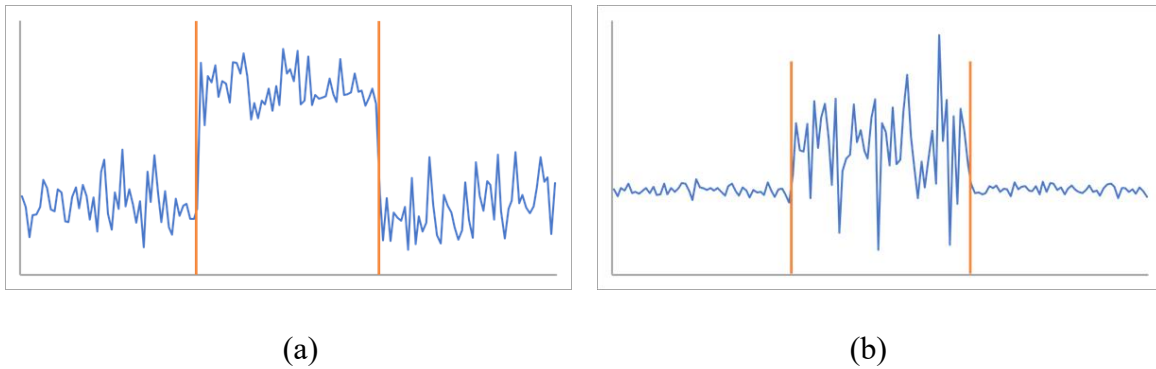


Figura 5.2. Detección de puntos de cambio: (a) Por un cambio en el valor de la media, (b) Por un cambio en el valor de la desviación estándar. Extraído de [65].

Aunque existen varios algoritmos que determinan automáticamente los puntos de cambio, hemos considerado que no es necesario llegar a profundizar en ellos, y que con una ventana deslizante era suficiente para la detección de eventos, al igual que se realizó en [56]. Debido al análisis estadístico anterior realizado, nosotros ya tenemos una idea de los umbrales a elegir.

B. Criterio de finalización de evento de perturbación.

Para que finalice un evento de perturbación, deberá de cumplirse que el valor de la frecuencia se haya normalizado, cumpliéndose con un “tiempo de estudio”, con el fin de evitar el tener falsos registros de múltiples eventos consecutivos, siendo que en la realidad se tratan del mismo evento. Según la clasificación de magnitud, así será el tiempo de estudio. Estos tiempos fueron sacados a base de prueba y error, concluyendo que:

- Para un evento de magnitud leve, el tiempo de estudio será de un minuto; durante ese periodo, el valor de la frecuencia no deberá de sobrepasar los 0.05 Hz.
- Para un evento de magnitud considerable, el tiempo de estudio será de cinco minutos; durante ese periodo, el valor de la frecuencia no deberá de sobrepasar los 0.12 Hz.
- Para un evento de magnitud crítica, el tiempo de estudio será de diez minutos; durante ese periodo, el valor de la frecuencia no deberá de sobrepasar los 0.12 Hz.

Como requisito final, deberá cumplirse que durante el último minuto del evento no se presentó ningún valor de la frecuencia superior a los 0.05 Hz.

5.1.3. Registro de información del evento

Una vez identificado el inicio y finalización de un evento de perturbación, se tiene que hacer un registro de los datos más relevantes obtenidos. Algunos datos que pueden ser registrados, teniendo identificado el inicio y finalización del evento, serían:

- El tiempo en el cual inició y finalizó el evento de perturbación de frecuencia y su duración.
- El valor mínimo y máximo de la frecuencia que fue identificado durante el evento.
- La clasificación del evento registrado.
- El o los medidores que identificaron el evento.

Existen otros parámetros que también son necesarios de calcular para poder realizar un registro completo, los cuales requieren una explicación ampliada. Esos parámetros serán introducidos posteriormente, según las necesidades existentes. Actualmente con los parámetros obtenidos, se puede realizar un estudio preliminar de los eventos de perturbación para lo que se necesita para este capítulo.

5.2. Descripción del comportamiento de los registros de eventos de perturbación.

5.2.1. Contabilización de eventos de perturbación en el periodo de estudio.

Para el caso de los eventos de perturbación, se tiene como tiempo de estudio desde el primer día del año 2022 hasta el último día del año 2023, gracias a los datos proporcionados por la Unidad de Transacciones.

Se puede realizar una descripción completa del registro de los eventos de perturbación realizado, en base a la sección 5.1. En total, se tiene reportado 36 309 eventos de perturbación durante los dos años de estudio. Estos eventos se distribuyen, según clasificación y tipo, de la siguiente forma:

Tabla XXI. Contabilización de eventos de perturbación, para el periodo de estudio, según clasificación y tipo. Realizado en base a mediciones del SCADA-UT y el FDR-UES 1628.

Clasificación	Underfrequency	Overfrequency	Total
Minor	12699	23118	35817
Moderate	155	276	431
Critical	60	1	61
Total	12914	23395	36309

Se puede apreciar lo siguiente:

- El 35.57% de los eventos de perturbación registrados son de infrafrecuencia, mientras que el 64.43% de los eventos registrados son de sobrefrecuencia. Se considera sobrefrecuencia cuando la desviación máxima reportada fue por un valor mayor al nominal, e infrafrecuencia cuando fue por un valor menor. Si durante el evento, existió una frecuencia mínima de 59.85 Hz, pero una frecuencia máxima de 60.20 Hz, se ha considerado que es un evento de sobrefrecuencia.
- El 98.64% de los eventos de perturbación son de magnitud leve, el 1.19% son de magnitud moderada y el 0.17% son eventos críticos. Los eventos de magnitud leve (desviaciones de frecuencia entre 0.05 Hz hasta 0.12 Hz) son algo cotidiano en la región centroamericana, mientras que los otros eventos de perturbación (moderados y críticos) se puede observar que estos son menos comunes pero frecuentes (esto se profundizará más adelante).

5.2.2. Diagrama de dispersión de los eventos de perturbación.

Un gráfico de dispersión es una representación gráfica de la relación entre dos variables cuantitativas. Cada punto representa un par de valores, uno para cada variable. Mediante los diagramas de dispersión se puede facilitar el encontrar relaciones existentes entre dos variables u observar el comportamiento de un fenómeno.

Se realizó un diagrama de dispersión en el que se graficó cada registro de perturbación de frecuencia, haciendo uso de la magnitud de la desviación es el valor máximo reportado y también la duración de los eventos de perturbación.

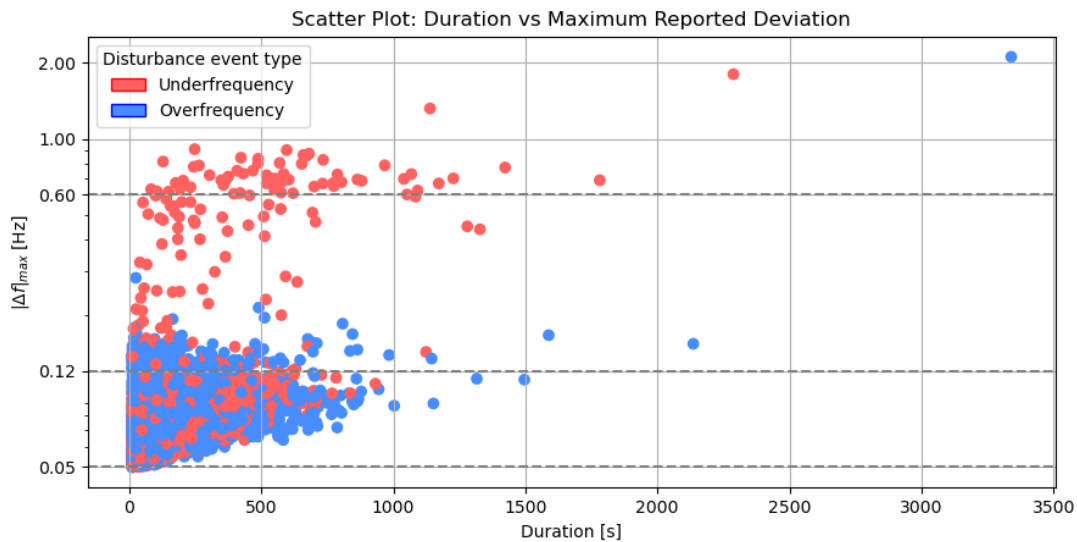


Figura 5.3. Gráfico de dispersión de los eventos de perturbaciones de frecuencia registradas, de la magnitud máxima reportada y la duración de los eventos: Escala semilogarítmica (eje vertical).

Existen varios aspectos que podemos destacar con respecto al gráfico de dispersión.

- Se pueden llegar a considerar que estos eventos de perturbación de frecuencia son algo cotidiano, observando que estos se concentran principalmente para bajas desviaciones de frecuencia y con duraciones que normalmente no llegan a superar los 1 000 segundos.
- Inclusive después de 0.12 Hz, existen desviaciones máximas reportadas en eventos en una cantidad considerable. Sin embargo, a medida que vamos subiendo el valor de las desviaciones se puede observar que la cantidad de eventos disminuye.

- Existe una cierta relación existente entre la duración de un evento y su magnitud; lo que podemos traducir que se debe a que, una vez ocurrido un evento, mientras más grande llegue a ser en su magnitud, más tiempo puede llegar a tardar la estabilización del sistema.

Para los primeros dos puntos mencionados, se puede visualizar de forma numérica este comportamiento mediante una tabla de distribución más especificada. Inicialmente puede apreciarse que la cantidad de eventos de sobrefrecuencia tienen mayor presencia que los eventos de infrafrecuencia, siendo que esta es una tendencia que sigue marcada hasta el intervalo de 0.15 a 0.18 Hz de desviación de la frecuencia. Posterior a ello, se puede apreciar que lo que existen son valores de infrafrecuencia en su gran mayoría de casos. A su vez, se puede contabilizar que los eventos de perturbación que han sido mayormente contados son bastante pequeños, y en la Tabla XXII se puede apreciar la tendencia, existiendo cerca de un 90% de todos los eventos únicamente en el intervalo de 0.05 a 0.08 Hz de desviación de la frecuencia. Aun así, el resto de cantidad de eventos que quedan es una cantidad significativa que prestar atención, en especial para las clasificaciones de eventos «moderados» y «críticos» existentes en el periodo de 24 meses.

Tabla XXII. Contabilización de valores máximos de la frecuencia durante eventos.

Intervals	Underfrequency	Overfrequency	Total
0.05 – 0.08	11479	21277	32756
0.08 – 0.12	1220	1841	3061
0.12 – 0.15	83	243	326
0.15 – 0.18	17	28	45
0.18 – 0.21	5	3	8
0.21 – 0.24	4	1	5
0.24 – 0.27	5	0	5
0.27 – 0.30	3	1	4
0.30 – 0.60	38	0	38
> 0.60	60	1	61
Total	12914	23395	36309

Para el segundo punto, que es determinar la correlación existente entre la magnitud máxima reportada de un evento con la duración de este, se puede hacer uso del coeficiente de correlación de Pearson [47], [66], [67]. Este coeficiente se calcula, considerando un parámetro X y Y que tienen la misma n cantidad de datos, de la siguiente forma [47]:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})/n}{\sigma_X \cdot \sigma_Y}$$

Donde ρ es el coeficiente de Pearson de correlación entre las variables X y Y , que en nuestro caso sería $X =$ la magnitud máxima reportada y $Y =$ la duración del evento; σ_{XY} es la covarianza de ambas variables X y Y , que indica como varían juntos estos valores; mientras que σ_X y σ_Y son las desviaciones estándar de cada una de las variables.

El resultado del análisis es un coeficiente de correlación que puede variar entre -1 y $+1$, indicando el tipo y la fuerza de la asociación. Si el valor que se obtiene es 0 , quiere significar que no existe relación alguna, si es $+1$ es que existe una relación directa, y si es -1 es que es una relación inversa [47], [66], [67].

Finalmente, se obtuvo un resultado de 0.462 . La interpretación de la magnitud del coeficiente de correlación de Pearson puede interpretarse según las sugerencias de Cohen [66], [67]:

Tabla XXIII. Interpretación de la magnitud del coeficiente de correlación, según sugerencias de Cohen. Extraído de [66], [67].

Rango de valores de $ \rho $	Interpretación
$0.00 \leq \rho < 0.10$	Correlación nula
$0.10 \leq \rho < 0.30$	Correlación débil
$0.30 \leq \rho < 0.50$	Correlación moderada
$0.50 \leq \rho < 1.00$	Correlación fuerte

Observando la tabla, se puede decir que existe una relación moderada entre lo que sería la magnitud máxima reportada, y la duración del evento. Se puede interpretar, usando un poco de lógica, que mientras de mayor magnitud sea el evento de perturbación, el tiempo que tomará el sistema a estabilizar el evento podría llegar a aumentar, en base a los criterios que nosotros hemos establecido para la identificación del inicio y finalización de evento.

5.2.3. Comportamiento de eventos de perturbación según su clasificación.

Dos características principales que se pueden evaluar para los eventos registrados, de forma estadística, serían la distribución de sus magnitudes máximas y sus duraciones.

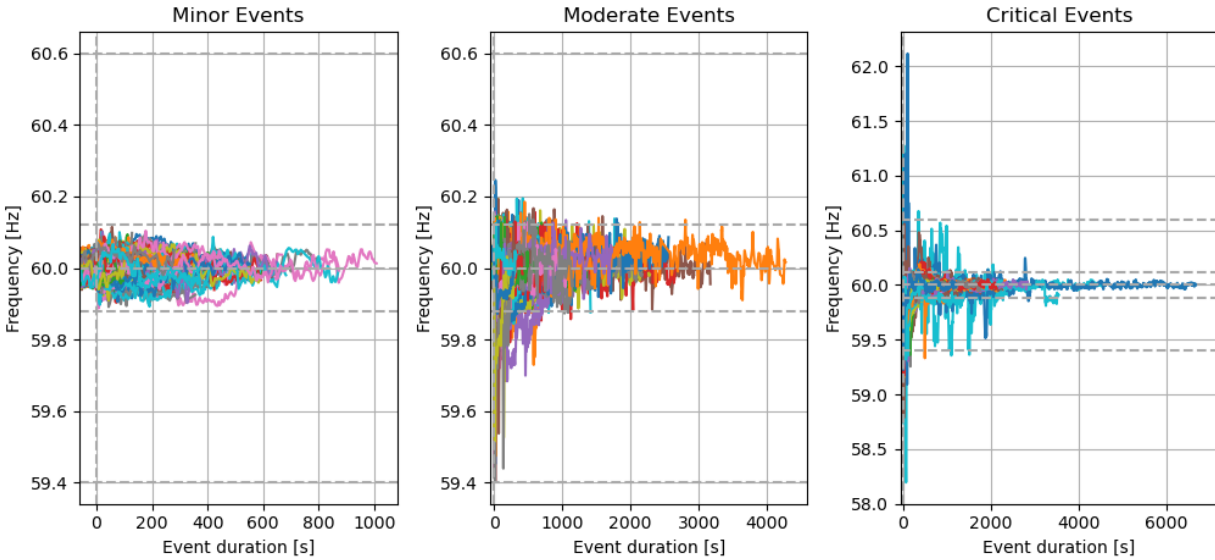


Figura 5.4. Gráfico del comportamiento de la frecuencia durante eventos de perturbación.

Para cada clasificación, observaremos brevemente la distribución de su magnitud máxima y su duración auxiliándonos de diagramas de cajas y bigotes,

A. Eventos de perturbación clasificados como menores

Dentro de los eventos registrados como menores, se puede observar los siguientes diagramas de cajas y bigotes para las desviaciones máximas reportadas y las duraciones.

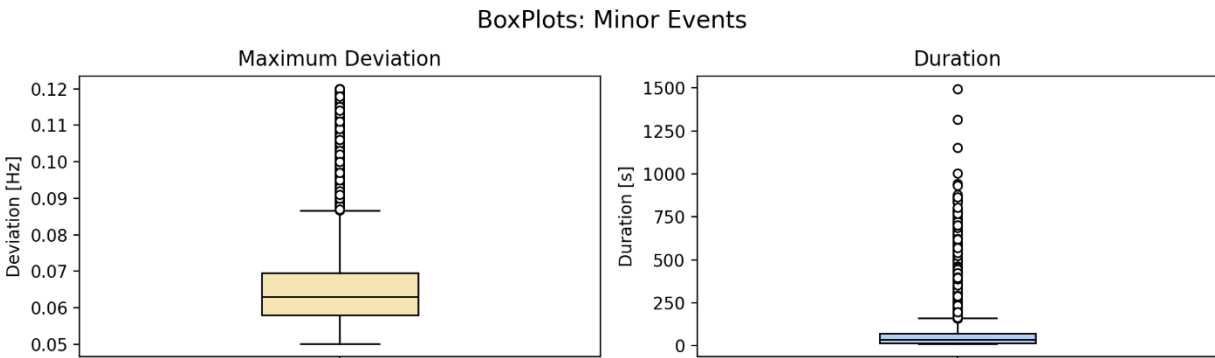


Figura 5.5. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación menores.

Podemos observar que el 75% de los valores de frecuencia se concentran en desviaciones inferiores a los 0.07 Hz, y que lo normal es que estas desviaciones de frecuencia que se presentan lleguen hasta valores de 0.09 Hz aproximadamente. El 50% inferior se concentra hasta 0.063 Hz, por lo que puede verse que buena cantidad de los eventos quedan agrupados en valores inferiores.

Para el caso de la duración de los eventos, se podrían presentar normalmente duraciones de hasta 160 segundos aproximadamente, siendo duraciones superiores los casos atípicos. Sin embargo, obsérvese que el 75% de los eventos inferiores duran hasta 73.4 segundos, y el 50% dura apenas 36 segundos, por lo que lo más común, en base a los criterios utilizados para este estudio, es esperar que muchos eventos duren normalmente menos de un minuto.

La siguiente es una tabla que contabiliza la cantidad de eventos ocurridos según la magnitud máxima reportada de cada uno de ellos.

Tabla XXIV. Contabilización de valores máximos de la frecuencia durante eventos menores.

Intervals	Underfrequency	Overfrequency	Total
0.05 – 0.06	4223	8078	12301
0.06 – 0.07	5357	9446	14803
0.07 – 0.08	1899	3753	5652
0.08 – 0.09	635	990	1625
0.09 – 0.10	321	402	723
0.10 – 0.11	170	250	420
0.11 – 0.12	94	199	293
Total	12699	23118	35817

A su vez, se resume la información de los diagramas de caja y bigote.

Tabla XXV. Parámetros de los diagramas de caja y bigotes para eventos menores.

	L1	Q1	Q2	Q3	L2
Max. Dev. [Hz]	0.0501	0.058	0.063	0.0695	0.0867
Duration [s]	10	15.9	35.9	73.4	159.6

B. Eventos de perturbación clasificados como moderados.

Dentro de los eventos registrados como moderados, se puede observar los siguientes diagramas de cajas y bigotes para las desviaciones máximas reportadas y las duraciones.

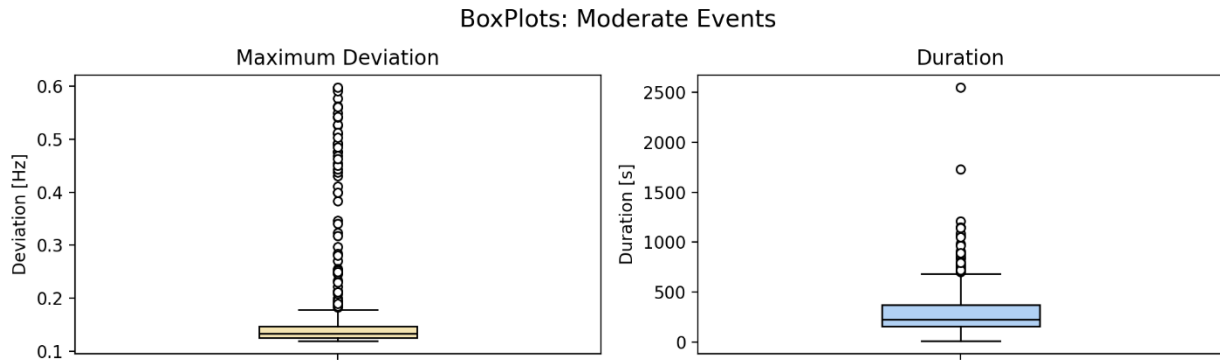


Figura 5.6. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación moderados.

El registro de eventos de perturbación moderados presenta el mismo comportamiento que los eventos menores: La cantidad de registros se presenta para el intervalo de 0.12 a 0.18 Hz, este porcentaje aumenta 86.07%. Todo lo dicho puede verse reflejado en el diagrama de caja de magnitudes y en la Tabla XXVII, donde el 75% de los valores de magnitud de los eventos moderados se concentran en valores inferiores a 0.1493 Hz.

La duración de estos eventos de perturbación moderados puede esperarse que rondan hasta los 528 segundos. Sin embargo, se puede apreciar que se concentran las duraciones en valores inferiores, siendo que el 75% inferior son duraciones menores a 233 segundos y el 50% inferior apenas son duraciones inferiores a los 108 segundos.

Una tabla que muestre la distribución de los valores reportados se muestra:

Tabla XXVI. Contabilización de valores máximos de la frecuencia durante eventos moderados.

Intervals	Underfrequency	Overfrequency	Total
0.12 – 0.18	100	271	371
0.18 – 0.24	9	4	13
0.24 – 0.30	8	1	9
0.30 – 0.36	4	0	4

Intervals	Underfrequency	Overfrequency	Total
0.36 – 0.42	4	0	4
0.42 – 0.48	10	0	10
0.48 – 0.54	9	0	9
0.54 – 0.60	11	0	11
Total	155	276	431

El diagrama de caja y bigotes se puede resumir en la Tabla XXVII.

Tabla XXVII. Parámetros de los diagramas de caja y bigotes para eventos moderados.

	L1	Q1	Q2	Q3	L2
Max. Dev. [Hz]	0.1201	0.12675	0.134	0.1493	0.183
Duration [s]	10.5	37.25	108	233.5	528

C. Eventos de perturbación clasificados como críticos.

Dentro de los eventos registrados como críticos, se puede observar los siguientes diagramas de cajas y bigotes para las desviaciones máximas reportadas y las duraciones.

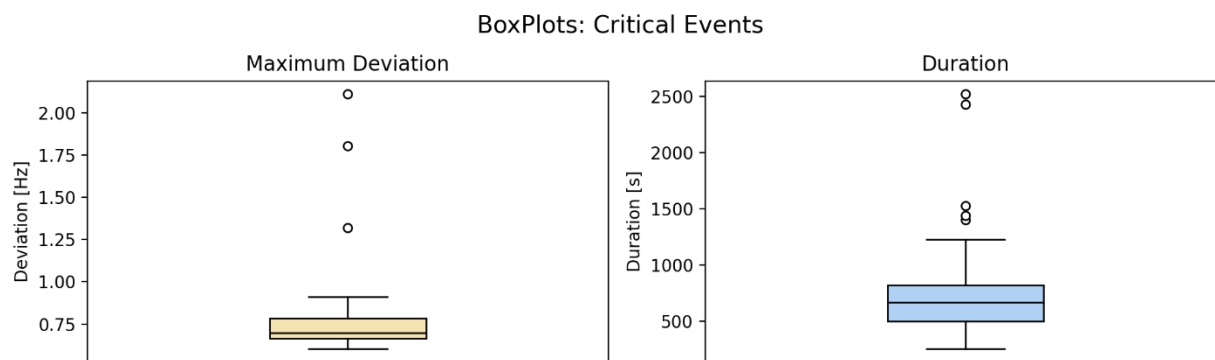


Figura 5.7. Diagrama de caja general de la magnitud máxima y duración de eventos de perturbación críticos.

Si observamos una tabla con la contabilización de los eventos, podemos observar que una buena parte de los eventos críticos, típicamente por pérdida de generación, se dan por pérdidas inferiores a los 0.70 Hz, no alcanzando a activar el esquema de desconexión de carga por baja frecuencia establecido en la región a partir de 59.30 Hz [29].

Se puede presentar todavía una cantidad significativa de eventos de perturbación para desviaciones superiores a 0.70, siendo que para esta clasificación predominan los eventos de pérdida de generación. Se puede observar aun así un evento de sobrefrecuencia en la contabilización, pero este caso podríamos decir que es atípico, y se profundizará más adelante en esta misma sección.

Si observamos los parámetros del diagrama de caja y bigotes, podemos observar que los eventos críticos reportados comúnmente se espera que estos tengan magnitudes menores a 0.91 Hz. Eventos de perturbación que generen mayores desviaciones de la frecuencia pueden ser considerados como casos atípicos, habiendo existido tres de ellos durante el periodo de estudio.

La duración de los eventos puede observarse que se podrían esperar duraciones de hasta 20 minutos, siendo valores superiores a estos considerados atípicos, pero posibles.

Tabla XXVIII. Contabilización de valores máximos de la frecuencia durante eventos críticos.

Intervals	Underfrequency	Overfrequency	Total
0.60 – 0.70	33	0	33
0.70 – 0.80	14	0	14
0.80 – 0.90	9	0	9
0.90 – 1.00	2	0	2
1.00 – 1.10	0	0	0
1.10 – 1.20	0	0	0
1.20 – 1.30	0	0	0
1.30 – 1.40	1	0	1
1.40 – 1.50	0	0	0
1.50 – 1.60	0	0	0
1.60 – 1.70	0	0	0
1.70 – 1.80	0	0	0
1.80 – 1.90	1	0	1
1.90 – 2.00	0	0	0
> 2.00	0	1	1
Total	60	1	61

A su vez, una tabla en la que muestra los parámetros de los diagramas de caja se muestra:

Tabla XXIX. Parámetros de los diagramas de caja y bigotes para eventos críticos.

	L1	Q1	Q2	Q3	L2
Max. Dev. [Hz]	0.602	0.665	0.696	0.7855	0.9101
Duration [s]	82.6	374.8	574.9	788.4	1227.4

Vale la pena mencionar que, durante el periodo de estudio, se registraron dos eventos de perturbación en las que se podría decir que existió una «oscilación» en el valor de la frecuencia eléctrica. Para el evento del 27 de mayo del 2022, se presentaron valores extremos de 59.40 Hz y a su vez un valor máximo superior 62 Hz, siendo este un caso atípico del cual no se encontró documentación. Para el 14 de febrero del 2023 se presentó un comportamiento similar.

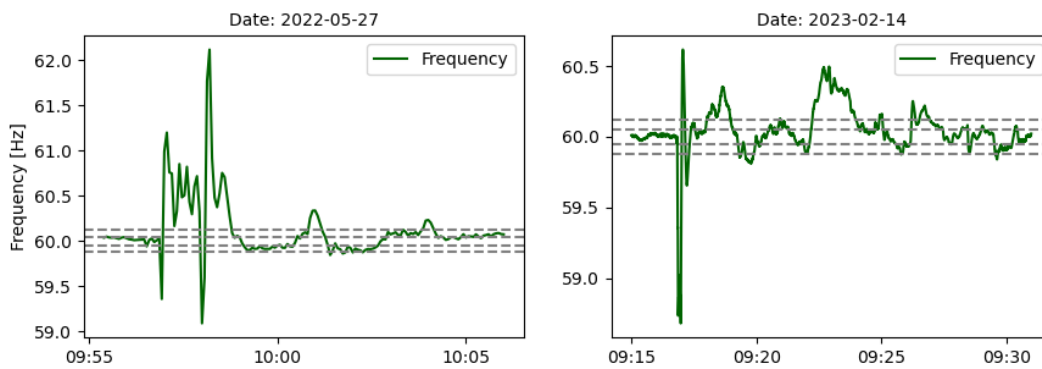


Figura 5.8. Comportamiento de dos eventos de perturbación críticos: Casos atípicos.

5.3. Tendencias en el tiempo

Para complementar al análisis, se procederá a observar las tendencias del registro de los eventos de perturbación a lo largo del tiempo. Considerando que los datos proporcionados por la Unidad de Transacciones y los del FDR-UES cubren dos años, se puede realizar el análisis de tendencias considerando: Cambios anuales, mensuales, para cada día de la semana e intervalo de horas.

5.3.1. Tendencias anuales.

Según la clasificación, se realiza la contabilización de los eventos que ha existido durante los dos años de estudio, los cuales se han presentado en un gráfico de barras.

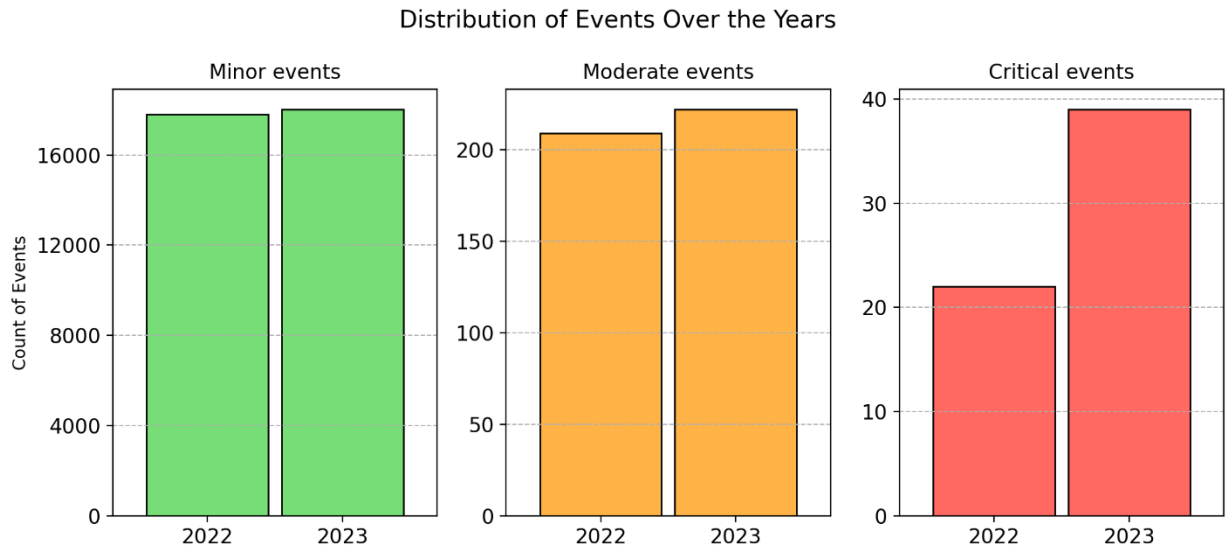


Figura 5.9. Contabilización de eventos de perturbación de frecuencia, por año.

Si observamos la cantidad de eventos reportados en 2022 y 2023, se puede apreciar que existe un aumento para todas las clasificaciones, notándose más para los eventos críticos. Estos resultados encontrados no pueden definir una tendencia, dado que estamos hablando de apenas dos años con los cuales se puede comparar., además que no poseemos datos anteriores a 2022.

Tabla XXX. Contabilización de eventos registrados, según su clasificación, para año

	Minor	Moderate	Critical
Year 2022	17800	209	22
Year 2023	18017	222	39

Otro aspecto que se puede observar es la evolución de las magnitudes máximas reportadas para cada periodo de tiempo. Para mayor simplicidad en el análisis, observaremos cual fue la media de los eventos por clasificación, y a su vez cual fue su desviación estándar existente.

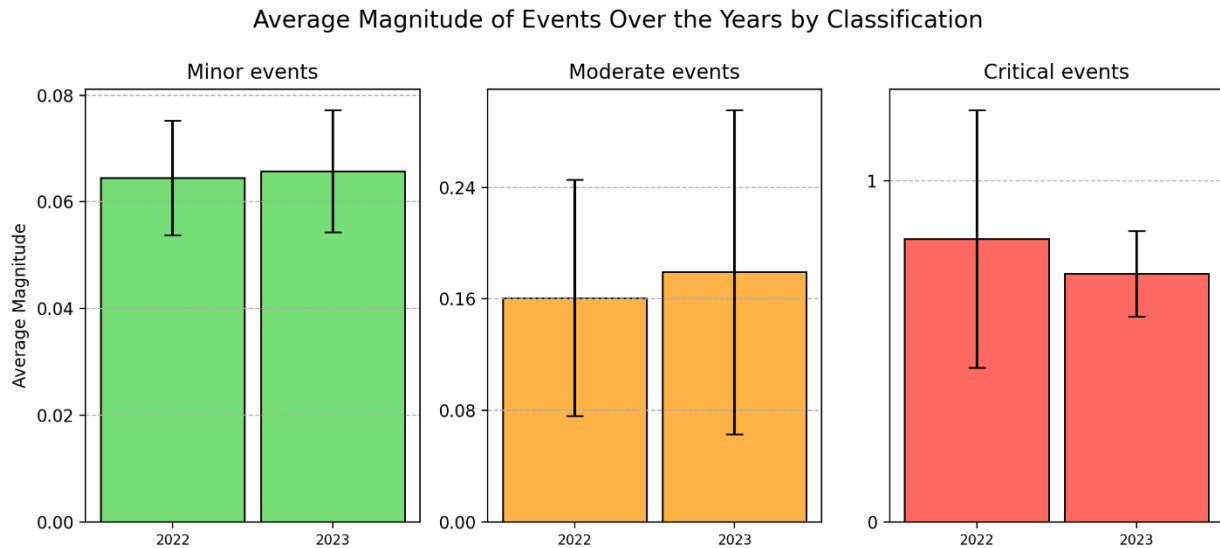


Figura 5.10. Promedio de desviación máxima de eventos de perturbación de frecuencia reportados, por año.

Para el caso de los eventos menores y moderados, podemos observar que sí existió un aumento en la media y en la desviación estándar, significando que para estas clasificaciones se dieron eventos de perturbación de mayor impacto en general, siendo esto más notable para el caso de los eventos de clasificación moderadas. Sin embargo, para los eventos de clasificación crítica fue el caso contrario, existiendo una notable disminución en la media de estos eventos como a su vez en la desviación estándar de estos mismos, por lo que se puede afirmar que, aunque los eventos críticos sí existieron en mayor cantidad durante el 2023, estos fueron eventos de magnitudes menores que los reportados en el año 2022. Los valores numéricos del promedio y la desviación estándar para cada uno de los años se pueden resumir en la siguiente tabla, mediante el formato $\bar{x} + \sigma$.

Tabla XXXI. Promedio de duraciones reportadas para eventos de perturbación de frecuencia, según su clasificación, para cada año del estudio.

Year	Minor	Moderate	Critical
2022	0.064514±0.010767	0.160746±0.084802	0.830586±0.377604
2023	0.065811±0.011437	0.179288±0.116355	0.728146±0.125513

5.3.2. Tendencias mensuales.

Según la clasificación, se realiza la contabilización de los eventos que ha existido durante los dos años de estudio, los cuales se han presentado en un gráfico de barras. La contabilización de los eventos ha sido apilada, estando 2023 en un tono más claro que los datos de 2022.

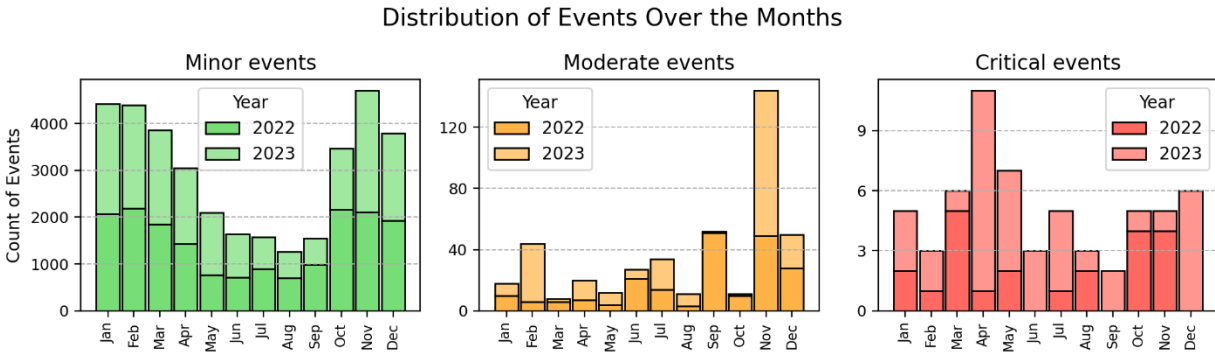


Figura 5.11. Contabilización de eventos de perturbación de frecuencia, por mes.

Para el caso de los eventos de perturbación leves, se puede apreciar que existe una tendencia general de una reducción en la cantidad de eventos que se hace evidente en el periodo de mayo hasta septiembre, coincidiendo bastante con el periodo mencionado para cuando se mencionó lo de la desviación estándar. Para el caso de los eventos de magnitud moderada y crítica, no existe ningún patrón que pueda ser evidentemente resaltado, incluso a pesar de tener datos de dos años. Al menos de forma mensual, no se puede visualizar que estas dos últimas clasificaciones presenten alguna tendencia evidente de la cual se pueda sacar alguna conclusión concreta.

Numéricamente se puede observar que la cantidad de eventos menores, en que la frecuencia alcanza y supera los 0.05 Hz, se promedia en 1500 eventos mensuales. Para el caso de los eventos moderados se promedia en 20 eventos mensuales. Finalmente, para los eventos críticos rondan un promedio de 2.54 mensuales.

Sin embargo, se puede observar que este no es un comportamiento que se mantenga constante a lo largo de los meses, siendo que existen meses donde los eventos leves pueden llegar a superar los 2000 eventos, mientras en otros meses no superan 1000. Para el caso de los eventos moderados y críticos, existen meses en que la cantidad de eventos ha sido cero, pero existen otros meses donde la cantidad de eventos llega a ser bastante más alto que otros meses.

Numéricamente se resume la contabilización de cada mes en la siguiente tabla:

Tabla XXXII. Contabilización de eventos registrados, según su clasificación, para cada mes y año.

	Minor	Moderate	Critical
January 2022	2070	10	2
February 2022	2183	6	1
March 2022	1853	6	5
April 2022	1431	7	1
May 2022	772	4	2
June 2022	717	21	0
July 2022	899	14	1
August 2022	698	3	2
September 2022	984	51	0
October 2022	2164	10	4
November 2022	2104	49	4
December 2022	1925	28	0
January 2023	2357	8	3
February 2023	2214	38	2
March 2023	2007	2	1
April 2023	1613	13	10
May 2023	1330	8	5
June 2023	921	6	3
July 2023	671	20	4
August 2023	569	8	1
September 2023	561	1	2
October 2023	1306	1	1
November 2023	2600	95	1
December 2023	1868	22	6

Y a su vez, para cada uno de los meses (sumados los resultados de cada año), se presenta:

Tabla XXXIII. Contabilización de eventos registrados, según su clasificación, para cada mes (sumados)

	Minor	Moderate	Critical
January	4427	18	5
February	4397	44	3
March	3860	8	6
April	3044	20	11
May	2102	12	7
June	1638	27	3
July	1570	34	5
August	1267	11	3
September	1545	52	2
October	3470	11	5
November	4704	144	5
December	3793	50	6

5.3.3. Tendencias para cada día de la semana

Según la clasificación, se realiza la contabilización de los eventos que ha existido durante los dos años de estudio, para cada uno de los días de la semana. Los resultados se presentan en la figura.

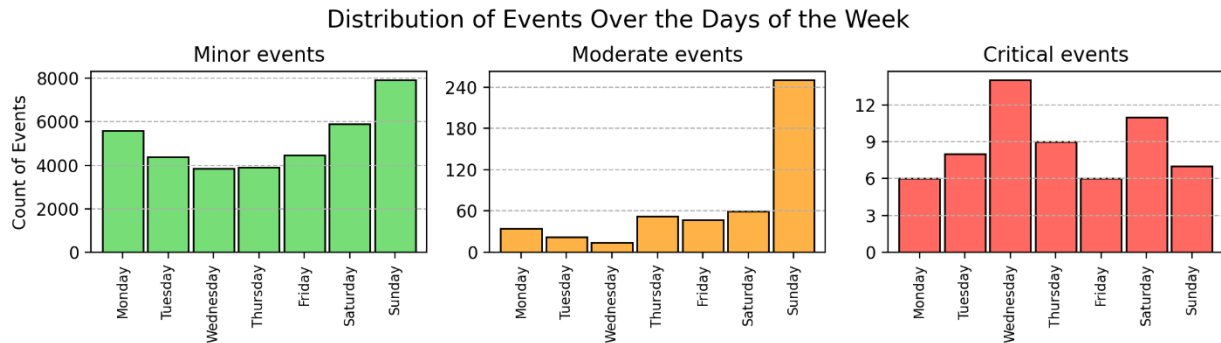


Figura 5.12. Contabilización de eventos de perturbación de frecuencia, por día de la semana.

Para el caso de los eventos de perturbación menores y moderados, se puede observar que la mayor cantidad de estos eventos se han dado durante los días domingo, siendo para los eventos menores una tendencia que ocurre de a poco, mientras para los eventos moderados es una tendencia brusca y absoluta. Para los eventos críticos, estos se presentaron con mayor frecuencia durante el miércoles, pero es difícil poder afirmar que esto es una tendencia marcada como lo que se pudiera afirmar con las otras dos clasificaciones.

Un resumen numérico es lo siguiente:

Tabla XXXIV. Contabilización de eventos registrados, según su clasificación, para cada día de la semana.

	Minor	Moderate	Critical
Monday	5575	35	6
Tuesday	4380	22	8
Wednesday	3852	14	14
Thursday	3916	52	9
Friday	4464	47	6
Saturday	5896	59	11
Sunday	7909	250	7

5.3.4. Tendencias para cada hora del día.

Cuando hablamos de tendencias horarias, nos referimos a analizar el comportamiento del valor de la frecuencia teniendo en consideración las 24 horas existentes en un día. Las variables temporales hacen referencia a que se tendrá en consideración la influencia que puede tener el realizar el análisis considerando únicamente los días laborales y fines de semana.

Algunas aclaraciones para evitar malentendidos, que serán utilizados para este documento:

- No es lo mismo decir «Son las 8:00 horas» que «Es la hora 8», al menos en este documento. Para el primer caso, nos estamos refiriendo a las 8 en punto, mientras que para el segundo nos referimos a toda la hora en sí, desde las 8:00 hasta las 8:59.
- Cuando hablemos de intervalos de horas, normalmente podríamos encontrarnos con un ejemplo el intervalo de las 14:00 hasta las 16:00 horas. Este intervalo refiere a que transcurre desde las 14:00 hasta las 15:59, no incluye a las 16:00.

Se establecen rangos de horarios para el análisis horario, principalmente para lo que sería la contabilización de los registros de eventos de perturbación de frecuencia. Estos intervalos serán de tres horas de duración, y abarcarán las siguientes horas:

- *Madrugada temprana*: Desde las 00:00 hasta las 03:00 horas.
- *Madrugada tardía*: Desde las 03:00 hasta las 06:00 horas.
- *Mañana temprana*: Desde las 06:00 hasta las 09:00 horas.
- *Mañana tardía*: Desde las 09:00 hasta las 12:00 horas.
- *Tarde temprana*: Desde las 12:00 hasta las 15:00 horas.
- *Tarde tardía*: Desde las 15:00 hasta las 18:00 horas.
- *Noche temprana*: Desde las 18:00 hasta las 21:00 horas.
- *Noche tardía*: Desde las 21:00 hasta las 24:00 horas.

Los intervalos propuestos para la complementación del análisis son únicamente propuestas realizados por nosotros, y no tienen que ser de necesario seguimiento en caso de que se desee realizar una investigación de esta índole. Se considera que el usar los intervalos de tres horas establecidos puede dar una buena idea del comportamiento del sistema eléctrico de potencia para distintos momentos de los días, evitando la sobresaturación de información que podría ser el manejar todo acerca de los eventos durante 24 horas individuales.

En base a todo lo aclarado anteriormente, se realiza la contabilización de los eventos que ha existido durante los dos años de estudio, para cada intervalo de horas para cada una de las clasificaciones de evento. Los resultados se presentan en la figura.

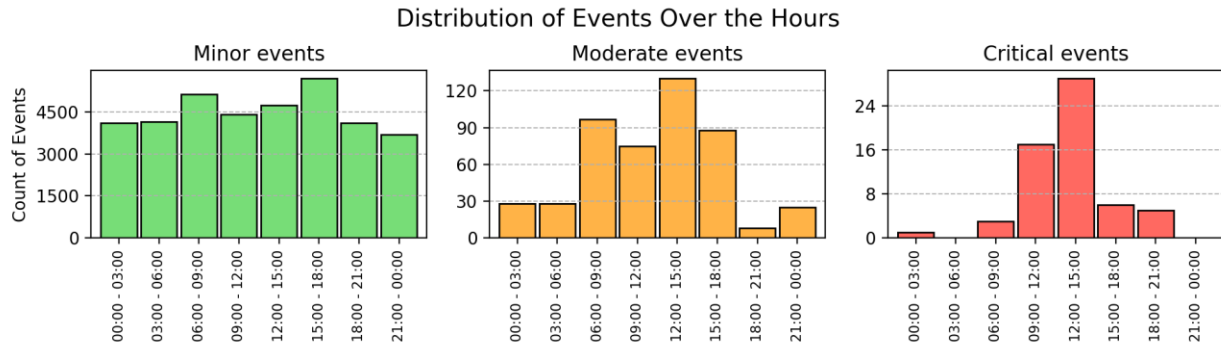


Figura 5.13. Contabilización de eventos de perturbación de frecuencia, por cada intervalo de horas del día.

Se puede visualizar que el momento en que se reportan más perturbaciones, independientemente de la clasificación, es durante el intervalo de la tarde. Aunque para el caso de la clasificación de eventos menores, los horarios en que se da la mayor cantidad de eventos son en la mañana temprana y en la tarde tardía. Para las otras dos clasificaciones (moderadas y críticas), buena parte de los eventos de perturbación se ha dado en el intervalo de las 12 a las 15 horas, aunque esto no exceptúa al resto de horarios a que ocurra un evento.

Tabla XXXV. Contabilización de eventos registrados, según su clasificación, para cada día de la semana.

	Minor	Moderate	Critical
00:00 - 03:00	4092	28	1
03:00 - 06:00	4147	28	0
06:00 - 09:00	5130	97	3
09:00 - 12:00	4404	75	17
12:00 - 15:00	4731	130	29
15:00 - 18:00	5686	88	6
18:00 - 21:00	4106	8	5
21:00 - 00:00	3696	25	0

CAPITULO 6. ANÁLISIS PRELIMINAR DE LA FRECUENCIA ELÉCTRICA Y LOS FLUJOS DE POTENCIA REGIONALES.

6.1. Contabilización de eventos reportados en el SER por el EOR.

El Ente Operador Regional realiza planificar la Red de Transmisión Regional (RTR), dirige y coordina la operación técnica del Sistema Eléctrico Regional (SER) [68], realizando un monitoreo en tiempo real de los flujos de potencia existentes en la red entre los países pertenecientes. El EOR tiene públicos, en su sitio web <https://www.enteoperador.org/>, los reportes preliminares de eventos de pérdida de generación o carga más relevantes que han existido durante 2022 y 2023. En base a esos reportes, se ha decidido observar uno por uno, y realizar una contabilización de eventos ocurridos por países del SER, además de extraer los datos más relevantes.

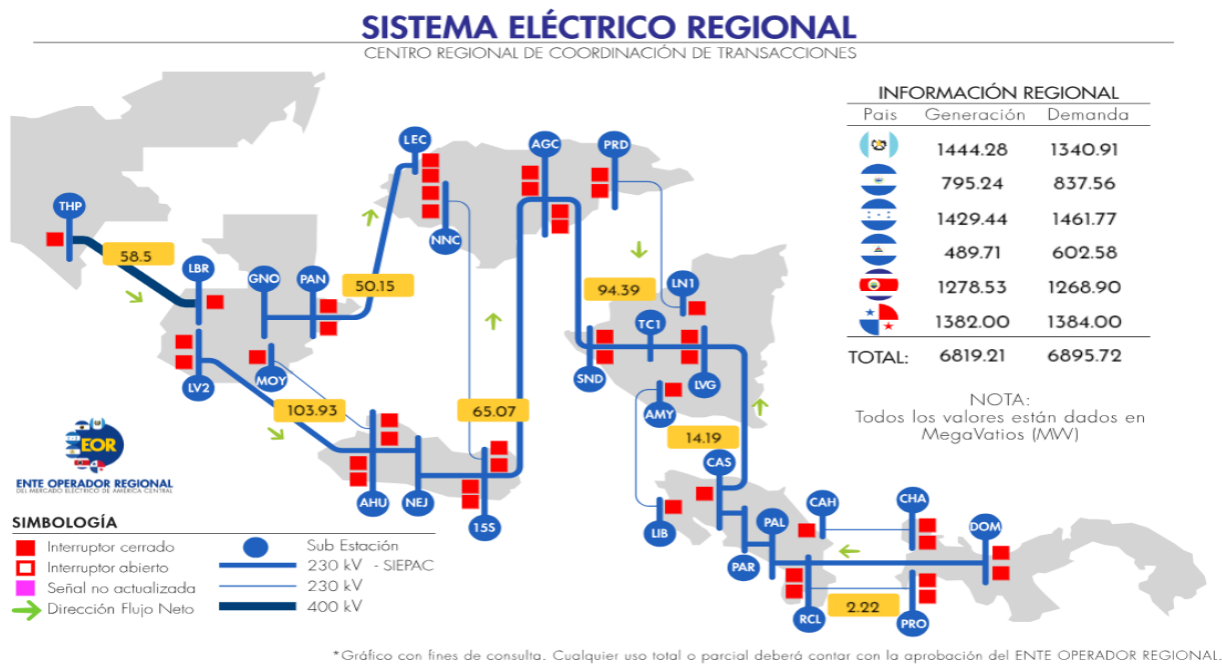


Figura 6.1. Esquema del Sistema Eléctrico Regional de América Central (SER). Extraído de [68].

Para el periodo de estudio, que comprende 2022 y 2023, se reportó la siguiente cantidad de eventos de perturbación por cada uno de los países pertenecientes al Sistema Eléctrico Regional. También se han contabilizado los eventos ocurridos en el Sistema Eléctrico Mexicano, dado que este tiene un impacto en la región (el cual se profundizará más adelante).

Nota: La etiqueta «SER» hace referencia a eventos provocados debido a pérdidas en múltiples países.

Tabla XXXVI. Contabilización de eventos reportados por el Ente Operador Regional, durante el periodo 2022 a 2023. Realizado en base a la contabilización de reportes preliminares, los cuales son públicos en su sitio web.

	GUA	SLV	HON	NIC	CRI	PAN	SER	MEX
Count	62	45	175	43	32	166	23	12

El total de eventos reportados por el EOR durante el periodo fue de 558, y se puede observar que la mayor cantidad de los eventos de pérdida de generación/carga se han dado por Honduras y Panamá, mientras que los eventos cuyo origen se haya dado en El Salvador solo representan menos del 10%. Este análisis no toma en cuenta aspectos como la generación instalada en cada país. Para todos estos eventos que sí tuvieron un impacto en el valor de la frecuencia, desde clasificación menor hasta crítico, se puede consultar en la siguiente tabla.

Tabla XXXVII. Contabilización de eventos reportados por el Ente Operador Regional, durante el periodo 2022 a 2023, teniendo en consideración solo los eventos en que se observaron perturbaciones de frecuencia.

Country	GUA	SLV	HON	NIC	CRI	PAN	SER	MEX	TOTAL
Minor	11	7	26	10	6	29	0	0	89
Moderate	15	0	4	6	0	10	10	10	55
Critical	12	3	13	3	3	12	13	2	61
TOTAL	38	10	43	19	9	51	23	12	205

Para la realización de la anterior tabla, se hizo un cruce de datos entre los registros obtenidos por el algoritmo desarrollado en el capítulo anterior y los eventos reportados por el EOR. Dado que cada reporte del EOR contiene el inicio del evento, y nuestro registro de eventos de perturbación de frecuencia tiene la hora de inicio y finalización, se hizo lo descrito en el siguiente pseudocódigo:

```
# Verificando si durante el evento de pérdida de generación/carga existió
perturbación en la frecuencia.
for event_EOR in events_EOR_list:
    time_EOR = event_EOR['Time']

    for event_freq in events_freq_list:
        start_time = event_freq['Start Time']
        end_time = event_freq['End Time']

        if (start_time - 60s) <= time_EOR <= (end_time + 60s):
            freq_event_cond = True
            Se procede a copiar los datos de la frecuencia
            break
```

Mediante este enfoque, Honduras y Panamá aún presentan las mayores cantidades de eventos, pero para se les une también Guatemala. Si observamos en El Salvador, solo se iniciaron 10 eventos de los 207 registrados por el Ente Operador Regional en que la frecuencia se vio afectada. El siguiente grafico de cajas y bigotes muestra la distribución que ha tenido el valor de la frecuencia para cada uno de los eventos reportados, según el origen de la pérdida de carga o generación reportada.

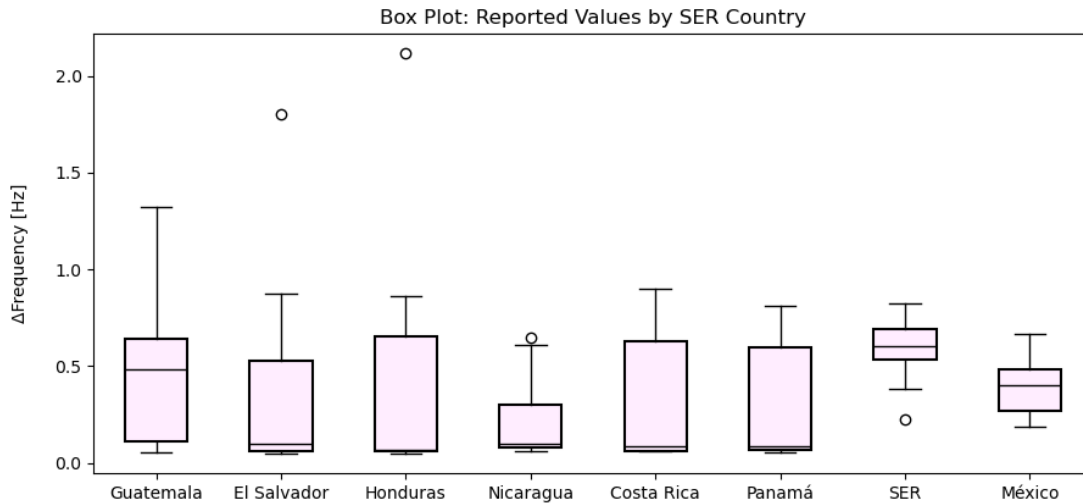


Figura 6.2. Diagrama de cajas y bigotes: Desviaciones máximas que ha presentado la frecuencia eléctrica en eventos de pérdida de generación/carga, según el origen. Realizado con reportes del EOR de 2022 y 2023.

Y la siguiente tabla muestra los datos:

Tabla XXXVIII. Parámetros de los diagramas de caja y bigotes.

Country	L1	Q1	Q2	Q3	L2
Guatemala	0.053	0.11195	0.487	0.640175	1.3203
El Salvador	0.051	0.058925	0.0972	0.52625	0.874
Honduras	0.051	0.061	0.07	0.65205	0.8637
Nicaragua	0.059	0.0784	0.0964	0.29945	0.61
Costa Rica	0.059	0.062	0.0891	0.6277	0.9027
Panamá	0.052	0.0646	0.0869	0.597	0.813
SER	0.383	0.5365	0.606	0.693	0.8255
México	0.19	0.266725	0.4	0.4816	0.669

Podemos advertir que los resultados vistos acá, utilizando los reportes del EOR, no es algo definitivo; existiendo información limitada.

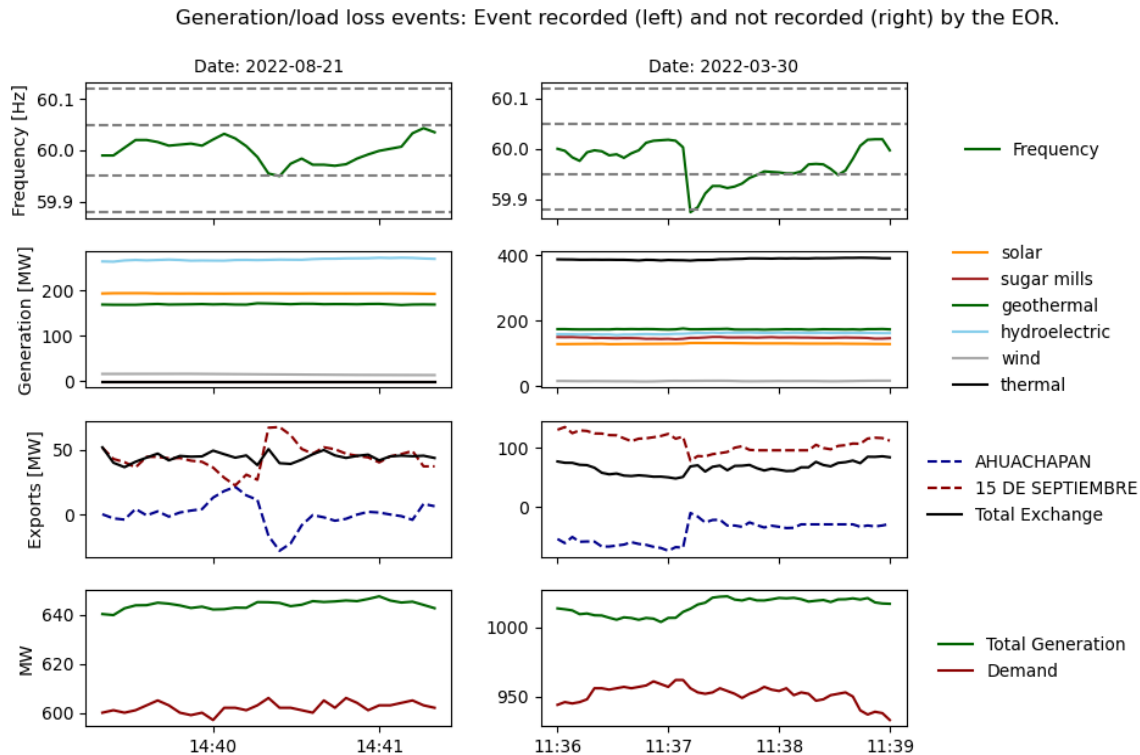


Figura 6.3. Visualización de eventos registrado y no registrado por el EOR, desde la perspectiva salvadoreña, utilizando datos brindados por la Unidad de Transacciones S.A de C.V.

En la figura se muestran dos casos:

- El 21 de agosto del 2022 es un evento registrado por el EOR, sin embargo, este no tuvo un impacto significativo en la frecuencia eléctrica. El EOR reportó un evento de pérdida ocurrido en territorio hondureño, según el reporte N°147-08-2022.
- El 30 de marzo del 2022 es un evento de perturbación de frecuencia de clasificación moderada. Puede observarse que en el momento existió un aumento en la exportación de energía a Guatemala (desde SE Ahuachapán) y reducción hacia Honduras (desde SE 15 de septiembre). El EOR no tiene un informe público sobre dicho evento.

Nota: La única clasificación en la que se observa que se tiene informe para todos los eventos es para las perturbaciones de frecuencia **críticas**.

6.2. Interconexión Sistema Eléctrico Regional – Sistema Eléctrico Mexicano.

Las transferencias de potencia desde México a través de la interconexión Los Brillantes, revisten de gran importancia para la estabilidad del Sistema Eléctrico Regional, afirma el Ente Operador Regional en [69]. Durante eventos de desbalance de los flujos de potencia, principalmente el déficit de generación, el flujo existente desde el Sistema Eléctrico Mexicano (SEM) juega un papel importante en la estabilidad del Sistema Eléctrico Regional. El EOR autorizó la puesta en servicio de la interconexión Guatemala-México, a partir del 26 de febrero de 2010 [70].

En el año 2019, se hizo un estudio para determinar los límites de transferencia de potencia desde el Sistema Eléctrico de México hacia el Sistema Eléctrico Regional, para la época húmeda [71]. Siendo este el último estudio realizado, el cual está público en su portal web, destacaremos las siguientes conclusiones y especificaciones técnicas mencionados, los cuales son útiles para contextualizar la importancia de la interconexión y sus limitantes.

- De conformidad con los resultados del presente estudio de actualización de los valores límite de transferencia de potencia desde el SEM hacia el SER, se presenta de escenarios que durante la demanda mínima (22:00 – 05:59) el límite será de 220 MW, mientras que para demanda media y demanda máxima (06:00 – 21:59) será de 240 MW.
- Existe una sola línea de interconexión SEM-SER, correspondiente a MEX-GUA, que transporta todo el flujo de México a los países del SER, el cual está conformado por dos bancos de transformadores en paralelo. Existen limitaciones identificadas cuando se aplican los Criterios de Calidad, Seguridad y Desempeño (CCSD) del Reglamento del Mercado Eléctrico Regional (RMER).
- Existen dos Esquemas de Control Suplementario (ECS) que inciden en la determinación de los límites seguros de las transferencias de potencia México-SER:
 - Cuando la transferencia de potencia México-SER aumenta por encima de 300 MW y el voltaje de la subestación mexicana Tapachula sea igual inferior de 0.97 p.u, Ambas condiciones deben de cumplirse simultáneamente, durante 11 ciclos.
 - Cuando se detectan oscilaciones inter-área inestables, para las oscilaciones inter-área pobremente amortiguadas.

En base a los reportes públicos del Ente Operador Regional, para el periodo 2022 – 2023, se ha contabilizado las ocasiones en que la interconexión SER-SEM ha sido abierto, y a su vez las condiciones en que fue abierto. La apertura se puede dar por:

- Apertura automática, debido a que la transferencia de potencia es superior a 300 MW y el voltaje en Tapachula (México) se ha visto reducido.
- Apertura manual, que son maniobras de sincronización, debido a oscilaciones existentes en el Sistema Eléctrico Regional.
- Aperturas debido a fallas en las líneas de transmisión de la interconexión SER-SEM, como lo podría ser alguna caída de árboles o fallas en infraestructura.
- Aperturas debido a mantenimientos programados.

La contabilización se ha realizado en la siguiente tabla:

Tabla XXXIX. Contabilización de aperturas en la interconexión SER-SEM, para el periodo 2022 - 2023.
Elaborado mediante reportes públicos del Ente Operador Regional.

	OPEN	SYNC	FAULT	MAINTENANCE	CLOSED	TOTAL
Count	62	9	34	6	447	558

Se puede observar que la cantidad de veces en que se abra la interconexión SER-SEM es considerable, y a su vez es considerable la cantidad de fallas que ha existido durante este periodo. La cantidad de eventos de perturbación de frecuencia ocurridos, según clasificación, se contabilizan de la siguiente forma:

Tabla XL. Contabilización de aperturas en la interconexión SER-SEM, para el periodo 2022 – 2023, y su impacto en la frecuencia eléctrica.

	OPEN	SYNC	FAULT	MAINTENANCE	CLOSED	TOTAL
Minor	1	0	1	0	88	89
Moderate	10	9	24	5	7	55
Critical	51	0	9	0	1	61
TOTAL	62	9	34	5	96	205

Se puede hacer evidente que existen mayores perturbaciones en el valor de la frecuencia cuando se han dado aperturas de la interconexión SER-SEM, que cuando esta apertura no ocurrió.

Si observamos un diagrama de caja y bigotes, se puede hacer evidente que se presentan mayores desviaciones de frecuencia cuando está abierta la interconexión que cuando se encuentra cerrado. El mayor impacto ocurre cuando es abierto debido a desbalances de generación-carga existentes en el Sistema Eléctrico Regional (aplicando el primer esquema de control), seguido por problemas existentes en las líneas de transmisión, maniobras de sincronización, quedando que el menor impacto ha sido cuando la línea ha estado en mantenimiento o directamente cerrada.

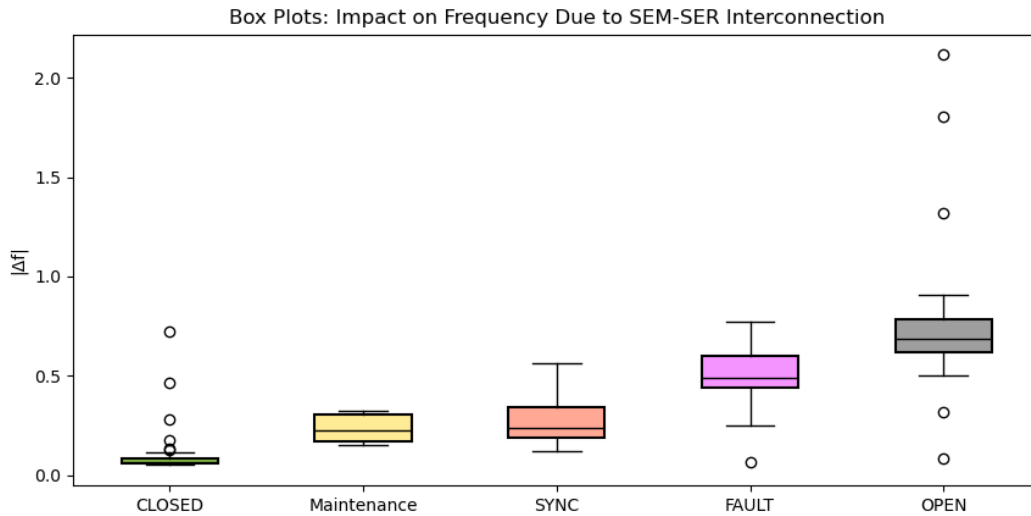


Figura 6.4. Distribución de la frecuencia eléctrica (medida en El Salvador) debido a la interconexión SEM-SER.

Los valores numéricos se muestran en la siguiente tabla:

Tabla XLI. Parámetros de los diagramas de caja y bigotes.

	L1	Q1	Q2	Q3	L2
CLOSED	0.051	0.059	0.0646	0.082	0.1148
MAINT.	0.154	0.1728	0.2276	0.3046	0.3238
SYNC	0.1224	0.189	0.2391	0.3415	0.561
FAULT	0.25	0.438	0.4912	0.602	0.775
OPEN	0.503	0.621	0.68745	0.785375	0.9101

6.3. Contabilización de eventos reportados en El Salvador por el EOR: Causas.

Para el caso de El Salvador, el EOR contabilizó un total de 45 eventos de pérdida de generación/carga en el periodo de 2022 a 2023. De estos, únicamente 10 tuvieron un impacto registrado en el valor de la frecuencia eléctrica.

Lo que nos interesa a continuación es tener estadísticas de las fallas existentes durante el periodo de estudio, teniendo registro del origen de la falla (en una generadora, problemas en las líneas de transmisión, etc.), y a su vez el registro de impacto que ha tenido este tipo de eventos. Para ello, se observaron los 45 reportes de pérdida de generación/carga del EOR, los cuales hacen referencia a fallas ocurridas en territorio salvadoreño.

Observando los reportes públicos, uno por uno, se pueden tabular los datos correspondientes. En ello, se han identificado: El ID del evento asignado por el EOR, el tipo de falla, la localización, la carga y generación desconectada y el cambio en el valor de la frecuencia dentro del SER. Los datos utilizados son exclusivamente los dados por el EOR en cada reporte, y hacen referencia a la perspectiva salvadoreña.

Tabla XLII. Resumen de pérdida de generación/carga iniciados en El Salvador, durante el periodo 2022 – 2023, reportados por el Ente Operador Regional.

ID EOR	Fault type	Localication	ΔMW_{load}	ΔMW_{gen}	Δf
02-01-2022	hydroelectric	Cerrón Grande	0.00	78.00	-0.02
04-01-2022	Transmission	SE Soyapango	70.00	0.00	-0.017
10-01-2022	hydroelectric	15 de septiembre	0	80	-0.06
11-01-2022	hydroelectric	15 de septiembre	0	90	-0.122
15-02-2022	Transmission	SE Santo Tomas / SE Pedregal	74.10	52.70	-0.053
28-03-2022	thermal	Termopuerto	0.00	70.00	-0.048
29-03-2022	Transmission	LTX Nejapa – San Martín	89.00	0.00	0.1
36-03-2022	thermal	Energía del Pacífico	0	72.8	-0.041
38-03-2022	thermal	Energía del Pacífico	0	69	-0.03

ID EOR	Fault type	Localication	ΔMW_{load}	ΔMW_{gen}	Δf
44-03-2022	thermal	Energía del Pacífico	0	148.5	-0.739*
61-04-2022	Transmission	SE Nuevo Cuscatlán	66.21	0	0.05
65-05-2022	Transmission	SE Nejapa	60	0	0.09
69-05-2022	Transmission	SE San Bartolo	60	0	0.028
72-05-2022	thermal	Energía del Pacífico	0	100	-0.022
74-05-2022	Transmission	SE Acajutla	0	561.4	-2.011**
90-06-2022	thermal	Energía del Pacífico	0	90	-0.06
105-06-2022	hydroelectric	C.H 15 de septiembre	0	90	-0.03
106-07-2022	Transmission	LTX Nejapa - Nejapa Power	83	0	0.071
111-07-2022	hydroelectric	C.H 15 de septiembre	0	60	-0.02
161-09-2022	hydroelectric	C.H15 de septiembre	0	90	-0.042
164-09-2022	Transmission	SE Nuevo Cuscatlán	100	0	0.053
180-09-2022	hydroelectric	C.H 15 de septiembre	0	90	-0.047
181-09-2022	Transmission	SE Santo Tomas	60	0	0
211-10-2022	Transmission	LTX Santo Tomás – Pedregal	68	0	0
217-10-2022	hydroelectric	C.H Cerrón Grande	0	84.5	0
239-11-2022	Transmission	SE 15 de septiembre / SE Berlín	12	125	-0.105
241-11-2022	hydroelectric	C.H 15 de septiembre	0	90	-0.033
244-11-2022	hydroelectric	C.H 15 de septiembre	0	90	-0.071
255-12-2022	Transmission	LTX Sonsonate – Acajutla	65	0	0
256-12-2022	hydroelectric	C.H Cerrón Grande	0	84	0
09-01-2023	hydroelectric	C.H Cerrón Grande	0	65	-0.078
14-02-2023	Transmission	SE Nejapa	0	120	0.05

ID EOR	Fault type	Localication	ΔMW_{load}	ΔMW_{gen}	Δf
48-03-2023	thermal	Energía del Pacífico	0	170	-0.045
88-05-2023	thermal	Nejapa Power	0	120	-0.034
89-05-2023	geothermal	LaGeo Ahuachapan	0	81.29	0
91-05-2023	thermal	Energía del Pacífico	0	340	-0.904*
102-05-2023	Transmission	LTX Nejapa – EDESAL/ANDA	103	0	0
142-06-2023	Transmission	LTX Sn Antonio Abad – Volcán	140	0	0.068
165-07-2023	Transmission	SE Ateos	65	60	0.037
171-07-2023	hydroelectric	C.H 15 de septiembre	0	90	-0.053
184-08-2023	hydroelectric	C.H 15 de septiembre	0	62	-0.042
213-09-2023	Transmission	LTX 5 de nov. - Cerrón Grande	0	60	-0.054
223-10-2023	Transmission	LTX 15 de sep. - 3 de feb.	0	60	-0.011
229-10-2023	Transmission	LTX 15 de sep. - 3 de feb.	0	60	-0.048
239-10-2023	Transmission	LTX 15 de sep. - 3 de feb.	0	60	-0.049

*Se dio apertura a la interconexión SER-MEX.

**Se dio apertura a las interconexiones entre países del SER. El valor de frecuencia corresponde a la isla de El Salvador.

La cuantificación se realiza en la siguiente tabla. Las tecnologías que no aparecen (eólico) no han presentado ninguna falla durante el periodo de estudio.

Tabla XLIII. Tipo de falla ocurrido para las pérdidas de generación/cargas reportadas en El Salvador por el Ente Operador Regional, para el periodo 2022-2023.

Transmission	Hydroelectric	Thermal	Geothermal
21	14	9	1

Una gran parte de las perdidas reportadas fueron por fallas en las líneas de transmisión, reportándose 21 en total, del total de 45 existentes. Una de estas fallas en la transmisión provocó el mayor evento de perturbación en la frecuencia reportado en El Salvador durante el periodo de estudio, correspondiente al evento N°74-05-2022 reportado por el Ente Operador Regional.

Gráficos de caja y bigotes pueden ilustrar de mejor forma el impacto que ha existido cada tipo de falla, reflejando la pérdida de generación/carga que ha existido. Se ha utilizado el valor máximo, es decir, si existió una pérdida de generación de 100MW y una pérdida de carga de 20MW, se ha optado por utilizar 100MW como dato para la realización del grafico siguiente.

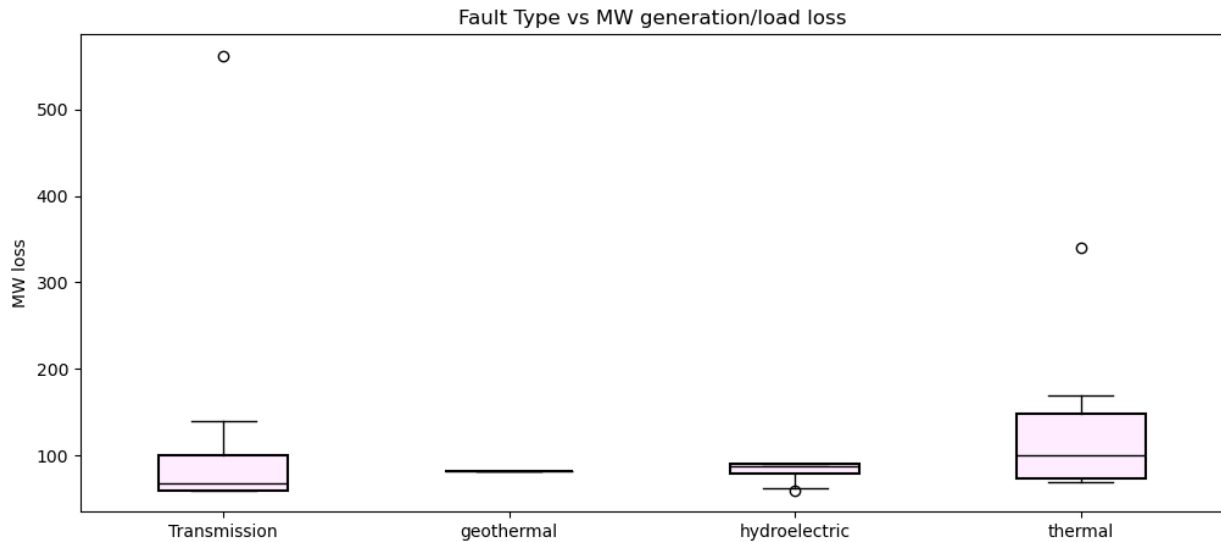


Figura 6.5. Distribución de pérdidas de generación/cargas reportadas por tipo de tecnología, según reportes del Ente Operador Regional, durante 2022 - 2023.

Visualmente puede verse que las fallas en la transmisión y en la generación térmica tienen límites superiores de valores más altos, si se le compara con el resto de las tecnologías. Las fallas en la generación hidroeléctrica, a pesar de ser contabilizadas en 14, se asemejan a las de otras tecnologías de generación como la geotérmica y solar.

Los datos numéricos se resumen en la siguiente tabla:

Tabla XLIV. Parámetros de los diagramas de caja y bigotes.

	L1	Q1	Q2	Q3	L2
Hydroelectric	62	78.5	87.25	90	90
Transmission	60	60	68	100	140
Thermal	69	72.8	100	148.5	170
Geothermdal	81.29	81.29	81.29	81.29	81.29

CAPITULO 7. IMPLEMENTACIÓN DE MODELO PREDICTIVO CON REDES NEURONALES LSTM PARA LA PREDICCIÓN DE FRECUENCIA ELÉCTRICA USANDO DATOS DEL FDR-1628

7.1. Definición de modelo predictivo

Antes de pasar a la implementación del modelo predictivo, es necesario explicar conceptos básicos como; modelo predictivo, saber que es una red neuronal y sus diferentes variaciones para comprender el funcionamiento de estas.

Un modelo predictivo es una de las ramas de un análisis avanzado que hace predicciones sobre, posibles resultados futuros. Su base principal son los datos estadísticos. La lógica de los modelos predictivos consiste en que, recopilando datos, predecirá lo que pueda pasar en un futuro o sugerir acciones para obtener resultados óptimos y proponer solución a cualquier inconveniente [72].

Un modelo predictivo puede beneficiar en lo siguiente:

- ✓ Proporciona herramientas para estimar datos desconocidos y que requieren de procesos analíticos demasiado extensos.
- ✓ Fortalece decisiones a largo plazo.
- ✓ Nos ayuda a encontrar anomalías y anticipar comportamientos.

7.2. Breve introducción sobre las Redes Neuronales.

7.2.1. Red neuronal básica

Una red neuronal es un modelo de aprendizaje automático que imita el funcionamiento del cerebro humano para reconocer patrones y tomar decisiones. Está organizada en capas de nodos interconectados: una capa de entrada, una o más capas ocultas, y una capa de salida. Cada nodo tiene su propia ponderación y umbral, y funciona recibiendo valores de entrada, realizando cálculos internos y generando un valor de salida. [73], [74] , [75]. Internamente, una neurona realiza una suma ponderada de los valores de entrada, con cada entrada teniendo un peso que determina su influencia. La salida se compara con un umbral; si la supera, el nodo se activa y transmite información a la siguiente capa. Si no, no se activa [73], [74] [75].

Matemáticamente, las operaciones de una neurona se expresan como una ecuación lineal, similar a un modelo de regresión lineal. Cada entrada se multiplica por su peso correspondiente y se suma un término de sesgo (bias), una constante que desplaza la función lineal. El sesgo se considera una conexión adicional con un valor de entrada constante y su propio peso ajustable durante el entrenamiento [73], [74], [75]. La ecuación que puede representar matemáticamente como:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

Donde:

- y : Es la salida de la neurona.
- f : Es la función de activación.
- x_i : Son las entradas de la neurona.
- w_i : Son los pesos asociados a cada entrada.
- b : Es el sesgo bias.

En la siguiente imagen se muestra una red neuronal y sus respectivas partes, así mismo se visualiza el diagrama de su funcionamiento:

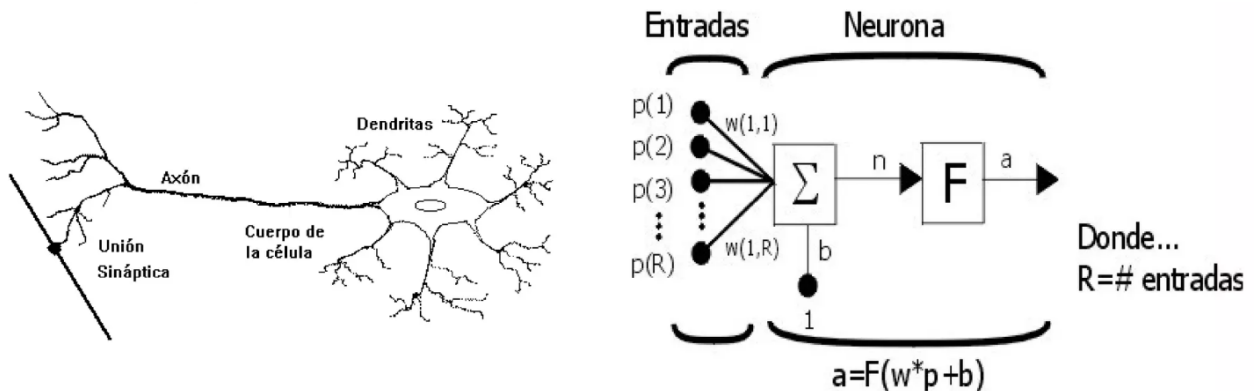


Figura 7.1. Estructura de una red neuronal. Extraído de la red.

Componentes de una red neuronal

Una red neuronal tiene los siguientes componentes fundamentales, los cuales son:

- ✓ Capas: Capas de entrada, Capas ocultas y Capas de salida.
- ✓ Neuronas.
- ✓ Función de activación.
- ✓ Pesos y sesgos.

1. Capas: En una red neuronal existen tres tipos de capas y estas son [74] [76], [77].:

- A) **Capa de entrada:** Aquí se reciben las entradas, conocidas como variables independientes o predictores X. Estos datos pueden provenir de diversas fuentes externas, como texto, imágenes, archivos de audio o video
- B) **Capas ocultas:** Se encuentran entre la entrada y la salida, y son responsables de procesar y extraer características relevantes del modelo. Pueden variar en número; una red puede tener una sola capa oculta o múltiples capas, dependiendo de la complejidad del problema.
- C) **Capa de salida:** Esta capa genera los resultados de la red neuronal. En problemas de regresión, puede ser un valor numérico; en problemas de clasificación, puede ser una categoría binaria o multicategoría.

En la siguiente imagen se visualizará un diagrama de cómo está compuesta una red neuronal.

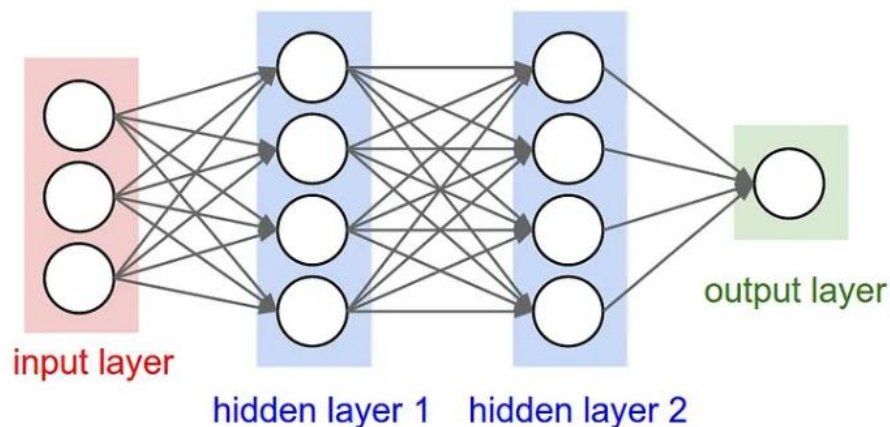


Figura 7.2. Composición de capas de una red neuronal. Extraído de [77].

2. Neuronas: Las neuronas son las unidades fundamentales de procesamiento en una red neuronal. Su función principal es recibir datos, realizar cálculos básicos y transmitir el resultado a las siguientes neuronas [74].

3. Pesos y sesgos: En una red neuronal, las conexiones entre las neuronas se establecen mediante pesos y sesgos. Los pesos determinan la fuerza de la conexión entre neuronas, similar a como las sinapsis en el cerebro biológico indican la fuerza entre las neuronas. Cada conexión entre las neuronas tiene un peso que controla cuanto afecta la salida de una neurona a la entrada de la siguiente capa. Los pesos se ajustan durante el entrenamiento para mejorar el rendimiento de la red al minimizar errores [74].

El sesgo, por otro lado, es una entrada constante añadida a cada neurona de una capa, independiente de las entradas de la capa anterior. Actúa como un término intercepto que permite a la red neuronal activarse incluso cuando todas las entradas son cero. Durante el entrenamiento, tanto los pesos como los sesgos se optimizan para mejorar la capacidad de la red para predicciones precisas [74].

4. Función de activación: La función de activación es muy importante porque introduce la no linealidad, es una red neuronal al aplicar una transformación a la salida de cada neurona. Esto es esencial porque permite a la red aprender y modelar relaciones complejas y no lineales entre las entradas y salidas [74], [77].

Una neurona biológica puede estar activa o inactiva esto quiere decir que tiene un estado de activación. Las neuronas artificiales también tienen diferentes estados de activación, otras solo tienen dos, al igual que las biológicas, pero otras pueden tomar cualquier valor dentro de un conjunto determinado [78], [77].

La función de activación es responsable de calcular el estado de actividad de una neurona en una red neuronal, eso quiere decir que, decide si una neurona debe de activarse y, si es así, en qué medida. Toma la entrada neta de la neurona, resta el umbral, y luego transforma este valor en un estado de activación. El resultado de esta transformación suele estar en un rango específico, generalmente entre 0 y 1, o entre -1 y 1. Esto permite representar neuronas que pueden estar completamente inactivas (0 o -1) o completamente activas [78], [77].

7.2.2. Red neuronal Recurrente

Una red neuronal recurrente (RNN) es un tipo de red neuronal con una "memoria" interna que le permite recordar información de entradas anteriores, lo cual es útil para tareas que involucran secuencias de datos. Las RNN realizan la misma operación en cada paso de la secuencia, considerando cálculos previos para determinar la salida actual y reenviando esta salida a la red para influir en futuras decisiones [79], [80]. En contraste, una red neuronal artificial estándar de propagación hacia adelante procesa la entrada actual sin considerar la secuencia temporal de los datos. La información se mueve en una sola dirección a través de las capas, lo que significa que estas redes no pueden retener información sobre las entradas anteriores ni predecir eventos futuros de manera efectiva. Solo utilizan lo aprendido durante su fase de entrenamiento [80], [81].

7.2.3. Red neuronal de corto y largo plazo (LSTM)

En una RNN, la salida también depende de las salidas anteriores en la secuencia, gracias a las conexiones recurrentes que envían información de vuelta a la red. Mientras que en una red tradicional cada neurona procesa las entradas y envía el resultado al siguiente nodo, en una RNN los nodos funcionan como pequeñas unidades de memoria que mantienen y procesan continuamente la información [80], [81]. Si una RNN produce una predicción incorrecta, ajusta sus parámetros mediante un proceso llamado retropropagación, mejorando su precisión [81].

Las redes neuronales recurrentes (RNN) básicas presentan limitaciones significativas en términos de memoria a largo plazo. Una RNN básica toma como entradas el dato actual y el estado oculto anterior, y produce dos salidas: la predicción y el estado oculto actualizado. Esta estructura permite a la RNN analizar secuencias de datos de entrada y generar predicciones [82], [83].

Sin embargo, la capacidad de memoria de corto plazo de las RNN básicas limita su efectividad en el manejo de secuencias largas. Esto se debe a que el efecto del estado oculto inicial se diluye a medida que la secuencia se alarga, lo que hace que la influencia de la memoria a largo plazo sea mínima [82], [83].

Para superar esta limitación, se utilizan las Redes de Memoria a Corto y Largo Plazo (LSTM). Las LSTM son una variante de las RNN diseñadas específicamente para aprender y retener dependencias a largo plazo. Son efectivas en una amplia variedad de problemas y representan el tipo de RNN más utilizado en la actualidad [82], [83].

A diferencia de las RNN simples, que utilizan una única capa tangente hiperbólica (tanh) para implementar la recurrencia, las LSTM emplean una estructura más compleja que involucra cuatro capas. Estas capas interactúan de manera específica para gestionar el estado oculto, permitiendo a las LSTM retener información relevante a lo largo de secuencias más largas. Esto las capacita para superar las limitaciones de memoria de las RNN básicas [82], [83].

A continuación, se presenta un diagrama que ilustra las transformaciones aplicadas al estado oculto en el paso de tiempo t en una LSTM, destacando cómo estas capas contribuyen a la gestión efectiva de la memoria a largo plazo [82].

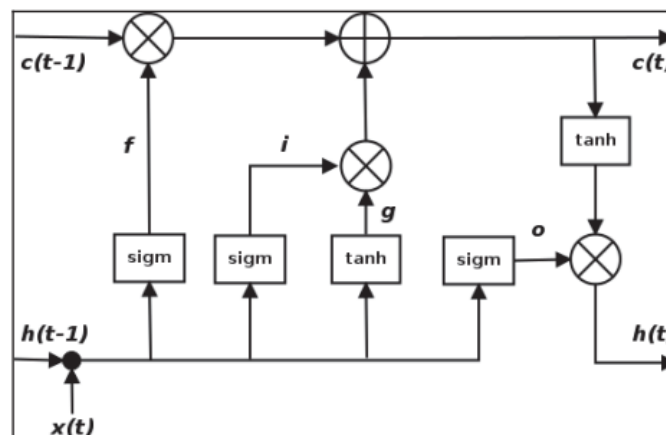


Figura 7.3. Arquitectura de una Red Neuronal tipo LSTM. Extraído de [82]

En la figura anterior se puede representar como está estructurada una red LSTM. En la parte superior del diagrama, se encuentran el estado oculto de la celda (c), que representa la memoria interna de la unidad. La línea que atraviesa la parte inferior es el estado oculto (h). Las puertas (i , f , o y g) son los mecanismos que la celda LSTM utiliza para resolver el problema del gradiente evanescente. Durante el entrenamiento, el LSTM aprende los parámetros de estas puertas [82].

Para comprender cómo estas puertas modulan el estado oculto del LSTM, consideremos las ecuaciones que muestran cómo se calcula el estado oculto h_t en el tiempo t a partir del estado oculto h_{t-1} del paso de tiempo anterior:

1. *Puerta de entrada (i_t):* La puerta de entrada decide cuánta información de la entrada actual x_t y del estado oculto anterior h_{t-1} debe ser almacenada en la memoria interna. La función sigmoide (σ) garantiza que los valores estén entre 0 y 1 [82].

$$i = \sigma(W_i h_{t-1} + U_i x_t)$$

2. *Puerta de olvido (f_t)*: La puerta del olvido determina cuanta información de la memoria interna anterior (c_{t-1}) debe ser olvidada. Al igual que la puerta de entrada, utiliza una función sigmoide para limitar los valores entre 0 y 1 [82].

$$f = \sigma(W_f h_{t-1} + U_f x_t)$$

3. *Puerta de salida (o_t)*: La puerta de salida controla cuanto del valor de la memoria interna debe ser usado para calcular el estado oculto siguiente. Esta puerta también utiliza la función sigmoide para regular valores [82].

$$o = \sigma(W_o h_{t-1} + U_o x_t)$$

4. *Nueva memoria candidata (g_t)*: Esta ecuación calcula un nuevo vector de memoria candidata que puede ser añadido a la memoria interna. La función tangente hiperbólica asegura que los valores estén entre -1 y 1 [82].

$$g = \tanh(W_g h_{t-1} + U_g x_t)$$

5. *Memoria interna actualizada (c_t)*: La memoria interna se actualiza combinando la memoria interna anterior, escalada por la puerta de olvido f_t y la nueva memoria candidata, escalada por la puerta de entrada i [82].

$$c_t = (c_{t-1} \otimes f) \oplus (g \otimes i)$$

6. *Estado oculto actualizado (h_t)*: El estado oculto se actualiza utilizando la memoria interna actualizada, modulada por la puerta de salida O_t . la memoria interna se pasa a través de una función tangente hiperbólica antes de ser multiplicada por el valor de la puerta de salida [82].

$$h_t = \tanh(c_t) \otimes O$$

7.3. Preparación del entorno

Para el desarrollo de este modelo, se han utilizado el entorno de desarrollo spyder, proporcionado por anaconda. El enfoque propuesto se implementó en Python donde se utilizaron varios paquetes para el manejo de datos, entrenar el modelo y la visualización de resultados. Los paquetes utilizados incluyeron: PyTorch, numpy, pandas, sklearn, matplotlib, tqdm. Los datos utilizados para el modelo predictivo fueron recolectados de la unidad #1628 y comprenden el periodo desde septiembre del 2022 hasta junio del 2023.

Al verificar los datos de frecuencia observamos que estaban registrados en series de 100ms, proporcionando una resolución muy alta que permite un análisis detallado. Sin embargo, se consumía mucho recurso de memoria que afectaba al tiempo de ejecución.

Para facilitar el manejo y el análisis de estos, hemos optado por realizar un preprocesamiento de estos, los cuales se detallan a continuación:

1. *Re-muestreo de los datos*: Los datos originales se han re-muestreado a intervalos de 6 segundos y también de una hora, con el fin de ayudar a reducir la complejidad computacional (en el consumo de recursos como memoria).
2. *Conversión de los datos a archivo parquet*: Los archivos parquet son un formato de almacenamiento de datos en columnas diseñado para ser eficiente en cuanto espacio y velocidad, se utilizan comúnmente en entornos de Big Data para almacenar grandes volúmenes de datos de manera eficiente.
3. *Normalización de datos*: Los datos han sido normalizados utilizando un escalador, para escalar los valores de frecuencia que comprenda entre valores de un rango de 0 a 1. Esto con el fin de garantizar que el modelo LSTM funcione correctamente y las predicciones sean precisas.
4. *División en conjunto de entrenamiento y prueba*: Los datos re-muestrados se dividieron en conjuntos de entrenamiento y prueba en una proporción de 70% y 30%. El conjunto de entrenamiento, que representa el 70% de los datos, se utiliza para entrenar el modelo LSTM. El conjunto de prueba, que representa el 30% restante, se utiliza para evaluar el rendimiento del modelo. Esta división asegura que el modelo se evalúe con datos que no ha visto durante el entrenamiento, proporcionando una medida más realista de su capacidad predictiva.

7.4. Diseño e implementación de una red LSTM

7.4.1. Configuración de la Red Neuronal

El modelo LSTM que hemos construido utiliza una configuración específica diseñada para capturar y predecir los patrones en los datos de frecuencia eléctrica. A continuación, se describen los componentes y la configuración del modelo:

1. *Capas de Entrada (input_size)*: Esta es la cantidad de características en cada paso de tiempo de la secuencia de entrada. En nuestro caso, cada registro de tiempo contiene solo la frecuencia eléctrica, por lo que su valor es de 1.
2. *Capas de estado oculto (hidden_size)*: Esta es la cantidad de unidades en la capa oculta de la red LSTM. Hemos configurado en un valor de 64, lo que permite a la red capturar patrones complejos en los datos.
3. *Capas de salida (output_size)*: Esta es la cantidad de valores que la red debe predecir. En nuestro caso, es la frecuencia eléctrica en el siguiente paso de tiempo, por lo que su valor corresponde a 1.
4. *Épocas (epochs)*: Una época se refiere a un ciclo completo de entrenamiento donde el modelo pasa por todos los datos de entrenamiento una vez, en cada época los pesos del modelo se actualizan para mejorar su rendimiento. Un mayor número de épocas permite que el modelo se ajuste mejor a los datos, en nuestro caso hemos establecido un número de épocas de 800 para los datos con un remuestro de 6 segundos y 4000 para los datos remuestros a 1 hora.
5. *Función de pérdida*: La función de pérdida mide la diferencia entre las predicciones del modelo y los valores reales. En este caso utilizamos el error cuadrático medio MSE, que es adecuado para problemas de regresión.
6. *Optimizador*: Es un algoritmo que ajusta los pesos y sesgos de una red neuronal durante un proceso de entrenamiento con el fin de minimizar la función de pérdida. En esta ocasión hemos optado utilizar el optimizador Adam debido a que es uno de los optimizadores más comunes en el entrenamiento de redes neuronales.

7.4.2. Construcción del modelo

Ahora que hemos abordado conceptos relevantes del modelo LSTM podemos pasar a lo que es la construcción de este, para ello se han establecido una serie de pasos que nos ayudara a estructurar de la mejor manera nuestro modelo LSTM.

1. *Importación de librerías:* En este primer paso, realizamos la importación de librerías necesarias para la construcción de nuestro modelo,

```
import torch
import torch.nn as nn
import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt
from tqdm import tqdm
from sklearn.metrics import mean_squared_error, r2_score
```

2. *Establecimiento del directorio y cargar los datos:* En esta parte, establecemos el directorio raíz, donde se encuentran almacenados los datos con los que trabajaremos, anteriormente estos datos se convirtieron a un archivo parquet para una mejor optimización.

```
data = pd.read_parquet("data/cleanedData/datos_2022_2023.parquet")
```

3. *Realizar la partición de datos en 70% de entrenamiento y 30% prueba:* En esta parte realizamos lo que es la partición de datos para el entrenamiento y evaluación del modelo. Los datos se dividen en dos conjuntos principales, los cuales son: el conjunto de entrenamiento y el conjunto de prueba. Asignamos el 70% al conjunto de entrenamiento y 30% al conjunto de datos de prueba. En el entrenamiento el modelo se entrena con ese porcentaje de datos, aprende patrones y las relaciones de los datos. En el conjunto de prueba se ocupan para evaluar el rendimiento del modelo una vez que haya sido entrenado, después del entrenamiento el modelo se aplica al conjunto de prueba y se evalúan sus predicciones utilizando métricas de rendimiento.

```
train_data, test_data = train_test_split(data, test_size=0.3, shuffle=False)
```

4. *Escalar los datos:* En esta parte se realiza el escalado de datos, es importante realizar un escalamiento de datos para que los modelos de aprendizaje funcionen en condiciones óptimas. Cuando las características de los datos tienen diferentes escalas, los algoritmos de aprendizaje tienen dificultades para optimizar de una forma adecuada.

```
scaler = MinMaxScaler()
scaled_train_data = scaler.fit_transform(train_data)
scaled_test_data = scaler.transform(test_data)
```

5. *Crear las secuencias para el modelo:* Se crea una función de secuencia para preparar los datos de entrada y salida. El objetivo de esta función es transformar la serie temporal en múltiples secuencias de longitud fija, que posteriormente serán usadas para entrenar el modelo. Aquí establecemos una secuencia de 100 pasos, debido a que nos dará una mayor ventaja a la hora de que el modelo se entrene.

```
def create_sequences(data, seq_length):
    X, y = [], []
    for i in range(len(data) - seq_length):
        X.append(data[i:i + seq_length])
        y.append(data[i + seq_length])
    return np.array(X), np.array(y)

seq_length = 100

X_train, y_train = create_sequences(scaled_train_data, seq_length)
X_test, y_test = create_sequences(scaled_test_data, seq_length)
```

6. *Convertir a tensores:* En este paso se realiza la conversión a tensores, dado que son la estructura de datos utilizada para realizar cálculos.

```
X_train_tensor = torch.tensor(X_train, dtype=torch.float32)
y_train_tensor = torch.tensor(y_train, dtype=torch.float32)
X_test_tensor = torch.tensor(X_test, dtype=torch.float32)
y_test_tensor = torch.tensor(y_test, dtype=torch.float32)
```

7. *Definir la arquitectura del modelo:* Definir la arquitectura del modelo es construir una red neuronal recurrente que pueda aprender y capturar patrones en datos secuenciales. La arquitectura incluye una capa LSTM para manejar las secuencias de datos y una capa lineal para generar la predicción final.

```

class LSTM(nn.Module):
    def __init__(self, input_size, hidden_size, output_size, num_layers):
        super(LSTM, self).__init__()
        self.lstm = nn.LSTM(input_size, hidden_size, num_layers, batch_first=True)
        self.linear = nn.Linear(hidden_size, output_size)

    def forward(self, x):
        lstm_out, _ = self.lstm(x)
        out = self.linear(lstm_out[:, -1, :])
        return out

```

8. *Entrenar el modelo:* El objetivo del entrenamiento del modelo es ajustar los parámetros de la red para que pueda predecir con precisión los valores futuros de la frecuencia eléctrica. Esto se logra mediante un proceso iterativo donde el modelo hace predicciones, calcula el error (perdida), y ajusta sus parámetros para minimizar ese error.

La monitorización de la pérdida a lo largo de las épocas permite verificar si el modelo está aprendiendo correctamente y ajustar el número de épocas o la tasa de aprendizaje si es necesario.

Hay que tener en cuenta que, mediante el incremento de número de épocas mayor será la precisión del modelo, porque el aprenderá durante ese recorrido.

```

for epoch in tqdm(range(num_epochs)):
    model.train()
    optimizer.zero_grad()
    output = model(X_train_tensor)
    loss = criterion(output, y_train_tensor)
    loss.backward()
    optimizer.step()
    losses.append(loss.item()) # Guardar La pérdida en La lista
    if (epoch+1) % (num_epochs // 10) == 0:
        print(f'Epoch {epoch+1}/{num_epochs}, Loss: {loss.item():.4f}')

```

9. *Evaluar el modelo:* El objetivo de la evaluación del modelo es medir su desempeño en el conjunto de datos de prueba, es decir, en datos que no ha visto durante el entrenamiento. Esto permite verificar si el modelo ha aprendido a generalizar correctamente y si se puede hacer predicciones precisas sobre nuevos datos.

```

model.eval()
with torch.no_grad():
    y_predicted = model(X_test_tensor).numpy()

```

7.5. Simulación de modelo

7.5.1. Resultados de simulación

En esta sección, presentamos los resultados de la simulación utilizando un remuestreo de 6 segundos y 1 hora.

El objetivo es evaluar la precisión del modelo para predecir los valores futuros de frecuencia eléctrica. Los resultados se dividen en las siguientes secciones:

- ✓ Muestra de datos reales con cada intervalo de remuestreo.
- ✓ Los datos de frecuencia predicha después del entrenamiento del modelo.
- ✓ Las pérdidas después del entrenamiento.
- ✓ El pronóstico de valores de frecuencia a futuro.

Datos reales

El primer gráfico utiliza un remuestreo de 6 segundos, lo que permite observar las fluctuaciones y patrones de la frecuencia en detalle. Este intervalo corto proporciona una visión más detallada de los datos, destacando cambios rápidos y eventos transitorios que pueden ser cruciales para el entrenamiento del modelo.

El segundo gráfico emplea un remuestreo de 1 hora, lo que proporciona una perspectiva a mayor escala de las variaciones de frecuencia. Este intervalo de tiempo más amplio permite identificar tendencias y comportamientos a largo plazo, facilitando una comprensión más clara de la dinámica general de la frecuencia a lo largo del tiempo.

Sin embargo, el conjunto de datos es tan grande que resulta imposible procesarlo en su totalidad con los recursos disponibles. Procesar todos los datos consumiría una cantidad excesiva de memoria, lo que supera las capacidades de nuestro sistema, por esta razón, hemos decidido utilizar solo una muestra del 10% del total de los datos. Esta reducción nos permite manejar y analizar la información de manera eficiente sin sobreajustar la memoria del sistema, aunque representa una limitación en términos de la cantidad de datos que podemos cargar y procesar simultáneamente.

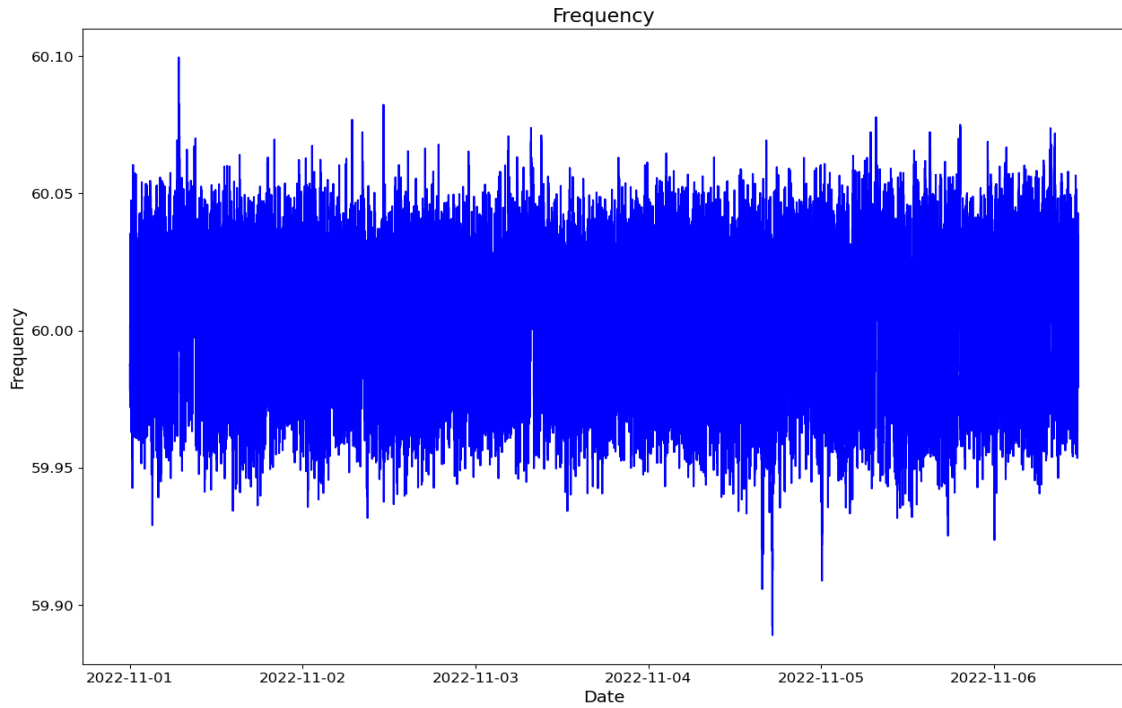


Figura 7.4. Datos reales de frecuencia eléctrica, con un resamdeo de seis segundos.

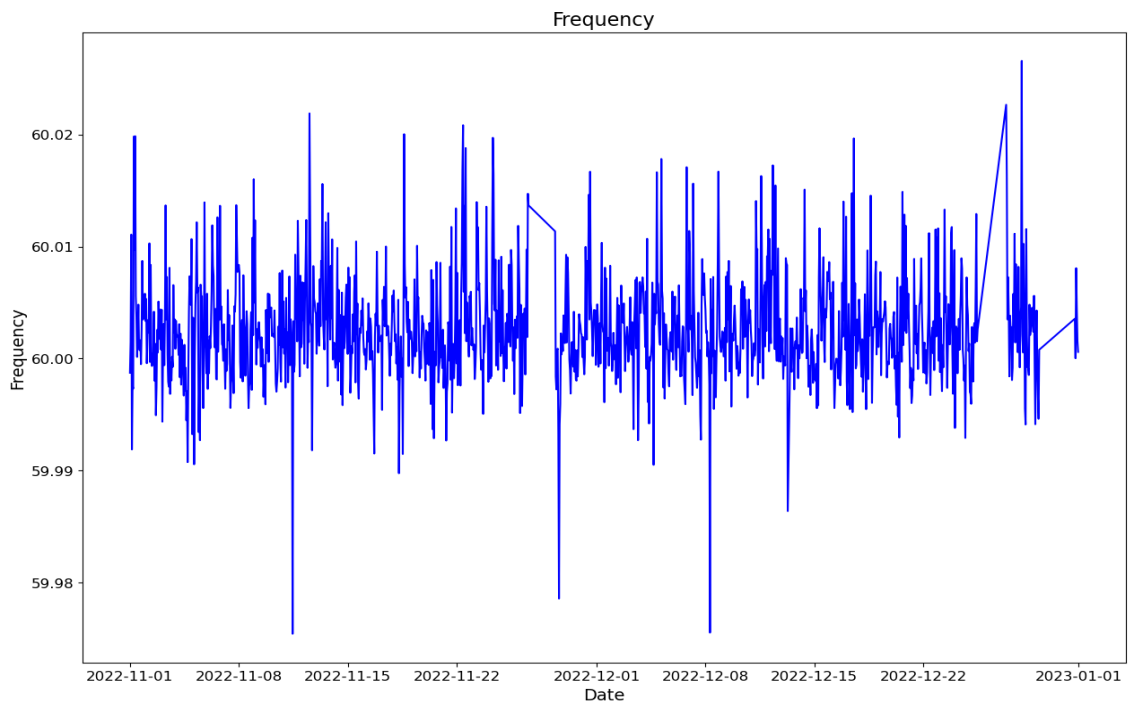


Figura 7.5. Datos reales de frecuencia eléctrica, con un resamdeo de una hora.

Es importante aclarar que, a medida que realizamos un remuestreo a intervalos de tiempo más grandes, hay una disminución considerable en el uso de recursos de memoria, lo cual es una ventaja significativa. Sin embargo, esta técnica tiene la desventaja de que reduce el nivel de detalle en el análisis. En lugar de observar las variaciones instante a instante, vemos una visión más aproximada y general de los datos, lo que puede ocultar patrones y fluctuaciones importantes que ocurren en intervalos más cortos.

Datos de frecuencia predicha

Podemos visualizar los datos de frecuencia predicha por el modelo con los datos reales de prueba, acá se comparan las frecuencias reales y las predichas.

En la primera grafica podemos observar una buena concordancia entre los valores reales con los valores predichos, obsérvese que la curva de color naranja es similar a la curva de color azul (mostrada al fondo de la gráfica).

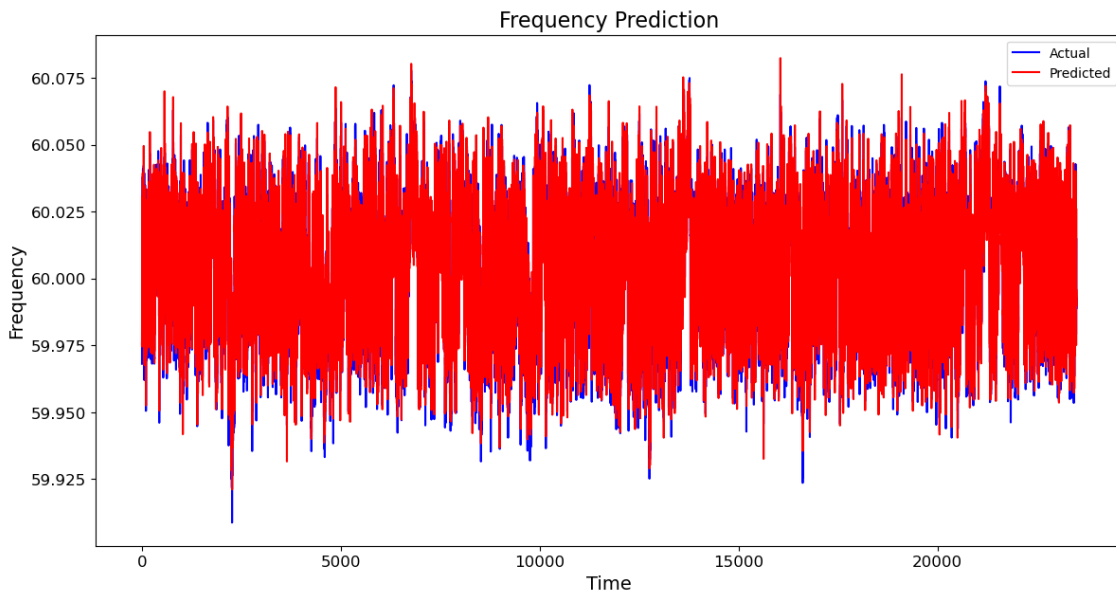


Figura 7.6. Datos reales con datos predichos de frecuencia eléctrica con un resamplio de seis segundos.

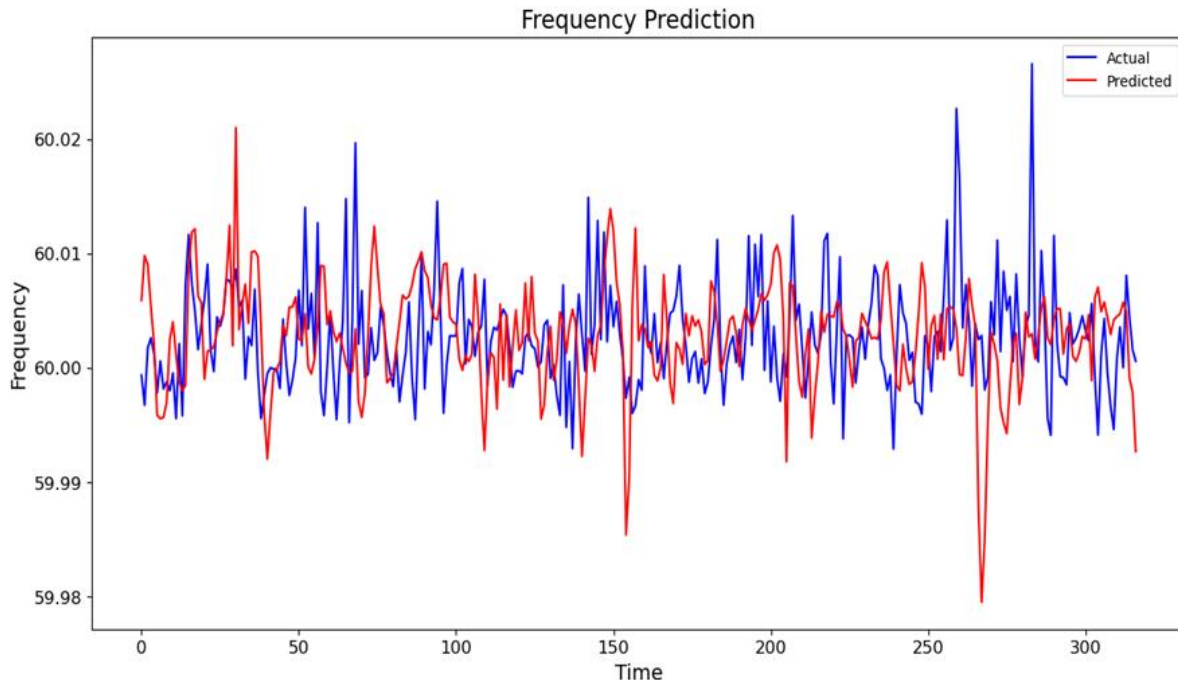


Figura 7.7. Datos reales con datos predichos de frecuencia eléctrica con resamdeo de una hora.

En la Figura 7.7 podemos observar una menor precisión con respecto a los datos de frecuencia, esto es debido a que estamos analizando datos con un remuestreo de 1 hora, el modelo podría no estar captando patrones importantes, debido a que, analizar datos a esos intervalos podría estar omitiendo sucesos importantes que pudieron haber ocurrido en lapsos de tiempo más pequeños.

Perdida a través de las épocas

En esta sección presentamos la pérdida del modelo y el valor de la pérdida del Error Cuadrático Medio (MSE) durante el proceso de entrenamiento. La pérdida es una medida de que también el modelo está ajustando los datos de entrenamiento. Como mencionamos anteriormente, el optimizador utilizado fue Adam y se estableció una tasa de aprendizaje de 0.001 para calcular el valor de pérdida optimizado. En el siguiente gráfico, se muestran dos curvas:

- **Curva de color azul:** Representa el valor de la perdida durante el entrenamiento.
- **Curva de color rojo:** Representa la pérdida del MSE a lo largo del entrenamiento.

Se observa que ambas curvas son casi similares, lo que indica que el modelo se está entrenando de manera correcta durante las épocas. Esta similitud sugiere que el modelo no solo está aprendiendo

adecuadamente, sino que también mantiene una consistencia en la minimización de la pérdida, lo cual es un buen indicativo de su desempeño.

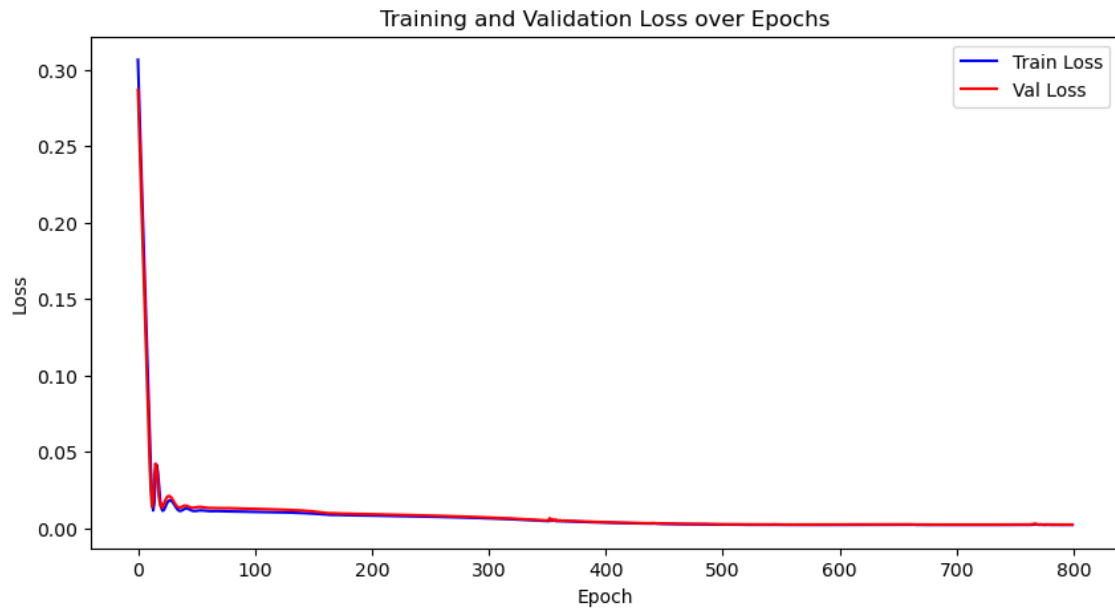


Figura 7.8. Pérdidas durante el entrenamiento con remuestro de seis segundos.

De igual manera, en la siguiente imagen se muestran la pérdida y el valor del MSE durante el entrenamiento con intervalos de 1 hora. Se puede observar que a grafica sigue una tendencia similar en ambas métricas, aunque hay una pequeña desviación al final de las curvas. Esta desviación no es preocupante, ya que ambas curvas han mostrado una tendencia reductora consistente a lo largo del entrenamiento.

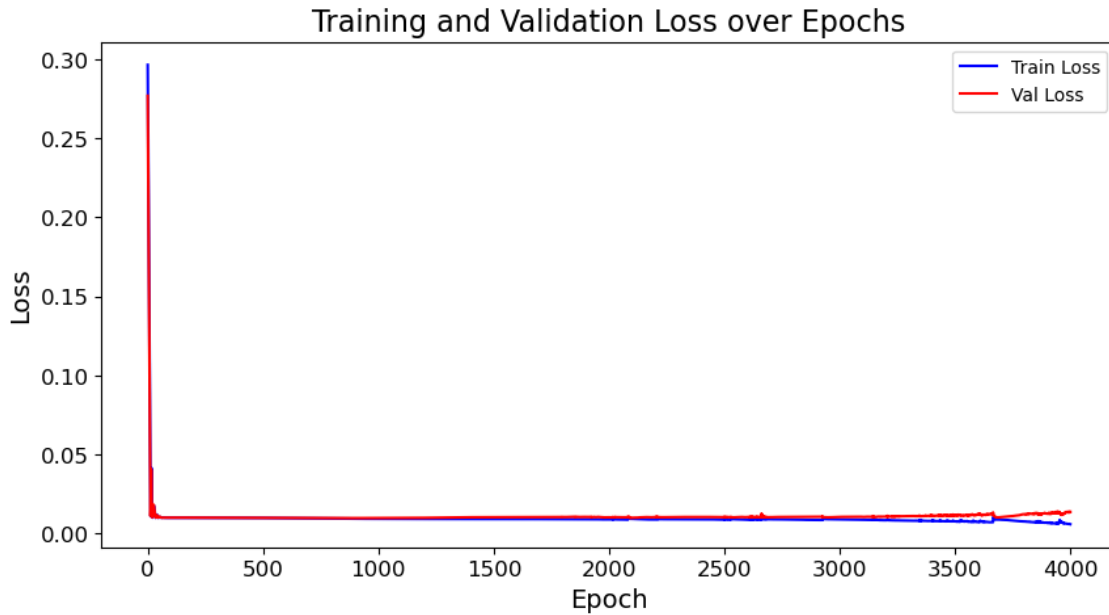


Figura 7.9. Pérdidas durante el entrenamiento con remuestro de una hora.

En las siguientes imágenes, presentaremos las épocas, el tiempo de ejecución, la pérdida y el valor del MSE durante el entrenamiento. En la primera imagen, se muestra el progreso del entrenamiento del modelo a lo largo de las épocas. Se puede notar que el tiempo de ejecución fue aproximadamente 5 horas. Esto se debe a que estábamos analizando datos en intervalos de 6 segundos, que como habíamos mencionado anteriormente, hace que el sistema consuma más recursos, ralentizando el proceso de entrenamiento.

Finalmente, se observan los valores de las pérdidas, que reflejan como el modelo está ajustando los datos de entrenamiento a lo largo del tiempo.

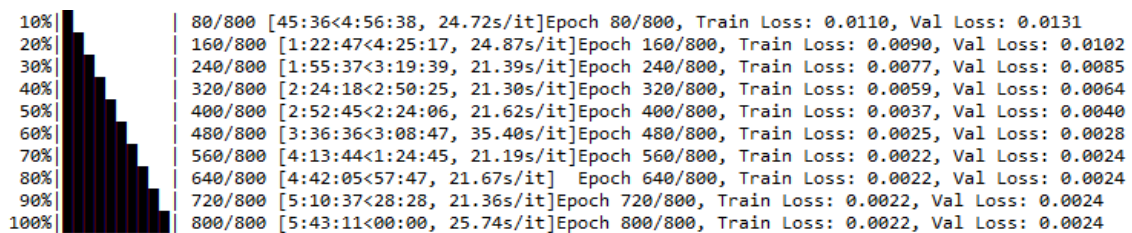


Figura 7.10. Representación del periodo de entrenamiento de datos (remuestro de seis segundos).

En esta otra imagen se muestra lo que anteriormente habíamos explicado, pero con la diferencia de que el entrenamiento se realizó con intervalos de una hora. Es importante notar que en este caso se necesitaron más épocas para que las pérdidas se ajustaran. Además, el tiempo de ejecución fue

de aproximadamente 47 minutos, lo que se debe a que estamos consumiendo menos recursos del sistema con este intervalo de remuestreo más amplio.

Observamos también que las pérdidas disminuyen a lo largo del entrenamiento, lo cual es un indicativo de que el modelo está aprendiendo y mejorando su capacidad para predecir correctamente los datos.

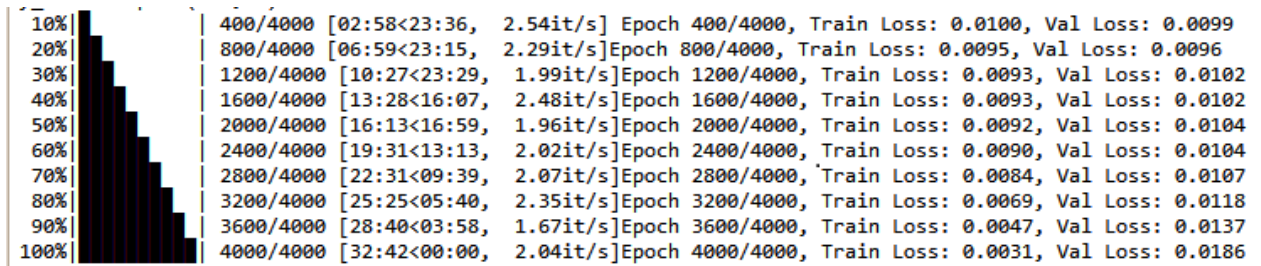


Figura 7.11. Representación del periodo de entrenamiento de datos (remuestreo de una hora).

Forecasting con datos de frecuencia.

En esta sección se mostrará el forecasting realizado con los datos de frecuencia. Presentaremos dos imágenes con los intervalos de tiempo que hemos venido trabajando anteriormente. Estos gráficos representan el valor de la frecuencia para un determinado tiempo. Durante el entrenamiento, el modelo aprendió patrones a partir de los datos, y ahora se utiliza para generar datos futuros. Tenemos tres curvas:

- **Curva de color azul:** Representa los datos originales que se utilizaron tanto para el entrenamiento como para la prueba del modelo. Esta curva proporciona una referencia de los valores reales de la frecuencia.
- **Curva de color naranja:** Muestra los datos predichos por el modelo durante el entrenamiento. Esta curva indica cómo el modelo se ajustó a los datos históricos durante el proceso de aprendizaje.
- **Curva de color verde:** Muestra el pronóstico de los datos que el modelo ha generado conforme al conjunto de datos. Esta curva representa las predicciones futuras basadas en el modelo entrenado.

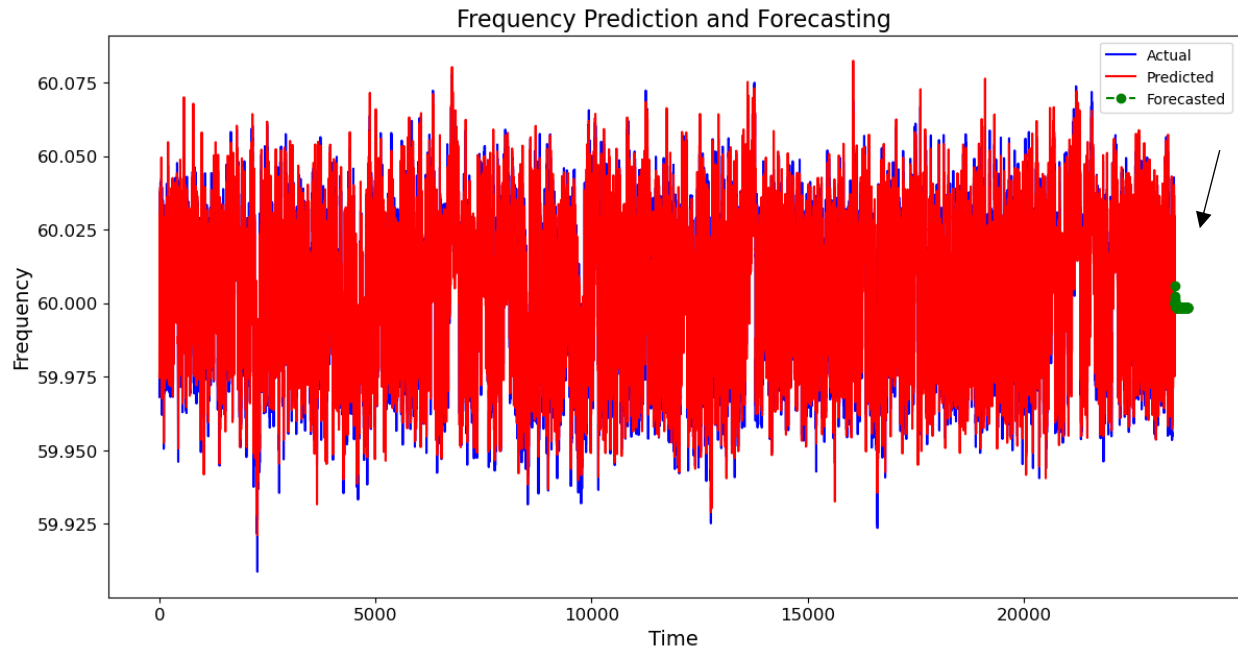


Figura 7.12. Forecasting de frecuencia eléctrica con resamdeo de seis segundos. Pronostico realizado para los próximos 120 segundos.

En la figura 8.9 se observa que la curva de color naranja sigue la misma tendencia que la curva de color azul, esto significa que el modelo está siguiendo el patrón de los datos reales. Esto es debido a que el modelo está estudiando los datos a periodos más reducidos, lo que le permite captar o recibir información más detallada sobre cómo cambia la frecuencia conforme el tiempo. Sin embargo, la limitante es que solo estamos usando el 10 por ciento de los datos y eso es una desventaja debido a que dificulta al modelo a que aprenda y en base a esos valores aprendidos, poder pronosticar valores a futuros.

En la segunda imagen podemos visualizar de mejor manera lo que sería el pronóstico a través del tiempo, debido a que es un intervalo muy amplio el de análisis. Así mismo, podemos notar que el modelo tiene un mayor repertorio de datos para aprender y así, pronosticar valores a futuros.

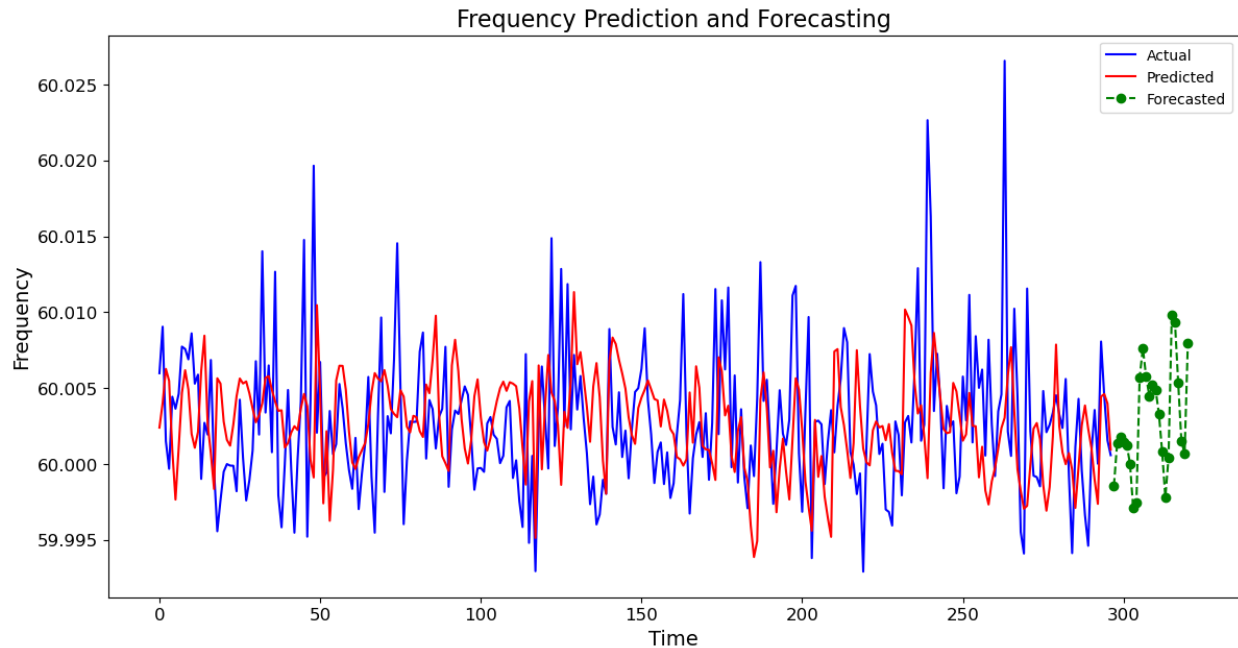


Figura 7.13. Forecasting de valores de frecuencia con resamdeo de una hora. Pronóstico realizado para las próximas 120 horas.

Para la imagen anterior, se realizó un pronóstico de las próximas 24 horas. En los gráficos se observa cómo el modelo predice los valores de frecuencia para este periodo de tiempo, utilizando los patrones aprendidos durante el entrenamiento.

En la parte izquierda de la gráfica, notamos que las líneas de color azul y roja no están exactamente juntas; de hecho, en algunos momentos se separan mucho. Esto indica que, aunque el modelo este capturando la tendencia general de los datos, no está prediciendo los valores exactos con gran precisión. Esto sugiere que el modelo puede estar teniendo dificultades para captar algunas fluctuaciones o variaciones rápidas de frecuencia.

Por otro lado, en la parte derecha de la gráfica, aparecen los puntos verdes que representan el pronóstico del modelo para el futuro. A medida que avanzamos hacia la derecha, vemos que los puntos verdes muestran una mayor variabilidad, con algunos puntos bastante alejados entre sí. Esto indica que el modelo predice una mayor inestabilidad o cambios más bruscos en la frecuencia a futuro.

Métricas de evaluación

Bien como hemos visto antes, los datos de frecuencia predicha y en el forecasting el modelo presenta una buena precisión con respecto al manejo de datos. Sin embargo, es necesario comprobar la precisión del modelo. Para ello, hemos hecho uso de métricas de evaluación como: El Error Medio Cuadrático y el Coeficiente de Determinación.

Tabla XLV. Métricas de evaluación de modelo elaborado.

Métrica	Descripción	Remuestreado 6s	Remuestreado 1h
Error medio cuadrático (MSE)	Mide la diferencia promedio al cuadrado entre los valores predichos y reales.	0.01	0.0001
Coeficiente de Determinación (R^2)	Indica que tan bien se ajusta el modelo a los datos reales.	0.8375	-0.3558

- 1) **Datos remuestreados a 6 segundos:** El modelo presentó un Error Cuadrático Medio (MSE) de 0.01, un valor muy cercano a cero. Esto indica que las predicciones de la frecuencia están muy próximas a los valores reales, lo que sugiere un buen rendimiento del modelo en este contexto. Además, el Coeficiente de Determinación (R^2) fue de 0.8375, un valor cercano a 1, lo que refuerza la idea de que el modelo tiene una buena precisión en la predicción. Un R^2 cercano a 1 implica que el modelo explica bien la variabilidad de los datos, mientras que valores alejados de 1 indican un desempeño deficiente.
- 2) **Datos remuestreados a 1 hora:** En este caso, el MSE fue de 0.0001, un valor prácticamente nulo, lo que a primera vista podría sugerir una predicción casi perfecta. Sin embargo, el Coeficiente de Determinación (R^2) fue de -0.3558, un valor significativamente alejado de 1. Esto indica que el modelo no está generalizando correctamente, sino que podría estar memorizando los datos, lo que se traduce en un rendimiento pobre. En este intervalo de remuestreo, el modelo es totalmente impreciso, ya que un R^2 negativo sugiere que el modelo es peor que simplemente usar la media de los datos para hacer predicciones.

7.5.2. Observaciones Clave.

1. Comparación de Resamples:

Resamplero de 6 segundos: La curva generada con este resamplero muestra detalles finos y fluctuaciones rápidas en los datos de frecuencia, capturando variaciones a corto plazo.

Resamplero de 1 hora: Esta curva, al tener un intervalo de tiempo más amplio, proporciona una perspectiva a mayor escala, destacando tendencias y patrones a largo plazo.

2. Tendencias Razonables:

Remuestreo de 6 segundos: Las predicciones siguen tendencias razonables y coherentes con los datos históricos, lo que sugiere una buena precisión en este intervalo.

Remuestreo de 1 hora: A pesar de que la curva puede parecer razonable a simple vista, las métricas sugieren que el modelo no está capturando correctamente las relaciones subyacentes en los datos, lo que podría indicar una falta de precisión en este contexto.

3. Evaluación del Modelo:

La observación de las gráficas de pérdida durante el entrenamiento revela que, en general, el modelo está aprendiendo correctamente, con una reducción consistente del Error Cuadrático Medio (MSE) a lo largo del tiempo para el remuestreo de 6 segundos. Sin embargo, para el remuestreo de 1 hora, un MSE extremadamente bajo acompañado de un coeficiente de determinación negativo indica que el modelo podría estar sobreajustando o memorizando los datos, en lugar de aprender patrones generales.

4. Validación de la Precisión:

Aunque las predicciones con el remuestreo de 6 segundos parecen ser razonables y reflejan bien los patrones de los datos originales, las predicciones con el remuestreo de 1 hora no demuestran la misma coherencia. La capacidad del modelo para mantener la precisión en diferentes intervalos de remuestreo está en duda, lo que sugiere que el modelo puede ser más fiable con remuestreos más cortos, donde las fluctuaciones y detalles son más evidentes y se pueden capturar mejor.

CONCLUSIONES

En base a todo lo visto en la investigación, se tiene que:

1. Se logró configurar exitosamente una entrada de flujo de datos y su almacenamiento para las mediciones de frecuencia provenientes de la unidad FDR-UES en un ordenador de la Escuela de Ingeniería Eléctrica. Se realizaron mediciones durante todo el mes de marzo, pero el equipo #1703 presentó fallas técnicas fuera de nuestro control, lo que interrumpió la recolección de datos. Sin embargo, se ha elaborado un manual detallado con instrucciones para configurar el sistema, listo para cuando se reemplace la unidad defectuosa.
2. Se desarrolló un criterio para identificar eventos de perturbación en la frecuencia eléctrica, mediante la teoría de la detección de puntos de cambio para identificar estos eventos, registrando información clave como los valores mínimos, máximos y la duración de cada evento. Se hicieron tres clasificaciones de eventos:
 - a. *Eventos menores/leves*: Desviaciones de frecuencia superiores a 0.05 Hz. Se ha optado por el valor de 0.05 Hz debido a que instituciones como el NERC o IEEE han hecho mención que desviaciones inferiores no son relevantes, y que además en el RMER se ha dicho que la frecuencia eléctrica puede estar dentro de este umbral.
 - b. *Eventos moderados*: Desviaciones de frecuencia superiores a 0.12 Hz. Se ha optado por este valor ya que es el umbral en que la Unidad de Transacciones considera que se está dentro de una condición normal, y a su vez es un valor recomendado por la Comisión Electrotécnica Internacional (IEC).
 - c. *Eventos críticos*: Desviaciones de frecuencia superiores a 0.60 Hz. Se ha optado por ese valor, dado que en muchos sistemas se considera una variación del 1% de la frecuencia como lo tolerable, además que el rango de operación de emergencia establecido por la UT cubre desde 59.88 a 59.40 Hz.
3. Utilizando los datos de frecuencia eléctrica recopilados desde septiembre 2022 hasta junio 2023 y los proporcionados por la Unidad de Transacciones, se observó que la frecuencia eléctrica se mantiene cercana a los 60 Hz, con una desviación estándar de 0.025 Hz. Se

pudo observar que la distribución existente de la frecuencia no es gaussiana, pero sí puede ser aproximada a una. Se registraron los eventos de perturbación de frecuencia mediante el algoritmo automatizado, destacando que hubo más eventos de sobrefrecuencia, aunque de magnitud menor, mientras los eventos de infrafrecuencia podían llegar a mayor magnitud.

4. Se observó que la frecuencia eléctrica muestra variaciones en su comportamiento a lo largo del tiempo, siendo esto reflejado en la desviación estándar y la cantidad de eventos menores. La desviación estándar y la cantidad de eventos menores disminuyó entre junio y septiembre, aunque para las otras clasificaciones no se encontró una tendencia clara. También puede destacarse que, para los eventos menores y moderados, existe una mayor concentración, bastante notable, de estos durante los domingos. A su vez, y esto aplica para las tres clasificaciones, se observó que existía una gran presencia de estos durante horario laboral, principalmente en las tardes.
5. Analizando los reportes públicos del Ente Operador Regional (EOR) sobre pérdidas de generación/carga, para el periodo 2022-2023, se encontró que muchas perturbaciones en la frecuencia eléctrica se debieron a pérdidas ocurridas en otros países del Sistema Eléctrico Regional (SER), especialmente en Honduras y Panamá. También se observó que la interconexión con el Sistema Eléctrico Mexicano (SEM) tiene una gran influencia en la estabilidad de la frecuencia eléctrica; la apertura de esta interconexión puede provocar caídas significativas en la frecuencia, lo que pone en riesgo de activar los Esquemas de Desconexión de Carga por Baja Frecuencia en los países del SER. Finalmente, los reportes del EOR indican que muchos eventos de pérdida de generación o carga en El Salvador se deben a problemas de transmisión, ya sea en líneas o subestaciones eléctricas, seguidos por problemas en la generación hidroeléctrica y térmica. Todo lo dicho puede ser considerado como un estudio preliminar, dado que los datos utilizados son bastante limitados.
6. Se implementó un modelo predictivo basado en redes neuronales LSTM para la predicción de la frecuencia eléctrica. Aunque el modelo puede predecir tendencias generales, no es la misma situación para eventos específicos, de magnitudes moderadas o críticas. En base al análisis realizado en el Capítulo 6, se puede entender que sería prácticamente imposible

esta tarea, dado que muchas causas de estos eventos son factores no cuantificables, a menudo externos a la red eléctrica salvadoreña. A pesar de no alcanzar una precisión óptima, el modelo ha demostrado una capacidad razonable para capturar tendencias generales en los datos.

RECOMENDACIONES

Con base en la experiencia obtenida en la investigación y los resultados obtenidos, se recomienda lo siguiente:

1. Se recomienda instalar el equipo medidor (FDR) en un lugar seguro, ya que en la ubicación anterior se sospecha que podría haber problemas eléctricos que afectaron el medidor.
2. Aunque el medidor es muy preciso, ocasionalmente puede proporcionar datos anómalos de frecuencia que podrían interpretarse como errores de medición. Sería ideal contar con al menos dos medidores para garantizar la redundancia de los datos y la fiabilidad de los resultados. Los medidores podrían instalarse en diferentes áreas de la Facultad de Ingeniería y Arquitectura o de la Universidad de El Salvador. A su vez, el proyecto FNET/GridEye podría ser muy beneficioso si se pudieran instalar más medidores en otros países de América Central, permitiendo obtener datos para investigaciones más completas.
3. El criterio de detección del inicio y finalización de eventos de perturbación de frecuencia puede mejorarse significativamente. Se debería evaluar la aplicación de técnicas de Machine Learning para aumentar la precisión en esta detección. Además, aunque los criterios de clasificación actuales funcionan para este análisis, podrían necesitar ajustes según otros contextos o necesidades específicas.
4. Sería beneficioso realizar análisis más profundos que permitan, además de clasificar los eventos, identificar el tipo de evento, como oscilatorio, pérdida de generación, etc. Existen investigaciones que abordan este tema utilizando mediciones sincrofasoriales. En esta investigación, debido a que se utilizaron datos del SCADA-UT, no fue posible aplicar estos métodos, limitándonos al análisis estadístico y de tendencias.
5. A medida que se recopilen más datos de las unidades FDR-UES y del proyecto FNET/GridEye, se podrán identificar tendencias en el comportamiento de la frecuencia en períodos de tiempo más largos, como anuales o quinquenales. También se podría considerar observar el comportamiento de la frecuencia con relación a otros factores.

6. La interconexión entre el Sistema Eléctrico Regional (SER) y el Sistema Eléctrico Mexicano (SEM) merece un estudio detallado, dada su importancia para la estabilidad de la frecuencia eléctrica en Centroamérica. Durante la investigación se identificó la relevancia de esta interconexión para el SER, pero aún falta explorar qué sucede con la frecuencia eléctrica en el lado mexicano, utilizando registros históricos de frecuencia. Los reportes públicos del EOR pueden proporcionar una visión preliminar sobre los momentos en que la interconexión se apertura, aunque lo recomendable sería solicitar datos más detallados a las correspondientes entidades, tanto en México como en Centroamérica (si es que existen). Otros estudios se pueden enfocar en la identificación de riesgos existentes en el SER que puedan causar la apertura de la interconexión con el SEM.
7. La implementación y entrenamiento de modelos LSTM requieren recursos computacionales significativos. Es recomendable invertir en infraestructura de hardware más avanzada, lo que permitirá realizar experimentos más amplios y obtener mejores resultados en cuanto a la precisión del modelo. Una de las principales limitaciones de este estudio fue la cantidad y calidad de los datos disponibles. Se recomienda ampliar la recolección de datos, abarcando periodos más largos y diferentes condiciones operativas. Datos adicionales pueden proporcionar una mejor base para entrenar el modelo, mejorando su capacidad predictiva y en su precisión. Además, una modificación al enfoque del uso de modelos predictivos, para la frecuencia eléctrica, podría ser útil, enfocándose en la realización de predicciones dada ciertas condiciones.

REFERENCIAS

- [1] A. G. Phadke y J. S. Thorp, “History and applications of phasor measurements”, en *2006 IEEE PES Power Systems Conference and Exposition, PSCE 2006 - Proceedings*, 2006. doi: 10.1109/PSCE.2006.296328.
- [2] Z. Zhong *et al.*, “Power system frequency monitoring network (FNET) implementation”, *IEEE Transactions on Power Systems*, vol. 20, núm. 4, 2005, doi: 10.1109/TPWRS.2005.857386.
- [3] S. D. Anton, D. Fraunholz, C. Lipps, F. Pohl, M. Zimmermann, y H. D. Schotten, “Two decades of SCADA exploitation: A brief history”, en *2017 IEEE Conference on Applications, Information and Network Security, AINS 2017*, 2017. doi: 10.1109/AINS.2017.8270432.
- [4] F. Milano y Á. O. Manjavacas, *Frequency Variations in Power Systems*. 2020. doi: 10.1002/9781119551881.
- [5] IEEE, “IEEE - Mission & Vision”. Consultado: el 12 de marzo de 2023. [En línea]. Disponible en: <https://www.ieee.org/about/vision-mission.html>
- [6] IEC, “About us | IEC”. Consultado: el 13 de marzo de 2023. [En línea]. Disponible en: <https://www.iec.ch/about-us>
- [7] SIGET, “Marco Institucional - Superintendencia General de Electricidad y Telecomunicaciones”. Consultado: el 12 de abril de 2023. [En línea]. Disponible en: <https://www.siget.gob.sv/marco-institucional/>
- [8] Unidad de Transacciones, “Marco regulatorio”. Consultado: el 12 de abril de 2023. [En línea]. Disponible en: <https://portal.ut.com.sv/web/guest/marco-regulatorio>
- [9] Comisión Regional de Interconexión Eléctrica (CRIE), “¿Qué es el MER?”, <https://crie.org.gt/>, 2023, Consultado: el 20 de julio de 2024. [En línea]. Disponible en: <https://crie.org.gt/mer/que-es-el-mer/>

- [10] Comisión Regional de Interconexión Eléctrica (CRIE), “Mercado Eléctrico Regional (MER)”, <https://crie.org.gt/>. Consultado: el 20 de julio de 2024. [En línea]. Disponible en: <https://crie.org.gt/mer/>
- [11] EOR, “Reglamento de Organización y Funciones”, Managua, sep. 2021.
- [12] IEEE, “IEEE Recommended Practice for Monitoring Electric Power Quality”, 2019.
- [13] IEC, “Electromagnetic compatibility (EMC) – Part 2-5: Environment – Description and classification of electromagnetic environments”, *IEC TR 61000-2-5:2017*, ene. 2017.
- [14] Unidad de Transacciones (UT), “ANEXO – NORMAS DE CALIDAD Y SEGURIDAD OPERATIVAS”, *REGLAMENTO DE OPERACIÓN DEL SISTEMA DE TRANSMISIÓN Y DEL MERCADO MAYORISTA - ROBCP*, feb. 2022, Consultado: el 4 de marzo de 2024. [En línea]. Disponible en: <https://www.ut.com.sv/documents/10100/279097/ROBCP-Anexos.pdf/f5ec6c02-8426-4cb6-5a61-e7cba80ff34b?t=1679083382997>
- [15] Comisión Regional de Interconexión Eléctrica (CRIE), “REGLAMENTO DEL MERCADO ELÉCTRICO REGIONAL -RMER-”, <https://crie.org.gt/>, 2020, Consultado: el 10 de abril de 2024. [En línea]. Disponible en: <https://crie.org.gt/wp-content/uploads/2020/07/RMER-17062020.pdf>
- [16] K. Dehghanpour y S. Afsharnia, “Electrical demand side contribution to frequency control in power systems: a review on technical aspects”, *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 1267–1276, ene. 2015, doi: 10.1016/J.RSER.2014.09.015.
- [17] “EPRI Power System Dynamics Tutorial”, Palo Alto, California, jul. 2009. Consultado: el 9 de mayo de 2024. [En línea]. Disponible en: <https://www.epri.com/research/products/000000000001016042>
- [18] Unidad de Transacciones (UT), “Boletín Estadístico Anual 2022”, Nuevo Cuscatlán, La Libertad, El Salvador, ene. 2023. Consultado: el 13 de junio de 2024. [En línea]. Disponible en: <https://www.ut.com.sv/documents/10100/279981/Estadistico+Anual+Enero+a+Diciembre+2022.pdf/cd649900-e52e-7fac-fdaf-1fd1c694b6aa?t=1675094724615>

- [19] Unidad de Transacciones (UT), “Boletín Estadístico Anual 2023”, Nuevo Cuscatlán, La Libertad, El Salvador, ene. 2024. Consultado: el 13 de junio de 2024. [En línea]. Disponible en: <https://www.ut.com.sv/documents/10100/279981/Informe+Estadistico+Anual+Enero-Diciembre+2023.pdf/6b6500fb-56bb-e2b9-bebf-e32cba485295?t=1706657853923>
- [20] A. von Meier, *Electric Power Systems*. Wiley, 2006. doi: 10.1002/0470036427.
- [21] “Definition and Classification of Power System Stability IEEE/CIGRE Joint Task Force on Stability Terms and Definitions”, *IEEE Transactions on Power Systems*, vol. 19, núm. 3, 2004, doi: 10.1109/tpwrs.2004.825981.
- [22] N. Hatziargyriou *et al.*, “Definition and Classification of Power System Stability - Revisited & Extended”, *IEEE Transactions on Power Systems*, vol. 36, núm. 4, 2021, doi: 10.1109/TPWRS.2020.3041774.
- [23] “IEEE Guide for Abnormal Frequency Protection for Power Generating Plants”, *IEEE Std C37.106-2003 (Revision of ANSI/IEEE C37.106-1987)*, pp. 1–40, 2004, doi: 10.1109/IEEESTD.2004.94434.
- [24] “Glossary of Terms Used in NERC Reliability Standards”, Palo Alto, California, mar. 2024. Consultado: el 12 de junio de 2024. [En línea]. Disponible en: https://www.nerc.com/pa/Stand/Glossary%20of%20Terms/Glossary_of_Terms.pdf
- [25] IEEE Power & Energy Society, “Dynamic Models for Turbine-Governors in Power System Studies”, *Technical Report PES-TR1*, 2013.
- [26] P. Denholm, T. Mai, R. W. Kenyon, B. Kroposki, y M. O. Malley, “Inertia and the Power Grid: A Guide Without the Spin”, *National Renewable Energy Laboratory*, núm. May, 2020.
- [27] P. Tumino, “Frequency Control in a Power System”, EE Power. Consultado: el 1 de julio de 2024. [En línea]. Disponible en: <https://eepower.com/technical-articles/frequency-control-in-a-power-system/#>
- [28] Schneider Electric, “Protección contra infrafrecuencia/sobrefrecuencia (código ANSI 81)”, Unidad de control MasterPact MTZ MicroLogic X.

- [29] Unidad de Transacciones (UT), “Informe de revisión de los esquemas de desconexión de carga: Por baja frecuencia y por bajo voltaje - El Salvador”, Nuevo Cuscatlán, La Libertad, El Salvador, dic. 2023.
- [30] Ente Operador Regional, “INFORME DE REGULACIÓN DEL MERCADO ELÉCTRICO REGIONAL (Ordinario) - IRMER-O-02-2019”, <https://www.enteoperador.org/>, jul. 2019.
- [31] “Frequency Response Standard Background Document”, Atlanta, Georgia, dic. 2012.
- [32] “Balancing and Frequency Control Reference Document”, Atlanta, Georgia, may 2011.
- [33] P. Etingov, D. Kosterev, y T. Dai, “Frequency Response Analysis Tool”, Richland, Washington, dic. 2014.
- [34] “2022 Frequency Response Annual Analysis”, Atlanta, Georgia, nov. 2022.
- [35] R. Cummings, “NERC Reliability Initiatives and NERC Reliability Initiatives and Smart Grid”, Calgary, Alberta, Canadá, jul. 2009. Consultado: el 12 de junio de 2024. [En línea]. Disponible en: https://www.smartgrid.gov/document/nerc_reliability_initiatives_and_smart_grid
- [36] Ente Operador Regional (EOR), “REPORTE DE EVENTO N° 91-05-23”, San Salvador, El Salvador, may 2023. Consultado: el 27 de febrero de 2024. [En línea]. Disponible en: https://www.enteoperador.org/?red_fm_connect%3Dtrue%26front%3Duser%26fid%3D175%26defaults%3D0%26access_all%3D0%26cmd%3Dfile%26target%3D11_MjAyMy8wNS1tYXlvLzlxLTA1LUV2ZW50b19TQUwwN01BWU8yMDIzXzEwNTQucGRm%26_t%3D1683578203
- [37] University of Tennessee y Oak Ridge National Laboratory, “About FNET/GridEye”. Consultado: el 7 de marzo de 2024. [En línea]. Disponible en: <https://fnetpublic.utk.edu/index.html>
- [38] Power Information Technology (FNET) Lab, “Frequency Disturbance Recorder (FDR): Installation Guide”, Knoxville, ago. 2012.
- [39] O. R. N. L. University of Tennessee, “FNET/GridEye Frequency Display”. Consultado: el 7 de marzo de 2024. [En línea]. Disponible en: <https://fnetpublic.utk.edu/tabledisplay.html>

- [40] J. L. Zárate Gómez y M. R. Arrieta Paternina, “Documentación: Instrucciones para acceder a los PMUs e instalar ZeroTier.”, *Desarrollo de Herramientas en Tiempo-Real para el Análisis del Desempeño Dinámico del SEN mediante su Implementación en Sistemas de Monitoreo de Área Amplia (WAMS)*, jun. 2021.
- [41] J. L. Zárate Gómez y M. R. Arrieta Paternina, “Documentación: openPDC y PMU connection tester”, *Desarrollo de Herramientas en Tiempo-Real para el Análisis del Desempeño Dinámico del SEN mediante su Implementación en Sistemas de Monitoreo de Área Amplia (WAMS)*, jun. 2021.
- [42] J. Azor, “La Transformada Wavelet”, *Revista De La Universidad De Mendoza*, p. 4, 2015, [En línea]. Disponible en: <http://dspace.ups.edu.ec/handle/123456789/14201>
- [43] D. Frau *et al.*, “Reducción del ruido en señales electrocardiográficas mediante la transformada Wavelet”, may 2000.
- [44] D. Tinoco Varela, F. Gudiño Peñaloza, J. L. Garza Rivera, A. Espinoza Godínez, R. D. Cruz Morales, y C. J. Villaseñor Rodelas, “Transformada wavelet”, Acervo para el mejoramiento del aprendizaje de alumnos de ingeniería, en inteligencia artificial. Consultado: el 10 de mayo de 2024. [En línea]. Disponible en: https://virtual.cuautitlan.unam.mx/intar/?page_id=43
- [45] Zach Bobbitt, “MAE vs. RMSE: Which Metric Should You Use?”, Statology. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: <https://www.statology.org/mae-vs-rmse/>
- [46] Ajitesh Kumar, “MSE vs RMSE vs MAE vs MAPE vs R-Squared: When to Use?”, Analytics Yogi. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: <https://vitalflux.com/mse-vs-rmse-vs-mae-vs-mape-vs-r-squared-when-to-use/>
- [47] R. E. Walpole, R. H. Myers, S. L. Myers, y K. Ye, *Probabilidad y estadística para ingeniería y ciencias*, 9a ed. Pearson Educación de México, SA de CV, 2012.
- [48] W3Schools, “SciPy Statistical Significance Tests”, SciPy Tutorial. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: https://www.w3schools.com/python/scipy/scipy_statistical_significance_tests.php

- [49] DATAtab Team, “Prueba t para muestras independientes”, DATAtab: Tutoriales. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: <https://datatab.es/tutorial/unpaired-t-test>
- [50] Zach Bobbitt, “Welch’s t-test: When to Use it + Examples”, Statology. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: <https://www.statology.org/welchs-t-test/>
- [51] Jim Frost, “Understanding Significance Levels in Statistics”, Statistics By Jim. Consultado: el 4 de mayo de 2024. [En línea]. Disponible en: <https://statisticsbyjim.com/hypothesis-testing/significance-levels/>
- [52] “Asimetría y kurtosis”, <https://www.odiolaestadistica.com/>. Consultado: el 16 de abril de 2024. [En línea]. Disponible en: <https://www.odiolaestadistica.com/estadistica-python/asimetria/>
- [53] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, y A. Keshri, “Descriptive statistics and normality tests for statistical data”, *Ann Card Anaesth*, vol. 22, núm. 1, 2019, doi: 10.4103/aca.ACA_157_18.
- [54] IBM, “Summarize Statistics”, SPSS Statistics. Consultado: el 19 de abril de 2024. [En línea]. Disponible en: <https://www.ibm.com/docs/en/spss-statistics/saas?topic=summarize-statistics>
- [55] H.-Y. Kim, “Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis”, *Restor Dent Endod*, vol. 38, núm. 1, 2013, doi: 10.5395/rde.2013.38.1.52.
- [56] P. Shaw y M. Kumar Jena, “A Novel Event Detection and Classification Scheme Using Wide-Area Frequency Measurements”, *IEEE Trans Smart Grid*, vol. 12, núm. 3, 2021, doi: 10.1109/TSG.2020.3039274.
- [57] J. Ilic, D. Wilson, M. Prica, S. Rabiei, y J. Goellner, “Frequency Instability Problems in North American Interconnections”, *National Energy Technology Laboratory*, 2011.
- [58] B. Kirby *et al.*, “Frequency Control Concerns In The North American Electric Power System”, may 2003.

- [59] D. Hammerstrom *et al.*, “Pacific Northwest GridWise™ Testbed Demonstration Projects; Part II. Grid Friendly™ Appliance Project,” May 2007, doi: 10.2172/926122.
- [60] G. Gajjar y S. A. Soman, “Auto detection of power system events using wide area frequency measurements”, en *2014 18th National Power Systems Conference, NPSC 2014*, 2015. doi: 10.1109/NPSC.2014.7103879.
- [61] “IEEE/IEC International Standard - Measuring relays and protection equipment - Part 118-1: Synchrophasor for power systems - Measurements”, *IEC/IEEE 60255-118-1:2018*, pp. 1–78, 2018, doi: 10.1109/IEEESTD.2018.8577045.
- [62] Parra Francisco, “Estadística y Machine Learning con R”, *Bookdown*, 2019.
- [63] N. Peach, M. Basseville, y I. V. Nikiforov, “Detection of Abrupt Changes: Theory and Applications.”, *J R Stat Soc Ser A Stat Soc*, vol. 158, núm. 1, 1995, doi: 10.2307/2983416.
- [64] P. F. R. Killick y I. A. Eckley, “Optimal Detection of Changepoints With a Linear Computational Cost”, *J Am Stat Assoc*, vol. 107, núm. 500, pp. 1590–1598, 2012, doi: 10.1080/01621459.2012.737745.
- [65] “Análisis de patrones espacio-temporales: Cómo funciona la detección de puntos de cambio.”, ArcGIS Pro 3.2. Consultado: el 11 de mayo de 2024. [En línea]. Disponible en: <https://pro.arcgis.com/es/pro-app/latest/tool-reference/space-time-pattern-mining/how-change-point-detection-works.htm>
- [66] P. Schober y L. A. Schwarte, “Correlation coefficients: Appropriate use and interpretation”, *Anesth Analg*, vol. 126, núm. 5, pp. 1763–1768, may 2018, doi: 10.1213/ANE.0000000000002864.
- [67] J. D. Hernández Lalinde *et al.*, “Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones On the proper use of the Pearson correlation coefficient: definitions, properties and assumptions”, *Archivos Venezolanos de Farmacología y Terapéutica*, vol. 37, núm. 5, 2018.
- [68] Ente Operador Regional, “Historia”, ENTE OPERADOR REGIONAL - Del Sistema Eléctrico Regional.

- [69] Ente Operador Regional, “AUMENTO DEL LÍMITE DE LAS TRANSFERENCIA DE POTENCIA DEL SISTEMA ELÉCTRICO DE MÉXICO AL SISTEMA ELÉCTRICO REGIONAL (SER)”, San Salvador, El Salvador, 2019. Consultado: el 30 de junio de 2024. [En línea]. Disponible en: https://www.enteoperador.org/wp-content/uploads/2019/06/COMUNICADO-Aumento-de-las-Max-Cap-Transf-MEX-GUA_VF.pdf
- [70] Instituto Nacional De Electrificación, “¿Cómo surgió la interconexión eléctrica entre Guatemala y México?”, <https://www.inde.gob.gt/>, oct. 2020, Consultado: el 10 de julio de 2024. [En línea]. Disponible en: <https://www.inde.gob.gt/como-surgio-la-interconexion-electrica-entre-guatemala-y-mexico/>
- [71] Gerencia de Planificación y Operación del EOR, “Límite de Transferencia de Potencia entre el Sistema Eléctrico de México y el Sistema Eléctrico Regional”, San Salvador, El Salvador, jun. 2019. Consultado: el 5 de julio de 2024. [En línea]. Disponible en: https://www.enteoperador.org/wp-content/uploads/2019/06/Estudio-Seguridad-Operativa-Actualizaci%C3%B3n_E_T_M%C3%A9xico-SER_%C3%A9poca_h%C3%BAmeda_2019.pdf
- [72] Alejandro Nieto Jeux, “Algoritmos de Aprendizaje Automático. Un Estudio de su Difusión y Utilización”, Trabajo de Fin de Grado, Universidad Politécnica de Madrid, Madrid, 2021. [En línea]. Disponible en: https://oa.upm.es/68484/1/TFG_ALEJANDRO_NIETO_JEUX.pdf
- [73] “¿Qué es una red neuronal? | IBM”. Consultado: el 28 de julio de 2024. [En línea]. Disponible en: <https://www.ibm.com/es-es/topics/neural-networks>
- [74] “Fundamentals of Neural Networks & Deep Learning | AnalytixLabs”. Consultado: el 28 de julio de 2024. [En línea]. Disponible en: <https://www.analytixlabs.co.in/blog/fundamentals-of-neural-networks/>
- [75] “(162) ¿Qué es una Red Neuronal? Parte 1: La Neurona | DotCSV - YouTube”. Consultado: el 28 de julio de 2024. [En línea]. Disponible en: <https://www.youtube.com/watch?v=MRIv2IwFTPg>

- [76] “Components of Neural Network | Neural Network Layers & Neurons”. Consultado: el 28 de julio de 2024. [En línea]. Disponible en: <https://www.analyticsvidhya.com/blog/2021/03/basics-of-neural-network/>
- [77] “Redes neuronales. Programa de Visión... | by Bootcamp AI | Medium”. Consultado: el 29 de julio de 2024. [En línea]. Disponible en: <https://bootcampai.medium.com/redes-neuronales-13349dd1a5bb>
- [78] C. Alberto Ruiz Marta Susana Basualdo Autor y D. Jorge Matich, “Cátedra: Informática Aplicada a la Ingeniería de Procesos-Orientación I Redes Neuronales: Conceptos Básicos y Aplicaciones”.
- [79] Aditti Mittal, “Understanding RNN and LSTM”, Medium. [En línea]. Disponible en: <https://aditi.mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e>
- [80] Niklas Donges, “A Complete Guide to Recurrent Neural Networks (RNNs)”, builtin. [En línea]. Disponible en: <https://builtin.com/data-science/recurrent-neural-networks-and-lstm>
- [81] “Recurrent Neural Networks and LSTM: Overview and Uses | Turing”. Consultado: el 28 de julio de 2024. [En línea]. Disponible en: <https://www.turing.com/kb/recurrent-neural-networks-and-lstm>
- [82] Antonio. Gulli y Sujit. Pal, *Deep Learning with Keras: Implement various deep-learning algorithms in Keras and see how deep-learning can be used in games*. Packt Publishing, 2017.
- [83] Miguel Sotaquirá, “¿Qué son las Redes LSTM?”, codificandobits. Consultado: el 29 de julio de 2024. [En línea]. Disponible en: <https://www.codificandobits.com/blog/redes-lstm/>