Red neuronal convolucional para el cálculo rápido del azimut de un terremoto

Convolutional Neural Network for fast calculation of earthquake azimuth

AUTORES: Cutié Mustelier, Manuel [1]
Salgado Castillo, Antonio [2]
Palau Clares, Raúl [3]

Dirección para correspondencia: manuelcutie@gmail.com

Fecha de recepción: febrero 12, 2024

Fecha de aceptación: septiembre 02, 2024

DOI: https://doi.org/10.61154/mrcm.v10i3.3303

RESUMEN

El proceso de automatización del Servicio Sismológico Nacional de Cuba, se ha convertido en una tarea de primer orden para el país. Lo anterior se justifica, con el objetivo de lograr ganar en efectividad y eficiencia en la detección y localización de sismos en el territorio nacional, de ahí la importancia de esta investigación. El objetivo fue el diseño e implementación de una red neuronal artificial, para determinar de manera automática el azimut de un terremoto al realizar su localización con una estación sismológica de tres componentes. Se consideró para el desarrollo de la investigación un enfoque cuantitativo. Las señales se tomaron de la estación sismológica, Río Carpintero. La investigación se abordó como un problema de regresión múltiple y se propuso un modelo de Red Convolucional 1D. Se empleó Python como lenguaje de programación. Para la adquisición y pre procesamiento de las señales sísmicas se utilizó la herramienta de código abierto basado en Python: Obspy. El modelo de red neuronal se implementó utilizando los marcos de desarrollo para aprendizaje de máquinas Keras y Tensor flow. Se utilizaron en total

[1] Ingeniero Electrónico, Tecnólogo. Centro Nacional de Investigaciones Sismológicas (CENAIS), Santiago de Cuba, Cuba. E-mail: manuelcutie@gmail.com Código ORCID https://orcid.org/0000-0001-7909-8925
[2] Licenciado en Ciencia de la Computación. Master en Neurociencias. Doctor en Ciencias. Profesor e Investigador Auxiliar. Centro Nacional de Investigaciones Sismológicas (CENAIS), Santiago de Cuba, Cuba. E-mail: asalgadocastillo@gmail.com Código ORCID https://orcid.org/0000-0003-3214-7587
[3] Licenciado en Física. Master en Ciencias de la Computación. Investigador Auxiliar. Centro Nacional de Investigaciones Sismológicas (CENAIS), Santiago de Cuba, Cuba. E-mail: rpalau@cenais.cu, Código ORCID https://orcid.org/0000-0003-3864-2159

49233 señales de terremotos, de ellas 39386 para el entrenamiento y 9847 para la validación. La red diseñada e implementada fue capaz de determinar automáticamente el azimut de un terremoto con una efectividad del 90 %, evidenciado en los valores del coeficiente de determinación R^2. Las pruebas y validaciones realizadas indicaron la posibilidad de extender la metodología y el uso de la herramienta computacional a otras estaciones sismológicas de Cuba.

PALABRAS CLAVE: Red neuronal convolucional, azimut; señales sísmicas; terremotos; sismo; Inteligencia artificial. (Tesauro de la UNESCO)

ABSTRACT

The automation process of the National Seismological Service of Cuba has become a task of the first order for the country. This is justified by the objective of achieving greater effectiveness and efficiency in the detection and location of earthquakes in the national territory, hence the importance of this research. The objective was the design and implementation of an artificial neural network to automatically determine the azimuth of an earthquake by locating it with a three-component seismological station. A quantitative approach was considered for the development of the research. The signals were taken from the seismological station, Río Carpintero. The research was approached as a multiple regression problem and a 1D Convolutional Network model was proposed. Python was used as the programming language. For the acquisition and pre-processing of the seismic signals, the open source tool based on Python: Obspy was used. The neural network model was implemented using the development frameworks for machine learning Keras and Tensor flow. A total of 49,233 earthquake signals were used, of which 39,386 were used for training and 9,847 for validation. The designed and implemented network was able to automatically determine the azimuth of an earthquake with 90% effectiveness, as evidenced by the values of the coefficient of determination R^2. The tests and validations carried out indicated the possibility of extending the methodology and the use of the computational tool to other seismological stations in Cuba.

KEYWORDS: Artificial convolutional network, azimuth; seismic signals; Seisms; earthquake; Artificial intelligence**.** (UNESCO Thesaurus)

INTRODUCCIÓN.

Los terremotos clasifican entre los fenómenos naturales capaces de causar daños económicos y humanos significativos. Por tal motivo, el estudio detallado de los mismos es fundamental para prevenir y mitigar sus efectos destructores. Un desafío clave en la investigación de los terremotos es su localización ya que es de vital importancia para la estimación de posibles riesgos geológicos donde destacan: sismos, deslizamientos, colapsos, entre otros (Rosabal et al., 2024).

En este orden de ideas, se debe precisar, que Cuba forma parte de la placa de Norteamérica, y que la parte sur oriental del archipiélago se ubica en el límite entre las placas de Norteamérica y el Caribe. Esta gran falla, transformante es responsable de más del 90 % de la sismicidad y de los mayores terremotos observados a lo largo del margen sur cubano (Palau et al., 2023).

Consecuentemente, el monitoreo y seguimiento de la sismicidad es de vital importancia teniendo en cuenta que son bien conocidos los graves impactos que pueden tener los sismos (Diez et al., 2022). Esta tarea suele realizarse por medio de redes sismológicas que se distribuyen a lo largo de las zonas de interés; las mismas se componen de sismómetros o acelerómetros que transmiten en tiempo real diversos parámetros hacia una estación central donde los datos son procesados (Salgado y Palau, 2023).

Por ende, es esencial que estas redes estén dispuestas en una geometría, típicamente circular, de manera que abarque toda la falla o área sismogénica previamente identificada, como históricamente responsable de la generación de sismos significativos. Esto determina el alcance o umbral de detección de la red, como se destacó en el estudio de Diez et al. (2022).

Este requerimiento ha traído consigo la inclusión e instalación de nuevas y modernas estaciones en el territorio nacional. Es así, que el Servicio Sismológico Nacional de Cuba (SSNC), se ha propuesto entre sus objetivos principales, el perfeccionamiento de los procesos de detección y determinación de los parámetros de los terremotos que ocurren en Cuba y en áreas cercanas (Iturralde y Arango, 2020). Sin embargo, en la actualidad se confrontan insuficiencias, en la detección epicentral en las zonas central y occidental de Cuba, ya que no han sido suficientemente cubiertas.

Asimismo, algunas partes de la región oriental presentan gaps sísmicos durante el proceso de detección, que obliga en muchas ocasiones a utilizar una sola estación para realizar la

determinación epicentral de un sismo, lo que trae consigo errores debido al procedimiento manual utilizado (Salgado et al., 2023).

Consecuentemente, el proceso de automatización del SSNC, se ha convertido en una tarea de primer orden para los investigadores del Centro Nacional de Investigaciones Sismológicas (CENAIS). Lo anterior se justifica, con el objetivo de lograr ganar en efectividad y eficiencia en la detección y localización de sismos en el territorio nacional, de ahí la importancia de esta investigación.

Contextualizando la situación existente en Cuba, se puede afirmar que se cuenta con una red de estaciones sismológicas de banda ancha y de corto periodo, que permiten la localización de cualquier sismo que ocurra en el país y en áreas cercanas. Sin embargo, debido a que la principal actividad sísmica es moderada, caracterizada con sismos de baja magnitud en su mayoría y la red no es homogénea, existe un gran número de eventos que, si bien están registrados en el catálogo del SSNC, no están localizados correctamente. Ya que generalmente son detectados por una o dos estaciones, lo que conlleva que el proceso de localización epicentral sea inexacto. Esta situación se agrava más en las regiones central y occidental de Cuba donde hay muy pocas estaciones sismológicas (Cutié, Salgado y Palau, 2023).

Ahora bien, la localización de un sismo utilizando una sola estación, constituye un desafío, en cuanto a las posibilidades de minimizar los errores y la incertidumbre de los parámetros epicentrales. Esto se evidencia si se contrasta con los resultados que se obtienen al utilizar tres o más estaciones con métodos como el Hypo71, HYPOINVERSE. Investigadores como Noda et al. (2012) han referido estos problemas, por lo que han tratado de combinarlo con el método de Análisis de Componentes Principales (PCA: siglas en inglés), en un intento por determinar eficientemente la onda de arribo P.

Así mismo, Eisermann et al. (2015), demostraron las dificultades que presentan las técnicas usadas para determinar el azimut de sismos mediante una sola estación, al intentar crear un sistema de alerta temprana para el sur de California, EE.UU. Estos autores plantearon que los errores vienen dados principalmente por la componente humana que está implícita en el marcaje de las ondas cuando el análisis se hace con una sola estación.

Sobre la base de estas limitaciones, manifestadas por investigadores de diversos países y la experiencia de los autores, se realizó una revisión del Catálogo Sismológico de Cuba desde 1998

al 2021, así como una entrevista a investigadores, técnicos y tecnólogos que se desempeñan en el SSNC. Del procesamiento de la información obtenida se pudo concluir que existen; dificultades al marcar la onda P, así como imprecisiones en la estimación del azimut hacia el epicentro (back azimuth) al realizar el análisis manual utilizando el registro de solo una estación.

Considerando lo anterior, se definió como problema principal, las limitaciones en la estimación del azimut de un terremoto, mediante una sola estación sismológica, lo que limita su correcta determinación epicentral.

Entre las causas principales de este problema, se encuentran una insuficiente automatización del proceso de marcaje de las ondas P de la señal sísmica y predominio de sesgos en las técnicas que se emplean para la estimación del azimut mediante una sola estación sismológica. Estas causas indican la necesidad de perfeccionar el proceso de determinación epicentral de un terremoto usando una sola estación sismológica, el cual se asumió como objeto de la investigación.

Este proceso ha sido investigado por numerosos autores, entre los que se destacan Siervo (2019), quien afirma que, en la actualidad, los softwares de detección temprana de eventos sísmicos requieren de la identificación automática de estas fases (principalmente la fase P por ser la primera en llegar). El algoritmo más utilizado para desarrollar esta tarea se conoce como STA/LTA, y consiste en el procesamiento de las señales a través de filtros en diferentes ventanas de tiempo para identificar anomalías de amplitud y así identificar las fases sísmicas. Con independencia de lo planteado, se debe señalar que el mencionado algoritmo en ocasiones se demora y no es preciso, provocando errores en la determinación epicentral.

Por otra parte, Havskov y Ottemöller (2023), quienes consideran que, en el procedimiento para la localización manual de sismos, lo más usual es que una persona experta en el análisis de estas señales identifique el arribo de las fases P y S en el sismograma. Con el tiempo de estos arribos por lo menos en tres estaciones se efectúa la localización. Sin embargo, es criterio del postulante que este método, aunque factible, puede ser inexacto pues depende en gran medida de la geometría de la red y puede incluir errores humanos.

De igual manera, Cutié, Salgado y Palau (2023) afirman que las técnicas comúnmente utilizadas presentan limitaciones importantes que pueden resultar en la no detección de sismos y en la generación de falsas detecciones. Aunque este procedimiento suele ser refinado posteriormente

por los expertos, ya que las identificaciones automáticas habitualmente presentan importantes imprecisiones. Entre éstas, las más importantes son, su falta de éxito reconociendo sismos que tienen una razón señal-ruido baja. Otro problema con este método es que los filtros para su ajuste deben ser constantemente modificados manualmente, debido a los cambios en el ruido de las estaciones. Este procedimiento es impreciso pues la metodología usada es de ensayo y error.

Ahora bien, con el desarrollo de las Tecnologías de la Información y las Comunicaciones (TICs) diversas ciencias se han integrado en pos de explotar sus potencialidades, este es el caso de las Geociencias (GS) y la Inteligencia Artificial (IA). Es así que a decir del investigador Siervo (2019), en el año 2018 ocurrió un despegue en las investigaciones de esta temática. Literalmente la comunidad científica se "lanzó" a socializar sus resultados en las más prestigiosas revistas.

Consecuentemente, un gran número de artículos en este último campo de estudio fueron publicados, destacando Cuéllar, Suárez, y Espinosa (2018) los que propusieron la determinación rápida de sismos con una estación usando redes neuronales.

De igual forma, Perol, Gharbi y Denolle (2018) usaron redes neuronales convolucionales para la detección de sismos y Ross et al. (2018) para la identificación del arribo de la onda P y su polaridad, todos con una sola estación. Siguiendo con esta línea de propuestas, Lomax, Michelini y Jozinovic (2019), usaron formas de onda para la caracterización rápida de terremotos mediante redes neuronales. Estas investigaciones aun cuando lograron su propósito, no automatizaron por completo el proceso de determinación epicentral.

En otro orden de ideas, Ochoa, Vargas y Niño (2019) implementaron máquinas de soporte vectorial, para la determinación rápida de coordenadas geográficas de terremotos en una sola estación. Estos autores lograron resultados positivos pues a diferencia de los anteriores utilizaron la morfología de las formas de ondas registradas por una estación, aspecto este con el que coinciden los autores.

Con otro enfoque, Böse et al. (2017) propusieron un framework probabilístico para detectar sismicidad usando una sola estación, al igual que Mousavi y Beroza (2020), quienes se apoyaron en aprendizaje bayesiano profundo para estimar la localización de un terremoto con las observaciones de una sola estación. También se deben destacar, los resultados de Seula et al. (2023), quienes propusieron la determinación automática del azimut para un sistema de alerta temprana de terremotos usando una sola estación. Por su parte, Abdalzaher et al. (2024),

desarrollaron varias máquinas de aprendizaje para la determinación de un sismo en tiempo cuasi real y realizaron comparaciones estadísticas para seleccionar la de mejor desempeño.

Estas propuestas, constituyen referentes para la temática que se aborda, aunque no llegaron a integrar completamente todo el proceso de determinación epicentral usando una sola estación. En resumen, la sistematización realizada ha revelado las limitaciones que presentan las principales propuestas existentes para la determinación del azimut de un terremoto usando una estación sismológica y técnicas de Inteligencia artificial. De aquí que resulte necesario profundizar en sus particularidades.

En tal sentido, las perspectivas de análisis para abordar la determinación del azimut de un terremoto usando una sola estación sismológica y técnicas de inteligencia artificial se fundamentan en una línea principal denominada Maquinas de Aprendizaje (Machine Learning: ML, por sus siglas en inglés), está línea a su vez se subdivide en tres enfoques principales, el primero orientado a las Máquinas de Soporte Vectorial (SVM: siglas en inglés), el segundo a las Redes Neuronales Artificiales (ANN: siglas en inglés) y el tercero, menos representado en la literatura, a la utilización de otras máquinas de aprendizaje, tales como: árboles de decisión (decision tree), impulso de gradiente extremo (Extreme Gradient boosting), bosque aleatorio (Random Forest), algoritmo del vecino más cercano (K-Nearest Neighbors: KNN), entre otros. Dentro de los partidarios del primer enfoque destacan (Barros y Barros, 2016; Zhang et al., 2016; Asim et al., 2018; Ochoa, Niño y Vargas, 2019) quienes consideran que el problema a tratar tiene su solución en la optimización de los hiperplanos de decisión que brindan las SVMs.

Es así, que en este enfoque destaca la propuesta de los investigadores Barros y Barros (2016), quienes propusieron un método automático basado en SVM para la detección de eventos sísmicos usando una sola estación, localizada en un área de alta sismicidad provocada por la actividad minera. Los autores refieren que los resultados no tuvieron mucha efectividad, pero el tiempo de ejecución fue muy bueno. Al respecto se considera que esta propuesta no logra el objetivo principal referido a mejorar la determinación epicentral.

Por otra parte, Zhang et al. (2016) desarrollaron un sistema de alerta temprana de terremotos usando SVM para la región de Fujian, China. Este sistema usa una estación localizada en esa región. Los resultados evidenciaron una alta efectividad en la localización de los sismos. Sin embargo, la demora computacional fue grande, lo que hace que esta propuesta no sea práctica.

En este mismo orden de ideas, Asim et al. (2018), propusieron un modelo de predicción de terremotos usando SVM el cual funciona como un algoritmo de regresión que permite estimar posibles ocurrencias de sismos. Sin embargo, aunque este modelo es muy novedoso, sus resultados no fueron buenos en comparación con los referidos por la literatura.

De igual manera, los investigadores colombianos Ochoa, Vargas y Niño (2019) desarrollaron un nuevo enfoque para estimar el azimut de llegada de terremotos utilizando registros sismológicos de la estación "El Rosal", cercana a la ciudad de Bogotá, Colombia, mediante la aplicación de SVM. Sin embargo, los resultados de las estimaciones no fueron buenos comparados con otros.

Ahora bien, como partidarios del segundo enfoque aparecen (Cuéllar, Suárez y Espinosa, 2018; Perol, Gharbi y Denolle, 2018; Ross et al. 2018; Lomax, Michelini y Jozinovic, 2019) quienes proponen utilizar como vía de solución al problema de la determinación epicentral con una estación sismológica, las potencialidades que brindan las redes neuronales artificiales (ANN) principalmente su paralelismo, para el procesamiento de la clasificación y reconocimiento de patrones. Hay que señalar que la mayoría de los resultados que aparecen publicados en la literatura especializada utilizan este enfoque.

Consecuentemente, Cuéllar, Suárez y Espinosa (2018) propusieron la determinación rápida de sismos a partir de una estación mediante ANN que dividen en pequeños clusters para explotar mejor su paralelismo. Se debe señalar que los autores obtuvieron resultados favorables, pero la demanda computacional fue muy alta por lo que en la práctica se hace difícil su introducción.

Por otro lado, los investigadores Perol, Gharbi y Denolle (2018) implementaron una red neuronal convolucional para la determinación y localización de sismos, la que usa múltiples capas profundas que le permiten obtener muy buenos resultados en la estimación del azimut y otros parámetros epicentrales. Sin embargo, el tiempo de cómputo y respuesta fue grande de acuerdo a las necesidades que se plantea el problema de la determinación epicentral.

También, Ross et al. (2018) propusieron una generalización de la detección de las fases sísmicas con aprendizaje profundo usando redes neuronales convolucionales. Esto fue efectivo pues estimó el azimut del sismo y una vez entrenada su tiempo de ejecución fue despreciable. De igual manera, Lomax, Michelini y Jozinovic (2019) utilizando la morfología de las ondas de una sola estación, lograron una caracterización eficiente y eficaz de los terremotos mediante ANN.

Estos autores integraron las potencialidades de las redes con la optimización de los parámetros de entrada creando vectores característicos de cada señal según la forma de onda. Hay que precisar que se coincide con esta propuesta y se utilizará como referente principal en esta investigación.

Como representantes, del tercer enfoque aparecen (Böse et al., 2017; Mousavi y Beroza, 2020; Seula et al., 2023; Abdalzaher et al., 2024) quienes proponen la utilización de maquinas de aprendizaje alternativas para enfrentar el problema de la localización epicentral de un terremoto. Por ejemplo, Böse et al. (2017) abordan el problema de la localización con una única estación desarrollando un marco probabilístico que combina estimaciones de localización a partir de múltiples algoritmos para estimar la función de densidad de probabilidad (PDF) para la distancia epicentral, el azimut posterior y el tiempo de origen. Esta estrategia, aunque acertada, solo funciona para sismos de grandes magnitudes, por lo que no es aplicable a los datos de la presente investigación donde la mayoría son eventos pequeños.

En el caso de Abdalzaher et al. (2024) incorporan varias máquinas de aprendizaje (árboles de decisión, bosque aleatorio, algoritmo del vecino más cercano) para analizar formas de onda de movimiento débil. La idea clave es utilizar solo una ventana de 3 s de la forma de onda (1 s antes de la onda p y 2 s después), simplificando la implementación y asegurando la efectividad. Al extraer características relevantes de la forma de onda sísmica, el sistema puede tomar decisiones rápidas y confiables utilizando técnicas ML. Hay que señalar que esta estrategia les permite a los autores obtener buenos resultados. Sin embargo, presenta el inconveniente que necesita un número muy grande de datos para el proceso de entrenamiento.

En resumen, al profundizar en las principales tendencias relacionadas con la inteligencia artificial y la sismología, se evidenció que los trabajos están dirigidos a tratar el problema tanto de clasificación como de regresión y el algoritmo más usado son las redes neuronales. Lo anterior posibilitó considerar como objetivo de la investigación, el diseño e implementación de una red neuronal artificial para la determinación automática del azimut de un terremoto usando una estación sismológica.

METODOLOGÍA

Para dar cumplimiento al objetivo, y considerando que se utilizará en la investigación un enfoque cuantitativo, se trazaron dos etapas, la cuales se explican a continuación. En la primera se realizó

la extracción de los datos del catálogo de terremotos del Servicio Sismológico Nacional de Cuba (SSNC), así como el preprocesamiento de los mismos para hacerlos compatible con las entradas de la red neuronal que se propone. En la segunda se procedió al diseño, implementación, entrenamiento y validación de la red neuronal convolucional 1D.

Extracción y preprocesamiento de los datos.

La investigación realizada toma como base la estación sismológica de Rio Carpintero (RCC) situada al sureste de la ciudad de Santiago de Cuba. La cercanía a la principal zona generadora de terremotos del país, las características geológicas del sitio en el que se encuentra ubicada, así como los años de explotación de la misma, hacen de esta una de las más importantes de la red de estaciones del SSNC (Salgado et al., 2021).

Se propone realizar un entrenamiento supervisado, donde la etiqueta o variable a predecir es el azimut inverso de la localización del terremoto y las variables de entrada se toman a partir de la señal registrada por las componentes de la estación sismológica (Vertical, Norte-Sur, Este-Oeste). Los datos empleados se obtuvieron a partir del Catálogo de Terremotos del SSNC, soportado sobre el programa de análisis de datos sismológicos Seisan (Havskov y Ottemöller, 2023).

Se extraen de la estructura de directorios del Seisan terremotos localizados con alta confiabilidad, y para garantizar esto se establecieron varios criterios de selección (ver Tabla 1), cantidad de estaciones, error cuadrático medio (RMS), periodo de tiempo y área geográfica.

Tabla 1. Parámetros empleados para la selección de los terremotos utilizados del Catálogo de Cuba. Fuente: Autores.

Parámetros	Valor
Base de Datos	CUBA_
Tiempo inicial (aaaa/mm/dd/hh:mm:ss)	2011/01/01/00:00:00
Tiempo final (aaaa/mm/dd/hh:mm:ss)	2021/12/31/23:59:59
Mínimo número de estaciones	3
Latitud mínima (grados)	19.000
Latitud máxima (grados)	21.500
Longitud mínima (grados)	-78.500
Longitud máxima (grados)	-73.000
Máximo Error Cuadrático Medio (RMS) de la solución	0.7

Estaciones Sísmicas que deben estar presente en la localización	RCC (Rio Carpintero)

La creación de la base de datos final se lleva a cabo empleando el lenguaje de programación Python, con el respaldo de la biblioteca Obspy, que ofrece un entorno destinado al procesamiento de datos sismológicos. Es válido señalar, que esta base de datos contiene eventos normalizados y organizados de manera que pueden ser de utilidad para futuras investigaciones que incluyan a la Estación de Rio Carpintero.

Es así que, como resultado se obtienen 16411 eventos sísmicos que cumplen con los parámetros de selección establecidos. Para garantizar una mayor homogeneidad de los datos y eliminar valores aislados no representativos que pudieran alterar el proceso de aprendizaje de la red neuronal, se realizó un segundo proceso de selección y se tomaron solo aquellos eventos localizados a una distancia no mayor a 100 Km, magnitud local menor o igual a 3 grados Richter y profundidad no mayor de 50 Km.

La Figura 1 presenta la correlación entre las variables resultantes, destacando la mayor uniformidad de los datos. Sin embargo, se observa un desequilibrio en el conjunto de datos con respecto a la variable que se debe predecir, el azimut. Algunos valores de azimut tienen una cantidad reducida de datos, mientras que otros tienen una cantidad significativamente mayor. Para abordar esta situación, se ha aplicado la técnica de aumento de datos mediante el escalamiento. De esta manera, se han generado nuevas señales para los valores de azimut que cuentan con un número limitado de muestras, manteniendo los patrones de las señales originales a partir de las cuales se derivaron. La Figura 2 ilustra un ejemplo de esta estrategia.

Como resultado de aplicar las técnicas de aumento de datos se obtiene un nuevo conjunto de 49233 eventos. Este conjunto se divide en 39386 eventos para el entrenamiento que representan el 80% de los datos y los restantes 9847 eventos (20 %) para la validación.

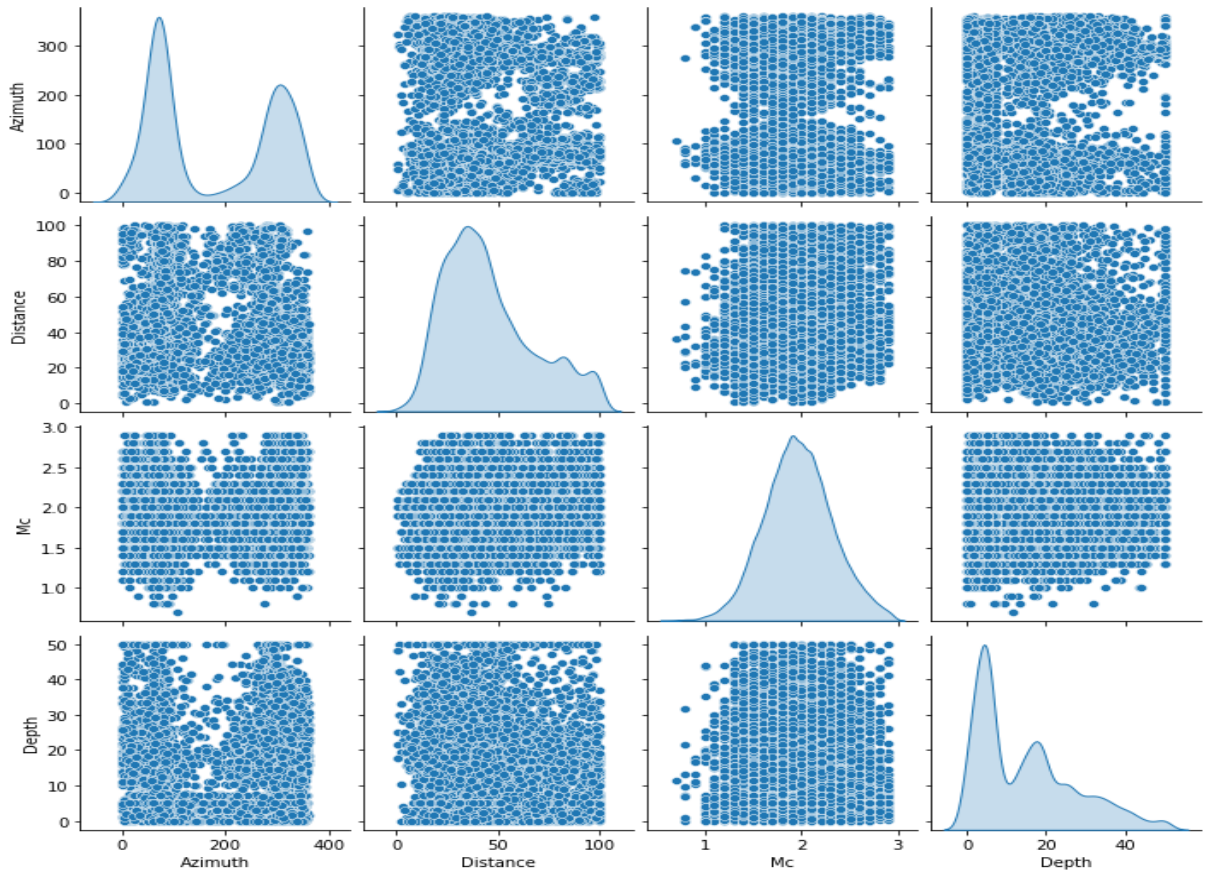
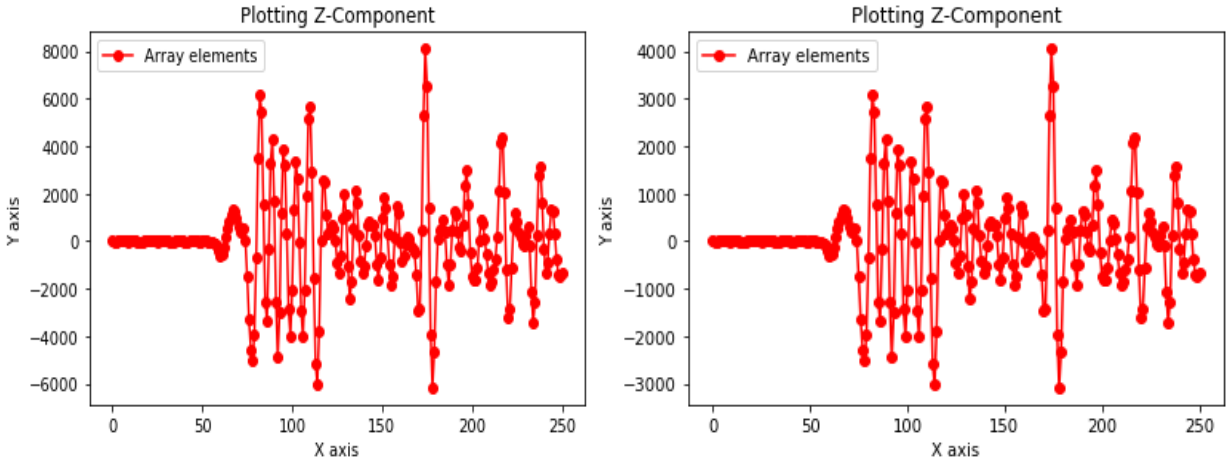


Figura 1. Correlación entre las principales variables del conjunto de datos, azimut, Distancia, Magnitud Local, Profundidad.



a) b)

Figura 2. Ejemplo de señal obtenida a partir de uso de la técnica de escalamiento para el aumento de datos. a) Señal original correspondiente a la componente vertical (Z). b) Señal aumentada a partir de un factor de escalamiento de 0.5.

Conformación de los conjuntos de datos de entrenamiento y de evaluación

La correcta división del conjunto de datos en subconjuntos de entrenamiento y validación es un paso crucial en el proceso de desarrollo de modelos de aprendizaje automático. En esta investigación, se aborda esta tarea con el propósito de garantizar que tanto el conjunto de entrenamiento como el conjunto de validación presenten una distribución balanceada de características relevantes, además de la variable a predecir, que en este caso es el azimut. Esta estrategia de división se fundamenta en el uso de la técnica StratifiedShuffleSplit de la librería scikit-learn.

En lugar de realizar una partición aleatoria simple, se busca mantener una proporción constante entre las variables críticas en ambos conjuntos. Estas variables; para este caso el azimut, la Distancia, la Magnitud y la Profundidad, pueden determinar el tipo de patrón de ondas de un terremoto que se recibe en una estación sismológica. Para lograr este objetivo, se implementó la metodología de StratifiedShuffleSplit, que permite distribuir de manera equitativa las observaciones en función de estas características clave. Esta técnica asegura que tanto el conjunto de entrenamiento como el de validación conserven una representación similar de las condiciones reales presentes en los datos.

La partición se realiza en una proporción donde el 80% de los datos se destina al conjunto de entrenamiento y el 20% restante al conjunto de validación. Se obtiene un conjunto de entrenamiento de 39386 datos y un conjunto de validación de 9847 datos.

RESULTADOS

Conformación de los vectores de entrada y salida

El modelo de red neuronal propuesto posee dos entradas y dos salidas (Mousavi y Beroza, 2020). Por una de las entradas recibe una estructura matricial compuesta por tres vectores distintos. Estos vectores representan los canales correspondientes al registro sísmico en sus componentes Vertical, Norte-Sur y Este-Oeste. Cada uno de estos canales se compone de 200 muestras, derivada de la consideración de tomar una ventana temporal de 2 segundos. Esta ventana abarca

desde 0.5 segundos previos a la llegada de la onda P hasta los 1.5 segundos posteriores, y la señal es muestreada a una frecuencia de 100 muestras por segundo.

Por la otra entrada, simultáneamente la red neuronal recibe una matriz 7x3 conformada a partir de la matriz de covarianza de las señales de los canales Vertical, Norte-Sur y Este-Oeste, en la ventana de tiempo de 2 segundos junto al vector y valor propios correspondientes.

El problema de esta investigación se aborda como un problema de regresión, siendo el valor a predecir el azimut. Es importante resaltar que la predicción de ángulos en un rango de 0 a 360 grados no puede ser tratada de la misma manera que otros tipos de variables, debido a su naturaleza circular (Mousavi y Beroza, 2020).

Por consiguiente, la utilización de la función de pérdida del error cuadrático medio (MSE) comúnmente utilizada en la regresión no debe ser utilizada para abordar problemas de este tipo (Mousavi y Beroza, 2020). En este trabajo, en lugar de tratar directamente con el azimut, se trabaja con sus equivalentes cos(azimut) y sin(azimut), que juntos forman las componentes de un vector unitario proyectado. Mediante esta técnica, la secuencia de valores angulares que abarca el intervalo de 0 a 2π resultará en la construcción de una circunferencia con un radio unitario como se muestra en la Figura 3.

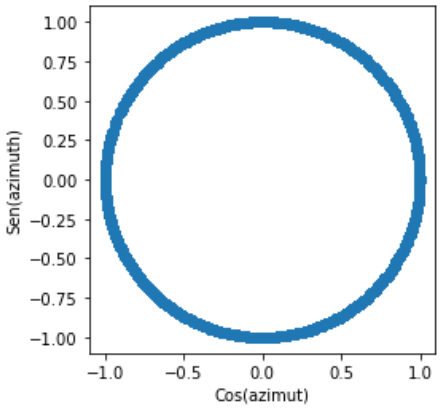


Figura 3 Proyección de los valores cos(azimut) *vs* seno(azimut) para valores de azimut de 0 a 2π

Diseño e implementación de la Red Neuronal

Se propone una red convolucional 1D con una entrada y dos salidas. Como herramienta utilizada en su implementación se utiliza Tensorflow en su versión 2.8.

La Figura 4 muestra el esquema de la red neuronal 1D. Donde la capa de entrada recibe una matriz de dimensiones 200 x 3 que corresponde a la ventana de la señal sísmica. A continuación,

se disponen cuatro capas de Convolución 1D; las dos primeras de ellas utilizan 40 filtros cada una, y los dos siguientes 64 filtros cada una. Todas las capas de convolución 1D emplean un kernel con un tamaño de 3, un paso (stride) de 1 y utilizan la función de activación RELU.

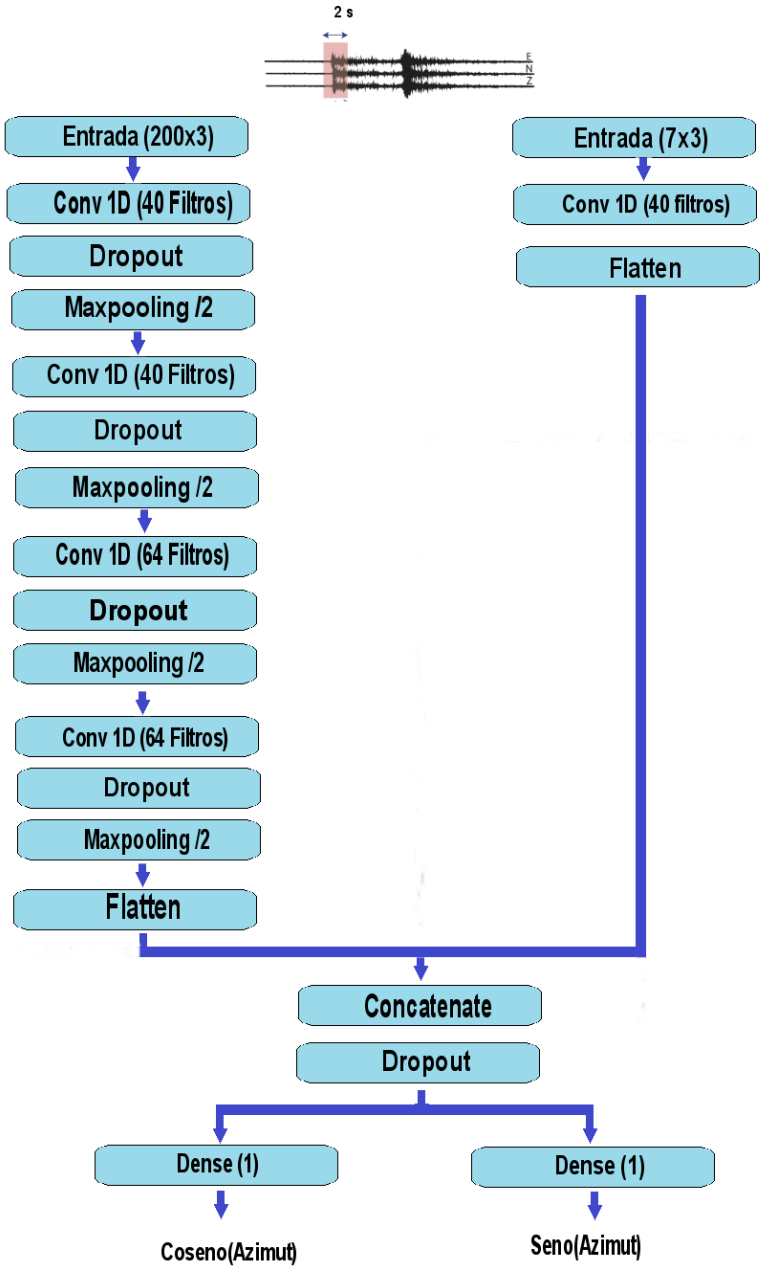


Figura 4. Esquema de la red neuronal convolucional diseñada e implementada. Fuente:

Después de cada capa de convolución, se inserta una capa de dropout, con el propósito de prevenir el sobreajuste durante el proceso de entrenamiento. Para cada capa de convolución, se

aplica también una capa de agrupamiento del tipo max-pooling con un tamaño de agrupación (pool_size) de 2. Posterior a la cuarta etapa, se coloca una capa de aplanamiento (Flatten). La segunda entrada recibe una matriz 7x3 obtenida a partir de la matriz de covarianza, el vector propio y el valor propio de los tres canales en la misma ventana de tiempo. Esta se introduce en una capa de convolución 1D de 40 filtros, kernel 3 y función de activación RELU, seguida de una capa de aplanamiento.

La salida de ambas ramas, tomadas a partir de sus respectivas capas de aplanamiento se concatenan en una capa Densa a partir de las cuales se obtienen las respectivas salidas cos (azimut) y sen (azimut).

El modelo utiliza dos funciones de pérdida, una para cada salida de predicción, ambas utilizando el Error Cuadrático Medio (MSE). El optimizador utilizado es Adam, con una tasa de aprendizaje de $1e^{-4}$. Además, se ha implementado una función personalizada para calcular el Coeficiente de Determinación R^2, que junto con el Error Absoluto Medio (MAE), son utilizados como métricas durante el proceso de entrenamiento.

La Figura 5 muestra la curva de entrenamiento del modelo. En la misma se puede apreciar la calidad del ajuste durante el entrenamiento.

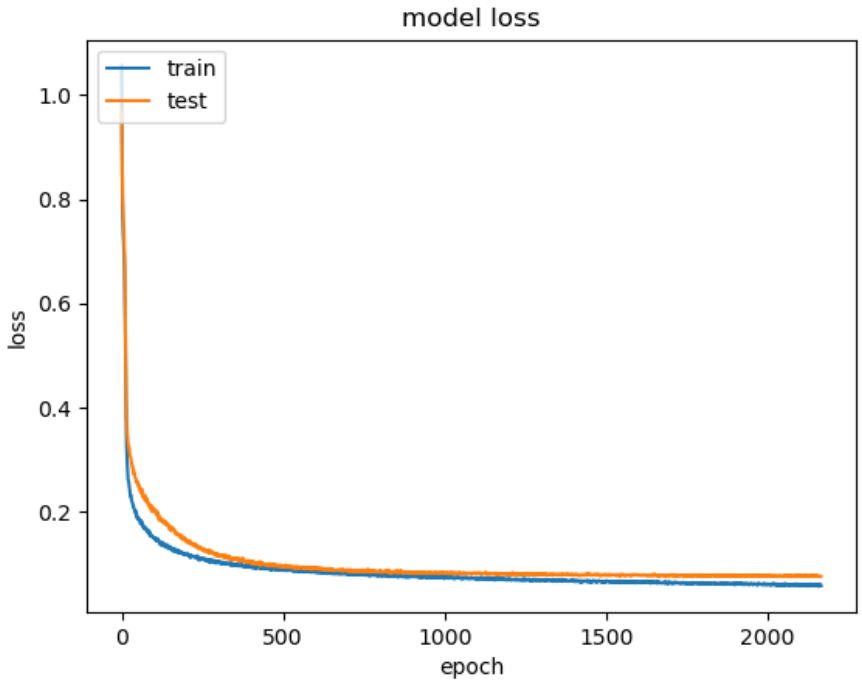


Figura 5. Gráfica de pérdidas (Error Cuadrático Medio) en función de las épocas del entrenamiento.

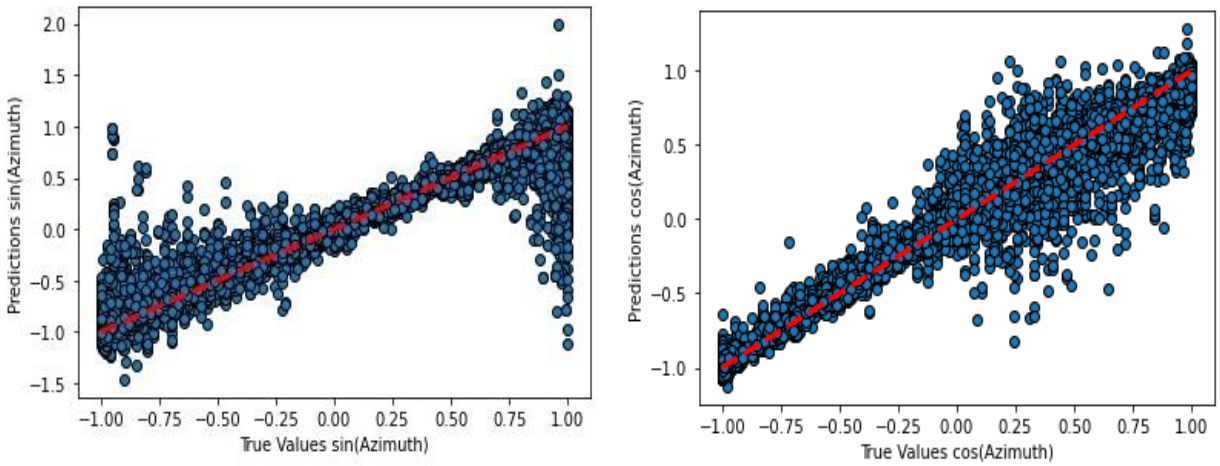
La Tabla 2 resume los resultados del entrenamiento. Se presentan los valores de pérdida, Error Cuadrático Medio (MSE), correspondientes al modelo de regresión propuesto, incluyendo tanto su evaluación general como de las salidas individuales de coseno y seno del azimut. Además, se exponen las métricas de entrenamiento, como el Error Absoluto Medio y el Coeficiente de Determinación (R^2). Esta comparativa proporciona una visión integral de cómo el modelo se comporta en datos inéditos y cómo se alinean sus resultados con las métricas de evaluación establecidas; se pone de manifiesto la capacidad de generalización del modelo entrenado.

Tabla 2. Comparativa de los resultados del entrenamiento y validación. Fuente: Autores.

Parámetro	Entrenamiento	Validación	Error
Pérdida (*Loss*) (MSE)	**0.1085**	**0.1059**	**0.0026**
Pérdida de la salida cos_azimuth (MSE)	0.0431	0.0440	0.0009
Pérdida de la salida sin_azimuth (MSE)	0.0654	0.0618	0.0036
Salida cos (MAE)	0.1488	0.1582	0.0094
Salida sin (MAE)	0.1600	0.1695	0.0098
Salida cos Coeficiente de determinación	0.8999	0.9062	0.0063
Salida sin Coeficiente de determinación	0.8689	0.8726	0.0037

De igual forma, se utiliza el modelo para generar predicciones basadas en los datos del conjunto de validación. Estas predicciones se comparan con los valores reales previamente establecidos en el mismo conjunto. En la Figura 6 (a y b), se pueden observar los gráficos que presentan la relación entre las predicciones y los valores reales para las salidas asociadas al seno y al coseno del azimut, respectivamente. En ambas figuras, se destaca cómo las predicciones se agrupan alrededor de la línea punteada, que representa el resultado ideal de alineación. Este patrón señala la capacidad del modelo para realizar predicciones cercanas a los valores reales.

Usando las predicciones obtenidas para el coseno y el seno del azimut, se calcula el valor del azimut, que es el objetivo central de esta investigación. La Figura 7 ilustra la correlación entre estas predicciones calculadas y los valores reales del azimut, que son conocidos en el conjunto de validación. Además, se evidencia cómo las predicciones se agrupan alrededor de la línea punteada, la cual representa la alineación ideal. Este patrón subraya la capacidad del modelo para realizar pronósticos que se acercan a los valores reales con eficacia.



a) b)

Figura 6. a) Grafica de las Predicciones de la Salida seno(azimut) vs los valores reales. b) Grafica de las Predicciones de la Salida coseno(azimut) vs los valores reales. Fuente: Autores

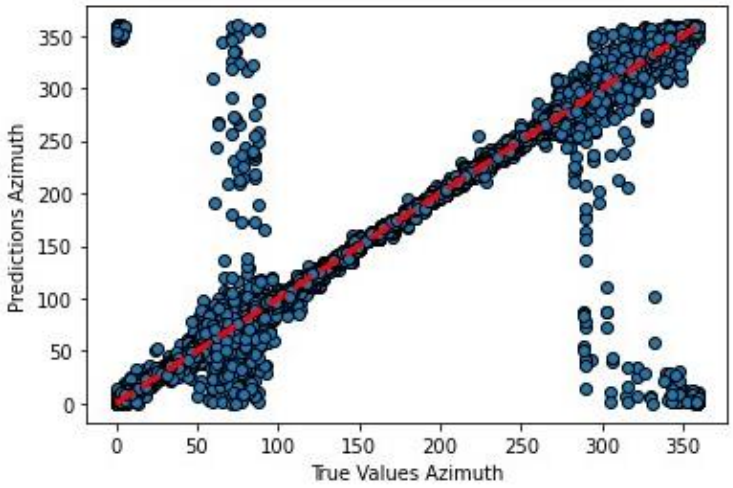


Figura 7. Grafica de las Predicciones del azimut *vs* los valores reales. Fuente: Autores.

DISCUSIÓN

En esta sección, se analizan los hallazgos del enfoque propuesto y se muestra su eficacia entre los existentes relacionados en la literatura. De hecho, es criterio de los autores, y se puede observar, que, al emplear la metodología propuesta, se logró entrenar el modelo de manera efectiva sin que se presentaran problemas de sobreajuste a los datos utilizados. Coincidiendo estos resultados con los de Lomax, Michelini y Jozinovic (2019). Esto se aseguró en gran medida gracias al uso de

dropouts como una técnica de regularización en varias etapas de la red neuronal, así como la implementación de la técnica de aumento de datos para equilibrar el conjunto de entrenamiento. Esta combinación de ambas técnicas constituye un aporte de esta investigación, pues en la sismología si bien, es común que haya pocos datos, no abundan las investigaciones en las que se realicen completamientos para lograr conjuntos homogéneos. Por lo general, los investigadores trabajan con los datos exactos con los que cuentan. De ahí, la importancia de estos resultados.

Por otra parte, el coeficiente de determinación R^2 para la regresión alcanzó 0.9, lo que implica que aproximadamente el 90 % de la variabilidad en la variable azimut puede ser explicada por el modelo de regresión empleado. A juicio de los autores, esta proporción señala la eficacia del modelo para capturar la relación entre las variables independiente y dependiente. Es importante destacar que este valor es coherente con hallazgos previos en investigaciones similares, como el estudio llevado a cabo por Mousavi y Beroza (2020), donde se obtuvo un R^2 del 87%.

De igual manera, estos resultados superan los obtenidos por Dickey, Borghetti y Junek (2021) que fueron de un 60% para una red neuronal profunda que toma un sismograma de tres componentes y produce tanto una predicción de azimut como la medida de certeza correspondiente. Más importante aún, en el presente estudio, es que cada estimación está acompañada de una medida de certeza robusta, lo que permite la selección de predicciones de alta confianza para ser transmitidas al algoritmo de localización epicentral. Lo que, a criterio de los autores, constituye una mejora considerable para este tipo de problema en sismología.

En este mismo orden de ideas, la red convolucional 1D que se implementó, permite capturar de manera efectiva las características clave de los trazos de las formas de onda, pues el modelo propuesto proporciona una estimación confiable de la intensidad del terremoto, logrando una alta tasa de precisión basada en las tres componentes. Estos resultados son comparables con los de Abdalzaher et al. (2024) quienes logran un 99,05 % de precisión a partir de cualquiera de las tres componentes usando una sola estación sismológica. Lo que sugiere su posible utilidad, por ejemplo, en los sistemas de alerta temprana.

Consecuentemente, cuando se evalúa el modelo en el conjunto de validación, que comprende el 20 % de todos los datos y, por lo tanto, consiste en datos previamente no utilizados por el sistema, se observan resultados consistentes con los obtenidos durante el entrenamiento. Esto

resalta la habilidad del modelo propuesto para generalizar datos no vistos anteriormente, lo que evidencia la calidad de la red 1D diseñada e implementada. Algo similar ocurre en relación a que la metodología desarrollada, logra mejorar la confiabilidad de la estimación de azimut como un esquema complementario a otros métodos. Aspecto este que coincide con los trabajos de Weijia, Hrvoje y Qingya (2024) y Abdalzaher et al. (2024), pues hay que precisar que el objetivo de estas herramientas no es suplantar los métodos de determinación epicentral, sino ofrecer de manera rápida y automática la estimación de algunos parámetros, en este caso el azimut, para aumentar la precisión de la localización.

Otro aspecto a destacar, es que se logra la determinación rápida y automática del azimut de un sismo y posteriormente su localización a partir de una estación con muy poca demanda computacional y tiempo de cómputo. Además, el enfoque desarrollado se puede utilizar para advertencias rápidas e incorporarlo a sistemas de alerta temprana, similar a la propuesta de Seula et al. (2023). Lo anterior, hace esta propuesta novedosa y la diferencia de los resultados de Cuéllar, Suárez y Espinosa (2018) y Perol, Gharbi y Denolle (2018), quienes, si bien obtuvieron buenos resultados, las herramientas desarrolladas requieren una mayor demanda computacional y tiempo de respuesta.

CONCLUSIONES

Se diseñó e implementó una red neuronal convolucional 1D que, a partir de una estación sismológica con tres componentes, es capaz de determinar de manera automática el azimut de un terremoto con una efectividad del 90 %, evidenciado en los valores del coeficiente de determinación R^2.

La investigación, considerando las pruebas y validaciones realizadas, proporciona una base para futuros avances en este campo. El trabajo posterior se centrará en mejorar la metodología entrenándola con formas de onda del resto de las estaciones sismológicas de Cuba, lo que podría mejorar su escalabilidad y aplicabilidad. Este paso es particularmente importante para las regiones con limitado cubrimiento de sensores sísmicos, como las regiones central y occidental, ya que aprovecha los datos existentes de sismos de baja energía y magnitud para expandir la cobertura y la precisión de la red.

Como resultado del preprocesamiento de los datos a utilizar en el entrenamiento y validación de la red neuronal, se obtuvo una base de datos que contiene eventos normalizados. Estos eventos están organizados de manera que pueden ser de utilidad para futuras investigaciones.

REFERENCIAS BIBLIOGRÁFICAS

Abdalzaher, M.S., Soliman, M.S., Krichen, M., Alamro, M.A., & Fouda, M.M. (2024). Employing Machine Learning for Seismic Intensity Estimation Using a Single Station for Earthquake Early Warning. *Remote Sens.* 2024, *16*, 2159. https://doi.org/10.3390/rs16122159

Asim, K.M., Idris, A., Iqbal, T., & Martínez, F. (2018). Seismic indicators-based earthquake predictor system using Genetic Programming and AdaBoost classification. Soil Dynamics and Earthquake Engineering, 111, 1–7. https://doi.org/10.1016/j.soildyn.2018.04.020

Barros, V., & Barros, L. (2016). Support vector machine model for automatic detection and classification of seismic events. In EGU General Assembly Conference Abstracts (pp. EPSC2016-9256). https://doi.org/10.1029/2019JB018132

Böse, M., Clinton, J.F., Ceylan, S., Euchner, F., van Driel, M., Khan, A., Giardini, D., Lognonné, P., &, Banerdt, W.B. (2017). A probabilistic framework for single-station location of seismicity on Earth and Mars. Physics of the Earth and Planetary Interiors, 262, 48–65. https://doi.org/10.1016/j.pepi.2016.11.003

Cuéllar, A., Suárez, G., & Espinosa, J. M. (2018). A fast earthquake early warning algorithm based on the first 3 s of the P-wave coda. Bulletin of the Seismological Society of America, 108(4), 2068–2079. https://doi.org/10.1785/0120180079

Cutié, M., Salgado, A., & Palau, R. (2023). Determinación automática del azimut de un terremoto usando redes neuronales artificiales. In Digital Scientific Journal: Proceedings of the X International Cuban Convention of Earth Sciences. “Geociencias´2023”, pág. 17. https://drive.google.com/drive/folders/14fNGNawt3JLcGsYRxnIpyew02HTA0-vs

Dickey, J., Borghetti, B., & Junek, W. (2021). BazNet: A deep neural network for confident three-component backazimuth prediction. Pure and Applied Geophysics, 178(7), 2459–2473. https://doi.org/10.1007/s00024-020-02578-x

Diez, E.D., Priolo, E., Sandron, D., Poveda, V., Cattaneo, M., Marzorati, S., & Palau R. (2022). Evaluation of the Event Detection Level of the Cuban Seismic Network. *Seismological Research Letters*; 93 (4): 2048–2062. doi: https://doi.org/10.1785/0220220016

Eisermann, A.S., Ziv, A., & Wust, G.H. (2015). Real-time back azimuth for earthquake early warning. Bulletin of the Seismological Society of America, 105(4), 2274–2285. https://doi.org/10.1785/0120140298

Havskov, J., & Ottemoller, L. (2023). Routine data processing in earthquake seismology: With sample data, exercises and software. Update. (2023a ed.). Springer.

Iturralde, M., & Arango, E.D. (2020). Ciudades de Cuba en alerta sísmica. Minería y Geología, 36(4), 366–376. https://revista.ismm.edu.cu/index.php/revistamg/article/view/art1_No4_2020/1534

Lomax, A., Michelini, A., & Jozinović, D. (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. Seismological Research Letters, 90(2A), 517–529. https://doi.org/10.1785/0220180311

Mousavi, S.M., &, Beroza, G.C. (2020). Bayesian-deep-learning estimation of earthquake location from single-station observations. IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society, 58(11), 8211–8224. https://doi.org/10.1109/tgrs.2020.2988770

Noda, S., Yamamoto, S., Sato, S., Iwata, N., Korenaga, M., & Ashiya, K. (2012). Improvement of back-azimuth estimation in real-time by using a single station record. Earth, planets, and space: EPS, 64(3), 305–308. https://doi.org/10.5047/eps.2011.10.005

Ochoa, L. H., Vargas, C. A., & Niño, L. F. (2019). Fast estimation of earthquake arrival azimuth using a single seismological station and machine learning techniques. Earth sciences research journal, 23(2), 103–109. https://doi.org/10.15446/esrj.v23n2.70581

Palau, R., González O.F., Salgado, A., Arango, E.D., Diez, E.R., Cutié, M., Montenegro, C., & Poveda, V. (2023). Estudio de los movimientos recientes de la corteza en el margen

suroriental de Cuba. In Digital Scientific Journal: Proceedings of the X International Cuban Convention of Earth Sciences. "Geociencias´2023", pág. 36. https://drive.google.com/drive/folders/14fNGNawt3JLcGsYRxnIpyew02HTA0-vs

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. Science Advances, 4 (2), e1700578. https://doi.org/10.1126/sciadv.1700578

Rosabal, S. Y., Salgado, A., Palau, R., & Leyva, M. (2024). Morfotectónica del sector Retiro-Sitio Campo de la región suroriental de Cuba. Minería y Geología, 39(3), 161–173. https://revista.ismm.edu.cu/index.php/revistamg/article/view/2394

Ross, Z. E., Meier, M.A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. Bulletin of the Seismological Society of America, 108(5A), 2894–2901. https://doi.org/10.1785/0120180080

Salgado A., & Palau, R. (2023). New macroseismic catalog of perceptible earthquakes in Cuba 2017 – 2020. Revista Internacional Mapping, 31 (207), 30–41. https://doi.org/10.59192/mapping.422

Salgado, A., Palau, R., Leyva, M., & Cutié, M. (2023). Learning machines evaluation for seismic signals classification. In Digital Scientific Journal: Proceedings of the X International Cuban Convention of Earth Sciences. "Geociencias´2023". pág. 22. https://drive.google.com/drive/folders/14fNGNawt3JLcGsYRxnIpyew02HTA0-vs

Salgado, A., Vicic, B., Palau, R., & Leyva, M. (2021). Red Neuronal artificial para la clasificación de señales sísmicas según su zona de origen. In Digital Scientific Journal: Proceedings of the IX International Cuban Convention of Earth Sciences. "Geociencias´2021". https://drive.google.com/drive/folders/15qwDMR7kJMXBFfoy5LWTa_wi2J4JLe6H

Seula Jung, Jung-Ho Park, Yun Jeong Seong, & Dong-Hoon Sheen (2023). Automatic Determination of Back Azimuth Based on a Single Mini Array for an Earthquake Early Warning System. *Bulletin of the Seismological Society of America* 2023; 113 (3): 1311–1323. doi: https://doi.org/10.1785/0120220129

Siervo, D. D. (2019). Uso de técnicas de aprendizaje automático en la detección de sismos. Universidad Internacional de La Rioja (UNIR) Escuela Superior de Ingeniería y Tecnología.

Weijia Sun, Hrvoje Tkalčić, & Qingya Tang (2024). Single-Station Back-Azimuth Determination with the Receiver Function Rotation Technique Validated by the Locations of Earthquakes, Impacts, and Explosions. *Seismological Research Letters* 2024; 95 (5): 2925–2938. doi: https://doi.org/10.1785/0220240117

Zhang, H., Jin, X., Wei, Y., Li, J., Kang, L., Wang, S., Huang, L., &, Yu, P. (2016). An earthquake early warning system in Fujian, China. Bulletin of the Seismological Society of America, 106(2), 755–765. https://doi.org/10.1785/0120150143

©2024 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional (CC BY-NC-SA 4.0) (https://creativecommons.org/licenses/by-nc-sa/4.0/).