



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES CARRERA EN TELECOMUNICACIONES

“Creación de algoritmos basados en técnicas de detección de la actividad de la voz aplicado a eventos sísmicos en el volcán Llaima”

Autor: Ortega Noroña Oscar Alfonso

Director del Proyecto: Ing. Román Alcides Lara Cueva, PhD.

6 de septiembre de 2023

VERSIÓN: 1.1



Agenda

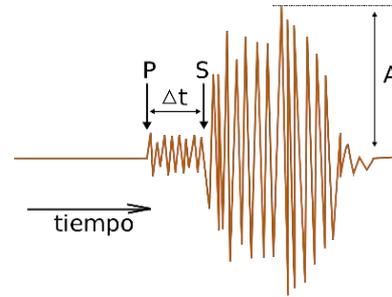
- 1.- **Introducción**
- 2.- **Materiales y Métodos**
- 3.- **Pruebas y Resultados**
- 4.- **Conclusiones y Trabajos futuros**



1. Introducción

Motivación e Importancia

Erupciones volcánicas



Señal microsísmica

Los desastres naturales que tienen gran relevancia y ocurren frecuentemente en diferentes lugares del mundo.

Desarrollar sistemas para predecir la magnitud y el alcance de la actividad volcánica.

El análisis y detección de señales para el desarrollo de sistemas de alerta temprana.



Volcán Llaima

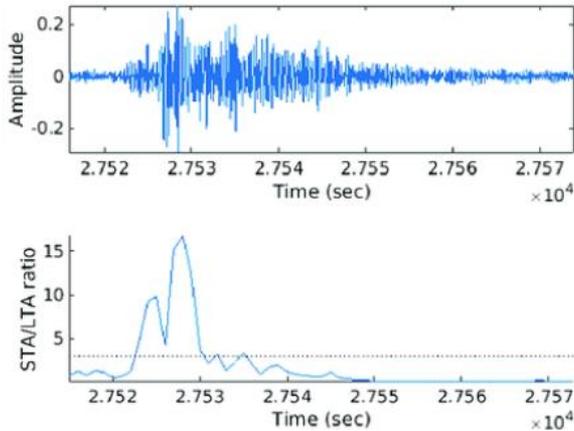


Vigilancia y Monitorización permanente



1. Introducción

Antecedentes



Algoritmo STA/LTA
99% Exactitud
0,011 BER

Algoritmo VAD
98,4% Exactitud
0,0114 BER



1. Introducción

Objetivos y Actividades

- OBJETIVO

Crear algoritmos basados en técnicas de detección de la actividad de la voz aplicado a eventos sísmicos en el volcán Llaima

- ACTIVIDADES

Actividad 1. - Identificación de las propiedades y variables del problema a tener en cuenta

Actividad 2. - Reunión de la base de datos representativa, con datos provistos por el Observatorio Volcanológico de los Andes del Sur.

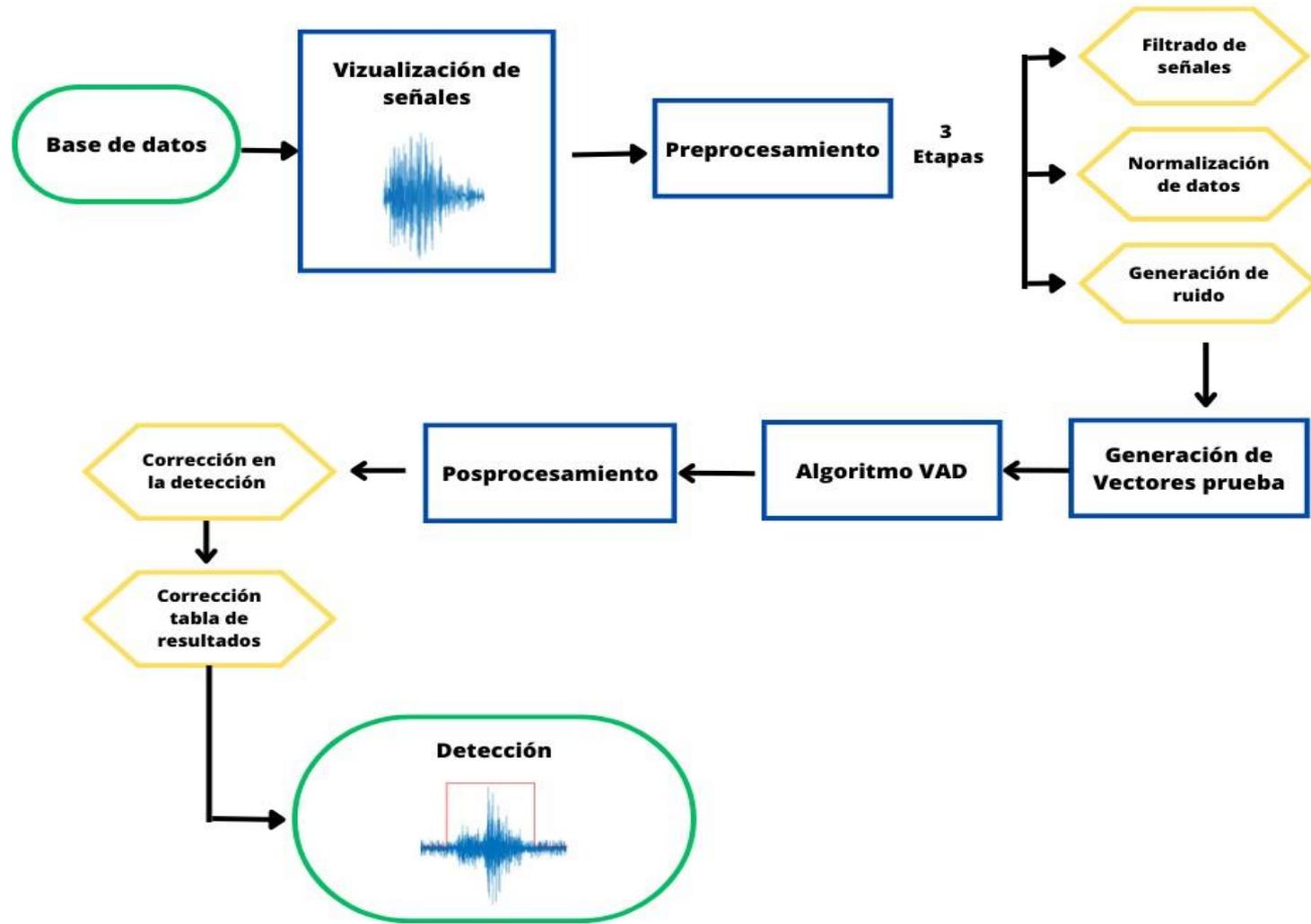
Actividad 3. - Creación de una estructura de datos (temporales, espaciales, otros.) con un soporte común.

Actividad 4. - Identificación de la información a priori a incluir en los algoritmos, y creación de algoritmos de Machine Learning tradicional adaptados al problema.

Actividad 5. - Pruebas y evaluación del desempeño.



2. Materiales y Métodos



2. Materiales y Métodos

Base de datos Volcán Llaima

En los datos proporcionados se tiene una base de datos la cual tiene cuatro tipos de eventos asociados a un microsismo

Largo Periodo (LP)



1310

Tectónicos (TC)



1488

Tremor (TR)



490

Volcano Tectónicos
(VT)

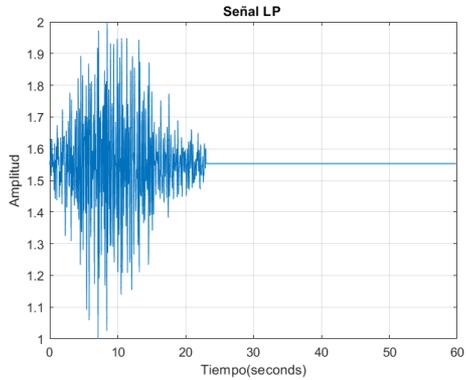


304

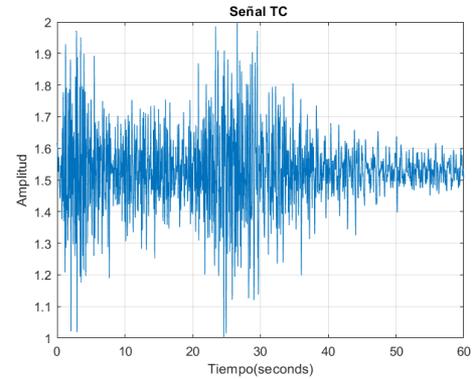
2. Materiales y Métodos

Base de datos Volcán Llaima

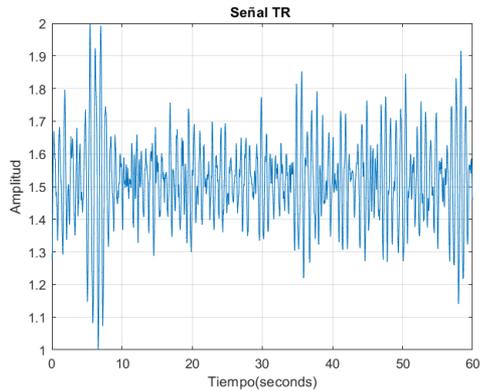
Largo Periodo
(LP)



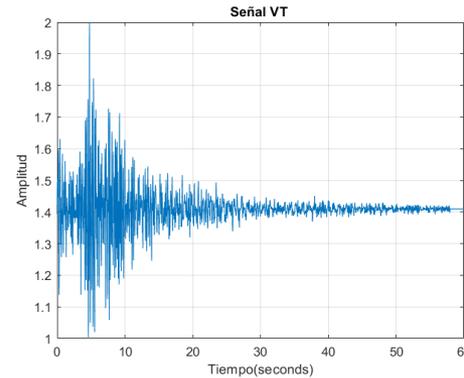
Tectónicos
(TC)



Tremor (TR)



Volcano
Tectónicos (VT)



2. Materiales y Métodos

Preprocesamiento



2. Materiales y Métodos – Preprocesamiento Filtrado

Se reúne un total de 3592 señales mezcladas entre los 4 tipos de microsismos.

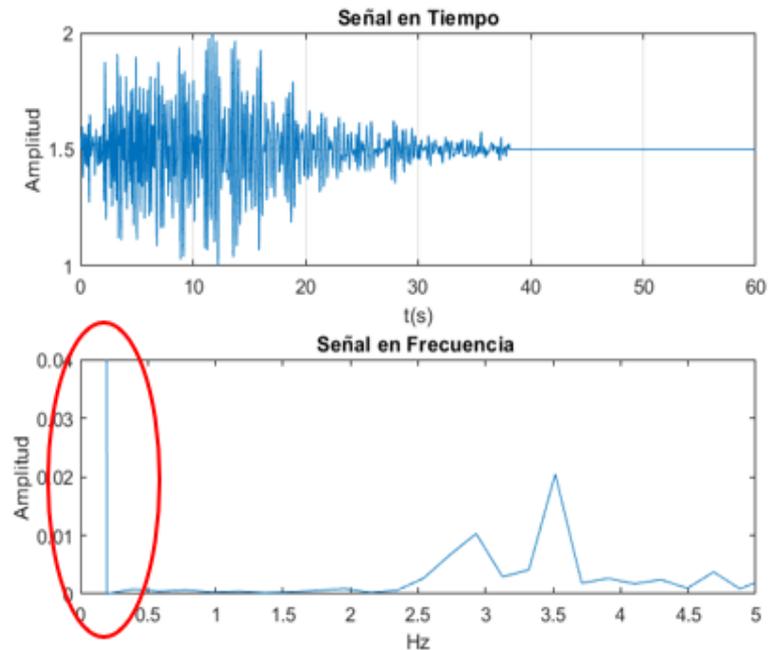
Representadas de la siguiente manera:

$$\{s_1^T, s_2^T, s_3^T, \dots, s_M^T\}^T,$$

A todas las señales se le pasa por un filtro pasa banda para eliminar una componente en 0,2 Hz.

$$H = \{h_1^T, h_2^T, h_3^T, \dots, h_M^T\}^T$$

Proceso que se almacena en una matriz H donde h_M corresponde a cada señal filtrada



Componente en
0,2 Hz

2. Materiales y Métodos – Preprocesamiento

Normalización

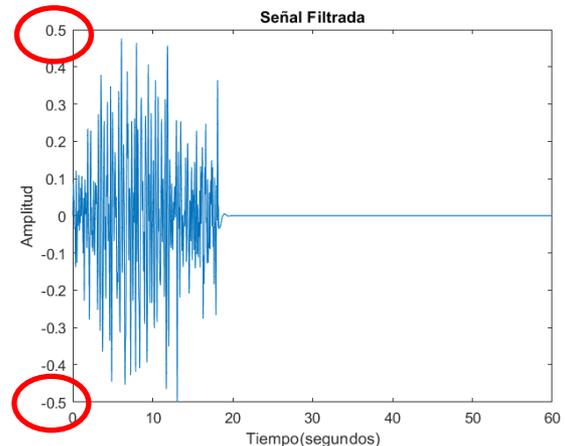
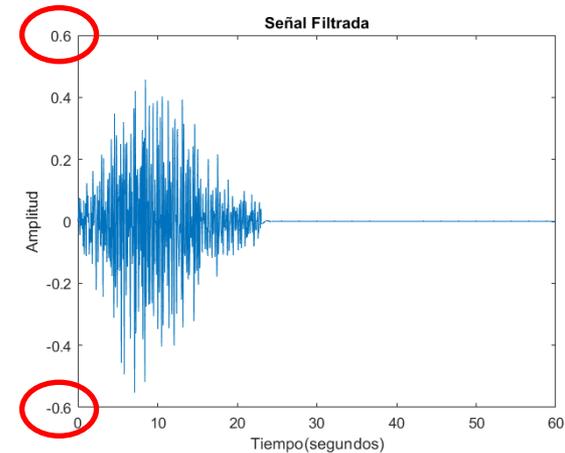
Posterior al Filtrado, se procede a realizar la normalización entre [1 -1] de cada una de las señales h_M , mediante la siguiente ecuación

$$z1_M = \frac{h_M}{\max|h_M|}$$

Almacenando el resultados de la normalización en la matriz Z1.

$$Z1 = \{z1_1^T, z1_2^T, z1_3^T, \dots, z1_M^T\},$$

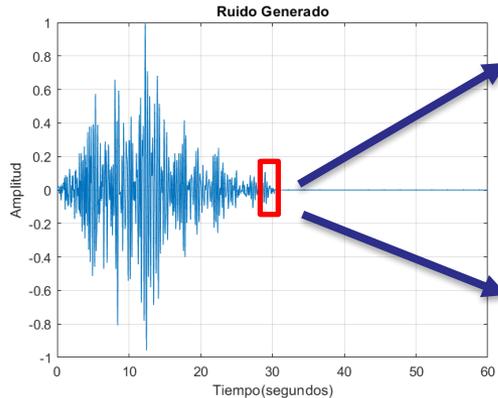
Donde $z1_M$, es la señal filtrada y normalizada.



2. Materiales y Métodos – Procesamiento

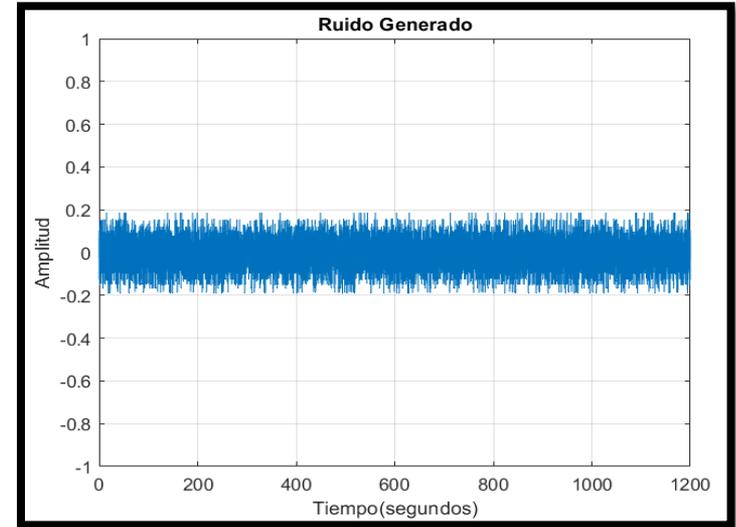
Ruido Generado

El vector de ruido se genera a partir de pequeñas partes de las señales o eventos que no son reconocidas y detectadas como actividad sísmica

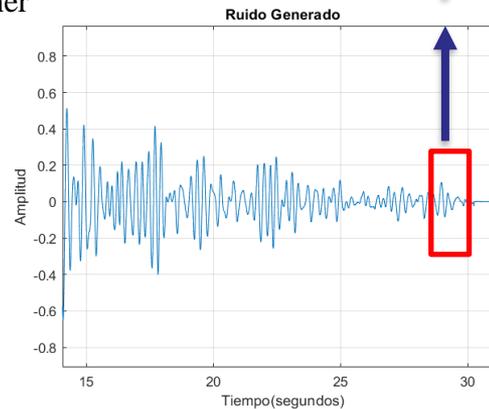


Se utiliza los puntos Fin para tomar muestras de ruido

Ventaneo de la señal para obtener ruido



Se toma en cuenta una ventana de:
 $0.2s = 20$ muestras

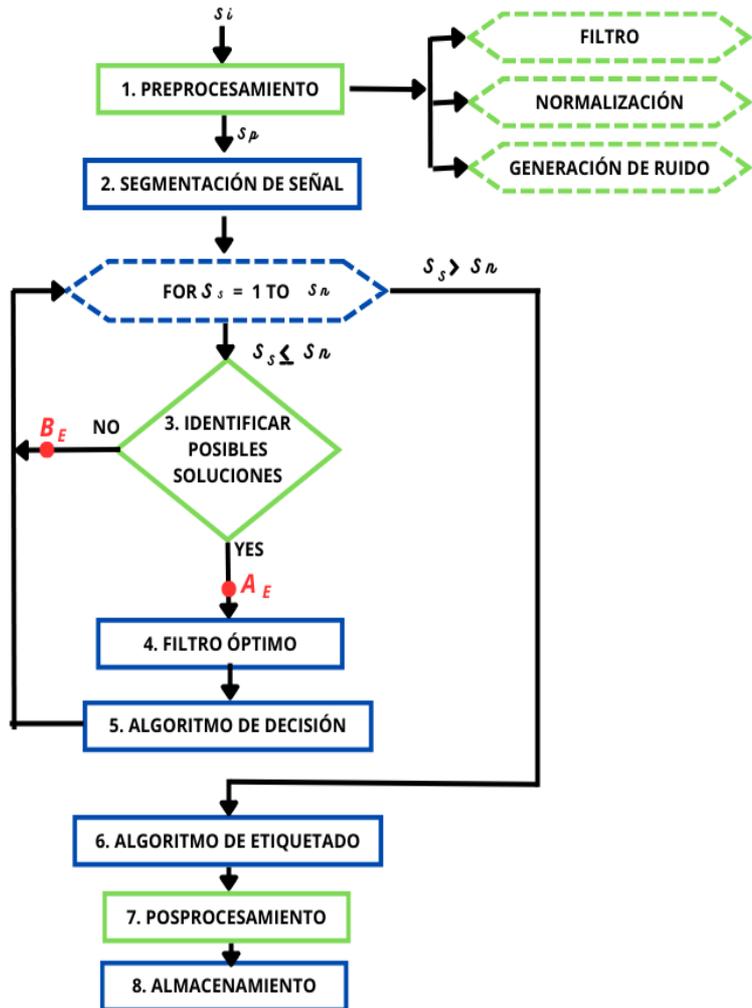


El ruido es un vector de 20 minutos o 1200 segundo

Un total de 120000 muestras

2. Materiales y Métodos

Algoritmo VAD



Segmentación de la señal

Se realiza el ventaneo, donde se separa la señal en segmentos, se realiza varias pruebas para encontrar el mejor.

Filtro Óptimo

Dos filtros que sirven para separar un microsismo del ruido de fondo que aún se puede encontrar presente en la señal con un filtro de bordes

Algoritmo de decisión

Discrimina los microsismos encontrados, se eliminan aquellos que por su corta duración no pueden ser considerados como eventos válidos

Algoritmo de etiquetado

Etiqueta a los eventos que detecto correctamente e incorrectamente, o que por el ventaneo detecto los mismo en dos tipos de ventana

2. Materiales y Métodos

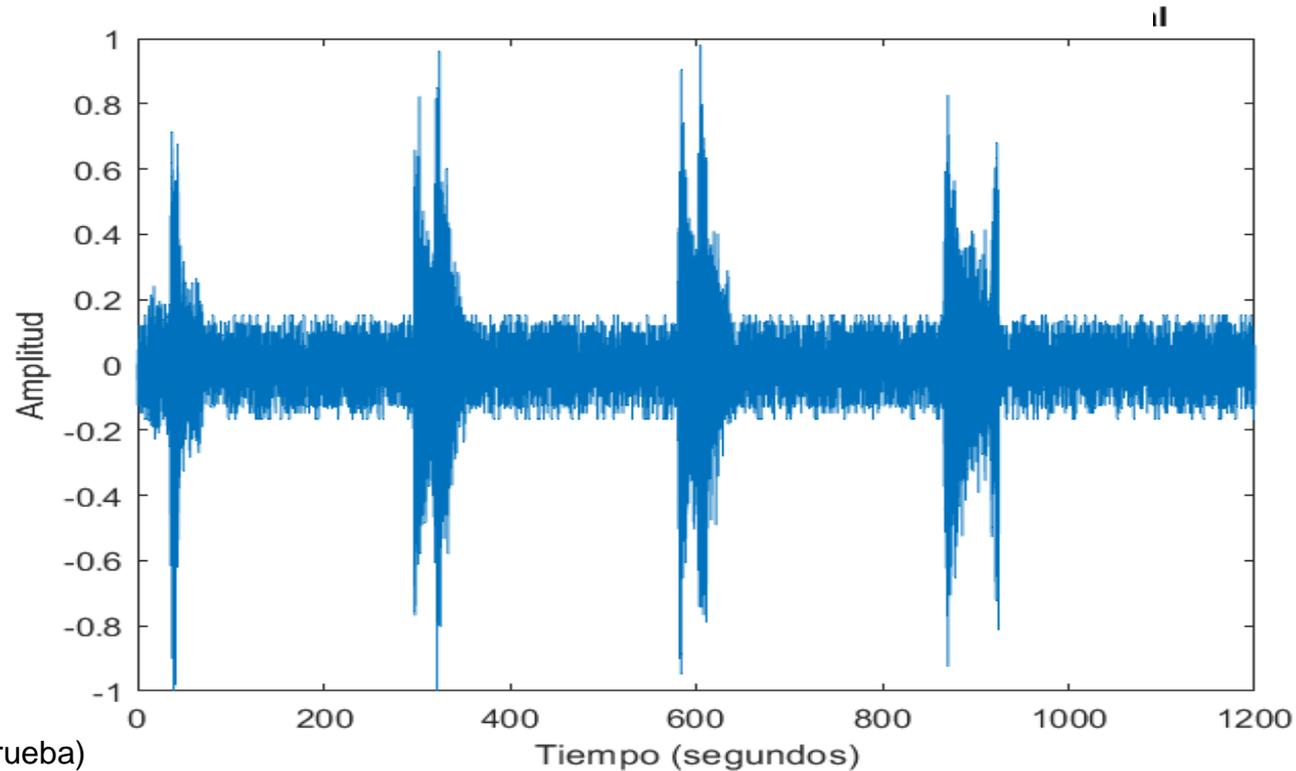
Ruido Generado

Se une el ruido junto con las señales microsísmicas.



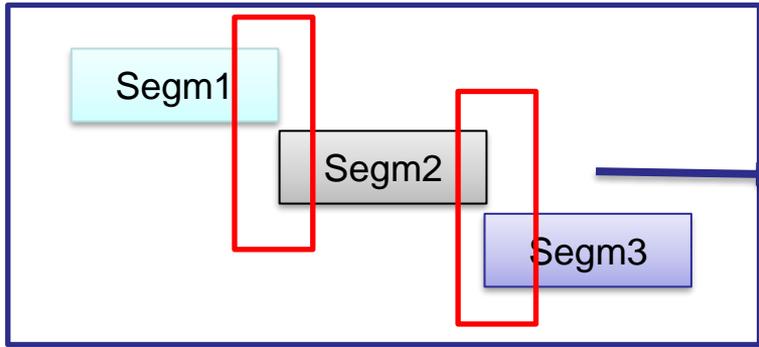
$$Y = z1_M + \eta$$

- Y .- Registro de 20 minutos (Para prueba)
- $z1_M$.- Señal normalizada
- η .- Ruido generado de 20 minutos.

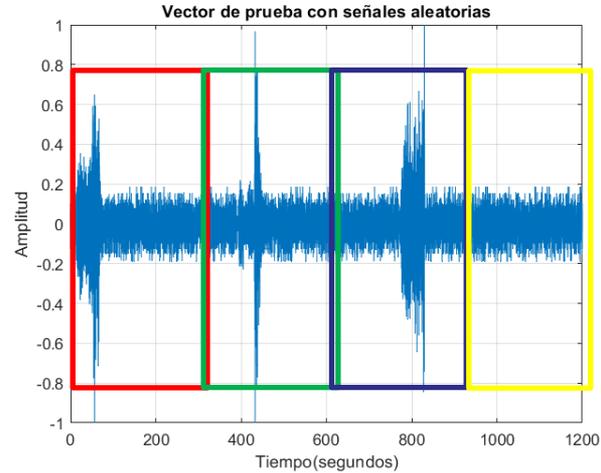


2. Materiales y Métodos

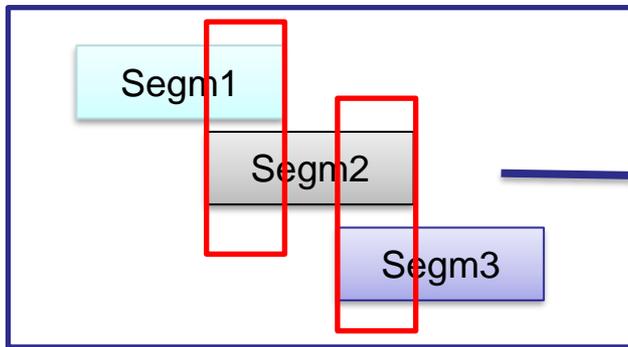
Ventaneo o Segmentación



Ventaneo sin solapamiento

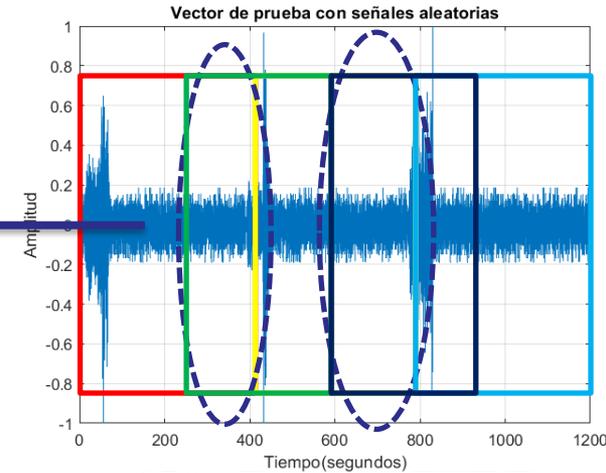


El número de experimentos y el porcentaje de efectividad determina el tamaño del solapamiento a utilizar



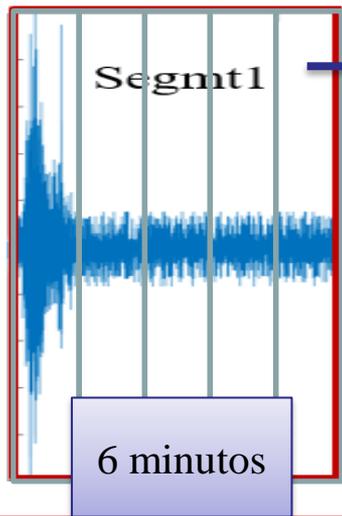
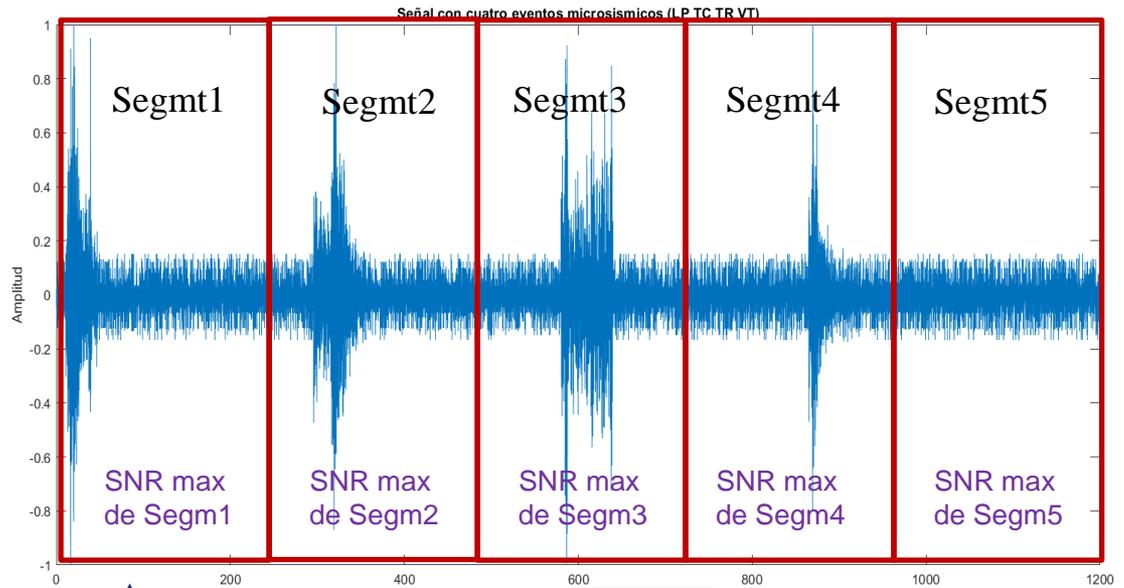
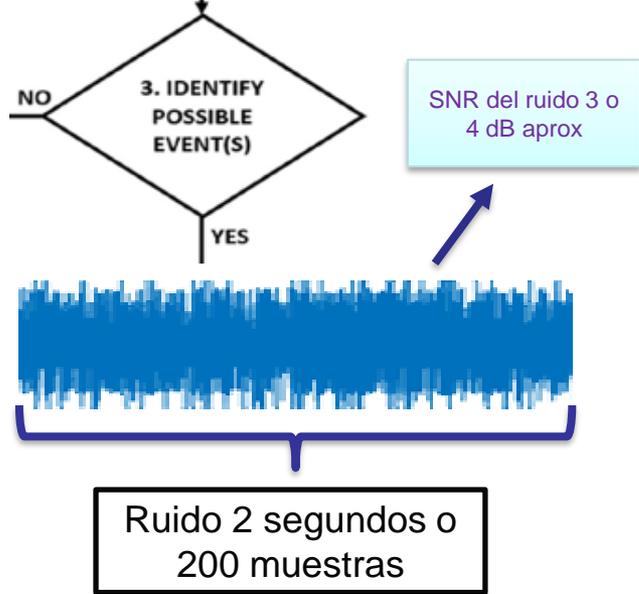
Ventaneo con solapamiento

El solapamiento varía dependiendo la efectividad de detección en distintos tamaños de ventana



2. Materiales y Métodos

Posible Evento



Segmentación interna de 2 segundos

SNR de cada segmento de 2 segundos

SNR más alta entre todos los segmentos

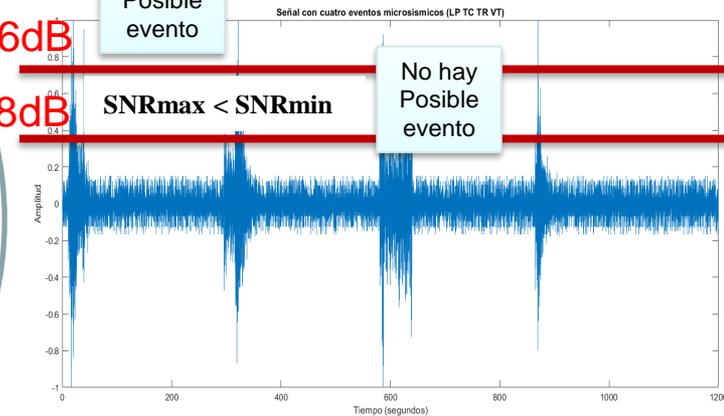
SNR de una señal aproximadamente 12-14



Existe Posible evento

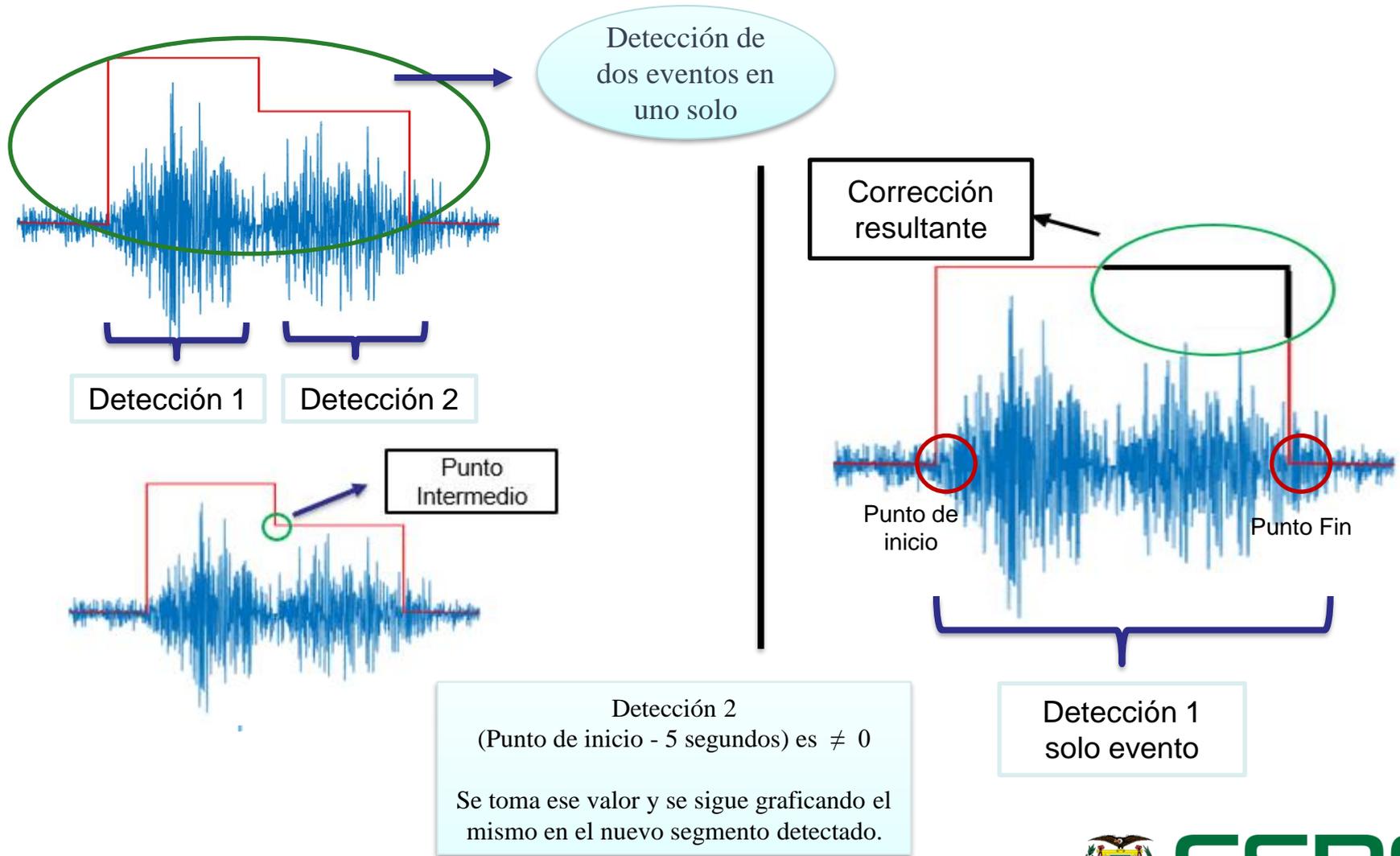
SNRmax < SNRmin

No hay Posible evento



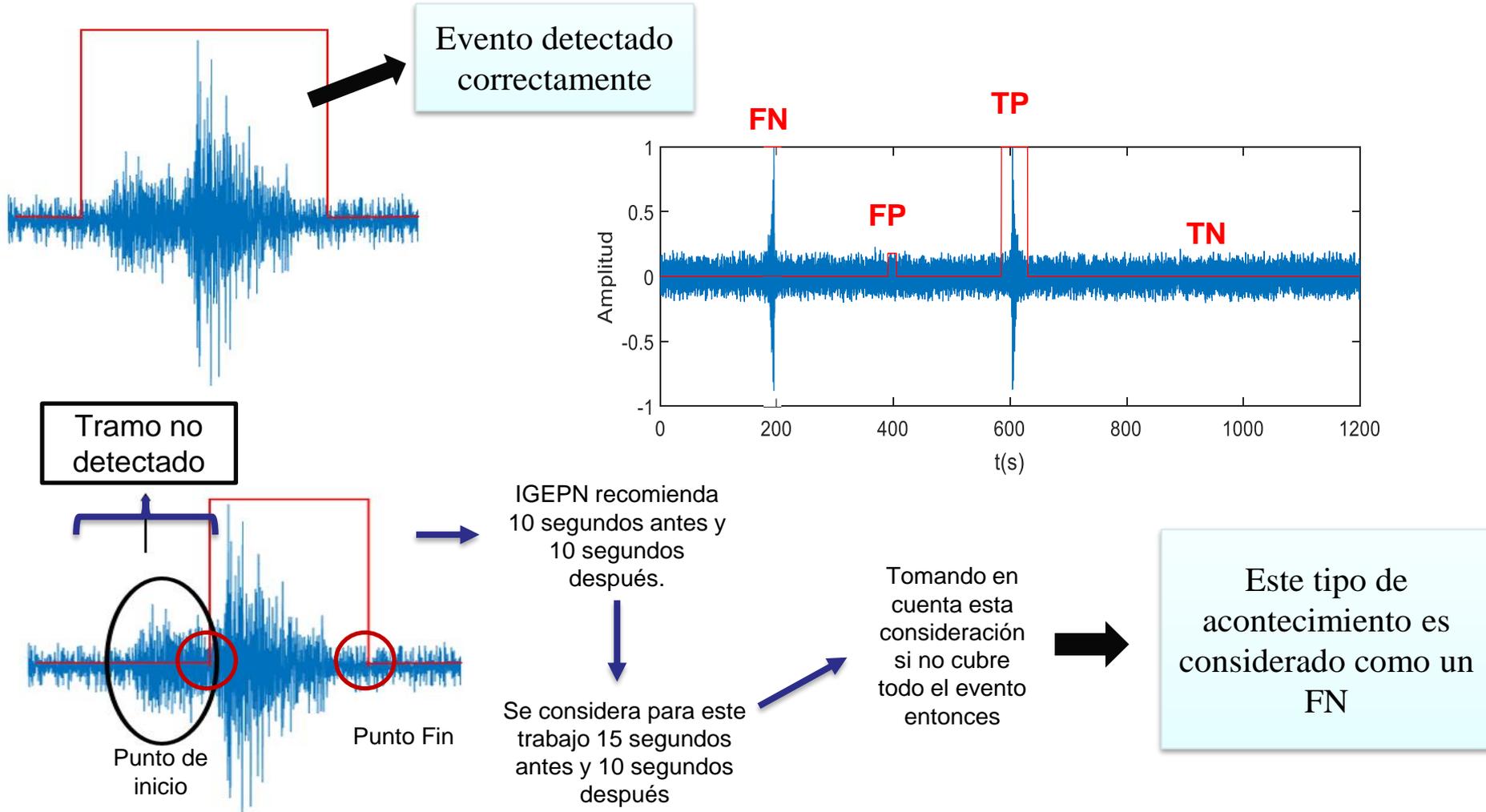
2. Materiales y Métodos

Posprocesamiento



2. Materiales y Métodos

Métricas de Rendimiento



2. Materiales y Métodos

Métricas de Rendimiento

Validación	
Tipo LP	Número de eventos
100Reg	405

Validación	
Tipo TC	Número de eventos
100Reg	475

Validación	
Tipo TR	Número de eventos
100Reg	487

Validación	
Tipo VT	Número de eventos
100Reg	469

Métrica de Desempeño	Fórmula
A (%)	$A = \frac{\text{Total microsismos Detectados}}{\text{Total de microsismos}} \times 100$
Ber	$Ber = 1 - \frac{A}{100}$

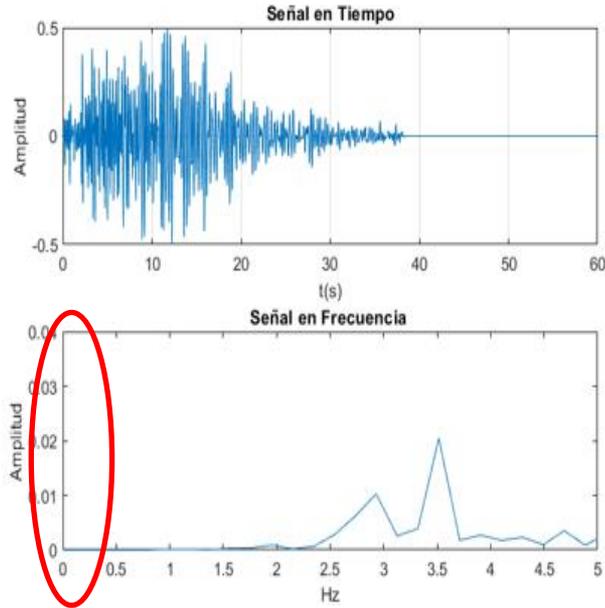
Nº Registros	Nº Señales
Registro 1	5
Registro 2	8
Registro 3	1
Registro 4	2
⋮	⋮
Registro 180	6
TOTAL	803

Se evalúa el 20% de la base de datos total, alrededor de 800 microsismos.

Métrica de Desempeño	Fórmula
Exactitud (%)	$A = \frac{TP + TN}{TP + TN + FP + FN} \times 100$
Precisión (%)	$P = \frac{TP}{TP + FP} \times 100$
Sensibilidad (%)	$R = \frac{TP}{TP + FN} \times 100$
Especificidad (%)	$S = \frac{TN}{TN + FP} \times 100$
BER	$BER = 1 - \frac{R + S}{200}$

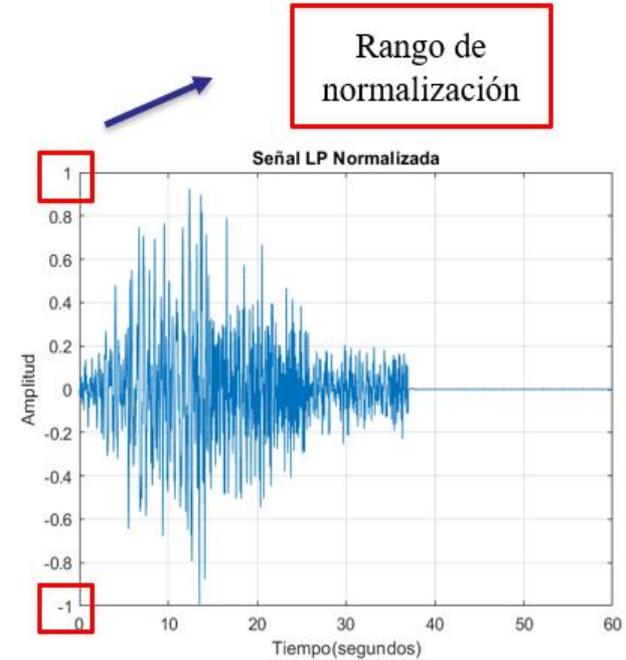
3. Pruebas y Resultados

Filtrado



Componente en
0,2 Hz
Desaparece

Normalización

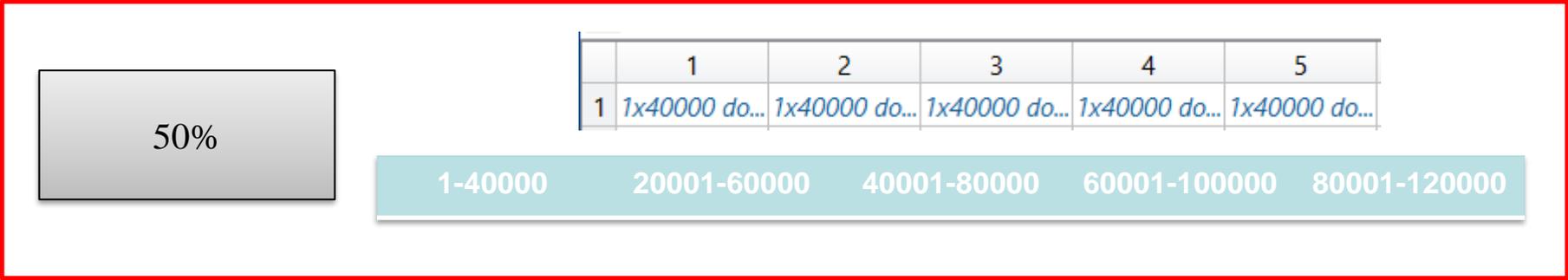


Ajustar o escalar
algo a una norma
o estándar
específico

3. Pruebas y Resultados

Segmentación

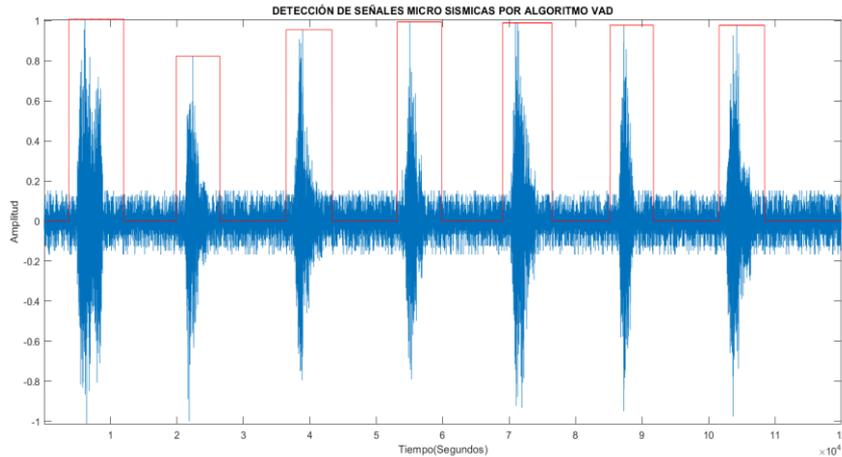
Mediante las pruebas realizadas se determinó para la detección un solapamiento del 50%



Cinco segmento de alrededor de 6 minutos del registro de 20 minutos

3. Pruebas y Resultados Validación

Eventos LP



A %

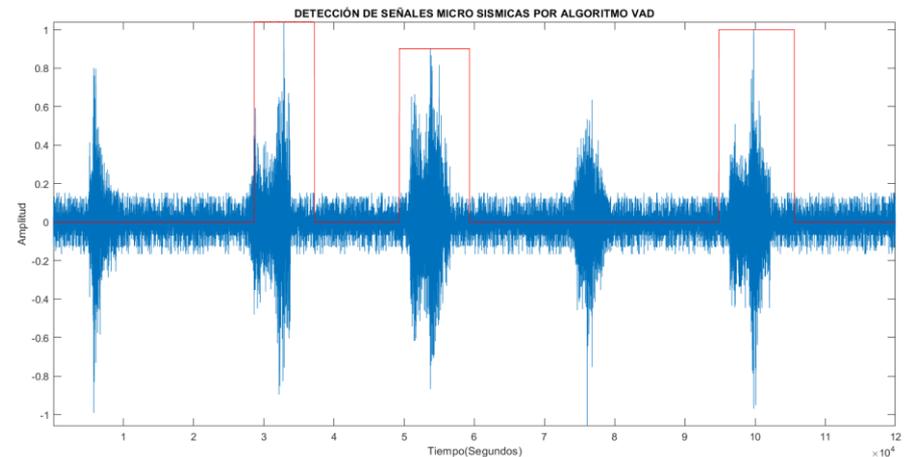
Ber

99.77

0,002

El algoritmo desarrollado no es capaz de detectar 2 en 1000 eventos puestos a prueba

Eventos TC



A %

Ber

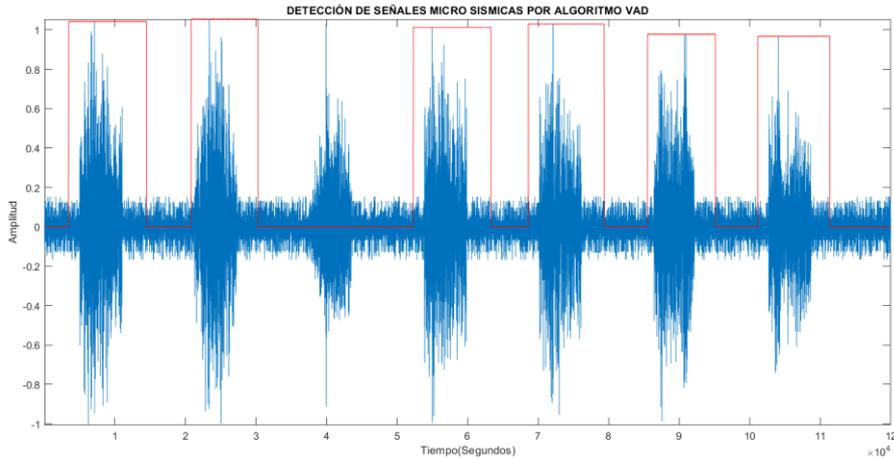
73,35

0.2

El algoritmo desarrollado no es capaz de detectar 2 de cada 10 eventos puestos a prueba

3. Pruebas y Resultados Validación

Eventos TR



A %

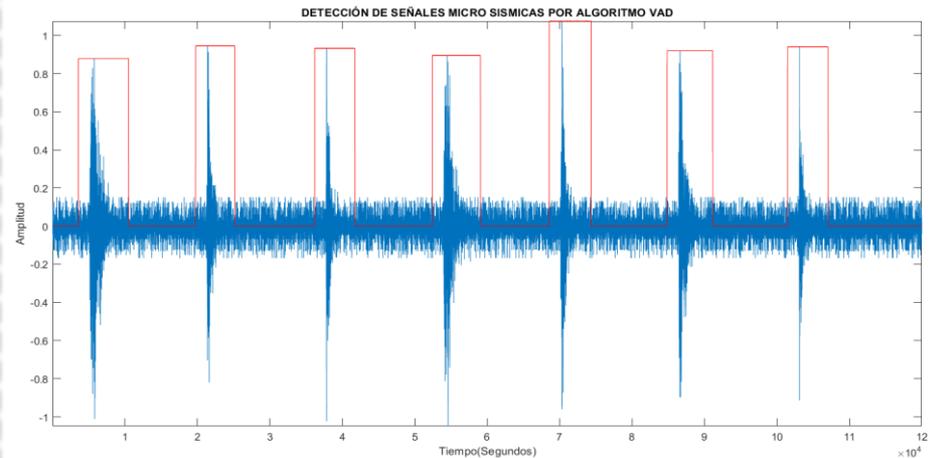
Ber

97.53

0,02

El algoritmo desarrollado no es capaz de detectar 2 en 100 eventos puestos a prueba

Eventos VT



A %

Ber

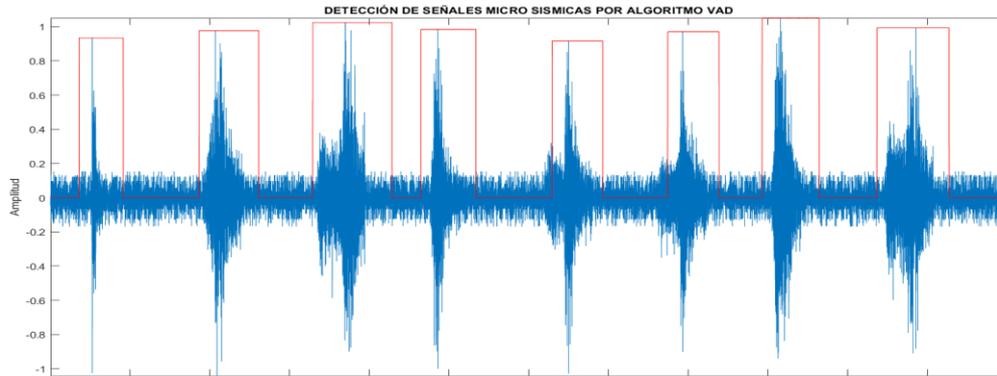
99.97

0,0003

El algoritmo desarrollado no es capaz de detectar 3 en 10000 eventos puestos a prueba

3. Pruebas y Resultados

A %	P %	R %	S %	BER
99.66	99.75	99.75	99.79	0.002



El algoritmo desarrollado no es capaz de detectar 2 en 1000 eventos puestos a prueba

	1	2	3	4	5	6	7	8	9	10
1 'Station'	'Fs (Hz)'	'Year'	'Month'	'NumberD...	'Start Point...	'End Point ...'	'Start Point...	'End Point (...'	'Duration (s)'	
2 'VAT'	100	2023	'Aug'	1	3660	14310	5160	13810	87	
3 'VAT'	100	2023	'Aug'	2	19710	28560	21210	28060	69	
4 'VAT'	100	2023	'Aug'	3	34280	40780	35780	40280	45	
5 'VAT'	100	2023	'Aug'	4	46420	53280	47920	52780	49	
6 'VAT'	100	2023	'Aug'	5	60790	67520	62290	67020	47	
7 'VAT'	100	2023	'Aug'	6	74720	83500	76220	83000	68	
8 'VAT'	100	2023	'Aug'	7	89490	99820	90990	99320	83	
9 'VAT'	100	2023	'Aug'	8	105310	111490	106810	110990	42	

En señales TC es en donde existe mayor dificultad en la detección

4. Conclusiones

- La ejecución exitosa del algoritmo VAD desarrollado en este estudio corrobora su capacidad para identificar señales microsísmicas, cumpliendo su propósito primordial. Este logro se consigue a través de la implementación de una base de datos más amplia y diversa en términos de eventos. **En esta investigación se llega a determinar un valor SNR mínima que actúa como umbral de (4 a 16 dB).** Mediante un ruido generado a partir de las mismas señales microsísmicas de 2 segundos.
- El algoritmo VAD desarrollado y los resultados obtenidos respaldan un nivel de **exactitud del 99.66% con un BER del 0.002.** Esta destacada eficacia proviene de la habilidad del algoritmo para acertar o detectar la gran mayoría de los eventos sometidos a prueba, con escasos o prácticamente nulos fallos. Sin embargo, al considerar la evaluación con eventos por separado, en donde con TC se obtiene el rendimiento más bajo con una **exactitud del 73.35% con un BER del 0.26,** por lo que no cumple con un nivel alto de confiabilidad.



4. Conclusiones

- Mediante la experimentación se consideró varios parámetros, entre ellos la longitud de los segmentos para el ventaneo, **el valor de SNR mínimo para la detección y también el tamaño de la ventana del ruido que para este trabajo se consideró de 2 seg.** De la de igual se mantuvo otros parámetros como el **solapamiento del 50%** al igual que en el trabajo (Rosero Jácome, 2020), donde se logra obtener la mayor cantidad de eventos microsísmicos detectados y el mejor porcentaje de rendimiento.
- Con los porcentajes obtenidos en este estudio se tiene espera **brindar una contribución significativa al continuo trabajo del IGEPN en la detección y prevención de eventos microsísmicos.** Este aporte se enfoca en la construcción de un sistema confiable que permita generar reportes e informes para la Secretaría General de Riesgos. Entidad que desempeña la función de emitir alertas tempranas, las cuales pretender ser un papel crucial en la protección de los ciudadanos, sus bienes materiales e incluso sus vidas.



4. Trabajos a futuro

- Se propone trabajar modificando el vector de ruido generado N_s , es decir generándolo de otra manera y teniendo en cuenta otras consideraciones como la amplitud, variar la longitud de este y realizar experimentos, lo cual hará que tome segmentos más grandes y mediante esto elegir los nuevos valores SNR mínimo, umbrales que son 4dB y 16dB, donde se encuentre mejores resultados que los obtenidos en esta investigación.
- Se propone realizar cambios en cuanto a la amplitud del ruido que tiene los registros de 20 minutos, es decir, al vector generado de ruido con señales microsísmicas mezcladas aleatoriamente. De esta manera poder observar el comportamiento del algoritmo frente a situaciones en donde el ruido se asemeja a la señal que se va a analizar.



4. Trabajos a futuro

- Otra propuesta del trabajo futuro se centra en la expansión de la base de datos, enfocándose en eventos más numerosos y específicamente vinculados al volcán Cotopaxi. El énfasis se depositará en los eventos TR y TC, que se han mostrado como desafíos dentro del algoritmo actual. Además, se contempla el análisis de un conjunto de señales aún más amplio, para superar las restricciones actuales y sumergiéndose en un rango más profundo de experimentación. Este enfoque no solo profundizará nuestra comprensión de los matices del proceso, sino que también promoverá el desarrollo de soluciones



GRACIAS

DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES CARRERA EN TELECOMUNICACIONES

**“Creación de algoritmos basados en técnicas de
detección de la actividad de la voz aplicado a
eventos sísmicos en el volcán Llaima”**

Autor: Ortega Noroña Oscar Alfonso

Director del Proyecto: Ing. Román Alcides Lara Cueva, PhD.

6 de septiembre de 2023

