

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA**

### **GENERACIÓN DE SEÑALES VOLCÁNICAS ARTIFICIALES DE TIPO LP (LONG-PERIOD) Y VT (VOLCANO-TECTONIC) A PARTIR DE UNA BASE DE DATOS DEL VOLCÁN COTOPAXI USANDO LA TÉCNICA DE BOOTSTRAPPING**

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO  
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE INGENIERA EN  
TELECOMUNICACIONES**

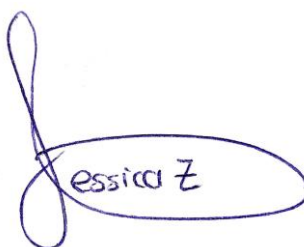
**JESSICA PAOLA ZAPATA YUGSI**  
**jessica.zapata@epn.edu.ec**

**DIRECTOR: Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO**  
**felipe.grijalva@epn.edu.ec**

**Quito, octubre 2022**

## CERTIFICACIONES

Yo, JESSICA PAOLA ZAPATA YUGSI declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A handwritten signature in blue ink. The signature is stylized, starting with a large loop on the left, followed by the letters 'essica Z' written in a cursive-like font. The entire signature is enclosed within a horizontal oval shape.

---

**JESSICA PAOLA ZAPATA YUGSI**

Certifico que el presente trabajo de integración curricular fue desarrollado por JESSICA PAOLA ZAPATA YUGSI, bajo mi supervisión.

A handwritten signature in blue ink. The signature is highly stylized and cursive, appearing to read 'Felipe' followed by a large, sweeping flourish that extends to the right.

---

**Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO**  
**DIRECTOR**

## **DECLARACIÓN DE AUTORÍA**

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el (los) producto(s) resultante(s) del mismo, son públicos y estarán a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

JESSICA PAOLA ZAPATA YUGSI

Ph.D. FELIPE LEONEL GRIJALVA ARÉVALO

## DEDICATORIA

A mis queridos padres, José y Laura quienes con mucho esfuerzo y sacrificio me brindaron la oportunidad de prepararme, a ustedes que con amor, fe y esperanza supieron inculcar en mi el carácter necesario para continuar durante todo el proceso.

A mi hermana quien siempre estuvo a mi lado brindándome su apoyo y comprensión cuando más lo necesitaba.

A ti, que estuviste en mis momentos tristes y felices, que me participaste una palabra de aliento cuando quise desfallecer en el intento.

A ti, que te habrás identificado al leer estas palabras.

A mis abuelos por derramar sobre mi bendiciones y oraciones.

***Jessica Paola Zapata Y.***

# AGRADECIMIENTOS

En primer lugar, agradezco infinitamente a Dios por no permitirme desfallecer en los momentos que pensé no poder continuar, por bendecirme y acompañarme en todo momento y lugar.

A mis padres José y Laura quienes siempre hicieron lo posible por brindarme su apoyo, compañía y comprensión durante toda mi etapa estudiantil. A ellos que supieron brindarme una voz de aliento y consejos para lograr superarme y culminar mi objetivo.

A Silvia, mi hermana, quien me brindo su compañía y apoyo incondicional cuando recurrí a ella.

A mi tutor, PhD. Felipe Leonel Grijalva quien en las aulas de la Escuela Politécnica Nacional supo impartir su intelecto y experiencia para llenarnos de conocimiento y prepararnos para enfrentar el campo laboral. A él que con su paciencia me acogió para realizar este proyecto de titulación, me guió y aconsejó durante la realización de toda mi tesis.

Finalmente, a la Escuela Politécnica Nacional y sus docentes por abrirme las puertas desde el día que elegí estudiar una carrera en tan prestigiosa entidad. Gracias por formarme y permitirme obtener mi título de tercer nivel.

# ÍNDICE DE CONTENIDO

CERTIFICACIONES .....	I
DECLARACIÓN DE AUTORÍA .....	II
DEDICATORIA .....	III
AGRADECIMIENTO .....	IV
ÍNDICE DE CONTENIDO.....	V
RESUMEN.....	VI
ABSTRACT .....	VIII
1. INTRODUCCIÓN .....	1
1.1. Objetivo general.....	3
1.2. Objetivos específicos.....	3
1.3. Alcance .....	3
1.4. Marco teórico .....	4
1.4.1. Volcán Cotopaxi.....	4
1.4.1.1. Fenómenos volcánicos.....	6
1.4.1.2. Monitoreo Volcán Cotopaxi.....	7
1.4.1.3. Zonas vulnerables por el Volcán Cotopaxi .....	8
1.4.2. Eventos Sismo-volcánicos .....	9
1.4.3. Procesamiento de Señales.....	10
1.4.3.1. Extracción de Características.....	10
1.4.4. Análisis Espectral .....	11
1.4.4.1. Transformada rápida Fourier STFT.....	11
1.4.4.2. Espectrograma.....	11
1.4.5. Herramienta Matlab.....	12
1.4.6. Técnica de Bootstrapping.....	13

1.4.6.1. Ventajas .....	14
1.4.6.2. Desventajas .....	14
2. METODOLOGÍA.....	15
2.1. BASE DE DATOS.....	15
2.2. PREPROCESAMIENTO.....	17
2.3. BOOTSTRAPPING .....	19
2.4. RECUPERACIÓN DE LA SEÑAL AL DOMINIO DEL TIEMPO .....	21
2.5. EXPERIMENTOS .....	22
3. RESULTADOS Y COMPARACIÓN.....	22
3.1. RESULTADOS.....	22
3.1.1. Resultados obtenidos al aplicar FD .....	23
3.2. COMPARACIÓN.....	24
4. CONCLUSIONES Y RECOMENDACIONES.....	27
4.1. CONCLUSIONES.....	27
4.2. RECOMENDACIONES.....	28
5. REFERENCIAS BIBLIOGRÁFICAS .....	29
ANEXOS .....	33

## RESUMEN

El estudio sismológico de un volcán es de vital importancia puesto que brinda amplia información acerca de su comportamiento y permite el análisis exhaustivo de su actividad volcánica interna. El monitoreo permanente de un volcán permite estimar, predecir futuros eventos, y es por esta razón, que una extracción de las características más relevantes de los eventos sísmicos es necesaria para accionar o crear estrategias de prevención ante catástrofes de esta índole.

Debido a la poca información procesada que se tiene se plantea usar la técnica de bootstrapping para generar eventos artificiales con similares características y así incrementar la cantidad de información nueva disponible. La información generada es muy útil ya que se genera en base a las necesidades de investigación requeridas, por ejemplo, un tipo de evento volcánico importante es un VT (Volcano-Tectonic); sin embargo, se dispone de muy pocos eventos etiquetados de este tipo considerando que los algoritmos de aprendizaje automático por lo general precisan de miles de ejemplos para el proceso de entrenamiento. Para aplicar la técnica de bootstrapping las señales originales son preparadas previamente en la etapa de preprocesamiento. El bootstrap es una herramienta estadística extremadamente poderosa y de amplia aplicación que se puede utilizar para determinar la incertidumbre asociada a un método de aprendizaje estadístico determinado. Se evalúa la similitud de las señales generadas mediante el uso de Frechet Distance (FD).

**PALABRAS CLAVE:** Base de datos, señales sintéticas, técnica bootstrapping, transformada rápida de fourier, volcán, Cotopaxi, distancia de frechet.



## ABSTRACT

The seismological study of a volcano is of vital importance since it provides ample information about its behavior and allows an exhaustive analysis of its internal volcanic activity. The permanent monitoring of a volcano allows estimating and predicting future events, and it is for this reason that an extraction of the most relevant characteristics of seismic events is necessary to activate or create prevention strategies in the face of catastrophes of this nature.

Due to the little processed information that is available, it is proposed to use the bootstrapping technique to generate artificial events with similar characteristics and thus increase the amount of new information available. The information generated is very useful since it is generated based on the required research needs, for example, a type of important volcanic event is a VT (Volcano-Tectonic); however, very few such labeled events are available considering that machine learning algorithms typically require thousands of examples for the training process. To apply the bootstrapping technique, the original signals are previously prepared in the preprocessing stage. The bootstrap is an extremely powerful and widely applicable statistical tool that can be used to determine the uncertainty associated with a given statistical learning method. The similarity of the generated signals is evaluated by using the Frechet Distance (FD).

**KEYWORDS:** Databases, Synthetic signals, bootstrapping technique, fast fourier transform, volcano, Cotopaxi, frechet distance.

# 1. INTRODUCCIÓN

Los sistemas de alerta temprana de volcanes se han convertido en un tema de investigación para preservar vidas humanas y pérdidas materiales. En este contexto, los criterios de detección de eventos sismo-volcánicos basados en la clasificación mediante técnicas de aprendizaje automático han demostrado su utilidad y se han propuesto varios sistemas en la literatura [1]. Es de gran importancia considerar que la información captada y registrada por los sensores es de tipo no procesada y requiere de un procesamiento y tratamiento minucioso, por lo que es necesario acudir a la destreza de un analista en vulcanología para la detección y clasificación de eventos, lo cual conlleva una inversión de tiempo y trabajo elevada. Por esta razón la necesidad de desarrollar una herramienta automática que permita fortalecer el procesamiento e informar de manera oportuna a los organismos competentes para una mitigación de los riesgos existentes frente a un posible proceso eruptivo [2].

El Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) se encarga de monitorear ininterrumpidamente a todos los volcanes del Ecuador [3] por medio de la red de sensores de alta precisión instalada en diferentes estaciones de monitoreo. En concreto, el volcán Cotopaxi es considerado uno de los volcanes más peligrosos del mundo debido a la frecuencia de sus erupciones, su estilo eruptivo, su relieve, su cobertura glaciaria y por la cantidad de poblaciones potencialmente expuestas a sus amenazas, su peligrosidad radica en que sus erupciones pueden dar lugar a la formación de enormes lahares (flujos de lodo y escombros) que viajarían por drenajes y afectarían directamente a la población más cercana, es por ello que a este volcán se dedican gran parte de los recursos para su frecuente monitoreo [4].

Considerando que los sensores instalados en las cercanías del volcán Cotopaxi captan y registran grandes cantidades de información no procesada y existe muy poca información de eventos detectados y clasificados [5]. La clave para lograr una optimización de este proceso es obtener una base de alta confiabilidad que permita un entrenamiento de modelos que clasifiquen y extraigan características representativas de las señales volcánicas.

Los datos de señales sismo-volcánicas captadas en las cercanías del volcán Cotopaxi son recopilados por el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN). Los registros de eventos sísmicos llevan información en el dominio del tiempo y frecuencia para el análisis y clasificación, en el dominio del tiempo se registra los eventos y su duración mientras que en el dominio de la frecuencia se identifica y clasifica el evento mediante el espectrograma. Evidentemente este proceso de clasificación conlleva mucho tiempo de tra-

bajo para los vulcanólogos y representa una limitante a la hora de procesar la información captada por los sensores cercanos al volcán Cotopaxi; por lo que el uso del aprendizaje de máquina disminuiría considerablemente el tiempo de procesamiento de la señal [6].

El procesamiento de señales vulcano-sísmicas marcan un punto clave para la detección temprana y prevención de desastres naturales por lo que la automatización de dicho procesamiento ahorraría tiempo [7]. Debido a la escasez de las señales de eventos sísmicos, la generación de señales con las mismas características de un evento de poca ocurrencia en el volcán Cotopaxi suplirá la ausencia de señales reales y permitirá diseñar un sistema de aprendizaje de máquina entrenado con estas señales generadas, disminuyendo la tarea de procesamiento e interpretación de datos reales.

La generación de nuevos eventos permitirá disponer de un conjunto amplio de datos para utilizar de forma óptima la gran cantidad de algoritmos de aprendizaje automático disponibles hoy en día. Además el personal del Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN) tendría grandes beneficios en su trabajo logrando que optimicen su tiempo en otro tipo de investigaciones, dejando el procesamiento de la información a la Inteligencia Artificial.

En este proyecto se pretende usar la técnica de bootstrapping para automatizar el monitoreo y solucionar los problemas ya descritos anteriormente, esta técnica tiene la finalidad de generar señales artificiales a partir de un conjunto de datos (señales disponibles) limitado [8]. La técnica de bootstrapping se encarga de la extracción de datos de muestra repetidamente con reemplazo de una fuente de datos para estimar, predecir un parámetro.

La actividad sismo-volcánica del Cotopaxi es registrada mediante los sensores existentes en su alrededor, son varias las señales captadas por estos instrumentos importantes [9] pero las más relevantes y que serán consideradas en este proyecto son:

- LP (Long period): Se caracterizan por presentar frecuencias bajas debido a la resonancia causada por la presión esporádica de los fluidos y gases en las grietas o en el conducto volcánico, involucran modos volumétricos de deformaciones. Las frecuencias generalmente son inferiores a los 5 Hz; ocurren con mayor frecuencia que los eventos Volcano Tectonic.
- VT (Volcano Tectonic): Se generan por el fracturamiento de las rocas adyacentes a los conductos volcánicos, tienen frecuencias predominantes entre los 5-15 Hz, aunque presentan en ocasiones frecuencias mayores [10].

## 1.1. Objetivo general

Generar señales volcánicas artificiales de tipo LP (Long-Period) y VT (Volcano-Tectonic) a partir de una base de datos del volcán Cotopaxi usando la técnica de Bootstrapping.

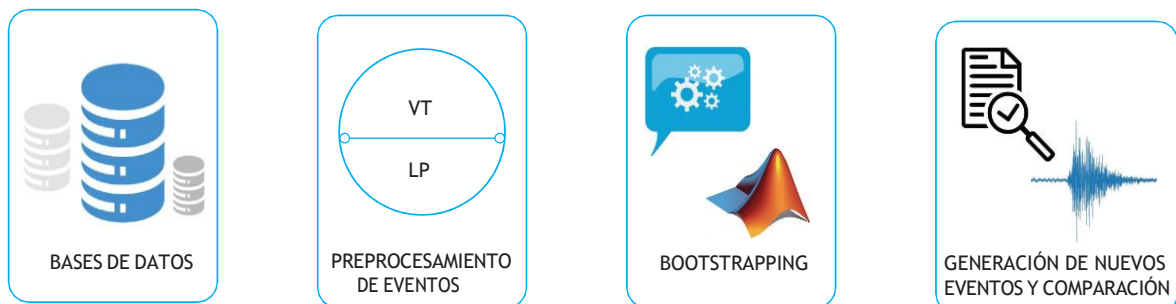
## 1.2. Objetivos específicos

- Estudiar las características generales de los eventos sismo-volcánicos.
- Estudiar la teoría relacionada con la técnica de bootstrapping.
- Usar la técnica de bootstrapping para generar señales sismo-volcánicas similares a las originales del volcán Cotopaxi.
- Evaluar el desempeño de la técnica de bootstrapping utilizando la distancia de Fréchet para proveer una métrica objetiva de la calidad de los nuevos eventos generados.

## 1.3. Alcance

Se desarrollará la implementación de una técnica conocida como Bootstrapping utilizando bases de datos previamente construidas por el Instituto Geofísico de la Escuela Politécnica Nacional disponible públicamente en el trabajo “E Seismic: Towards an Ecuadorian volcano seismic repository” [1]. Las estaciones de monitoreo registran varios tipos de eventos, pero para la generación de nuevos eventos sismo-volcánicos se considerarán los eventos VT (*Volcano-Tectonic*) y LP (*Long-Period*) que son los más importantes.

Para la construcción y evaluación de la técnica se distinguirán los procesos descritos en la Figura 1.6.



**Figura 1.1:** Generación nuevos eventos volcánicos

- **Base de datos:** Se utilizará la base de datos pública proporcionada por el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN). La base consta de varios tipos de eventos sismo-volcánicos pero solo se tomará en cuenta los que corresponden a eventos LP y VT.
- **Preprocesamiento de las señales a usarse:** Las señales que se encuentran en el dominio del tiempo serán filtradas, normalizadas a una determinada frecuencia de muestreo para finalmente remover la componente DC.
- **Diseño:** La técnica que se emplea en el presente proyecto es el Bootstrapping. Esta técnica permite remuestrear iterativamente un conjunto de datos con reemplazo lo cual permitirá la obtención de señales de iguales características a las señales del conjunto de datos previo.
- **Evaluación de la técnica de remuestreo:** Para evaluar la calidad de las señales generadas en comparación con las señales originales, se utilizará la distancia de Fréchet (Fréchet Distance, FD) que se encarga de medir que tanto se asemejan dos distribuciones.

El presente trabajo no tendrá producto final demostrable, únicamente se presentará la simulación correspondiente del uso de la técnica de Bootstrapping para la generación de nuevos eventos sismo-volcánicos del volcán Cotopaxi.

## 1.4. Marco teórico

En este capítulo se presenta a detalle los conceptos principales involucrados para tener una mejor comprensión de la finalidad del presente trabajo.

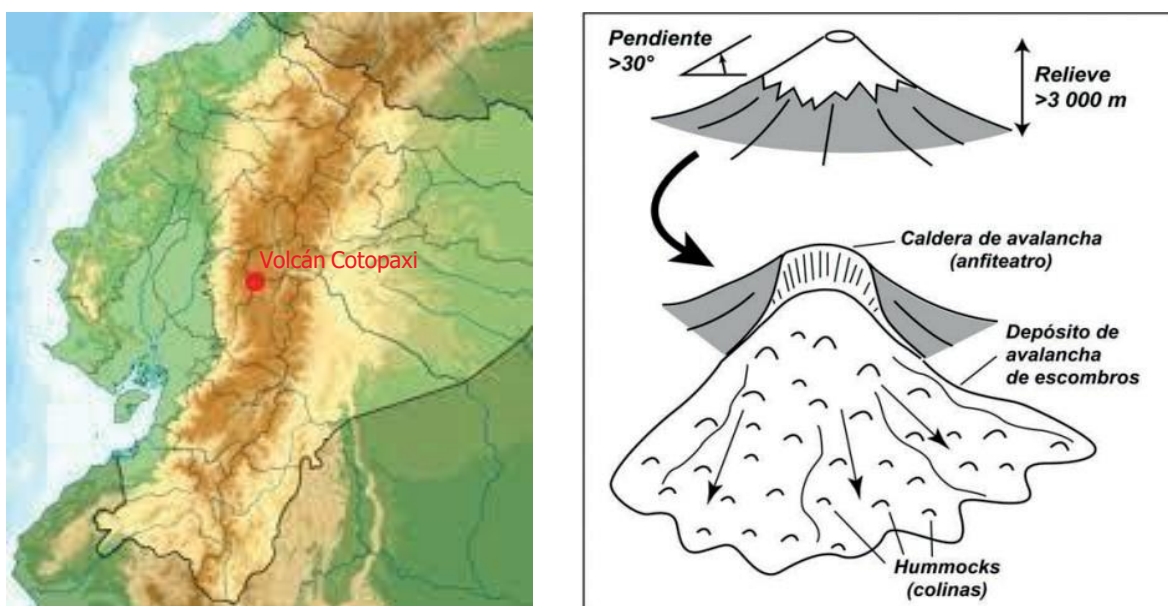
### 1.4.1. Volcán Cotopaxi

El Cotopaxi es un estrato volcán compuesto activo que se encuentra ubicado en Ecuador en la cordillera Oriental de los Andes, exactamente a 60 km al sur este de Quito y a 45 km al norte de Latacunga. El Cotopaxi es un volcán de alto riesgo que se encuentra cubierto de glaciares, presenta una vegetación de páramo que cubre sus laderas bajas, la parte superior tiene vegetación extremadamente escasa y existen extensos arenales, Figura 1.2. Este estrato volcán alcanza una altitud de 5897 metros sobre el nivel del mar, tiene una base de 16 x 19 km y un relieve de 3000 metros, además, los flancos altos del volcán presentan una inclinación de 30° y 35° [11]; esta descripción se muestra en la Figura 1.3.

Los deshielos del volcán alimentan tres importantes sistemas fluviales que son los ríos: Pita, Cutuchi y Tambo. Las planicies y páramos extensos cercanos al volcán Cotopaxi forman la zona de recreación turística y reserva animal y forestal conocida como Parque Nacional Cotopaxi [12].



**Figura 1.2:** Vista Flanco Norte del Volcán Cotopaxi



**Figura 1.3:** Descripción esquema volcánico Cotopaxi

La frecuencia eruptiva, estilo eruptivo y relieve del Cotopaxi sin lugar a duda hacen que sea

considerado uno de los volcanes más peligrosos del mundo; por lo que se utilizan gran cantidad de recursos disponibles para su monitoreo constante. Cabe mencionar que el Cotopaxi ha presentado cinco períodos eruptivos: 1532-1534, 1742-1744, 1766-1768, 1853-1854 y 1877-1880 en los cuales se han producido fenómenos volcánicos (se detallan en la subsección 1.4.1.1) extremadamente peligrosos [13].

#### **1.4.1.1. Fenómenos volcánicos**

Durante la erupción del volcán Cotopaxi se podrían producir los siguientes fenómenos volcánicos [14], a continuación se describen brevemente cada uno de ellos según su grado de peligrosidad:

- **Sismos Volcánicos:** Son detectados mediante un red sismográfica instalada en las cercanías del volcán, en pocas ocasiones este fenómeno es percibido por la población, no provocan daños importantes. La reactivación de un volcán siempre está acompañada por sismo volcánicos [15].
- **Gases Volcánicos:** Son de origen magmático, están constituidos mayormente por vapor de agua, aunque a veces también por otros gases peligrosos como: dióxido de azufre, ácido clorhídrico, dióxido de carbono o monóxido de carbono [16]. Una alta concentración de estos gases provoca daños en los ojos, boca y nariz, además, puede causar lluvias ácidas y perjudicar a los animales y vegetación.
- **Derrames de Lava:** Salen desde el cráter o fisuras del volcán y fluyen por quebradas [12]. Los derrames de lava es roca en estado líquido a una temperatura de 900°C, se presentan durante la fase eruptiva de un volcán cuando el contenido de gases es escaso.
- **Domos de Lava:** Aparecen cuando el magma es altamente viscoso y no puede fluir sino que se aglutina y enfría sobre el cráter por el cual está saliendo. Estudios indican que es un fenómeno poco probable en el volcán Cotopaxi [12].
- **Flujos Piroclásticos:** Son mezclas de gases, ceniza y fragmentos de roca que alcanzan temperaturas mayores a 500°C y descienden a grandes velocidades desde el cráter en forma de avalancha. Los flujos piroclásticos derriten grandes superficies de hielo rápidamente. En las erupciones del volcán Cotopaxi este fenómeno es muy común [17].

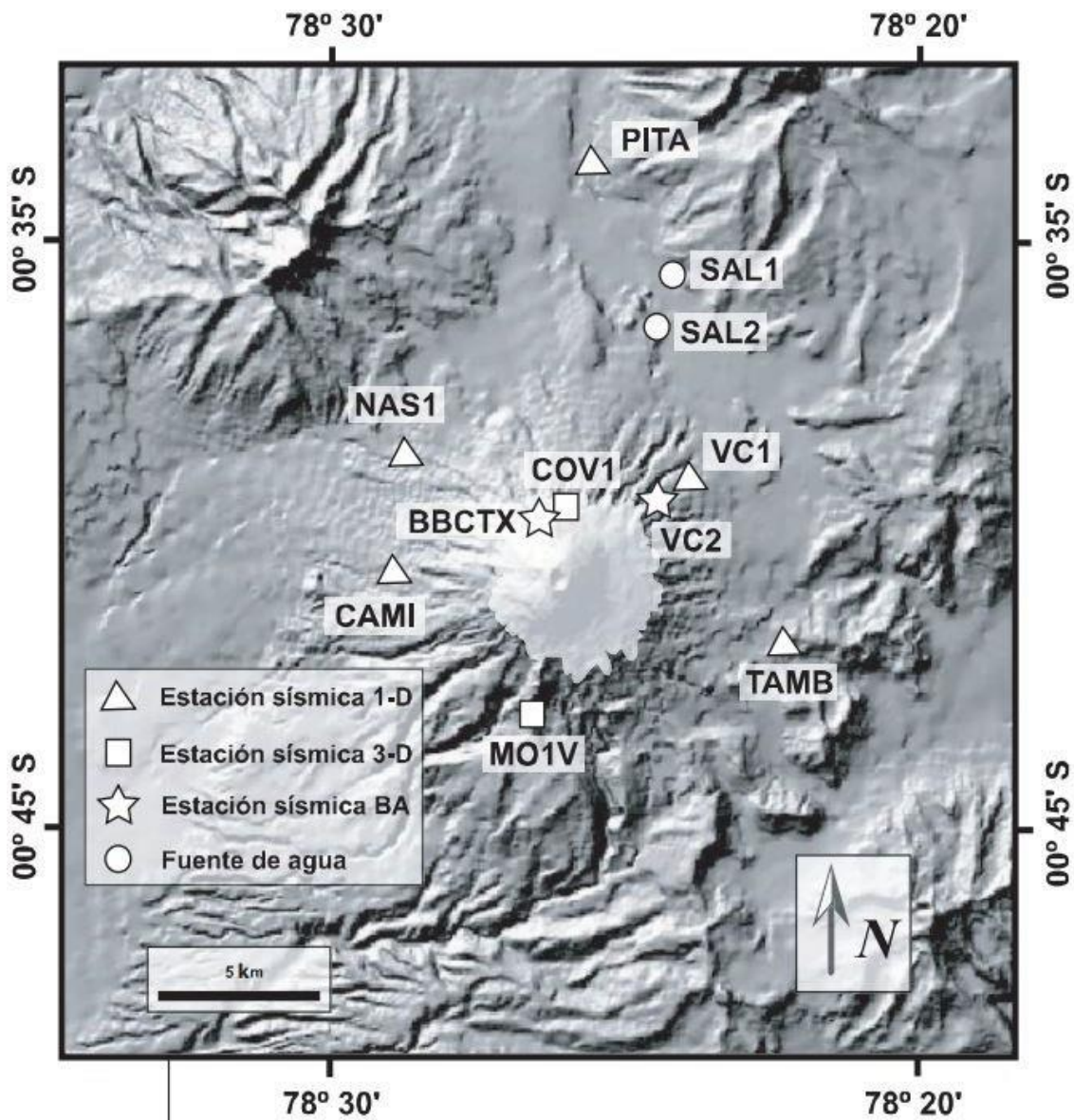
- **Ceniza:** La caída de ceniza es un fenómeno común en la erupción del Cotopaxi, esta se deposita en la superficie terrestre y su impacto depende del grosor de la ceniza acumulada [18]. La caída de ceniza afecta principalmente las vías respiratorias y causa irritación en los ojos.
- **Lahares:** Es la mezcla de rocas, piedra pómez, arena y agua que se transportan bajo la fuerza de la gravedad de la tierra a grandes velocidades, pueden alcanzar los 100 km/h. Son considerados altamente peligrosos ante la erupción de un volcán por su gran velocidad y densidad [19].
- **Avalanchas de Escombros:** Son derrumbes de rocas que se producen cuando los flancos de un volcán son inestables, las avalanchas viajan a altas velocidades y cubren extensas áreas [20]. Dada la peligrosidad y magnitud de este fenómeno todo lo que encuentra a su paso es destruido.

#### 1.4.1.2. Monitoreo Volcán Cotopaxi

El monitoreo permanente del volcán Cotopaxi es indispensable debido al alto riesgo volcánico que presenta. La finalidad de monitorear es utilizar algún medio que determine el estado continuo del volcán y percibir cambios de estado para predecir el apareamiento de una erupción volcánica futura [21]. Se conocen dos formas de monitorear el comportamiento constante de un volcán estas son: monitoreo por vigilancia visual y monitoreo instrumental.

- **Monitoreo Visual:** Este tipo de monitoreo esta basado en los sentidos humanos, no es un monitoreo asistido por instrumentos de medición y pueden ser detectados por vulcanólogos mediante la realización sistemática de observaciones permanentes.
- **Monitoreo Instrumental:** Está basado en el uso de instrumentos de medición, utilizan equipos desde los más elementales como los termómetros hasta los de más alta gama [22] como el uso de satélites y cámaras termográficas. Los instrumentos de medición determinan los siguientes parámetros: sismicidad, deformación, fumarolas, aguas termales, temperatura superficial y glaciares. La figura 1.4 muestra las estaciones que son monitoreadas en el volcán Cotopaxi. Cabe mencionar que el Instituto Geofísico de la Escuela Politécnica Nacional se encarga de monitorear la actividad de los volcanes del Ecuador 24 horas al día.





**Figura 1.4:** Estaciones sísmicas del volcán Cotopaxi.

### 1.4.1.3. Zonas vulnerables por el Volcán Cotopaxi

De acuerdo a estudios realizados [23] se puede asegurar que las zonas con mayor vulnerabilidad ante la erupción del Cotopaxi serían: el Valle Interandino entre Latacunga y Mulaló, el Valle de los Chillos en Quito. Esto debido a que durante la erupción del Cotopaxi se producen grandes lahares, se estimada que más de 300.000 personas viven en zonas amenazadas por lahares [19]. Además la caída de ceniza afectaría notablemente la región Costa y Sierra del Ecuador. La figura 1.5 muestra las zonas que se verían afectadas ante la erupción del Cotopaxi.

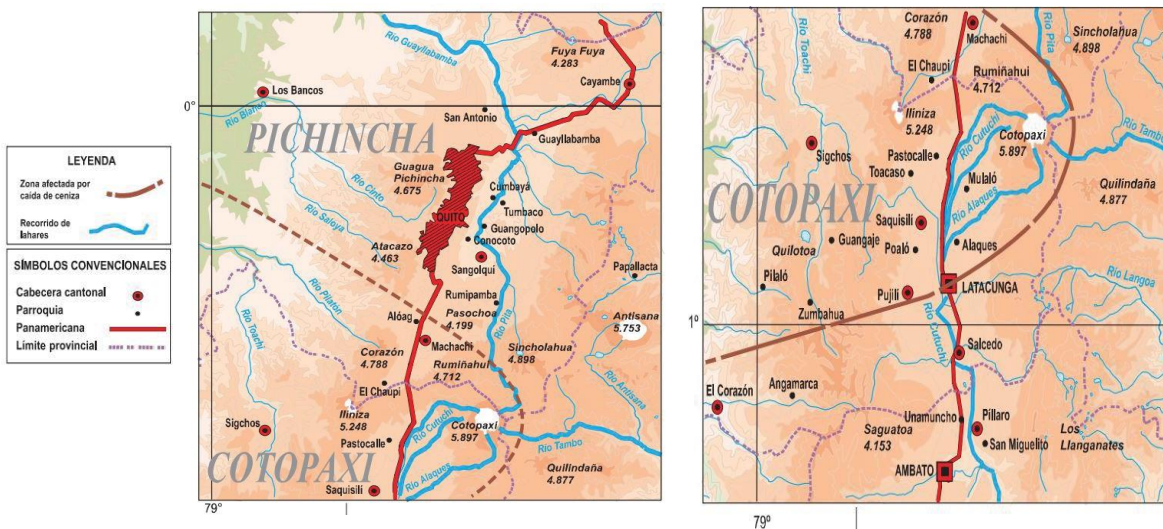


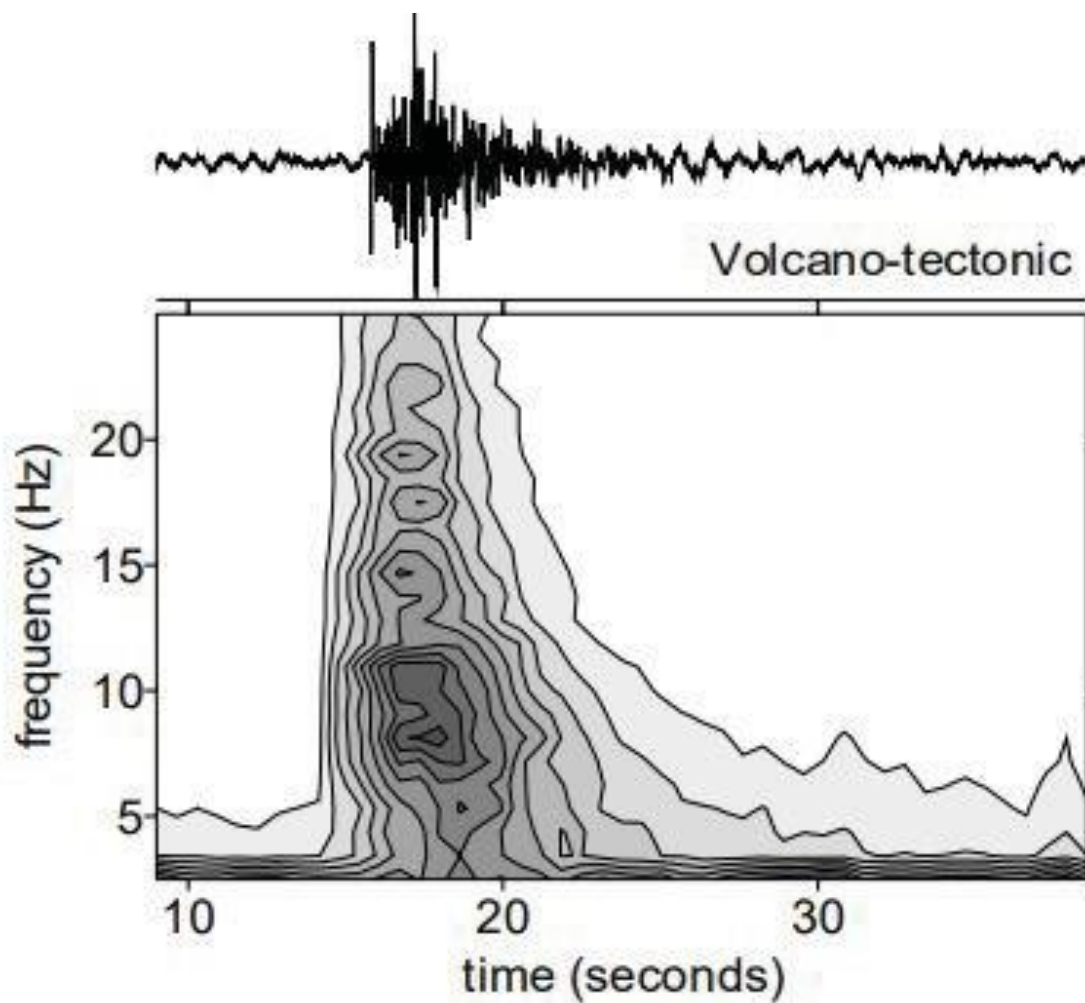
Figura 1.5: Zonas vulnerables ante la erupción del volcán Cotopaxi.

#### 1.4.2. Eventos Sismo-volcánicos

Al abordar el tema de eventos sismo volcánicos, viene a nuestra mente la idea de temblores y terremotos. La sismicidad abarca más que solo movimientos originados por placas tectónicas, sino que también de otras fuentes artificiales y naturales, como lo son los sismos de origen volcánico. Los eventos sismo-volcánicos pueden deberse a múltiples factores como desprendimientos de roca, movimientos de fluido o desgasificación de magma [24].

Los sismos volcánicos suceden a gran escala a profundidades de 1 a 9 km [25], y representan el comportamiento interno de un volcán. Los eventos mayormente relacionados a las señales sismo volcánicas son [24]:

- **VT-Eventos volcano tectónicos.-** Son caracterizados por ser una señal de duración variable, presenta un amplio contenido espectral, se suelen presentar en un amplio rango de profundidades, desde las decenas de kilómetros.
- **LP-Eventos de largo período.-** Son eventos conocidos también bajo el nombre de eventos de baja frecuencia, se caracterizan por presentar una duración de pocos segundos y un estrecho contenido espectral, ver Figura 1.7.
- **HB-Eventos híbridos.-** Se caracteriza por presentar altas frecuencias al inicio de la señal, además presenta una amplia banda espectral mayor a 10 Hz [26].
- **TR-Eventos tipo tremor.-** Es un evento caracterizado por mantener una amplitud constante durante un largo período de tiempo que puede oscilar entre varios minutos



**Figura 1.6:** Evento Volcano-Tectónico

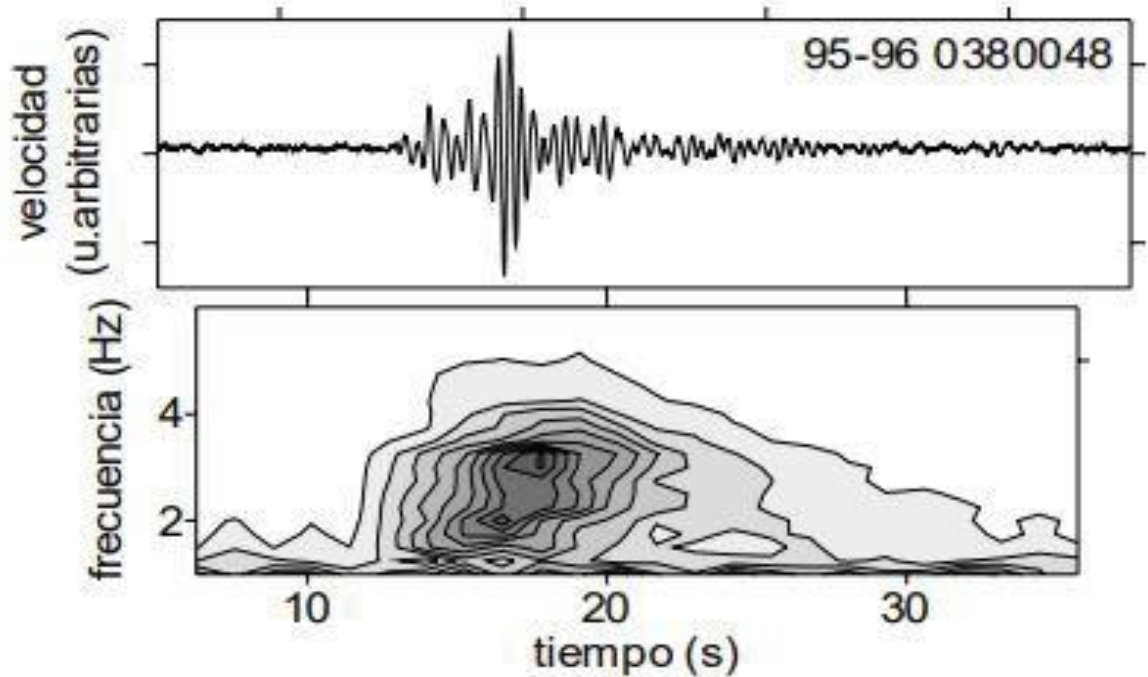
y horas [26]. Su contenido espectral se encuentra centrado en bandas de frecuencia estrecha.

### 1.4.3. Procesamiento de Señales

Área de la ciencia e ingeniería que se ha desarrollado a pasos agigantados durante los últimos años [27], permite analizar y optimizar las señales en tratamiento con el propósito de modificarlas o extraer información útil para su análisis. El procesamiento de señales se aplica en varias industrias, para ello, las señales son modeladas como funciones matemáticas. Se basa en la matemática aplicada y estadística.

#### 1.4.3.1. Extracción de Características

Proceso fundamental dentro de la etapa del procesamiento de señales que permite identificar las características más relevantes de las señales sismo-volcánicas, de forma secuencial o de escala [2].



**Figura 1.7:** Evento de largo período

#### 1.4.4. Análisis Espectral

El análisis espectral hace referencia al estudio de señales, en este caso a señales sismo-volcánicas en función de su comportamiento en el dominio de la frecuencia. Para lograr cambiar el dominio de trabajo de una señal (tiempo a tiempo-frecuencia) es necesario el uso de herramientas que lo permitan, usándose para este trabajo la STFT.

##### 1.4.4.1. Transformada rápida Fourier STFT

La transformada rápida de Fourier (STFT), conocida también como transformada corta de Fourier, es un algoritmo que permite el cálculo eficiente de la transformada discreta de Fourier y su inversa [28]. Dentro de los sistemas de procesamiento de señales brinda mayor eficiencia en tiempo y recursos, puesto que trabaja con funciones discretas o señales digitales. Su ecuación viene dada por:

$$X(n, k) = \sum_{m=-\infty}^{\infty} x[n]w[n-m]e^{-j\frac{2\pi}{N}kn} \quad (1.1)$$

##### 1.4.4.2. Espectrograma

El espectrograma, corresponde a una representación tiempo-frecuencia usado para realizar un análisis frecuencial exhaustivo en señales en este caso señales sismo-volcánicas [29].

El análisis tiempo - frecuencia utilizado en este trabajo corresponde al que utiliza la transformada discreta de Fourier (STFT, Short- Time Fourier Transform) con una ventana deslizante de Hamming [30]. La ecuación 1.1 define a la STFT en donde:  $x[n]$  es la señal y  $w[n-m]$  es la ventana definida [31]. El espectrograma se origina a partir de la magnitud cuadrada de las STFT y viene dado por:

$$S \{x(n)\} \equiv [X(n, k)]^2 \quad (1.2)$$

El uso de la ventana de Hamming en la STFT tiene la finalidad de limitar el tamaño de la secuencia a ser transformada, de tal manera que las características espectrales sean constantes en el intervalo de duración de la ventana. La separación de cada ventana viene dada por el overlap, valor que corresponde al número de puntos repetidos de la ventana anterior. Según lo descrito en la (Subsubsección 1.4.2) las señales sismo-volcánicas tienen frecuencia variable a lo largo del tiempo y sus variaciones son rápidas por lo que la ventana debe ser corta.

La selección acertada de longitud de la ventana de trabajo asegura la precisión de las frecuencias y su ubicación en el tiempo [32]. Una ventana pequeña brinda mayor determinación de variaciones en el tiempo, pero un decrecimiento en la resolución en frecuencia. Por el lado contrario si la ventana es amplia la resolución en frecuencia es amplia y no permite la resolución en tiempo.

#### 1.4.5. Herramienta Matlab

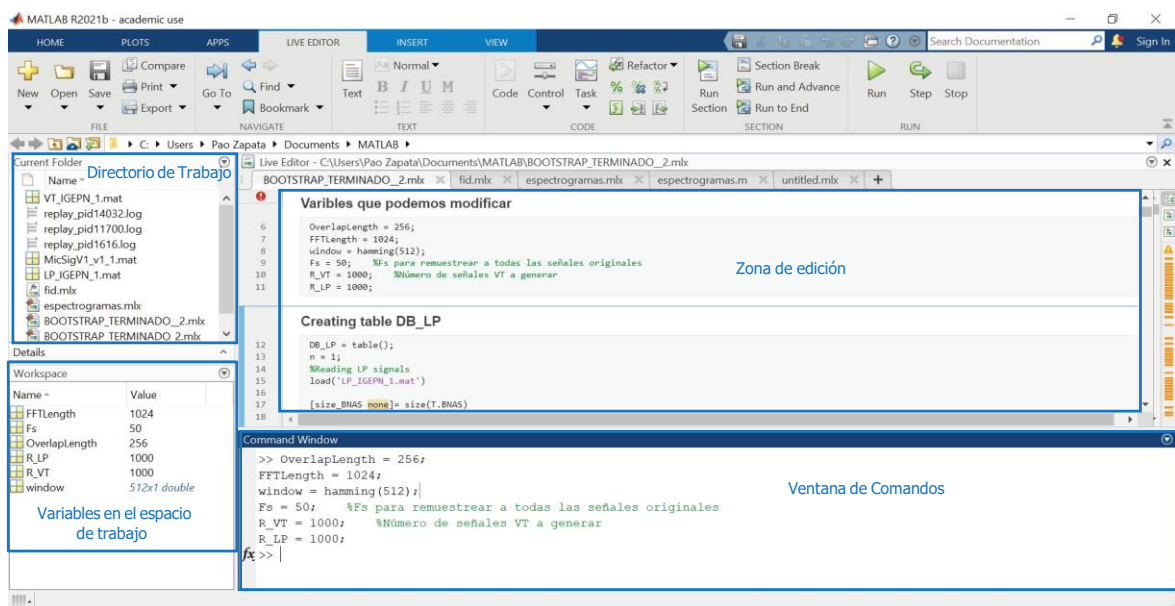
Proviene del acrónimo MATrix Laboratory , es un software de simulación que permite realizar cálculos complejos, está basado en modelo matemático de cálculos sobre matrices. Matlab comprende una serie de funciones y librerías conocidas como *toolboxes* que permiten resolver problemas, desarrollar y diseñar técnicas que faciliten tareas cotidianas [33]. El software Matlab es actualizado continuamente, por lo que frecuentemente aparecen nuevas versiones [34]. El sistema Matlab consta principalmente de:

- **Espacio de trabajo.**- Permite realizar cálculos y almacenar datos para manipularlos de acuerdo a la necesidad del usuario.
- **Entorno gráfico.**- Permite visualizar, representar los datos almacenados en dos y tres dimensiones.
- **Lenguaje de programación.**- Permite manejar los datos en forma escalar, vectorial y matricial.



- **Librerías.-** Permiten el uso de una variedad de funciones para lograr realizar operaciones útiles.
- **Toolboxes.-** Son herramientas en forma de aplicaciones ejecutables desde Matlab para elevar su potencialidad.
- **Interfaz de Aplicación.-** Permite la comunicación con otras aplicaciones para ampliar el uso de las características que ofrece Matlab.

A continuación la Figura 1.8 muestra la Interfaz global de Matlab.



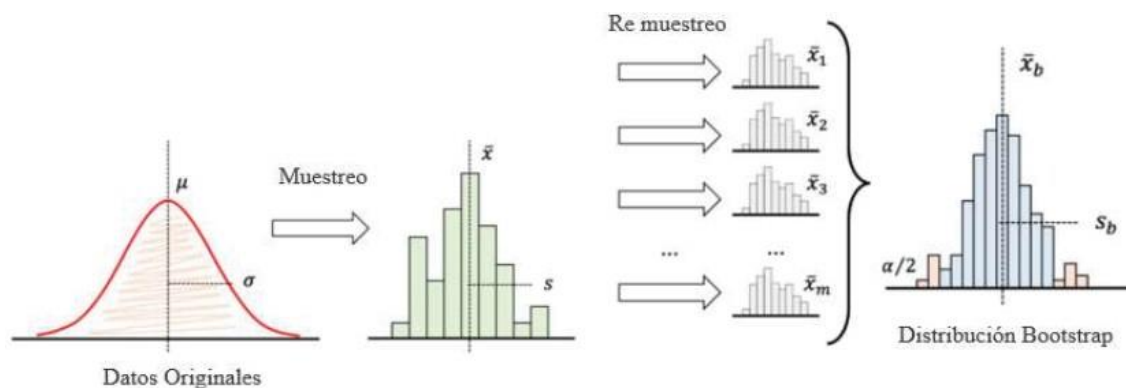
**Figura 1.8:** Interfaz global de la herramienta Matlab.

Matlab, se encuentra disponible para los sistemas operativos más usados como lo son: Windows, Apple y Unix, variando solamente la interfaz y el aspecto gráfico entre cada una de las versiones. También está disponible para aplicaciones móviles iOS y Andorid.

#### 1.4.6. Técnica de Bootstrapping

El tiempo de análisis de una gran cantidad de datos en la actualidad ha disminuido notablemente gracias a la técnica de remuestreo Bootstrapping [35]. El bootstrapping es una técnica de remuestreo de datos que resuelve problemas de estimación de intervalos de confianza. Esta técnica se encarga de estimar probabilidades empíricas a partir de datos previos, es decir, extrae muestras aleatorias significativas de los datos originales y proporciona estimaciones del error estadístico, utilizando pocas restricciones sobre las variables aleatorias analizadas [36].

El bootstrapping únicamente requiere un ordenador capaz de simular el proceso de remuestreo aleatorio de datos para analizar fácilmente un conjunto extenso de muestras de una o varias bases de datos. La figura 1.9 muestra un breve detalle de la técnica de remuestreo utilizada.



**Figura 1.9:** Esquema técnico del método Bootstrapping.

#### 1.4.6.1. Ventajas

- La generalidad en la formulación del método de Bootstrapping conforma una de las ventajas más importantes, puesto que, la metodología básica es usada en una gran variedad de situaciones, sin necesidad de acudir a modelos, expresiones o fórmulas específicas. Es un método poco restrictivo a la hora de compararlo con técnicas convencionales [37].
- El procedimiento a seguir a la hora de aplicar la técnica de bootstrapping es relativamente simple de entender y ponerlo en marcha ya que utiliza la muestra original del conjunto de datos [38].
- El bootstrapping proporciona conceptos estadísticos concretos que favorecen el aprendizaje activo.

#### 1.4.6.2. Desventajas

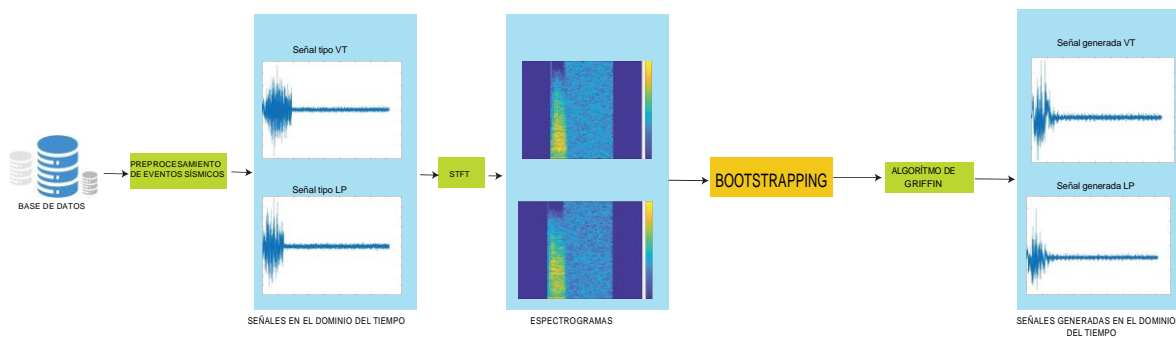
- La determinación del tamaño mínimo de la muestra y el número de réplicas requerido para lograr la aproximación de las distribuciones aun intervalo de confianza representan un verdadero desafío [39].
- Debido a que las varianzas originales son demasiado pequeñas, las varianzas estimadas tienden a ser cero, motivo por el cual la técnica es inconsistente y poco fiable según [40]

## 2. METODOLOGÍA

En el presente capítulo se describe el procedimiento para la generación de nuevas señales sismo volcánicas a partir de la base de datos (Subsección 2.1). La generación de nuevas señales sismo volcánicas se logra luego de atravesar varias etapas. Una de las más importantes es la etapa de preprocesamiento (Subsección 2.2). Aquí es en donde se preparan las señales originales existentes en el dominio del tiempo para obtener señales de la misma dimensión tanto en tiempo como en amplitud, descartando así señales que contengan demasiado ruido.

Luego, se transforman las señales sismo volcánicas al dominio de la frecuencia para lograr obtener una representación tiempo-frecuencia de cada señal. El resultado de este cálculo es una imagen bidimensional. A esta imagen se normaliza para regresar al dominio del tiempo mediante el algoritmo de griffin-lim. Este algoritmo permite la reconstrucción de las señales a través del uso de la representación tiempo-frecuencia obtenida. Finalmente, se evalúa la calidad de las señales generadas en comparación a las señales originales al usar la técnica de bootstrapping mediante la distancia de Fréchet (FD).

A continuación, se muestra en la Figura 2.1 la metodología general utilizada en el presente proyecto y se describe cada subsección detalladamente.



**Figura 2.1:** Proceso de generación de señales sismo volcánicas

### 2.1. BASE DE DATOS

Para utilizar la técnica de bootstrapping se considerarán dos bases de datos públicas proporcionadas por el Instituto Geofísico de la Escuela Politécnica Nacional (IGEPN). Para efectos de estudio las denominaremos D1 y D2 a las mencionadas bases de datos.

D1 consta de 20309 señales volcánicas sin procesar pertenecientes a cuatro estaciones

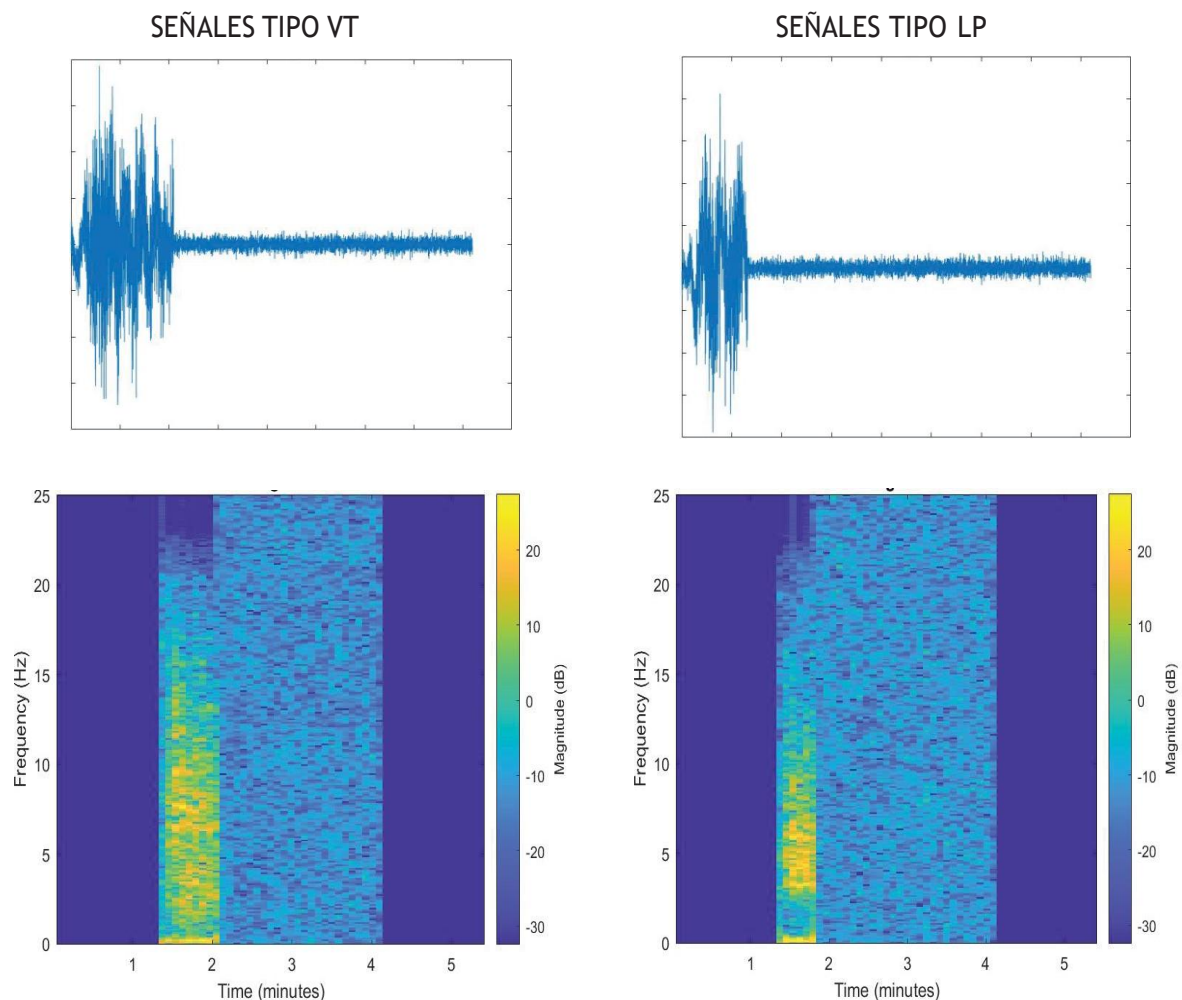


**Tabla 2.1:** Distribución de los eventos sismo- volcánicos correspondientes a las dos bases de datos en cuestión para aplicar la técnica de Bootstrapping.

Bases de Datos	Eventos LP (Long-Period)	Eventos VT (Volcano-Tectonic)
D1	11553	8756
D2	1086	101

(BNAS, BREF, BTAM y BVC2), de las cuales se tomará información del canal BHZ. [1]. La segunda base de datos D2, consta de 1187 señales sin procesar. Es importante mencionar que se considerarán los tipos de eventos LP (*Long-Period*) y VT (*Volcano-Tectonic*) los mismos que se muestran en la Figura 2.2 junto con su respectivo espectrograma.

Para el proceso de generación se seleccionó la representación tiempo- frecuencia de los eventos para las dos bases datos pertenecientes a las cuatro estaciones de acuerdo a la tabla 2.1. Se usan las cuatro estaciones con la finalidad de proporcionar mayor cantidad de información a la técnica de bootstrapping.



**Figura 2.2:** Eventos LP y VT registrados en el volcán Cotopaxi

## 2.2. PREPROCESAMIENTO

El proceso de preparación y adecuación del conjunto de datos original cumple un papel fundamental dentro del proyecto, puesto que facilita la aplicación de la técnica de bootstrapping. Las señales volcánicas presentan una estructura compleja y variable en el dominio del tiempo por lo que pueden representar un verdadero desafío al aplicar la técnica de remuestreo. Por dicha razón, en este trabajo, se genera una representación tiempo-frecuencia de cada una de las señales originales  $\mathbf{x}$  considerando los eventos tipo VT y LP del conjunto de datos original y se procede a remover el desplazamiento de las mismas.

En primer lugar, puesto que las estaciones de las bases de datos originales están a diferentes frecuencias de muestreo (50Hz y 100Hz), se normalizó la frecuencia de muestreo a 50Hz mediante la técnica de interpolación lineal. Una vez normalizadas las señales se procede a tomar la información del canal BHZ de las cuatro estaciones mencionadas en la (Subsección 2.1).

Se elimina el nivel DC de las señales restando la media, se remueve el offset y se rellenan las señales con ruido de fondo hasta alcanzar una longitud fija  $N$ . La Figura 2.3 muestra la longitud fija definida tanto para eventos LP y VTs. La amplitud del vector  $N$  se normalizó en el rango de  $-1$  a  $1$  para evitar inconvenientes en el proceso de generación de nuevas señales volcánicas.

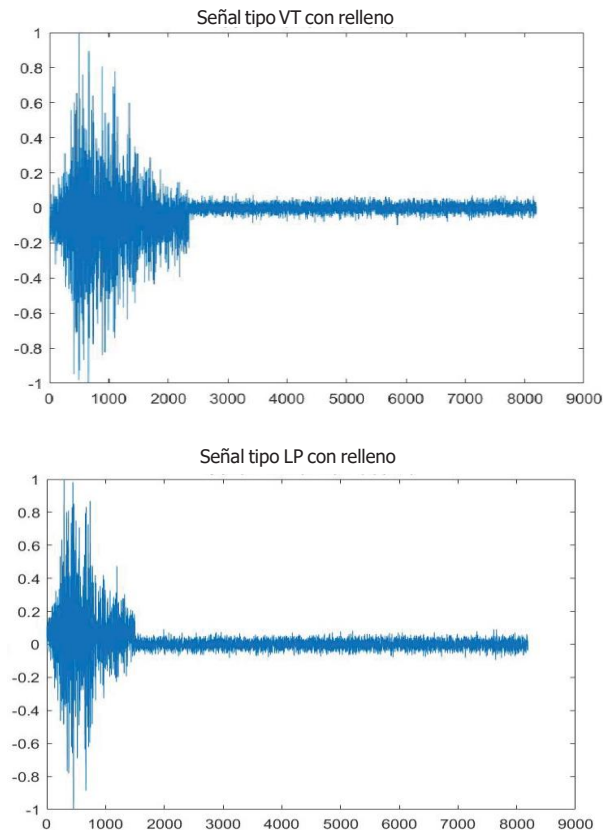
Entonces se obtiene el vector en tiempo  $\mathbf{x} \in \mathbb{R}^N$  que será la entrada para el proceso del cálculo de los espectrogramas de las señales.

Para calcular los espectrogramas de las señales y lograr una representación tiempo-frecuencia se utiliza la transformada rápida de Fourier (STFT) [41] [42], para ello se definen tres parámetros. Los parámetros definidos son una ventana de hamming de 512, overlaplength de 256 y una FFT de 1024 puntos que la denominaremos  $L$ . La ventana de hamming permite el suavizado y cambia el espectro en la frecuencia de las señales. El cálculo del espectrograma de cada una de las señales  $\mathbf{x} \in \mathbb{R}^N$  asegura que las frecuencias negativas obtenidas con información redundante sean descartadas.

La cantidad de puntos definidos determina la resolución en frecuencia, siendo:

$$A = f_s/L \quad (2.1)$$

La resolución de la señal en tiempo esta dada por  $N = 8192$  muestras. Este número fijo de muestras se transforma a tiempo resultando  $B$  minutos. Al aplicar la transformada rápida de

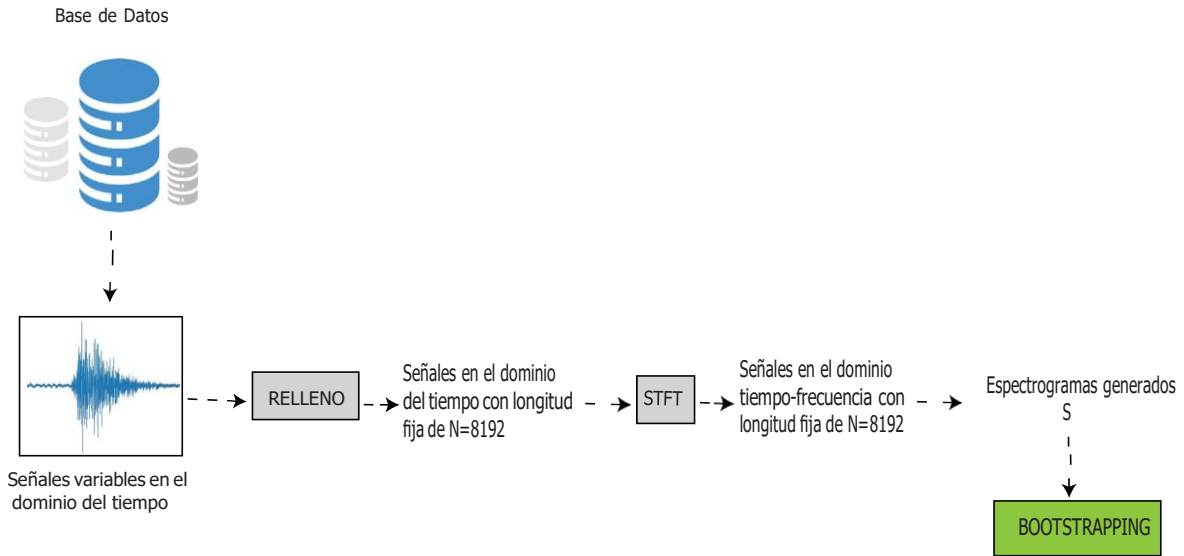


**Figura 2.3:** Longitud fija de  $N = 8192$  muestras

Fourier (STFT) [42] [43] con los parámetros definidos anteriormente los espectrogramas  $\mathbf{S}$  obtenidos para los eventos LP y VT presentan una dimensión de:

$$\mathbf{S} = L * B \quad (2.2)$$

En donde  $L$  hace referencia a la frecuencia y  $B$  al tiempo, obteniéndose así la representación tiempo-frecuencia esperada. Finalmente, los espectrogramas calculados se almacenan en la matriz  $\mathbf{X} \in \mathbb{R}$ . A continuación en la Figura 2.4 se presenta un resumen de la etapa de preprocesamiento de las señales volcánicas. Cabe mencionar que el proceso de preprocesamiento de eventos sísmo volcánicos se realiza en Matlab.



**Figura 2.4:** Preprocesamiento de señales sísmicas volcánicas

## 2.3. BOOTSTRAPPING

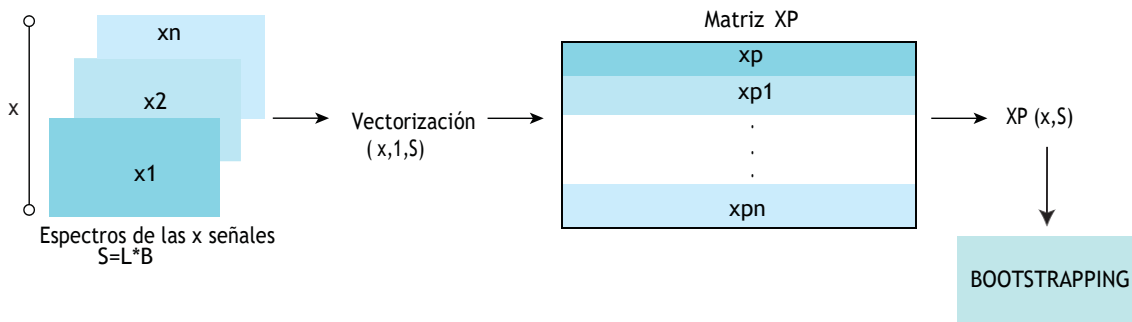
La técnica usada en este proyecto consiste en realizar el re-muestreo de cada una de las señales existentes en las bases de datos para proporcionar nuevas señales con características similares. Para esta etapa, una vez generados los espectrogramas  $\mathbf{S} \in \mathbb{R}$  (Subsección 2.2) de todas las señales nativas de las bases de datos, estos espectrogramas se vectorizan para lograr una entrada factible a la etapa del Bootstrapping. El conjunto de vectores obtenido a partir del cálculo del espectrograma tienen los siguientes parámetros.

$$(x, 1, s) \tag{2.3}$$

En donde  $x$  representa el número total de señales a ser vectorizadas. Una vez que se vectorizan las señales se forma la matriz  $\mathbf{XP}$  que tiene los parámetros:

$$\mathbf{XP} = (x, s) \tag{2.4}$$

La descripción de este proceso se muestra en la Figura 2.5. Es importante aclarar que el proceso de vectorización de las señales se realiza para los dos tipos de eventos, es decir, para los eventos tipo VT y LP respectivamente. Por lo tanto, se tendrían dos matrices que son la entrada a la etapa de la técnica de Bootstrapping. Para cada uno de los eventos en cuestión, se tendrían las siguientes matrices:



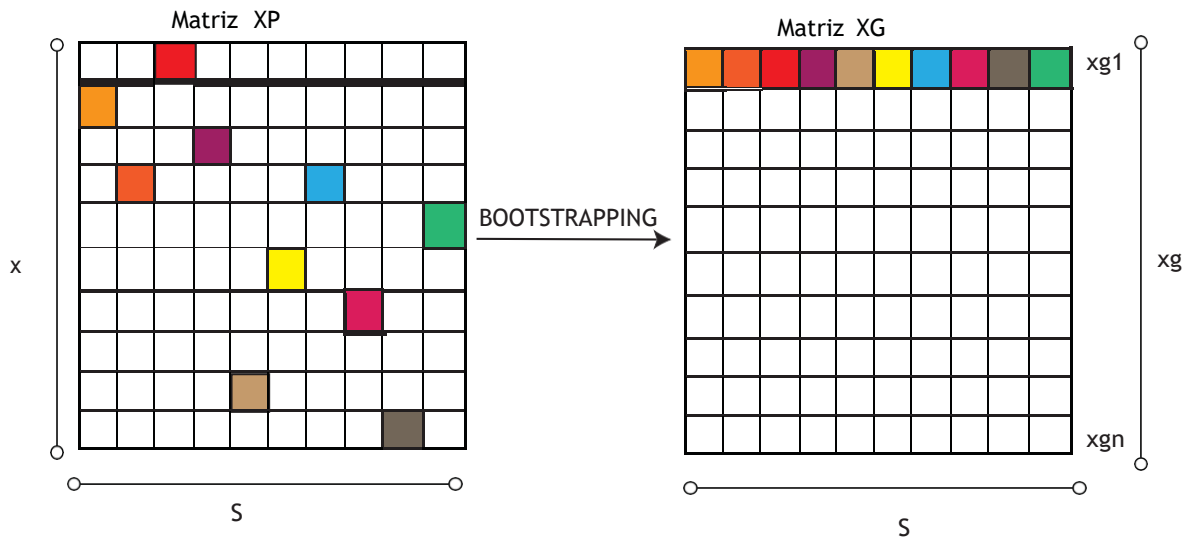
**Figura 2.5:** Vectorización de los espectros de las señales sismo- volcánicas VT y LP

$$\mathbf{XPvt} = (xvt, S) \quad (2.5)$$

$$\mathbf{XPlp} = (xlp, S) \quad (2.6)$$

Finalmente, se inicia con el re-muestreo (técnica de bootstrapping), para ello se toma aleatoriamente una celda de cada columna de la matriz  $\mathbf{XP}$  obtenida y con la ayuda de un lazo *for* se construye un nuevo vector generado tal como se evidencia en la Figura 2.6. Resultando así una nueva matriz  $\mathbf{XG}$  que está constituida por el número de señales nuevas generadas  $xg$  y  $S$ , este proceso se realiza tanto para señales VT y LP.

$$\mathbf{XG} = (xg, S) \quad (2.7)$$



**Figura 2.6:** Descripción gráfica del proceso de Bootstrapping

## 2.4. RECUPERACIÓN DE LA SEÑAL AL DOMINIO DEL TIEMPO

El proceso de recuperación de la señal al dominio del tiempo se logra al aplicar la técnica de griffin-lim. Esta técnica devuelve una señal real  $X$  en el dominio del tiempo reconstruida, estimada a partir de la magnitud  $M$  de la transformada rápida de Fourier (STFT). La función supone que  $M$  se calculó utilizando la longitud  $nfft$  de la transformada discreta de Fourier (DFT). [41]. Para ello los  $L$  puntos son reducidos a la mitad para lograr simetría y así poder aplicar la transformada.

Luego, se agrega un cero al comienzo de  $N$  para recuperar la componente DC eliminada en la etapa de preprocesamiento (Subsección 2.2). Finalmente, se aplica la STFT inversa al espectro complejo para obtener un evento sismo volcánico en el dominio del tiempo de las nuevas señales generadas, para ello es indispensable que las señales sean simétricas. Por lo tanto, los  $L$  puntos (1024) detallados en la (Subsección 2.2) se reducen a la mitad resultando así:

$$P = L/2 \quad (2.8)$$

En donde los  $P$  puntos (512) son únicos y los demás son un espejo.

## 2.5. EXPERIMENTOS

Para evaluar la similitud entre las señales originales y las nuevas señales generadas se utiliza como métrica de evaluación a la distancia de Fréchet. Para lograr aplicar la distancia de Fréchet se consideran las bases de datos D1 y D2. Entonces, primero se almacenan en dos tablas diferentes las señales originales y las señales generadas. Seguido de ello se obtiene la covarianza mediante el cálculo de la media. Luego, se calcula la suma al cuadrado de la diferencia entre las medias y la raíz del producto entre las covarianzas obtenidas anteriormente. Finalmente, se procede con la obtención de la FD (distancia de Fréchet) realizando los siguientes experimentos para eventos LP y VT respectivamente:

- FD entre señales originales y generadas con D1.
- FD entre señales generadas y ruido con D1.
- FD entre señales originales y generadas con D2.
- FD entre señales generadas y ruido con D2.

Los resultados obtenidos de los experimentos realizados se muestran detalladamente en la (Subsección 3.1).

## 3. RESULTADOS Y COMPARACIÓN

En este capítulo se presentan los resultados de los dos tipos de eventos generados LP (*Long-Period*) y VT (*Volcano-Tectonic*) para las dos bases de datos descritas en la Subsección 2.1) al ser sometidos a la técnica de bootstrapping, también se realiza una comparación de los resultados obtenidos.

### 3.1. RESULTADOS

En este apartado se muestran todos los resultados obtenidos luego de realizar los experimentos descritos en la (Subsección 2.5). El diseño e implementación de la técnica de Bootstrapping en la herramienta Matlab fue decisiva en el proceso de generación de nuevos eventos volcánicos. En el campo vulcanológico la forma de onda de un evento (VT y LP) es impredecible y variable, por lo que fue un tanto laborioso diseñar una técnica de remuestreo que plasme las mismas características del conjunto de datos original.

**Tabla 3.1:** Resultados de FD calculados sobre 1000 espectrogramas generados para D1 para las señales LP y VT, se hizo uso del conjunto de datos completo como referencia para la distribución real.

<b>Descripción Señales</b>	<b>FD (Frecuencia)</b>
LP Real vs LP Generado	45,5099
LP Real vs Ruido	16,7645
LP Generado vs Ruido	59,3686
VT Real vs VT Generado	115.7306
VT Real vs Ruido	36.4983
VT Generado vs Ruido	141.5271

**Tabla 3.2:** Resultados de FD calculados sobre 1000 espectrogramas generados para D2 para eventos tipo LP.

<b>Descripción Señales</b>	<b>FD (Frecuencia)</b>
LP Real vs LP Generado	59.1842
LP Real vs Ruido	15.8604
LP Generado vs Ruido	65.4625

### **3.1.1. Resultados obtenidos al aplicar FD**

Una vez realizadas las combinaciones evidenciadas en la Subsección 2.5), se obtuvieron los siguientes resultados al aplicar la métrica de evaluación. Primero, se muestran los resultados de FD (Frechet Distance) para 1000 espectrogramas generados, esto tanto para eventos tipo VT y LP respectivamente. Las tablas 3.1 y 3.2 muestran los resultados para las bases de datos D1 y D2.

Luego, se realizaron experimentos tomando 2000 espectros generados para los dos tipos eventos (VT y LP), de donde resultó lo mostrado en las tablas 3.3 y 3.4.

Finalmente, se obtuvieron los resultados para 10000 espectrogramas generados, esto con la finalidad de comprobar el correcto funcionamiento de la implementación de nuestra métrica de evaluación, los resultados se muestran en las tablas 3.5 y 3.6. Además, para afianzar los resultados se obtuvo la FD (Frechet Distance) de las señales originales dando como respuesta un valor que tiende a cero, lo cual indica una vez más que la métrica de evaluación arroja resultados verídicos según las señales a ser evaluadas.



**Tabla 3.3:** Resultados de FD calculados sobre 2000 espectrogramas generados para la base de datos D1.

<b>Descripción Señales</b>	<b>FD (Frecuencia)</b>
LP Real vs LP Generado	65,6473
LP Real vs Ruido	13,45
LP Generado vs Ruido	76,4601
VT Real vs VT Generado	146.5996
VT Real vs Ruido	28.6604
VT Generado vs Ruido	167.8687

**Tabla 3.4:** Resultados de FD calculados sobre 2000 espectrogramas generados para eventos tipo LP de la bases de datos D2.

<b>Descripción Señales</b>	<b>FD (Frecuencia)</b>
LP Real vs LP Generado	79.3216
LP Real vs Ruido	12.5504
LP Generado vs Ruido	82.554

### 3.2. COMPARACIÓN

Esta sección muestra la similitud de las señales generadas respecto a las señales originales, para ello en primer lugar se grafican las señales originales versus las señales generadas en el dominio del tiempo. Los gráficos mostrados a continuación corresponden a los eventos tipo VT y LP para las dos bases de datos en cuestión.

Ahora, también cabe recalcar que los resultados obtenidos al aplicar la distancia de frechet (FD) como métrica de evaluación se comprueba que las señales generadas son similares a las originales, muestra de ello es que el valor de FD obtenido al comparar las señales originales con las señales generadas es menor al valor obtenido al comparar las señales con el ruido generado.

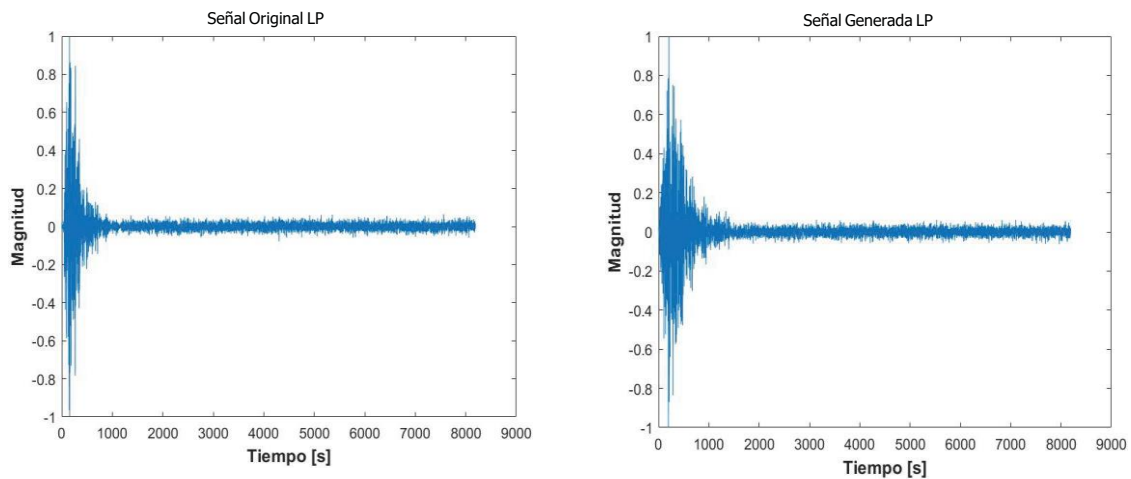
**Tabla 3.5:** Resultados de la métrica de evaluación FD calculados sobre 10000 espectrogramas generados para eventos VT y LP para D1.

<b>Descripción Señales</b>	<b>FD (Frecuencia)</b>
LP Real vs LP Generado	302.864
LP Real vs Ruido	9.7515
LP Generado vs Ruido	136.732
VT Real vs VT Generado	246.952
VT Real vs Ruido	26.2049
VT Generado vs Ruido	210.7328

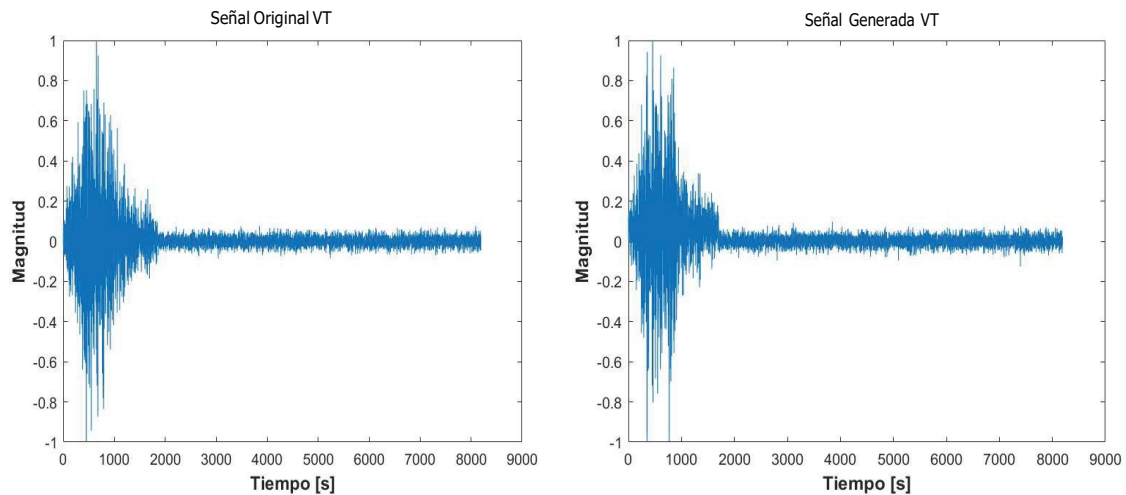
**Tabla 3.6:** Resultados de FD obtenidos sobre 10000 espectrogramas generados para las bases de datos D1 y D2, se usaron los datos de las bases originales.

Descripción Señales	FD (Frecuencia)
LP D1 vs LP D2	318.7652
VT D1 vs VT D2	257.5087

Las Figuras 3.1 y 3.2 muestran la gráfica obtenida a partir de la base de datos D1 para eventos LP (Long-Period) y VT (Volcano-Tectonic).

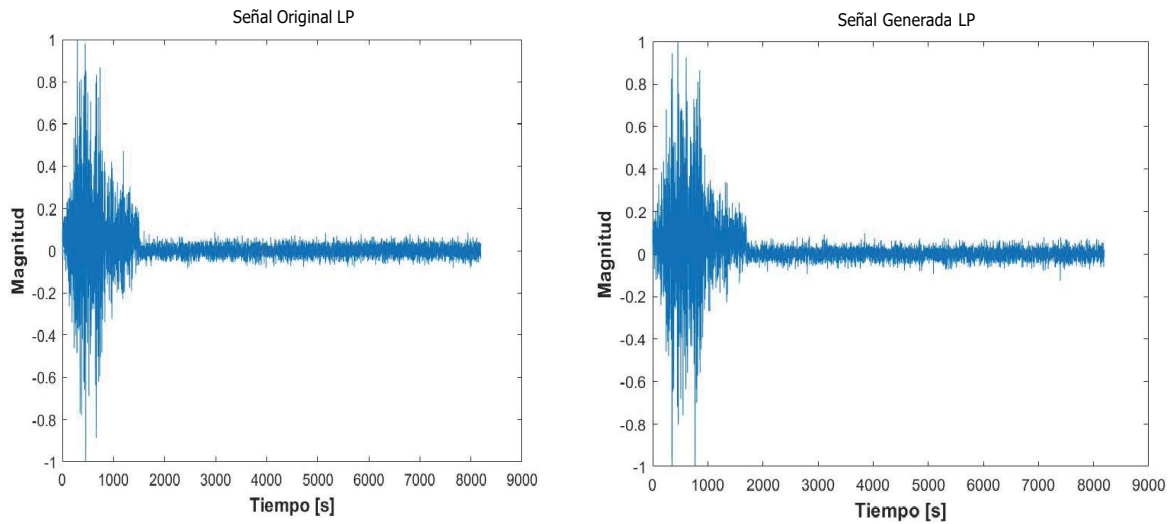


**Figura 3.1:** Eventos originales y generados tipo LP para D1

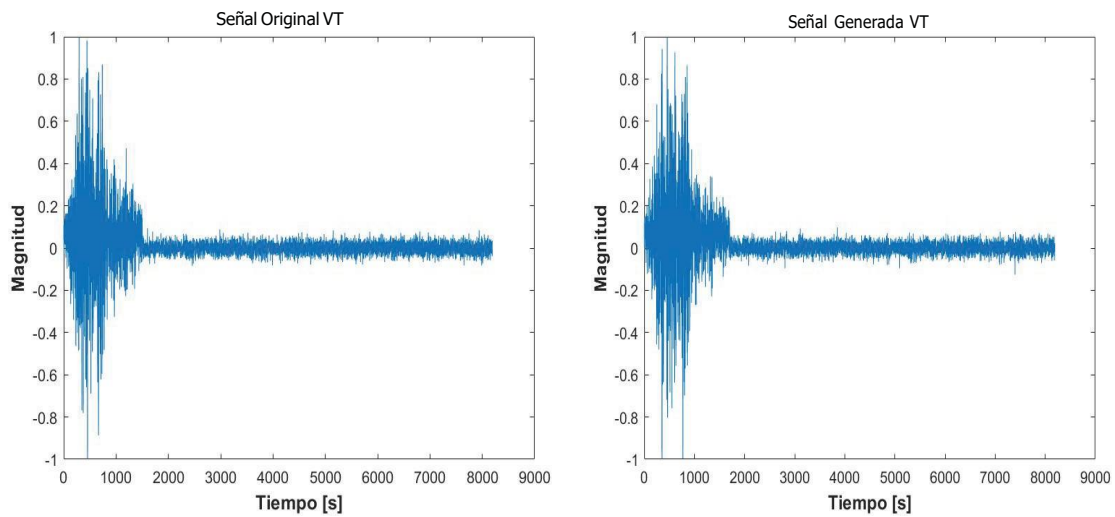


**Figura 3.2:** Eventos originales y generados tipo VT para D1

Las Figuras 3.3 y 3.4 muestran la gráfica obtenida a partir de la base de datos D2 para eventos LP (Long-Period) y VT (Volcano-Tectonic).



**Figura 3.3:** Eventos originales y generados tipo LP para D2



**Figura 3.4:** Eventos originales y generados tipo VT para D2

Evidentemente de acuerdo a las figuras presentadas las señales generadas respecto a las originales son visualmente similares, lo que asegura que nuestra técnica de remuestreo implementada tiene buenos resultados.

## 4. CONCLUSIONES Y RECOMENDACIONES

### 4.1. CONCLUSIONES

En este trabajo, se implementó la técnica de bootstrapping que fue usada para generar señales volcánicas para dos tipos de eventos VT (*Volcano-Tectonic*) y LP (*Long-Period*). La técnica se basa en la toma de muestras con reemplazo repetidamente de las bases de datos dadas, para ello se obtuvieron los espectrogramas de las señales originales para tener una representación tiempo-frecuencia, finalmente utilizando la transformada rápida de Fourier se generan nuevas señales volcánicas en el dominio del tiempo.

Dentro de las etapas más importantes de nuestro trabajo está la etapa de preprocesamiento, en la cual todos los eventos sismo-volcánicos pertenecientes a las bases de datos D1 y D2 respectivamente fueron preparados y clasificados para crear una entrada viable a la etapa del bootstrapping. Cabe indicar que se tomó información del canal BHZ de las cuatro estaciones (BNAS, BREF, BTAM y BVC2) existentes en el volcán Cotopaxi para la extracción de las características. Las señales generadas obtenidas al aplicar la técnica de bootstrapping visualmente son idénticas a las señales originales de las bases de datos existentes.

Se evaluó la técnica de bootstrapping usando la inspección visual y la métrica de evaluación FD (Frechet Distance), resultando favorable la aplicación de la técnica de bootstrapping para la generación de nuevos eventos sismo-volcánicos, puesto que ayuda significativamente a incrementar la cantidad de información nueva disponible.

## 4.2. RECOMENDACIONES

- En este estudio se obtuvieron buenos resultados al aplicar la técnica de bootstrapping usando una representación tiempo - frecuencia, por lo tanto, se recomienda usar esta representación en futuras técnicas de remuestreo. En cuanto a lo que respecta a lenguaje de programación se recomienda el uso de Matlab, puesto que es una herramienta que cuenta con numerosas funciones y documentación detallada.
- Antes de utilizar un método de generación de datos, se recomienda analizar las características de los datos existentes para contar con información de calidad en las bases de datos originales. Además, es de vital importancia contar con bases de datos equilibradas para evitar obtener resultados erróneos y sesgados.
- El análisis exhaustivo en tiempo - frecuencia permite la obtención de resultados interesantes e importantes de cualquier tipo de evento, por ello sería importante incursionar en mejorar técnicas que involucren este tema para aplicar en los diferentes campos.
- Se recomienda el uso de la distancia de frechet (FD) como métrica de evaluación para comparar y comprobar la similitud de las señales originales con las nuevas generadas, puesto que esta métrica brinda buenos resultados.

## 5. REFERENCIAS BIBLIOGRÁFICAS

- [1] N. Pérez, D. Benítez, F. Grijalva, R. Lara-Cueva, M. Ruiz, and J. Aguilar, "Eseismic: Towards an ecuadorian volcano seismic repository," *Journal of Volcanology and Geothermal Research*, vol. 396, p. 106855, 2020.
- [2] R. A. L. Cueva, V. Paillacho, and M. Villalva, "Detección de eventos del volcán cotopaxi empleando clasificación supervisada," *RECI Revista Iberoamericana de las Ciencias Computacionales e Informática*, vol. 5, no. 9, pp. 123–143, 2016.
- [3] E. Viracucha and J. De la Bastida, "Sistema informático para el análisis y procesamiento de señales sísmicas de volcanes en el ecuador." *Revista Politécnica*, vol. 33, no. 1, 2014.
- [4] J. Egred, "Instituto geofísico de la escuela politécnica nacional," 2015.
- [5] M. J. Proaño Portilla, "Análisis comparativo de grupos de características para la clasificación de señales sísmicas vt y lp del volcán cotopaxi." B.S. thesis, Quito: EPN, 2022., 2022.
- [6] R. Lara-Cueva, D. Benítez, E. Carrera, M. Ruiz, and J. Rojo-Álvarez, "Feature selection of seismic waveforms for long period event detection at cotopaxi volcano," *Journal of Volcanology and Geothermal Research*, vol. 316, pp. 34–49, 2016.
- [7] R. Lara-Cueva, V. Paillacho-Salazar, and M. Villalva-Chaluisa, "Hacia un sistema de detección automática de señales del volcán cotopaxi," *Dyna*, vol. 84, no. 200, pp. 176–184, 2017.
- [8] V. M. Zobin, *Introduction to volcanic seismology*. Elsevier, 2011, vol. 6.
- [9] P. Venegas, N. Pérez, D. Benítez, R. Lara-Cueva, and M. Ruiz, "Combining filter-based feature selection methods and gaussian mixture model for the classification of seismic events from cotopaxi volcano," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 6, pp. 1991–2003, 2019.
- [10] H. Tavera and Y. T. Antayhua Vera, "Señales sísmicas de volcanes activos: ejemplos del volcán sabancaya, arequipa (perú)," 2000.

- [11] C. A. Brusil Cruz, "Análisis comparativo entre aprendizaje supervisado y aprendizaje semi-supervisado para la clasificación de señales sísmicas vulcanológicas del volcán cotopaxi," B.S. thesis, Quito, 2020., 2020.
- [12] D. Andrade, M. Hall, P. Mothes, L. P. Troncoso Salgado, J.-P. Eissen, P. Samaniego, J. Egred, P. Ramón, D. Rivero, and H. Yepes, *Los peligros volcánicos asociados con el Cotopaxi*. Quito: Corporación Editora Nacional, CEN/Escuela Politécnica Nacional ..., 2005.
- [13] E. E. Y. T. DE INGENIERA, "Departamento de eléctrica, electrónica y telecomunicaciones."
- [14] T. Wolf, "Crónica de los fenómenos volcánicos y terremotos en el Ecuador con algunas noticias sobre otros países de la América Central y Meridional desde 1535 hasta 1797," 1904.
- [15] R. Ortiz, *Riesgo volcánico*. Servicio de Publicaciones, Excmo. Cabildo de Lanzarote, 1996.
- [16] J. B. Córdova, "La gestión preventiva y reactiva de riesgos, para afrontar una posible erupción del volcán cotopaxi, por parte del COE del gobierno municipal de Rumiñahui," *Revista de Ciencias de Seguridad y Defensa*, vol. 2, no. 4, pp. 20–20, 2017.
- [17] A. A. LAS AUTORIDADES, "Mapa preliminar de amenazas potenciales del volcán cotopaxi-zona oriental."
- [18] B. Bernard, M. Encalada, A. Proaño, D. Narváez, E. Gaunt, J.-L. Le Pennec, U. Kuepers, K. Acuña, E. Telenchana, S. Santamaría *et al.*, "Caracterización de las emisiones de ceniza asociadas a la actividad de los volcanes Tungurahua y Cotopaxi," in *Congreso Anual de Meteorología y Calidad del Aire*, 2016.
- [19] L. C. Gallegos and L. Rodrigo, "Simulación de lahares del volcán cotopaxi en sus drenajes del lado norte con el uso del programa HEC-RAS," 2013.
- [20] J. Toapaxi, C. Torres, K. Hamad, P. Vera, and J. Méndez, "Actualización de la modelación numérica del flujo de escombros producto de una erupción del volcán cotopaxi-flanco norte," *Revista Politécnica*, vol. 44, no. 1, pp. 7–14, 2019.
- [21] B. Cáceres *et al.*, "Determinación del volumen del casquete de hielo del volcán cotopaxi," 2004.

- [22] R. A. Lara, P. Vizcaino, J. D. Alvarez, and F. R. Acosta, "Caracterización de un sistema de monitoreo del volcán cotopaxi empleando sensores inalámbricos," in *CIRAD. ESPE, Sangolquí*, 2009.
- [23] D. Salazar and R. D'Ercole, "Percepción del riesgo asociado al volcán cotopaxi y vulnerabilidad en el valle de los chillos (ecuador)," *Bulletin de l'Institut français d'études andines*, no. 38 (3), pp. 849–871, 2009.
- [24] M. Malfante, M. Dalla Mura, J.-P. Métaixian, J. I. Mars, O. Macedo, and A. Inza, "Machine learning for volcano-seismic signals: Challenges and perspectives," *IEEE Signal Processing Magazine*, vol. 35, no. 2, pp. 20–30, 2018.
- [25] H. Sigurdsson, B. Houghton, S. McNutt, H. Rymer, and J. Stix, *The encyclopedia of volcanoes*. Elsevier, 2015.
- [26] T. D. E. SÍSMICOS, "Sismicidad volcánica."
- [27] C. O. J. Tintaya, "Procesamiento digital de señales sísmicas con matlab," *Revista de Investigación de Física*, vol. 10, no. 02, pp. 45–50, 2007.
- [28] M. J. Spilsbury and A. Euceda, "Transformada rápida de fourier," *Revista de la Escuela de Física*, vol. 4, no. 2, pp. 45–52, 2016.
- [29] A. P. C. RIOBAMBA and E. G. R. NARVAEZ, "Caracterización de señales volcánicas para la clasificación supervisada de eventos sísmicos."
- [30] R. Donoso Meneses, "Procesamiento digital de señales utilizando el paquete matlab," B.S. thesis, Quito: EPN, 1995., 1995.
- [31] G. Cortés Moreno *et al.*, "Reconocimiento de señales sismo-volcánicas mediante canales específicos basados en modelos ocultos de markov," 2016.
- [32] A. Pedroza, "Paso 4\_ realizar simulaciones aplicando herramientas tipo software."
- [33] Ó. R. García, L. M. J. García, L. P. Castelló, A. G. Aparicio, and A. P. Vidal, *MATLAB: conceptos básicos y descripción gráfica*. Universidad Miguel Hernández, 2018.
- [34] A. Gilat, *Matlab: Una introducción con ejemplos prácticos*. Reverté, 2006.
- [35] S. Abney, "Bootstrapping," in *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 2002, pp. 360–367.



- [36] R. Ledesma, "Introducción al bootstrap. desarrollo de un ejemplo acompañado de software de aplicación," *Tutorials in quantitative methods for psychology*, vol. 4, no. 2, pp. 51–60, 2008.
- [37] R. Ledesma, G. Macbeth, and N. CORTADA DE KOHAN, "Tamaño del efecto: revisión teórica y aplicaciones con el sistema estadístico vista," *Revista latinoamericana de psicología*, vol. 40, no. 3, pp. 425–439, 2008.
- [38] "Análisis de arranque multinivel con suposiciones violadas."
- [39] I. C. Ramírez, C. J. Barrera, and J. C. Correa, "Efecto del tamaño de muestra y el número de réplicas bootstrap," *Ingeniería y Competitividad*, vol. 15, no. 1, pp. 93–101, 2013.
- [40] "Bootstrap y pruebas de permutación: capítulo complementario 18 de la "práctica de las estadísticas comerciales"."
- [41] L. De la Fraga, "La transformada discreta de fourier y la transformada rápida de fourier," 2001.
- [42] A. L. Schmidt, "Fft: Transformada rápida de fourier," *Estudiante de Ingeniería en Sistemas de Computación. Universidad Nacional de Sur. Bahía Blanca, Argentina. Marzo 2013*, 2013.
- [43] F. M. Rodas Posso, "La transformada rápida de fourier," B.S. thesis, QUITO/EPN/1977, 1977.

## **ANEXOS**

Para la elaboración del presente trabajo se utilizó el software de simulación Matlab R2021b.

ANEXO A. Código fuente.

**ORDEN DE EMPASTADO**