



Universidad Nacional
SAN LUIS GONZAGA



Atribución-NoComercial-SinDerivadas 4.0 Internacional

Esta licencia es la más restrictiva de las seis licencias principales Creative Commons, permitiendo a otras solo descargar sus obras y compartirlas con otras siempre y cuando den crédito, pero no pueden cambiarlas de forma alguna ni usarlas de forma comercial.

<http://creativecommons.org/licenses/by-nc-nd/4.0>



Recibo de pago N° 753934

Visto el Informe N° 017-2025-PIEO-UI-FIMEE-UNSLG, emitido la operaria del sistema de antiplagio se emite la siguiente constancia:

N° 014-2025

CONSTANCIA

El que suscribe, director de la Unidad de Investigación de la Facultad de Ingeniería Mecánica Eléctrica y Electrónica, hace constar que se ha realizado el análisis con el software de verificación de similitud de la **Tesis** cuyo título es:

DETECCIÓN DE ANOMALÍAS EN DIVERSAS REDES DE SENSORES UTILIZANDO MACHINE LEARNING

Presentado por:

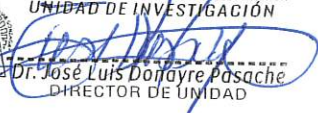
MIRANDA ARANGO, YERSON ELMER

BACHILLER de la Facultad INGENIERÍA MECÁNICA ELÉCTRICA Y ELECTRÓNICA – Escuela Profesional de INGENIERÍA ELECTRÓNICA. El resultado obtenido es un porcentaje de UNO POR CIENTO (1%), por el cual se le otorga el calificativo de:

APROBADO

Se adjunta al presente, el reporte de evaluación con el software de verificación de originalidad.

Ica, 03 de Febrero del 2025

UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA Y ELECTRÓNICA
UNIDAD DE INVESTIGACIÓN

Dr. José Luis Donayre Pasache
DIRECTOR DE UNIDAD

UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"
VICERRECTORADO DE INVESTIGACIÓN
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA Y
ELECTRÓNICA



Detección de anomalías en diversas redes de sensores utilizando machine learning

Línea de investigación

Ciencias naturales, ingeniería y tecnologías sostenibles

INFORME FINAL DE TESIS

Autor:

YERSON ELMER MIRANDA ARANGO

Asesor:

Mag. HIPOLITO MARTIN RODRIGUEZ CASAVILCA

Ica, Perú

2023

Dedicatoria

A mi valiente mamá. Esta tesis es el resultado de tu amor, apoyo y sacrificio en mi viaje educativo. Tus palabras de aliento, tu perseverancia y tu ejemplo constante han sido mi inspiración. Cada día que trabajaste incansablemente y cada vez que me brindaste tu cariño son tesoros que valoro profundamente. También quiero agradecer a mis amigos por sus palabras de aliento, su compañía y su ayuda en los momentos más difíciles. A mis profesores y asesores por su valiosa orientación y experiencia, gracias por compartir su conocimiento conmigo.

Agradecimientos

Deseo expresar mi más sincero agradecimiento a Dios por darme salud, fuerza y habilidades. También quiero extender mi gratitud a todos mis profesores de educación superior, quienes me brindaron las herramientas para mi formación profesional.

Estas palabras van dirigidas a ustedes. A mi madre, por todo su amor, comprensión y apoyo, pero sobre todo por su infinita paciencia conmigo. No tengo palabras para agradecerles las innumerables veces que me brindó su apoyo en todas las decisiones que he tomado a lo largo de mi vida, algunas buenas, otras malas, otras locas. Gracias por darme la libertad de desarrollarme como ser humano.

Quiero agradecer a mi tutor, quien dedicó su tiempo y conocimientos para guiarme en este proceso y siempre estuvo dispuesto a aclarar mis dudas y orientarme en el camino correcto.

Índice

Dedicatoria.....	ii
Agradecimientos	iii
Índice.....	iv
Índice de tablas.....	vii
Índice de figuras.....	ix
Resumen.....	xiii
Abstract.....	xiv
I. Introducción.....	1
1.1. Realidad problemática	1
1.1.1. Problema general.....	1
1.1.2. Problemas específicos	1
1.2. Antecedentes	1
1.2.1. Antecedentes internacionales	1
1.2.2. Antecedente nacional	3
1.3. Justificación e importancia de la investigación.....	3
1.3.1. Justificación.....	3
1.3.2. Importancia.....	4
1.4. Objetivos de la investigación	4
1.4.1. Objetivo general	4
1.4.2. Objetivos específicos.....	4
1.5. Hipótesis y variables de la investigación	4
1.5.1. Hipótesis.....	5
1.5.2. Variables.....	5
1.6. Contenido de la investigación	5
II. Estrategia metodológica	6
2.1. Estudio de predicción de sismos	6
2.2. Estudios de detección de precursores de sismos	7
2.3. Técnica de detección de precursores de sismos basada en modelos	9

2.4. Centro Sismológico Nacional (CENSIS).....	10
2.5. Una técnica de interpolación TEC espacio-temporal.....	11
2.6. Generación de la señal de detección de precursores de sismos	14
2.7. Generación de la señal de confiabilidad TEC	15
2.8. Generación adaptativa de los umbrales de detección de anomalías.....	17
2.8.1. Umbrales de detección de precursor de sismos espaciales.....	18
2.8.2. Umbrales temporales de detección de precursores de sismos	20
2.9. Detectores de precursores de sismos.....	21
2.9.1. Control de falsa alarma	22
2.9.2. Clasificación de las decisiones precursoras.....	23
2.9.3. Técnicas de fusión de decisión.....	25
2.10. Rendimiento de la técnica propuesta de detección de sismos basada en el modelo propuesto	30
III. Resultados	38
3.1. Recopilación y procesamiento de datos	38
3.2. Técnica de detección de precursores de sismos basada en aprendizaje automático: EQ-PD	42
3.2.1. Interpolación TEC espacio-temporal.....	43
3.2.2. Generación de señales de detección de anomalías diurnas y nocturnas	46
3.2.3. Generación adaptativa de los umbrales de detección de anomalías	48
3.2.3.1. Umbrales de detección de anomalías espaciales	49
3.2.3.2. Umbrales de detección espacio-temporales	52
3.2.4. Generación de señal de detección de precursores de ecualización.....	57
3.2.5. Generación de detecciones de precursores de EQ.....	59
3.3. Rendimiento de detección de precursores de la técnica EQ-PD propuesta.....	61
3.3.1. Resultados de pruebas	62
3.3.2. Análisis de fiabilidad.....	68
3.4. Técnica de detección de precursores de sismos basada en aprendizaje profundo	74
3.5. Recopilación y procesamiento de datos	75
3.6. Técnica de detección de precursores EQ basada en aprendizaje profundo: DL-PD	75

3.6.1. Preprocesamiento GIM	76
3.6.2. Predicción GIM basada en redes neuronales profundas.....	79
3.6.3. Detector de precursor de sismos.....	82
3.7. Rendimiento de la técnica propuesta de detección de precursores de EQ basada en DL.....	84
IV. Discusión	92
V. Conclusiones	93
VI. Recomendaciones	94
VII. Referencias bibliográficas	95

Índice de tablas

Tabla I: Error RMS para las medidas de TEC y sus estimaciones en la estación de referencia "datc" en dos fechas diferentes: 16.10.2023 y 17.10.2023.....	13
Tabla II: 21 sismos que han tenido lugar alrededor de Ica en 2023.....	31
Tabla III: Entrenamiento promedio de validación cruzada de 10 veces y rendimiento de prueba general del primero, el segundo y el tercer mejor punto de operación de decisión del precursor del sismo $Fused_{\Delta_3}$	36
Tabla IV: Número de sismos y días sin actividad sísmica en cada pliegue. Entrenamiento y pruebas de desempeño de las terceras mejores decisiones precursoras de sismos $Fused_{\Delta_3}$ para cada pliegue	37
Tabla V: Ubicaciones geodésicas y regiones de las estaciones de referencia elegidas: "AQUI", "VEN1", "TORI" y "MATE".....	40
Tabla VI: Para dos fechas diferentes: 11 de noviembre de 2023 y 12 de noviembre de 2023, diarios AE, A_p, Dst, K_p y SSN	42
Tabla VII: Medias de cluster generadas para los días entre agosto y diciembre del 2023 y valor medio de cada parámetro geomagnético	54
Tabla VIII: Información de entrenamiento, validación y conjunto de prueba	62
Tabla IX: Para los puntos ROC de validación y prueba elegidos: número de precursores de EQ detectados, número de todos los EQ, número de detecciones de precursores falsos, número de días sin actividad sísmica (días de falsa alarma), parámetros de entrenamiento de SVM y distancia promedio de detección de precursores para todos los EQ $\sum M$, 11 EQs débiles con magnitud inferior a 4,3, 10 ecualizadores moderados con magnitud entre 4,3 y 4,6 y 3 EQs fuertes con magnitudes superiores a 4,6 en la escala de Richter	67
Tabla X: Para los puntos ROC de validación y prueba elegidos: $TP_{EQ}, N_{Pos,EQ}, FN_{EQ}, TP_{day}, N_{Pos,day}, FN_{day}, FP_{day}, TN_{day}$ y $N_{Neg,days}$ parámetros de espacio ROC, media y desviación estándar de la puntuación MCC de los experimentos aleatorios μ_{RG} y σ_{RG} , puntuación MCC del punto ROC relacionado (MCC_R), score-Z de MCC_R	73
Tabla XI: 21 EQs diarios diferentes que han tenido lugar en la región elegida.....	74
Tabla XII: Para la red propuesta, formas de salida de cada capa. El número de muestras de entrenamiento se representa como $N_{train,batch}$	81

Tabla XIII: Información de datos GIM-TEC procesada	84
Tabla XIV: Parámetros de entrenamiento del modelo	85
Table XV: Para los puntos ROC elegidos: número de precursores de EQ detectados, número de todos los EQ, número de decisiones de precursores falsos, número de decisiones de precursores falsos que resultaron en la detección de precursores de EQ más débiles, número de días sin actividad sísmica (días de falsa alarma), SSIM umbral de índice y magnitud mínima de EQ en la escala de Richter	89
Tabla XVI: Para los puntos ROC elegidos: distancias medias de detección de precursores de EQ ($\sum d_{prec}$), desviación estándar de las distancias de detección de precursores de EQ ($\sigma(\sum d_{prec})$) y magnitud mínima de EQ en la escala de Richter	89
Tabla XVII: 32 EQ diferentes diarios que han tenido lugar en la región de Anatolia	91

Índice de figuras

Figura N°1: Ubicaciones de las estaciones de referencia del CENSIS	10
Figura N°2: Mediciones diarias de TEC en la estación "datc" en dos fechas diferentes: (a) 16.10.2023 y (b) 17.10.2023.....	11
Figura N°3: Mediciones diarias de TEC y sus estimaciones en la estación "datc" en dos fechas diferentes: (a) 16.10.2023 y (b) 17.10.2023	13
Figura N°4: Visualización de señales de detección de precursores de sismos para dos fechas diferentes: (a) 29 con mancha solar número 34 y (b) 30 con mancha solar número 51	15
Figura N°5: Visualización de señales de confiabilidad de TEC para dos fechas diferentes: (a) 29 con mancha solar número 34 y (b) 30 con mancha solar número 51	17
Figura N°6: Variación local de TEC Distribuciones acumulativas negativas de Pareto para las estaciones "Fethiye" (roja), "Eskisehir" (naranja) y "Samsun" (azul).....	19
Figura N°7: Umbrales de detección de precursores de sismos espaciales con diferentes probabilidades de cola de variación local de TEC: (a) 0.05, (b) 0.005 y (c) 0.0005	20
Figura N°8: Detección temporal de precursores de sismos para los primeros 60 días de 2023 con diferentes parámetros β	21
Figura N°9: (derecha) ajustando la P_{FA} del detector espacial ajustando la probabilidad de cola de variación local $P_{tail;TEC_{\Delta}}$ de TEC, (izquierda) ajustando la P_{FA} del detector temporal ajustando el parámetro β en (ec. 23).....	23
Figura N°10: Decisiones de precursores del detector de precursores de sismos espaciales separados por parámetros de clasificador lineal predefinidos	25
Figura N°11: Series de tiempo de decisión espacial y temporal para 5 días con series de tiempo de decisión fusionadas para diferentes parámetros Δ_d	26
Figura N°12: Funciones de barrera ROC con $\gamma = 3$ y $P_{FA;MIN} = [0.025, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75]$ (de izquierda a derecha).....	29
Figura N°13: Diagrama de flujo de la técnica propuesta de detección de precursores de sismos	30
Figura N°14: Las medidas de TEC y las medidas de TEC nocturnas en la estación "datc" en os fechas diferentes: (a) 16.10.2023 y (b) 17.10.2023 (c) función de ventana de medida de TEC	32

Figura N°15: (a) Curvas ROC para decisiones de precursores de sismos $Fused_{\Delta_0}$ (naranja), $Fused_{\Delta_1}$ (verde), $Fused_{\Delta_2}$ (azul) y $Fused_{\Delta_3}$ (rosa).....	34
Figura N°16: 10 pliegues diferentes sin intersección generados para los días de 2023 con sus etiquetas de actividad sísmica y sin actividad sísmica. El número promedio de días de falsa alarma (días sin clase de actividad sísmica) es de alrededor de 21,1 y el número promedio de sismos es de alrededor de 2,3 por cada pliegue	35
Figura N°17: Las posiciones de las estaciones de referencia CENSIS en Perú.....	39
Figura N°18: Ubicaciones geodésicas de las estaciones de referencia AQUI, VEN1, TORI y MATE con epicentros EQ registrados durante días entre agosto y diciembre del 2023	41
Figura N°19: Diagrama de flujo de la técnica EQ-PD propuesta. Las señales de detección se generan en función de los datos TEC regionales y los umbrales de detección se eligen de forma adaptativa en función de los índices geomagnéticos.....	43
Figura N°20: Estimaciones diarias de TEC $\hat{x}_{u,d;R_r}$ y mediciones de TEC $x_{u,d}$ para la estación de referencia "MATE" para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023	45
Figura N°21: Estimaciones TEC nocturnas x y medidas TEC X para la estación de referencia "MATE" para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023	46
Figura N°22: Señales de detección de anomalías durante todo el día para las estaciones de referencias elegidas en la figura N°17 para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023. Las estaciones de referencia con TEC_{Δ}^{w-day} disponible están marcadas con cuadrados negros y las estaciones de referencia sin TEC_{Δ}^{w-day} están marcadas con cuadrados verdes.....	48
Figura N°23: (a) Distribuciones acumulativas negativas de Pareto de las variaciones locales de TEC de las estaciones de referencia AQUI, VEN1 y TORI para los días entre agosto y diciembre del 2023. (b) Distribuciones acumuladas negativas de Pareto de TEC diurno completo y TECV nocturno de las variaciones locales de TEC de la estación de referencia TORI para los días entre agosto y diciembre del 2023	51
Figura N°24: Umbrales de detección de anomalías espaciales generados para las probabilidades de cola de TEC elegidas (a) 0,005 y (b) 0,05.....	52
Figura N°25: (a) Distribuciones acumulativas de Pareto truncado superior negativo TEC_{Δ}^{w-day} de día completo de la estación de referencia TORI para los tres grupos de día	

diferentes, (b) Distribuciones acumulativas de Pareto truncado superior negativo de $TEC_{\Delta}^{\text{night}}$ nocturno de la estación de referencia TORI para los tres grupos de día diferentes	55
Figura N°26: Umbrales de detección de anomalías espacio-temporales generados para los diferentes grupos y probabilidades de cola de $TEC_{\Delta}^{\text{w-day}}$ elegidas (a) Cluster-1 0.005, (b) Cluster-1 0.05, (c) Cluster-2 0.005, (d) Cluster-2 0.05, (e) Cluster-3 0.005, (f) Cluster-3 0.05.....	56
Figura N°27: Relación entre la probabilidad de cola de TEC y P obtenida de los días sin actividad sísmica durante agosto y diciembre de 2023	58
Figura N°28: Señal de detección de precursores de ecualización	59
Figura N°29: Curvas ROC de los conjuntos de validación (azul) y prueba (naranja) y los puntos ROC de validación (círculo azul) y prueba (círculo naranja) elegidos para (a) Validación1 y Prueba1, (b) Validación2 y Prueba2, (c) Validación3 y Prueba3, (d) Validación4 y Prueba4.....	64
Figura N°30: EQ Histogramas de detección de precursores para puntos ROC (a) Prueba ₁ , (b) Prueba ₂ , (c) Prueba ₃ , (d) Prueba ₄	66
Figura N°31: Histogramas de puntuación de MCC obtenidos de simulaciones aleatorias de Monte Carlo para los puntos ROC elegidos: (a) Validación ₁ , (b) Validación ₂ , (c) Validación ₃ y (d) Validación ₄ . La puntuación MCC del punto ROC elegido está marcada con una flecha en cada figura	71
Figura N°32: Histogramas de puntuación de MCC obtenidos de simulaciones aleatorias de Monte Carlo para los puntos ROC elegidos: (a) Test ₁ , (b) Test ₂ , (c) Test ₃ y (d) Test ₄ . La puntuación MCC del punto ROC elegido está marcada con una flecha en cada figura.....	72
Figura N°33: Mapa GIM-TEC del mundo para la fecha: 2023-08-19T11:00	75
Figura N°34: Diagrama de flujo de la técnica propuesta de detección de precursores de sismos basada en aprendizaje profundo, DL-PD	76
Figura N°35: Mapa recortado del mundo GIM-TEC para la fecha 2023-08-19T11:00.....	77
Figura N°36: (a) 2° (2023-08-18T22:00), (b) 4° (2023-08-19T00:00), (c) 6° (2023-08-19T02:00), (d) 8° (2023-08-19T04:00), (e) 10° (2023-08-19T06:00), (f) 12° (2023-08-19T08:00) y (g) 14° (2023-08-19T10:00) data frames GIM-TEC de una secuencia GIM-TEC de 15 pasos de tiempo para la fecha: 2023-08-19T11:00 (frame 15). (h) <i>ground truth</i> de esta secuencia según GIM-TEC (2023-08-19T12:00).....	78

Figura N°37: Propuesta de red neuronal profunda con múltiples capas espacio-temporales ConvLSTM.....	79
Figura N°38: La capa ConvLSTM con sus entradas espacio-temporales x_t controla las puertas i_t, f_t, o_t , los estados ocultos h_t y los estados de celda c_t . Las conexiones de mirilla también se visualizan en azul.....	80
Figura N°39: Estimación de GIM-TEC y verdad de campo de GIM-TEC para la secuencia de GIM-TEC presentada en la figura N°36	81
Figura N°40: Tiempo de actividad sísmica, índice SSIM, umbral SSIM y decisiones precursoras de EQ para los últimos 100 días de 2023.....	83
Figura N°41: Registros de entrenamiento del modelo propuesto (a) Pérdida de entrenamiento, (b) Pérdida de validación, (c) Entrenamiento PSNR y (d) Validación PSNR	86
Figura N°42: Curvas ROC basadas en el rendimiento de detección de precusores de EQ para EQ con magnitudes (M) (a) $M \geq 5$, (b) $M \geq 5.2$, (c) $M \geq 5.4$ y (d) $M \geq 5.6$. Puntos ROC elegidos (a) ROC_1 , (b) ROC_2 , (c) ROC_3 y (d) ROC_4	87
Figura N°43: Decisiones de tiempo de actividad sísmica, índice SSIM, umbral SSIM y precursor de EQ para los últimos 500 días de 2023. El tiempo sísmicamente activo se genera con respecto a 7 EQ diferentes con $M \geq 5.4$	90
Figura N°44: Tiempo de actividad sísmica, índice SSIM, umbral SSIM, detecciones de precusores de EQ y decisiones falsas de precusores de EQ durante los últimos 25 días antes del EQ que tuvo lugar el 20 de agosto de 2023 con una magnitud $M = 6.6$ en la región de Anatolia.....	90

Resumen

Los sismos han logrado causar daños catastróficos en las infraestructuras urbanas y rurales resultando en pérdida de cientos de vidas humanas, deslizamientos, desprendimientos, y pérdida económica, hechos inevitables pese al avance tecnológico como la predicción y estimación del epicentro, de la magnitud y el tiempo del sismo; por ello, el propósito del presente estudio es desarrollar técnicas para la detección de anomalías en diversas redes de sensores utilizando machine learning. Para estudiar la detección de precursores de un sismo se aplicaron técnicas de detección de precursores de un sismo basada en modelos para la generación de señales de detección de precursores de un sismo, y de fiabilidad del contenido total de electrones. Los resultados obtenidos muestran una alta precisión en la detección de precursores de un sismo, con una tasa de detección superior al 70% para sismos de magnitud superior a 5.4 en la escala de Richter, mientras que se mantiene una baja tasa de falsas alarmas. Además, se logró una distancia promedio de detección de precursores de aproximadamente 2300 km, lo que sugiere que el método es capaz de detectar anomalías ionosféricas relacionadas con actividad sísmica en un área geográfica amplia.

Palabras claves: precursor de un sismo, detección, machine learning.

Abstract

Earthquakes have managed to cause catastrophic damage to urban and rural infrastructure resulting in loss of hundreds of lives, landslides, landslides, and economic loss, inevitable facts despite technological advances such as prediction and estimation of the epicenter, magnitude and time of the earthquake; therefore, the purpose of this study is to develop techniques for the detection of anomalies in various sensor networks using machine learning. To study the detection of earthquake precursors, model-based earthquake precursor detection techniques for the generation of earthquake precursor detection signals and total electron content reliability were applied. The results obtained show a high accuracy in the detection of earthquake precursors, with a detection rate above 70% for earthquakes of magnitude greater than 5.4 on the Richter scale, while maintaining a low false alarm rate. In addition, an average precursor detection distance of approximately 2300 km was achieved, suggesting that the method is capable of detecting ionospheric anomalies related to seismic activity over a wide geographical area.

Keywords: earthquake precursor, detection, machine learning.

I. INTRODUCCIÓN

1.1. Realidad problemática.

La detección de anomalías en diversas redes de sensores utilizando machine learning se presenta como una herramienta crucial en la optimización de sistemas y procesos en el Perú. Sin embargo, la realidad problemática en este ámbito revela una serie de desafíos que afectan la eficacia de estas tecnologías y su implementación efectiva.

En primer lugar, es importante señalar que el Perú enfrenta diversas problemáticas relacionadas con la infraestructura y conectividad. Muchas regiones del país carecen de una infraestructura sólida para la implementación de redes de sensores, lo que limita la capacidad de monitoreo y detección de anomalías. La falta de acceso a tecnologías de última generación en áreas rurales y remotas dificulta la implementación de sistemas avanzados de detección. Otro desafío significativo radica en la falta de conciencia y capacitación en el uso de machine learning para la detección de anomalías. La brecha en la formación de profesionales en este campo impide aprovechar plenamente el potencial de estas tecnologías. Es crucial invertir en programas educativos y de capacitación para que los especialistas peruanos estén equipados con las habilidades necesarias para implementar y mantener sistemas de detección de anomalías de manera efectiva. Además, la diversidad geográfica y climática del Perú agrega complejidad a la detección de anomalías en las redes de sensores. Las condiciones extremas, como sismos, deslizamientos de tierra o eventos climáticos intensos, pueden generar señales confusas o interferencias en los datos recopilados. Es esencial adaptar los algoritmos de machine learning a estas condiciones específicas para garantizar una detección precisa y oportuna.

La falta de recursos financieros y la inversión insuficiente en investigación y desarrollo también son obstáculos importantes. La implementación de sistemas avanzados de detección de anomalías requiere una inversión significativa en tecnología y desarrollo de software. Sin un compromiso adecuado por parte del gobierno y el sector privado, es difícil superar estos desafíos y lograr una implementación efectiva en todo el país.

1.1.1. Problema general.

¿Las técnicas permitirán detectar anomalías en diversas redes de sensores utilizando machine learning?

1.1.2. Problemas específicos.

¿Podrán las técnicas machine learning lograrán minimizar el umbral de las anomalías en las diversas redes de sensores?

¿Una red neuronal deep learning permitirá la detección de sismos premonitores basada en el Global Ionosphere Maps?

1.2. Antecedentes.

1.2.1. Antecedentes internacionales.

En el año 2022, Chen, et al. [1] proponen un novedoso algoritmo numerical differential artificial natural killer cell algorithm (NDANKA) inspirado en la optimización inmunológica para la predicción de sismos. Desarrollaron un sistema de pruebas acústicas y electromagnéticas basado en las señales recopiladas por un sistema desarrollado de pruebas acústicas y electromagnéticas todo en uno (AETA), el registro histórico y los datos de los sismos precursores se combinan como la fuente de datos del modelo propuesto.

En el año 2022, Dong, Li, y Liu [2] recopilaron los datos físicos de la localidad o relacionada al espacio de la superficie de Wenchuan a través del sistema de información geográfica (SIG) para las pruebas de rendimiento, la data es recopilada de los registros de eventos de las réplicas del sismo en Wenchuan, entre ellos, utilizaron 14 000 eventos sísmicos como un conjunto de datos original y utilizaron 8800 eventos de réplica como un conjunto de datos de prueba. Los datos sísmicos involucran la detección de múltiples cantidades físicas, lo que hace que los datos sísmicos aumenten gradualmente, muchos datos no tienen obviamente una relación lineal y los métodos tradicionales de detección difícilmente cumplen con los requisitos de detección.

En el año 2021, Huang, et al. [3] investigaron un sensor electromagnético para evaluar los sismos con anticipación mediante una recopilación de señales sísmicas. En la actualidad, la principal predicción de las magnitudes de los sismos comprende dos métodos, por un lado, la mayoría de los geofísicos o expertos en análisis de datos extraen una serie de características básicas de las señales de sismos precursores para una clasificación sísmica; por otro lado, los datos obtenidos relacionados con las actividades terrestres como el sismógrafo o el satélite espacial se utilizan directamente en las redes de clasificación. Los investigadores proponen una red neuronal convolucional y un diseño en 3D de un mapa de características que pueda utilizar en la resolución de problemas para la clasificación de magnitudes de los sismos al combinar las ventajas de las características superficiales y los datos de alta dimensión. En el 2023, Akyol, Arıkan y Arıkan [4] desarrollaron una nueva técnica machine learning basada en la detección de sismos precursores (EQ-PD) para la detección de sismos precursores en tiempo real con base a los datos del sistema de posicionamiento global y del contenido total de electrones junto con índices geomagnéticos diarios. La técnica EQ-PD propuesta utiliza un clasificador de máquina de vectores de soporte (SVM) para decidir si una anomalía espaciotemporal observada está relacionada con un sismo precursor o no, los datos alimentados al clasificador se componen de un mapa de variabilidad espaciotemporal de una región.

En el 2020, Balestriero, et al. [5] desarrollaron un framework de machine learning no supervisado para detectar y agrupar las señales sísmicas en registros sísmicos continuos. El enfoque propuesto por los investigadores combina una red profunda de dispersión y un modelo de mixtura gaussiana para la agrupación en segmentos de señales sísmicas y la detección de estructuras novedosas. Para un mejor entendimiento analizaron los datos sísmicos adquiridos durante el deslizamiento de tierra en Nuugaatsiaq, Groenlandia durante junio de 2017.

1.2.2. Antecedente nacional.

En el 2021, Hashimoto M., et al. [6] estudiaron eventos de desastres en los que la intensidad logre modelar mediante simulación numérica y/o instrumentación mediante dos procedimientos totalmente automatizados para la detección de edificios gravemente dañados. El primero supone fundamentalmente que las muestras que están ubicadas en áreas con baja intensidad de desastre representan principalmente edificios no dañados, además, las áreas con intensidades de desastre moderadas a fuertes probablemente contengan edificios dañados y no dañados. El segundo se basa en el uso de dos parámetros de regularización para definir la máquina de vector de soporte estos framework evitan la recopilación de muestras en edificios etiquetados a través de un estudio de campo y/o inspección visual de imágenes ópticas, lo que requiere una cantidad significativa de tiempo.

1.3. Justificación e importancia de la investigación.

1.3.1. Justificación.

La investigación se justifica socialmente por abordar problemáticas que impactan la seguridad, la resiliencia ante desastres, la protección del medio ambiente y el desarrollo económico. Económicamente, por ofrecer oportunidades para la eficiencia, la reducción de pérdidas, la innovación tecnológica y el atractivo para inversiones, contribuyendo así al desarrollo económico sostenible. Teóricamente, por su contribución al avance del conocimiento en áreas específicas como machine learning aplicado, ciencia de datos ambientales, desarrollo de modelos teóricos innovadores y aplicación ética de la tecnología. Técnicamente, por abordar desafíos específicos relacionados con la implementación de sistemas de detección de anomalías, optimizando algoritmos, integrando datos heterogéneos, desarrollando sistemas eficientes y aplicando tecnologías avanzadas para mejorar la precisión y la utilidad de estos sistemas en la práctica. Prácticamente, se centra en los impactos tangibles que puede tener en la seguridad, la economía, la agricultura, la gestión de recursos naturales y otros aspectos cruciales para el bienestar y el desarrollo sostenible. Ambientalmente, su capacidad para contribuir activamente a la conservación del medio ambiente, la gestión sostenible de recursos naturales y la reducción de impactos

ambientales negativos. La aplicación efectiva de la detección de anomalías en redes de sensores puede ser una herramienta valiosa para fortalecer los esfuerzos de preservación y sostenibilidad ambiental.

1.3.2. Importancia.

La importancia de la propuesta de investigación se centra en dos enfoques, el primero a los sismos como uno de los desastres naturales más antiguos y mortales en nuestra región y en el Perú, logrando causar daños catastróficos en las infraestructuras urbanas y rurales resultando en pérdida de cientos de vidas humanas, deslizamientos, desprendimientos, tsunamis, incendios y pérdida económica, hechos que son inevitables pese al avance tecnológico en infraestructura, construcción y conectividad de las zonas urbanas y rurales; y un segundo enfoque a la predicción del sismo que tiene como objetivo la estimación del epicentro, de la magnitud y el tiempo del sismo; el valor de la propuesta de investigación se centra en las limitaciones de los investigadores, quienes no han logrado desarrollar una técnica de predicción de ecualización que pueda estimar con precisión el epicentro, la magnitud y el tiempo de las ecualizaciones; el valor de la propuesta de investigación se centra en las limitaciones de los investigadores, quienes no han logrado desarrollar una técnica de predicción de ecualización que pueda estimar con precisión el epicentro, la magnitud y el tiempo de las ecualizaciones, desde hace décadas, la predicción de sismos es una de las áreas de investigación más desafiantes ya que pretende revelar las incertidumbres con respecto a los sismos; y la pertinencia de la propuesta de investigación se centra en los estudios de predicción de sismos todavía hay debates en curso sobre la previsibilidad de los métodos sísmicos de prospección (ecualizadores) centrado en la oscuridad de la actividad sísmica subyacente y en los mecanismos de preparación del sismo así como la confiabilidad de las posibles técnicas de predicción basadas en sismos precursores.

1.4. Objetivos de la investigación.

Según el proyecto propuesto, los objetivos son:

1.4.1. Objetivo general.

Desarrollar técnicas para la detección de anomalías en diversas redes de sensores utilizando machine learning.

1.4.2. Objetivos específicos.

Aplicar las técnicas machine learning para minimizar el umbral de las anomalías en las diversas redes de sensores.

Configurar una red neuronal deep learning para predecir la detección de sismos premonitores basada en el Global Ionosphere Maps.

1.5. Hipótesis y variables de la investigación.

1.5.1. Hipótesis.

Hipótesis general.

Las técnicas permitieron la detección de anomalías en diversas redes de sensores utilizando machine learning.

Hipótesis específicas.

Las técnicas machine learning han logrado minimizar el umbral de las anomalías en las diversas redes de sensores.

La red neuronal deep learning permite la detección de sismos premonitores basada en el Global Ionosphere Maps.

1.5.2. Variables.

Variable independiente.

Diversas redes de sensores

Anomalías

Variable dependiente.

Detección

Machine learning

1.6. Contenido de la investigación.

Para una mejor comprensión en el capítulo I, se hace la introducción describiendo la realidad problemática, los antecedentes, la justificación e importancia, y los objetivos de la investigación. Respecto al capítulo II, se explica la metodología empleada en la investigación con base a la revisión del estado del arte, modelado del sistema, protocolos de enrutamiento, y análisis de las interferencias. De acuerdo a ello, en el capítulo III se explica los resultados obtenidos tras realizar cálculos y simulaciones. En virtud de ello, en el capítulo IV, se discuten los resultados de la investigación. Finalmente, en el capítulo V, se formula las conclusiones de la tesis, y en el capítulo VI, se propone las recomendaciones para futuras investigaciones.

II. ESTRATEGIA METODOLÓGICA

Los sismos (EQs) son uno de los desastres naturales más antiguos y mortíferos. Pueden causar daños colosales en edificios, estructuras y viviendas que resultan en la pérdida de cientos de vidas humanas, lesiones, incendios y pérdidas económicas. Este es un hecho inevitable de que los EQ todavía tienen un impacto peligroso en las vidas humanas a pesar de los avances en las tecnologías de construcción de los edificios. Los efectos perjudiciales de los EQ son difíciles de prevenir, ya que el epicentro de EQ, la magnitud de EQ y el tiempo de inicio son inciertos para la humanidad. La predicción de los EQ tiene como objetivo estimar el epicentro, la magnitud y el tiempo de los EQs [7]. Lamentablemente, aún no se ha desarrollado una técnica de predicción de ecualización que pueda estimar con precisión el epicentro, la magnitud y el tiempo de las ecualizaciones. Por lo tanto, desde hace décadas, la predicción de EQ es una de las áreas de investigación más desafiantes que tiene como objetivo revelar las incertidumbres con respecto a los EQ. En las siguientes secciones, se detallan los estudios de predicción de EQ y detección de precursores de EQ.

2.1. Estudios de predicción de sismos. Todavía hay debates en curso sobre la previsibilidad de los ecualizadores. Estas discusiones se centran en la oscuridad de la actividad sísmica subyacente y los mecanismos de preparación de EQ y la confiabilidad de las posibles técnicas de predicción basadas en precursores de EQ [8, 9]. A pesar de que el fenómeno desencadenado por la actividad sísmica subyacente no se comprende completamente, la discusión en curso aún no ha llegado a una conclusión. Además, la literatura relacionada con la predicción de EQ consiste en estudios de casos empíricos que se llevan a cabo para detectar señales precursoras de EQ después de que se haya producido un EQ fuerte. Debido a la falta de estudios exhaustivos que investiguen el comportamiento de las señales precursoras durante intervalos de tiempo más largos, aún existen críticas relacionadas con la confiabilidad y repetibilidad de estas técnicas [10]. Las técnicas de predicción de EQ en la literatura se agruparán en técnicas de predicción basadas en modelos y basadas en precursores. Las técnicas basadas en modelos tienen como objetivo ajustar un modelo matemático, de aprendizaje automático o de aprendizaje profundo para simular un comportamiento de recurrencia entre actividades sísmicas pasadas o ecualizaciones con las próximas ecualizaciones. Algunos de los modelos matemáticos de predicción de EQ elegidos en la literatura se enumeran a continuación. Un modelo de reaparición de EQ propuso una relación entre las fuerzas relacionadas con la deformación de la línea de falla y la periodicidad de reaparición de EQ [11]. Otro modelo que utiliza distribuciones de Poisson y factores de distancia afirma que todas las ocurrencias de EQ en una zona de falla están relacionadas y deben predecirse [12]. De manera similar, se propone un modelo probabilístico para predecir la ubicación, la magnitud y el tiempo de los próximos EQ [13].

Se ajusta un modelo Fibonacci Lucas Dual a los tiempos de ocurrencia de EQs pasados para predecir los próximos tiempos de recurrencia de EQs [14]. En los últimos años, también se han propuesto varios modelos basados en aprendizaje automático para predecir ubicaciones, magnitudes o tiempos de aparición de EQ. Estos modelos de aprendizaje automático se basan en: técnicas de impulso como AdaBoost, árboles de decisiones y bosques aleatorios, red difusa de información multiobjetivo, técnicas no supervisadas como k-vecinos más cercanos, Support Vector Machine (SVM) y redes neuronales artificiales [15, 16, 17, 18, 19]. Además, los regresores SVM y las redes neuronales híbridas están capacitados para predecir EQ en un catálogo de sismos elegido [20]. También se proponen modelos basados en aprendizaje profundo [21, 22, 23, 24, 25].

2.2. Estudios de detección de precursores de sismos. Los estudios relacionados con la detección de precursores de sismos tienen como objetivo detectar anomalías en varias señales precursoras antes de fuertes ecualizadores. Estas señales precursoras se generarán a partir de emisiones de gas radón [26, 27], fenómeno ionosférico [28, 29, 30, 31, 32], variaciones de temperatura [33, 34, 35], composición química del agua subterránea [36, 27] e incluso comportamiento animal inusual [37] o luces extrañas [38]. Hay estudios que sugieren los posibles mecanismos de acoplamiento entre las actividades sísmicas y el estado de la ionosfera antes de fuertes ecualizaciones. En [39] se propone un posible mecanismo de acoplamiento entre la litosfera, la atmósfera y la ionosfera de la Tierra. El modelo Lithosphere Atmosphere Ionosphere Coupling (LAIC) afirma que la actividad sísmica o los movimientos de las placas tectónicas generan emisiones de gas de radón desde la corteza terrestre a través de la atmósfera. Los gases de radón liberados ionizan el aire en la atmósfera, lo que resulta en cambios en la ionosfera. Estos cambios se observan en la frecuencia crítica de la capa máxima de ionización (foF2), la altura máxima de ionización (hmF2) y los parámetros de Contenido Total de Electrones (TEC) de la ionosfera. Otro mecanismo de acoplamiento físico se propone en [40]. La propuesta afirma que la baja altura de la capa límite planetaria (CLP) provoca la acumulación de gases radón que se liberan de la corteza terrestre. Esta acumulación de gas radón da como resultado una alta ionización en la ionosfera durante la noche. Además, el propuesto también afirmó que las anomalías de densidad de electrones positivos resultantes tienen una característica única. Por lo tanto, es posible crear un patrón (máscara) que puede servir como una señal precursora de ecualización para próximas ecualizaciones fuertes. Se realizan investigaciones exhaustivas sobre las posibles correlaciones entre los parámetros ionosféricos de la capa F2 y las actividades sísmicas y los ecualizadores. Se realiza un análisis estadístico preliminar de las frecuencias críticas de foF2 que se miden en la estación terrestre de Chung-Li (isla de Taiwán) para los ecualizadores con magnitudes $M \geq 4$ que tuvieron lugar cerca de esta

estación terrestre durante el período de 1978 a 1986 [41]. Los datos ionosféricos se procesan durante 6 días antes de cada hora de ecualización elegida. Los resultados de este estudio afirman que los ecualizadores de enfoque superficial ($z < 60$ km) apenas generan un precursor de ecualización detectable, mientras que los ecualizadores de enfoque profundo ($60 \text{ km} < z < 300 \text{ km}$) provocan un aumento significativo en la respuesta ionosférica antes de los próximos ecualizadores con magnitudes $M \geq 4$. Se lleva a cabo otro estudio sobre los EQs que se llevaron a cabo en el área de Taiwán durante el período de 1997 a 1999 [42]. Los resultados de este estudio sugieren que es posible observar anomalías de foF2 y TEC de 1 a 7 días antes de los EQs investigados. De manera similar, se realiza otra investigación estadística relacionada con foF2 ionosférico para los 184 EQ con una magnitud $M \geq 5.0$ durante 1994 - 1999 en el área de Taiwán [43]. El estudio informó evidencias empíricas de anomalías ionosféricas previas al sismo (PEIA) de 1 a 5 días antes de los sismos investigados de 184 $M \geq 5.0$. Se lleva a cabo otro estudio integral para investigar la relación entre las anomalías TEC ionosféricas y las actividades sísmicas. En el estudio, los datos de TEC se extraen de Global Ionosphere Maps (GIM) [44]. Durante la investigación, se seleccionan 736 $M \geq 6.0$ sismos en todo el mundo durante 2002 - 2010. Las anomalías de TEC se definen en los datos de TEC elegidos con respecto a los ecualizadores elegidos. Se calculan las estadísticas de las tasas de ocurrencia de anomalías de 1 a 21 días antes de los EQs (PE) y los días de fondo sin actividad sísmica (PN). Los resultados de la investigación revelaron que la tasa de anomalías de TEC comienza a aumentar durante 10 días antes de los EQ. Además, la probabilidad de observar una anomalía de TEC durante 1 a 10 días antes de un EQ fuerte es de 1.8 a 3.6 veces más probable en comparación con la observación de estas anomalías de TEC durante los días de fondo elegidos. En otro estudio estadístico, las anomalías ionosféricas previas al EQ se investigan en los datos TEC obtenidos de GIM para los EQ elegidos en Japón [45]. Los resultados del estudio sugieren el hecho de que la probabilidad de observar anomalías GIM-TEC aumenta entre 1 y 5 días antes de los EQ investigados. Las investigaciones estadísticas presentadas [41, 42, 43, 44, 45] realizaron sus investigaciones durante los períodos de tiempo sísmicamente activos en los que las anomalías ionosféricas detectadas se atribuyen a actividades sísmicas. Por lo tanto, las estadísticas de recurrencia de anomalías ionosféricas cuando no hay actividad sísmica siguen siendo inciertas. También hay otros estudios relacionados con precursores ionosféricos que informan anomalías ionosféricas antes de EQs fuertes en Italia [46, 47, 48, 49], Taiwán [50, 51, 52], Indonesia [53], Japón [29, 54, 55, 32], Chile [31, 56], India [57], Nepal [58], China [59, 29], México [30, 60, 61], El-Salvador [62], EE. UU. [63, 64] y Turquía [65, 66, 67]. En estos estudios, se han practicado varios métodos estadísticos para comprender la relación entre las variaciones ionosféricas y las actividades sísmicas, tales como: análisis de variación y diferencia de TEC [31, 29, 54, 55, 56, 62, 65], corrección ionosférica [56], análisis

intercuartil y percentil [32, 50, 51, 52, 58, 59], análisis de correlación entre TEC y foF2 [30, 64, 61, 60], análisis de anomalías de TEC ecuatorial [53], generación de patrones de precursores basados en parámetros ionosféricos [49], análisis de desviación relativa de TEC [47, 63] e investigación de variación sismo-ionosférica [46] y detección [48].

2.3. Técnica de detección de precursores de sismos basada en modelos. Las redes GPS con posiciones de receptor conocidas con precisión se implementan para proporcionar información de geolocalización en tiempo real de alta precisión. Estas redes de GPS ubicadas permanentemente también generarán mediciones indirectas de TEC ionosféricas para monitorear el estado de la ionosfera. En este trabajo, las mediciones TEC de Centro Sismológico Nacional (CENSIS) se procesan y presentan en la sección 2.5. Existen múltiples fuentes de ionización en la ionosfera, como la fuerte radiación solar, la emisión de gas radón, la química ionosférica y las fuertes actividades sísmicas. Las mediciones de TEC varían tanto espacial como temporalmente debido a la ionización en la atmósfera. La técnica propuesta de detección de precursores de sismos basada en un modelo se basa en la detección de anomalías TEC locales que son provocadas por fuertes actividades sísmicas en la corteza terrestre. La técnica de detección propuesta consta de cinco etapas de procesamiento secuencial. En primer lugar, las mediciones de TEC se estiman utilizando una técnica de estimación espacio-temporal como se detalla en la sección 2.6. Luego, la detección de precursores de sismos y las señales de confiabilidad de TEC se generan como se presenta en la sección 2.7 y la sección 2.8, respectivamente. La señal de detección de precursores de sismos generada tiene un umbral para declarar una decisión de precursor mediante el uso de umbrales de detección de precursores como se describe en la sección 2.9. Las decisiones de precursores declarados son validadas por detectores de anomalías que utilizan la señal de confiabilidad de TEC y los parámetros de detección para evaluar la confiabilidad de las decisiones de anomalías como se presentó en la sección 2.10. Finalmente, como se detalla en la sección 2.10.3, las decisiones precursoras validadas de los detectores precursores se fusionan para generar detecciones precursoras sísmicas finales que tienen mayor robustez y rendimiento en términos de índices de falsas alarmas más bajos.

2.4. Centro Sismológico Nacional (CENSIS). Se trata de un servicio creado por el Instituto Geofísico del Perú con el propósito de vigilar la aparición de sismos en el país, con el objetivo de proporcionar información completa y oportuna tanto a los miembros del SINAGERD como a la población en general.

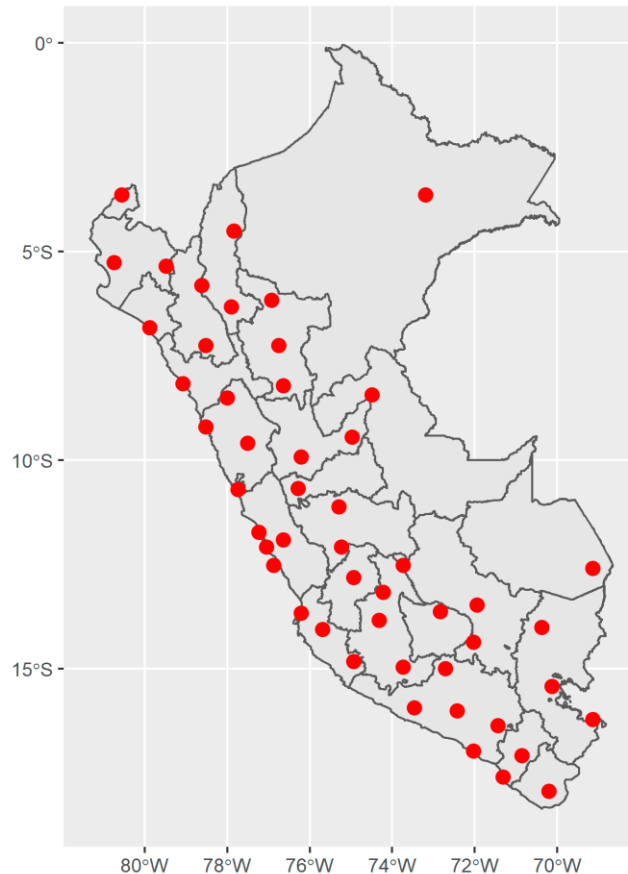


Figura N°1: Ubicaciones de las estaciones de referencia del CENSIS.

Fuente: Elaboración propia.

Los datos brutos del CENSIS que se registran en cada estación de referencia operativa son procesados por los servicios IONOLAB-TEC e IONOLAB-BIAS [63, 64]. Las mediciones de TEC preprocesadas se generarán con una separación de tiempo de 0,5 segundos. En este trabajo se postprocesan medidas de TEC con separación de tiempo de 2.5 minutos. Las muestras de los datos TEC procesados se presentan en la figura N°2.

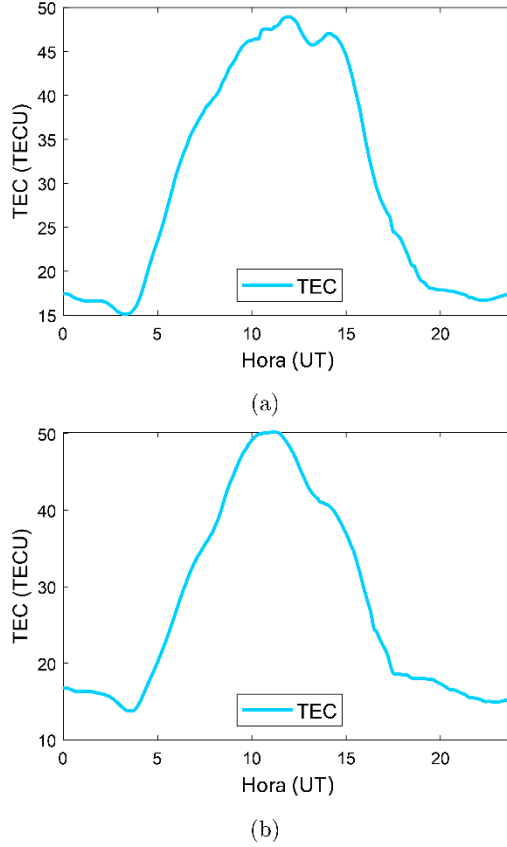


Figura N°2: Mediciones diarias de TEC en la estación "datc" en dos fechas diferentes:
(a) 16.10.2023 y (b) 17.10.2023.
Fuente: Elaboración propia.

Para dos fechas diferentes de 16.10.2023 y 17.10.2023, se visualizan las mediciones diarias de TEC preprocesadas en la estación de referencia con el nombre de código "datc". Como se ve en la figura N°2, a pesar de que las mediciones diarias de TEC se registran en dos días consecutivos, existen diferencias temporales entre las grabaciones de TEC por hora.

2.5. Una técnica de interpolación TEC espacio-temporal. Los comportamientos anómalos en las mediciones de TEC se detectarán generando estimaciones para las mediciones de TEC y comparando estas estimaciones con las mediciones de TEC. La estimación de TEC de una estación de referencia se generará utilizando las mediciones de TEC de estaciones de referencia vecinas (estimación de TEC espacial), mediciones de TEC pasadas de la misma estación de referencia (estimación de TEC temporal) o ambas mediciones de TEC pasadas de la estación de referencia y sus vecinas (estimación TEC espacio-temporal). En este trabajo se implementa una técnica de estimación espacio-temporal de TEC [65, 60].

Sea $\mathbf{x}_{u,d}$ ($(N_{u,d} \times 1)$ columna de vector) definido como los datos TEC disponibles en la estación de referencia u el día d :

$$\mathbf{x}_{u,d} = [x_{u,d}(1) \dots x_{u,d}(n) \dots x_{u,d}(N_{u,d})]^T \quad (1)$$

donde $N_{u;d}$ es el número total de mediciones de TEC, $x_{u;d}(n)$, es la n -ésima medición de TEC de $N_{u;d}$ mediciones de TC consecutivas. Las estaciones de referencia vecinas de la estación de referencia u se elegirán con respecto a un círculo donde el centro del círculo es la ubicación de la estación de referencia u y el radio del círculo es R_r km. Los datos TEC de la v -ésima estación de referencia que se encuentra dentro del círculo se definirán como $\mathbf{x}_{v;d;R_r}$ ($(N_{u;d} \times 1)$ vector) y el número total de estaciones de referencia vecinas se definirá como $N_{u;R_r}$. Luego, se obtendrá una estimación espacio-temporal de TEC de la estación de referencia u elegida en el día d , $\hat{\mathbf{x}}_{u;d;R_r}$ como

$$\hat{\mathbf{x}}_{u;d;R_r} = \sum_{v=1}^{N_{u;R_r}} \alpha_{u;d;R_r}(v) \mathbf{x}_{v;d;R_r}, \quad (2)$$

donde $\alpha_{u;d;R_r}(v)$ es el peso de la v -ésima estación vecina. Por lo tanto, el vector de pesos $\alpha_{u;d;R_r}(N_{u;R_r} \times 1)$ se definirá como la única solución de la siguiente minimización:

$$\alpha_{u;d;R_r} = \underset{\alpha_{u;d;R_r}(v)}{\operatorname{argmin}} \sum_{d_i}^{d_s} \left\| \mathbf{x}_{u;d_n} - \sum_{v=1}^{N_{u;R_r}} \alpha_{u;d;R_r}(v) \mathbf{x}_{v;d_n;R_r} \right\|_2^2 \quad (3)$$

$\alpha_{u;d;R_r}$ también se obtendrá en forma cerrada como:

$$\underline{\alpha}_{u;d;R_r} = \left(\sum_{d_n=d_i}^{d_s} \mathbf{X}_{u;d_n;R_r}^T \mathbf{X}_{u;d_n;R_r} \right)^{-1} \left(\sum_{d_n=d_i}^{d_s} \mathbf{b}_{u;d_n;R_r} \right) \quad (4)$$

donde $\mathbf{X}_{u;d_n;R_r}$ matriz ($N_{u;d} \times N_{u;R_r}$) se forma combinando múltiples mediciones TEC de R_r km. estaciones de referencia vecinas como:

$$\mathbf{X}_{u;d_n;R_r} = [\mathbf{x}_{1;d_n;R_r} \cdots \mathbf{x}_{v;d_n;R_r} \cdots \mathbf{x}_{N_{u;R_r};d_n;R_r}] \quad (5)$$

y $\mathbf{b}_{u;d_n;R_r}$ vector ($N_{u;R_r} \times 1$) se define como:

$$\mathbf{b}_{u;d_n;R_r} = \mathbf{X}_{u;d_n;R_r}^T \mathbf{x}_{u;d_n} \quad (6)$$

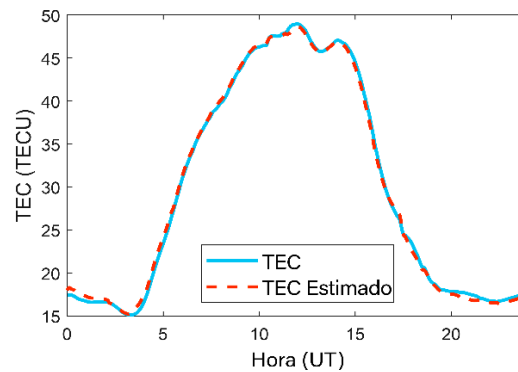
La densidad de electrones de la ionosfera y la fuerte actividad magnética del sol basada en la radiación solar están altamente correlacionadas. El número de manchas solares (SSN) es un índice ampliamente utilizado para monitorear el sol y la actividad solar diariamente [72]. Dado que la técnica de estimación de TEC espacio-temporal propuesta depende de observaciones de TEC anteriores, los días dentro de $[d_i, d_s]$ en (ec. 4) se agrupan con respecto a su SSN correspondiente. Los datos de TEC anteriores que están en el mismo grupo con el día de estimación se procesan durante la minimización en (ec. 4). Por lo tanto, el proceso de minimización se correlaciona temporalmente con la actividad magnética del sol.

Utilizando los pesos obtenidos $\underline{\alpha}_{u;d_n;R_r}$, $\hat{\mathbf{x}}_{u;d;R_r}$ puede calcularse y compararse con $\mathbf{x}_{u;d}$ para detectar las anomalías locales de TEC. El rendimiento de la estimación de TEC de la técnica de interpolación espacio-temporal de TEC se presenta en la tabla I y la figura N°3 para la estación de referencia "datc" en dos días 16.10.2023 y 17.10.2023 diferentes que también se ilustran en la figura N°2. Error de estimación de Root Mean Square TEC, $TEC_{RMS,u,d}$ se define en (ec. 7), y se calcula para la estación de referencia y las fechas elegidas. Como se muestra en la tabla I, la raíz del error cuadrático medio (RMS) de ambas fechas de estimación es cercana a cero.

$$\hat{\mathbf{x}}_{RMS,u,d,R_r} = \sqrt{\frac{1}{N_{u,d}} \|\mathbf{x}_{u,d} - \hat{\mathbf{x}}_{u;d;R_r}\|_2^2} \quad (7)$$

Tabla I
Error RMS para las medidas de TEC y sus estimaciones en la estación de referencia "datc" en dos fechas diferentes: 16.10.2023 y 17.10.2023.

16.10.2023 "Datça" Sunspot Number: 60	17.10.2023 "Datça" Sunspot Number: 74
$\hat{\mathbf{x}}_{RMS,u,d,R_r} = 0.4230$ TECU	$\hat{\mathbf{x}}_{RMS,u,d,R_r} = 0.4214$ TECU



(a)

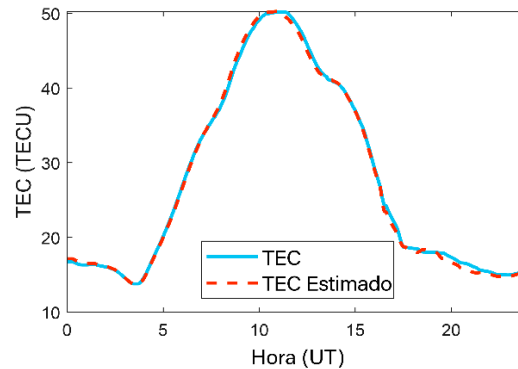


Figura N°3: Mediciones diarias de TEC y sus estimaciones en la estación "datc" en dos fechas diferentes: (a) 16.10.2023 y (b) 17.10.2023.

Fuente: Elaboración propia.

2.6. Generación de la señal de detección de precusores de sismos. Una anomalía se definirá como una "variación de la norma" [73]. La detección de una anomalía de TEC requiere comparar las mediciones de TEC con la norma o, por ejemplo, las estimaciones de TEC. El resultado de esta comparación es la puntuación de anomalía de las mediciones de TEC comparadas o, por ejemplo, la variación local de TEC. La cantidad de variación local de TEC en una medición de TEC y su estimación se calculará utilizando la distancia simétrica de Kullback-Leibler (SKLD) entre la $\mathbf{x}_{u,d}$ normalizada y su estimación normalizada $\hat{\mathbf{x}}_{u,d;R_r}$ [74, 75]:

$$SKLD(P_{u,d}; \hat{P}_{u,d;R_r}) = KLD(\hat{P}_{u,d;R_r}|P_{u,d}) + KLD(P_{u,d}|\hat{P}_{u,d;R_r}) \quad (8)$$

donde KLD es la distancia Kullback-Leibler dada por:

$$KLD(\hat{P}_{u,d;R_r}|P_{u,d}) = \left(\sum_{n=1}^{N_{u,d}} \hat{P}_{u,d;R_r}(n) \ln \left(\frac{\hat{P}_{u,d;R_r}(n)}{P_{u,d}(n)} \right) \right) \quad (9)$$

$$KLD(P_{u,d}|\hat{P}_{u,d;R_r}) = \left(\sum_{n=1}^{N_{u,d}} P_{u,d}(n) \ln \left(\frac{P_{u,d}(n)}{\hat{P}_{u,d;R_r}(n)} \right) \right) \quad (10)$$

donde $P_{u,d}$ y $\hat{P}_{u,d;R_r}$ son las $\mathbf{x}_{u,d}$ y $\hat{\mathbf{x}}_{u,d;R_r}$ y Z normalizadas como:

$$P_{u,d} = \frac{\mathbf{x}_{u,d}}{\|\mathbf{x}_{u,d}\|_1} \quad (11)$$

$$\hat{P}_{u,d;R_r} = \frac{\hat{\mathbf{x}}_{u,d;R_r}}{\|\hat{\mathbf{x}}_{u,d;R_r}\|_1} \quad (12)$$

La señal de detección de precusores de sismos que puede servir para detectar anomalías en las variaciones locales de TEC se construye mediante la interpolación 2D de todos los SKLD calculados para todas las estaciones en el CENSIS [76, 77]. Para dos fechas consecutivas: 29.08.2023 y 30.08.2023, se generan señales de detección de precusores de sismos y se presentan en la figura N°4. Como se muestra en la figura N°4, a pesar de que las fechas de las señales de detección de precusores generadas son de dos días consecutivos, muestran diferencias de variación de TEC locales espaciales y temporales.

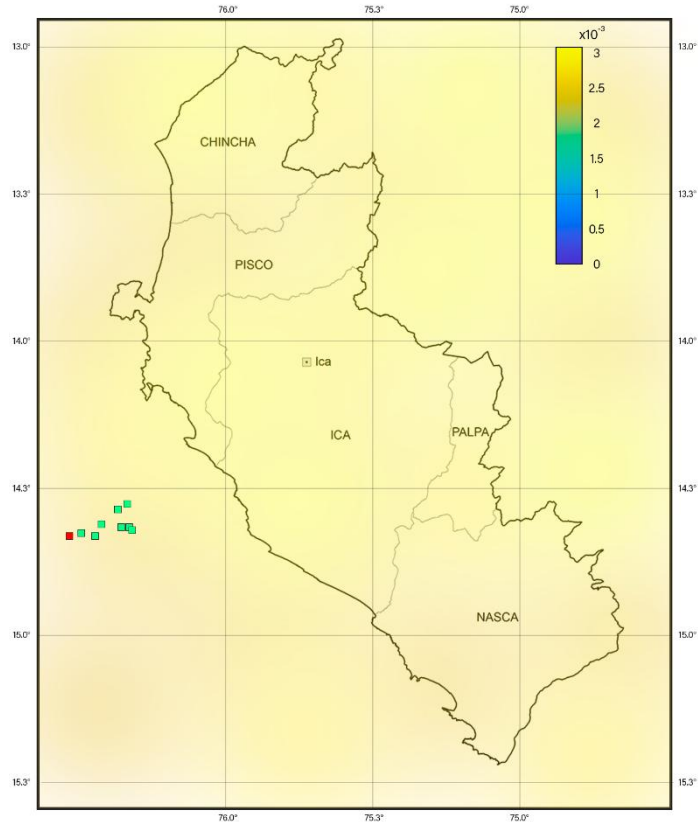


Figura N°4: Visualización de señales de detección de precusores de sismos para dos fechas diferentes: (a) 29 con mancha solar número 34 y (b) 30 con mancha solar número 51.

Fuente: Elaboración propia.

2.7. Generación de la señal de confiabilidad TEC. Las perturbaciones TEC ionosféricas no solo son provocadas por la radiación solar y las actividades sísmicas, sino que también se ven afectadas por las tormentas geomagnéticas. Para una detección robusta y confiable de anomalías TEC locales, las anomalías TEC causadas por las tormentas geomagnéticas deben tenerse en cuenta al monitorear los parámetros relacionados con las tormentas geomagnéticas. D_{st} , K_p y el índice de tormentas ionosféricas planetarias W_p son algunos de los parámetros ampliamente utilizados para identificar las tormentas geomagnéticas [78]. W_p se puede obtener fácilmente mediante el uso de mediciones TEC y puede servir para identificar las tormentas ionosféricas incluso en condiciones de tiempo de quietud magnética [79, 80]. Por lo tanto, W_p es un candidato razonable para evaluar la confiabilidad de las mediciones de TEC. El parámetro W_p es un parámetro discreto que se puede obtener clasificando la variabilidad de la ionosfera, $\mathbf{w}_{u,d}$. $\mathbf{w}_{u,d}$ puede tomar valores negativos y positivos que indican tormentas geomagnéticas negativas y positivas, ya que sus magnitudes son proporcionales a $\mathbf{w}_{u,d}$ respectivamente. A medida que $\mathbf{w}_{u,d}$ se acerca a cero, el estado de la ionosfera se vuelve más tranquilo. Para representar la confiabilidad de la variación local

en TEC para la detección de anomalías locales debidas a actividades sísmicas, $\mathbf{w}_{u;d}$ se definirá como:

$$\mathbf{w}_{u;d} = \log_{10}(\mathbf{x}_{u;d}/x_{u;d;med}) \quad (13)$$

donde la variabilidad local de la ionosfera, $\mathbf{w}_{u;d}$ es una versión no cuantificada del parámetro W_p ionosférico y $x_{u;d;med}$ se define como la mediana de los 7 días TEC anteriores al día d , en la estación u . En este trabajo, las señales de detección de precursores de sismos se generan calculando las variaciones locales de TEC basadas en SKLD para todas y cada una de las estaciones de la red GPS. De manera similar, las señales de confiabilidad de TEC se generan calculando las variabilidades locales de la ionosfera, $\mathbf{w}_{u;d}$ para todas y cada una de las estaciones de la red GPS. Posteriormente, se genera una estimación de las variabilidades locales de la ionosfera, $\hat{\mathbf{w}}_{u;d}$, mediante el algoritmo espacio-temporal descrito en la sección 2.6 para reducir la variación espacial de las variabilidades locales de la ionosfera. La puntuación de fiabilidad del TEC, el $TEC_{\alpha;u;d}$ para la estación de referencia u y el día d se definirán de la siguiente manera:

$$TEC_{\alpha;u;d} = \max(\hat{\mathbf{w}}_{u;d}) \quad (14)$$

De manera similar, la señal de confiabilidad de TEC que servirá para evaluar la confiabilidad de las mediciones de TEC también se genera mediante la interpolación 2D de todos los valores de confiabilidad de TEC calculados para todas las estaciones en el CENSIS. La figura N°5 ilustra las señales de confiabilidad de TEC generadas para los mismos días que también se presentan en la figura N°4.

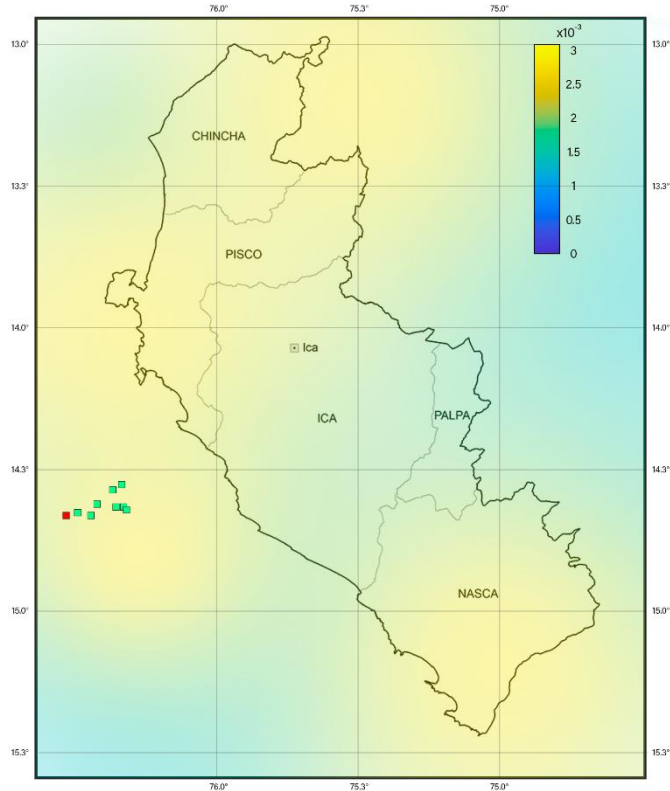


Figura N°5: Visualización de señales de confiabilidad de TEC para dos fechas diferentes:
 (a) 29 con mancha solar número 34 y (b) 30 con mancha solar número 51.

Fuente: Elaboración propia.

Como se presenta en la figura N°5, las regiones de Chincha e Ica muestran puntajes de confiabilidad de TEC más altos en comparación con otras regiones de Ica. Este problema debe atribuirse al hecho de que estas regiones se encuentran en el límite de la placa tectónica de Eurasia y la placa tectónica de Arabia. Por lo tanto, estas regiones pueden exhibir actividades sísmicas frecuentes que dan como resultado puntajes de confiabilidad más altos. A continuación, se presentan dos detectores de precursores de sismos diferentes que funcionan con estas señales precursoras y de fiabilidad.

2.8. Generación adaptativa de los umbrales de detección de anomalías. Un límite de decisión se definirá como un hiperplano que separa el espacio de observación en múltiples regiones de decisión [81]. Las señales de detección de precursores de sismos espacio-temporales generadas son el espacio de observación de la técnica de detección de anomalías propuesta. Por lo tanto, los umbrales de detección de anomalías deben separar las variaciones anómalas de TEC de las variaciones regulares de TEC en las señales de detección de precursores de sismos generadas. Estos umbrales de detección de anomalías se generarán con respecto a las variaciones espaciales, temporales o espacio-temporales de las señales de detección de precursores sísmicos generadas. En este trabajo, se generan umbrales de detección de precursores de sismos espaciales que son constantes en el tiempo para detectar anomalías

TEC espaciales y umbrales de detección de precursores de sismos temporales que son constantes en el espacio para detectar anomalías TEC temporales. Cada día clasificado en dos clases no superpuestas denominadas: "actividad sísmica" y "sin actividad sísmica" y las estadísticas de variación de TEC de los días sin clase de actividad sísmica se utilizan durante el proceso de generación del umbral.

2.8.1. Umbrales de detección de precursores de sismos espaciales. Los umbrales de detección de precursores de sismos espaciales se generarán estimando el comportamiento de variación de TEC de cada estación de referencia en los días de clase sin actividad sísmica [82]. El comportamiento de variación de TEC de cada estación de referencia se estima mediante distribuciones acumulativas negativas de Pareto. Cada distribución acumulativa negativa de Pareto es única para una estación de referencia CENSIS y asigna cada variación local de TEC de esta estación de referencia a una probabilidad de cola. Por lo tanto, elegir una probabilidad de cola constante corresponde a un valor de variación de TEC local diferente para cada estación de referencia CENSIS.

Las distribuciones de Pareto truncadas superiores proporcionan una caracterización estadística adecuada en las mediciones de TEC densas, ya que la medición de TEC y sus estimaciones generadas tienen un valor positivo. Para una variable aleatoria Z de Pareto truncada en la parte superior, su probabilidad de cola viene dada por el siguiente modelo de tres parámetros en ρ, γ y ϑ [76, 77]:

$$P(Z > z) = \frac{\gamma^\rho (z^{-\rho} - \vartheta^{-\rho})}{1 - \left(\frac{\gamma}{\vartheta}\right)^\rho}, 0 < \gamma \leq z \leq \vartheta < \infty \quad (15)$$

Las estimaciones de máxima verosimilitud (ML) para los parámetros de distribución son:

$$\hat{\gamma} = \min\left(TEC_{\Delta;u}(1), \dots, TEC_{\Delta;u}(N_{ns})\right), \quad (16)$$

$$\hat{\vartheta} = \max\left(TEC_{\Delta;u}(1), \dots, TEC_{\Delta;u}(N_{ns})\right) \quad (17)$$

y $\hat{\rho}$ se obtiene como solución de:

$$\frac{N_{ns}}{\hat{\rho}} + \frac{N_{ns} \left(\frac{\hat{\gamma}}{\hat{\vartheta}}\right)^{\hat{\rho}} \log\left(\frac{\hat{\gamma}}{\hat{\vartheta}}\right)}{1 - N_{ns} \left(\frac{\hat{\gamma}}{\hat{\vartheta}}\right)^{\hat{\rho}}} = \sum_{i=1}^{N_{ns}} \left[\log\left(TEC_{\Delta;u}(i)\right) - \log(\hat{\gamma}) \right] \quad (18)$$

donde N_{ns} es el número de días sin actividad sísmica y $TEC_{\Delta;u}(i), 0 \leq i \leq N_{ns}$ es una serie temporal de las variaciones locales de TEC obtenidas a partir de los días sin actividad sísmica para la estación de referencia u [82. 83].

Los parámetros $\hat{\gamma}$, $\hat{\vartheta}$ y $\hat{\rho}$ se obtendrán utilizando (ec. 16), (ec. 17) y (ec. 18), respectivamente. Posteriormente, se estimará la distribución probabilística de la probabilidad de cola de la variación local de TEC, $P_{tail;TEC_{\Delta}}$, para todas las estaciones en el CENSIS para los días sísmicamente inactivos. Como demostración de ejemplo del rendimiento de la estimación, en la figura N°6 se ilustran tres distribuciones de probabilidad de cola de variación de TEC locales diferentes para tres estaciones de referencia diferentes "Fethiye", "Eskisehir" y "Samsun".

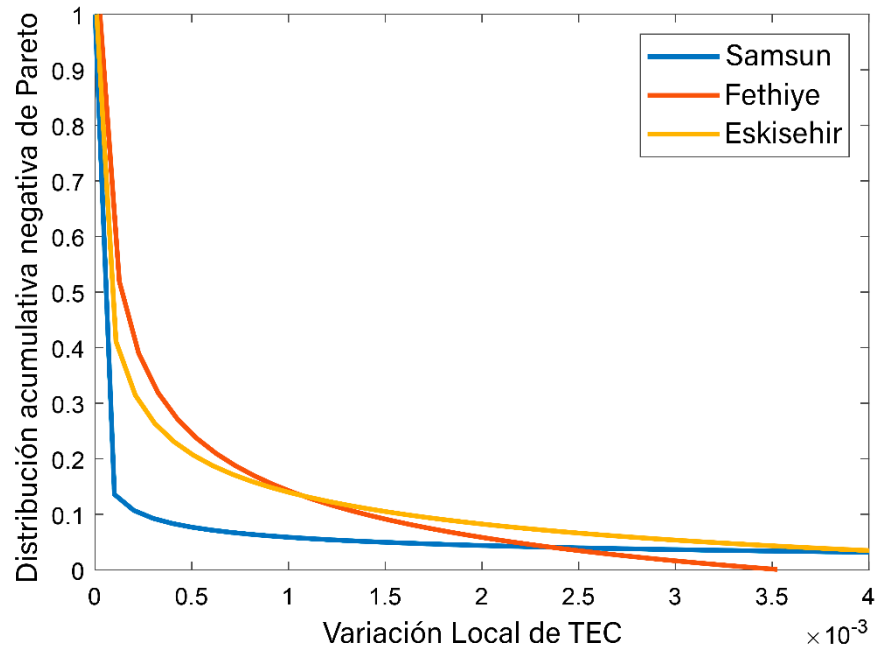


Figura N°6: Variación local de TEC Distribuciones acumulativas negativas de Pareto para las estaciones "Fethiye" (roja), "Eskisehir" (naranja) y "Samsun" (azul).

Fuente: Elaboración propia.

Como se presenta en la figura N°6, las distribuciones de probabilidad de cola de variación local estimadas de TEC muestran una variación espacial que las distribuciones varían con respecto a las ubicaciones de las estaciones de referencia. Cuando se elige una probabilidad de cola constante como $P_{tail;TEC_{\Delta}} = 0.1$, la probabilidad de cola corresponde a una variación de TEC local diferente en las estaciones de referencia "Fethiye", "Eskisehir" y "Samsun" como 0,0013, 0,0015 y 0,0002, respectivamente. Si este procedimiento se repite para todas las estaciones de referencia CENSIS y las correspondientes variaciones locales de TEC se eligen como, umbrales espaciales, entonces se generará un umbral espacial de detección de precursores de sismos que corresponda a la probabilidad de cola elegida. En la figura N°7 se ilustran tres umbrales de detección de precursores de sismos espaciales diferentes para tres probabilidades de cola de variación de TEC diferentes. Las

probabilidades de cola se eligen como 0.05 en la figura N°7a, 0.005 en la figura N°7b y 0.0005 en la figura N°7c.

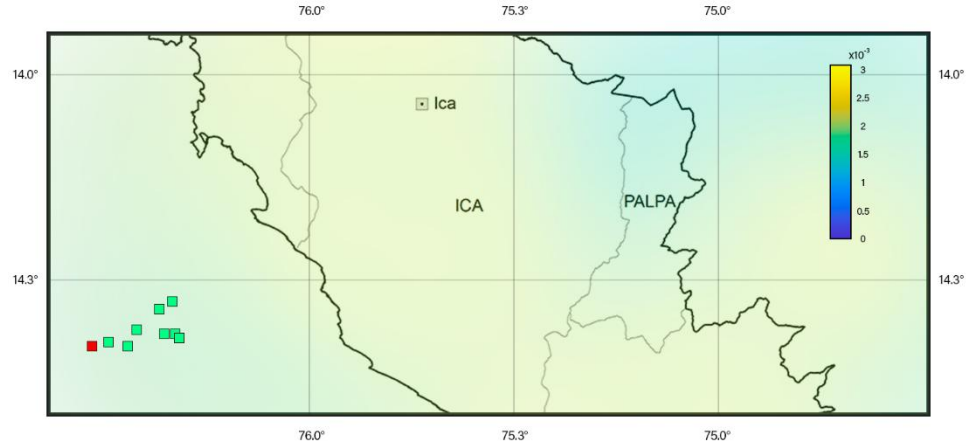


Figura N°7: Umbrales de detección de precusores de sismos espaciales con diferentes probabilidades de cola de variación local de TEC: (a) 0.05, (b) 0.005 y (c) 0.0005.

Fuente: Elaboración propia.

2.8.2. Umbrales temporales de detección de precusores de sismos. Los umbrales temporales de detección de precusores de sismos se generan utilizando todas las variaciones TEC locales diarias $TEC_{\Delta;v;d}$ obtenidas de todas las estaciones de referencia $v = 1, \dots, N_s$, donde N_s es el número de todas las estaciones en TNGPN y d es el día elegido. Sea $TEC_{\Delta;d;sort}$ el vector ordenado ($N_s \times 1$) que contiene todas las variaciones locales de TEC para el día d :

$$TEC_{\Delta;d;sort} = f \left([TEC_{\Delta;1;d} \dots TEC_{\Delta;v;d} \dots TEC_{\Delta;N_s;d}]^T \right) \quad (19)$$

donde la función $f(\cdot)$ es una función de clasificación vectorial. La mediana de la variación local de TEC del día elegido d , Med_d se definirá de la siguiente manera:

$$Med_d = \frac{TEC_{\Delta;d;sort}^{N_s/2} + TEC_{\Delta;d;sort}^{N_s/2+1}}{2} \quad (20)$$

donde $TEC_{\Delta;d;sort}^{N_s/2}$ es el $N_s/2$ -ésimo elemento del vector $TEC_{\Delta;d;sort}$. De manera similar, la media M_d y la desviación estándar S_d de las variaciones diarias locales de TEC del día d se definirán de la siguiente manera:

$$M_d = \frac{1}{N_s} \sum_{v=1}^{N_s} TEC_{\Delta;v;d} \quad (21)$$

$$S_d = \sqrt{\frac{1}{N_s - 1} \sum_{v=1}^{N_s} (TEC_{\Delta;v;d} - M_d)^2} \quad (22)$$

Por lo tanto, el umbral temporal de detección de precursoros de sismos TET_d para el día d elegido será generado por:

$$TET_d(\beta) = Med_d + \beta S_d \quad (23)$$

donde el parámetro β es un parámetro de control que ajusta el rango esperado de variaciones diarias de TEC. Elegir un parámetro β alto corresponde a aumentar las variaciones diarias esperadas de TEC que resultan en menos detecciones de precursoros de sismos para el día d . Por otro lado, elegir un parámetro β alto también resultará en una menor cantidad de decisiones de anomalías de TEC falsas.

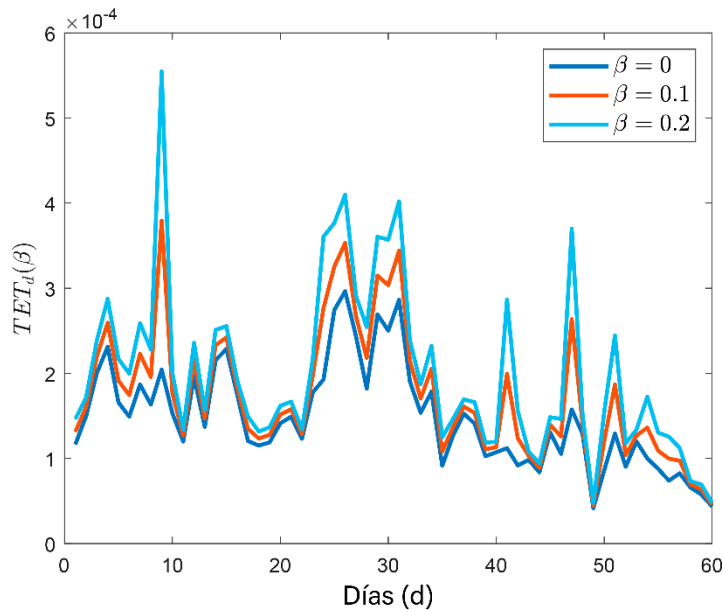


Figura N°8: Detección temporal de precursoros de sismos para los primeros 60 días de 2023 con diferentes parámetros β .

Fuente: Elaboración propia.

Diferentes umbrales temporales de detección de precursoros de sismos se ilustran en la figura N°8. Los umbrales se generan para los primeros 60 días de 2023 con tres parámetros β diferentes elegidos como: 0, 0.1 y 0.2. Como se ve en la figura N°8, los umbrales de detección de precursoros de sismos temporales generados muestran variaciones temporales mientras permanecen constantes en el espacio.

2.9. Detectores de precursoros de sismos. A diferencia de la investigación anterior, se obtienen detectores de precursoros de sismos espaciales y temporales que pueden operar tanto con señales de detección y confiabilidad de precursoros de sismos generados como con la generación de estas señales descritas en la sección 2.6 y la sección 2.7 [84]. Durante la operación, estos detectores requieren parámetros de detección y los umbrales de detección de precursoros de sismos que se generan con base en las técnicas detalladas en la sección 2.8. Dado que los detectores de precursoros de sismos espaciales y temporales pueden declarar

una decisión de precursor sobre la señal de detección con respecto a los parámetros de detección, la determinación de un conjunto apropiado de parámetros del detector tiene una importancia significativa en el rendimiento de la técnica de detección de precursores de sismos. A continuación, se presentará una técnica de clasificación de precursores de sismos basada en parámetros de detectores que hace uso del control de falsas alarmas de estos detectores.

2.9.1. Control de falsa alarma. Para demostrar el desempeño del control de falsas alarmas de los detectores propuestos, las falsas alarmas generadas por los detectores precursores de sismos propuestos se investigan en los días sin actividad sísmica. Para cada sismo presentado en la tabla II, se define una ventana de tiempo única para fines de etiquetado de días. Los días de actividad sísmica están etiquetados con respecto a estas ventanas de tiempo que comienzan desde los 9 días anteriores a un sismo y terminan con el día del sismo relacionado. Por lo tanto, 154 días de 2023 están etiquetados como días sísmicamente activos y pertenecen a la clase de actividad sísmica. Los 211 días restantes en 2023 están etiquetados como en la clase de actividad no sísmica. Las estadísticas de variación local de TEC que se obtienen de los días sin actividad sísmica se utilizan para generar umbrales espaciales de detección de precursores de sismos. La relación entre la probabilidad de cola de la variación local de TEC $P_{tail,TEC_{\Delta}}$ y la probabilidad de falsa alarma, P para los días sin actividad sísmica, se presenta en la figura N°9. Como se muestra en la figura N°9, se puede lograr un nivel deseado de P_{FA} obteniendo la probabilidad de cola correspondiente. Una vez que se elige una probabilidad de cola, los umbrales de detección de precursores de sismos espaciales se pueden obtener utilizando la distribución de Pareto correspondiente para cada estación de referencia. A diferencia de los umbrales de detección de precursores de sismos espaciales, los umbrales de detección de precursores de sismos temporales se generan utilizando las estadísticas que se obtienen de los días de actividad sísmica y sin actividad sísmica. La relación entre el parámetro β en (ec. 23) y P_{FA} para los días sin actividad sísmica también se presenta en la figura N°9. Como se muestra en la figura N°9, P se puede establecer en el nivel deseado ajustando el parámetro β . Una vez que se elige un parámetro β , los umbrales temporales de detección de precursores de sismos se generarán utilizando (ec. 23).

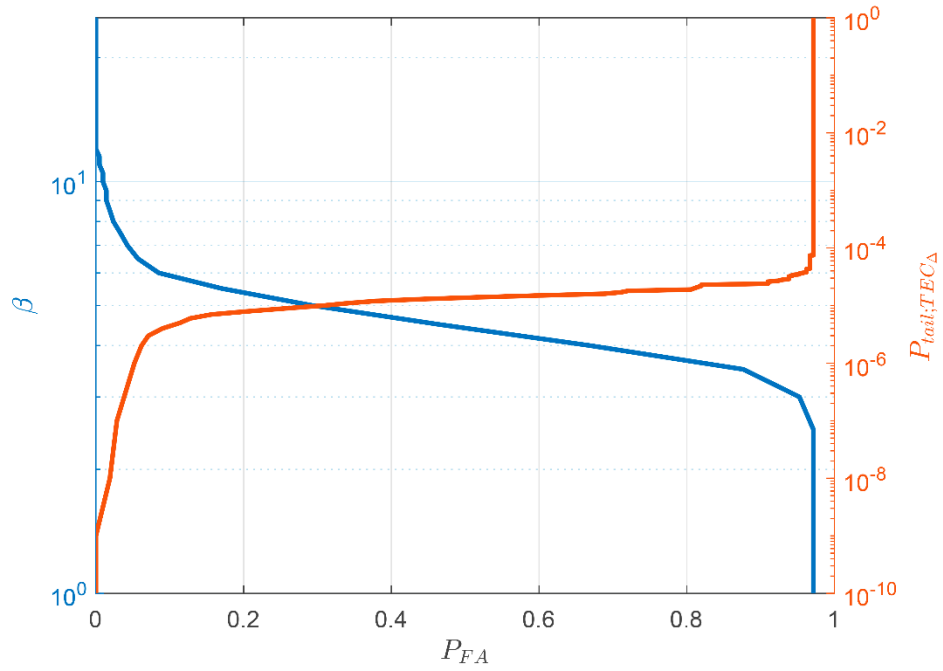


Figura N°9: (derecha) ajustando la P_{FA} del detector espacial ajustando la probabilidad de cola de variación local $P_{tail;TEC_{\Delta}}$ de TEC, (izquierda) ajustando la P_{FA} del detector temporal ajustando el parámetro β en la ec. 23.

Fuente: Elaboración propia.

Los detectores de precusores de sismos espaciales y temporales deben ajustar los parámetros $P_{tail;TEC_{\Delta}}$ y β para obtener umbrales de detección de precusores de sismos espaciales y temporales apropiados para un nivel deseado de P_{FA} . Si se superan los umbrales de detección de precusores de sismos espaciales o temporales en la señal de detección de precusores de sismos generada, el detector genera una decisión de precursor que podría estar relacionada con un próximo sismo. Las decisiones de precusores generados se clasificarán con respecto a sus grados de confiabilidad de TEC. La clasificación de decisiones valida las decisiones precursoras con un grado aceptable de confiabilidad de TEC y rechaza las decisiones precursoras restantes. En la siguiente sección se detalla una técnica de clasificación de decisiones.

2.9.2. Clasificación de las decisiones precursoras. Las señales de detección de precusores de sismos y las señales de confiabilidad de TEC se sincronizarán de manera que todas y cada una de las estaciones de referencia en el CENSIS tengan una variación local de TEC, $TEC_{\Delta;u;d}$ y puntaje de confiabilidad de TEC, $TEC_{\alpha;u;d}$ para un día d elegido y una estación de referencia u . La señal sincronizada representa cada medida de TEC en términos de su variación y confiabilidad. Por lo tanto, es posible evaluar cada variación de TEC con respecto a su confiabilidad. Para distinguir las variaciones locales de TEC con respecto a su confiabilidad, se utilizará un clasificador lineal. Para una estación de referencia u elegida y un día d , un clasificador lineal se definirá como:

$$TEC_{\Delta;u;d} \leq g_{LC} \times TEC_{\alpha;u;d} + C_{TEC_{\Delta}} \quad (24)$$

Donde g_{LC} es el gradiente del clasificador lineal y $C_{TEC_{\Delta}}$ es el punto de intersección del clasificador lineal y el eje de variación de TEC_{Δ} . Los parámetros g_{LC} y $C_{TEC_{\Delta}}$ pueden considerarse como el par de parámetros del detector para un detector de precusores de sismos. Para demostrar la ventaja particular del clasificador lineal propuesto, las decisiones precursoras generadas por un detector espacial para un $P_{FA} = 0.1517$ elegido y los parámetros del detector ($g_{LC} = -0.25, C_{TEC_{\Delta}} = 0.029$) se ilustran en la figura N°10. Como se muestra en la figura N°10, el clasificador lineal puede reducir la P_{FA} al rechazar 7 decisiones de precusores que resultan en falsas alarmas mientras preserva 20 decisiones de precusores que resultan en detecciones de precusores positivas verdaderas. A pesar de que la mejora del rendimiento observada puede estar limitada a este caso particular, no se garantiza una mejora del rendimiento con parámetros constantes del detector. Por lo tanto, los parámetros del detector de precusores de sismos espaciales y temporales requieren ajuste para un P_{FA} elegido. Tenga en cuenta que elegir un clasificador lineal también es ventajoso en términos de complejidad de cálculo. Para elegir un límite de decisiones cuadrática univariante, se debe resolver un problema de optimización tridimensional para la ecuación $TEC_{\Delta;u;d} \leq a \times TEC_{\alpha;u;d}^2 + b \times TEC_{\alpha;u;d} + c$. Sin embargo, el límite de decisión resultante se abrirá hacia arriba o hacia abajo solo porque la precisión de clasificación de este límite de decisión podría verse limitada. De manera similar, proponer un límite de decisión bivariado requiere resolver un problema de optimización de 6 dimensiones que dificulta el problema de optimización. Si los límites de decisión de los detectores temporales y espaciales se optimizan simultáneamente para límites de decisión cuadráticos univariados y bivariados, el problema de optimización resultante será de 6 o 12 dimensiones, respectivamente. En la siguiente sección se presentará una técnica de fusión de decisiones con ajuste de parámetros para los detectores precusores de sismos individuales.

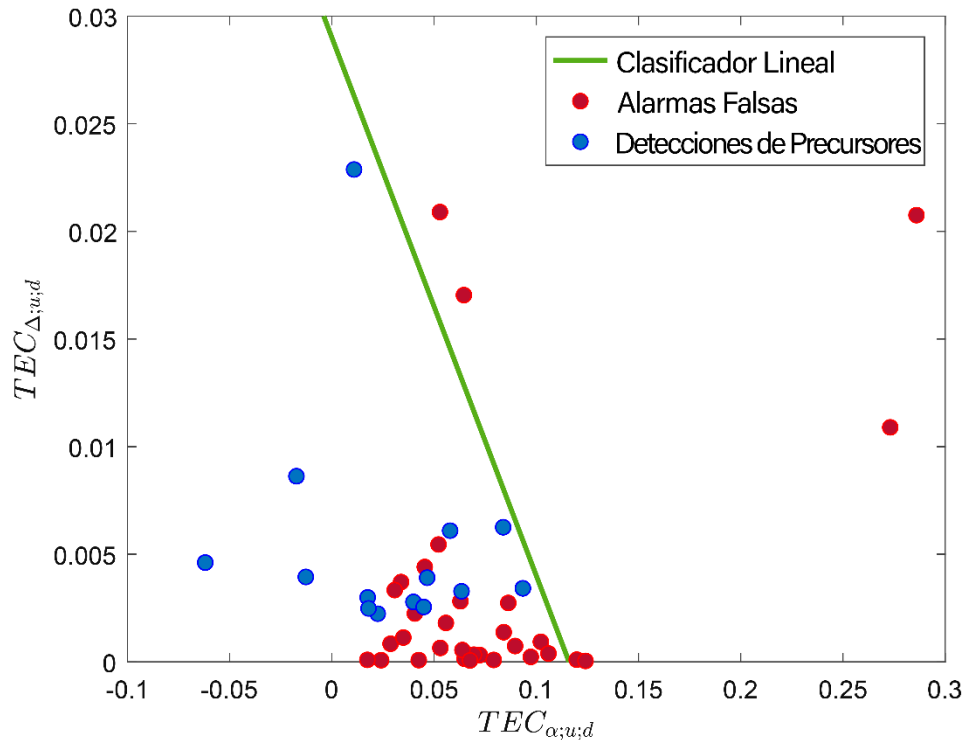


Figura N°10: Decisiones de precursores del detector de precursores de sismos espaciales separados por parámetros de clasificador lineal predefinidos.

Fuente: Elaboración propia.

2.9.3. Técnica de fusión de decisión. La técnica de fusión de decisiones propuesta se basa en generar una decisión de precursores fusionados que combina decisiones de precursores de dos o más detectores simultáneamente, de modo que la decisión de precursores fusionados sea más confiable que las decisiones de precursores individuales [85]. El rendimiento de detección de las decisiones de precursores fusionados depende en gran medida de la diversidad y el número de decisiones individuales. En este trabajo se analizan las detecciones de precursores de sismos generadas por disímiles detectores espaciales y temporales. Las decisiones de precursores que generan estos detectores se combinan para formar detecciones de precursores de sismos fusionados para lograr un mejor rendimiento en términos de menos falsas alarmas y tasas de detección más altas.

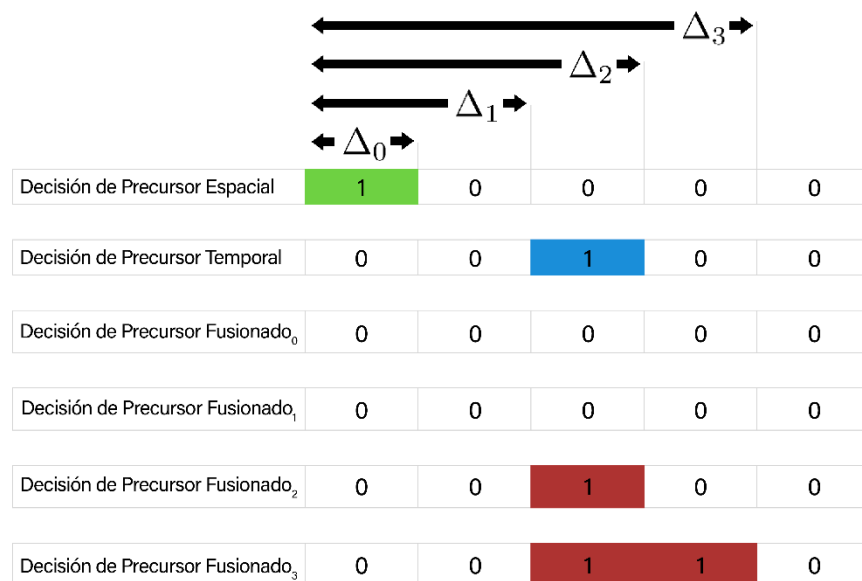


Figura N°11: Series de tiempo de decisión espacial y temporal para 5 días con series de tiempo de decisión fusionadas para diferentes parámetros Δ_d .
Fuente: Elaboración propia.

Las decisiones diarias sobre precursores de sismos que son generadas por detectores de precursores de sismos espaciales y temporales pueden considerarse decisiones difíciles. Estas decisiones difíciles se pueden agrupar en dos clases distintas, de modo que las próximas decisiones precursoras del sismo se marcan como 1 y los días restantes sin decisiones precursoras se marcan como 0. Por lo tanto, las decisiones diarias de precursores de sismos de un detector de precursores de sismos se pueden representar como una serie de tiempo que consta de solo 1 y 0. Estas decisiones precursoras de sismos se pueden fusionar mediante el uso de dos etapas de procesamiento. Primero, las decisiones precursoras que están marcadas como 1 en la serie de tiempo de decisión se repetirán durante Δ_d días para cada serie de tiempo de decisión. A partir de entonces, se aplica la operación AND lógica para las dos series de tiempo de decisión con eco diferentes para obtener decisiones de precursores de sismos $Fused_{\Delta_d}$. Para demostrar la técnica de fusión de decisión, se generan dos series de tiempo de decisión espacial y temporal diferentes durante 5 días y sus decisiones fusionadas correspondientes se presentan como en la figura N°11. Para un parámetro Δ_d elegido y una P_{FA} dada, cada detector de precursores de sismos generará decisiones duras sobre precursores de sismos sin considerar los parámetros del detector y la evaluación de confiabilidad. Como se discutió en la sección 2.10.2, la aplicación de la clasificación de decisión con los parámetros del detector es una mejora prometedora del rendimiento, mientras que requiere el ajuste de los parámetros del detector. En este

trabajo, los parámetros del detector tanto espaciales como temporales se ajustan mediante el uso de la técnica de optimización de enjambre de partículas (PSO) [86]. PSO es una técnica de optimización iterativa basada en la simulación del comportamiento de partículas en movimiento en un espacio de búsqueda de parámetros con respecto a restricciones predefinidas.

La técnica tiene como objetivo obtener el mejor conjunto de parámetros en el espacio de búsqueda representado por una partícula que comienza desde la primera iteración hasta la última. Los parámetros del detector espacial, $(g_{LC;S}, C_{TEC_{\Delta};S})$ y el detector temporal, $(g_{LC;T}, C_{TEC_{\Delta};T})$ forman un espacio de búsqueda de 4 dimensiones como $(g_{LC;S}, C_{TEC_{\Delta};S}, g_{LC;T}, C_{TEC_{\Delta};T})$ donde el rendimiento de cada partícula en PSO se evalúa sujeta a este espacio. El rendimiento de la i -ésima partícula x_i se representará como un punto de operación en un espacio de características operativas del receptor (ROC) que tiene la misma probabilidad de detección de precusores $P_{D;x_i;\Delta_d}$ y probabilidad de falsa alarma $P_{FA;x_i;\Delta_d}$ que las decisiones de precusores de sismos $Fused_{\Delta_d}$ lograrán. El pseudocódigo de la técnica PSO aplicada se presenta en el algoritmo 1.

Algoritmo 1 Pseudocódigo de optimización de enjambre de partículas

```
// Definición de parámetros //
 $N_{iter} \leftarrow$  Número máximo de iteraciones
 $N_{particles} \leftarrow$  Número de partículas
 $P_{FA;MIN} \leftarrow P_{FA}$  mínimo permisible para todas las partículas
 $\Delta_d \leftarrow$  Parámetro de eco de fusión de decisión

// Inicialización //
for  $i = (1 \text{ a } N_{particles})$  do
     $x_i \leftarrow$  inicializar la posición de la  $i$ -ésima partícula
     $v_i \leftarrow$  inicializar la velocidad de la  $i$ -ésima partícula
     $L_{best,i} \leftarrow$  establecer el mejor local inicial de la  $i$ -ésima partícula
    if  $L_{best,i} > G_{best}$  then
         $G_{best} = L_{best,i}$ 
    end if

end for

// Optimización operativa //
for  $t = (1 \text{ a } N_{iter})$  do
    for  $i = (1 \text{ a } N_{particles})$  do
         $v_i \leftarrow$  Actualizar la velocidad de la  $i$ -ésima partícula
         $x_i(t) = x_i(t - 1) + v_i \leftarrow$  Actualizar la posición de la  $i$ -ésima partícula
        if  $P_{D;x_i;\Delta_d} \leq \exp^{-\gamma(P_{FA;x_i;\Delta_d} - P_{FA;MIN})}$  then
            if  $P_{D;x_i;\Delta_d} > L_{best,i}$  then
                 $L_{best,i} = P_{D;x_i;\Delta_d}$ 
            end if
        end if
    end for
end for
```

```

if  $L_{best,i} > G_{best}$  then
     $G_{best} = L_{best,i}$ 
end if

end if

end if

end for

end for

```

Tenga en cuenta que el proceso de optimización de parámetros del detector se dividirá en 2 técnicas de optimización paralelas, de modo que los parámetros del detector espacial y temporal se optimizarán por separado mediante el uso de 2 técnicas de PSO separadas. Sin embargo, la convergencia a los mínimos globales del espacio de búsqueda de 4 dimensiones no está garantizada en tal proceso de optimización por etapas. A pesar de que las técnicas de optimización separadas pueden converger a sus respectivos mínimos globales para espacios de búsqueda de 2 dimensiones, la combinación de estos mínimos globales de 2 dimensiones en una solución de 4 dimensiones no garantiza la obtención de un mínimo global en un espacio de búsqueda de 4 dimensiones. Por lo tanto, los parámetros del detector se obtienen simultáneamente en un espacio de 4 dimensiones, en lugar de optimizar estos parámetros en dos espacios de búsqueda de 2 dimensiones separados. Dado que el rendimiento de cada partícula se representa en un espacio ROC bidimensional, pueden moverse libremente en todas las direcciones. Para obligar a las partículas a moverse a través de direcciones con menor P_{FA} y mayor P_D , el movimiento de las partículas se puede restringir a un cierto rango de P_{FA} mediante el uso de funciones de barrera espacial ROC. Estas funciones de barrera pueden limitar el rendimiento de las partículas en términos de menor probabilidad de falsas alarmas y se definirán de la siguiente manera:

$$P_{D;x_i;\Delta_d} \leq \exp^{-\gamma(P_{FA;x_i;\Delta_d} - P_{FA;MIN})} \quad (25)$$

donde $P_{FA;MIN}$ es la probabilidad mínima permisible de falsa alarma para todas las partículas y γ es un parámetro de transición de la función de barrera. Como se muestra en el algoritmo 1, la función de barrera es una restricción para el cálculo de la aptitud en la parte de optimización iterativa. La figura N°12 ilustra las funciones de barrera para diferentes valores de P. Como se muestra en la figura N°12, el espacio ROC está limitado por funciones de barrera y no se considerará el rendimiento de detección de las partículas que se mueven fuera de esta región. Dado que el desempeño de las

partículas está restringido en el sentido de generación de falsas alarmas, el desempeño de una partícula se evaluará sujeto únicamente a su probabilidad de detección.

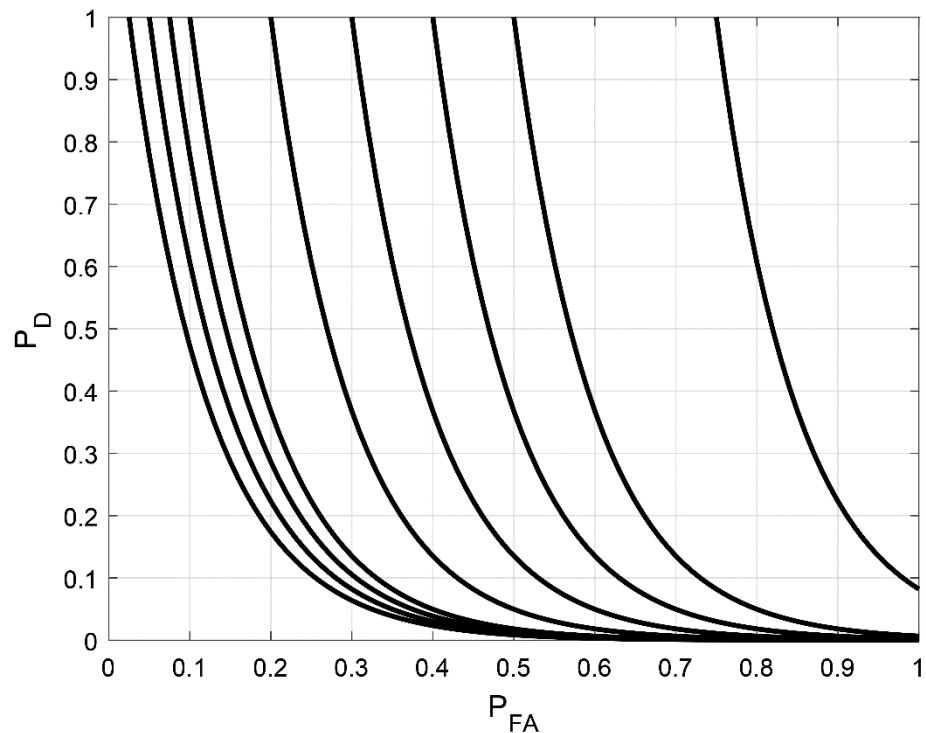


Figura N°12: Funciones de barrera ROC con $\gamma = 3$ y $P_{FA;MIN} = [0.025, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75]$ (de izquierda a derecha).

Fuente: Elaboración propia.

A continuación, se evaluará el rendimiento de la técnica de detección de precursores de sismos fusionada fusionando las decisiones de precursores de sismos de los detectores de precursores de sismos espaciales y temporales para los días de 2023. El diagrama de flujo de la técnica propuesta de detección de precursores de sismos basada en TEC se presenta en figura N°13.

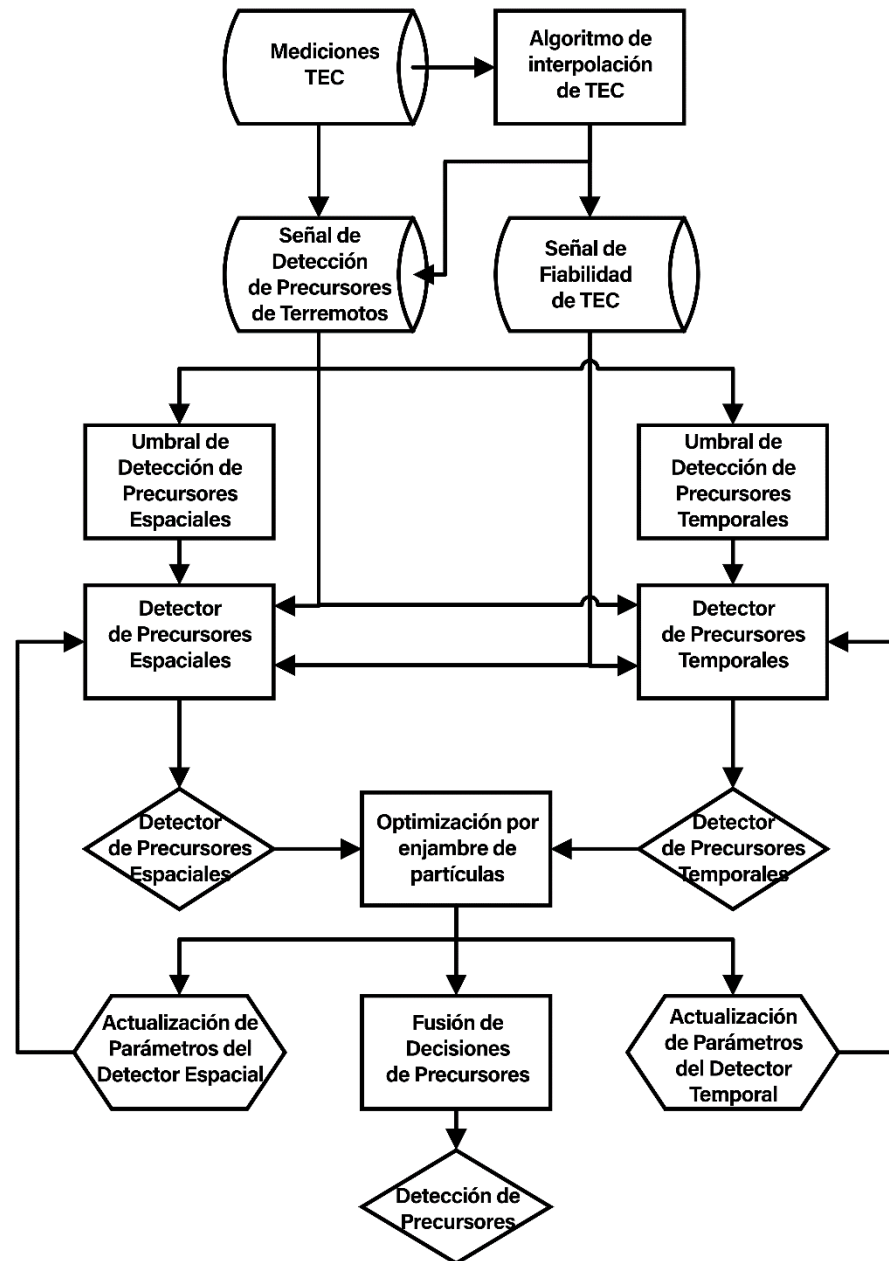


Figura N°13: Diagrama de flujo de la técnica propuesta de detección de precusores de sismos.

Fuente: Elaboración propia.

2.10. Rendimiento de la técnica propuesta de detección de sismos basada en el modelo propuesto. El rendimiento de detección de precusores de sismos de la técnica propuesta se investiga en los sismos que se presentan en la tabla II. Como se muestra en la tabla II, hay 21 sismos diferentes diarios que han tenido lugar alrededor de Ica en 2023 con magnitudes superiores a 4 en la escala de Richter. Todas las simulaciones de evaluación del desempeño se realizan con respecto a la actividad sísmica y las clases de actividad no sísmica que se generan a partir de la información temporal obtenida de la tabla II.

Tabla II

21 sismos que han tenido lugar alrededor de Ica en 2023.

Fecha (dd.mm.aaaa)	Hora (hh.mm)	Epicentro del sismo		Mw (Richter)	Z (km)
		Latitud	Longitud		
11.12.2023	13:58:23	15.25°S	75.61°O	4.3	27.81
15.12.2023	12:53:07	14.55°S	76.61°O	4	25.42
14.12.2023	10:55:50	14.45°S	76.49°O	4.1	8.99
14.12.2023	02:21:44	15.21°S	75.9°O	4.1	24.46
10.12.2023	07:14:12	14.41°S	76.57°O	4	13.42
10.12.2023	10:43:28	14.35°S	76.52°O	4.1	18.27
10.12.2023	07:14:12	14.41°S	76.57°O	4	6.97
10.12.2023	06:52:46	14.39°S	76.67°O	5.8	15.5
10.12.2023	06:40:30	14.39°S	76.51°O	5.3	5.00
10.12.2023	02:35:12	14.39°S	76.63°O	4.5	35.05
09.12.2023	22:38:19	14.36°S	76.58°O	4.6	7.18
06.12.2023	02:06:20	14.87°S	76.61°O	4.3	40.43
18.11.2023	23:16:06	15.19°S	75.8°O	5	5.00
18.11.2023	17:39:05	13.78°S	75.9°O	3.7	19.02
03.11.2023	10:05:35	14.07°S	75.84°O	4	17.27
15.10.2023	22:49:31	15.01°S	75.61°O	4	17.44
02.10.2023	11:24:02	14.1°S	76.07°O	4	21.61
25.09.2023	20:30:43	15.2°S	75.08°O	4	8.36
22.09.2023	16:17:51	14.61°S	75.52°O	4.2	21.47
14.08.2023	15:13:50	14.72°S	75.75°O	3.9	23.32
02.08.2023	02:55:31	14.76°S	75.83°O	4.3	8.00

Se pueden generar diferentes señales de detección de precursores de sismos considerando el hecho de que la cantidad de variaciones de TEC cambia con respecto a la hora del día. Los estudios informan que se observarán variaciones de TEC de hasta un 25 % durante el día y un 30 % durante la noche alrededor de las zonas de falla sísmica [87]. Por lo tanto, la ionosfera diurna y nocturna muestra una diferencia distinguible en términos de variaciones de TEC. Esta diferencia se puede atribuir a la fuerte radiación solar que provoca una alta variabilidad de TEC durante el día y una baja variabilidad de TEC durante la noche. Es posible disociar la variación local de TEC $TEC_{\Delta;u;d}$ en diferentes fuentes de variaciones de TEC como;

$$TEC_{\Delta;u;d} \simeq TEC_{\Delta;solar} + TEC_{\Delta;seismic} + TEC_{\Delta;other} \quad (26)$$

donde $TEC_{\Delta;solar}$ es la variación de TEC desencadenada por una fuerte radiación solar, $TEC_{\Delta;seismic}$ es la variación de TEC desencadenada por actividades sísmicas y $TEC_{\Delta;other}$ es la variación de TEC con diferentes fuentes. Debido a la fuerte radiación solar observada

durante el día, $TEC_{\Delta;solar}$ puede atenuar las variaciones de TEC provocadas por la actividad sísmica, $TEC_{\Delta;seismic}$. Por otro lado, podría ser posible observar variaciones de TEC provocadas por la actividad sísmica durante la noche, ya que el efecto perturbador de la fuerte radiación solar es menos efectivo durante la noche. Por lo tanto, el efecto de la fuerte radiación solar se eliminará de las mediciones de TEC para investigar las variaciones de TEC provocadas por la actividad sísmica, $TEC_{\Delta;seismic}$ durante la noche. Se aplicará una función de ventana de medición de TEC en una medición de TEC elegida para eliminar el efecto de la fuerte radiación solar de la medición. Una ventana de ejemplo se presenta en la figura N°14c y se aplica a las mediciones de TEC que se presentan en la figura N°2. Las mediciones de TEC en ventana se denominan mediciones de TEC nocturnas y se ilustran en la figura N°14.

En este trabajo, se generan las señales diurnas y nocturnas de detección de precursores de sismos para todas las mediciones diarias de TEC diurnas y nocturnas de todas las estaciones CENSIS en 2023. Las estimaciones de TEC para las mediciones diurnas y nocturnas de TEC se obtienen como se detalla en la sección 2.8. A partir de entonces, las variaciones locales diurnas y nocturnas de TEC se miden mediante la métrica SKLD y las señales de detección de precursores de sismos se generan como se explica en la sección 2.6.

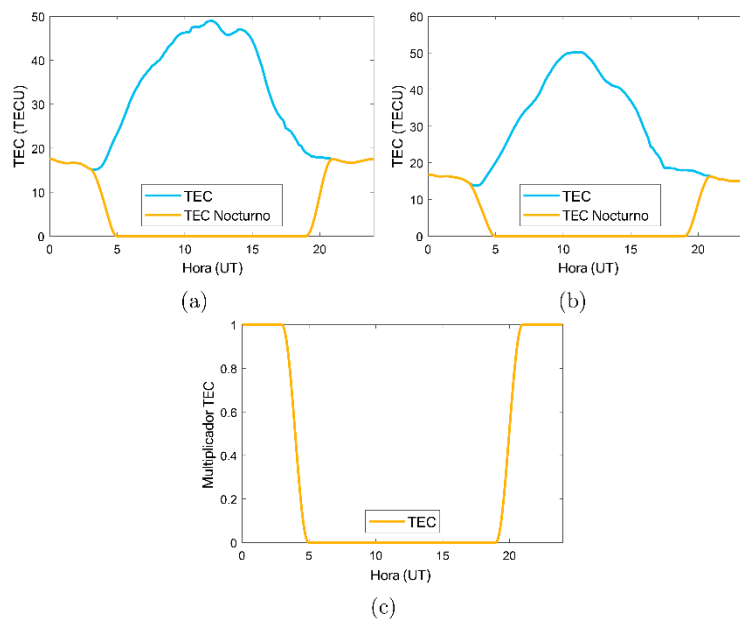


Figura N°14: Las medidas de TEC y las medidas de TEC nocturnas en la estación "date" en os fechas diferentes: (a) 16.10.2023 y (b) 17.10.2023 (c) función de ventana de medida de TEC.

Fuente: Elaboración propia.

Las señales de confiabilidad de sismos durante el día se generan utilizando todas las mediciones de TEC durante el día de todas las estaciones CENSIS para los días de 2023. El puntaje de confiabilidad de TEC de cada medición de TEC se examina mediante el valor

máximo de las variabilidades estimadas de la ionosfera local, y las señales de confiabilidad de sismos se generan como se discutió en la sección 2.8. Las señales de confiabilidad de sismos se sincronizan con las señales de detección de precursores de sismos para evaluar la confiabilidad de las variaciones locales de TEC durante el proceso de optimización en los detectores de precursores de sismos. Investigaciones anteriores han revelado que las señales de detección de precursores de sismos durante la noche con detectores de precursores de sismos temporales y las señales de detección de precursores de sismos durante el día con detectores de precursores de sismos espaciales logran un mejor rendimiento en términos de tasas de detección más altas con menos tasas de falsas alarmas [66],[84]. En este trabajo, el detector de precursores de sismos temporales se opera con señales de detección de precursores nocturnos con una señal de confiabilidad basada en el día y el detector de precursores de sismos espaciales se opera con señales de detección de precursores diurnos con la misma señal de confiabilidad diurna. También se generan umbrales apropiados de detección de precursores de sismos y las decisiones espaciales y temporales de precursores de sismos se generan en base a estos umbrales. Rendimiento de la técnica de detección de precursores de sismos propuesta evaluada en términos de decisiones de precursores fusionados que se generan al combinar las decisiones de precursores de sismos temporales nocturnos con las decisiones de precursores de sismos espaciales diurnos. La figura N°15 visualiza las curvas ROC para diferentes decisiones de precursores de sismos fusionados. Como se ve en la figura N°15, las decisiones de precursores de sismos fusionados detectan 14 precursores de sismos antes de 21 sismos y generan 5 falsas alarmas durante 211 días sin actividad sísmica.

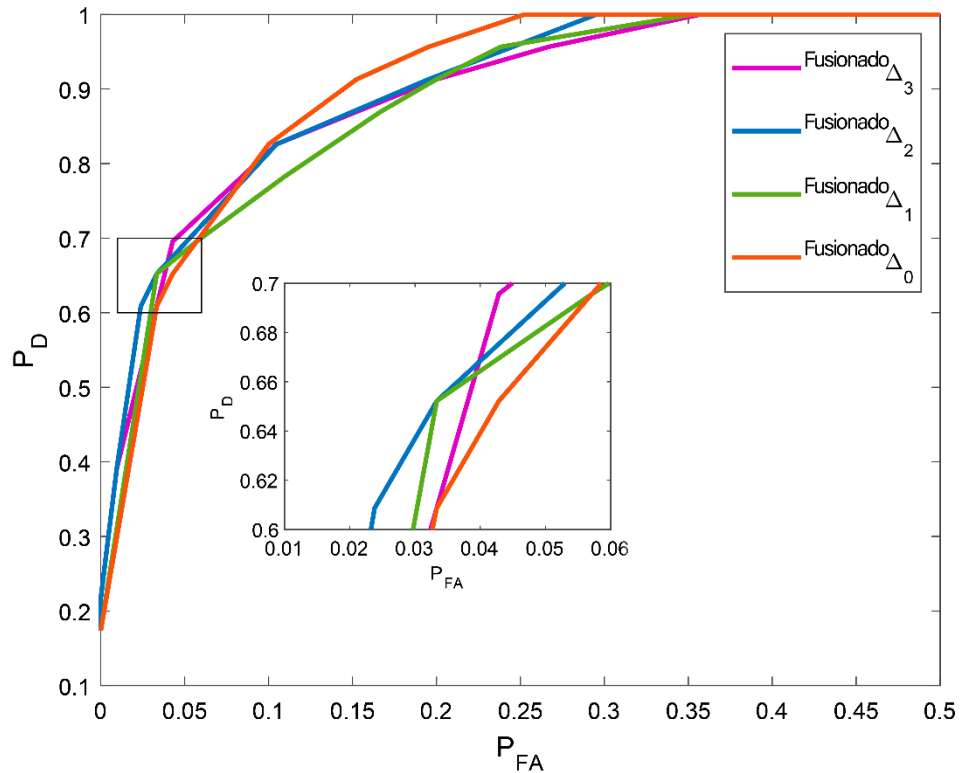


Figura N°15: (a) Curvas ROC para decisiones de precursores de sismos $Fused_{\Delta_0}$ (naranja), $Fused_{\Delta_1}$ (verde), $Fused_{\Delta_2}$ (azul) y $Fused_{\Delta_3}$ (rosa).

Fuente: Elaboración propia.

A pesar del rendimiento de detección significativo de las decisiones de precursores $Fused_{\Delta_2}$, la implementación casi en tiempo real de la técnica propuesta de detección de precursores de sismos podría ser limitada. Por lo tanto, también se debe evaluar el rendimiento casi en tiempo real de la propuesta. Las técnicas de validación cruzada se utilizan ampliamente para evaluar el rendimiento de las técnicas basadas en el aprendizaje automático en diferentes conjuntos de datos [88]. Entre las técnicas de validación cruzada, la técnica de validación cruzada k-fold es un buen candidato para evaluar el rendimiento de detección de precursores casi en tiempo real de la técnica de detección de precursores de sismos propuesta. La técnica de validación se basa en entrenar un modelo de aprendizaje automático por k veces y probar el modelo entrenado por k veces dividiendo el conjunto de datos en k pequeños conjuntos de datos, pliegues, mientras se reserva un pliegue para probar y usar los pliegues restantes para entrenar k veces. Cuando se considera la técnica de detección de precursores de sismos propuesta, los pliegues son las pequeñas particiones de señales de detección y confiabilidad de precursores de sismos con diferentes duraciones de tiempo. La fase de entrenamiento de la técnica de detección de precursores sísmicos propuesta puede considerarse como la obtención del mejor par de umbrales de detección de precursores con sus parámetros de detección apropiados. El rendimiento de la fase de prueba de la técnica de detección de precursores de sismos propuesta se puede evaluar mediante el uso de umbrales de detección

de precursores y parámetros del detector que se obtienen en la fase de entrenamiento en el conjunto de datos de prueba elegido. En este trabajo, se aplica una validación cruzada de 10 pliegues al dividir las señales de confiabilidad y detección de precursores de sismos en 10 pliegues diferentes. Cada pliegue contiene 36,5 días en promedio y el número de sismos y el número de días sin clase de actividad sísmica se presentan en la figura N°16 y también se tabulan en la primera y segunda columna de la tabla IV, respectivamente.

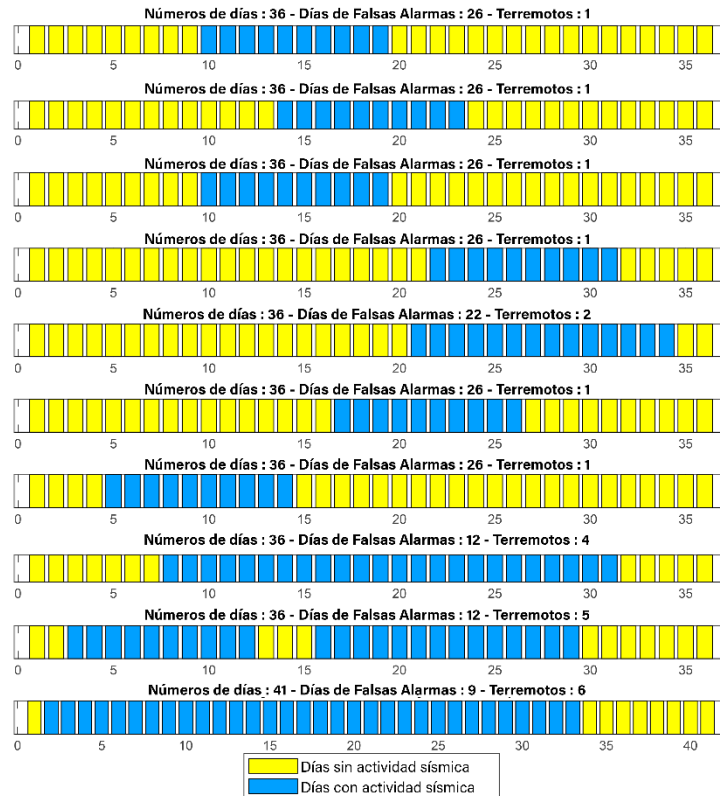


Figura N°16: 10 pliegues diferentes sin intersección generados para los días de 2023 con sus etiquetas de actividad sísmica y sin actividad sísmica. El número promedio de días de falsa alarma (días sin clase de actividad sísmica) es de alrededor de 21,1 y el número promedio de sismos es de alrededor de 2,3 por cada pliegue.

Durante la fase de entrenamiento de la técnica de detección de precursores sísmicos propuesta, la técnica propuesta podría memorizar señales de detección y confiabilidad de precursores sísmicos de tal manera que cualquier variación de estas señales de entrenamiento podría resultar en errores en las señales de prueba. Este fenómeno se conoce como sobreajuste en el aprendizaje automático [89]. Hay varios métodos para hacer frente a los inconvenientes del problema de sobreajuste. Una de las soluciones más utilizadas para este problema en el aprendizaje automático es la poda [90]. La poda reduce la complejidad del modelo de modo que se imponen implicaciones más generales en el conjunto de datos de entrenamiento. En este trabajo, se implementan dos técnicas de poda diferentes para evitar

que la técnica propuesta se sobreajuste en las señales de detección y confiabilidad de precursores de sismos. La primera técnica se basa en limitar el punto operativo ROC de entrenamiento de las decisiones precursoras de sismos $Fused_{\Delta_d}$ de modo que se acepten los puntos operativos con un cierto rango de falsa alarma, $P_{FA;range}$ y un rango de detección, $P_{D;range}$. Otra técnica es elegir múltiples puntos operativos en la curva ROC de entrenamiento para obtener la segunda y la tercera mejor curva ROC de entrenamiento. También se utiliza una función de fitness para elegir entre posibles puntos operativos en la curva ROC de entrenamiento [91]:

$$fitness(K, D, P_{FA}, P_D) = \frac{1}{K(P_{FA} - 0) + D(1 - P_D)} \quad (27)$$

donde K y D son parámetros de compensación para controlar las penalizaciones de error de punto ROC en términos de probabilidad de falsas alarmas, P_{FA} y probabilidad de detecciones, P_D , respectivamente. Para demostrar el rendimiento operativo casi en tiempo real de la técnica de detección de precursores de sismos propuesta, se implementa una técnica de validación cruzada de 10 veces para la detección de precursores de sismos y la señal de confiabilidad se pliega con el parámetro de fusión de decisión elegido como, $\Delta_d = 3$, los parámetros de función de barrera elegidos como $\gamma = 20$ y $P_{FA;MIN} = [0.01 : 0.01 : 0.1]$. Durante el entrenamiento, los puntos de operación ROC están restringidos a $P_{FA,Range} = [0, 0.07]$ y $P_{D,Range} = [0.4, 0.6]$. La tabla III ilustra el entrenamiento promedio (media de 10 datos de entrenamiento diferentes) y los rendimientos de prueba generales (rendimiento acumulativo de todos los pliegues de prueba) de los puntos de operación ROC elegidos con respecto a (ec. 27), limitados con respecto a $P_{FA,Range}$ y $P_{D,Range}$. Como se muestra en la tercera fila de la tabla III, los rendimientos de entrenamiento y prueba están razonablemente cerca uno del otro, de modo que el efecto del sobreajuste casi se elimina con las técnicas de poda. En la tabla IV también se presentan los rendimientos de entrenamiento y prueba individuales de las terceras mejores decisiones de precursores de sismos fusionados para 10 pliegues diferentes.

Tabla III

Entrenamiento promedio de validación cruzada de 10 veces y rendimiento de prueba general del primero, el segundo y el tercer mejor punto de operación de decisión del precursor del sismo $Fused_{\Delta_3}$.

	Rendimiento de entrenamiento promedio		Rendimiento general de la prueba	
	P_D	P_{FA}	P_D	P_{FA}
1ro mejor	12.1/20.7	5/189.9	10/23	12/211
2do mejor	12/20.7	5.7/189.9	8/23	9/211

3ro mejor	12/20.7	7.2/189.9	14/23	8/211
-----------	---------	-----------	-------	-------

Tabla IV

Número de sismos y días sin actividad sísmica en cada pliegue. Entrenamiento y pruebas de desempeño de las terceras mejores decisiones precursoras de sismos $Fused_{\Delta_3}$ para cada pliegue.

Información de pliegue		Rendimiento de entrenamiento		Prueba de rendimiento	
Número de sismos	Días sin actividad sísmica	P_{FA}	P_D	P_{FA}	P_D
1	26	6/185	12/22	3/26	0/1
1	26	8/185	13/22	0/26	1/1
1	26	6/185	13/22	1/26	0/1
1	26	7/185	13/22	0/26	1/1
2	22	6/189	12/21	0/22	2/2
1	26	8/185	13/22	2/26	0/1
1	26	9/185	13/22	0/26	1/1
4	12	7/199	11/19	0/12	3/4
5	12	7/199	10/18	2/12	1/5
6	9	8/202	10/17	0/9	5/6

Durante la fase de entrenamiento de la técnica propuesta de detección de precursores sísmicos, la probabilidad promedio alcanzada de falsa alarma y la probabilidad promedio de detección de precursores son 0.0380 y 0.5794 con desviaciones estándar de 0.0057 y 0.0167, respectivamente. Dado que estas desviaciones estándar son menores que los valores medios, la técnica de detección de precursores de sismos propuesta deberá elegir un punto de operación similar cada vez durante la fase de entrenamiento. Además, la técnica propuesta de detección de precursores sísmicos alcanza una probabilidad media de falsa alarma de 0,0397 y una probabilidad media de detección de precursores de 0,5783 durante la fase de prueba. A diferencia de la fase de entrenamiento, la desviación estándar de la probabilidad de detección de precursores es 0.4657 debido a la distribución no uniforme del número de sismos para cada pliegue. Como se muestra en la sexta columna de la tabla IV, no hay detecciones de precursores para tres pliegues con solo un sismo, se calcula que la P_D de la detección de precursores es alta. La desviación estándar de la probabilidad de falsa alarma se logra como 0.0603 durante la fase de prueba que es comparablemente cercana a la desviación estándar de la probabilidad de falsa alarma que se logra durante la fase de entrenamiento.

III. RESULTADOS

Para determinar la importancia estadística de la detección de precursores en tiempo real de EQ, se llevó a cabo una investigación para obtener el rendimiento de detección de precursores y generación de falsas alarmas de una técnica de detección de precursores de EQ basada en TEC ionosférica [66, 67]. Los resultados de estas investigaciones ilustran el hecho de que existe una interacción entre las anomalías TEC ionosféricas y las actividades sísmicas intensas. En este estudio, se propone una técnica de detección de precursores de EQ ionosférica basada en aprendizaje automático, EQ-PD, para detectar efectos anómalos de esta interacción en TEC ionosférica. EQ-PD tiene las siguientes contribuciones clave a la literatura de detección de precursores de EQ. Una de las primeras contribuciones de la propuesta es que EQ-PD genera detecciones de precursores con técnicas de aprendizaje automático. La técnica EQ-PD propuesta se entrena, valida y prueba en datos TEC ionosféricos para generar detecciones de precursores de EQ. Por lo tanto, es posible investigar su desempeño de detección de precursores de una manera robusta y confiable. Otra contribución de la propuesta es que EQ-PD funcionará con cualquier red GNSS elegida que registrará datos TEC diarios en sus estaciones de referencia. EQ-PD requiere índices geomagnéticos diarios y datos TEC ionosféricos diarios para generar decisiones precursoras de EQ diarias. La contribución final a la literatura es que EQ-PD se definirá como una de las pocas técnicas de refuerzo en la literatura. EQ-PD genera múltiples decisiones precursoras de EQ y las combina para obtener decisiones precursoras de EQ con mayor robustez y rendimiento. En este trabajo, los datos de TEC obtenidos de estaciones de referencia ubicadas alrededor de Italia con una duración de casi 3 años se procesan para evaluar el rendimiento de detección de precursores de EQ de la técnica propuesta. Los datos TEC obtenidos se dividen y se generan conjuntos de datos de entrenamiento, validación y prueba que no se superponen. La evaluación del rendimiento de detección de precursores de EQ del EQ-PD se realiza solo en los conjuntos de datos de prueba. este trabajo consta de múltiples etapas de procesamiento secuencial. Los datos de TEC procesados por la técnica se presentan en la sección 3. La generación de la señal de detección de anomalías, los umbrales de detección de anomalías, las decisiones de anomalías, la señal de detección de precursores y las detecciones de precursores de EQ se detallan en la sección 1. El rendimiento de detección propuesto se investiga en la sección 3.6. Finalmente, el trabajo concluye con comentarios finales y direcciones de investigación futuras en la sección 4.

3.1. Recopilación y procesamiento de datos. Las redes GNSS constan de múltiples estaciones de referencia, de modo que la ubicación de estas estaciones de referencia se estima y actualiza con los retrasos de fase ionosférica de las señales del satélite [92]. Estos retrasos de fase no solo se utilizan para estimar las posiciones de la estación de referencia, sino que también se utilizan para estimar el contenido total de electrones inclinados (STEC). STEC se definirá

como el número total de electrones encapsulados por un cilindro con una sección transversal de 1 m^2 y se extiende entre el satélite y la estación de referencia. También se generará una estimación del contenido total de electrones verticales (VTEC) o (TEC) combinando todas las estimaciones de STEC que se obtienen de un satélite diferente [93]. La unidad de medida del TEC estimado se definirá como TECU y 1 TECU corresponde a 10^{16} electrones/ m^2 .

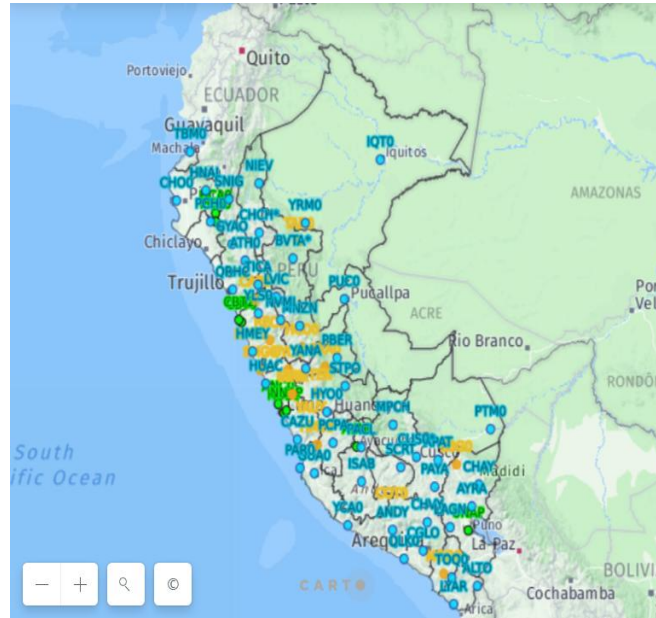


Figura N°17: Las posiciones de las estaciones de referencia CENSIS en Perú.

Fuente: Elaboración propia.

Los datos TEC de una estación de referencia se controlarán con una separación temporal de 2,5 minutos. Los datos TEC de múltiples estaciones de referencia o una red GNSS también se monitorearán generando mapas TEC 2-D de un área elegida con la misma separación de tiempo de 2,5 minutos [94, 95, 70]. Por lo tanto, es posible monitorear de manera constante y confiable los datos TEC de una región elegida, si ya se ha implementado una red GNSS en la región elegida. En este trabajo, se elige un subconjunto de estaciones de referencia de la red GNSS permanente (EPN) de EUREF para fines de recopilación y procesamiento de datos TEC [96]. Las estaciones de referencia elegidas se distribuyen por Italia y se presentan como círculos naranjas en la figura N°17. EPN es una red GNSS y es operada por la Subcomisión del Marco de Referencia Regional para Europa de la Asociación Internacional de Geodesia (EUREF). Más de 100 organismos y universidades contribuyen al funcionamiento de la EPN con lineamientos y estándares bien definidos [96]. La EPN está formada por más de 300 Estaciones de Referencia GNSS de Funcionamiento Continuo (CORS) desplegadas por toda Europa. En este trabajo, los datos TEC de las estaciones de referencia seleccionadas de EPN son procesados por los servicios IONOLAB-TEC e IONOLAB-BIAS para obtener datos TEC más confiables y refinados. IONOLAB-TEC proporciona estimaciones GPS-TEC fiables y precisas para una estación de referencia GPS seleccionada [93, 95, 97, 98].

IONOLAB-TEC combina diferentes datos STEC obtenidos de todos los satélites GPS que se encuentran por encima del límite de horizonte de la estación de referencia elegida. A partir de entonces, se estima una TEC vertical para cada satélite GPS y se combina para formar una estimación TEC final de modo que se reduzcan los efectos de trayectos múltiples de las posiciones de los satélites GPS [99]. Se implementa el método IONOLAB-BIAS para el proceso de estimación de STEC. El método incluye sesgos de código diferencial (DCB) para la estimación de STEC y reduce el ruido en las estimaciones [98]. En este trabajo, las variables temporales y espaciales de las mediciones de TEC se visualizan mediante las estaciones de referencia elegidas con nombres en clave "AQUI" en la región de L'Aquila, "VEN1" en la región de Venecia, "TORI" en la región de Turín y "MATE" en la región de Matera de Italia. Las ubicaciones geodésicas de estas estaciones de referencia elegidas se presentan en la tabla V y se muestran en la figura N°18. Como se muestra en la figura N°18, las estaciones de referencia "AQUI", "VEN1" y "TORI" están ubicadas en regiones sísmicamente activas y la estación de referencia "MATE" está ubicada en una región sísmicamente inactiva en comparación con otras ubicaciones de estaciones de referencia. Por lo tanto, las estaciones de referencia "AQUI", "VEN1" y "TORI" se eligen para visualizar las variabilidades espaciales de las mediciones de TEC en regiones sísmicamente activas. La estación de referencia "MATE" se elige para visualizar las variabilidades temporales de las mediciones de TEC en una región sísmicamente inactiva.

Tabla V

Ubicaciones geodésicas y regiones de las estaciones de referencia elegidas: "AQUI", "VEN1", "TORI" y "MATE".

Nombre del código de la estación	Ubicación		País	Ciudad
	Latitud	Longitud		
AQUI	42.37°S	13.35°O	Perú	Ica
VEN1	45.43°S	12.35°O	Perú	Ica
TORI	45.06°S	7.66°O	Perú	Ica
MATE	40.65°S	16.7°O	Perú	Ica

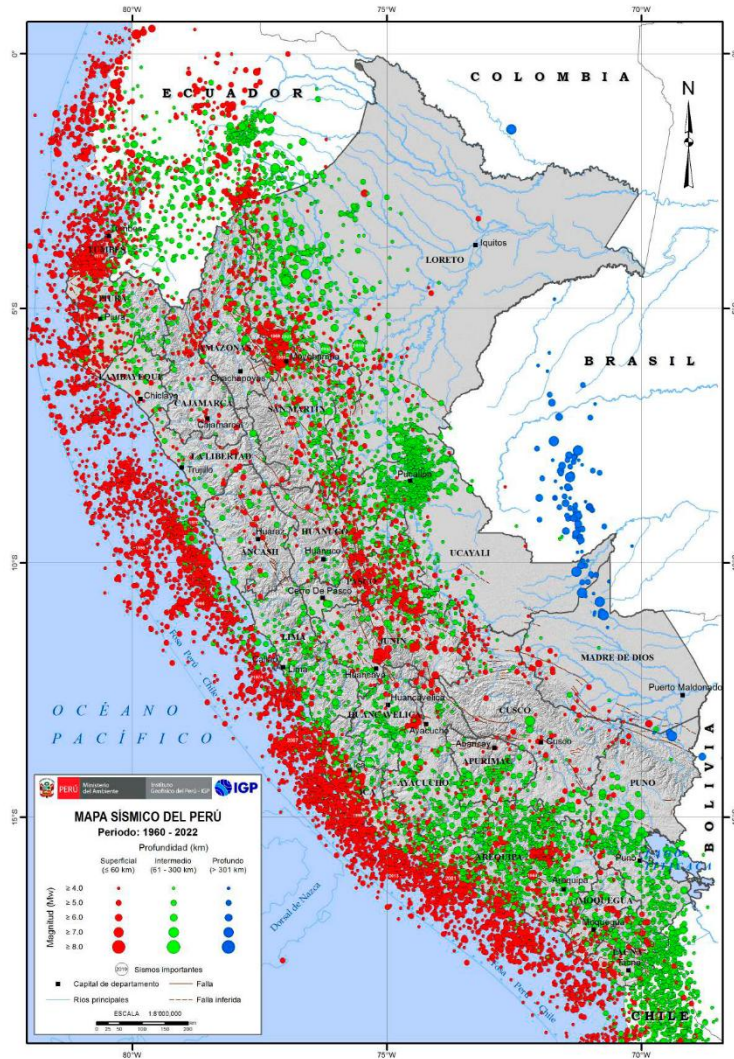


Figura N°18: Ubicaciones geodésicas de las estaciones de referencia AQUI, VEN1, TORI y MATE con epicentros EQ registrados durante días entre agosto y diciembre del 2023.

Fuente: Elaboración propia.

Las actividades geomagnéticas perturban múltiples regiones de la ionosfera [100, 101]. Dado que las actividades geomagnéticas causan perturbaciones o anomalías ionosféricas, las anomalías relacionadas con la actividad geomagnética deben desvincularse de las otras anomalías ionosféricas [102, 103, 104]. La actividad geomagnética y las perturbaciones ionosféricas relacionadas con las tormentas geomagnéticas se clasificarán mediante el seguimiento de los índices geomagnéticos. Los índices geomagnéticos ampliamente utilizados se enumerarán como: índice de electrochorro auroral (AE), índice de actividad magnética planetaria (A_p), índice de tiempo de tormenta perturbadora (Dst), índice de tormenta geomagnética global (K_p) y número de manchas solares (SSN) [94, 66, 99]. Se podrá acceder a los promedios por hora, diarios, de 27 días y anuales de AE, A_p, Dst, K_p y SSN a través del servicio OMNIWeb del Centro de Vuelo Espacial

Goddard de la NASA [106]. Entre estos índices geomagnéticos, los índices Dst y K_p se usan para identificar tormentas geomagnéticas, SSN se usa para estimar la cantidad de radiación solar, AE se usa para medir la actividad magnética de la zona auroral y los índices A_p se usan para medir la actividad geomagnética promedio.

Los datos de TEC se visualizan en la ec. 31. Como se muestra en la ec. 31, dos mediciones diarias consecutivas de TEC de una estación de referencia muestran una diferencia temporal observable. Los índices geomagnéticos diarios de estos días también se tabulan en la tabla VI.

Tabla VI
Para dos fechas diferentes: 11 de noviembre de 2023 y 12 de noviembre de 2023, diarios AE , A_p , Dst , K_p y SSN .

Fecha (día mes año)	AE	A_p	Dst (nT)	K_p (nT)	SSN (nT)
11 de noviembre de 2023	268	8	-19	20	85
12 de noviembre de 2023	176	12	-11	27	63

Se accede a los datos relacionados con EQ a través del catálogo completo de sismos del Sistema Nacional Sísmico Avanzado (ANSS ComCat) [101]. Los datos relacionados con EQ accedidos: número de EQ y fecha de EQ diferentes diariamente (primera columna), hora de EQ (segunda columna), epicentro de EQ (tercera y cuarta columna) y tipo de magnitud de EQ (quinta columna) se presentan en la tabla XI.

3.2. Técnica de detección de precursores de sismos basada en aprendizaje automático: EQ-PD. Las redes GNSS recopilan de manera constante y confiable datos relacionados con la ionosfera, como TEC, foF2 y hmF2. Los datos recopilados muestran anomalías antes de las actividades sísmicas fuertes, por lo que estas anomalías se definirán como precursores de EQ si se detectan antes de los EQ [108]. En este trabajo, se propone la técnica de detección de precursores EQ basada en TEC ionosférica y aprendizaje automático, EQ-PD. El propuesto deberá detectar precursores de EQ diariamente utilizando los datos TEC diarios que se recopilan de estaciones de referencia distribuidas espacialmente o de una red GNSS. los datos recopilados de TEC exhiben variaciones de TEC tanto espaciales como temporales. Por lo tanto, la detección de anomalías TEC desencadenadas por actividad sísmica o precursores de EQ requiere una técnica de detección de precursores que aprenda y detecte variaciones espaciales y temporales de TEC en datos TEC, simultáneamente.

Este trabajo está organizado de la siguiente manera. En la sección 3.2.1, se presenta el resultado del rendimiento de la estimación de TEC de una técnica de interpolación espacio-temporal de TEC. En la sección 3.2.2, se generan señales de detección de anomalías que

varían espacial y temporalmente. Luego, en la sección 3.2.3 se generan espaciales y temporales. En la sección 3.2.4, las decisiones de anomalías se obtienen al establecer un umbral de señales de detección de anomalías y las decisiones de anomalías obtenidas se combinan para formar una señal de detección de precursores de ecualización. Los clasificadores de la Máquina de vectores de soporte (SVM) se entrenan, validan y prueban en las señales de detección de precursores de EQ y, finalmente, se generan detecciones de precursores de EQ como se describe en la sección 3.2.5. El diagrama de flujo de la técnica EQ-PD propuesta se visualiza en la figura N°19.

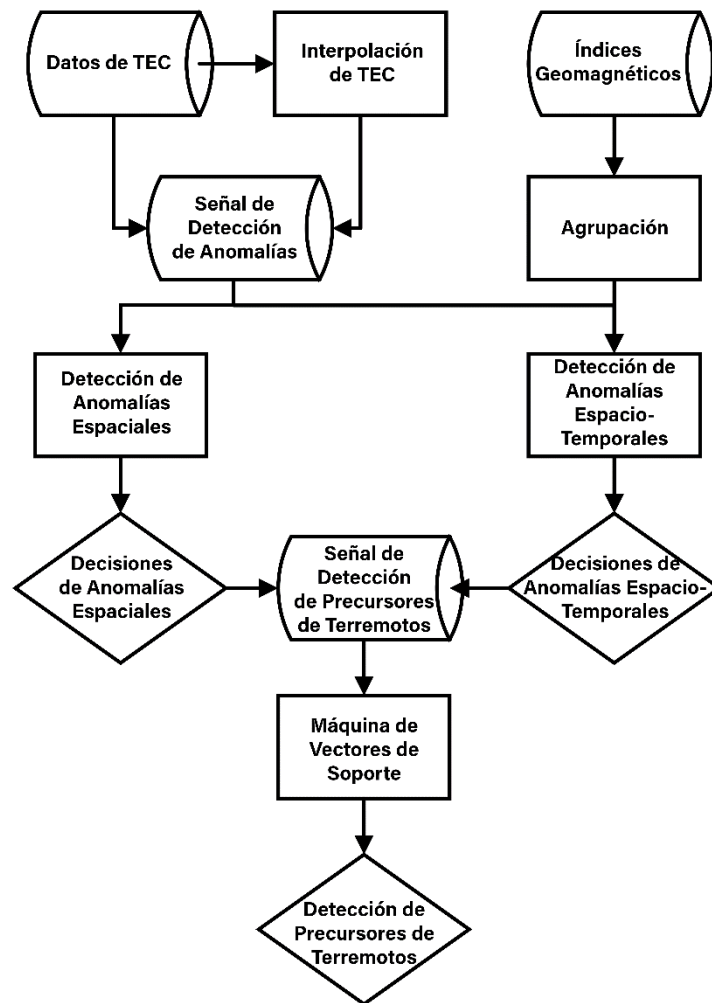


Figura N°19: Diagrama de flujo de la técnica EQ-PD propuesta. Las señales de detección se generan en función de los datos TEC regionales y los umbrales de detección se eligen de forma adaptativa en función de los índices geomagnéticos.

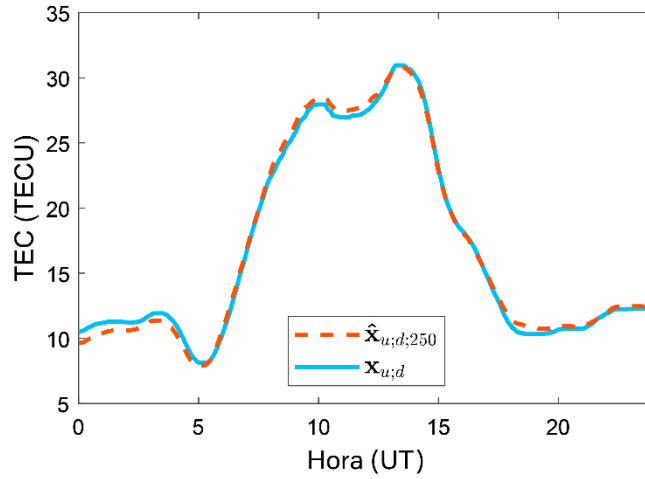
Fuente: Elaboración propia.

3.2.1. Interpolación TEC espacio-temporal. Como se discutió en la sección 2.6, es posible obtener una estimación de TEC, $\hat{\mathbf{x}}_{u;d;R_r}$ para la medición de TEC de $\mathbf{x}_{u;d}$ usando el vector de peso $\underline{\alpha}_{u;d_n;R_r}$ de la estación de referencia vecina en (ec. 2). Similitud entre la estimación de TEC obtenida, $\hat{\mathbf{x}}_{u;d;R_r}$ y la medición de TEC se medirá $\mathbf{x}_{u;d}$ para

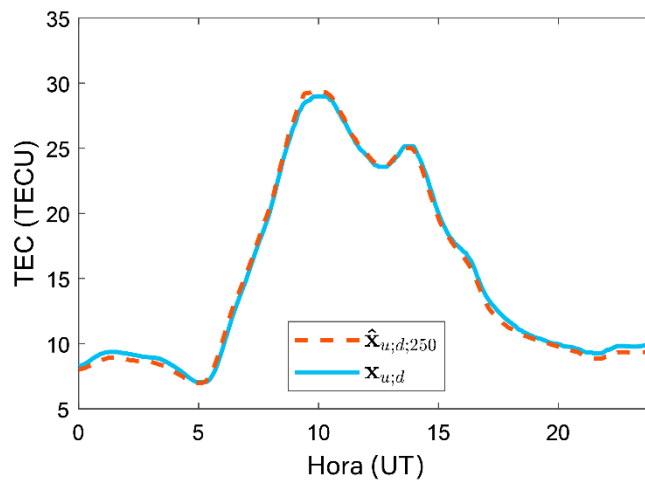
generar la variación local de TEC para el día d y la estación de referencia u . Estimaciones de TEC, $\hat{\mathbf{x}}_{u;d;R_r}$ y mediciones de TEC, $\mathbf{x}_{u;d}$ de la estación de referencia "MATE" para las fechas elegidas en la tabla VI se visualizan en la figura N°20. Como se ilustra en la figura N°20, la cantidad de variaciones de TEC muestra diferencias distinguibles durante el día y durante la noche. Las fuentes de variabilidad de TEC se separarán en diferentes categorías, tales como:

$$TECV_{u;d} \simeq TECV_{geo;u;d} + TECV_{seis;u;d} + TECV_{oth;u;d} \quad (28)$$

Donde $TECV_{u;d}$ es la variabilidad de TEC de una estación de referencia u para el día d . Para la estación de referencia y el tiempo seleccionados, $TECV_{geo;u;d}$ se define como la fuente de variabilidad de TEC generada por actividades geomagnéticas. De manera similar, $TECV_{seis;u;d}$ se define como la fuente de variabilidad de TEC generada por actividades sísmicas en la corteza terrestre en curso. Finalmente, $TECV_{oth;u;d}$ es la fuente de variabilidad de TEC generada por otras fuentes de interacciones entre la ionosfera y el clima espacial. Debido a la fuerte actividad solar durante el día, $TECV_{geo;u;d}$ muestra mayores variabilidades de TEC durante el día en comparación con la noche. Sin embargo, la cantidad de variabilidad de TEC causada por $TECV_{seis;u;d}$ depende de la magnitud de la actividad sísmica y es independiente del momento del día. Por lo tanto, la detección de variabilidades de TEC nocturnas causadas por $TECV_{geo;u;d}$ será más fácil en comparación con las variabilidades de TEC diurnas debido a la ausencia de $TECV_{geo;u;d}$ durante la noche. La ventana de variación de TEC que se presenta en la figura N°14c se aplicará a las mediciones de TEC, $\mathbf{x}_{u;d}$ para generar mediciones de TEC nocturnas, $\mathbf{x}_{u;d}^{\text{night}}$. De manera similar, las estimaciones nocturnas de TEC, $\hat{\mathbf{x}}_{u;d;R_r}^{\text{night}}$ también se generarán con la técnica de estimación espacio-temporal como se presenta en la sección 2.6. En la figura N°21 se visualizan las estimaciones nocturnas de TEC, medida de X y TEC, Z de la estación de referencia "MATE" para las fechas elegidas en la tabla VI.



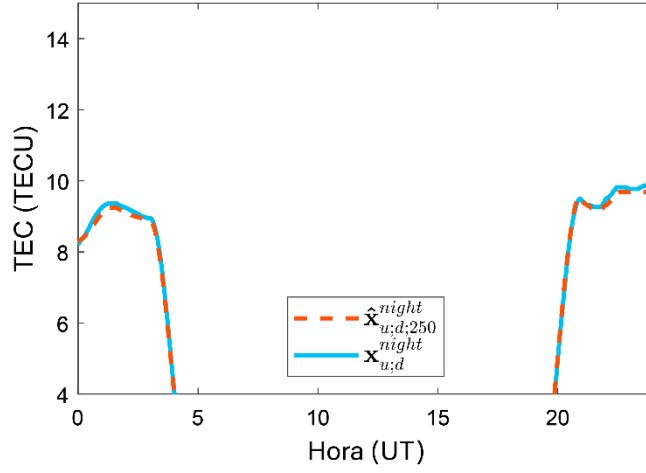
(a)



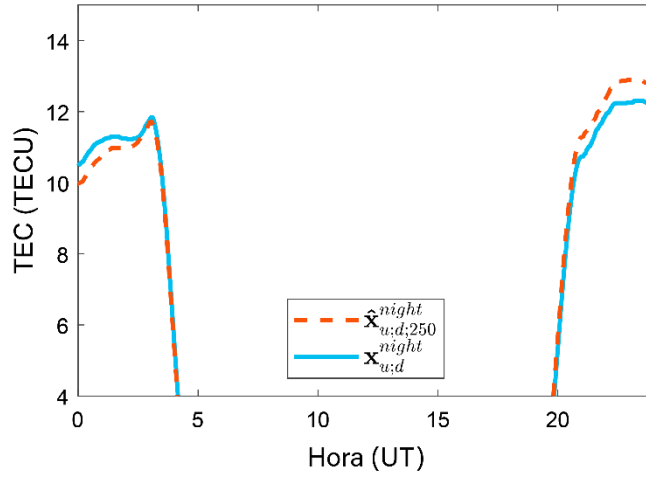
(b)

Figura N°20: Estimaciones diarias de TEC $\hat{\mathbf{x}}_{u;d;R_T}$ y mediciones de TEC $\mathbf{x}_{u;d}$ para la estación de referencia "MATE" para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023.

Fuente: Elaboración propia.



(a)



(b)

Figura N°21: Estimaciones TEC nocturnas \hat{x} y medidas TEC X para la estación de referencia "MATE" para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023.

Fuente: Elaboración propia.

3.2.2. Generación de señales de detección de anomalías diurnas y nocturnas. La detección de anomalías relacionadas con la TEC ionosférica requiere la generación de señales de detección de anomalías. Las señales de detección de anomalías se generarán midiendo la distancia entre las mediciones de TEC de todas las estaciones de referencia con sus estimaciones de TEC. La divergencia simétrica de Kullback-Leible (SKLD) se utilizará como una métrica apropiada para medir la distancia métrica entre las estimaciones $\hat{x}_{u,d;R_r}$ de TEC y las medidas $x_{u,d}$ de TEC, como también se detalla en la sección 2.7 [74, 75, 109, 110]. Para una estación de referencia u y el día d , la variación de TEC de un día completo basada en SKLD, TEC_{Δ}^{w-day} se definirá de la siguiente manera.

$$TEC_{\Delta}^{w-day} = KLD(\hat{\mathbf{P}}_{u,d;R_r} | \mathbf{P}_{u,d}) + KLD(\mathbf{P}_{u,d} | \hat{\mathbf{P}}_{u,d;R_r}) \quad (29)$$

Donde KLD es la abreviación de Knnullback-Leibler Divergence y se define como:

$$KLD(\hat{\mathbf{P}}_{u;d;R_r} | \mathbf{P}_{u;d}) = \left(\sum_{n=1}^{N_{u;d}} \hat{\mathbf{P}}_{u;d;R_r}(n) \log \left(\frac{\hat{\mathbf{P}}_{u;d;R_r}(n)}{\mathbf{P}_{u;d}(n)} \right) \right) \quad (30)$$

$$KLD(\mathbf{P}_{u;d} | \hat{\mathbf{P}}_{u;d;R_r}) = \left(\sum_{n=1}^{N_{u;d}} P_{u;d}(n) \log \left(\frac{P_{u;d}(n)}{\hat{\mathbf{P}}_{u;d;R_r}(n)} \right) \right) \quad (31)$$

donde $\mathbf{P}_{u;d}$ ($(N_{u;d} \times 1)$ columna de vectores) es el vector de medición de TEC normalizado diario y $\hat{\mathbf{P}}_{u;d;R_r}$ ($(N_{u;d} \times 1)$ columna de vectores) es el vector de estimación de TEC normalizado diario. $\mathbf{P}_{u;d}$ y $\hat{\mathbf{P}}_{u;d;R_r}$ se definirán como sigue:

$$\mathbf{P}_{u;d} = \frac{\mathbf{x}_{u;d}}{\|\mathbf{x}_{u;d}\|_1} = \mathbf{x}_{u;d} \left(\sum_{n=1}^{N_{u;d}} x_{u;d}(n) \right)^{-1} \quad (32)$$

$$\hat{\mathbf{P}}_{u;d;R_r} = \frac{\hat{\mathbf{x}}_{u;d;R_r}}{\|\hat{\mathbf{x}}_{u;d;R_r}\|_1} = \hat{\mathbf{x}}_{u;d;R_r} \left(\sum_{n=1}^{N_{u;d}} \hat{x}_{u;d;R_r}(n) \right)^{-1} \quad (33)$$

observe que los vectores $\mathbf{P}_{u;d}$ y $\hat{\mathbf{P}}_{u;d;R_r}$ no tienen unidades ya que TECU está normalizado en (32) y (33). De manera similar, la variación nocturna de TEC basada en SKLD, $TEC_{\Delta}^{\text{night}}$ se definirá de la siguiente manera:

$$TEC_{\Delta}^{\text{night}} = KLD(\hat{\mathbf{P}}_{u;d;R_r}^{\text{night}} | \mathbf{P}_{u;d}^{\text{night}}) + KLD(\mathbf{P}_{u;d}^{\text{night}} | \hat{\mathbf{P}}_{u;d;R_r}^{\text{night}}) \quad (34)$$

donde $\mathbf{P}_{u;d}^{\text{night}}$ es el vector de medición de TEC normalizado durante la noche y $\hat{\mathbf{P}}_{u;d;R_r}^{\text{night}}$ es el vector de estimación de TEC normalizado durante la noche. $\mathbf{P}_{u;d}^{\text{night}}$ y $\hat{\mathbf{P}}_{u;d;R_r}^{\text{night}}$ se definirán de la siguiente manera:

$$\mathbf{P}_{u;d}^{\text{night}} = \frac{\mathbf{x}_{u;d}^{\text{night}}}{\|\mathbf{x}_{u;d}^{\text{night}}\|_1} = \mathbf{x}_{u;d}^{\text{night}} \left(\sum_{n=1}^{N_{u;d}} x_{u;d}^{\text{night}}(n) \right)^{-1} \quad (35)$$

$$\hat{\mathbf{P}}_{u;d;R_r}^{\text{night}} = \frac{\hat{\mathbf{x}}_{u;d;R_r}^{\text{night}}}{\|\hat{\mathbf{x}}_{u;d;R_r}^{\text{night}}\|_1} = \hat{\mathbf{x}}_{u;d;R_r}^{\text{night}} \left(\sum_{n=1}^{N_{u;d}} \hat{x}_{u;d;R_r}^{\text{night}}(n) \right)^{-1} \quad (36)$$

Las señales de detección de anomalías tanto diurnas como nocturnas se visualizarán mediante la interpolación kriging de todos los $TEC_{\Delta}^{\text{w-day}}$ y $TEC_{\Delta}^{\text{night}}$ para las estaciones de referencia elegidas que se presentan en la figura N°17 y para una fecha elegida d [76, 77]. Para las dos fechas consecutivas presentadas en la tabla VI, las

señales de detección de anomalías de todo el día se generan y visualizan en la figura N°22. Como se muestra en la figura N°22, las señales de detección de anomalías generadas durante todo el día muestran diferencias espaciales y temporales.

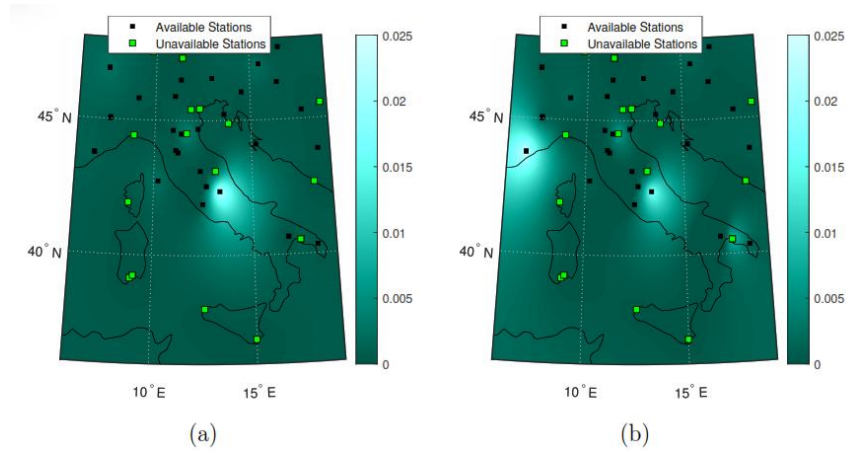


Figura N°22: Señales de detección de anomalías durante todo el día para las estaciones de referencia elegidas en la figura N°17 para (a) 11 de noviembre de 2023 y (b) 12 de noviembre de 2023. Las estaciones de referencia con TEC_{Δ}^{w-day} disponible están marcadas con cuadrados negros y las estaciones de referencia sin TEC_{Δ}^{w-day} están marcadas con cuadrados verdes.

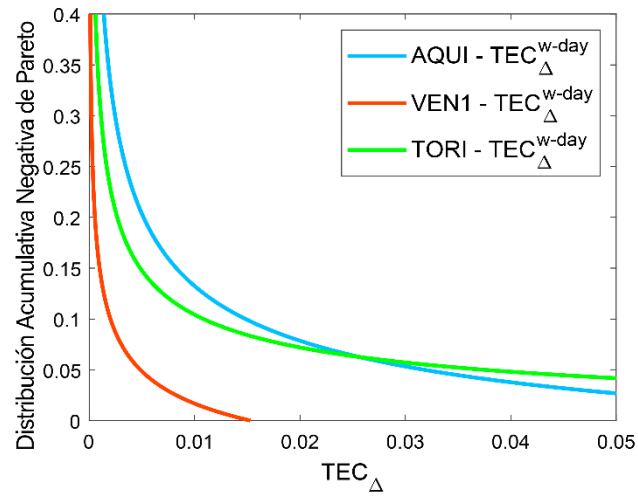
Fuente: Elaboración propia.

3.2.3. Generación adaptativa de los umbrales de detección de anomalías. Las variaciones del TEC local ionosférico se detectarán utilizando umbrales de detección de anomalías. Estos umbrales de detección de anomalías se generarán utilizando estadísticas de variaciones de TEC pasadas que se obtendrán a partir de señales de detección de anomalías pasadas. Los días de las señales de detección de anomalías se dividen en tres intervalos de tiempo no superpuestos denominados: conjunto de días de entrenamiento, validación y prueba. Las estadísticas de variaciones pasadas de TEC se generan mediante el uso de señales de detección de anomalías de conjuntos de días de entrenamiento y validación. La señal de detección de anomalías con un conjunto de fechas de prueba se reserva para la evaluación del rendimiento de la técnica EQ-PD propuesta que utiliza estadísticas de variación de TEC obtenidas de conjuntos de días de entrenamiento y validación. Luego de obtener las estadísticas de variación de TEC, se generarán y aplicarán umbrales de detección de anomalías para detectar comportamientos de anomalías en TEC_{Δ}^{w-day} observado y TEC_{Δ}^{night} en (ec. 29) y (ec. 34), respectivamente. Como se presenta en la figura N°22, las señales de detección de anomalías de TEC varían tanto en el espacio como en el tiempo. Por lo tanto, es posible umbralizar variaciones temporales, variaciones espaciales o variaciones tanto espaciales como temporales de las señales de detección de anomalías. En este estudio, se generan umbrales de detección de anomalías espaciales que son constantes en el

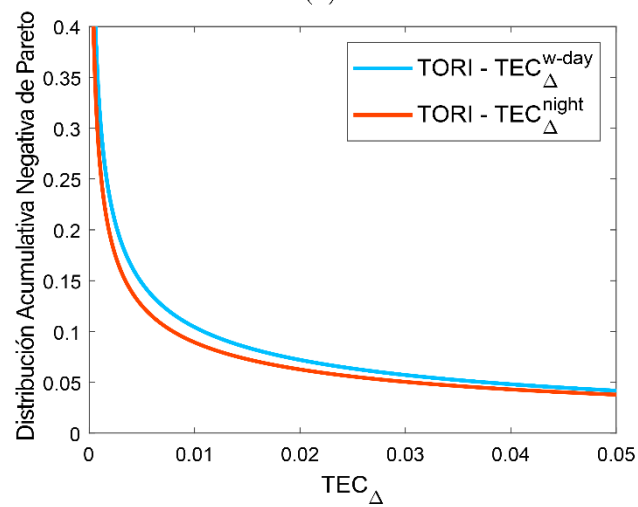
tiempo para detectar anomalías de TEC que varían espacialmente y se generan umbrales de detección de anomalías espaciotemporales para detectar anomalías de TEC que varían espacial y temporalmente. Las señales de detección de anomalías de los conjuntos de días de entrenamiento y validación se dividen además en dos subclases no superpuestas denominadas: días con actividad sísmica y días sin actividad sísmica. Las estadísticas de variación de TEC se obtienen de los días de entrenamiento y validación sin actividad sísmica.

3.2.3.1. Umbrales de detección de anomalías espaciales. Como se discutió en la sección 2.9.1, es posible generar umbrales de detección de anomalías espaciales modelando el comportamiento de variación de TEC para cada estación de referencia en una región seleccionada. El comportamiento de variación de TEC de cada estación de referencia en una región elegida. El comportamiento de variación de TEC de cada estación de referencia se estimará mediante una única distribución acumulativa negativa de Pareto. Cada distribución acumulativa negativa de Pareto mapea cada variación local de TEC de su estación de referencia relacionada con una probabilidad de cola. Las distribuciones acumulativas negativas de Pareto truncadas en la parte superior se modelarán estimando los parámetros $\hat{\gamma}$, $\hat{\vartheta}$ y $\hat{\rho}$ utilizando (ec. 16), (ec. 17) y (ec. 18), respectivamente. También es posible estimar la distribución acumulativa de Pareto para una estación de referencia elegida utilizando estos parámetros de distribución. La distribución estimada representa la probabilidad condicional de TEC_{Δ} dada una estación de referencia elegida en una región elegida y días sísmicamente inactivos durante un intervalo de tiempo elegido. Para las tres estaciones de referencia elegidas: AQUI, TORI y VEN1 ubicadas en la región presentada en la figura N°17 y el intervalo de tiempo entre agosto y diciembre del 2023, las distribuciones acumulativas negativas de Pareto truncadas en la parte superior estimadas de TEC se presentan en la figura N°23a. Como se presenta en la figura N°23a, una probabilidad de cola de variación de TEC elegida de 0,1 corresponde a diferentes valores de umbral espacial de 0,015, 0,012 y 0,003 en las estaciones de referencia AQUI, TORI y VEN1, respectivamente. Además, se repetirá el mismo procedimiento de estimación para las variaciones locales nocturnas del TEC, TEC. Para la estación de referencia TORI y el mismo intervalo de tiempo que en la figura N°23a, en la figura N°23b se presentan las distribuciones acumulativas negativas de Pareto truncadas en la parte superior estimadas de TEC y $TEC_{\Delta}^{\text{night}}$. Como se ilustra en la figura N°23b, se estiman

diferentes distribuciones acumulativas negativas de Pareto para las variaciones locales de TEC durante todo el día y la noche de la misma estación de referencia. Para cada estación de referencia EPN en el área elegida se estimará una distribución acumulativa negativa de Pareto truncada en la parte superior. Para una probabilidad de cola elegida, también se obtendrá una variación local de TEC, TEC_{Δ} para cada estación de referencia de interés. Se forma un umbral de detección de anomalías espaciales seleccionando todas las variaciones de TEC locales elegidas como umbrales. Se generan dos umbrales de detección de anomalías espaciales para todo el día eligiendo probabilidades de cola de TEC_{Δ}^{w-day} de 0,005 y 0,05. Los umbrales generados se presentan en la figura N°24a y la figura N°24b, respectivamente. Como se ilustra en la figura N°24a y la figura N°24b, a medida que las probabilidades de cola elegidas se hacen más pequeñas, las variaciones locales de TEC resultantes se vuelven más altas. Además, cambiar las probabilidades de cola da como resultado umbrales de detección de anomalías espaciales con diferentes características espaciales.



(a)



(b)

Figura N°23: (a) Distribuciones acumulativas negativas de Pareto de las variaciones locales de TEC de las estaciones de referencia AQUI, VEN1 y TORI para los días entre agosto y diciembre del 2023. (b) Distribuciones acumuladas negativas de Pareto de TEC diurno completo y TECV nocturno de las variaciones locales de TEC de la estación de referencia TORI para los días entre agosto y diciembre del 2023.

Fuente: Elaboración propia.

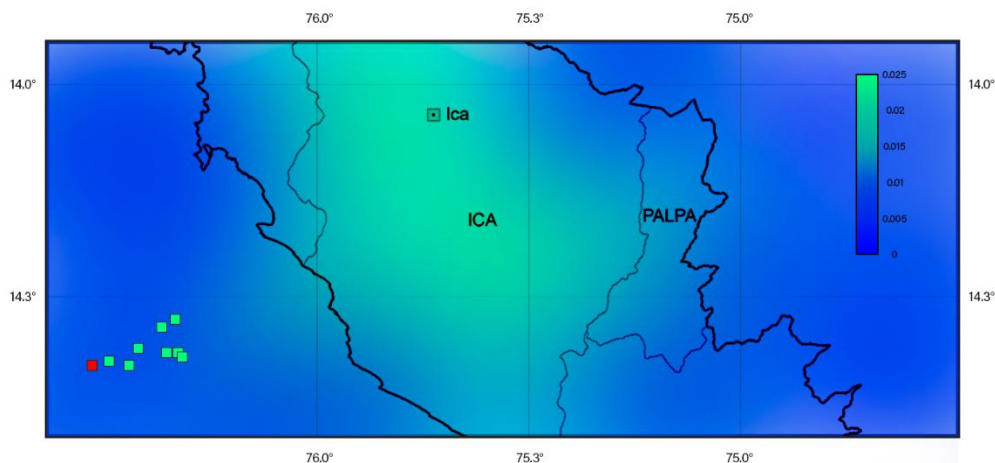


Figura N°24: Umbrales de detección de anomalías espaciales generados para las probabilidades de cola de TEC elegidas (a) 0,005 y (b) 0,05.

Fuente: Elaboración propia.

Los umbrales de detección de anomalías espaciales generados extraen la distribución de variación de TEC subyacente de cada estación de referencia independientemente de que estos umbrales se generen de forma espacialmente adaptable para cualquier estación o región de referencia determinada.

3.2.3.2. Umbrales de detección de anomalías espacio-temporales. Mediante el uso de parámetros geomagnéticos diarios, cada día de conjuntos de entrenamiento y validación se agrupará con respecto a su actividad geomagnética. Después de la agrupación de los días basada en la actividad geomagnética, es posible generar un umbral de detección de anomalías espaciales para cada agrupación de días diferente. Se generará un umbral de detección de anomalías espacio-temporales que varíe tanto en el espacio como en el tiempo aplicando un umbral de detección de anomalías espaciales diferente para cada día diferente. Las técnicas de aprendizaje automático no supervisadas no requieren datos etiquetados para funcionar. La agrupación en clústeres de K-means es una de las técnicas de aprendizaje automático no supervisadas más utilizadas. La técnica se basa en encontrar k medias de conglomerados diferentes en m espacio dimensional, iterativamente. El pseudocódigo de la técnica de agrupamiento de k -means aplicada se presenta en el algoritmo 2. Para demostrar el rendimiento de agrupamiento de la técnica propuesta, datos de parámetros geomagnéticos diarios, g_d se define para el día d como un vector de cinco dimensiones que consta de parámetros geomagnéticos diarios de índices A_p, AE, K_p, Dst y SSN . Los datos de parámetros geomagnéticos se almacenan para cada día entre agosto y diciembre del 2023. Cada dimensión o parámetro geomagnético de G se normaliza con respecto a sus valores

mínimos y máximos. La técnica de agrupamiento de k -means se ejecuta en el g_d normalizado para diferentes valores de k que oscilan entre 1 y 20. Para cada ejecución de la técnica, se calcula y almacena una puntuación de silueta media. El valor k óptimo que maximiza la puntuación media de la silueta se elige como $k = 3$ [111, 112]. Las medias de conglomerados generadas para los datos de parámetros geomagnéticos se presentan en la tabla VII con el valor medio de cada dimensión de g_d . Cuando se trata de la operación en tiempo real de la propuesta, la propuesta se ejecuta en una operación en tiempo casi real debido al requisito de procesamiento adicional en algunos de los parámetros geomagnéticos elegidos. La técnica EQ-PD requiere algunos datos geomagnéticos como índices Dst y AE que pueden ser provisionales y no definitivos para el momento actual. Estos parámetros geomagnéticos requieren un procesamiento adicional para eliminar el posible ruido artificial y los cambios de línea de base [94]. A pesar de este requisito de procesamiento de los datos, la técnica EQ-PD es inmune a posibles errores en los datos. El efecto del ruido artificial y los cambios de línea de base en la técnica de agrupación en clústeres de k -means es mínimo ya que la técnica procesa datos en el orden de los años. Por lo tanto, las ubicaciones medias de los conglomerados se desplazarán ligeramente debido a estos errores.

Algoritmo 2 Pseudocódigo de agrupamiento k -means

```
// Definición de parámetros //
 $N_{iter} \leftarrow$  Número máximo de iteraciones
 $N_{samples} \leftarrow$  Número de muestras del día
 $k \leftarrow$  Número de ubicaciones de clusters

// Inicialización //
for  $i = (1 \text{ a } K)$  do
     $\mu_i \leftarrow$  Inicializar la media para el  $i$ -ésimo cluster
end for

// Búsqueda iterativa //
for  $t = (1 \text{ a } N_{iter})$  do
    for  $d = (1 \text{ a } N_{samples})$  do
         $z_d = \operatorname{argmin}_i \|\mu_i - g_d\| \leftarrow$  Asigne la muestra del día  $d$  a la ubicación media más cercana
    end for
    for  $i = (1 \text{ a } k)$  do
         $\mu_i = \operatorname{MEAN}(g_d : z_d = k) \leftarrow$  Vuelve a calcular la media del  $i$ -ésimo cluster
    end for
end for
```

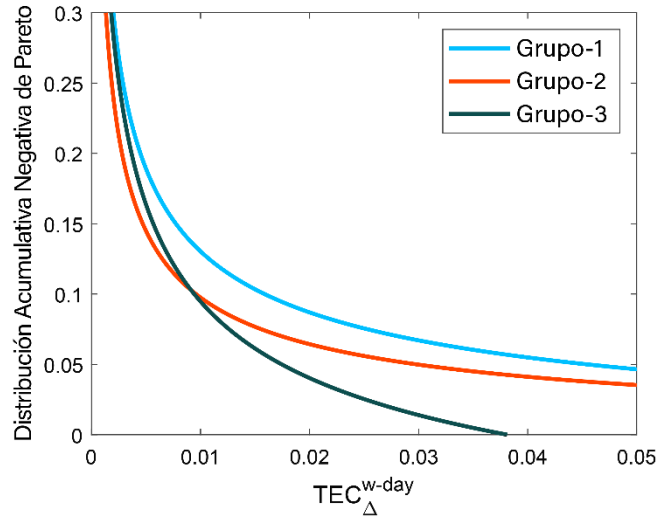
Además, el proceso de detección de precursores de EQ también es inmune a estos errores, ya que la detección de precursores de EQ se realiza mediante la técnica Support Vector Machine, que apenas se ve afectada por valores atípicos en los datos.

Tabla VII
Medias de cluster generadas para los días entre agosto y diciembre del 2023 y valor medio de cada parámetro geomagnético.

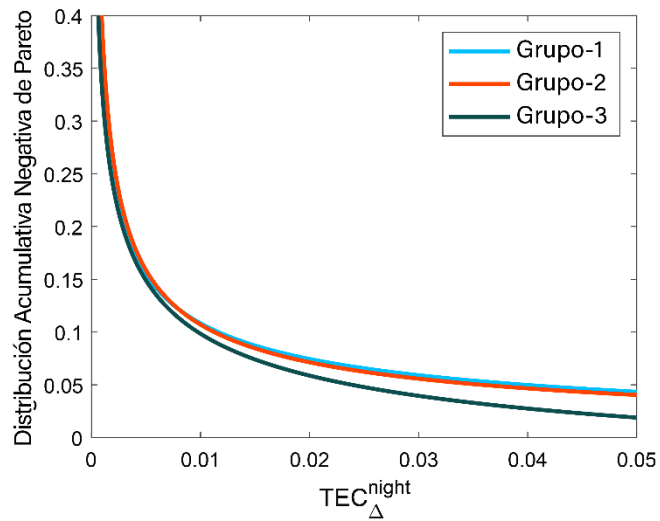
k -means cluster	K_p	SSN	Dst (nT)	A_p (nT)	AE (nT)
1	3.19481	51.83144	-25.8080	21.7639	350.6083
2	1.12953	20.9920	-5.4621	4.8475	93.5610
3	1.40000	112.5771	-5.8790	6.2715	127.5974
Mean	1.58834	50.9175	-9.3896	8.3969	150.7634

Como se muestra en la tabla VII, la técnica de agrupamiento de k -means obtuvo el puntaje de silueta promedio más alto cuando $k = 3$ y generó tres promedios de agrupamiento diferentes basados en los parámetros geomagnéticos elegidos de los índices A_p, AE, K_p, Dst . Cada una de las medias de grupos de días representa un nivel diferente de actividad geomagnética ionosférica. Cuando los niveles medios de actividad geomagnética del grupo se comparan entre sí, el cluster-1 y el cluster-3 representan los días con actividades geomagnéticas relativamente más altas. El cluster-2 representa los días con actividades geomagnéticas relativamente más bajas en la ionosfera. Una vez obtenidas estas medias de conglomerados, cada día siguiente se clasificará con respecto a su actividad geomagnética midiendo la distancia entre los parámetros geomagnéticos de un día elegido d, g_d y las medias de cluster. El día elegido se asignará al clúster con menor distancia. Los umbrales de detección de anomalías espacio-temporales se generarán estimando k diferentes distribuciones acumulativas de Pareto que se obtienen a partir de días agrupados de conjuntos de entrenamiento y validación. Los días agrupados se clasifican con respecto a las k -means de agrupamiento diferentes que se presentan en la tabla VII. Para una región de elección, se estimarán las k distribuciones acumulativas de Pareto diferentes para cada estación de referencia. Para la estación de referencia TORI ubicada en la región presentada en la figura N°17, el intervalo de tiempo entre agosto y diciembre del 2023 y tres estados de actividad diferentes de la ionosfera, las

distribuciones acumulativas negativas de Pareto truncadas en la parte superior estimadas de TEC_{Δ}^{w-day} se presentan en la figura N°25a.



(a)



(b)

Figura N°25: (a) Distribuciones acumulativas de Pareto truncado superior negativo TEC_{Δ}^{w-day} de día completo de la estación de referencia TORI para los tres grupos de día diferentes, (b) Distribuciones acumulativas de Pareto truncado superior negativo de TEC_{Δ}^{night} nocturno de la estación de referencia TORI para los tres grupos de día diferentes.

Fuente: Elaboración propia.

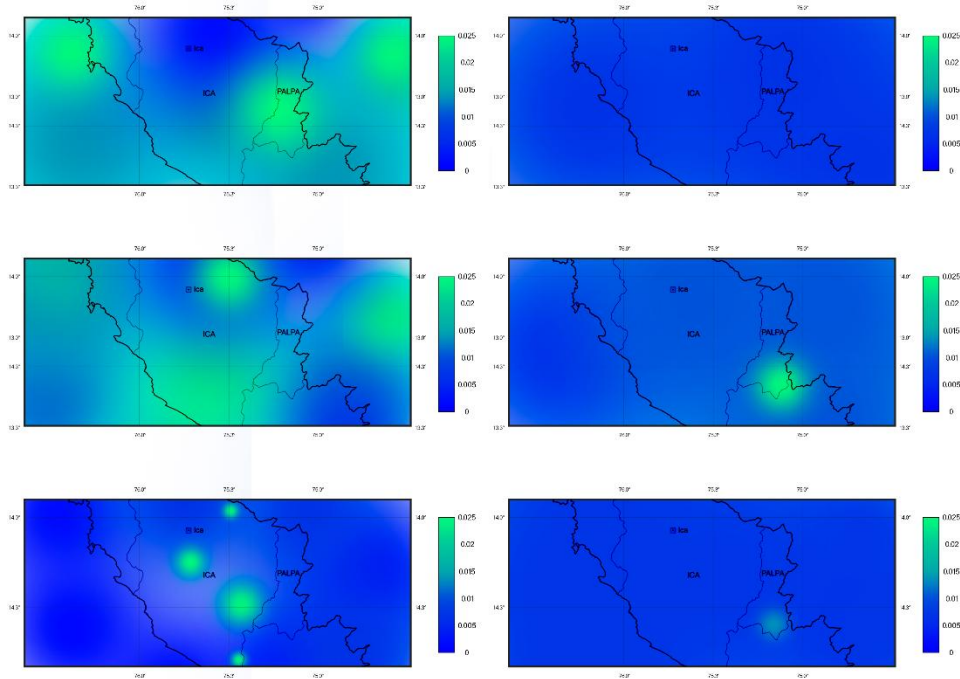


Figura N°26: Umbrales de detección de anomalías espacio-temporales generados para los diferentes grupos y probabilidades de cola de TEC_{Δ}^{w-day} elegidas (a) Cluster-1 0.005, (b) Cluster-1 0.05, (c) Cluster-2 0.005, (d) Cluster-2 0.05, (e) Cluster-3 0.005, (f) Cluster-3 0.05.

Fuente: Elaboración propia.

Como se presenta en la figura N°25a, una probabilidad de cola de variación de TEC elegida de 0,05 corresponde a tres valores de umbral espacial de 0,045, 0,027 y 0,016 en la estación de referencia TORI para el cluster-1, el cluster-2 y el cluster-3, respectivamente. Además, se repetirá el mismo procedimiento de estimación para las variaciones locales de TEC durante la noche, TEC_{Δ}^{night} como se muestra en la figura N°25b. Después de elegir una probabilidad de cola de TEC_{Δ} adecuada, es posible obtener k umbrales de TEC_{Δ} diferentes de cada estación de referencia. Por lo tanto, se genera un umbral de detección de anomalías espacio-temporales que es posible aplicar un umbral espacial diferente para cada tipo de día diferente. La figura N°26 presenta un umbral de detección de anomalías espacio-temporales obtenidos del conjunto de días que pertenecen al cluster-1, cluster-2 y cluster-3 en la tabla VII al elegir la probabilidad de cola de TEC_{Δ} como 0.05 y 0.005. Como se presenta en la figura N°26, los umbrales de detección de anomalías espacio-temporales generados son espacialmente diferentes entre sí. Los umbrales de detección de anomalías espacio-temporales generados extraen las distribuciones de variación de TEC subyacentes de cada estación de referencia para diferentes estados de actividad de la ionosfera de forma independiente.

Por lo tanto, estos umbrales se generarán de forma adaptable espacial y temporalmente para cualquier región o estación de referencia determinada.

3.2.4. Generación de señal de detección de precursores de ecualización. El EQ-PD

propuesto se basa en la detección de anomalías más allá del TEC_{Δ} esperado, provocadas por fuertes actividades sísmicas. TEC_{Δ} , se examinan los comportamientos de las estaciones de referencia para conocer los TEC_{Δ} esperados, comportamientos cuando no hay actividad sísmica en la ionosfera. Por lo tanto, es posible identificar más allá de las anomalías TEC_{Δ} esperadas cuando hay actividad sísmica en la ionosfera. Para un TEC_{Δ} dado, probabilidad de cola, se puede generar un umbral de detección de anomalías al encontrar el TEC_{Δ} correspondiente en las distribuciones de Pareto estimadas para cada estación de referencia de interés. Dado que las distribuciones de Pareto estimadas se obtienen de días sin actividad sísmica, representan el comportamiento TEC_{Δ} de cada estación de referencia para los días sin actividad sísmica. Por lo tanto, cada probabilidad de cola de TEC_{Δ} diferente corresponde a una probabilidad diferente de falsa alarma (P_{FA}) o tasa de detecciones de anomalías de TEC_{Δ} para la clase de días sin actividad sísmica.

Varios estudios informan alteraciones importantes en TEC antes de 1 a 9 días antes de EQS fuerte [29, 109, 110, 113]. Varias perturbaciones relacionadas con TEC se informan y analizan estadísticamente en [44]. Como se presenta en [44], es de 1,8 a 3,6 veces más probable observar perturbaciones TEC antes de 1 a 9 días de EQ fuertes que observar estas perturbaciones cuando no hay actividad sísmica. Debido a estos hechos, los días de clase de días de actividad sísmica se eligen con respecto a los 9 días anteriores a cada EQ fuerte, incluidos los días de EQ. Todos los días sísmicamente inactivos restantes se incluyen en la clase de días sin actividad sísmica.

Para demostrar la relación entre las probabilidades de cola de TEC_{Δ} y P_{FA} , de agosto a diciembre del 2023 se divide en días de actividad sísmica y clases de días sin actividad sísmica. Dado que se observan perturbaciones significativas de TEC 9 días antes de un EQ, estos días, incluido el día EQ, se seleccionan para la clase de días de actividad sísmica. La clase de días de actividad sísmica consta de 24 intervalos de tiempo diferentes, incluidos 9 días antes de los 24 EQ que se presentan en la tabla XI. La relación entre la probabilidad de cola de TEC_{Δ} y P_{FA} obtenida a partir de una señal de detección de anomalías de día completo para la actividad no sísmica la clase de días se presenta en la figura

Nº27. Como se muestra en la figura Nº27, se pueden obtener diferentes comportamientos estadísticos para la detección de anomalías espaciales y los umbrales de detección de anomalías espacio-temporales ajustando las probabilidades de cola de TEC_{Δ} en las distribuciones de Pareto estimadas.

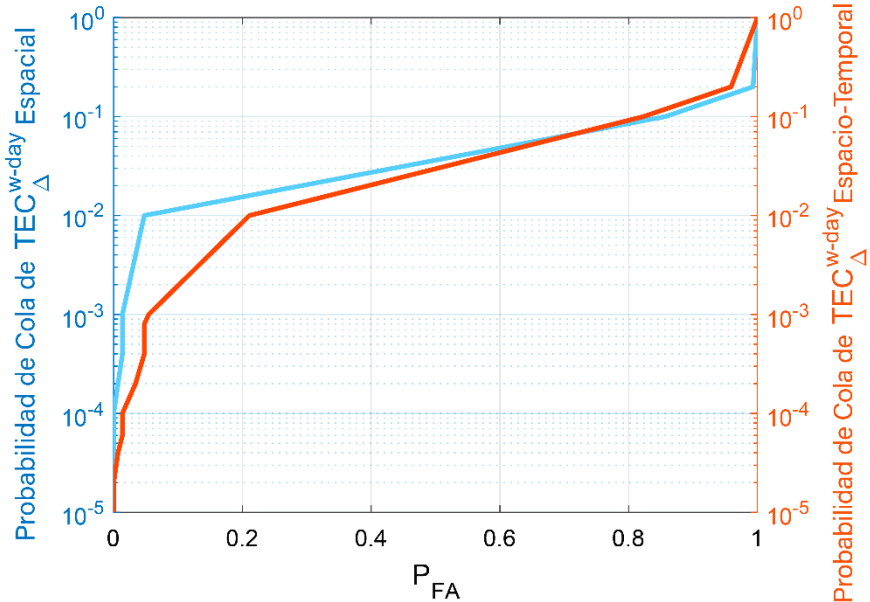


Figura Nº27: Relación entre la probabilidad de cola de TEC y P obtenida de los días sin actividad sísmica durante agosto y diciembre de 2023.

Fuente: Elaboración propia.

Para una probabilidad de cola de TEC_{Δ} , se pueden tomar decisiones sobre anomalías aplicando los umbrales de detección de anomalías correspondientes en la señal de detección de anomalías. Las señales de detección de precursores de EQ se pueden generar recopilando todas las decisiones de anomalías diarias obtenidas de diferentes combinaciones de señales de detección de anomalías y umbrales de detección de anomalías. En este estudio, se generan 6 señales de detección de anomalías diferentes. Se generan tres señales diferentes de detección de anomalías durante todo el día y tres durante la noche alterando el parámetro de radio R_r de la técnica de interpolación TEC espaciotemporal descrita en la sección 2.6 y la sección 2.7. Los umbrales de detección de anomalías espaciales y espacio-temporales se generan para estas seis señales de detección de anomalías diferentes, como se detalla en la sección 3.2.3.1 y la sección 3.2.3.2. Durante el proceso de generación de umbrales, se eligen 23 probabilidades de cola de TEC_{Δ} diferentes para cada umbral de detección de anomalías. Para 6 señales de detección de anomalías diferentes, 2 tipos de umbral de detección de anomalías diferentes y 23 niveles de umbrales diferentes se generarán $6 \times 2 \times 23 = 276$ decisiones de anomalía diferentes para cada día de interés. Cada decisión de anomalía se marca como 1 si hay

una superación del umbral en una señal de detección de anomalía elegida y un umbral de detección de anomalía. Si no se excede el umbral, la decisión de anomalía relacionada se marca como 0. La señal de detección de precursores de EQ diaria se forma apilando las 276 decisiones de anomalía diferentes y se visualiza en la figura N°28.

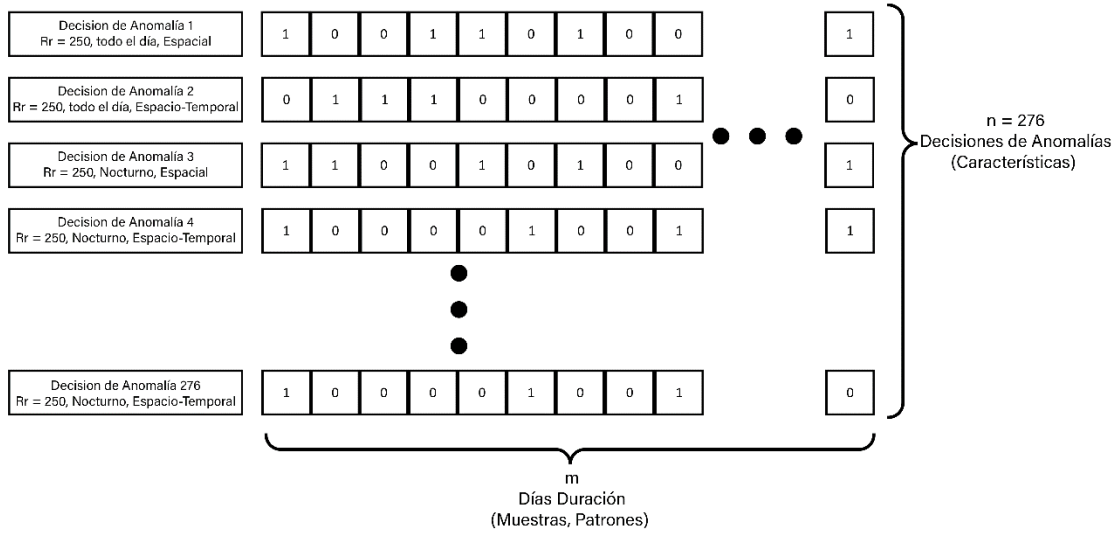


Figura N°28: Señal de detección de precursores de ecualización.

Fuente: Elaboración propia.

3.2.5. Generación de detecciones de precursores de EQ. Support Vector Machine (SVM)

es uno de los métodos de predicción más robustos y ampliamente utilizados en el aprendizaje automático [114, 115, 116]. SVM es una técnica de aprendizaje automático supervisado que realizará una clasificación lineal binaria o una clasificación no lineal binaria con el uso de kernels. En este trabajo, SVM se elige para clasificar los días de la señal de detección de precursores de EQ en días de actividad sísmica y clases de días sin actividad sísmica. Cada conjunto de días de capacitación y validación de días se etiqueta como $y = 1$, si el día pertenece a la clase de días de actividad sísmica. Todos los días restantes de la clase de actividad no sísmica se etiquetan como $y = 0$. El objetivo de SVM es obtener los coeficientes de límite de decisión óptimos, θ , que minimicen la siguiente función de costo:

$$\min_{\theta} C_{reg} \sum_{d=1}^m y^{(d)} C_{1;h}(\theta^T s^{(d)}) + (1 - y^{(d)}) C_{0;h}(\theta^T s^{(d)}) + \frac{1}{2} \sum_{j=1}^m \theta_j^2 \quad (37)$$

donde n son las n decisiones de anomalías diarias apiladas en la señal de detección de precursores de EQ, m es la duración del día de la señal de detección de precursores de EQ, C_{reg} es la constante de regularización. Las funciones de pérdida de bisagra de

$C_{0;h}$ y $C_{1;h}$ representan el costo de clasificar cuando las etiquetas de día son $y = 0$ y $y = 1$, respectivamente, y se definen de la siguiente manera:

$$C_{0;h}(\boldsymbol{\theta}^T \mathbf{s}^{(d)}) = \max\left(0, (1 + \boldsymbol{\theta}^T \mathbf{s}^{(d)})\right) \quad (38)$$

$$C_{1;h}(\boldsymbol{\theta}^T \mathbf{s}^{(d)}) = \ln\left(0, (1 - \boldsymbol{\theta}^T \mathbf{s}^{(d)})\right) \quad (39)$$

$\mathbf{s}^{(d)}$ representa un vector de muestra dimensional m que se obtendrá aplicando un kernel gaussiano con parámetro σ en la muestra del día d -ésimo de la señal de detección del precursor EQ de la siguiente manera:

$$\mathbf{s}^{(d)} = [\mathbf{s}_1^{(d)} \dots \mathbf{s}_r^{(d)} \dots \mathbf{s}_m^{(d)}]^T \quad (40)$$

$$\mathbf{s}_r^{(d)} = \exp\left(-\frac{\|\mathbf{P}_E^{(d)} - \mathbf{P}_E^{(r)}\|^2}{2\sigma^2}\right) = \exp\left(-\frac{\sum_{j=1}^n (P_{E;j}^{(d)} - P_{E;j}^{(r)})^2}{2\sigma^2}\right) \quad (41)$$

donde $P_{E;j}^{(d)}$ representa la j -ésima decisión de entrada o anomalía de la señal de detección de precursores de EQ para el día d . Durante el entrenamiento del SVM, cada vector de muestra, $\mathbf{s}^{(d)}$, se multiplica por un peso de muestra, $w^{(d)}$. Los pesos de muestra se eligen inicialmente como 1, $w_{init}^{(d)} = 1$ para la muestra d -ésima y se normalizan con respecto al número de muestras en cada clase para aproximar probabilidades previas de las clases. La probabilidad previa de la clase sin actividad sísmica o Clase 0, $p_D(0)$ y la probabilidad previa de la clase de actividad sísmica o Clase 1, $p_D(1)$ se definirán como:

$$\mathbf{p}_D = \begin{bmatrix} p_D(0) \\ p_D(1) \end{bmatrix} = \frac{1}{m_0 + m_1} \begin{bmatrix} m_0 \\ m_1 \end{bmatrix} \quad (42)$$

donde m_0 es el número de muestras en la Clase 0 y m_1 es el número de muestras en la Clase 1. Además, las probabilidades previas de estas clases se modificarán para entrenar el SVM de manera sensible a los costos. El vector de probabilidad anterior, \mathbf{p}_D , se actualizará multiplicando el vector con una matriz de costos de la siguiente manera:

$$\mathbf{p}_{DC} = \begin{bmatrix} 0 & Cost_{FA} \\ Cost_{MD} & 0 \end{bmatrix} \mathbf{p}_D \quad (43)$$

donde $Cost_{FA}$ es el costo de generar una falsa alarma y $Cost_{MD}$ es el costo de generar una detección errónea. El peso de la muestra de la muestra d -ésima se obtendrá utilizando el vector de probabilidad anterior de actualización, \mathbf{p}_{DC} para $d \in \text{Class } 0$ como en:

$$w^{(d)} = \frac{w_{init}^{(d)}}{\sum_{\forall d \in Class 0} w_{init}^{(d)}} p_{DC}(0) \quad (44)$$

Y cuando $d \in Class 1$:

$$w^{(d)} = \frac{w_{init}^{(d)}}{\sum_{\forall d \in Class 1} w_{init}^{(d)}} p_{DC}(1) \quad (45)$$

Los pesos de muestra actualizados incorporan el costo de la detección errónea y el costo de las falsas alarmas en la función objetivo de SVM en (ec. 37). Por lo tanto, SVM genera el límite de decisión, θ con respecto a estos costos de que el entrenamiento de SVM será sensible a los costos. Como ejemplo, elegir $Cost_{MD} = 1$ y $Cost_{FA} > 1$ aplicará la SVM entrenada para reducir la cantidad de falsas alarmas generadas.

El entrenamiento del modelo SVM propuesto requiere el ajuste de hiperparámetros de que los parámetros C_{reg}, σ y $Cost_{FA}$ afectan el rendimiento de detección de precursores de EQ-PD. Por lo tanto, el entrenamiento se realizará para diferentes combinaciones de parámetros C_{reg}, σ y $Cost_{FA}$ en el conjunto de días de entrenamiento. A partir de entonces, para cada coeficiente de límite de decisión de combinación elegido, se debe obtener θ . Entre todos los conjuntos de coeficientes de límite de decisión, los que tengan el rendimiento más alto en el conjunto de días de validación deben elegirse como los coeficientes de límite de decisión óptimos. El rendimiento del EQ-PD se evaluará con este conjunto de coeficientes límite de decisión en el conjunto de días de prueba.

3.3. Rendimiento de detección de precursores de la técnica EQ-PD propuesta. El rendimiento de la técnica EQ-PD propuesta se investiga en conjuntos de datos de entrenamiento, validación y prueba que proporcionan una evaluación imparcial del modelo SVM elegido. Para lograr una evaluación imparcial de la técnica, se generan y presentan en la tabla VIII tres señales de detección de precursores de EQ que no se cruzan para el entrenamiento, la validación y la prueba del modelo. Como se muestra en la tabla VIII, se presentan las fechas de inicio y finalización, el número de días y el número de EQ diferentes diarios en estos períodos de tiempo para los conjuntos de entrenamiento, validación y prueba. El rendimiento de detección de precursores de la técnica EQ-PD se evalúa en el conjunto de días de prueba. Durante el conjunto de pruebas de fechas y la región elegida, se habían producido 24 EQ diferentes diarios y se presentan en la tabla XI. Todos los EQ presentados tenían una magnitud superior a 4 en la escala de Richter [107]. Si se han realizado múltiples ecualizadores el mismo día en la región elegida, se elige el ecualizador con la magnitud más

alta para la evaluación del rendimiento de la prueba. Como se muestra en la tabla XI, se investigan diariamente 24 EQ diferentes de un total de 43 EQ.

Tabla VIII
Información de entrenamiento, validación y conjunto de prueba.

Data set	Fecha (día mes año)		Número de días	Números de EQ diarios diferentes
	Inicio	Fin		
Entrenamiento	1 agosto de 2023	1 agosto de 2023	456	35
Validación	2 agosto de 2023	23 setiembre de 2023	266	21
Test	24 noviembre de 2023	30 diciembre de 2023	282	24

Además, se realizan experimentos de adivinanzas aleatorias para mostrar la importancia estadística y la fiabilidad del rendimiento de detección de precursores de la técnica EQ-PD. Los resultados de los experimentos se comparan con los rendimientos de detección de precursores de EQ-PD para probar su significación estadística. Los resultados de las pruebas y el análisis de confiabilidad de lo propuesto se presentan en subsecciones posteriores.

3.3.1. Resultados de pruebas. En la evaluación del rendimiento de la prueba, se generan 3 estimaciones diferentes de TEC de día completo y 3 de noche diferentes ajustando el parámetro de radio de estimación de TEC, R_r como 250 km, 300 km y 350 km como se detalla en la sección 2.6. Las estimaciones de TEC se generan para las mediciones de TEC de las estaciones de referencia ubicadas en la región elegida para los días entre agosto de 2023 (primer día del conjunto de entrenamiento) y diciembre de 2023 (último día del conjunto de prueba). TEC_{Δ}^{w-day} diurno completo TEC_{Δ}^{night} diurno y nocturno se calculan las variaciones locales de TEC y se generan señales de detección de anomalías diurnas y nocturnas para todas las estaciones de referencia ubicadas en la región elegida durante el mismo tiempo, como se detalla en la sección 3.2.2. Como se detalla en la sección 3.2.3.1, los comportamientos de TEC_{Δ} espacial de las señales de detección de anomalías generadas se obtienen mediante la estimación de TEC_{Δ} . Distribuciones acumulativas de Pareto para los días sin actividad sísmica entre las fechas de agosto de 2023 (primer día del conjunto de entrenamiento) y diciembre de 2023 (último día del conjunto de entrenamiento). De manera similar, como se detalla en la sección 3.2.3.2, los comportamientos espacio-temporales de TEC_{Δ} de las señales de detección de anomalías generadas se obtienen agrupando los días de clase sin actividad sísmica en $k = 3$ grupos de días diferentes y estimando las distribuciones acumulativas de Pareto de TEC_{Δ} para cada uno de estos grupos de días para los días de clase de actividad no sísmica entre las fechas de agosto de 2023 (primer día del

conjunto de entrenamiento) y diciembre 2023 (último día del conjunto de entrenamiento). Además de los comportamientos de TEC_{Δ} , las ubicaciones medias de los grupos también se obtienen del conjunto de días de entrenamiento. Los próximos días de validación y el conjunto de días de prueba se agrupan con respecto a estas $k = 3$ ubicaciones medias de conglomerados. Finalmente, los umbrales de detección de anomalías espaciales y espacio-temporales también se generan al seleccionar 23 probabilidades de cola de TEC_{Δ} diferentes en las distribuciones acumulativas de Pareto de TEC_{Δ} relacionadas.

La señal precursora de ecualización del conjunto de entrenamiento se obtendrá mediante la umbralización de las señales de detección de anomalías con los umbrales de detección de anomalías apropiados y el apilamiento resultante de 276 decisiones de anomalías diarias diferentes para el conjunto de fechas de entrenamiento presentado en la tabla VIII. Para el conjunto de fechas de validación y prueba que se presenta en la tabla VIII, las señales precursoras de EQ se generan al establecer el umbral de las futuras señales de detección de anomalías con los umbrales de detección de anomalías apropiados que se obtienen del conjunto de entrenamiento.

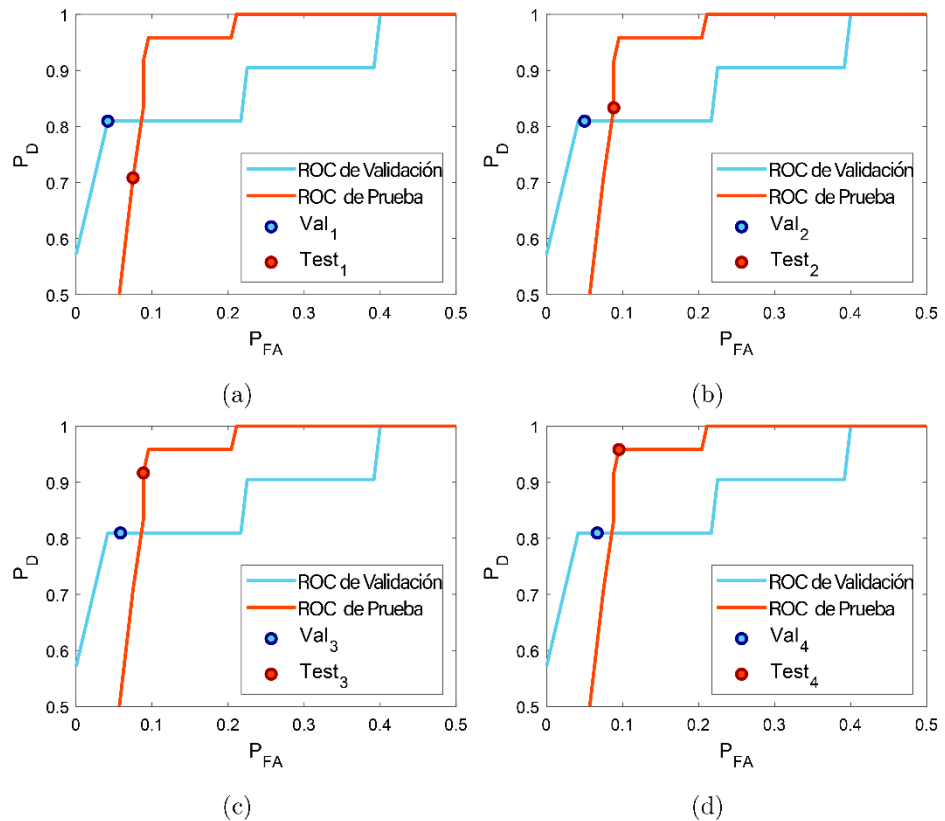


Figura N°29: Curvas ROC de los conjuntos de validación (azul) y prueba (naranja) y los puntos ROC de validación (círculo azul) y prueba (círculo naranja) elegidos para (a) Validación1 y Prueba1, (b) Validación2 y Prueba2, (c) Validación3 y Prueba3, (d) Validación4 y Prueba4.

Fuente: Elaboración propia.

Para diferentes combinaciones de los parámetros del modelo C_{reg} , σ y $Cost_{FA}$, se entrenan diferentes modelos SVM en la señal de detección de precursores de EQ que se genera para el conjunto de días de entrenamiento. Durante el entrenamiento de los modelos, se elige el parámetro de regularización de SVM, C_{reg} del rango 0.01 y 10000, el parámetro del kernel Gaussiano, se elige el σ del rango 0.001 y 1000 y el parámetro costo de generar una falsa alarma, $Cost_{FA}$ se elige del rango 1 y 2.1. Costo de generar un parámetro de detección errónea, $Cost_{MD}$ se elige como 1 durante el entrenamiento. SVM Coeficientes límite de decisión, θ se obtienen para cada combinación de parámetros de modelo diferente. Cada límite de decisión de SVM obtenido a partir de los coeficientes de límite de decisión se valida en la señal de detección de precursores de EQ que se genera para el conjunto de días de validación. La curva de características operativas del receptor (ROC) del conjunto de validación se obtiene eligiendo los modelos SVM validados con la mayor probabilidad de detección (P_D) para un P_{FA} determinado en el conjunto de días de validación. De manera similar, la curva ROC del conjunto de prueba también se obtiene aplicando los modelos SVM validados elegidos en la señal de detección de precursores de EQ

que se genera para el conjunto de días de prueba. Las curvas ROC de validación y prueba de los modelos SVM elegidos con diferentes parámetros se presentan en la figura N°29. Cuando se tienen en cuenta las curvas ROC generadas en la figura N°29, estas curvas exhiben un comportamiento escalonado en el eje P_D , ya que los valores de P_D se calculan con respecto a un pequeño número de precursores de EQ en los conjuntos de validación y prueba. Hay 21 precursores de EQ en el conjunto de validación y 24 precursores de EQ en el conjunto de prueba que P_D tomará 21 y 24 valores distintos de cero durante las evaluaciones de rendimiento de los conjuntos de validación y prueba, respectivamente. Además, el eje P_{FA} se controlará ajustando las probabilidades de cola de TEC de los umbrales de detección de anomalías, como se ilustra en la figura N°27. Por lo tanto, los umbrales de detección de anomalías generados no tienen control sobre el eje P_D , ya que el aumento de la P_{FA} no siempre puede dar lugar a un mayor número de detecciones de precursores de EQ. Finalmente, las curvas ROC generadas se visualizan para $0.5 \leq P_D \leq 1$ y $0 \leq P_{FA} \leq 0.5$ que los rendimientos de los puntos ROC con $P_D \leq 0.5$ o $P_{FA} \geq 0.5$ no se tienen en cuenta durante la visualización. Para determinar el rendimiento de detección de precursores de EQ del propuesto, se eligen 4 puntos ROC diferentes de la curva ROC de validación y se nombran como: Validación1, Validación2, Validación3 y Validación4. El rendimiento de detección de precursores de eualización de estos puntos ROC con sus correspondientes puntos ROC de prueba también se presenta en la figura N°29. Además, el número de precursores EQ detectados, el número de detecciones de precursores falsos y los parámetros del modelo SVM se presentan en la tabla IX. para los puntos ROC elegidos. Como se muestra en la figura N°29 y la tabla IX, los puntos ROC de validación logran una menor cantidad de falsas alarmas, mientras que sus correspondientes puntos ROC de prueba generan una mayor cantidad de falsas alarmas.

Las distancias de detección de precursores de EQ se calcularán asignando la estación de referencia con el nivel más alto de anomalía de TEC_{Δ} a cada día. Por lo tanto, cada detección de precursor de EQ se asignará a una estación de referencia. Para un epicentro de EQ elegido, la distancia de detección del precursor de EQ se obtendrá promediando las distancias entre el epicentro y las ubicaciones de las estaciones de referencia donde se detectó el precursor relacionado. Las distancias de detección de precursores de EQ también se promediarán en todos los EQ para generar distancias de detección de precursores promedio. Para los puntos ROC de prueba que se presentan en la figura N°29, las distancias promedio de detección de precursores de EQ se presentan en la tabla IX. Como se muestra en la tabla IX, Test₁ ROC point logra la distancia de detección de precursor promedio más pequeña, mientras que Test₄ tiene

la distancia de detección de precursor de EQ promedio más alejada. Los histogramas de detección de precursores de estos puntos ROC también se visualizan en la figura N°30.

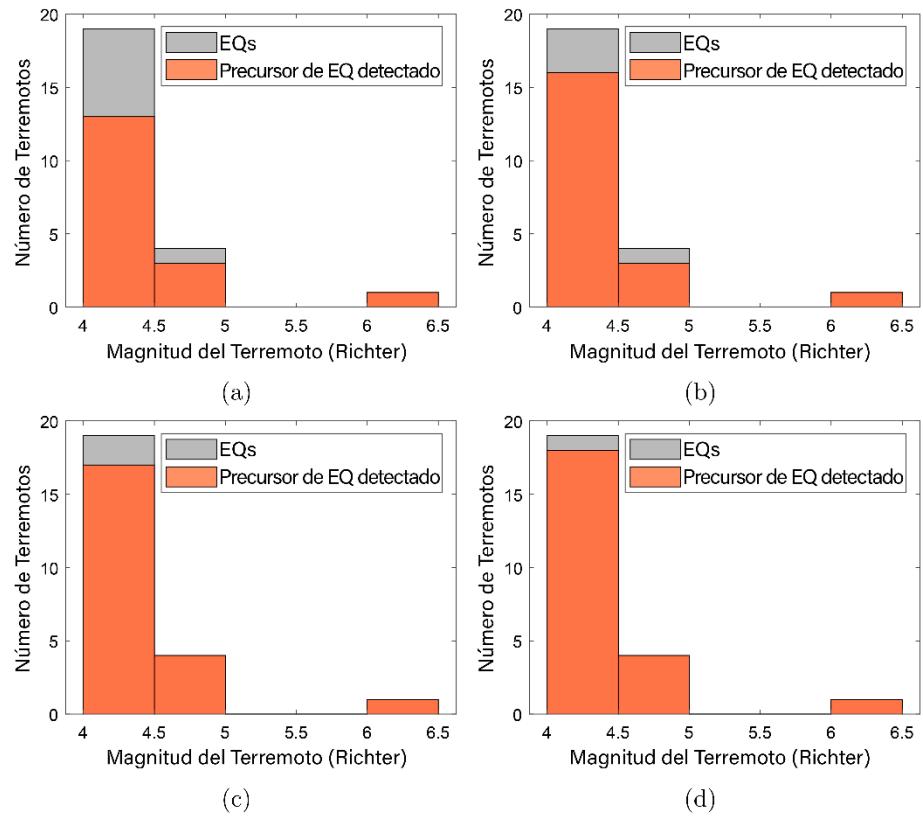


Figura N°30: EQ Histogramas de detección de precursores para puntos ROC (a) Prueba₁, (b) Prueba₂, (c) Prueba₃, (d) Prueba₄.

Fuente: Elaboración propia.

Tabla IX

Para los puntos ROC de validación y prueba elegidos: número de precursores de EQ detectados, número de todos los EQ, número de detecciones de precursores falsos, número de días sin actividad sísmica (días de falsa alarma), parámetros de entrenamiento de SVM y distancia promedio de detección de precursores para todos los EQ $\sum M$, 11 EQs débiles con magnitud inferior a 4,3, 10 ecualizadores moderados con magnitud entre 4,3 y 4,6 y 3 EQs fuertes con magnitudes superiores a 4,6 en la escala de Richter.

Punto ROC	Precursor de EQ detectado	Números de EQs	Generó falsas alarmas	Días sin actividad sísmica	Parámetros de entrenamiento			Distancia media de detección de precursores (km)			
					C_{reg}	σ	$Cost_{FA}$	$\sum M$	$M < 4.3$	$4.3 \leq M < 4.6$	$4.6 \leq M$
Validación ₁	17	21	5	120	80	500	1.6	-			
Test ₁	17	24	11	147	-			380	389	415	280
Validación ₂	17	21	6	120	0.2	30	1.2	-			
Test ₂	20	24	13	147	-			426	495	402	280
Validación ₃	17	21	7	120	1	40	1.8	-			
Test ₃	22	24	13	147	-			442	508	431	280
Validación ₄	17	21	8	120	0.2	20	1.5	-			
Test ₄	23	24	14	147	-			466	545	443	280

El proceso de ajuste de hiperparámetros de SVM decide si los coeficientes de límite de decisión resultantes se sobreajustan a la señal de detección de precursores de EQ del conjunto de días de entrenamiento o no. A medida que el parámetro de regularización de SVM, C_{reg} aumenta, la regularización del coeficiente de límite de decisión se reducirá y dará como resultado un sobreajuste del modelo entrenado. Además, como el parámetro del núcleo gaussiano, σ se vuelve más pequeño, aumenta la varianza de las características y también da como resultado un sobreajuste del modelo entrenado. A que se muestra en la tabla IX, los modelos SVM entrenados no se ajustan ni se ajustan demasiado a los datos de entrenamiento. Como resultado, la técnica EQ-PD detecta 22 precursores de EQ de 24 EQ y genera 13 falsas alarmas diferentes en 147 días de actividad sin actividad sísmica en 2023.

3.3.2. Análisis de fiabilidad. Con el fin de mostrar la importancia estadística y la fiabilidad del rendimiento de detección de precursores de la técnica EQ-PD, se realizan 100000 simulaciones aleatorias de Monte Carlo con respecto a los parámetros del espacio ROC de los puntos ROC de validación y prueba elegidos. El rendimiento de la propuesta se investiga en términos de precursores de EQ detectados y genera falsas alarmas diarias en la sección 3.3.1. Sin embargo, las simulaciones de adivinanzas aleatorias se realizarán utilizando tanto detecciones de precursores diarias como detecciones de precursores falsos. Para realizar las simulaciones de adivinanzas aleatorias, los parámetros del espacio ROC se definen de la siguiente manera:

- TP_{EQ} , número de precursores de EQ detectados
- FN_{EQ} , Número de precursores de EQ mal detectados
- $N_{Pos,EQ}$, número de EQS
- TP_{day} , número de precursores diarios detectados
- FN_{day} , número de precursores diarios mal detectados,
- $N_{Pos,day}$, número de días sísmicamente activos,
- FP_{day} , número de falsas alarmas diarias generadas,
- TN_{day} , Número de verdaderos negativos diarios,
- $N_{Neg,day}$, número de posibles días de falsa alarma.

TP_{EQ} , FN_{EQ} y $N_{Pos,EQ}$ son los parámetros del espacio ROC que están relacionados con la detección de un precursor de EQ. Los restantes,

TP_{day} , FN_{day} , $N_{Pos,day}$, FP_{day} , TN_{day} y $N_{Neg,day}$ son los parámetros del espacio ROC que están relacionados con las decisiones precursoras diarias. Tenga en cuenta que múltiples decisiones de precursores diarios pueden conducir a la detección de un solo precursor de EQ. Para los puntos ROC elegidos en la tabla IX, todos los parámetros del espacio ROC se presentan en la tabla X. El pseudocódigo de la simulación de adivinanzas aleatorias se presenta en el algoritmo 3. Para cada simulación de adivinanzas aleatorias, el rendimiento de la detección de precursores de EQ se calcula utilizando la puntuación del coeficiente de correlación de Matthews (MCC) definida de la siguiente manera, [117]:

$$MCC(TP, TN, FP, FN) = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (46)$$

La puntuación MCC es una métrica de medición unidimensional para puntos espaciales ROC bidimensionales. Es posible comparar la puntuación MCC de un punto ROC con un experimento de adivinanza aleatorio. $MCC = 1$ corresponde al caso en que se detectan todos los precursores de EQ sin generar una sola falsa alarma. $MCC = 0$ corresponde al caso en el que el rendimiento de detección de precursores de EQ no es mejor que la adivinación aleatoria. Finalmente, $MCC = -1$ corresponde al caso en el que se pierden todos los precursores de EQ y se generan falsas alarmas para cada día de clase de días sin actividad sísmica. Las puntuaciones MCC de los puntos ROC elegidos en la tabla IX se definen como MCCR y se presentan en la ec. 33.

Algoritmo 3

```
// Definición de parámetros //
 $N_{MC} \leftarrow$  Número de simulaciones de Monte Carlo
 $N_{Pos,EQ} \leftarrow$  Número de EQs
 $N_{Precursor} = TP_{day} + FP_{day} \leftarrow$  Número total de detecciones de precursores
 $N_{days} \leftarrow$  Número total de días
 $MCC_{vec} \leftarrow$  Un vector para almacenar las puntuaciones del coeficiente de correlación de Matthews

for n = (1 a  $N_{MC}$ ) do

// Inicialización //
 $P_{date} \leftarrow$  Asigne cada detección de precursor a una fecha única
 $EQ_{date} \leftarrow$  Asigne cada EQ a una fecha única
Definir días de clase de actividad sísmica con respecto a EQ
for d = (1 a  $N_{days}$ ) do
```

```

if  $d \notin$  Días de clase de actividad sísmica
     $N_{Neg,day} = N_{Neg,day} + 1$ 
end if
end for

// Encuentre precursores de ecalización detectados y falsas alarmas //
for EQ = (1 a  $EQ_{date}$ ) do
    for P = (1 a  $P_{date}$ ) do
        if P related to EQ then
             $TP_{EQ} = TP_{EQ} + 1$ 
            break
        end if
    end for
end for

for P = (1 a  $P_{date}$ ) do
    if P  $\notin$  días de clase de actividad sísmica then
         $FP_{day} = FP_{day} + 1$ 
    end if
end for

 $FN_{EQ} = N_{Pos,EQ} - TP_{EQ}$ 
 $TN_{day} = N_{Neg,day} - FP_{day}$ 
 $MCC_{vec}(n) = MCC_R(TP_{EQ}, FN_{EQ}, TN_{day}, FP_{day})$ 

end for

```

Además, el MCC_R de los puntos ROC relacionados y los histogramas de puntaje MCC de simulaciones de Monte Carlo de adivinación aleatoria para los puntos ROC de validación y prueba elegidos se presentan en la figura N°31 y la figura N°32, respectivamente. Las desviaciones media y estándar de estas simulaciones de adivinanzas aleatorias también se definen como μ_{RG} y σ_{RG} y se presentan en la tabla X. Como se presenta en la figura N°31 y la figura N°32. $MCC_R \geq \mu_{RG} + 2 \times \sigma_{RG}$ se mantiene para todos los MCC_R visualizados. Además, se obtendrá la puntuación Z de MCC_R con respecto a μ_{RG} y σ_{RG} . La puntuación Z resultante indicará la importancia estadística del rendimiento de detección de precursores de EQ del punto ROC relacionado y se presentará para cada ROC elegido punto en la tabla X. Como se presenta en la tabla X, todos los puntajes z son superiores a 2, por lo que el rendimiento de detección de precursores de EQ-PD es superior a las simulaciones de adivinación aleatoria para todos los puntos ROC elegidos.

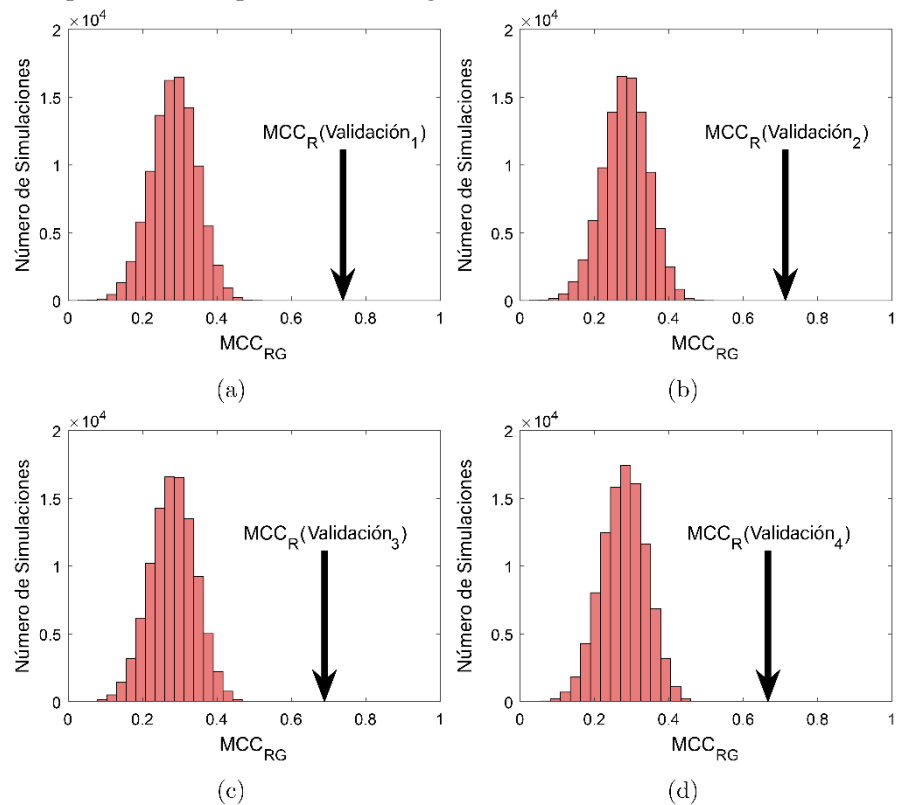


Figura N°31: Histogramas de puntuación de MCC obtenidos de simulaciones aleatorias de Monte Carlo para los puntos ROC elegidos: (a) Validación₁, (b) Validación₂, (c) Validación₃ y (d) Validación₄. La puntuación MCC del punto ROC elegido está marcada con una flecha en cada figura.

Fuente: Elaboración propia.

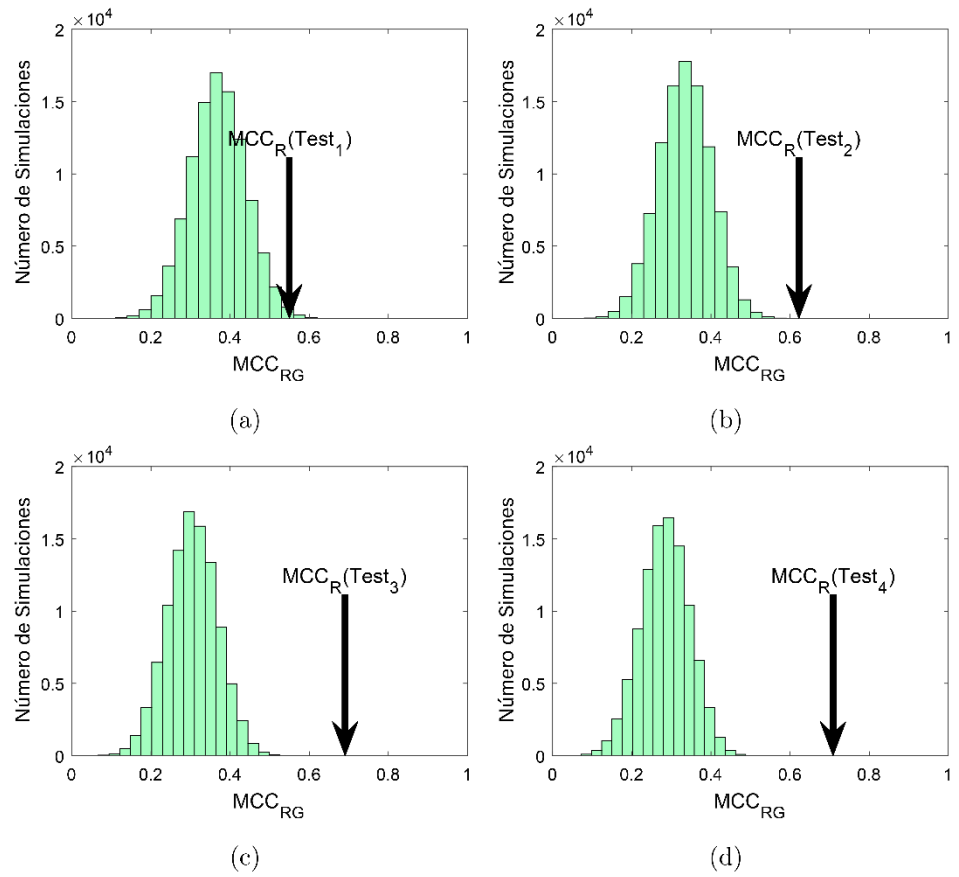


Figura N°32: Histogramas de puntuación de MCC obtenidos de simulaciones aleatorias de Monte Carlo para los puntos ROC elegidos: (a) Test₁, (b) Test₂, (c) Test₃ y (d) Test₄. La puntuación MCC del punto ROC elegido está marcada con una flecha en cada figura.

Fuente: Elaboración propia.

Tabla X

Para los puntos ROC de validación y prueba elegidos: TP_{EQ} , $N_{Pos,EQ}$, FN_{EQ} , TP_{day} , $N_{Pos,day}$, FN_{day} , FP_{day} , TN_{day} y $N_{Neg,days}$ parámetros de espacio ROC, media y desviación estándar de la puntuación MCC de los experimentos aleatorios μ_{RG} y σ_{RG} , puntuación MCC del punto ROC relacionado (MCC_R), score-Z de MCC_R .

Punto de ROC	Validation ₁	Test ₁	Validation ₂	Test ₂	Validation ₃	Test ₃	Validation ₄	Test ₄
$TP_{EQ}/N_{Pos,EQ}$	17/21	17/24	17/21	20/24	17/21	22/24	17/21	23/24
$FN_{EQ}/N_{Pos,EQ}$	4/21	7/24	4/21	4/24	4/21	2/24	4/21	1/24
$TP_{day}/N_{Pos,days}$	123/146	81/135	123/146	99/135	123/146	115/135	123/146	125/135
$FN_{day}/N_{Pos,days}$	23/146	54/135	23/146	36/135	23/146	20/135	23/146	10/135
$FP_{day}/N_{Neg,days}$	5/120	11/147	6/120	13/147	7/120	13/147	8/120	14/147
$TN_{day}/N_{Neg,days}$	115/120	136/147	114/120	134/147	113/120	134/147	112/120	133/147
$N_{Pos,days} + N_{Neg}$	266	282	266	282	266	282	266	282
Mean (μ_{RG}) MCC score (RG)	0.28476	0.36973	0.28279	0.33400	0.28092	0.30479	0.27894	0.28388
Std (σ_{RG}) MCC score (RG)	0.06053	0.07059	0.06039	0.06597	0.06014	0.06332	0.06008	0.06171
$MCC_R(TP_{EQ}, FN_{EQ}, TN_{day}, FP_{day})$	0.75335	0.59466	0.73187	0.65564	0.71163	0.71304	0.69248	0.72806
$Z = (MCC_R - \mu_{RG})/\sigma_{RG}$	7.7410	3.1859	7.4362	4.8751	7.1613	6.4471	6.8825	7.1969

Tabla XI

21 EQs diarios diferentes que han tenido lugar en la región elegida.

Sismo # - Fecha (# - Día Mes Año)	Tiempo (hh:mm:ss)	Epicentro del sismo		Mw (Richter)
		Latitud	Longitud	
1 – 11 diciembre 2023	13:58:23	15.25°S	75.61°O	4.3
2 – 15 diciembre 2023	12:53:07	14.55°S	76.61°O	4
3 – 14 diciembre 2023	10:55:50	14.45°S	76.49°O	4.1
4 – 14 diciembre 2023	02:21:44	15.21°S	75.9°O	4.1
5 – 10 diciembre 2023	07:14:12	14.41°S	76.57°O	4
5 – 10 diciembre 2023	10:43:28	14.35°S	76.52°O	4.1
5 – 10 diciembre 2023	07:14:12	14.41°S	76.57°O	4
5 – 10 diciembre 2023	06:52:46	14.39°S	76.67°O	5.8
5 – 10 diciembre 2023	06:40:30	14.39°S	76.51°O	5.3
5 – 10 diciembre 2023	02:35:12	14.39°S	76.63°O	4.5
6 – 09 diciembre 2023	22:38:19	14.36°S	76.58°O	4.6
7 – 06 diciembre 2023	02:06:20	14.87°S	76.61°O	4.3
8 – 18 noviembre 2023	23:16:06	15.19°S	75.8°O	5
8 – 18 noviembre 2023	17:39:05	13.78°S	75.9°O	3.7
9 – 03 noviembre 2023	10:05:35	14.07°S	75.84°O	4
10 – 15 octubre 2023	22:49:31	15.01°S	75.61°O	4
11 – 02 octubre 2023	11:24:02	14.1°S	76.07°O	4
12 – 25 septiembre 2023	20:30:43	15.2°S	75.08°O	4
13 – 22 septiembre 2023	16:17:51	14.61°S	75.52°O	4.2
14 – 14 agosto 2023	15:13:50	14.72°S	75.75°O	3.9
15 – 02 agosto 2023	02:55:31	14.76°S	75.83°O	4.3

3.4. Técnica de detección de precursores de sismos basada en aprendizaje profundo.

Estudios recientes reportan el hecho de que las anomalías TEC ionosféricas han estado ocurriendo antes de las fuertes actividades sísmicas y EQs [66, 67, 68]. En este estudio, se propone una técnica de detección de precursores de ecualización ionosférica basada en aprendizaje profundo, DL-PD, para detectar anomalías TEC desencadenadas por actividad sísmica. DL-PD tiene las siguientes contribuciones clave a la literatura de detección de precursores de EQ. Una de las contribuciones de la técnica DL-PD propuesta es que genera estimaciones GIM-TEC de forma adaptativa tanto espacial como temporalmente. Por lo tanto, la capacidad de estimación espacio-temporal de TEC de DL-PD es mucho más completa en comparación con las técnicas basadas en modelos. Otra contribución de la propuesta es que la técnica DL-PD es una técnica de aprendizaje automático no supervisado. DL-PD funcionará independientemente de la información de la actividad sísmica. Por lo tanto, genera estimaciones GIM-TEC esperadas que se compararán con el GIM-TEC observado para detectar anomalías espacio-temporales.

La contribución final a la literatura es que DL-PD funcionará con cualquier duración de tiempo y región del mundo elegidos. DL-PD requiere datos diarios de GIM-TEC para generar decisiones precursoras de ecualización diarias u horarias. En este trabajo, los datos

de GIM-TEC obtenidos del Laboratorio de Propulsión a Chorro de la NASA para la región que cubre la mayor parte de Europa, la región de Anatolia y Medio Oriente con una duración de casi 1,5 años se procesan para generar estimaciones de GIM-TEC. Las estimaciones generadas se comparan con los datos reales de GIM-TEC para obtener anomalías TEC ionosféricas. El rendimiento de detección de precursores de EQ de las decisiones de anomalía TEC generadas se evalúa durante 500 días consecutivos. Este trabajo consta de múltiples etapas de procesamiento.

3.6. Recopilación y procesamiento de datos. En este estudio, se procesan mapas ionosféricos globales basados en TEC (GIM-TEC). Los datos TEC de las estaciones de referencia de la red del servicio GNSS internacional (IGS) son preprocesados por Jet Propulsion Laboratory para generar datos GIM-TEC [118]. IONOLAB seguirá procesando los datos de GIM-TEC para mejorar su fiabilidad y solidez [70]. Los mapas GIM-TEC obtenidos tienen una resolución espacial de 1° de latitud y 1° de longitud con una resolución temporal de 15 minutos. Por lo tanto, se almacena un cubo de datos de $180 \times 360 \times 96$ dimensiones que contiene 96 mapas GIM-TEC temporalmente diferentes para cada día individual. Un ejemplo de GIM-TEC se presenta en la figura N°33 para una fecha elegida de 2023-08-19T11:00. Se accedió a los datos relacionados con EQ a través del catálogo completo de sismos del Sistema Nacional Sísmico Avanzado (ANSS ComCat) [107]. Los datos relacionados con EQ accedidos: número de EQ y fecha de EQ diferentes diariamente (primera columna), hora de EQ (segunda columna), epicentro de EQ (tercera y cuarta columnas) y tipo de magnitud de EQ (quinta columna) se presentan en la tabla XVII.

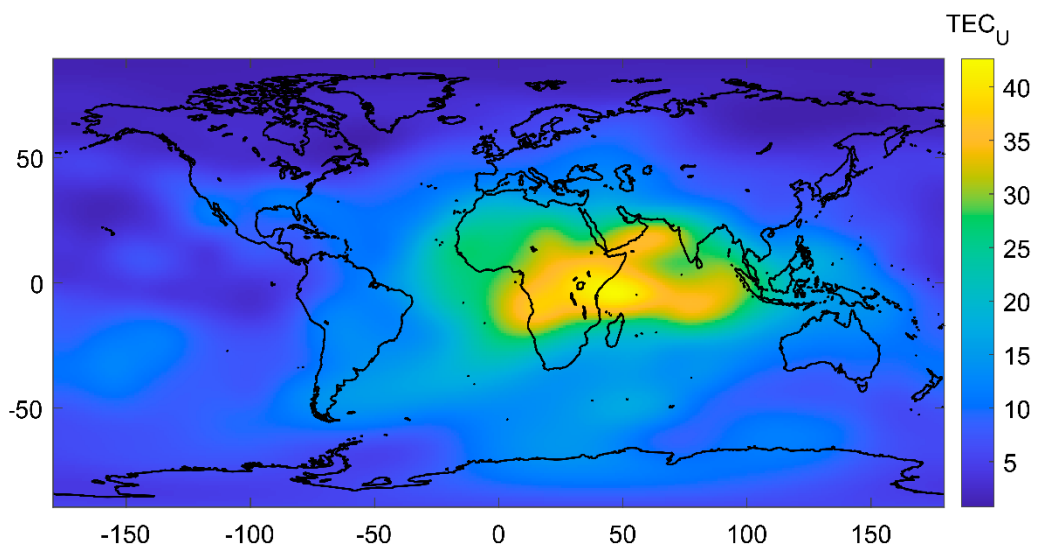


Figura N°33: Mapa GIM-TEC del mundo para la fecha: 2023-08-19T11:00.

Fuente: Elaboración propia.

3.6. Técnica de detección de precursores EQ basada en aprendizaje profundo: DL-PD. Los mapas GIM-TEC recopilados presentarán anomalías ionosféricas detectables que están

relacionadas con las próximas actividades sísmicas fuertes y EQ [44, 108]. En este trabajo, se propone la técnica de detección de precursores de EQ basada en aprendizaje profundo (DL) y GIM-TEC ionosférico, DL-PD. El propuesto detectará precursores de EQ diariamente utilizando los datos diarios de GIM-TEC. Los datos recopilados de GIM-TEC exhiben variaciones de TEC tanto espaciales como temporales. Por lo tanto, se requiere una técnica de detección de anomalías que se adapte a los datos espacio-temporales. La técnica DL-PD propuesta genera estimaciones espacio-temporales de mapas GIM-TEC que hacen posible la detección de anomalías ionosféricas basadas en TEC. Esta sección está organizada de la siguiente manera. En la sección 3.7.1, se presentan los pasos de preprocesamiento de los datos GIM-TEC obtenidos. En la sección 3.7.2, se presenta una red neuronal profunda con capacidades de estimación espacio-temporal. A partir de entonces, las detecciones de precursores de EQ se generan con respecto a las puntuaciones de anomalías ionosféricas, como se explica en la sección 3.7.3. El diagrama de flujo de la técnica DL-PD propuesta se visualiza en la figura N°34.

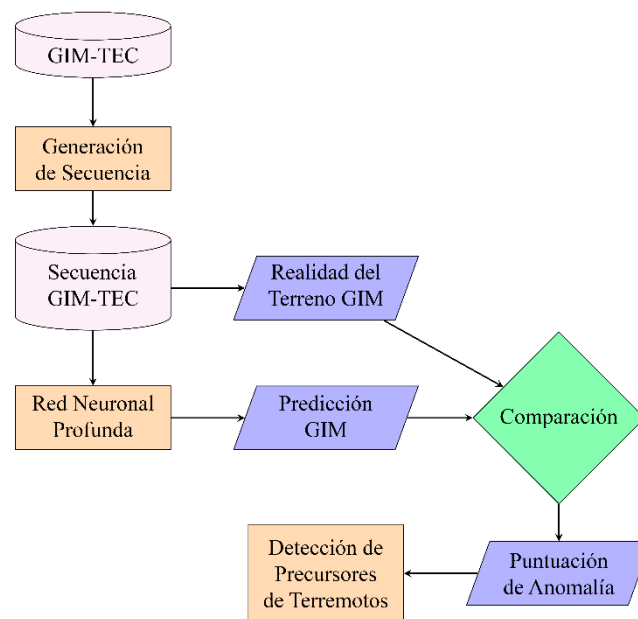


Figura N°34: Diagrama de flujo de la técnica propuesta de detección de precursores de sismos basada en aprendizaje profundo, DL-PD.

Fuente: Elaboración propia.

3.6.1. Preprocesamiento GIM. Cada mapa GIM-TEC se recorta espacialmente con respecto a un área elegida que se presenta en la figura N°35. El área elegida está limitada de -7,5° a 75,5° de longitud y de 22,5° a 57,5° de latitud. La región de Anatolia reside en el centro de esta área elegida y se extiende desde los 20,5° hasta los 47,5° de longitud y desde los 34,5° hasta los 45,5° de latitud. El mapa GIM-TEC del área elegida se ilustra en la figura N°35 para el mapa GIM-TEC de la fecha elegida en la figura N°33.

Los epicentros de EQ de los EQ con magnitudes mayores o iguales a 5,4 para las fechas 01/01/2000 y 31/12/2018 también se representan en la figura N°35.

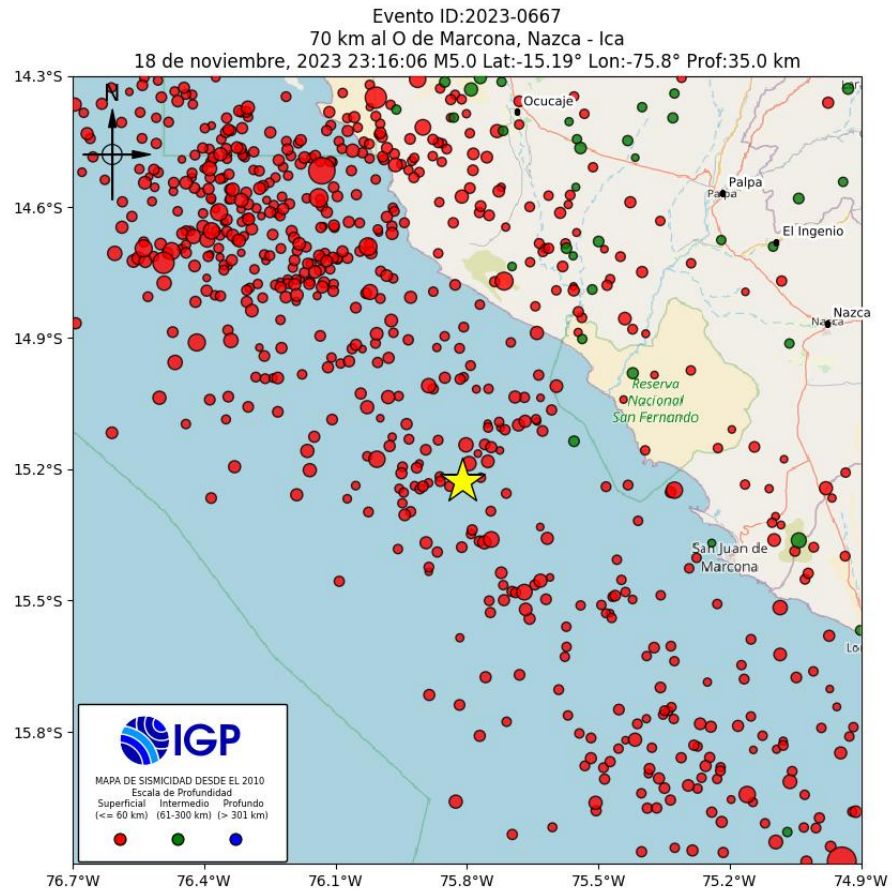


Figura N°35: Mapa recortado del mundo GIM-TEC para la fecha 2023-08-19T11:00. EQs con magnitudes mayores o iguales a 5.4 para las fechas 01.01.2000 y 31.12.2018.

La región de Anatolia está marcada con un rectángulo rojo discontinuo.

Fuente: Elaboración propia.

Como se muestra en la figura N°35, los datos GIM-TEC del área elegida corresponden a una matriz dimensional de 36×84 para cada período de tiempo. Los datos GIM-TEC por hora se generan promediando los datos GIM-TEC con 15 minutos de separación de tiempo y se generan datos GIM-TEC diarios de $36 \times 84 \times 24$ dimensiones. Cuando el número total de días se define como N_{days} . Los datos de GIM-TEC se representarán como un cubo de datos de 4 dimensiones con dimensiones $36 \times 84 \times 24 \times N_{days}$.

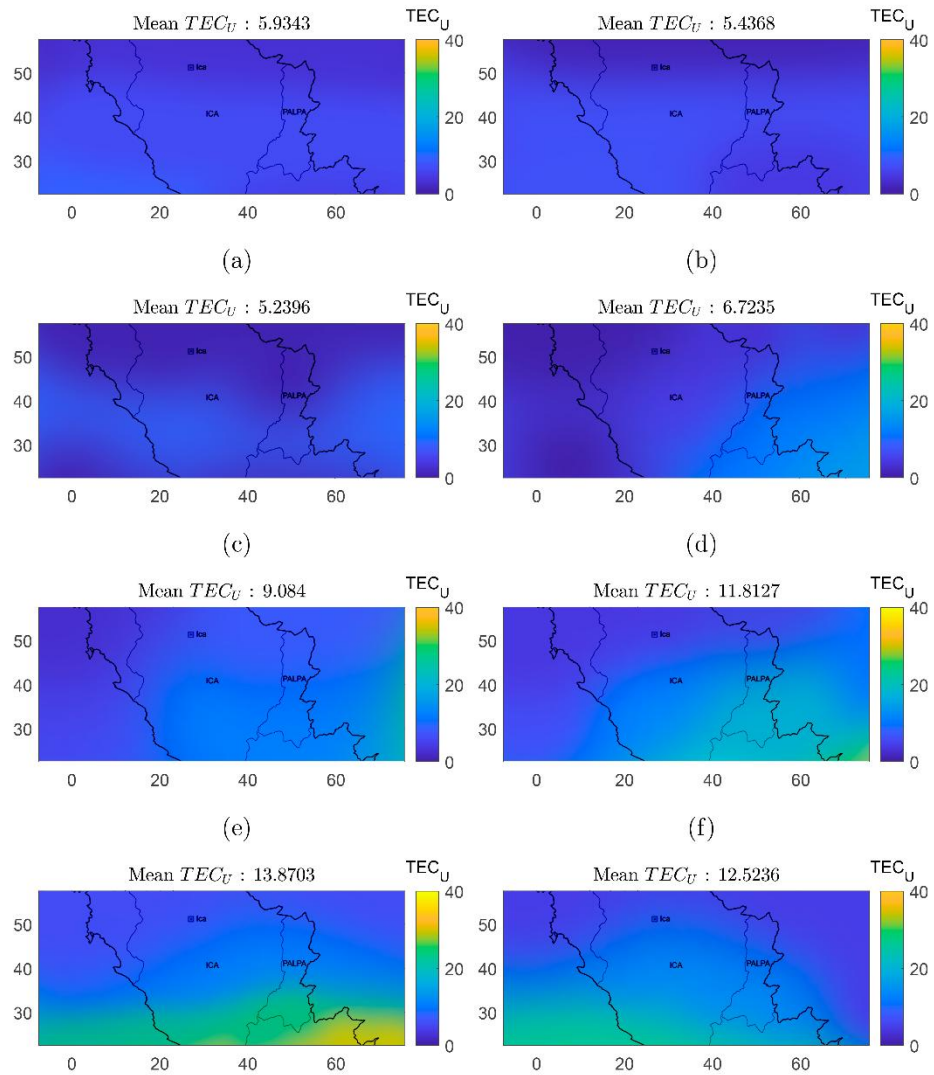


Figura N°36: (a) 2° (2023-08-18T22:00), (b) 4° (2023-08-19T00:00), (c) 6° (2023-08-19T02:00), (d) 8° (2023-08-19T04:00), (e) 10° (2023-08-19T06:00), (f) 12° (2023-08-19T08:00) y (g) 14° (2023-08-19T10:00) data frames GIM-TEC de una secuencia GIM-TEC de 15 pasos de tiempo para la fecha: 2023-08-19T11:00 (frame 15). (h) Ground truth de esta secuencia según GIM-TEC. (2023-08-19T12:00).

Fuente: Elaboración propia.

El entrenamiento de redes neuronales recurrentes requiere datos de entrada de modo que cada muestra de entrenamiento de los datos debe incluir una dimensión temporal o intervalos de tiempo. La dimensión de intervalos de tiempo define una secuencia de datos temporales para todas las muestras espaciales. Los datos GIM-TEC por hora se procesarán para generar secuencias GIM-TEC por hora. Cuando el número de pasos de tiempo se define como N_{time} , las tramas de datos GIM-TEC horarias de N_{time} anteriores se fusionarán para generar una secuencia GIM-TEC para una hora y fecha elegidas. Por lo tanto, una secuencia GIM-TEC se definirá como un cubo de datos tridimensional con $36 \times 84 \times N_{time}$ dimensiones. Además, el próximo marco de

datos GIM-TEC con una dimensión de 36×84 se etiquetará como verdad sobre el terreno para esta secuencia GIM-TEC generada. Los datos de secuencia de GIM-TEC generados por hora son un cubo de datos dimensional de $36 \times 84 \times N_{time} \times 24N_{days}$ con datos *ground truth* según GIM-TEC de $36 \times 84 \times 24N_{days}$ dimensionales. En la figura N°36, se visualiza un número par de marcos de datos GIM-TEC de una secuencia GIM-TEC de 15 pasos de tiempo para la fecha elegida en la figura N°33. La realidad básica de esta secuencia GIM-TEC también se visualiza en la figura 36h [113]. Cada secuencia horaria GIM-TEC generada se normaliza con respecto al TEC_U máximo registrado durante N_{time} duración y el valor mínimo constante de 0. La verdad de campo GIM-TEC correspondiente de estas secuencias generadas también se normaliza con respecto a los mismos valores TEC_U mínimo y máximo.

3.6.2. Predicción GIM basada en redes neuronales profundas. Se propone un modelo de red neuronal de aprendizaje profundo para estimar el próximo GIM-TEC por hora mediante el uso de una secuencia GIM-TEC con N_{time} muestras temporales y 36×84 muestras espaciales. El modelo propuesto se presenta en la figura N°37. Como se muestra en la figura N°37, el modelo propuesto está compuesto por capas de entrada, LSTM convolucional (ConvLSTM) y red neuronal convolucional 2-D con un total de 282 661 parámetros entrenables. La capa ConvLSTM es un tipo de capa LSTM que adoptará datos espacio-temporales como GIM-TEC [120], [121]. La capa ConvLSTM cuenta con una celda de memoria para el tiempo t , c_t que acumula la información de estados de la capa. Las puertas de control deben acceder, borrar o borrar la memoria de la celda.

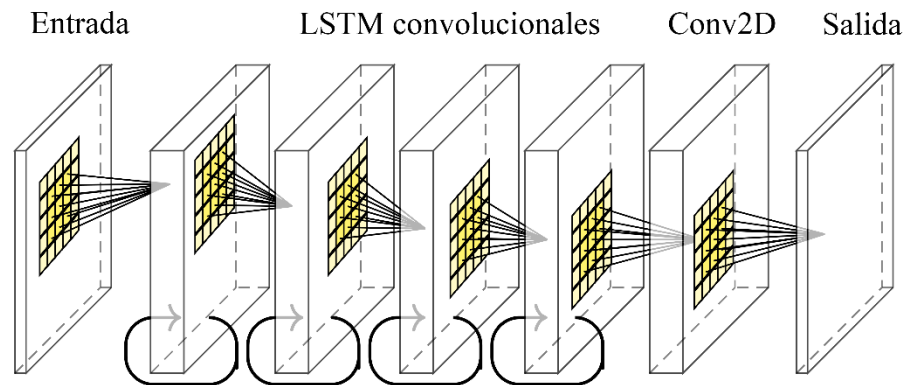


Figura N°37: Propuesta de red neuronal profunda con múltiples capas espacio-temporales ConvLSTM.

Fuente: Elaboración propia.

Cada una de estas puertas tendrá sus propios parámetros de entrenamiento. Cuando se alimenta una nueva muestra de entrenamiento al ConvLSTM, se activa el *input gate* i_t . *Forget gate*, f_t eliminará el estado de celda anterior c_{t-1} durante el entrenamiento.

Output gate, o_t decide si el último estado de celda c_t debe propagarse a través del estado oculto final h_t o no. Entrada x_1, \dots, x_t , estado de celda $c_1 \dots c_t$, estados ocultos $h_1 \dots h_t$, puertas de control i_t, f_t y o_t contienen información espacio-temporal 3D para cada marco de tiempo t . Las ecuaciones de activaciones de estado de celda se presentan en (ec. 33).

$$\begin{aligned}
 i_t &= \sigma(W_{x,i} * x_t + W_{h,i} * h_{t-1} + W_{c,i} \circ c_{t-1} + b_i) \\
 f_t &= \sigma(W_{x,f} * x_t + W_{h,f} * h_{t-1} + W_{c,f} \circ c_{t-1} + b_f) \\
 o_t &= \sigma(W_{x,o} * x_t + W_{h,o} * h_{t-1} + W_{c,o} \circ c_{t-1} + b_o) \\
 c_t &= f_t \circ c_{t-1} + i_t \circ \tanh(W_{c,c} * x_t + W_{h,c} * h_{t-1} + b_c) \\
 h_t &= o_t \circ \tanh(c_t)
 \end{aligned} \tag{47}$$

donde $*$ denota convolución, \circ denota producto de Hadamard, b_i, b_f, b_o denotan sesgos de *input*, *forget* y *output gate*. $W_{x,i}, W_{x,f}, W_{x,o}, W_{h,i}, W_{h,f}, W_{h,o}, W_{c,c}$ y $W_{h,c}$ denotan los núcleos de convolución para el *control gate* relacionada. Un ejemplo de capa ConvLSTM se visualiza en la figura N°38 con núcleos de convolución se representan como W_x y W_h por simplicidad.

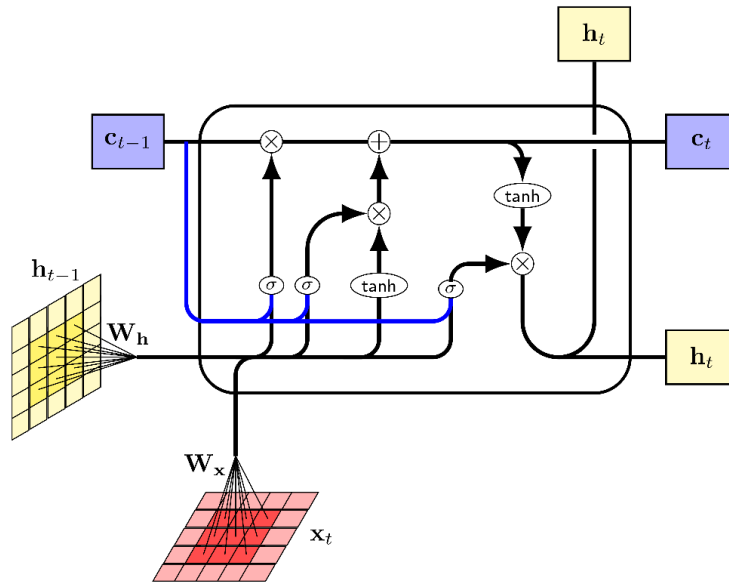


Figura N°38: La capa ConvLSTM con sus entradas espacio-temporales x_t controla las puertas i_t, f_t, o_t , los estados ocultos h_t y los estados de celda c_t . Las conexiones de mirilla también se visualizan en azul.

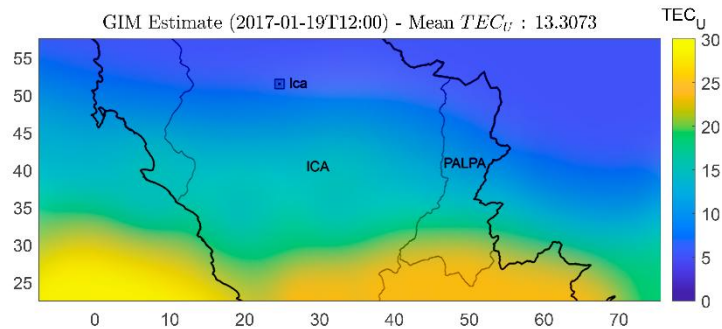
Fuente: Elaboración propia.

Cada capa ConvLSTM de la red propuesta contiene 20 filtros de convolución diferentes, cada uno de estos filtros tiene una dimensión de 5×5 . Las primeras tres capas de ConvLSTM generan salidas como secuencias, mientras que la última capa

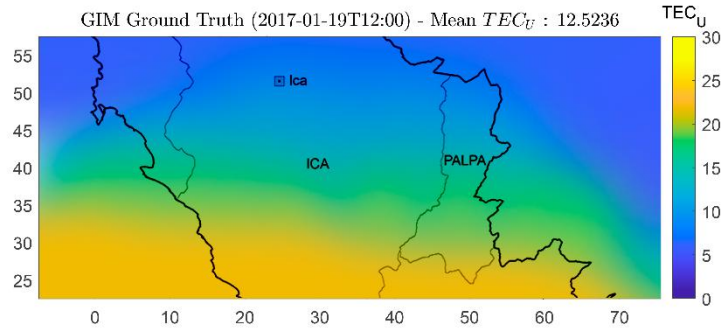
de ConvLSTM solo devuelve las salidas del último período de tiempo. La capa Conv2D contiene solo un filtro de convolución de tamaño 3×3 y tiene una función de activación sigmoidea. La forma de salida de cada capa que se presenta en la figura N°37 también se presenta en la tabla XII. Como se muestra en la tabla XII, la red neuronal profunda propuesta tomará secuencias GIM-TEC y generará una verdad básica GIM-TEC para un total de $N_{train, batch}$ muestras. Además, la normalización por lotes se realiza después de cada capa ConvLSTM para reducir el cambio de covariable interno entre los intervalos de tiempo [122].

Tabla XII
Para la red propuesta, formas de salida de cada capa. El número de muestras de entrenamiento se representa como $N_{train, batch}$.

Capa	Forma de salida
Entrada	$N_{train, batch} \times N_{time} \times 36 \times 86 \times 1$
ConvLSTM ₁ , ..., ConvLSTM ₃	$N_{train, batch} \times N_{time} \times 36 \times 86 \times 20$
ConvLSTM ₄	$N_{train, batch} \times 36 \times 86 \times 20$
Conv2D	$N_{train, batch} \times 36 \times 86 \times 1$



(a)



(b)

Figura N°39: Estimación de GIM-TEC y verdad de campo de GIM-TEC para la secuencia de GIM-TEC presentada en la figura N°36.

Fuente: Elaboración propia.

El rendimiento de la estimación GIM-TEC de la técnica propuesta se visualiza en la figura N°39 para la secuencia GIM TEC presentada en la figura N°36. Como se ilustra en la figura N°39, el GIM-TEC estimado se asemeja a la verdad del terreno de GIM-TEC de la próxima hora en que el RMSE entre la *ground truth* de GIM-TEC y la estimación de GIM-TEC es 1.0065 TEC_U . Tenga en cuenta que la red neuronal profunda propuesta se entrenará sin supervisión. Por lo tanto, no se requieren etiquetas de clase explícitas durante el entrenamiento del propuesto.

3.6.3. Detector de precursor de sismos. La detección de una anomalía de GIM-TEC requiere comparar los datos reales de GIM-TEC con la norma o, por ejemplo, las estimaciones de GIM-TEC. El resultado de esta comparación es la puntuación de anomalía de los datos GIM-TEC comparados. La puntuación de anomalía entre estos datos GIM-TEC se calculará utilizando la Medida del índice de similitud estructural (SSIM) [123]. El SSIM de dos imágenes I_1 y I_2 se calculará multiplicando los términos de luminancia, $l(I_1, I_2)$, contraste, $c(I_1, I_2)$ y estructura, $s(I_1, I_2)$ de la siguiente manera:

$$l(I_1, I_2) = \frac{2\mu_{I_1}\mu_{I_2} + K_1}{\mu_{I_1}^2 + \mu_{I_2}^2 + K_1} \quad (48)$$

$$c(I_1, I_2) = \frac{2\sigma_{I_1}\sigma_{I_2} + K_2}{\sigma_{I_1} + \sigma_{I_2} + K_2} \quad (49)$$

$$s(I_1, I_2) = \frac{2\sigma_{I_1 I_2} + K_3}{\sigma_{I_1}^2 \sigma_{I_2}^2 + K_3} \quad (50)$$

$$SSIM(I_1, I_2) = l(I_1, I_2) \times c(I_1, I_2) \times s(I_1, I_2) \quad (51)$$

donde μ_{I_1}, μ_{I_2} son los valores medios de píxeles, $\sigma_{I_1}, \sigma_{I_2}$ son las desviaciones estándar de píxeles y $\sigma_{I_1 I_2}$ es la covarianza cruzada para las imágenes I_1 y I_2 . K_1, K_2 y K_3 son las constantes para estabilizar la división por denominador débil. $K_1 = (0.01L)^2, K_2 = (0.03L)^2$ y $K_3 = K_2/2$. L es el rango dinámico de la imagen y para una secuencia GIM-TEC normalizada que cambia entre $[0, 1]$, el valor del parámetro L se elegirá como 1.

El índice SSIM es una medida de similitud que satisface las condiciones de simetría $SSIM(I_1, I_2) = SSIM(I_2, I_1)$, acotación $SSIM(I_1, I_2) \leq 1$ y máximo único $SSIM(I_1, I_2) = 1 \Leftrightarrow I_1 = I_2$. Por lo tanto, el valor mínimo del índice SSIM es 0 y el valor máximo del índice SSIM es 1. El índice SSIM de la estimación GIM-TEC y la verdad del terreno GIM-TEC presentados en la figura N°39 se obtendrá como 0,9566. Cuando el valor del índice SSIM es cercano a 1, la similitud entre la estimación de GIM-TEC y la realidad del terreno de GIM-TEC es alta.

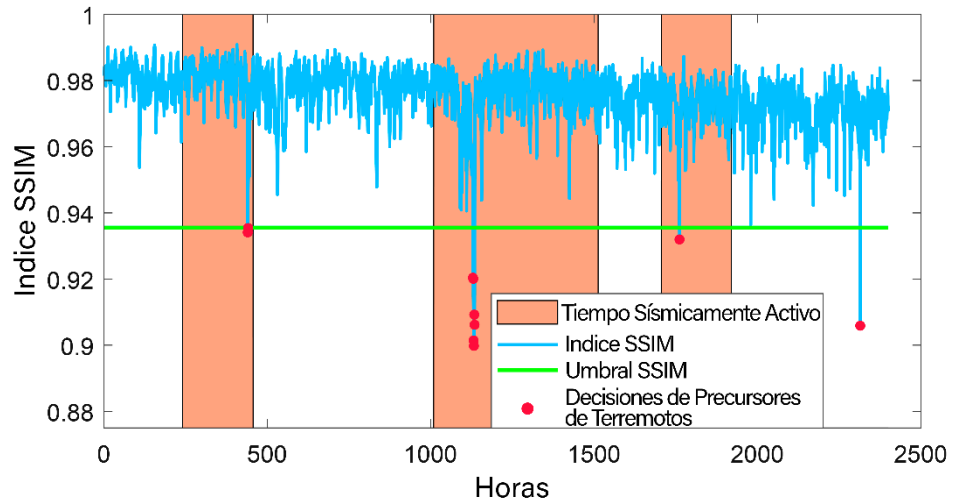


Figura N°40: Tiempo de actividad sísmica, índice SSIM, umbral SSIM y decisiones precursoras de EQ para los últimos 100 días de 2023.

Fuente: Elaboración propia.

Para investigar la relación de las anomalías GIM-TEC y los precusores de EQ, los últimos 100 días de 2023 se dividen en días de actividad sísmica y clases de días sin actividad sísmica. Los días de actividad sísmica se separan respecto a 6 EQs con magnitudes mayores o iguales a 5 y 8 días previos a estos diarios se seleccionan diferentes EQs inclusive con los días EQ para los días de actividad sísmica. La red neuronal basada en DL propuesta genera estimaciones GIM-TEC por hora. Por lo tanto, las horas de cada día de actividad sísmica también se etiquetan como tiempo de actividad sísmica y se visualizan con el índice SSIM entre las verdades del terreno GIM-TEC y las estimaciones para los últimos 100 días de 2023 en la figura N°40. Como se ilustra en la figura N°40, se elegirá un umbral SSIM para generar decisiones precursoras de EQ. Los índices SSIM que son más pequeños que el umbral elegido, se etiquetarán como anomalías GIM-TEC. Las anomalías GIM-TEC generadas se asociarán con los tiempos de actividad sísmica en que las anomalías GIM-TEC resultantes se denominarán decisiones precursoras de EQ. Las decisiones de precursor de EQ visualizadas en la figura N°40 detectan 4 de 6 precusores de EQ mientras generan 1 decisión de precursor de EQ falsa de 61 días sin actividad sísmica.

3.7. Rendimiento de la técnica propuesta de detección de precursores de EQ basada en DL.

El rendimiento de detección de precursores de la técnica DL-PD propuesta se investiga en datos GIM-TEC de 500 días de duración para la región elegida presentada en la figura N°35. Las fechas de inicio y finalización, el número de días y el número de EQ diferentes diarios en este período de tiempo se presentan en la tabla XIII. Las decisiones precursoras de EQ se generan para los EQ en la región de Anatolia como se presenta en la figura N°35. A medida que la región espacial se haga más grande, las anomalías GIM-TEC se detectarán de forma más fiable. Sin embargo, las regiones espaciales más grandes pueden dar lugar a EQ simultáneos o consecutivos en los que el número de días con actividad sísmica será mucho mayor que el número de días sin actividad sísmica. Por lo tanto, la evaluación del rendimiento se realiza con datos GIM-TEC de una región elegida más grande con epicentros EQ de una región de Anatolia más pequeña.

Tabla XIII
Información de datos GIM-TEC procesada.

Fecha (día mes año)		Número de días	Número de EQs diarios diferentes
Inicio	Fin		
19 de agosto de 2023	31 diciembre de 2023	500	32

Durante la duración del tiempo elegido y la región de Anatolia, se habían realizado 32 EQ diferentes diarios y se presentan en la tabla XVII. Todos los EQs presentados tenían una magnitud superior a 5 en la escala de Richter [107]. Si se han realizado varios EQ en el mismo día, se elige el EQ con la mayor magnitud para la evaluación del rendimiento. Como se muestra en la tabla XVII, se investigan diariamente 32 EQ diferentes de un total de 37 EQ. Como se detalla en la sección 3.7.1, los datos GIM-TEC obtenidos se procesan previamente para generar secuencias GIM-TEC por hora para la duración elegida de 500 días. El número de pasos de tiempo N_{time} se elige como 15 que generaron datos de secuencia de GIM-TEC y los datos *ground truth* de GIM-TEC tienen las dimensiones de $36 \times 84 \times 15 \times 12000$ y $36 \times 84 \times 12000$, respectivamente. tenga en cuenta que los datos de secuencia GIM-TEC generados tienen un tamaño de $\approx 2 GB$ cuando cada píxel GIM se representa con un número de coma flotante de precisión única (4 bytes). Para la misma cantidad de pasos de tiempo y días de duración, el tamaño de los datos de las secuencias GIM-TEC de todo el mundo será $\approx 43 GB$. La red neuronal profunda que se propone en la sección 3.7.2 está entrenada para las secuencias GIM-TEC generadas y los datos de verdad del terreno GIM-TEC de forma no supervisada. Los parámetros de entrenamiento de la propuesta se tabulan en la tabla XIV. El modelo propuesto se implementa en el lenguaje de

programación Python y el modelo propuesto se entrena en una computadora con una sola GPU NVIDIA Quadro RTX 8000.

Tabla XIV
Parámetros de entrenamiento del modelo.

Parámetro de entrenamiento	Valor
Tasa de aprendizaje	0.0002, 0.0005, 0.001
Porcentaje de validación	%10
Número de épocas	200
Tamaño del lote	32
Optimizador	Adam($\beta_1 = 0.9, \beta_2 = 0.999$)
Pérdida	L2 al cuadrado
Parada temprana	Decaimiento de pérdida de validación, paciencia = 20
Métrica de rendimiento	PSNR

Como se muestra en la tabla XIV, la Relación pico-signal a ruido (*peak signal-to-noise ratio* – PSNR, en inglés) se elige cómo la métrica de evaluación de rendimiento adecuada. PSNR es una relación entre la potencia máxima posible de los píxeles de la imagen, MAX_I y el error cuadrático medio, $MSE(I_1, I_2)$ entre las imágenes comparadas I_1, I_2 y definidas a continuación:

$$PSNR(I_1, I_2) = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE(I_1, I_2)} \right) \quad (52)$$

Tanto el índice SSIM como el PSNR son métricas de evaluación de calidad de imagen ampliamente utilizadas con diferentes sensibilidades de calidad de imagen [124]. Los registros de entrenamiento de la red propuesta se presentan en la figura N°41.

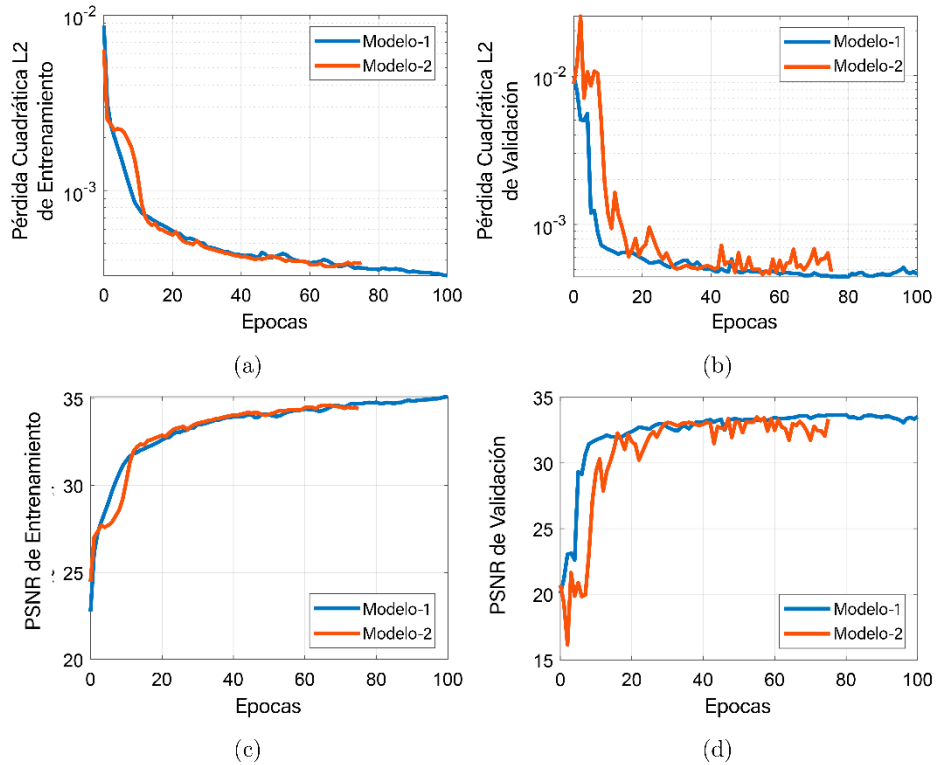


Figura N°41: Registros de entrenamiento del modelo propuesto (a) Pérdida de entrenamiento, (b) Pérdida de validación, (c) Entrenamiento PSNR y (d) Validación PSNR. Fuente: Elaboración propia.

En la figura N°41, los modelos cuyo valor de PSNR de validación supera los 32 se eligen y visualizan entre todos los modelos entrenados. Además, el Modelo-1 y el Modelo-2 se comparan para la validación. El Modelo-1 es superior al Modelo-2, ya que logra pérdidas de entrenamiento y validación más bajas y valores de PSNR más altos durante el entrenamiento y la validación. Por lo tanto, Modelo-1 es el modelo validado para evaluaciones de desempeño.

Tenga en cuenta que los datos de GIM-TEC que se presentan en la tabla XIII también se dividen en 3 intervalos de entrenamiento, validación y prueba que no se cruzan. Los últimos 15 días del 2023 están separados para el conjunto de datos de prueba. Los 475 días restantes se utilizan para el entrenamiento con una división de validación del 10%. El valor medio del índice SSIM para los conjuntos de datos de entrenamiento y validación es 0,9723. De manera similar, el índice SSIM medio es 0,9653 para el conjunto de datos de prueba. Las evaluaciones del rendimiento de detección de precursores de EQ se realizan para cada día en la tabla XIII. Durante estas evaluaciones, los días no se dividen con respecto a sus respectivos conjuntos de datos, ya que los rendimientos de estimación de GIM-TEC de los conjuntos de

datos de entrenamiento, validación y prueba son muy similares entre sí. Como se detalla en la sección 3.7.3, las decisiones de precusores de EQ se generan para las fechas presentadas en la tabla XIII y las EQ presentadas en la tabla V mediante el umbral del índice SSIM que sirve como una puntuación de anomalía GIM-TEC. Las curvas ROC de detección de precusores de EQ resultantes se presentan en la figura N°42.

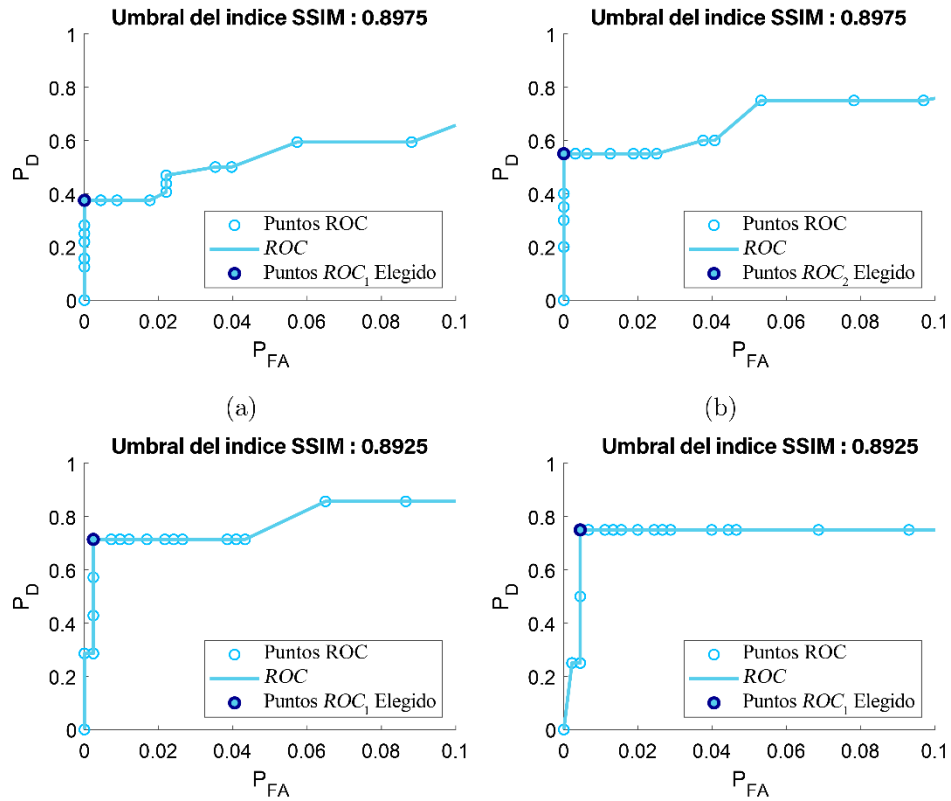


Figura N°42: Curvas ROC basadas en el rendimiento de detección de precusores de EQ para EQ con magnitudes (M) (a) $M \geq 5$, (b) $M \geq 5.2$, (c) $M \geq 5.4$ y (d) $M \geq 5.6$. Puntos ROC elegidos (a) ROC_1 , (b) ROC_2 , (c) ROC_3 y (d) ROC_4 .

Fuente: Elaboración propia.

Cuando las curvas ROC generadas en la figura N°42, estas curvas exhiben secciones lineales escalonadas y de tornillo de banco en el eje P_D . Las secciones escalonadas en el eje P_D deben atribuirse al hecho de que los valores de P_D se calculan con respecto a un pequeño número de precusores de sismos. Como se presenta en la tabla XV, los rendimientos de detección de precusores se evalúan para 32 $M \geq 5$ precusores EQ, 20 $M \geq 5.2$ precusores EQ, 7 $M \geq 5.4$ precusores EQ y 4 $M \geq 5.6$ precusores EQ. Por lo tanto, P_D tomará 32, 20, 7 y 4 valores distintos de cero durante las evaluaciones de desempeño en la figura N°42a, figura N°42b, la figura N°42c y figura N°42d, respectivamente. Como se muestra en la figura N°42, las secciones lineales de tornillo de banco en el eje P_D se deben atribuir al hecho de que los puntos ROC operativos se interpolan para generar las curvas ROC. Por lo tanto, las secciones lineales de tornillo de banco de las curvas ROC no son operativas excepto los vértices de estas secciones lineales. Finalmente, las curvas ROC generadas se visualizan para $0 \leq P_D \leq$

1 y $0 \leq P_{FA} \leq 0.1$, de modo que los rendimientos de los puntos ROC con $P_{FA} \geq 0.1$ no se tienen en cuenta durante la visualización. Tenga en cuenta que el número de posibles decisiones de precursores falsos es mayor que el número de todos los precursores de EQ cuando $P_{FA} = 0.1$ y la magnitud mínima de EQ $M \geq 5.2$. Como se muestra en la figura N°43, se obtienen cuatro curvas ROC diferentes para 4 grupos diferentes de EQ y 4 puntos ROC diferentes, ROC_1, ROC_2, ROC_3 y ROC_4 , respectivamente. El número de precursores de EQ detectados, el número de todos los EQ, el número de decisiones de precursores falsos, el número de días sin actividad sísmica se presentan para los puntos ROC elegidos en la tabla XV. Como se presenta tanto en la figura N°42 como en la tabla XV, la técnica DL-PD propuesta detecta 5 de 7 precursores de EQ mientras genera 1 decisión falsa de precursor de EQ de 416 días sin actividad sísmica. El umbral de SSIM se seleccionará entre 0,90 y 0,89 para etiquetar una estimación de GIM-TEC como una anomalía. Además, la probabilidad de detección de precursores aumenta a medida que aumenta la magnitud de los ecualizadores deseados, mientras que las decisiones de precursores de ecualizadores falsos generados siguen siendo bajas. Por tanto, la propuesta alcanza su mayor rendimiento cuando se detectan los precursores de EQ que pertenecen a EQs con magnitudes mayores o iguales a 5.4. Tenga en cuenta que, el posicionamiento confiable de las ubicaciones de detección de precursores requiere más técnicas de procesamiento espacial. Para correlacionar la detección de un precursor con un epicentro de EQ específico, las variaciones espaciales GIM-TEC desencadenadas por la actividad sísmica deben procesarse y detectarse alrededor del epicentro de EQ. Dicha técnica de detección requiere el procesamiento de múltiples regiones GIM-TEC de forma independiente y la combinación de sus puntuaciones de anomalías basadas en SSIM para posicionar la ubicación precisa de detección de precursores. En este trabajo, dicha técnica de procesamiento espacial no se implementa y se deja como un trabajo futuro. En su lugar, se implementa una técnica de posicionamiento de distancia de detección de precursor basada en diferencia de TEC simple.

Cada precursor de EQ es detectado por decisiones de precursor anteriores. La ubicación de la decisión precursora se obtendrá comparando GIM-Estimate con la verdad de GIM Ground de la fecha en que se genera la decisión precursora. El píxel de ubicación de decisión precursora se ubica obteniendo un píxel que tiene la mayor diferencia de TEC entre las imágenes reales de GIM-Estimate y GIM Ground. La ubicación de decisión del precursor se obtendrá convirtiendo esta ubicación de píxel en coordenadas geodésicas. La distancia de detección del precursor, d_{prec} para un EQ, se calculará midiendo la distancia entre el epicentro del EQ y la información sobre la ubicación de la decisión del precursor. Además, las decisiones de múltiples precursores contribuirán a la detección de un precursor de EQ. Por lo tanto, también se obtendrán múltiples distancias de detección de precursores. Cuando

se generan múltiples distancias de detección de precursores, la media de estas distancias, $\sum d_{prec}$ representa la distancia de detección del precursor. Las distancias medias de detección de precursores, $\sum d_{prec}$ y la desviación estándar de estas distancias de detección de precursores, $\sigma(\sum d_{prec})$ se presentan en la tabla XVI. Como se muestra en la tabla XVI, los precursores de EQ con $M \geq 5.4$ se detectarán desde distancias más largas con desviaciones estándar de distancia de detección de precursores más pequeñas.

Tabla XV

Para los puntos ROC elegidos: número de precursores de EQ detectados, número de todos los EQ, número de decisiones de precursores falsos, número de decisiones de precursores falsos que resultaron en la detección de precursores de EQ más débiles, número de días sin actividad sísmica (días de falsa alarma), SSIM umbral de índice y magnitud mínima de EQ en la escala de Richter.

ROC point	Detected EQ Precursor	Number of EQs	False Prec. Detections	Weaker EQ Prec. Detections	No-seismic activity days	SSIM index Threshold	Min EQ magnitude
ROC ₁	12	32	0	0	227	0.8975	5
ROC ₂	11	20	0	0	320	0.8975	5.2
ROC ₃	5	7	1	1 $M \geq 5.2$	416	0.8925	5.4
ROC ₄	3	4	2	1 $M \geq 5.2$ 1 $M \geq 5.4$	452	0.89	5.6

Tabla XVI

Para los puntos ROC elegidos: distancias medias de detección de precursores de EQ ($\sum d_{prec}$), desviación estándar de las distancias de detección de precursores de EQ ($\sigma(\sum d_{prec})$) y magnitud mínima de EQ en la escala de Richter.

ROC point	Distancia de detección del precursor (km) $\sum d_{prec}$	Distancia de detección del precursor (km) $\sigma(\sum d_{prec})$	Magnitud mínima del EQ
ROC1	2042.8887	833.9594	5
ROC2	2031.674	916.2285	5.2
ROC3	2300.6086	632.8927	5.4
ROC4	2235.5288	807.1093	5.6

El número de decisiones de precursores falsos generados que dan como resultado la detección de precursores de EQ más débiles se investiga más a fondo para los puntos ROC elegidos. Los puntos ROC_1 y ROC_2 no generan decisiones precursoras falsas. Hay una decisión precursora falsa del punto ROC_3 . Como se presenta en la tabla XV, la decisión de precursor es una decisión de precursor falso para los EQ con $M \geq 5.4$ mientras que es una detección de precursor para los EQ con $M \geq 5.2$. Del mismo modo, 2 decisiones de precursores falsos de ROC_4 dan como resultado la detección de 1 precursor de EQ con EQ $M \geq 5.4$ y otro precursor de EQ con EQ $M \geq 5.2$. Por lo tanto, estas decisiones falsas de precursores de EQ se atribuirán a precursores de EQ que pertenecen a EQ más débiles.

Además, las decisiones precursoras de EQ del punto ROC elegido, ROC, se visualizan en la figura N°43. Como se ilustra en la figura N°43, hay 7 períodos de tiempo sísmicamente activos distintos que representan 7 EQ diferentes con $M \geq 5.4$. Las decisiones precursoras falsas entre estas decisiones precursoras también se visualizan en la figura N°44. Hay tres decisiones precursoras falsas diferentes por hora que son 7, 10 y 13 horas antes del tiempo sísmicamente activo del EQ que tuvo lugar el 20 de agosto de 2023 con una magnitud $M = 6.6$ alrededor de la región de Anatolia. Dado que estas decisiones falsas de precursor de EQ por hora se toman el mismo día, se marcan como 1 decisión falsa diaria de precursor de EQ para el punto ROC_3 en la tabla XV.

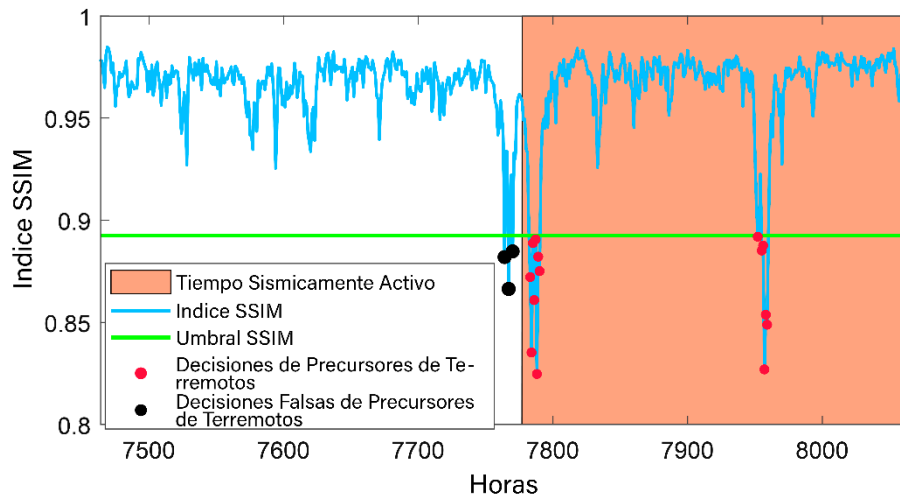


Figura N°43: Decisiones de tiempo de actividad sísmica, índice SSIM, umbral SSIM y precursor de EQ para los últimos 500 días de 2023. El tiempo sísmicamente activo se genera con respecto a 7 EQ diferentes con $M \geq 5.4$.

Fuente: Elaboración propia.

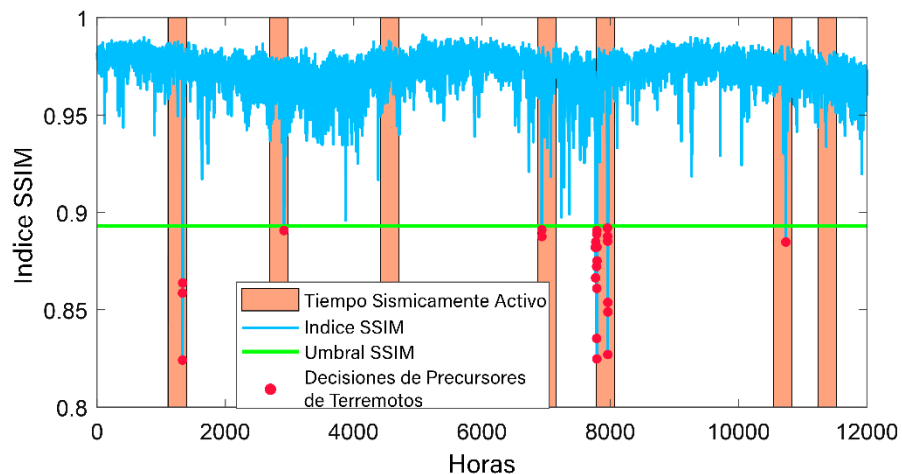


Figura N°44: Tiempo de actividad sísmica, índice SSIM, umbral SSIM, detecciones de precursoras de EQ y decisiones falsas de precursoras de EQ durante los últimos 25 días antes del EQ que tuvo lugar el 20 de julio de 2013 con una magnitud $M = 6.6$ en la región de Anatolia.

Fuente: Elaboración propia.

Tabla XVII
21 EQ diferentes diarios que han tenido lugar en la región de Ica.

Sismo - Fecha (Día Mes Año)	Tiempo (hh:mm:ss)	Epicentro del sismo		Mw (Richter)
		Latitud	Longitud	
1 – 11 diciembre 2023	13:58:23	15.25°S	75.61°O	4.3
2 – 15 diciembre 2023	12:53:07	14.55°S	76.61°O	4
3 – 14 diciembre 2023	10:55:50	14.45°S	76.49°O	4.1
4 – 14 diciembre 2023	02:21:44	15.21°S	75.9°O	4.1
5 – 10 diciembre 2023	07:14:12	14.41°S	76.57°O	4
5 – 10 diciembre 2023	10:43:28	14.35°S	76.52°O	4.1
5 – 10 diciembre 2023	07:14:12	14.41°S	76.57°O	4
5 – 10 diciembre 2023	06:52:46	14.39°S	76.67°O	5.8
5 – 10 diciembre 2023	06:40:30	14.39°S	76.51°O	5.3
5 – 10 diciembre 2023	02:35:12	14.39°S	76.63°O	4.5
6 – 09 diciembre 2023	22:38:19	14.36°S	76.58°O	4.6
7 – 06 diciembre 2023	02:06:20	14.87°S	76.61°O	4.3
8 – 18 noviembre 2023	23:16:06	15.19°S	75.8°O	5
8 – 18 noviembre 2023	17:39:05	13.78°S	75.9°O	3.7
9 – 03 noviembre 2023	10:05:35	14.07°S	75.84°O	4
10 – 15 octubre 2023	22:49:31	15.01°S	75.61°O	4
11 – 02 octubre 2023	11:24:02	14.1°S	76.07°O	4
12 – 25 septiembre 2023	20:30:43	15.2°S	75.08°O	4
13 – 22 septiembre 2023	16:17:51	14.61°S	75.52°O	4.2
14 – 14 agosto 2023	15:13:50	14.72°S	75.75°O	3.9
15 – 02 agosto 2023	02:55:31	14.76°S	75.83°O	4.3

IV. DISCUSIÓN

La investigación actual presenta un enfoque basado en redes neuronales convolucionales y datos de contenido total de electrones (TEC) para la detección de precursores de sismos. Este método difiere de los enfoques anteriores, que se basaban principalmente en técnicas tradicionales de análisis de señales sísmicas y electromagnéticas. El nuevo enfoque aprovecha la capacidad de las redes neuronales para capturar patrones complejos en datos de alta dimensión, lo que permite una detección más precisa de anomalías ionosféricas relacionadas con actividad sísmica inminente. Los resultados muestran una alta precisión en la detección de precursores, superando a los métodos tradicionales en términos de probabilidad de detección y reducción de falsas alarmas.

A diferencia de los enfoques previos que utilizaban algoritmos de optimización basados en la naturaleza, como el algoritmo de células asesinas naturales, la investigación actual emplea redes neuronales profundas para la estimación y predicción de mapas GIM-TEC. Este enfoque de aprendizaje automático permite capturar las complejas relaciones espacio-temporales presentes en los datos de TEC, lo que resulta en una mejor capacidad de detección de anomalías ionosféricas precursoras. Además, el método propuesto no requiere la recopilación manual de muestras de entrenamiento etiquetadas, lo que simplifica el proceso y reduce los costos asociados.

Mientras que algunos estudios anteriores se centraban en el análisis de señales sísmicas y electromagnéticas locales, la investigación actual utiliza datos de TEC globales obtenidos de mapas GIM. Esta estrategia permite una cobertura espacial más amplia y una detección de precursores más generalizada, no limitada a una región específica. Sin embargo, también plantea desafíos en cuanto al procesamiento y análisis de grandes cantidades de datos. El uso de redes neuronales profundas ayuda a abordar este desafío al permitir un procesamiento eficiente de los datos de alta dimensión. Los resultados muestran una alta precisión en la detección de precursores de sismos de diferentes magnitudes, lo que resalta la efectividad del enfoque propuesto.

V. CONCLUSIONES

El enfoque propuesto de detección de precursores de sismos basado en redes neuronales profundas y datos de TEC ha demostrado ser altamente efectivo. Los resultados obtenidos muestran una alta precisión en la detección de precursores, con una tasa de detección superior al 70% para sismos de magnitud superior a 5.4 en la escala de Richter, mientras que se mantiene una baja tasa de falsas alarmas. Además, se logró una distancia promedio de detección de precursores de aproximadamente 2300 km, lo que sugiere que el método es capaz de detectar anomalías ionosféricas relacionadas con actividad sísmica en un área geográfica amplia.

El uso de redes neuronales profundas, en particular las redes ConvLSTM, ha demostrado ser una elección acertada para el modelado y predicción de datos de TEC. Estas redes han podido capturar eficazmente las complejas relaciones espacio-temporales presentes en los datos, lo que ha permitido una estimación precisa de los mapas GIM-TEC futuros. Esto, a su vez, ha facilitado la detección de anomalías ionosféricas al comparar las estimaciones con los datos reales. Además, el enfoque no supervisado ha simplificado el proceso al no requerir datos de entrenamiento etiquetados manualmente.

Si bien los resultados obtenidos son prometedores, aún existen áreas de mejora. Por ejemplo, el posicionamiento preciso de las ubicaciones de detección de precursores sigue siendo un desafío, ya que requiere el procesamiento adicional de las variaciones espaciales de los datos de TEC. Además, la investigación se centró en una región específica (Anatolia), por lo que sería beneficioso explorar el rendimiento del método en otras regiones sísmicas. En general, la investigación ha abierto nuevas posibilidades en el campo de la detección de precursores de sismos, sentando las bases para futuros avances y aplicaciones prácticas.

VI. RECOMENDACIONES

Se recomienda explorar técnicas avanzadas de procesamiento espacial y análisis de patrones para mejorar el posicionamiento preciso de las ubicaciones de detección de precursoros. Esto podría lograrse mediante el procesamiento de múltiples regiones GIM-TEC de forma independiente y la combinación de sus puntuaciones de anomalías basadas en SSIM. Además, la incorporación de otros datos geoespaciales, como mediciones terrestres y de satélite, podría proporcionar información complementaria útil para triangular con mayor precisión las ubicaciones de detección de precursoros.

Sería beneficioso ampliar la investigación a diferentes regiones sísmicas del mundo para evaluar el rendimiento del método propuesto en diversos entornos geológicos y condiciones ionosféricas. Esto permitiría ajustar y optimizar los parámetros del modelo, así como identificar posibles limitaciones o sesgos regionales. Además, podría explorarse la posibilidad de desarrollar modelos específicos para cada región, aprovechando las características únicas de los datos de TEC en diferentes ubicaciones.

Se recomienda explorar la incorporación de técnicas de aprendizaje profundo adicionales, como redes generativas antagónicas (GANs) o enfoques de aprendizaje por transferencia, para mejorar aún más la capacidad de detección de anomalías ionosféricas. Estas técnicas podrían ayudar a capturar patrones más sutiles y complejos en los datos de TEC, lo que podría llevar a una mayor precisión en la detección de precursoros y una reducción adicional de falsas alarmas. Además, podrían investigarse enfoques de aprendizaje semi-supervisado para aprovechar cualquier dato etiquetado disponible de manera óptima.

VI. REFERENCIAS BIBLIOGRÁFICAS

- [1] J. Chen, et al., “Immune optimization inspired artificial natural killer cell earthquake prediction method”, *The Journal of Supercomputing*, vol. 78, pp. 19478-19500, 2022.
- [2] F. Dong, Z. Li y J. Liu, “Application of Convolutional Neural Network to GIS and Physics”, *Computational Intelligence and Neuroscience*, vol. 8559343, pp.1-6, 2022.
- [3] P. Huang, et al., “A Deep Learning-Based Electromagnetic Signal for Earthquake MAgnitude Prediction”, *Sensors*, vol. 21, no. 4434, pp. 1-16, 2021.
- [4] A. Akyol, O. Arikan y F. Arikan, “A machine learning-based detection of earthquake precursors using ionospheric data”, *Radio Science*, vol. 55, no. 11, pp. 1-21, 2020.
- [5] R. Balestriero, et al., “Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning”, *Nature Communications*, vol. 11, no. 3972, pp. 1-12, 2020.
- [6] M. Hashimoto, et al., “Disaster Intensity-based Selection of Training Samples for Remote Sensing Building Damage Classification”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 10, pp. 8288-8304, 2021.
- [7] “FAQs, Can you predict earthquakes?, U.S. Geological Survey”, <https://www.usgs.gov/faqs/can-you-predict-earthquakes>.
- [8] M. Wyss, “Cannot earthquakes be predicted?”, *Science*, vol. 278, no. 5337, pp. 487-490, 1997.
- [9] R. J. Geller, “Cannot earthquakes be predicted”, *Responses, Science 278 N5337*, Citeseer, 1997.
- [10] R. J. Geller, “Earthquake prediction: a critical review”, *Geophysical Journal International*, vol. 131, pp. 425–450, 1997.
- [11] R. Bendick y R. Bilham, “Do weak global stresses synchronize earthquakes?”, *Geophysical Research Letters*, vol. 44, no. 16, pp. 8320–8327, 2017.
- [12] S. Kannan, “Innovative mathematical model for earthquake prediction”, *Engineering Failure Analysis*, vol. 41, pp. 89-95, 2014.
- [13] B. Papazachos y C. A. Papaioannou, “Long-term earthquake prediction in the aegean area based on a time and magnitude predictable model”, *pure and applied geophysics*, vol. 140, no. 4, pp. 593-612, 1993.
- [14] A. Boucouvalas, M. Gkasios, N. Tselikas y G. Drakatos, “Modifiedfibonacci-dual-lucas method for earthquake prediction”, *Proc. of SPIE Vol*, vol. 9535, pp. 95351A–1, 2015.

- [15] M. Moustra, M. Avraamides y C. Christodoulou, “Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals,” *Expert systems with applications*, vol. 38, no. 12, pp. 15032–15039, 2023.
- [16] J. Mahmoudi, M. A. Arjomand, M. Rezaei y M. H. Mohammadi, “Predicting the earthquake magnitude using the multilayer perceptron neural network with two hidden layers”, *Civil engineering journal*, vol. 2, no. 1, pp. 1-12, 2016.
- [17] G. Asencio-Cortés, F. Martínez-Álvarez, A. Morales-Esteban, y J. Reyes, “A sensitivity study of seismicity indicators in supervised learning to improve earthquake prediction”, *Knowledge-Based Systems*, vol. 101, pp. 15-30, 2016.
- [18] G. Asencio-Cortés, F. Martínez-Alvarez, A. Morales-Esteban, J. Reyes, and A. Troncoso, “Improving earthquake prediction with principal component analysis: application to chile”, *International Conference on Hybrid Artificial Intelligence Systems*, pp. 393-404, Springer, 2015.
- [19] M. Last, N. Rabinowitz y G. Leonard, “Predicting the maximum earthquake magnitude from seismic data in israel and its neighboring countries”, *PloS one*, vol. 11, no. 1, p. e0146101, 2016.
- [20] K. M. Asim, A. Idris, T. Iqbal y F. Martínez-Alvarez, “Earthquake pre-diction model using support vector regressor and hybrid neural networks”, *PloS one*, vol. 13, no. 7, p. e0199004, 2018.
- [21] M. Shakeel, K. Itoyama, K. Nishida y K. Nakadai, “Detecting earthquakes: a novel deep learning-based approach for effective disaster response”, *Applied Intelligence*, pp. 1-11, 2021.
- [22] R. Jena, B. Pradhan, A. Al-Amri, C. W. Lee y H.-j. Park, “Earthquake probability assessment for the indian subcontinent using deep learning”, *Sensors*, vol. 20, no. 16, p. 4369, 2020.
- [23] M. Ibrahim, J. Park y N. Athens, “Earthquake warning system: Detecting earthquake precursor signals using deep neural networks,” *Technical Report CS 230*, 2018.
- [24] Q. Wang, Y. Guo, L. Yu y P. Li, “Earthquake prediction based on spatiotemporal data mining: An lstm network approach”, *IEEE Transactions on Emerging Topics in Computing*, 2017.
- [25] A. Panakkat y H. Adeli, “Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators”, *Computer-Aided Civil and Infrastructure Engineering*, vol. 24, no. 4, pp. 280–292, 2009.

- [26] L. Allegri, F. Bella, G. Della Monica, A. Ermini, S. Improta, V. Sgrigna y P. Biagi, “Radon and tilt anomalies detected before the irpinia (south italy) earthquake of november 23, 1980 at great distances from the epicenter”, *Geophysical Research Letters*, vol. 10, no. 4, pp. 269-272, 1983.
- [27] J. Hartmann y J. K. Levy, “Hydrogeological and gasgeochemical earthquake precursors—a review for application,” *Natural Hazards*, vol. 34, no. 3, pp. 279-304, 2005.
- [28] T. Gulyaeva, F. Arikan y I. Stanislawski, “Persistent long-term (1944-2015) ionosphere-magnetosphere associations at the area of intense seismic activity and beyond,” *Advances in Space Research*, vol. 59, no. 4, pp. 1033-1040, 2017.
- [29] S. Karatay, F. Arikan y O. Arikan, “Investigation of total electron content variability due to seismic and geomagnetic disturbances in the ionosphere”, *Radio Science*, vol. 45, no. 5, 2010.
- [30] S. Pulnits, T. Gaivoronska, A. L. Contreras, L. Ciraolo, et al., “Correlation analysis technique revealing ionospheric precursors of earthquakes”, *Natural Hazards and Earth System Science*, vol. 4, no. 5/6, pp. 697–702, 2004.
- [31] A. Namgaladze, O. Zolotov, I. Zakharenkova, I. Shagimuratov y O. Martynenko, “Ionospheric total electron content variations observed before earthquakes: Possible physical mechanism and modeling”, *arXiv preprint arXiv:0905.3313*, 2009.
- [32] K.-I. Oyama, M. Devi, K. Ryu, C. Chen, J. Liu, H. Liu, L. Bankov y T. Kodama, “Modifications of the ionosphere prior to large earthquakes: report from the ionosphere precursor study group”, *Geoscience Letters*, vol. 3, no. 1, p. 6, 2016.
- [33] A. Tronin, P. Biagi, O. Molchanov, Y. Khatkevich y E. Gordeev, “Temperature variations related to earthquakes from simultaneous observation at the ground stations and by satellites in Kamchatka area”, *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 29, no. 4, pp. 501–506, 2004.
- [34] A. A. Tronin, M. Hayakawa y O. A. Molchanov, “Thermal IR satellite data application for earthquake research in Japan and China”, *Journal of Geodynamics*, vol. 33, no. 4, pp. 519-534, 2002.
- [35] S. Pulnits, D. Ouzounov, A. Karelin, K. Boyarchuk y L. Pokhmelnikh, “The physical nature of thermal anomalies observed before strong earthquakes”, *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 31, no. 4, pp. 143-153, 2006.

- [36] G. Asteriadis and E. Livieratos, “Pre-seismic responses of underground water level and temperature concerning a 4.8 magnitude earthquake in Greece on October 20, 1988,” *Tectonophysics*, vol. 170, no. 1-2, pp. 165–169, 1989.
- [37] R. A. Grant y T. Halliday, “Predicting the unpredictable; evidence of preseismic anticipatory behaviour in the common toad,” *Journal of Zoology*, vol. 281, no. 4, pp. 263–271, 2010.
- [38] C. Fidani, “The earthquake lights (EQL) of the 6 April 2009 Aquila earthquake, in Central Italy”, *Nat. Hazards Earth Syst. Sci.*, vol. 10, no. 5, pp. 967–978, 2010.
- [39] S. Pulnits y D. Ouzounov, “Lithosphere–atmosphere–ionosphere coupling (laic) model—an unified concept for earthquake precursors validation”, *Journal of Asian Earth Sciences*, vol. 41, no. 4-5, pp. 371–382, 2023.
- [40] S. Pulnits y D. Davidenko, “The nocturnal positive ionospheric anomaly of electron density as a short-term earthquake precursor and the possible physical mechanism of its formation”, *Geomagnetism and Aeronomy*, vol. 58, no. 4, pp. 559–570, 2018.
- [41] S. Pulnits, K. Boyarchuk, A. Lomonosov, V. Khegai y I. Y. Lyu, “Ionospheric precursors to earthquakes: a preliminary analysis of the foF2 critical frequencies at Chung-Li ground-based station for vertical sounding of the ionosphere (Taiwan island)”, *Geomagnetism and Aeronomy*, vol. 42, no. 4, pp. 508-513, 2002.
- [42] J.-Y. Liu, Y.-I. Chen, H.-K. Jhuang y Y.-H. Lin, “Ionospheric foF2 and TEC anomalous days associated with $M=5.0$ earthquakes in Taiwan during 1997-1999,” *Terrestrial Atmospheric and Oceanic Sciences*, vol. 15, no. 3, pp. 371–384, 2004.
- [43] J.-Y. Liu, Y. Chen, Y. Chuo y C.-S. Chen, “A statistical investigation of preearthquake ionospheric anomaly”, *Journal of Geophysical Research: Space Physics*, vol. 111, no. A5, 2006.
- [44] H. Le, J.-Y. Liu y L. Liu, “A statistical analysis of ionospheric anomalies before 736 $M \geq 6.0$ earthquakes during 2002–2010”, *Journal of Geophysical Research: Space Physics*, vol. 116, no. A2, 2023.
- [45] S. Kon, M. Nishihashi y K. Hattori, “Ionospheric anomalies possibly associated with $m \geq 6.0$ earthquakes in the Japan area during 1998–2010: Case studies and statistical study”, *Journal of Asian Earth Sciences*, vol. 41, no. 4-5, pp. 410–420, 2023.
- [46] S. Pulnits, “Seismic activity as a source of the ionospheric variability”, *Advances in Space Research*, vol. 22, no. 6, pp. 903–906, 1998.

- [47] S. Kouris, K. Polimeris y L. R. Cander, “Specifications of TEC variability”, *Advances in Space Research*, vol. 37, no. 5, pp. 983–1004, 2006.
- [48] M. Akhoondzadeh, “Decision tree, bagging and random forest methods detect tec seismo-ionospheric anomalies around the time of the chile,(mw= 8.8) earthquake of 27 february 2010”, *Advances in Space Research*, vol. 57, no. 12, pp. 2464–2469, 2016.
- [49] D. Davidenko y S. Pulinets, “Deterministic variability of the ionosphere on the eve of strong ($m \geq 6$) earthquakes in the regions of greece and italy according to long-term measurements data”, *Geomagnetism and Aeronomy*, vol. 59, no. 4, pp. 493–508, 2019.
- [50] J. Liu, Y. Chen, S. Pulinets, Y. Tsai y Y. Chuo, “Seismo-ionospheric signatures prior to $M \geq 6.0$ Taiwan earthquakes”, *Geophysical research letters*, vol. 27, no. 19, pp. 3113–3116, 2000.
- [51] J. Y. Liu, Y. Chuo, S. Shan, Y. Tsai, Y. Chen, S. Pulinets, S. Yu, et al., “Pre-earthquake ionospheric anomalies registered by continuous GPS TEC measurements”, *Annales Geophysicae*, vol. 22, pp. 1585–1593, 2004.
- [52] J. Liu, C. Chen, Y. Chen, W. Yang, K. Oyama y K. Kuo, “A statistical study of ionospheric earthquake precursors monitored by using equatorial ionization anomaly of gps tec in taiwan during 2001–2007”, *Journal of Asian Earth Sciences*, vol. 39, no. 1, pp. 76–80, 2010.
- [53] M. Devi, A. Medhi, A. J. D. S. Sarma y A. K. Barbara, “Growth and inhibition of equatorial anomaly prior to an earthquake (eq): case studies with total electron content (tec) data for major eqs of japan 2023 and indonesia 2012”, *Positioning*, vol. 2013, 2013.
- [54] D. Ouzounov, S. Pulinets, A. Romanov, A. Romanov, K. Tsybulya, D. Davidenko, M. Kafatos y P. Taylor, “Atmosphere-ionosphere response to the M9 Tohoku earthquake revealed by multi-instrument space-borne and ground observations: Preliminary results”, *Earthquake Science*, vol. 24, no. 6, pp. 557–564, 2023.
- [55] O. Zolotov, A. Namgaladze y B. Prokhorov, “Specific features of ionospheric total electron content variations in the periods of preparation of the earthquakes on march 11, 2023 (japan) and october 23, 2023 (turkey)”, *Russian Journal of Physical Chemistry B*, vol. 7, no. 5, pp. 599–605, 2013.
- [56] K. S. Yadav, S. P. Karia y K. N. Pathak, “Anomalous variation in gps based tec prior to the 5 earthquakes in 2009 and 2010”, *Positioning*, vol. 6, no. 04, p. 96, 2015.
- [57] A. Trigunait, M. Parrot, S. Pulinets y F. Li, “Variations of the ionospheric electron density during the bhuj seismic event”, *Annales Geophysicae*, vol. 22, pp. 4123-4131, Copernicus GmbH, 2004.

- [58] W. Li, J. Guo, J. Yue, Y. Yang, Z. Li y D. Lu, “Contrastive research of ionospheric precursor anomalies between calbuco volcanic eruption on april 23 and nepal earthquake on april 25, 2015”, *Advances in space research*, vol. 57, no. 10, pp. 2141–2153, 2016.
- [59] M. Zhang, J. Shi, X. Wang y S. Radicella, “Ionospheric variability at low latitude station: Hainan, China”, *Advances in Space Research*, vol. 34, no. 9, pp. 1860–1868, 2004.
- [60] S. Pulnits, “Ionospheric precursors of earthquakes; recent advances in theory and practical applications”, *Terrestrial Atmospheric and Oceanic Sciences*, vol. 15, no. 3, pp. 413-436, 2004.
- [61] S. Pulnits, A. Leyva-Contreras, G. Bisiacchi-Giraldi y C. Ciralo, “Total electron content variations in the ionosphere before the Colima, Mexico, earthquake of 21 January 2003”, *Geofisica Internacional*, vol. 44, no. 4, 2005.
- [62] V. Plotkin, “GPS detection of ionospheric perturbation before the 13 February 2001, El Salvador earthquake”, *Natural Hazards and Earth System Science*, vol. 3, no. 3/4, pp. 249-253, 1999.
- [63] S. Kouris y D. Fotiadis, “Ionospheric variability: a comparative statistical study”, *Advances in Space Research*, vol. 29, no. 6, pp. 977–985, 2002.
- [64] S. Pulnits, A. Kotsarenko, L. Ciralo y I. Pulnits, “Special case of ionospheric day-to-day variability associated with earthquake preparation”, *Advances in Space Research*, vol. 39, no. 5, pp. 970–977, 2007.
- [65] F. Arikani, M. Deviren, O. Lenk, U. Sezen y O. Arikani, “Observed ionospheric effects of 23 october 2023 van, turkey earthquake”, *Geomatics, Natural Hazards and Risk*, vol. 3, no. 1, pp. 1–8, 2012.
- [66] A. A. Akyol, “Investigation on the Reliability of Earthquake Prediction Based on Ionospheric Electron Content Variation”, Master’s thesis, Bilkent University, Ankara, 2013.
- [67] A. A. Akyol, O. Arikani, F. Arikani y M. N. Deviren, “Investigation on the reliability of earthquake prediction based on ionospheric electron content variation”, *Information Fusion (FUSION)*, 2013 16th International Conference on, pp. 1658-1663, IEEE, 2013.
- [68] A. A. Akyol, O. Arikani y F. Arikani, “A machine learning-based detection of earthquake precursors using ionospheric data”, *Radio Science*, vol. 55, no. 11, pp. 1-21, 2020.
- [69] U. Sezen, F. Arikani, O. Arikani, O. Ugurlu y A. Sadeghimorad, “Online, automatic, near-real time estimation of gps-tec: Ionolab-tec”, *Space Weather*, vol. 11, no. 5, pp. 297-305, 2013.
- [70] “IONOLAB, Ionospheric Research Laboratory.” <http://www.ionolab.org/>.

- [71] M. Deviren, F. Arıkan y O. Arıkan, “Spatio-temporal interpolation of total electron content using a gps network”, *Radio Science*, vol. 48, no. 3, pp. 302–309, 2013.
- [72] SILSO World Data Center, “The International Sunspot Number”, *International Sunspot Number Monthly Bulletin and online catalogue*, 2005-2016.
- [73] K. G. Mehrotra, C. K. Mohan y H. Huang, *Anomaly detection principles and algorithms*. Springer, 2017.
- [74] N. Turel y F. Arıkan, “Probability density function estimation for characterizing hourly variability of ionospheric total electron content”, *Radio Science*, vol. 45, no. 6, 2010.
- [75] S. Karatay, F. Arıkan y O. Arıkan, “Investigation of Hourly and Daily Patterns for Lithosphere-Ionosphere Coupling Before Strong Earthquakes”, *Recent Advances in Space Research*, 2009.
- [76] M. L. Stein, *Interpolation of spatial data: some theory for kriging*. Springer, 1999.
- [77] I. Stanislawska, G. Juchnikowski, L. R. Cander, L. Ciraolo, P. Bradley, Z. Zbyszynski y A. Swiatek, “The kriging method of TEC instantaneous mapping”, *Advances in Space Research*, vol. 29, no. 6, pp. 945–948, 2002.
- [78] M. Sugiura, T. Kamei, A. Berthelier y M. Menvielle, *Equatorial Dst Index: 1957-1986*. ISGI Publications Office, 1991.
- [79] T. Gulyaeva y I. Stanislawska, “Derivation of a planetary ionospheric storm index”, *Annales Geophysicae*, vol. 26, pp. 2645–2648, Copernicus GmbH, 2008.
- [80] A. Mikhailov, A. Depueva, V. Depuev, et al., “Daytime f2-layer negative storm effect: what is the difference between storm-induced and qdisturbance events?”, *Annales Geophysicae*, vol. 25, pp. 1531–1541, 2007.
- [81] C. Lee y D. A. Landgrebe, “Feature extraction based on decision boundaries”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 4, pp. 388–400, 1993.
- [82] I. B. Aban, M. M. Meerschaert y A. K. Panorska, “Parameter estimation for the truncated Pareto distribution”, *Journal of the American Statistical Association*, vol. 101, no. 473, pp. 270–277, 2006.
- [83] L. Zaninetti y M. Ferraro, “On the truncated Pareto distribution with applications”, *Central European Journal of Physics*, vol. 6, no. 1, pp. 1-6, 2008.

- [84] A. A. Akyol, O. Ankan y F. Ankan, “Generation of earthquake detection signal based on ionospheric electron content variation”, 2014 22nd Signal Processing and Communications Applications Conference (SIU), pp. 590-593, IEEE, 2014.
- [85] W. Elmenreich, “A review on system architectures for sensor fusion applications”, *Software Technologies for Embedded and Ubiquitous Systems*, pp. 547–559, Springer, 2007.
- [86] R. C. Eberhart y Y. Shi, “Particle swarm optimization: developments, applications and resources”, *evolutionary computation*, 2001. *Proceedings of the 2001 Congress on*, vol. 1, pp. 81–86, IEEE, 2001.
- [87] C. Kuo, J. Huba, G. Joyce y L. Lee, “Ionosphere plasma bubbles and density variations induced by pre-earthquake rock currents and associated surface charges”, *Journal of Geophysical Research*, vol. 116, no. A10, p. A10317, 2023.
- [88] R. Kohavi et al., “A study of cross-validation and bootstrap for accuracy estimation and model selection”, *Ijcai*, vol. 14, pp. 1137–1145, 1995.
- [89] T. Dietterich, “Overfitting and undercomputing in machine learning”, *ACM computing surveys (CSUR)*, vol. 27, no. 3, pp. 326–327, 1995.
- [90] J. Fürnkranz, “Pruning algorithms for rule learning,” *Machine Learning*, vol. 27, no. 2, pp. 139–172, 1997.
- [91] C. Scott, “Performance measures for neyman–pearson classification”, *Information Theory, IEEE Transactions on*, vol. 53, no. 8, pp. 2852–2863, 2007.
- [92] K. Davies y G. Hartmann, “Studying the ionosphere with the global positioning system”, *Radio Science*, vol. 32, no. 4, pp. 1695–1703, 1997.
- [93] F. Arikan, C. Erol y O. Arikan, “Regularized estimation of vertical total electron content from gps data for a desired time period”, *Radio Science*, vol. 39, no. 6, 2004.
- [94] I. Sayin, F. Arikan y O. Arikan, “Regional tec mapping with random field priors and kriging”, *Radio Science*, vol. 43, no. 5, 2008.
- [95] F. Arikan, C. Erol y O. Arikan, “Regularized estimation of vertical total electron content from Global Positioning System data”, *Journal of Geophysical Research*, vol. 108, no. A12, p. 1469, 2003.
- [96] C. Bruyninx, H. Habrich, W. Söhne, A. Kenyeres, G. Stangl. y C. Volksen, “Enhancement of the euref permanent network services and products”, *Geodesy for Planet Earth*, pp. 27-34, Springer, 2012.

- [97] F. Arikan, O. Arikan y C. B. Erol, “Regularized estimation of tec from gps data for certain midlatitude stations and comparison with the iri model”, *Advances in Space Research*, vol. 39, no. 5, pp. 867–874, 2007.
- [98] H. Nayir, F. Arikan, O. Arikan y C. Erol, “Total electron content estimation with regest”, *Journal of Geophysical Research: Space Physics*, vol. 112, no. A11, 2007.
- [99] F. Arikan, H. Nayir, U. Sezen y O. Arikan, “Estimation of single station interfrequency receiver bias using gps-tec”, *Radio Science*, vol. 43, no. 4, 2008.
- [100] M. Manda y M. Korte, *Geomagnetic observations and models*, vol. 5. Springer, 2010.
- [101] T. Fuller-Rowell, M. Codrescu y P. Wilkinson, “Quantitative modeling of the ionospheric response to geomagnetic activity”, *Annales Geophysicae*, vol. 18, pp. 766-781, Copernicus GmbH, 2000.
- [102] T. Gulyaeva, F. Arikan, I. Stanislawska y L. Poustovalova, “Global distribution of zones of enhanced risk for the ionospheric weather”, *Journal of Geography, Environment and Earth Science International*, vol. 4, no. 1, pp. 1–13, 2016.
- [103] T. Gulyaeva y F. Arikan, “Statistical discrimination of global postseismic ionosphere effects under geomagnetic quiet and storm conditions”, *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 509-524, 2017.
- [104] T. Gulyaeva, F. Arikan y I. Stanislawska, “Earthquake aftereffects in the equatorial ionization anomaly region under geomagnetic quiet and storm conditions”, *Advances in Space Research*, vol. 60, no. 2, pp. 406–418, 2017.
- [105] G. Rostoker, “Geomagnetic indices”, *Reviews of Geophysics*, vol. 10, no. 4, pp. 935-950, 1972.
- [106] “NASA OMNIWeb, SPDF Goddard Space Flight Center”, <https://omniweb.gsfc.nasa.gov/form/dx1.html>. Accessed: 2017-04-17.
- [107] “ANSS Comprehensive Catalog (ComCat), U.S. Geological Survey”, <https://earthquake.usgs.gov/earthquakes/search/>. Accessed: 2017-04-17.
- [108] D. Ouzounov, S. Pulinet, K. Hattori y P. Taylor, *Pre-earthquake processes: A multidisciplinary approach to earthquake prediction studies*, vol. 234. John Wiley & Sons, 2018.
- [109] F. Arikan, S. Shukurov, H. Tuna, O. Arikan y T. Gulyaeva, “Performance of gps slant total electron content and iri-plas-tec for days with ionospheric disturbance”, *Geodesy and Geodynamics*, vol. 7, no. 1, pp. 1-10, 2016.

- [110] M. Necat Deviren, T. Gulyaeva, U. Sezen, F. Arikan y O. Arikan, “Detection of seismic precursors using distance metrics between gps-tec and iri-plas”, 40th COSPAR Scientific Assembly, vol. 40, 2014.
- [111] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis”, *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.
- [112] L. Kaufman y P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*, vol. 344. John Wiley & Sons, 2009.
- [113] S. Karatay, F. Arikan y O. Arikan, “Investigation of hourly and daily patterns for lithosphere-ionosphere coupling before strong earthquakes”, *Recent Advances in Space Technologies*, 2009. RAST’09. 4th International Conference on, pp. 670–674, IEEE, 2009.
- [114] C. Cortes y V. Vapnik, “Support-vector networks”, *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [115] C. J. Burges, “A tutorial on support vector machines for pattern recognition”, *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [116] J. Platt, “Sequential minimal optimization: A fast algorithm for training support vector machines”, 1998.
- [117] B. W. Matthews, “Comparison of the predicted and observed secondary structure of t4 phage lysozyme”, *Biochimica et Biophysica Acta (BBA) Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.
- [118] J.-I. G. S. Data, “Crustal dynamics data information system (cddis)”, 2022. Last accessed 01 January 2022.
- [119] T. Cooijmans, N. Ballas, C. Laurent, C. . G“ul,cehre y A. Courville, “Recurrent batch normalization”, 2016.
- [120] S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong y W.-c. Woo, “Convolutional lstm network: A machine learning approach for precipitation nowcasting”, *Advances in neural information processing systems*, pp. 802–810, 2015.
- [121] A. Graves, “Generating sequences with recurrent neural networks”, 2013.
- [122] C. Laurent, G. Pereyra, P. Brakel, Y. Zhang y Y. Bengio, “Batch normalized recurrent neural networks”, 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2657–2661, IEEE, 2016.

- [123] Z. Wang, A. C. Bovik, H. R. Sheikh y E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE transactions on image processing, vol. 13, no. 4, pp. 600-612, 2004.
- [124] A. Hore y D. Ziou, "Image quality metrics: Psnr vs. ssim", 2010 20th international conference on pattern recognition, pp. 2366-2369, IEEE, 2010.