

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



**Optimización de Algoritmo STA/LTA Para la Detección de
Tiempos de Arribo de Ondas Sísmicas Utilizados por el
Instituto Nacional de Sismología, Vulcanología, Meteorología
e Hidrología de Guatemala**

Trabajo de graduación presentado por Diego Roberto Valdez Del Aguila
para optar al grado académico de Licenciado en Ingeniería Electrónica

Guatemala,

2021

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería




**Optimización de Algoritmo STA/LTA Para la Detección de
Tiempos de Arribo de Ondas Sísmicas Utilizados por el
Instituto Nacional de Sismología, Vulcanología, Meteorología
e Hidrología de Guatemala**

Trabajo de graduación presentado por Diego Roberto Valdez Del Aguila
para optar al grado académico de Licenciado en Ingeniería Electrónica

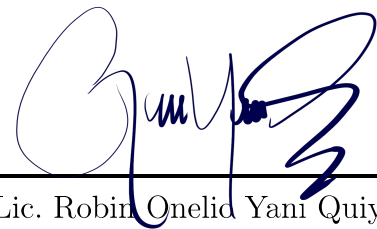
Guatemala,

2021

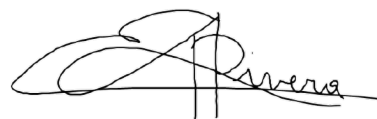
Vo.Bo.:

(f) 
Lic. Robin Onelio Yaní Quiyuch

Tribunal Examinador:

(f) 
Lic. Robin Onelio Yaní Quiyuch

(f) 
MSc. Carlos Alberto Esquit Hernández

(f) 
Dr. Luis Alberto Rivera Estrada

Fecha de aprobación: Guatemala, 25 de enero de 2021.

Con la realización del presente estudio se pretende ayudar al sector nacional público y privado de Guatemala, ya que ambos se ven beneficiados por los procedimientos y acciones del INSIVUMEH. Esta institución proporciona información altamente valiosa para la implementación de proyectos de obra civil, planes de contingencia y prevención de desastres. Me permite también, aplicar las capacidades técnicas y analíticas que he obtenido a lo largo de mis estudios en una institución que trabaja para el país, dándole, en lo personal, mayor relevancia al objeto de estudio.

Aunque el presente trabajo de graduación es de autoría y responsabilidad propia quisiera dedicar los siguientes párrafos a las personas que han formado parte directa o indirecta en la creación de este trabajo de graduación.

Primero, quisiera agradecer a mi abuela y abuelo, quienes en paz descansan, Ester Rodríguez de Del Aguila y Mario Del Aguila, gracias a su esfuerzo y sacrificio fue posible para mi continuar con mis estudios universitarios, además de su apoyo económico fue para mí una motivación continua el poder oportunamente devolver todo ese esfuerzo.

Segundo, quisiera agradecer a las personas del departamento de sismología del INSIVUMEH que formaron parte del trabajo de graduación que se presenta. Diego Castro, por proveerme las capacitaciones necesarias para el manejo de paquetes de software y herramientas utilizadas actualmente en esta institución, por que estuvo siempre dispuesto a resolver mis dudas y por proveerme material de estudio, bibliografía, además por su gran aporte en el material que se incluye en este documento. Ludwing Asturias, porque estuvo siempre dispuesto a resolver mis dudas, por sus aportes en la definición de objetivos, metodología, selección de catálogos, y sus aportes al material que se incluye en este documento.

Quisiera agradecer a mi tío, Luis Del Aguila, pues al momento de fallecimiento de mi abuela fue el quien me apoyo económicamente, también agradezco su paciencia, apoyo, consejos, sabiduría y por siempre luchar para que la familia salga adelante. Quisiera agradecer a mi madre, padre y hermanos, quienes indirectamente han formado parte de este estudio en forma de apoyo moral, dada la necesidad de este apoyo en un proyecto de alta demanda de dedicación, esfuerzo, organización y responsabilidad. Quisiera agradecer a Norma Rodríguez quien siempre ha estado dispuesta a apoyarme cuando ha sido necesario y sin quien no

hubiese sido posible este trabajo de graduación, ya que sin ella no hubiese sido posible la colaboración con el INSIVUMEH.

Quisiera agradecer a Anabela Molina, del departamento de bienestar estudiantil de la Universidad del Valle de Guatemala, quien durante mis años de estudio me asesoró tanto a nivel personal como a nivel académico y siempre me exhortó a finalizar mis estudios y para seguir esforzándome por cumplir mis objetivos. Por último y no menos importante quisiera agradecer a las personas del departamento de ingeniería electrónica, empezando por mi director de carrera Carlos Esquit, Estuardo Mancio y Willy Reyes quienes siempre estuvieron dispuestos a brindarme su apoyo tanto con equipo, como con materiales para la realización de proyectos, entre otros aportes a nivel personal.

| | |
|---|-------------|
| Prefacio | VI |
| Lista de figuras | XII |
| Lista de cuadros | XIV |
| Resumen | XV |
| Abstract | XVII |
| 1. Introducción | 1 |
| 2. Antecedentes | 3 |
| 3. Justificación | 5 |
| 4. Objetivos | 7 |
| 4.1. Objetivo general | 7 |
| 4.2. Objetivos específicos | 7 |
| 5. Alcance | 9 |
| 6. Marco teórico | 11 |
| 6.1. Sismicidad de Guatemala | 11 |
| 6.1.1. Red sismológica de Guatemala | 13 |
| 6.1.2. Fuentes sismogénicas | 14 |
| 6.1.3. Estadísticas de eventos sísmicos | 15 |
| 6.2. Análisis de datos | 16 |
| 6.2.1. Algoritmo STA/LTA | 16 |
| 6.2.2. Descripción de parámetros de algoritmo STA/LTA | 18 |
| 6.2.3. Archivos de onda | 20 |
| 6.2.4. Software y utilidades | 22 |

| | |
|--|-----------|
| 7. Catálogo de eventos sísmicos | 25 |
| 7.1. Creación, administración y selección de registros | 27 |
| 8. Clasificación binaria | 29 |
| 8.1. Curvas ROC | 30 |
| 8.2. Funciones de probabilidad | 31 |
| 9. Introducción a optimización de parámetros STA/LTA | 33 |
| 9.1. Banco de pruebas | 34 |
| 9.2. Proceso de análisis y criterios de selección | 34 |
| 9.3. Comparación entre algoritmo STA/LTA clásico y recursivo | 34 |
| 10. Optimización de parámetros STA/LTA estación FG12 | 37 |
| 10.1. Análisis de curvas ROC para FG12 | 37 |
| 10.2. Análisis histogramas en FG12 | 39 |
| 10.3. Análisis de probabilidades para FG12 | 39 |
| 11. Optimización de parámetros STA/LTA estación ESSG | 41 |
| 11.1. Análisis de curvas ROC para ESSG | 41 |
| 11.2. Análisis de histogramas en ESSG | 43 |
| 11.3. Análisis de probabilidades para ESSG | 43 |
| 12. Optimización de parámetros STA/LTA estación HUEH | 45 |
| 12.1. Análisis de curvas ROC para HUEH | 45 |
| 12.2. Análisis de histogramas para HUEH | 46 |
| 12.3. Análisis de probabilidades para HUEH | 47 |
| 13. Optimización de parámetros STA/LTA estación STG8 | 49 |
| 13.1. Análisis de curvas ROC para STG8 | 49 |
| 13.2. Análisis de histogramas para STG8 | 51 |
| 13.3. Análisis de probabilidades para STG8 | 51 |
| 14. Optimización de parámetros STA/LTA estación SMCA | 53 |
| 14.1. Análisis de curvas ROC para SMCA | 53 |
| 14.2. Análisis de histogramas para SMCA | 55 |
| 14.3. Análisis de probabilidades para SMCA | 56 |
| 15. Optimización de parámetros STA/LTA estación APG | 57 |
| 15.1. Análisis de curvas ROC para APG | 57 |
| 15.2. Análisis de histogramas para APG | 59 |
| 15.3. Análisis de probabilidades para APG | 60 |
| 16. Optimización de parámetros STA/LTA estación RETA | 63 |
| 16.1. Análisis de curvas ROC para RETA | 63 |
| 16.2. Análisis de histogramas para RETA | 65 |
| 16.3. Análisis de probabilidades para RETA | 66 |

| | |
|---|-----------|
| 17.Optimización de parámetros STA/LTA estación CHIE | 67 |
| 17.1. Análisis de curvas ROC para CHIE | 67 |
| 17.2. Análisis de histogramas para CHIE | 68 |
| 17.3. Análisis de probabilidades para CHIE | 69 |
| 18.Conclusiones | 71 |
| 19.Recomendaciones | 73 |
| 20.Bibliografía | 75 |
| 21.Anexos | 77 |
| 21.1. Implementación de STA/LTA clásico en Obspy | 77 |
| 21.2. Implementación de STA/LTA recursivo en Obspy | 78 |
| 21.3. Banco de pruebas | 79 |
| 21.4. Código para obtención de función característica STA/LTA | 80 |
| 21.5. Código para análisis estadístico | 82 |
| 21.6. Archivo de parámetros de RTQUAKE | 86 |
| 21.7. Código para selección de muestra | 86 |
| 21.8. Código para solicitudes y administración de base de datos | 89 |

Lista de figuras

| | | |
|-----|---|----|
| 1. | Mapa de placas tectónicas de Guatemala | 12 |
| 2. | Mapa de fallas principales de Guatemala | 12 |
| 3. | Mapa de estaciones de red sísmica nacional | 13 |
| 4. | Demostración de subducción entre placas tectónica | 14 |
| 5. | Función característica STA/LTA y variables de condiciones de activación STA/LTA. | 17 |
| 6. | Mapa de calor de eventos de subducción locales registrados, 2019. | 25 |
| 7. | Distribución de eventos locales por fuente sísmica genética, 2019. | 25 |
| 8. | Histograma de número eventos de subducción por magnitud, 2019. | 26 |
| 9. | Histograma de número eventos por magnitud, 2019. | 26 |
| 10. | Valoración de curvas ROC. | 30 |
| 11. | Curvas ROC para FG12 Grupo 1 | 38 |
| 12. | Curvas ROC para FG12 Grupo 2 | 38 |
| 13. | Curvas ROC para FG12 Grupo 3 | 38 |
| 14. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en FG12.BHZ | 39 |
| 15. | Histogramas de número de eventos para valores pico de función característica STA/LTA en FG12.BHZ | 39 |
| 16. | Curvas ROC para ESSG Grupo 1 | 42 |
| 17. | Curvas ROC para ESSG Grupo 2 | 42 |
| 18. | Curvas ROC para ESSG Grupo 3 | 42 |
| 19. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en ESSG.BHZ | 43 |
| 20. | Histogramas de número de eventos para valores pico de función característica STA/LTA en ESSG.BHZ | 43 |
| 21. | Curvas ROC para HUEH Grupo 1 | 46 |
| 22. | Curvas ROC para HUEH Grupo 2 | 46 |
| 23. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en HUEH.HHZ | 47 |

| | | |
|-----|--|----|
| 24. | Histogramas de número de eventos para valores pico de función característica STA/LTA en HUEH.HHZ | 47 |
| 25. | Curvas ROC para STG8 Grupo 1 | 50 |
| 26. | Curvas ROC para STG8 Grupo 2 | 50 |
| 27. | Curvas ROC para STG8 Grupo 3 | 50 |
| 28. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en STG8 | 51 |
| 29. | Histogramas de número de eventos para valores pico de función característica STA/LTA en STG8 | 51 |
| 30. | Curvas ROC para SMCA Grupo 1 | 54 |
| 31. | Curvas ROC para SMCA Grupo 2 | 54 |
| 32. | Curvas ROC para SMCA Grupo 3 | 54 |
| 33. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en SMCA | 55 |
| 34. | Histogramas de número de eventos para valores pico de función característica STA/LTA en SMCA | 55 |
| 35. | Curvas ROC para APG Grupo 1 | 58 |
| 36. | Curvas ROC para APG Grupo 2 | 58 |
| 37. | Curvas ROC para APG Grupo 3 | 58 |
| 38. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en APG | 59 |
| 39. | Histogramas de número de eventos para valores pico de función característica STA/LTA en estación APG | 59 |
| 40. | Curvas ROC para RETA Grupo 1 | 64 |
| 41. | Curvas ROC para RETA Grupo 2 | 64 |
| 42. | Curvas ROC para RETA Grupo 3 | 64 |
| 43. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en RETA | 65 |
| 44. | Histogramas de número de Eventos para valores pico de función característica STA/LTA en RETA | 65 |
| 45. | Curvas ROC para CHIE Grupo 1 | 68 |
| 46. | Curvas ROC para CHIE Grupo 2 | 68 |
| 47. | Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en CHIE | 69 |
| 48. | Histogramas de número de eventos para valores pico de función característica STA/LTA en CHIE | 69 |

Lista de cuadros

| | | |
|-----|---|----|
| 1. | Distribución de grupos por magnitud | 26 |
| 2. | Valores propuestos para optimizar el algoritmo STA/LTA por estación | 33 |
| 3. | Cantidad de eventos analizados para FG12.BHZ | 37 |
| 4. | Valores mínimos para $T(S[n])$ en FG12 | 40 |
| 5. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para FG12 | 40 |
| 6. | Cantidad de eventos analizados para ESSG.EHZ | 41 |
| 7. | Valores mínimos para $T(S[n])$ en ESSG | 44 |
| 8. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para ESSG | 44 |
| 9. | Cantidad de eventos analizados para HUEH.HHZ | 45 |
| 10. | Valores mínimos para $T(S[n])$ en HUEH | 47 |
| 11. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para HUEH | 48 |
| 12. | Cantidad de eventos analizados para STG8.BHZ | 49 |
| 13. | Valores mínimos para $T(S[n])$ en STG8 | 52 |
| 14. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para STG8 | 52 |
| 15. | Cantidad de eventos analizados para SMCA.BHZ | 53 |
| 16. | Valores mínimos para $T(S[n])$ en SMCA | 56 |
| 17. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para SMCA | 56 |
| 18. | Cantidad de eventos analizados para APG.BHZ | 57 |
| 19. | Valores mínimos para $T(S[n])$ en APG | 60 |
| 20. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para APG | 61 |
| 21. | Cantidad de eventos analizados para RETA.BHZ | 63 |

| | | |
|-----|--|----|
| 22. | Valores mínimos para $T(S[n])$ en RETA | 66 |
| 23. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para RETA | 66 |
| 24. | Cantidad de eventos analizados para CHIE.EHZ | 67 |
| 25. | Valores mínimos para $T(S[n])$ en CHIE | 69 |
| 26. | Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para CHIE | 70 |

En el presente trabajo de graduación se estudia la respuesta al algoritmo STA/LTA para 8 estaciones de la red nacional sísmica administrada por el INSIVUMEH. Estas corresponden a FG12, ESSG, HUEH, STG8, SMCA, APG, RETA Y CHIE ya que son las más utilizadas para localización hipocentral y los consecuentes procesamientos de la actividad sísmica en la región nacional. El algoritmo STA/LTA consta, principalmente, de 4 parámetros que permiten la detección de eventos, al activar automáticamente el sistema de registro. Estos son conocidos como longitud de ventana corta (STA - short time average), longitud de ventana larga (LTA - long time average), nivel de umbral de activación y nivel de umbral para desactivación. Se propone a través de estos, optimizar el algoritmo STA/LTA para la detección de tiempos de arribo de ondas sísmicas del INSIVUMEH, a manera de minimizar el tiempo de identificación del tiempo de llegada de eventos sísmicos y la ocurrencia de eventos de falsa alarma; identificar los parámetros óptimos para cada una de las 8 estaciones de estudio e identificar estadísticamente los resultados óptimos de acuerdo con cada método de estudio. Como herramientas estadísticas se utiliza un análisis de probabilidades, así como curvas ROC que permiten definir visualmente la capacidad de un algoritmo de clasificar entre objetos (en este caso, señales sísmicas con distintas características) para evaluar qué conjunto de parámetros optimizan la relación entre probabilidad de detección y la probabilidad de falsa alarma. Se utilizó además Python, obspy y numpy para el procesamiento de señales. Se creó tres grupos con eventos por su magnitud, el primero corresponde a eventos con magnitud mayor a 4.5, el grupo 2 son eventos menores a 4.1 y mayores a 3.3, inclusivos. El grupo 3 incluye eventos con magnitudes menores a 3.3. Debido a las similitudes entre la implementación del algoritmo STA/LTA de RTQUAKE, el cual corresponde al software de monitoreo sísmico utilizados en la actualidad por INSIVUMEH, y la implementación de la implementación de STA/LTA clásico de OBSPY se elige este como mejor método de estudio. Fue posible para cada una de las estaciones optimizar la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

This document presents the study of the response for the application of STA/LTA algorithm over 8 stations for the national seismic network. The stations that were subject to study are FG12, ESSG, HUEH, STG8, SMCA, APG, RETA Y CHIE due to their relevance for INSIVUMEH in the subsequent seismic signal processes, once the arrival of a P-wave has been determined. STA/LTA algorithm is primarily defined by four parameters, these being short-time average (STA), long-time average (LTA), on threshold and off threshold. The purpose is to optimize this algorithm for accurate triggering on the P-wave arrival time, in way to minimize detection time, and false alarm triggering and statistically identify the best combination of parameters for each station subject to this study. For this purpose the statistical and visual analysis is performed with ROC curves, which allow to identify an algorithm ability to classify objects, in this case, seismic signals with different characteristics such as noise and actual events. For this implementation the use of python, obspy and numpy were needed. The sample consisted of 3 groups of events. Where, group 1 contains events with magnitude greater than 4.5; group 2, consists of events with magnitude lesser than 4.1 and greater than 3.3, both inclusive; whilst, group 3, contains events with magnitude lesser than 3.3. Due to the similarity shown by RTQUAKE's implementation of STA/LTA algorithm with obspy's STA/LTA classic algorithm we use this function and obtain the STA/LTA characteristic function of seismic events sample, though the analysis using obspy's STA/LTA recursive was also implemented and it showed less precision in all cases. It was possible through this methods to determine the best possible combination of STA/LTA parameters in order to get the best relation between detection probability and false alarm probability in comparison with the set of parameters used today. This combination is $\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$.

En el presente trabajo de graduación se estudia la respuesta al algoritmo STA/LTA para 8 estaciones de la red nacional sísmica administrada por el INSIVUMEH. Estas corresponden a FG12, ESSG, HUEH, STG8, SMCA, APG, RETA Y CHIE ya que son las más utilizadas para localización hipocentral y los consecuentes procesamientos de la actividad sísmica en la región nacional

El algoritmo STA/LTA permite la activación automática de los sistemas de registro cuando un evento sísmico es detectado, este es el primer procesamiento que se realiza sobre las señales de actividad sísmica que están siendo constantemente monitoreadas por una red aún mayor de vigilancia sísmica la cual consiste en detección de eventos locales, regionales y distantes.

La región nacional está sujeta a una alta actividad sísmica debido las placas tectónicas de la región conocidas como placa de cocos, placa de Norteamérica y la placa del caribe y ha sido afectada históricamente por los desastres producidos por actividad sísmica.

El algoritmo STA/LTA consta, principalmente, de 4 parámetros que permiten la detección de eventos, al activar automáticamente el sistema de registro. Estos son conocidos como longitud de ventana corta (STA - short time average), longitud de ventana larga (LTA - long time average), nivel de umbral de activación y nivel de umbral para desactivación.

El primero se designa para vigilar la amplitud instantánea de una señal sísmica, mientras que el segundo vigila la amplitud promedio del ruido presente en una estación, de manera tal que la razón entre ellas, es decir, STA/LTA producen lo que se define como función característica STA/LTA y al ser comparada con el nivel de umbral de activación permite la ignición de los sistemas de registro, y cuando esta decae por debajo del nivel de umbral de desactivación el registro de la señal sísmica se detiene.

La utilización de este algoritmo permite el ahorro de espacio de almacenamiento y debido a los altos costos de un registro continuo por la gran cantidad de información sísmica de hoy en día, resulta en un ahorro económico a la misma vez. Si bien el registro continuo es la

única manera de eliminar por completo la probabilidad de falsas alarmas, su implementación es demasiado costosa lo que hace que esta opción sea descartada inmediatamente.

Para este estudio se utiliza el estándar SEED de intercambio de información sísmica además de Python con librerías de apoyo como Numpy, Obspy y matplotlib entre otras. Debido a que en el INSIVUMEH actualmente se utiliza el software RT Quake que permite la implementación del algoritmo STA/LTA se consideran las especificaciones de este durante el estudio.

Como ha sido mencionado antes, la Librería Obspy permite el procesamiento y manejo de datos sísmicos a través de Python. Por lo cual fue utilizado para obtención de la función característica STA/LTA. Es importante mencionar que este provee dos métodos de cálculo, conocidos como el STA/LTA clásico y el STA/LTA Recursivo. Dadas las diferencias de su cálculo matemático fue estudiado también cuál de los dos provee una mejor precisión en el procesamiento de señales sísmicas. La Librería Numpy fue utilizada para el análisis estadístico y matemático luego de haber procesado la información a través de OBSPY. Y por último, las gráficas y figuras de visualización incluidas en este documento fueron creadas con la utilización de la librería matplotlib.

Se utilizan métodos estadísticos para el análisis de la combinación de parámetros que resultan en la optimización de este algoritmo, entre ellos podemos mencionar la utilización de curvas ROC las cuales funcionan como una herramienta analítica matemática para la comparación entre algoritmos de clasificación, dado que el resultado esperado es la reducción de la ocurrencia de falsas alarmas y maximizar la detección de eventos sísmicos de subducción local con magnitud mayor a 4.5.

Withers M. *et al.* (1998) [1] realizan una comparación entre los algoritmos STA/LTA, Z-statistic, Transiente de frecuencias y el algoritmo de polarización para detección del tiempo de llegada de ondas P con el objetivo de medir y diferenciar las capacidades de estos algoritmos. Para este análisis utilizaron información sísmica sin procesamiento y de esto pudieron determinar que, aunque ninguno de los algoritmos propuestos proveía una solución claramente superior a las otras cuando la información viaja a través de distintas condiciones de canal, receptor, transmisión y condiciones de ruido el algoritmo STA/LTA, al utilizar ventanas de tiempo adaptivas, obtuvo la mejor calificación entre los otros métodos.

Vaezi, Y. y Van der Bann, M. (2015) [2] realizan un estudio con el propósito de comparar el algoritmo de densidad espectral de potencia (PSD) y el algoritmo STA/LTA para la detección de tiempo de llegada de ondas P de eventos de baja magnitud. Aunque el método PSD obtuvo mejor resultados para detección de eventos en aplicaciones remotas al detectar mayor cantidad de eventos a una relación buena de falsa alarma, el método STA/LTA presentó mejores resultados para aplicaciones locales, regionales o industriales.

Kumar, S. *et al.* (2018) estudian la viabilidad de la implementación de un sistema de alerta sísmica basados en el uso del algoritmo STA/LTA con el propósito de mitigar los grandes efectos económicos y sociales de los desastres provocados por terremotos. En este estudio presentan el diseño de su propuesta, y además validan los resultados cuando se ajusta cada elemento de la red con distintos parámetros según el ruido y la actividad sísmica del sector.

Sharma, B. *et al.* (2010) [3] estudian y comparan los algoritmos STA/LTA con sofisticados métodos de reconocimiento de patrones, métodos adaptivos y redes neuronales para la detección del tiempo de llegada de un evento sísmico. Determinaron que ninguno de los algoritmos propuestos es capaz de adaptarse óptimamente a todas las situaciones y variables a las que está sujeta una red sísmica.

De acuerdo a CEPREDENAC (2011) [4] la ocurrencia de sismos severos potencialmente causan daños y pérdidas materiales, así como pérdida de vidas humanas. Estos eventos también causan la interrupción temporal de los servicios de líneas vitales, tales como: agua potable, caminos, salud y educación, entre otros, incidiendo en forma negativa en la sociedad, la actividad económica y el ambiente. Es importante que los países que se encuentran localizados en regiones sísmicamente activas, como es el caso de América Central, conozcan la realidad sismotectónica y la cuantifiquen. Para ello es necesario elaborar estudios técnicos que permitan conocer la intensidad sísmica que se puede presentar en un sitio en particular. La información que generan los estudios es muy importante para que las autoridades competentes puedan implementar normas de diseño de edificaciones que aseguren un adecuado comportamiento sísmico y establezcan medidas de prevención y mitigación ante estos eventos.

El estudio de los terremotos es útil para disminuir los daños que estos producen, y aún cuando no es posible saber cuándo y dónde ocurrirá el próximo terremoto, sí estamos seguros que habrán más, según lo demuestra la evidencia geológica, los registros históricos y la información instrumental.

La elección del tiempo de llegada de ondas sísmicas es un tema crítico de estudio ya que el mismo influencia la efectividad de los siguientes procesos en el análisis de ondas sísmicas, tales como corrección estática, y el modelado de velocidad. Identificaciones erróneas en los tiempos de llegada puede provocar efectos significativos en la identificación de hipocentros. Sin embargo, las trazas sísmicas siempre contienen fuertes señales de ruido con condiciones complejas cerca de la superficie, lo cual provoca que la identificación automática de tiempos de llegada sea uno de los retos principales de estudio.

4.1. Objetivo general

Optimizar el algoritmo STA/LTA para la detección de tiempos de arribo de ondas sísmicas del INSIVUMEH.

4.2. Objetivos específicos

- Identificar estadísticamente el método con resultados óptimos para la identificación de eventos sísmicos entre los distintos métodos de las variantes del algoritmo STA/LTA así como la parametrización de los mismos.
- Identificar los parámetros óptimos de cada estación sísmica de estudio para cada uno de los métodos de estudio.
- Minimizar la ocurrencia de eventos de falsa alarma.
- Minimizar el tiempo de identificación del tiempo de llegada de eventos sísmicos.

En este estudio se optimiza el algoritmo STA/LTA para la red nacional administrada por INSIVUMEH, dado el tamaño de esta se consideran las estaciones más utilizadas para localización hipocentral, estas son FG12, ESSG, HUEH, STG8, SMCA, APG, RETA Y CHIE. Este algoritmo está definido por 8 parámetros,

- Longitud de ventana corta (STA - short time average)
- Longitud de ventana corta (STA - short time average)
- Nivel de umbral de activación
- Nivel de umbral para desactivación
- Esquema de votación
- Filtro pasa bandas
- Tiempo pre evento
- Tiempo post evento

A pesar de que estos ocho son importantes, los primeros cuatro son capaces de definir en su totalidad la respuesta, la probabilidad de detección y la probabilidad de falsa alarma. Mientras que los otros, funcionan como mecanismos para mejorar la calidad y captura de las señales. Es por esto que en este estudio nos basamos en la optimización de la longitud de ventana corta, longitud de ventana corta, nivel de umbral de activación y el nivel de umbral para desactivación.

Durante varios años el desarrollo de métodos para detección automática de eventos sísmicas ha sido objeto de estudio, debido a esto existen varios métodos tanto para la implementación del algoritmo STA/LTA como derivados y aquellos que no tienen relación con el mismo. Debido a la utilización de este y su facilidad de implementación, se estudia el algoritmo STA/LTA clásico y recursivo.

6.1. Sismicidad de Guatemala

Los siguientes párrafos describen la tectónica de Guatemala según lo presentado por INSIVUMEH en Marco Tectónico para Guatemala [5], de lo cual resulta la actividad sísmica nacional.

El territorio nacional está repartido en tres placas tectónicas: Norteamérica, Caribe y Cocos. Los movimientos relativos entre estas determinan los principales rasgos topográficos del país y la distribución de los terremotos y volcanes. El contacto entre las placas de Norteamérica y Caribe es de tipo transcurrente. Su manifestación en la superficie es el sistema de fallas de Chixoy-Polochic y Motagua.

El contacto entre las placas de Cocos y del Caribe es de tipo convergente, en el cual la Placa de Cocos se mete por debajo de la Placa del Caribe (fenómeno conocido como subducción). Este proceso da origen a una gran cantidad de temblores y formación de volcanes. El contacto entre estas dos placas está aproximadamente a 50 Km frente a las costas del Océano Pacífico. A su vez, estos dos procesos generan deformaciones al interior de la Placa del Caribe, produciendo fallamientos secundarios como: Jalpatagua, Mixco, Santa Catarina Pinula, etc. [5]. En la Figura 1, se muestra las placas tectónicas descritas en estos párrafos.

En la Figura 2 se muestra las fallas que afectan el territorio nacional de Guatemala, y están relacionadas a las interacciones entre las placas mencionadas con anterioridad.

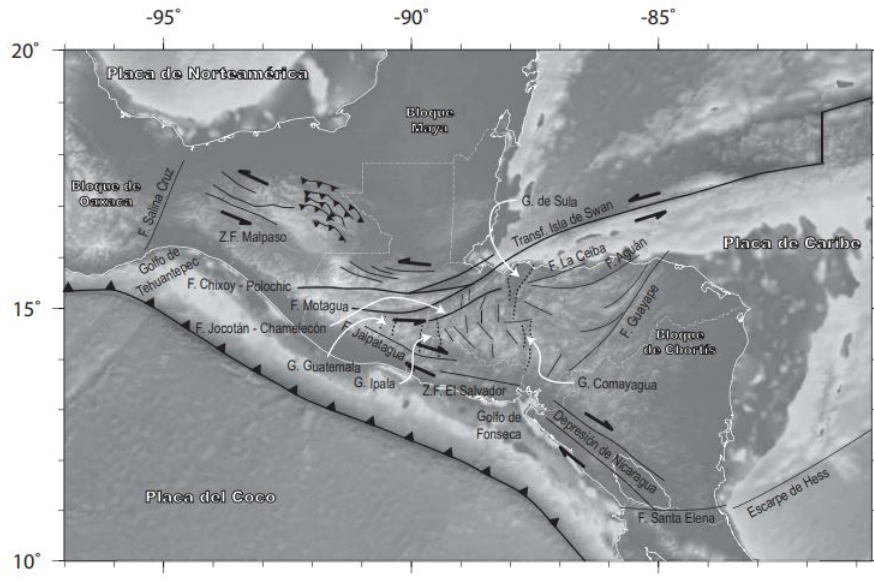


Figura 1: Mapa de placas tectónicas de Guatemala.

Álvarez, J. (2009) [6]

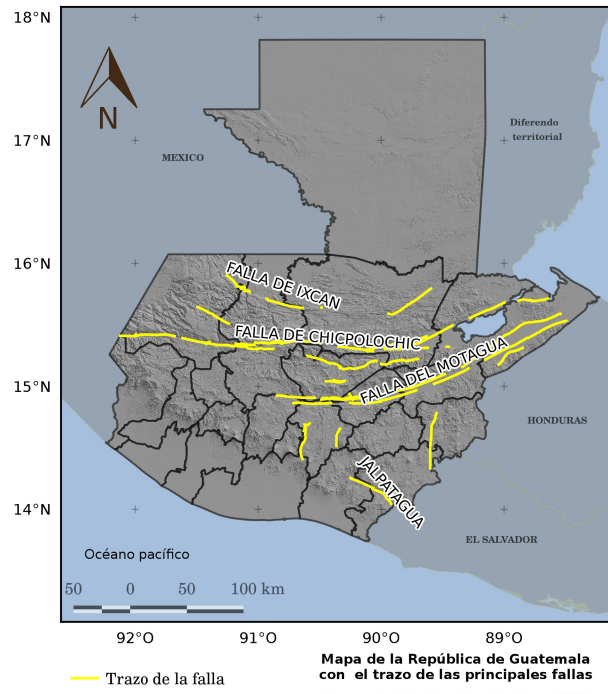


Figura 2: Mapa de fallas principales de Guatemala.

Elaborado por Diego Castro, INSIVUMEH, Guatemala

6.1.1. Red sísmológica de Guatemala

INSIVUMEH es una institución pública de la república de Guatemala a cargo del monitoreo y estudio de la actividad sísmica que afecta a la región nacional. Para ello, cuenta con estaciones sísmicas instaladas alrededor de la región nacional. Esta red empezó operaciones en el año 1976 contando con únicamente estaciones análogas. Desde el 2013 estas han sido reemplazadas por estaciones mas modernas, digitales. En la actualidad esta red consiste en 16 estaciones de banda ancha, 11 estaciones de periodo corto y 5 estaciones para movimiento fuerte. [7]

En la Figura 3 se puede observar las estaciones que forman parte de la red nacional.

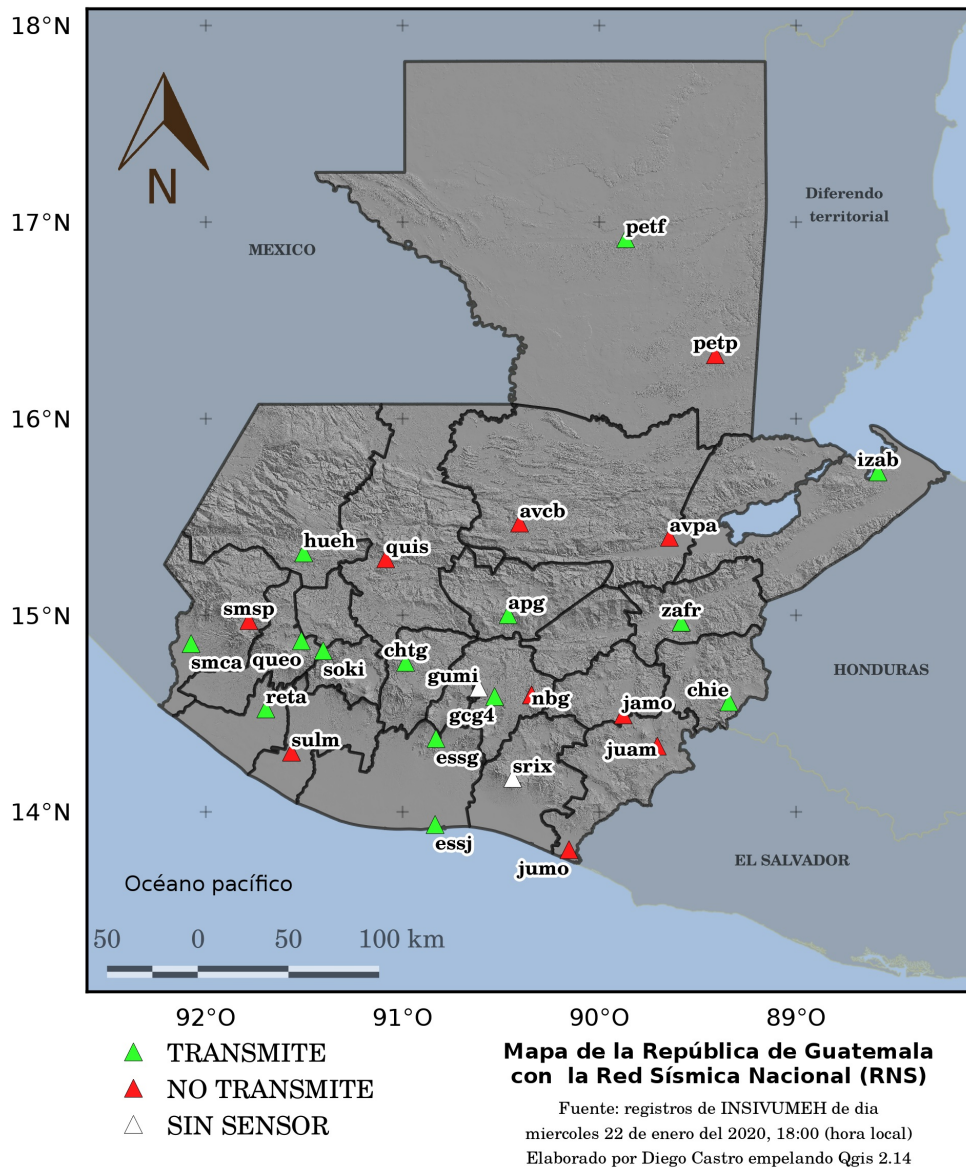


Figura 3: Mapa de estaciones de red sísmica nacional.
Elaborado por Diego Castro, INSIVUMEH, Guatemala

6.1.2. Fuentes sismogénéticas

Tavera, Hernando *et al.* (2014) [8] define una fuente sismo genética como aquella línea, área o volumen geográfico que presenta similitudes geológicas, geofísicas y sísmicas, a tal punto que puede asegurarse que su potencial sísmico es homogéneo en toda la fuente; es decir, que el (los) proceso de generación y recurrencia de sismos es espacial y temporalmente homogéneo. Definir la geometría de la fuente sismogénica es fundamental para la evaluación del peligro sísmico, debido a que proporcionan los principales parámetros físicos que controlan la sismo tectónica de la región en estudio.

Hemos presentado con anterioridad la tectónica de placas en la región nacional de Guatemala (ver Figura 1), y de esto podemos decir que las fuentes sismo genéticas más comúnmente presentes en la actividad sísmica en nuestra región son de subducción, siendo el segundo de tipo cortical según registros del INSIVUMEH, por lo cual a continuación se describe cada una de estas.

Subducción

De acuerdo a Boutelier, D. (2005) [9], las zonas de subducción son todas aquellas zonas en las que se dan los choques o colisiones entre las placas. En la Figura 4 se puede observar lo que ocurre en las zonas de subducción de forma tal que una placa que puede ser oceánica o continental choca con otra y esta se mete por debajo de la otra generando así diversas consecuencias. El 95 % de la energía sísmica del planeta se localiza en las zonas de Subducción y la mayor parte de los volcanes activos de la Tierra erupcionan en estas zonas.

La principal consecuencia que conlleva estas colisiones de placas son la provocación de terremotos que pueden ser de mayor o menor magnitud dependiendo del choque y de otros factores. Otra de las consecuencias de la subducción es la creación de volcanes que aparecen debido a la generación de magma durante la fusión parcial del manto terrestre. Además, esta colisión entre placas puede causar también grandes y profundas fosas oceánicas.

En la Figura 4 se puede observar el fenómeno geológico

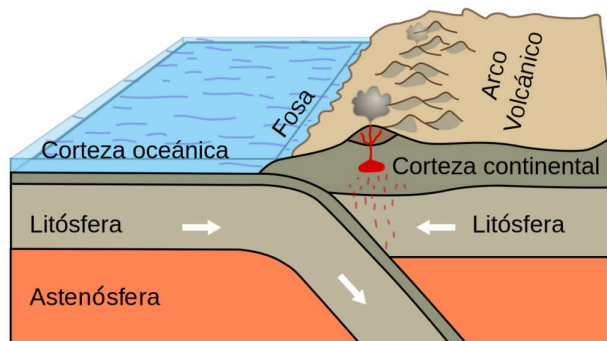


Figura 4: Demostración de subducción entre placas tectónicas.

Mejía, T. [10]

Cortical

Se define un sismo de tipo cortical como todos aquellos que se generan en la corteza, es decir a no más de 50 km. En Guatemala, la mayoría de la sismicidad denominada como cortical no supera los 25 km. Estos son producidos en los sistemas de fallas transcurrentes del norte y las fallas geológicas menores en el interior de la placa del Caribe.

6.1.3. Estadísticas de eventos sísmicos

Para evaluar la peligrosidad sísmica de una zona hay que conocer previamente la sismicidad de la misma. Los parámetros que se manejan comúnmente son los de localización y tamaño. Los parámetros que relacionan tamaño y destructividad de un terremoto son: magnitud, intensidad, aceleración, velocidad y desplazamiento del suelo.

Intensidad sísmica

Es un parámetro no instrumental. Se mide por escalas macrosísmicas que evalúan los efectos producidos en un lugar determinado, sobre las personas, edificios y terreno. Las escalas de intensidad, aunque han evolucionado presentan algunos defectos que hay que tener presentes cuando se realizan evaluaciones de intensidad.

Magnitud

La magnitud está relacionada con la energía sísmica liberada en el foco del terremoto. Este parámetro fue propuesto por Richter en 1935 y su definición original está dada en función de la amplitud máxima de la onda sísmica registrada en un sismógrafo Wood-Anderson a 100 kilómetros de distancia del epicentro.

Aceleración, velocidad y desplazamiento

Estos parámetros son determinados a partir de acelerogramas y acelerógrafos. Por otro lado, los datos sobre aceleraciones son escasos y limitados a regiones muy concretas. se ha utilizado, fundamentalmente, el parámetro aceleración pico, es decir, el valor máximo de la aceleración medida en el registro. Sin embargo, este valor máximo o pico ocurre en un instante y los daños producidos por un terremoto no están únicamente relacionados con estos valores máximos instantáneos, de los que quizá las estructuras podrían recuperarse, sino que un factor tan importante o más que los valores pico es la duración de la sacudida. Es más conveniente, por tanto, utilizar los valores de los dos parámetros, aceleración máxima y duración.

6.2. Análisis de datos

6.2.1. Algoritmo STA/LTA

El algoritmo de short-time-average/long-time-average (STA/LTA) trigger es usualmente utilizado en aplicaciones de movimiento débil que pretenden capturar y almacenar la mayor cantidad de eventos posibles. Este algoritmo es ampliamente utilizado y es casi estándar en paquetes de software para procesamiento de eventos sísmicos en tiempo real. Resulta ser un algoritmo útil para discriminación entre los distintos tipos de eventos sísmicos además de evitar falsas activaciones por ruido sísmico natural o provocado por el humano.

La configuración de los parámetros de este suele resultar en un intercambio entre varias consideraciones instrumentales y sismológicas. El objetivo para la selección de parámetros es a través de la selección de la más alta sensibilidad dadas las características de la señal de interés. Este algoritmo es muy beneficioso en lugares sísmicamente silenciosos donde el ruido natural es el tipo de ruido dominante o bien existe ruido continuo provocado por actividad humana.

Descripción de algoritmo STA/LTA

El algoritmo STA/LTA mantiene una observación continua de los siempre presentes cambios de amplitud de ruido sísmico en la estación y automáticamente ajusta la sensibilidad al verdadero nivel de ruido. Como resultado una significativa sensibilidad del sistema es lograda en largos periodos de baja actividad sísmica, reduciendo de esta manera la ocurrencia de falsa alarma. Este proceso es usualmente independiente en cada uno de los canales de la estación y es repetido continuamente en tiempo real.

El algoritmo procesa señales sísmicas filtradas en dos ventanas de tiempo en movimiento – una de ellas short-time average (STA) y la segunda corresponde a long-time average (LTA). La ventana STA mide la amplitud instantánea y vigilia en busca de actividad sísmica o bien terremotos. La ventana LTA toma en cuenta el nivel actual de ruido sísmico.

Primero, se toma la amplitud absoluta de cada muestra de la señal que ha sido digitalizada. Luego, el promedio de amplitudes absolutas dentro de ambas ventanas de tiempo es calculada. Por último, se calcula la razón entre ambas (STA/LTA). Esta razón se compara continuamente con un nivel de umbral de activación seleccionado por el usuario, si esta razón es mayor al nivel de umbral de activación la estación comienza a grabar en sus registros el evento sísmico.

Cuando consideramos una red sísmica y no solamente una estación, la mayoría de las redes sísmicas cuenta con un mecanismo de votos, que define cuantos canales y cuales deben hacer trigger antes de que la red comience la grabación del evento. Cuando la señal termina gradualmente y la razón STA/LTA es menor a un nivel de umbral de desactivación el canal termina la grabación. En adición a esto, las redes sismológicas usualmente agregan una cantidad de tiempo pre-evento (PEM) y post-evento (PET) que se desea grabar para su posterior almacenamiento.

En la Figura 5 se muestra un evento local típico y las variables de activación para el registro del evento. El gráfico (a) muestra una señal sísmica continua; la gráfica (b) muestra el promedio absoluto de la señal tanto en la ventana STA como en la ventana LTA. La gráfica (c) muestra la razón STA/LTA, el periodo PEM, el periodo PET, las líneas punteadas representan el nivel de umbral de activación y el de desactivación, respectivamente. Finalmente la gráfica (d) muestra la señal que fue capturada.

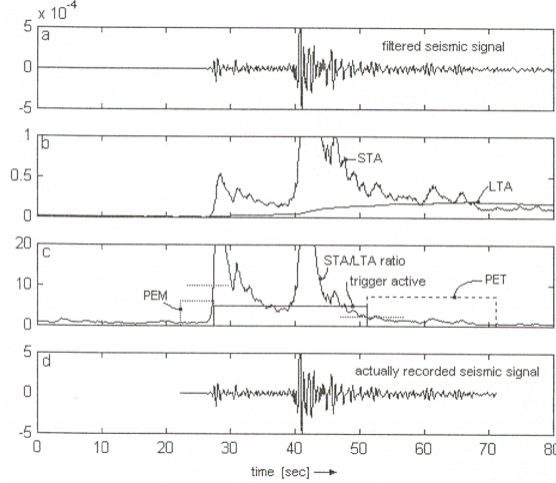


Figura 5: Función característica STA/LTA y variables de condiciones de activación STA/LTA. *Trnkoczy, A. (1999) [11]*

De lo anterior podemos decir que la idea detrás del algoritmo STA/LTA es evaluar de forma continua el valor de la función característica de una señal sísmica en dos ventanas de tiempo, para así, detectar un evento sísmico. La función característica usada puede ser en función de la energía, amplitud absoluta o bien una función descriptiva de trazas de microsismos. Independientemente a esta, cuando la razón entre el valor instantáneo de una señal sísmica (STA) y el valor promedio del ruido sísmico (LTA) excede el valor umbral de activación, las muestras siguientes a este suceso son marcadas como provenientes de un evento sísmico, lo cual a su vez marca el inicio del registro de esta, sí y solo sí todas las condiciones han sido cumplidas.

Algoritmo STA/LTA clásico

El método clásico para el cálculo de ambos valores STA y LTA está definido en las ecuaciones 1 y 2.

$$STA = \frac{1}{s} \left(\sum_{k=t-s}^t CF(x_k) \right) \quad (1)$$

$$LTA = \frac{1}{l} \left(\sum_{k=t-l}^t CF(x_k) \right) \quad (2)$$

Donde, s denota la longitud de la ventana de tiempo corta, en muestras. l denota la longitud de la ventana de tiempo larga, en muestras. t denota el tiempo de la última muestra, en muestras. x_1, x_2, \dots, x_n denota cada muestra tomada de una señal sísmica sin procesamiento previo y $CF(x_k)$ es la función característica de cada muestra. Como se puede observar en anexos 19.1 el paquete obspy, utilizado para el presente estudio define: $CF(x) = x^2$

Algoritmo STA/LTA recursivo

El método recursivo para el cálculo de ambos valores STA y LTA está definido en las ecuaciones 3 y 4.

$$STA_k = \frac{1}{s} (STA_{k-1}(s-1) + CF(x_k)) \quad (3)$$

$$LTA_k = \frac{1}{l} (LTA_{k-1}(l-1) + CF(x_k)) \quad (4)$$

Donde, s denota la longitud de la ventana de tiempo corta, en muestras. l denota la longitud de la ventana de tiempo larga, en muestras. x_1, x_2, \dots, x_n denota cada muestra tomada de una señal sísmica sin procesamiento previo y $CF(x_k)$ es la función característica de cada muestra. Como se puede observar en anexos 19.2 el paquete obspy, utilizado para el presente estudio define: $CF(x) = x^2$

Este método fue desarrollado por Evans, J y Allen S. (1983) [12] con el objetivo de proveer un algoritmo para grabadores sísmicos controlados por medio de microprocesadores. Este tipo de equipos son ampliamente utilizados en aplicaciones de campo, en áreas remotas. Son además muy limitados en su capacidad de almacenamiento y no cuentan con la habilidad de procesar datos en listas. Este método es óptimo para aplicaciones en las cuales la detección de eventos de tipo local no es un requerimiento.

Es importante decir que en los próximos capítulos se define a la función característica STA/LTA de una señal sísmica $x[n]$, según la ecuación 5

$$T(x) = \frac{STA}{LTA} \quad (5)$$

6.2.2. Descripción de parámetros de algoritmo STA/LTA

En esta sección se presentan los parámetros, así como su función dentro de la respuesta del algoritmo, de manera tal que se expliquen los efectos que cada uno de estos produce en la función característica STA/LTA de las señales digitalizadas, filtradas y luego procesadas por este algoritmo para la activación y desactivación de registro de eventos. Es importante resaltar que la selección de estos depende de las condiciones de ruido sísmico en las estaciones, las propiedades de las señales sísmicas de interés y del tipo de sensores.

Longitud de ventana STA

Este parámetro está dado en segundos y mide la amplitud instantánea de una señal sísmica y es la principal responsable en la detección de terremotos. Regularmente se decide una longitud de ventana STA mayor a unos cuantos periodos de la señal sísmica esperada. Dado que esta señal ha sido digitalizada podemos decir entonces que STA debe ser mayor a T_s , también se debe cumplir por teorema de muestreo de Nyquist-Shannon que $f_s > 2f_{max}$ de la señal esperada. Donde T_s es el periodo de muestreo, f_s es la frecuencia de muestreo y f_{max} es la frecuencia máxima esperada de la señal sísmica.

Este parámetro funciona como un filtro dado que un valor de STA pequeño provoca un mayor sensibilidad a eventos locales de duración corta. Mientras que un valor STA alto provoca una mayor sensibilidad a eventos de larga duración y menor frecuencia. De lo cual podemos decir entonces que un valor pequeño para la longitud de STA provocará mayor sensibilidad a eventos de falsa alarma.

Longitud de ventana LTA

Este parámetro esta dado en segundos y mide la amplitud promedio del ruido sísmico. Está dictada por unos cuantos periodos de las fluctuaciones típicas de ruido. Con este parámetro podemos controlar la sensibilidad a eventos alrededor del canal de una estación en un rango de 200 Km a 1500 Km. Valores altos de LTA producen una mayor razón STA/LTA lo que a su vez produce mayor sensibilidad a la activación de registro de ondas primarias débiles. En el caso en el que una onda primaria no es detectada puede contaminar la información sobre ruido antes del evento.

Nivel de umbral de activación

Este parámetro es adimensional puesto que es dependiente de la función característica STA/LTA la cual es también adimensional. Este depende altamente del ruido sísmico en las estaciones. Por lo cual un nivel de ruido sísmico alto demanda un nivel de umbral de activación igualmente alto para evitar la ocurrencia de falsa alarma. El valor de 4 es ampliamente utilizado en áreas donde el ruido provocado por actividad humana es sumamente bajo o bien inexistente. Mientras que valores mas altos, como 8, son utilizados ampliamente en áreas expuestas a niveles ruido sísmico altos que es provocado por el humano.

Nivel de umbral de desactivación

Este parámetro es adimensional y es dependiente del nivel de umbral de activación lo cual significa que es también seleccionado a partir de las características propias del ruido presente en una estación. Regularmente se utiliza un valor porcentual, regularmente entre 45% y 75%, del nivel de un umbral de activación. Una diferencia muy alta entre ambos puede provocar un registro continuo de actividad sísmica una vez la función característica STA/LTA ha sobrepasado el nivel de umbral de activación ya que si el ruido es mayor al de desactivación este podría tener una desactivación tardía o continua. Mientras que

diferencias muy bajas entre ambos pueden provocar que la finalización de un evento sísmico sea declarada de manera temprana.

Otros parámetros

Los parámetros que se mencionan a continuación no están en el alcance del estudio que se presenta en este documento, sin embargo, los mencionamos para proveer una mayor claridad sobre por qué estos no fueron considerados como parte la optimización del algoritmo STA/LTA.

- **Tiempo pre-evento:** este parámetro define la cantidad de tiempo que se desea registrar previo a la ocurrencia de un evento sísmico y no afecta los resultados sobre la activación o desactivación de la grabación de este. Por lo cual este parámetro depende de los estudios posteriores que se realizan sobre la actividad sísmica detectada, las cuales están fuera del alcance del presente estudio.
- **Tiempo post-evento:** este parámetro define la cantidad de tiempo que se desea registrar posterior a la ocurrencia de un evento sísmico y no afecta los resultados sobre la activación o desactivación de la grabación de este. Por lo cual este parámetro depende de los estudios posteriores que se realizan sobre la actividad sísmica detectada, las cuales están fuera del alcance del presente estudio.
- **Esquema de votación para activación:** este parámetro rige la activación o desactivación del registro de eventos en una red dado un conjunto de estaciones que han declarado un evento. Para este se define cuantas y cuales estaciones deben realizar activación para que el evento sea registrado.
- **Filtro pasa bandas:** este permite la eliminación de la componente DC de las señales sísmicas, así como delimitar las frecuencias a las que se encuentran las señales de interés eliminando efectos de ruido fuera de estos límites.

6.2.3. Archivos de onda

La sismología es, en gran extensión, una ciencia basada en observación, modelado y estudio de grandes cantidades de información. El proceso de estudio requiere herramientas de procesamiento simples, robustas y rápidas. Bajo este contexto el formato en el cual la información es almacenada e intercambiada es esencial, estos deben proveer la información necesaria para ser reproducibles además de contener las características de procedencia.

Dada una variedad de estándares para formatos de almacenamiento e intercambio de información sísmica se consideran las siguientes características fundamentales [13]:

- **Eficiencia:** describe la capacidad del formato para almacenar información con un bajo consumo de recursos.

- Organización de datos: describe la capacidad del formato para almacenar datos descriptivos, tales como: formas de onda, información sobre receptor y fuente e información genérica sobre eventos sísmicos.
- Intercambio de datos: describe la habilidad del formato para contener conjuntos de datos complejos en un estructura simple e intuitiva.
- Reproducibilidad: un aspecto crítico en investigación es la habilidad de reproducir resultados, por lo cual estos formatos deben proveer una total reproducción de los conjuntos de datos de origen.

SEED

Es un formato internacional estándar para el intercambio de información sísmica. SEED fue diseñado para el uso por parte de la comunidad de investigación de actividad sísmica, principalmente para el intercambio de data del movimiento de la tierra. Es un formato para data digital medida en un punto en el espacio a iguales intervalos de tiempo.

Este formato es ampliamente utilizado para transferencia de datos entre estación y el centro de recolección de datos, esta misma es luego transmitida a un centro de administración de datos con el objetivo final de ser presentado ante el usuario. Puede ser utilizado para transmisión y recepción de datos tanto tiempo real como para futuras solicitudes.

Este formato no fue diseñado para el intercambio de data sísmica procesada (ej. filtrada). Este formato requiere también de series de tiempo muestreadas a intervalos constantes de tiempo. Debido a sus características es ampliamente utilizado por individuos e instituciones para coleccionar, registrar, transmitir y visualizar data sísmica.

En sus implementaciones mas recientes el formato a tomado dos formas conocidas como miniSEED y dataless SEED [14]. El primero utilizado para el intercambio de formas de onda, mientras que el segundo es utilizado para metadatos. Esto sucede debido al amplio uso de bases de datos relacionales en las cuales se almacenan metadatos y formas de onda de forma separada cuando estos son colectados. La utilización del formato miniSEED ha visto una creciente utilización debido a su capacidad de ser combinada con dataless SEED.

A continuación, se detallan las características de este formato que se consideran, para el propósito del presente estudio, como las características más importantes (además de las mencionadas con anterioridad):

- Eficiente: minimiza la cantidad de espacio no utilizado para almacenamiento e intercambio.
- Portable: adecuado para fácil y eficaz lectura en cualquier computadora disponible comercialmente.
- Auto-correctible: aunque errores al momento del registro de datos sobre eventos sísmicos es inevitable, este formato provee medios para corrección entre volúmenes lógicos subsecuentes.

- Provee combinación de datos: los centros de colección y administración de datos pueden combinar grandes cantidades de volúmenes lógicos a uno nuevo con cambios mínimos en la información auxiliar y sin cambio en la estructura de bloques para la mayor parte de los encabezados de control.

Estas características son consideradas esenciales debido a las características propias de la red sísmológica del INSIVUMEH de Guatemala, así como de las herramientas disponibles para el presente estudio.

Las características de portabilidad y eficiencia se consideran dado los recursos computacionales con los que se cuenta actualmente dado que se utilizó y procesó grandes cantidades de información.

Las características de ser auto corregible y proveer combinación de datos se consideran de alta utilidad dado el presente objetivo de optimizar los algoritmos STA/LTA. Durante el estudio se puede encontrar errores de registro tales como gaps, estos son considerados como un registro no continuo o interrumpido de eventos sísmicos dados los parámetros de umbral de activación y desactivación ya que estos no se encuentran en su punto óptimo.

6.2.4. Software y utilidades

En esta sección se introduce a los paquetes de software y utilidades sobre los cuales se basan los análisis realizados en el presente estudio.

RT Quake

Este es un sistema para monitoreo, detección y registro de actividad sísmica provenientes de datos previamente digitalizados y en concordancia con el estándar de SeedLink. Pretende ser utilizado en su mayoría en redes locales y regionales. Está escrito en C y utiliza Python, OpenGL y GD para gráficos. Este sistema funciona de manera similar a Seiscomp3 sin embargo, no proporciona tantas utilidades como las que este proporciona.

Una de sus características principales es la capacidad para registrar eventos directamente en la base de datos SEISAN lo cual permite el procesamiento de datos inmediatamente. Para cada evento se registra un s-file el cual contiene todos los datos de parámetros así como del evento y los enlaces a los archivos que contienen las formas de onda.

Descripción de algoritmo de activación

1. Lectura de los archivos de parámetros en el servidor. (Este archivo con los parámetros óptimos para cada estación fue incluido en Anexos)
2. Se registra la información de cada canal en un arreglo de datos bidimensional.
3. Cada entrada nueva de datos es agregada al inicio del arreglo de datos, y el dato más antiguo eliminado del mismo.

4. Se verifica la activación del algoritmo a través de la implementación de un algoritmo STA/LTA estándar.

Es importante aclarar que según lo definido por Utheim, T. y Havskov J. (2017) [15] cada componente de la red puede tener diferentes condiciones de activación y aunque exista retraso por parte de las diferentes estaciones en una red ya sea por la propagación de ondas o por inestabilidad en la transmisión de datos los eventos pueden ser declarados.

Python

Es un lenguaje de programación de alto nivel y de propósito general creado por Guido van Rossum, en este se enfatiza la facilidad de su lectura debido a su notable uso de espacios en blanco. Es de tipo interprete, aunque en sus mas recientes versiones soporta un enfoque orientado a objetos.

Obspy

Es un proyecto open-source dedicado a proveer una estructura para el procesado de data sismológica. Contiene soporte para los formatos mas comúnmente usados por la comunidad así como funciones para el procesamiento de señales sismológicas que permiten la manipulación de series de tiempo.

Numpy

Numpy es un proyecto open- source que permite computación numérica con Python. Fue creado en 2005.

Matplotlib

Es una librería comprensiva para crear visualizaciones estáticas, animadas e interactivas en Python. Permite la creación de gráficos de calidad completamente personalizables para publicaciones científicas.

Catálogo de eventos sísmicos

Durante el año 2019 fueron registrados un total de 3446 eventos sísmicos por el departamento de sismología del INSIVUMEH. La Figura 6 muestra la distribución de eventos por subducción local registrados por mes y día. En la Figura 7 se muestra la distribución de eventos locales registrados según su fuente sismo genética, para la cual podemos observar un 57 % (1345 eventos) de subducción local contra un 43 % corticales (1020 eventos).

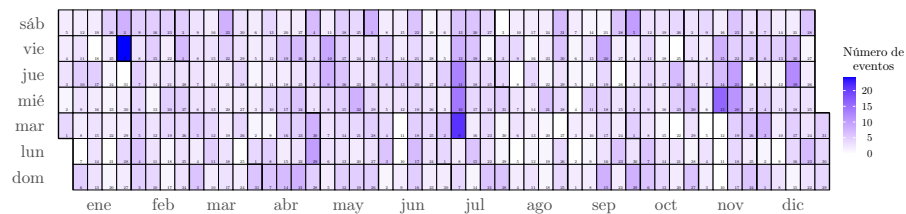


Figura 6: Mapa de calor de eventos de subducción locales registrados, 2019. Gráfico elaborado por Ludwing Asturias, INSIVUMEH, Guatemala.

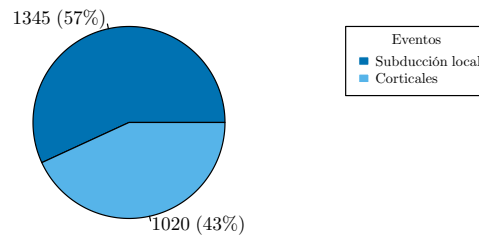


Figura 7: Distribución de eventos locales por fuente sismo genética, 2019. Gráfico elaborado por Ludwing Asturias, INSIVUMEH, Guatemala.

En la Figura 8 podemos observar el número total de eventos registrados según su magnitud mientras que en la Figura 9 podemos observar el total de eventos de subducción local registrados por magnitud. De las mismas es evidente la representatividad de los eventos de subducción registrados para una magnitud menor a 4.0.

Dadas las características, descritas en los párrafos anteriores, de la actividad sísmica en la región de Guatemala y la presencia de eventos de subducción local en 39 % del total de eventos registrados, se decide entonces trabajar con eventos por subducción locales.

De acuerdo con los cuartiles de magnitud tenemos que el 50 % de eventos por subducción locales registrados tienen una magnitud entre 3.5 y 4.0, con eventos menores a 3.5 y aquellos mayores a 4.0 con igual representatividad en 25 % cada uno. El promedio de magnitud es de 3.7 con desviación estándar de 0.4. De estas estadísticas se sugiere realizar tres grupos para el catálogo seleccionado. Estos grupos, así como su objetivo de estudio se muestran en el Cuadro 1.

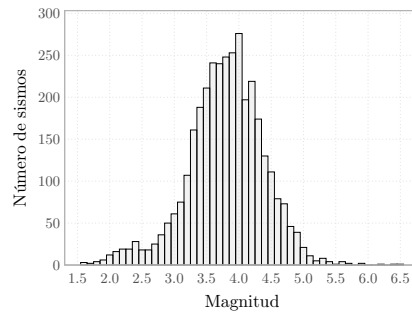


Figura 8: Histograma de número eventos de subducción por magnitud, 2019.
Gráfico elaborado por Ludwing Asturias, INSIVUMEH, Guatemala.

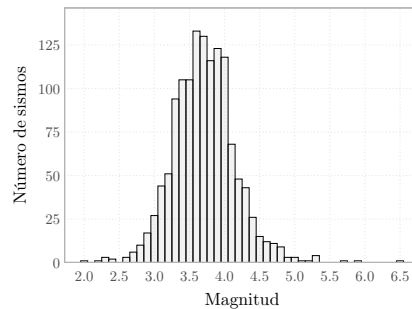


Figura 9: Histograma de número eventos por magnitud, 2019.
Gráfico elaborado por Ludwing Asturias, INSIVUMEH, Guatemala.

| Grupo | Característica | Muestra | Resultado esperado |
|-------|------------------------------|------------|-------------------------------------|
| 1 | $Magnitud > 4.5$ | 62 eventos | Activar el disparador siempre |
| 2 | $3.3 \leq Magnitud \leq 4.1$ | 48 eventos | Mínima perdida de eventos |
| 3 | $Magnitud < 3.3$ | 48 eventos | Minimizar número de falsos disparos |

Cuadro 1: Distribución de grupos por magnitud

7.1. Creación, administración y selección de registros

En la actualidad en el INSIVUMEH se utilizan las estaciones FG12, ESSG, HUEH, STG8, SMCA, APG, RETA Y CHIE para la localización hipocentral y los consecuentes procesamientos de la actividad sísmica en la región nacional. Es por esto que se realiza la optimización sobre las mismas.

Para la creación de la muestra se consideró contar con una representatividad de eventos durante todo el año, además al ser los eventos de grupo 1 en su mayoría pertenecientes al grupo de control, se utiliza los 62 eventos contenidos en la misma.

A partir de aquí se creó una base de datos relacional que contiene la información general sobre tiempos de inicio por estación, tiempo final, archivo que contiene el evento por fecha, hora y su identificación Alpha numérica para su posterior procesamiento. Los códigos de selección de muestra y de administración de solicitudes a la base de datos han sido incluidos en anexos.

Clasificación binaria

En este capítulo se presenta la metodología utilizada para la valoración y selección de los parámetros que logran proveer la mejor relación de falsos disparos contra detección de eventos reales.

Para caracterizar el rendimiento de estos parámetros se propone realizar una prueba de clasificación de hipótesis donde se distinga la presencia de un evento sísmico real, o bien la ausencia de este en la función característica del algoritmo STA/LTA. Para ellos consideramos lo siguiente:

Sea x_1, x_2, \dots, x_n una muestra de datos de un registro sísmico. Considere dos hipótesis simples:

$$H_0 : x[n] = w[n] \tag{6}$$

$$H_1 : x[n] = s[n] + w[n] \tag{7}$$

Donde $w[n]$ representa ruido sísmico y $s[n]$ representa señal sísmica. El objetivo es observar x y decidir cuál es la mejor hipótesis:

$$R_0 = \{x : \text{decidir } H_0\} \tag{8}$$

$$R_1 = \{x : \text{decidir } H_1\} \tag{9}$$

A lo largo de la aplicación del algoritmo STA/LTA nos encontraremos los siguientes casos.

- **Sensitividad:** cuando H_1 es verdadero y se decide H_1 . Este es el caso de un evento sísmico detectado. Su probabilidad será denotada por P_D .

- **Error Tipo I:** cuando H_0 es verdadero pero se decide H_1 . Este es el caso de una falsa alarma. Su probabilidad será denotada por P_{FA} .
- **Error Tipo II:** cuando se tiene H_1 y se decide H_0 . Representa el caso de un evento perdido. Su probabilidad es $1 - P_D$.
- **Especificidad:** cuando H_0 es verdadero y se decide H_0 . Aquí se ha rechazado correctamente que el ruido pueda ser un evento real. Su probabilidad es $1 - P_{FA}$

8.1. Curvas ROC

Un análisis ROC (Receiver Operating Characteristic), característica operativa del receptor, es una herramienta útil para evaluar la precisión de un modelo estadístico que clasifica muestras en 1 o 2 categorías evento sísmico real o falsa alarma. Las medidas fundamentales para medir la precisión del algoritmo de clasificación son sensibilidad y especificidad, ambas definidas para nuestro propósito en la sección anterior.

Una curva ROC es un gráfico que muestra la sensibilidad en el eje y versus el complementario de la especificidad, que para nuestro caso fue descrito como error tipo I en la sección anterior. Para obtener una clasificación perfecta se espera que la curva pase sobre la línea que conecta (0, 0) con (0, 1) y la línea que conecta (0, 1) con (1, 1). De la Figura 10 podemos ver el caso A donde el algoritmo es completamente capaz de catalogar entre eventos reales de falsas alarmas, el caso B donde el algoritmo declara falsas alarmas parcialmente y el C donde el algoritmo es incapaz de clasificar.

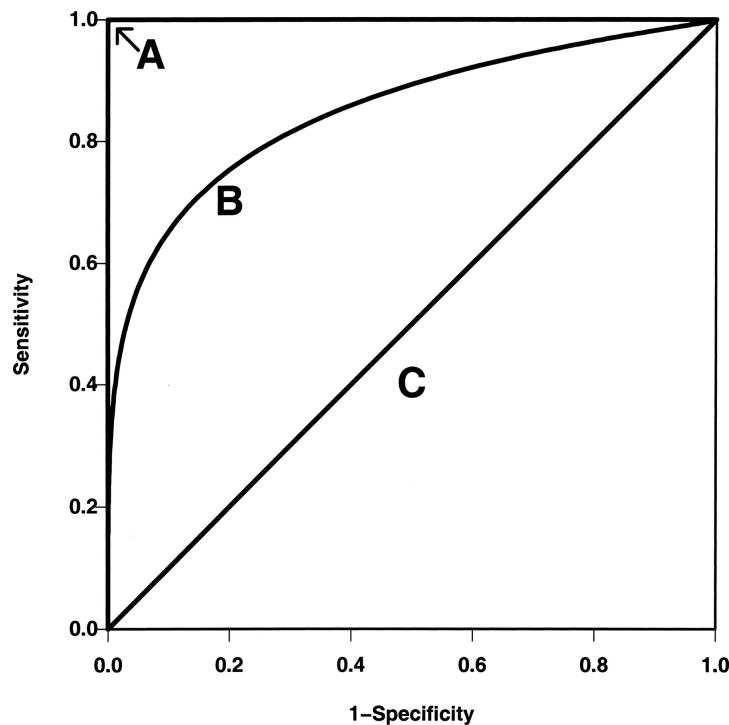


Figura 10: Valoración de curvas ROC.

Markham, K. (2014) [16]

8.2. Funciones de probabilidad

Bajo los conceptos de la sección anterior y retomando los conceptos en las ecuaciones 6, 7, 8, 9 definimos la probabilidad de detección y la probabilidad de falsa alarma por:

$$P_{FA} = \int_{x \in R_1} f(x|H_0 \text{ es verdadero})dx \quad (10)$$

$$P_D = \int_{x \in R_1} f(x|H_1 \text{ es verdadero})dx \quad (11)$$

Donde f es la función de distribución de probabilidad. Típicamente las regiones R_0 y R_1 están definidas usando una función $T(x)$ y un nivel umbral η : R_0 cuando $T(x) < \eta$ y R_1 cuando $T(x) > \eta$. En nuestro caso $T(x)$ será la función característica del algoritmo STA/LTA y η será el nivel umbral de activación STA/LTA. Denotaremos por $f_{H_0}(T)$ a la función de densidad de probabilidad de $T(x)$ cuando H_0 es verdadero y $f_{H_1}(T)$ a la función de densidad de probabilidad de $T(x)$ cuando H_1 es verdadero. De ahí se tiene:

$$P_{FA} = \int_{\eta}^{\infty} f_{H_0}(T)dT \quad (12)$$

$$P_D = \int_{\eta}^{\infty} f_{H_1}(T)dT \quad (13)$$

Para describir la compensación entre falsas alarmas y eventos reales del algoritmo STA/LTA graficaremos P_D contra P_{FA} para varios valores de η . Obteniendo de esta manera la curva ROC para un conjunto de valores longitud sta y longitud lta. Buscando de esta manera obtener la curva que se aproxime a el caso A de la Figura 10.

 Introducción a optimización de parámetros STA/LTA

Este capítulo presenta una breve descripción de las reglas generales utilizadas para la optimización de parámetros del algoritmo STA/LTA de las estaciones, de esta manera también se pretende introducir el conjunto de criterios utilizados para la toma de decisiones, las excepciones a estas reglas serán explicadas a detalle para cada una de las estaciones de estudio.

Dada la naturaleza del procesamiento de datos realizado por INSIVUMEH para localización hipocentral de un evento sísmico, se considera el canal vertical de cada estación, es decir BHZ según el estándar internacional, como el canal para el cual se desea optimizar la parametrización.

El Cuadro 2 muestra el resumen del mejor conjunto de parámetros para cada estación analizada. El proceso para su selección se encuentra en los siguientes capítulos. Recordamos entonces que en el capítulo “Clasificación Binaria” hemos definido que η corresponde a un nivel umbral, por lo cual a propósito del análisis de parámetros definiremos η_a como nivel de umbral de activación y a η_d como nivel de umbral de desactivación.

| Estación | STA (s) | LTA (s) | η_a | η_d |
|----------|---------|---------|----------|----------|
| FG12 | 0.5 | 70 | 4.0 | 2.6 |
| ESSG | 1.0 | 120 | 7.3 | 4.0 |
| HUEH | 1.0 | 70 | 3.4 | 2.2 |
| STG8 | 1.5 | 120 | 6.2 | 4.0 |
| SMCA | 1.0 | 70 | 5.6 | 3.6 |
| APG | 5.0 | 100 | 6.6 | 4.2 |
| RETA | 1.5 | 60 | 8.0 | 5.2 |
| CHIE | 1.0 | 120 | 4.2 | 2.7 |

Cuadro 2: Valores propuestos para optimizar el algoritmo STA/LTA por estación

9.1. Banco de pruebas

Cada una de las estaciones presentadas en el Cuadro 2 fue sometida a las simulaciones presentadas en anexos 21.3 y a partir del análisis visual y estadístico se obtuvo el conjunto de parámetros que presenta la mejor relación entre la probabilidad de detección y la probabilidad de falsa alarma. Además, los anexos 21.4 y 21.5 muestran los códigos de la implementación.

9.2. Proceso de análisis y criterios de selección

1. Seleccionar la combinación de parámetros STA, LTA que resultó en la mejor relación de probabilidad de falsa alarma y probabilidad declaración de eventos y el conjunto de parámetros utilizados en la actualidad, el cual corresponde a STA=2s, LTA=100s, Nivel de Umbral de Activación =3.5, Nivel de Umbral de Desactivación = 1.5.
2. Comparación de curvas ROC para cada conjunto de parámetros seleccionados con anterioridad con el propósito de demostrar cuál de los conjuntos de parámetros muestra una mejor r probabilidad de detección (P_D) versus la probabilidad de falsa alarma (P_{FA}) (Ver Figura 10).
3. Seleccionamos un valor $\eta_a < \min(T(S[n]))$ y $\eta_a \approx \min(T(S[n]))$ donde $S[n]$ es un conjunto de registros de valores de pico de eventos sísmicos en la estación de estudio.
4. Calculamos P_{FA} y P_D para ambos conjuntos de parámetros. Para el conjunto de parámetros utilizado en la actualidad se considera también su valor de P_{FA} así como P_D a manera de demostrar cuál de los dos muestra la mejor combinación entre probabilidad de detección (P_D) y probabilidad de falsa alarma (P_{FA}).
5. Por último, seleccionamos un η_d como un porcentaje de η_a que esté unas cuantas unidades por encima de la densidad de probabilidad de valores STA/LTA para la sección de ruido, a manera tal, que $\eta_d < \eta_a$.

9.3. Comparación entre algoritmo STA/LTA clásico y recursivo

Dada la naturaleza del algoritmo STA/LTA recursivo y el propósito para el cual este fue desarrollado vemos que no es el ideal para realizar esta optimización debido a las razones que enumeramos a continuación.

- Este fue desarrollado para aplicaciones en área remotas, mientras que nuestro enfoque de estudio es sobre una red nacional interconectada y con capacidad de procesamiento en tiempo real.
- Nuestro grupo de eventos objetivos para la optimización de parámetros es aquellos eventos de subducción local, mientras que el algoritmo recursivo fue diseñado y es óptimo cuando la detección de eventos de tipo local no son requeridos.

- La implementación del algoritmo STA/LTA recursivo provee una solución alternativa para aquellas aplicaciones donde el manejo de arreglos de datos no es una opción, mientras que el software RT QUAKE utiliza arreglos de datos para su implementación del algoritmo STA/LTA
- Dada la alta similitud entre la implementación del algoritmo STA/LTA en el software RTQUAKE y la implementación del algoritmo STA/LTA clásico en Obspy, se valora mejor la utilidad del mismo para el proceso de optimización que realizamos en los siguientes capítulos.

De lo anterior podemos resaltar que se elige realizar el análisis y procedimientos de optimización basados en la implementación del algoritmo STA/LTA clásico en obspy dada su alta similitud con la implementación de este en RTQUAKE. Por lo tanto, los conjuntos de parámetros que optimizan todos los resultados que serán presentados en los siguientes capítulos corresponden a la implementación del algoritmo STA/LTA clásico de obspy.

 Optimización de parámetros STA/LTA estación FG12

En el Cuadro 3 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación FG12 fue registrada para el 92 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 el cual se detalla en el Cuadro 1 la estación fue registrada en 73 % de los mismos, y a su vez fue registrada en 65 % de los eventos para el grupo 3 (ver Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 57 | 92 % |
| 2 | 35 | 73 % |
| 3 | 31 | 65 % |

Cuadro 3: Cantidad de eventos analizados para FG12.BHZ

Los conjuntos de parámetros analizados para la estación FG12 corresponden a $\{\text{STA} = 2\text{s}, \text{LTA} = 100\text{s}\}$ y $\{\text{STA} = 0.5\text{s}, \text{LTA} = 70\text{s}\}$

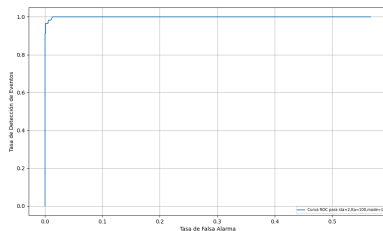
10.1. Análisis de curvas ROC para FG12

En la Figura 11a se muestra la respuesta de la función característica STA/LTA para eventos de grupo 1 dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 11b se muestra la respuesta, para el mismo grupo, de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Ambos parámetros resultan en una curva ROC considerada buena, lo cual denota que el grupo 1 esta distribuido de una forma tal que permite la captura completa de eventos con una mínima probabilidad

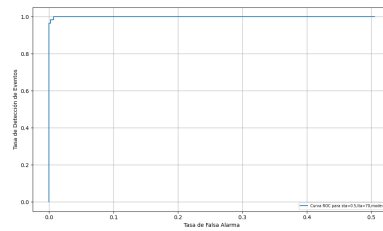
de falsa alarma.

La Figura 12 muestra la respuesta de la función característica STA/LTA para el grupo 2, es evidente entonces que la relación entre la probabilidad de detección y la probabilidad de falsa alarma mejora considerablemente con los parámetros propuestos.

Como definimos en el Cuadro 1 el grupo 3 nos servirá para evidenciar la mejora entre falsos disparos y eventos declarados debido a su proximidad en amplitud a la del ruido, por lo cual, al analizar la Figura 13 podemos ver que (b) se aproxima considerablemente mejor al punto óptimo definido en la Figura 10, esto denota que la probabilidad de detección versus la probabilidad de falsa alarma puede ser mejor.

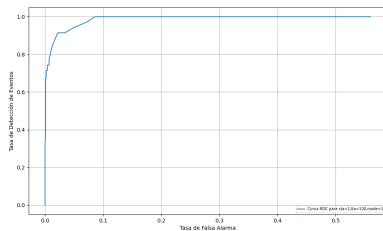


(a) STA = 2.0s, LTA = 100s

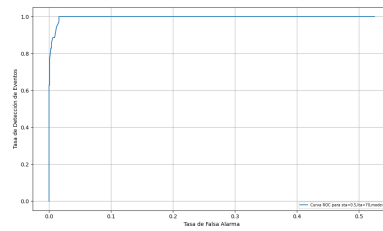


(b) STA = 0.5s, LTA = 70s

Figura 11: Curvas ROC para FG12 Grupo 1

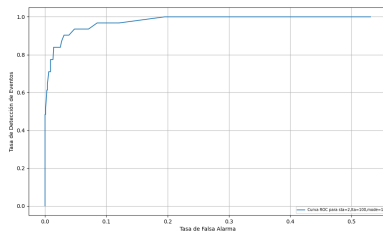


(a) STA = 2.0s, LTA = 100s

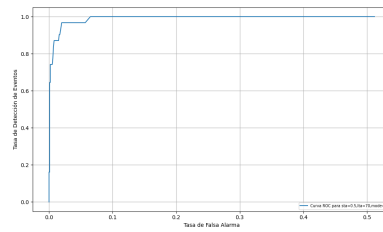


(b) STA = 0.5s, LTA = 70s

Figura 12: Curvas ROC para FG12 Grupo 2



(a) STA = 2.0s, LTA = 100s



(b) STA = 0.5s, LTA = 70s

Figura 13: Curvas ROC para FG12 Grupo 3

10.2. Análisis histogramas en FG12

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 14 Y 15.

Para la Figura 14 mostramos densidad de probabilidad de valores STA/LTA para la sección de ruido en FG12, es evidente entonces la similitud de las mismas, sin embargo, cuando analizamos 15 podemos ver que la Figura 15b muestra una respuesta mas impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor sera la respuesta de $T(x[n])$, esto nos permite distribuir de manera conveniente los valores STA/LTA pico pertenecientes a eventos de grupo 2 y 3 ya que estos se ve una mayor ocurrencia de amplitudes menores a las correspondientes a grupo 1.

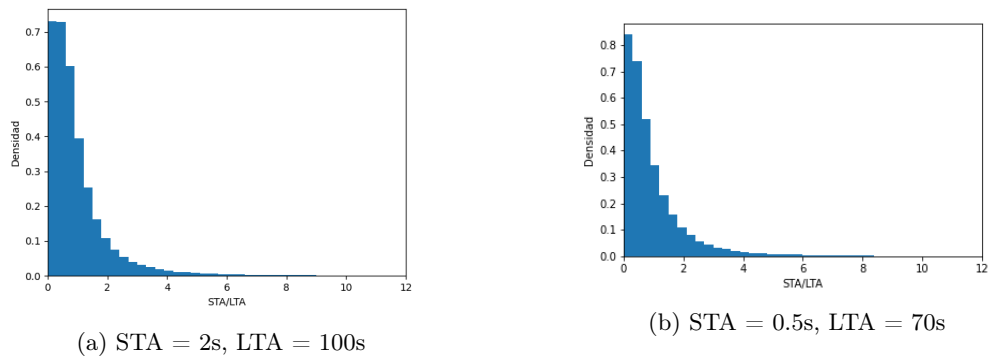


Figura 14: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en FG12.BHZ

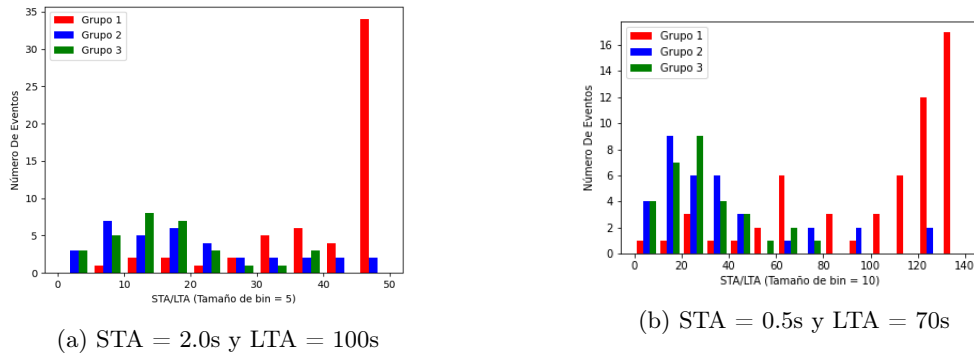


Figura 15: Histogramas de número de eventos para valores pico de función característica STA/LTA en FG12.BHZ

10.3. Análisis de probabilidades para FG12

En el Cuadro 4 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se

decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|------|
| 1 | 2.0 | 100 | 5.60 |
| 2 | 2.0 | 100 | 2.55 |
| 3 | 2.0 | 100 | 1.82 |
| 1 | 1.5 | 60 | 9.67 |
| 2 | 1.5 | 60 | 4.01 |
| 3 | 1.5 | 60 | 4.22 |

Cuadro 4: Valores mínimos para $T(S[n])$ en FG12

El Cuadro 5 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Al mismo tiempo tomamos los valores STA/LTA de la parametrización actual y acomodamos η_a a forma tal que nos permita la máxima probabilidad de detección. Es evidente entonces, que la combinación de parámetros propuesta, no solo maximiza P_D , sino también reduce P_{FA} en todos los casos.

Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 14b, se propone utilizar un nivel de umbral de desactivación del 65% del valor de η a forma de disminuir la probabilidad de registro continuo. El Cuadro 2 muestra los valores propuestos que optimizarán la captura de eventos en FG12. Este conjunto corresponde a $\{STA = 0.5s, LTA = 70s, \eta_a = 4.0, \eta_d = 2.6\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100 | 4.0 |
| 2 | 2.0 | 100 | 3.5 | 94.3 | 5.1 |
| 3 | 2.0 | 100 | 3.5 | 93.6 | 5.6 |
| 1 | 2.0 | 100 | 1.8 | 100 | 13.857 |
| 2 | 2.0 | 100 | 1.8 | 100 | 15.115 |
| 3 | 2.0 | 100 | 1.8 | 100 | 15.39 |
| 1 | 0.5 | 70 | 4.0 | 100 | 3.6 |
| 2 | 0.5 | 70 | 4.0 | 100 | 4.4 |
| 3 | 0.5 | 70 | 4.0 | 100 | 4.6 |

Cuadro 5: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para FG12

 Optimización de parámetros STA/LTA estación ESSG

En el Cuadro 6 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación ESSG fue registrada para el 80 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 la estación fue registrada en 42 % de los mismos, y a su vez fue registrada en 25 % de los eventos para el grupo 3 (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 50 | 80 % |
| 2 | 20 | 42 % |
| 3 | 12 | 25 % |

Cuadro 6: Cantidad de eventos analizados para ESSG.EHZ

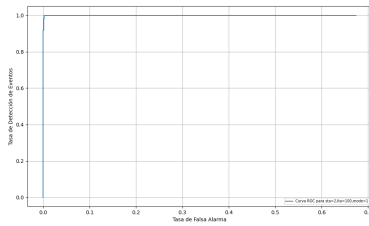
Los conjuntos de parámetros analizados para la estación ESSG corresponden a $\{\mathbf{STA} = 2\mathbf{s}, \mathbf{LTA} = 100\mathbf{s}\}$ y $\{\mathbf{STA} = 1\mathbf{s}, \mathbf{LTA} = 120\mathbf{s}\}$

11.1. Análisis de curvas ROC para ESSG

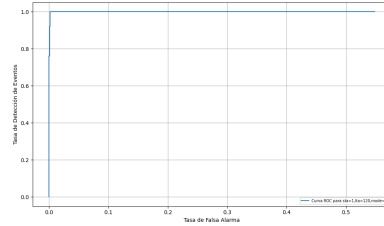
En la Figura 16a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 16b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Debido a que para ambos casos la curva se aproxima a la idealidad para los eventos de grupo 1 no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica STA/LTA solamente con esta observación.

La Figura 17 muestra la respuesta de la función característica para el grupo 2, tanto en (a) como en (b) se muestra una respuesta muy similar, sin embargo es evidente un ligera mejora en la relación entre la probabilidad de detección y la probabilidad de falsa alarma en la Figura 17b.

Como definimos en el Cuadro 1 el grupo 3 nos servirá para evidenciar la mejora entre falsos disparos y eventos declarados, por lo cual al analizar la Figura 18 podemos ver que (b) muestra una mejora evidente en la relación entre la probabilidad de detección y la probabilidad de falsa alarma.

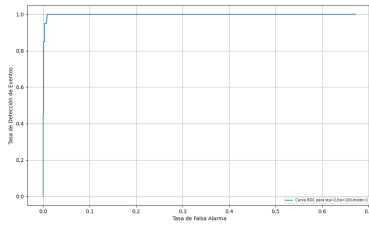


(a) STA = 2s, LTA = 100s

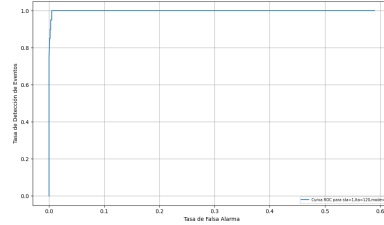


(b) STA = 1s, LTA = 120s

Figura 16: Curvas ROC para ESSG Grupo 1

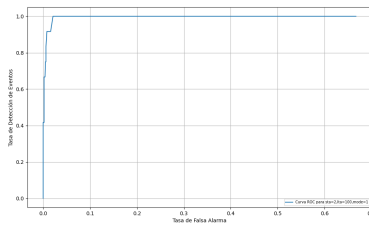


(a) STA = 2s, LTA = 100s

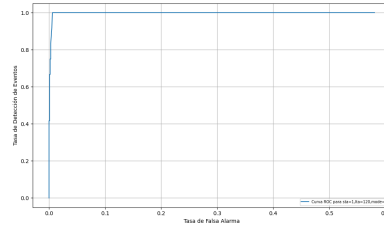


(b) STA = 1s, LTA = 120s

Figura 17: Curvas ROC para ESSG Grupo 2



(a) STA = 2s, LTA = 100s



(b) STA = 1s, LTA = 120s

Figura 18: Curvas ROC para ESSG Grupo 3

11.2. Análisis de histogramas en ESSG

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 19 Y 20.

Para la Figura 19a mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en ESSG para ambos conjuntos de parámetros, y es posible evidenciar la alta similitud entre ambas, sin embargo, es posible ver que obtenemos valores un poco mas altos que al utilizar el conjunto de parámetros por defecto. En la Figura 20 vemos que la respuesta para la combinación {STA = 1s y LTA = 120s} (Figura 20b) resulta ser más impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor sera la respuesta de $T(x[n])$, esto nos permite distribuir de manera conveniente los valores STA/LTA pico, atípicos en eventos de los grupos 2 y 3.

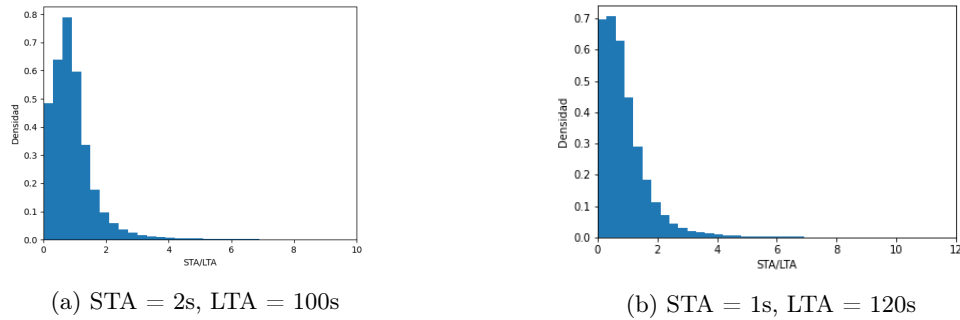


Figura 19: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en ESSG.BHZ

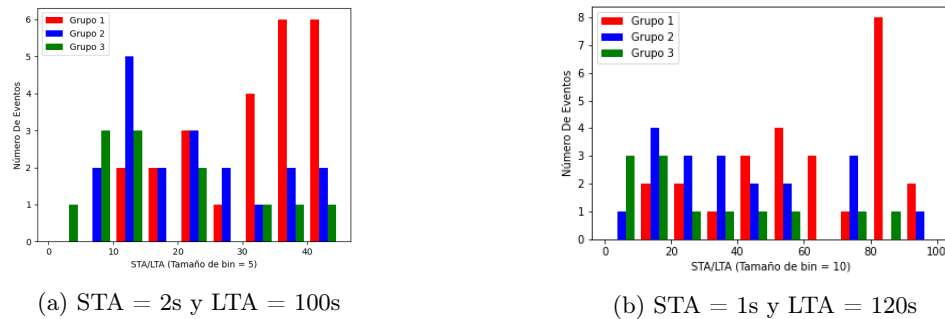


Figura 20: Histogramas de número de eventos para valores pico de función característica STA/LTA en ESSG.BHZ

11.3. Análisis de probabilidades para ESSG

En el Cuadro 7 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se

decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 12.26 |
| 2 | 2.0 | 100 | 6.10 |
| 3 | 2.0 | 100 | 4.71 |
| 1 | 1.0 | 120 | 16.23 |
| 2 | 1.0 | 120 | 9.30 |
| 3 | 1.0 | 120 | 7.43 |

Cuadro 7: Valores mínimos para $T(S[n])$ en ESSG

El Cuadro 8 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto STA = 1s, LTA = 120s, $\eta_a = 7.3$, se toma la decisión de utilizar aquel conjunto capaz de minimizar la probabilidad de falsa alarma y maximizar la probabilidad de detección para los tres grupos de eventos. (Ver Cuadro 1).

Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 19a, se propone utilizar un nivel de umbral de desactivación (η_d) del 55 % del valor de η_A debido a que como se discutió anteriormente esta combinación tiene una respuesta más impulsiva y a su vez el intervalo $[\eta_a, \eta_d]$ contiene una mínima densidad de probabilidad de valores STA/LTA para la sección de ruido.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en ESSG, este conjunto corresponde a $\{STA = 1s, LTA = 120s, \eta_a = 7.3, \eta_d = 4.0\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100.0 | 2.0 |
| 2 | 2.0 | 100 | 3.5 | 100.0 | 2.3 |
| 3 | 2.0 | 100 | 3.5 | 100.0 | 2.1 |
| 1 | 2.0 | 100 | 4.6 | 100.0 | 1.2 |
| 2 | 2.0 | 100 | 4.6 | 100.0 | 1.3 |
| 3 | 2.0 | 100 | 4.6 | 100.0 | 1.3 |
| 1 | 1.0 | 120 | 7.3 | 100.0 | 0.6 |
| 2 | 1.0 | 120 | 7.3 | 100.0 | 0.7 |
| 3 | 1.0 | 120 | 7.6 | 100.0 | 0.7 |

Cuadro 8: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para ESSG

 Optimización de parámetros STA/LTA estación HUEH

En el Cuadro 9 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 2.5 registrados para el año 2019 la estación HUEH fue registrada para el 80 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 (Cuadro 1) la estación fue registrada en 4.1 % de los mismos, y a su no estuvo presente en ninguno de los eventos de grupo 3 de los utilizados en la muestra para este estudio. (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 40 | 65 % |
| 2 | 2 | 4.1 % |
| 3 | 0 | 0 % |

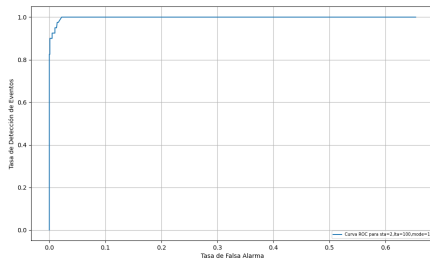
Cuadro 9: Cantidad de eventos analizados para HUEH.HHZ

Los conjuntos de parámetros analizados para la estación HUEH corresponden a $\{\text{STA} = 2\text{s}, \text{LTA} = 100\text{s}\}$ y $\{\text{STA} = 1\text{s}, \text{LTA} = 70\text{s}\}$

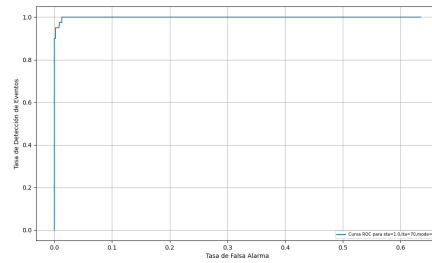
12.1. Análisis de curvas ROC para HUEH

En la Figura 21a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 21b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. En este caso es evidente una mejora representativa en la relación de la activación de eventos verdaderos versus aquellos de falsa alarma según el comportamiento representado en la Figura 21b.

La Figura 22 muestra la respuesta de la función característica para el grupo 2 para ambas combinaciones. En este caso podemos inferir en que la muestra disminuida de eventos de grupo 2 puede ser significativa en los resultados obtenidos, sin embargo, dada la representatividad de eventos de grupo 1 y los resultados demostrados para este, además analizando la mejora en la respuesta mostrada por la Figura 22b podemos decir que esta combinación de parámetros mejora la misma relación para el grupo 2, tal y como lo hace para el grupo 1.

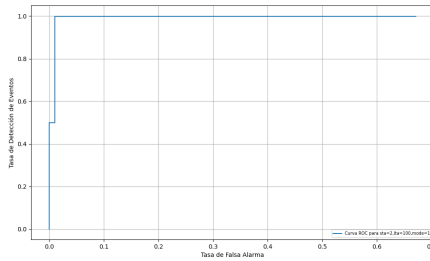


(a) STA = 2s, LTA = 100s

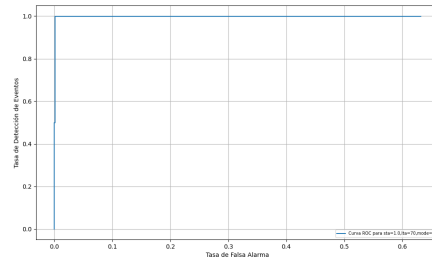


(b) STA = 1s, LTA = 70s

Figura 21: Curvas ROC para HUEH Grupo 1



(a) STA = 2s, LTA = 100s



(b) STA = 1s, LTA = 70s

Figura 22: Curvas ROC para HUEH Grupo 2

12.2. Análisis de histogramas para HUEH

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 23 Y 24.

Para la Figura 23 mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en ESSG para ambos conjuntos de parámetros, y es posible evidenciar la alta similitud entre ambas; sin embargo, en la Figura 24 vemos que la respuesta para la combinación $\{STA = 1s \text{ y } LTA = 70s\}$ (Figura 24b) resulta ser más impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor será la respuesta de $T(x[n])$, esto nos permite distribuir de manera conveniente los valores STA/LTA pico, atípicos en eventos de grupo 1.

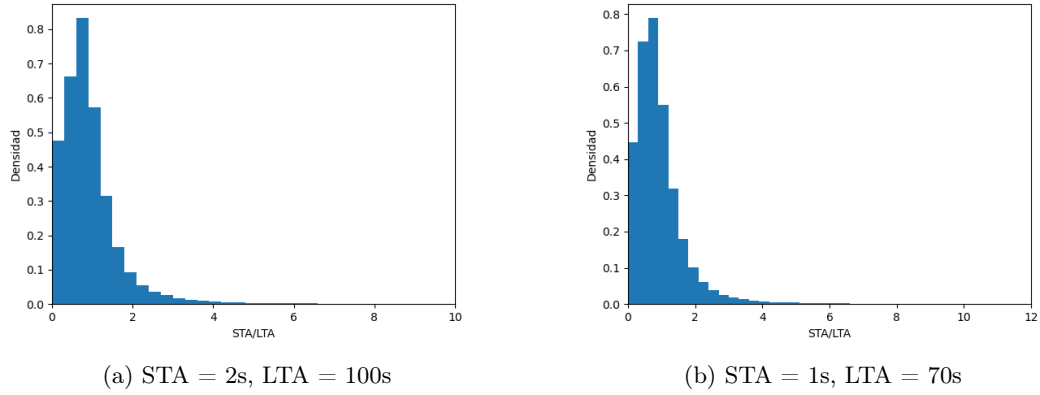


Figura 23: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en HUEH.HHZ

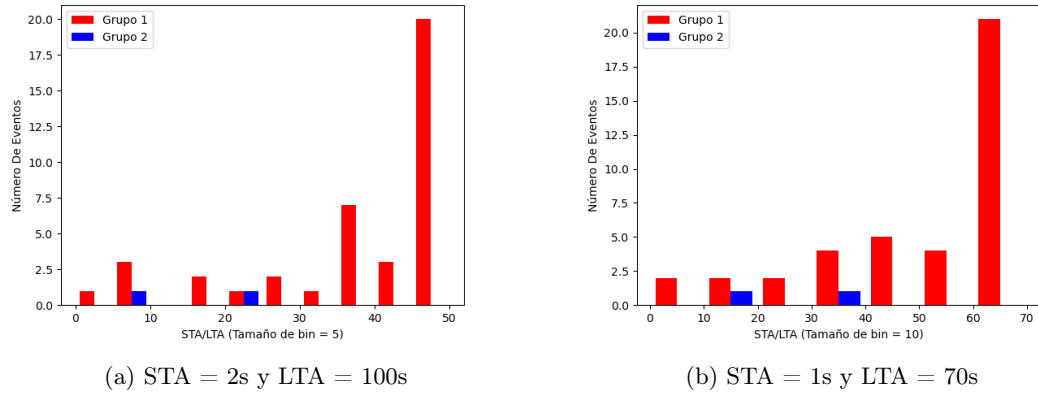


Figura 24: Histogramas de número de eventos para valores pico de función característica STA/LTA en HUEH.HHZ

12.3. Análisis de probabilidades para HUEH

En el Cuadro 10 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se decide tomar un valor $\eta - a$ cercano y menor al menor de los mínimos entre los conjuntos de grupo 1 y grupo 2 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 3.76 |
| 2 | 2.0 | 100 | 7.03 |
| 1 | 1.0 | 70 | 4.33 |
| 2 | 1.0 | 70 | 18.77 |

Cuadro 10: Valores mínimos para $T(S[n])$ en HUEH

El Cuadro 11 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto $STA = 10$, $LTA = 60$ y $\eta = 3.4$ para ambos grupos, como se mencionó con anterioridad para estos conjuntos de parámetros es posible la detección absoluta de ambos grupos.

Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 23b, se propone utilizar un nivel de umbral de desactivación del 65% del valor de η_a debido a que el conjunto de parámetros seleccionado como el que presenta la mejor respuesta es suave y de esta manera se puede disminuir la probabilidad de desactivación temprana o bien la posibilidad de registro continuo.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en HUEH, este conjunto corresponde a $\{STA = 1s, LTA = 70s, \eta_a = 4.2, \eta_d = 2.7\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100.0 | 2.1 |
| 2 | 2.0 | 100 | 3.5 | 100.0 | 2.2 |
| 1 | 2.0 | 100 | 3.7 | 100.0 | 1.9 |
| 2 | 2.0 | 100 | 3.7 | 100.0 | 2.0 |
| 1 | 1.0 | 70 | 4.2 | 100.0 | 1.4 |
| 2 | 1.0 | 70 | 4.2 | 100.0 | 1.4 |

Cuadro 11: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA , LTA y η_a para HUEH

 Optimización de parámetros STA/LTA estación STG8

En el Cuadro 12 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la estación STG8. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación STG8 fue registrada para el 71 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 (Cuadro 1) la estación fue registrada en 60 % de los mismos, y a su vez fue registrada en 19 % de los eventos para el grupo 3 (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 44 | 71 % |
| 2 | 29 | 60 % |
| 3 | 9 | 19 % |

Cuadro 12: Cantidad de eventos analizados para STG8.BHZ

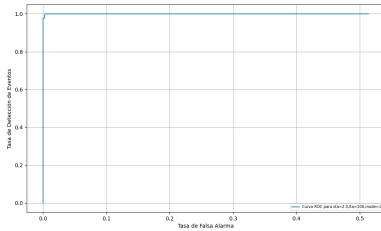
Los conjuntos de parámetros analizados para la estación ESSG corresponden a **{STA = 2s, LTA = 100s}** y **{STA = 1.5s, LTA = 120s}**

13.1. Análisis de curvas ROC para STG8

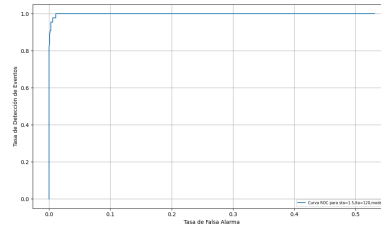
En la Figura 25a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 25b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Es posible evidenciar en esta una ligera mejora en la relación entre la probabilidad de detección contra la probabilidad de falsa alarma para eventos de grupo 1, ya que esta se aproxima de forma escalonar más rápido al punto ideal.

La Figura 26 muestra la respuesta de la función característica para el grupo 2, Es posible evidenciar en esta una mejora considerable en la relación entre la probabilidad de detección contra la probabilidad de falsa alarma para eventos de grupo 2, ya que en esta podemos decir por observación como esta se aproxima de forma considerablemente mejor al punto ideal.

Como definimos en el Cuadro 1 el grupo 3 nos servirá para evidenciar la mejora entre falsos disparos y eventos declarados, por lo cual al analizar la Figura 27 podemos ver que (b) muestra una mejora evidente en la relación entre la probabilidad de detección y la probabilidad de falsa alarma para este grupo.

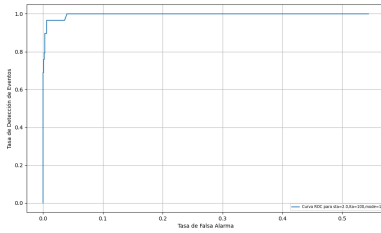


(a) STA = 2.0, LTA = 100

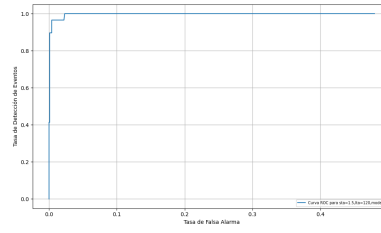


(b) STA = 1.5, LTA = 120

Figura 25: Curvas ROC para STG8 Grupo 1

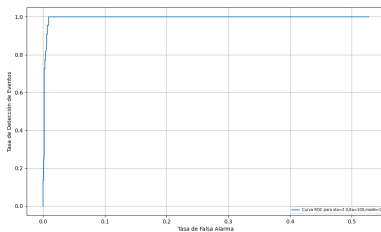


(a) STA = 2s, LTA = 100s

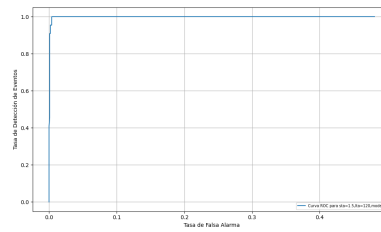


(b) STA = 1.5s, LTA = 120s

Figura 26: Curvas ROC para STG8 Grupo 2



(a) STA = 2s, LTA = 100s



(b) STA = 1.5s, LTA = 120s

Figura 27: Curvas ROC para STG8 Grupo 3

13.2. Análisis de histogramas para STG8

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 28 Y 29.

Para la Figura 28 mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en ESSG para ambos conjuntos de parámetros, y es posible evidenciar la alta similitud entre ambas, sin embargo, vemos que la respuesta para la combinación {STA = 1.5s y LTA = 120s} (29b resulta ser más impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor sera la respuesta de $T(x[n])$), esto nos permite distribuir de manera conveniente los valores STA/LTA pico, atípicos en eventos de los grupo 2.

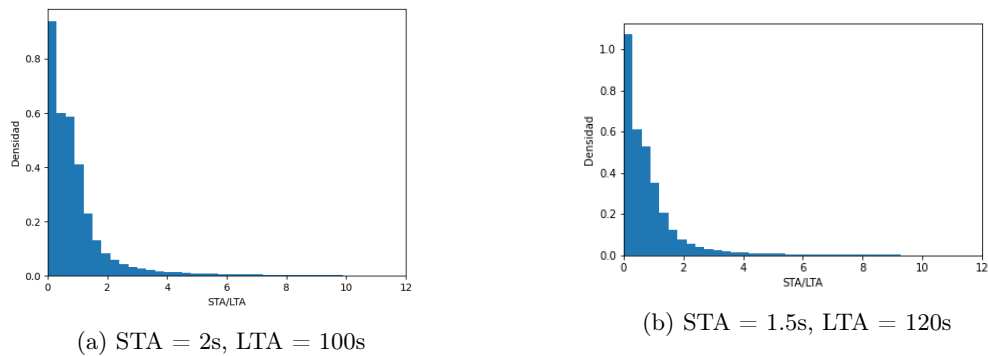


Figura 28: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en STG8

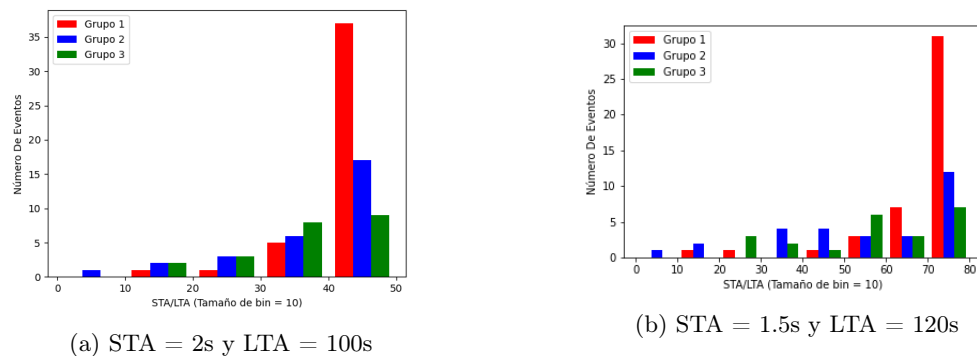


Figura 29: Histogramas de número de eventos para valores pico de función característica STA/LTA en STG8

13.3. Análisis de probabilidades para STG8

En el Cuadro 13 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos,

se decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 12.19 |
| 2 | 2.0 | 100 | 4.58 |
| 3 | 2.0 | 100 | 14.73 |
| 1 | 1.5 | 120 | 13.17 |
| 2 | 1.5 | 120 | 6.33 |
| 3 | 1.5 | 120 | 20.34 |

Cuadro 13: Valores mínimos para $T(S[n])$ en STG8

El Cuadro 14 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto STA = 1.5s, LTA = 120 s y $\eta_a = 6.2$.

Con la utilización de estos parámetros se cumplen las dos condiciones, 1 disminuir la probabilidad de falsa alarma, proveer detección de todos los eventos que pertenecen a los grupos 1, 2 y 3.

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido en la Figura 28b, se propone utilizar un nivel de umbral de desactivación próximo al 65% del valor de η_d ya que esto proveerá un mejora significativa en la detección de final de evento y disminuye la probabilidad de registro continuo.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en STG8. este conjunto corresponde a $\{STA = 1.5s, LTA = 120s, \eta_a = 6.2, \eta_d = 4.0\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100 | 5.4 |
| 2 | 2.0 | 100 | 3.5 | 100 | 6.3 |
| 3 | 2.0 | 100 | 3.5 | 100 | 6.6 |
| 1 | 2.0 | 100 | 4.5 | 100 | 3.8 |
| 2 | 2.0 | 100 | 4.5 | 100 | 4.5 |
| 3 | 2.0 | 100 | 4.5 | 100 | 4.5 |
| 1 | 1.5 | 120 | 6.2 | 100 | 2.6 |
| 2 | 1.5 | 120 | 6.2 | 100 | 3.1 |
| 3 | 1.5 | 120 | 6.2 | 100 | 3.1 |

Cuadro 14: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para STG8

 Optimización de parámetros STA/LTA estación SMCA

En el Cuadro 15 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación SMCA fue registrada para el 65 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 (Cuadro 1) la estación fue registrada en 29 % de los mismos, y a su vez fue registrada en 21 % de los eventos para el grupo 3 (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 40 | 65 % |
| 2 | 14 | 29 % |
| 3 | 10 | 21 % |

Cuadro 15: Cantidad de eventos analizados para SMCA.BHZ

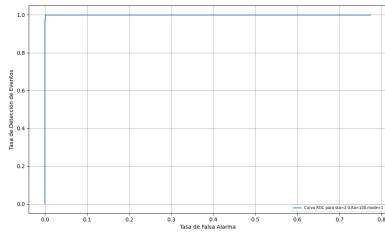
Los conjuntos de parámetros analizados para la estación ESSG corresponden a $\{\text{STA} = 2\text{s}, \text{LTA} = 100\text{s}\}$ y $\{\text{STA} = 1\text{s}, \text{LTA} = 70\text{s}\}$

14.1. Análisis de curvas ROC para SMCA

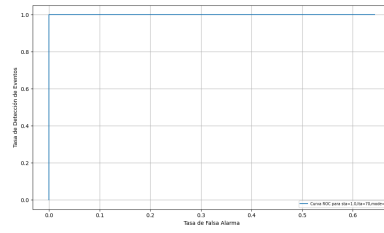
En la Figura 30a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 30b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Debido a que para ambos casos la curva se aproxima a la idealidad para los eventos de grupo 1 no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica solamente con esta observación.

La Figura 31 muestra la respuesta de la función característica para el grupo 2, tanto en (a) como en (b) se muestra una respuesta muy similar, sin embargo hay un mayor decaimiento en 31a lo que nos indica que habrá una mayor pérdida de eventos pertenecientes al grupo 2 con esta combinación de parámetros,

Como definimos en el Cuadro 1 el grupo 3 nos servirá para evidenciar la mejora entre falsos disparos y eventos declarados, por lo cual al analizar la Figura 32 es evidente que para ambos casos la curva se aproxima a la idealidad para los eventos de grupo 3, por lo tanto no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica solamente con esta observación.

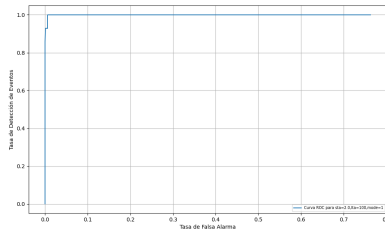


(a) STA = 2s, LTA = 100s

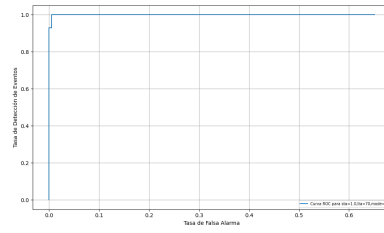


(b) STA = 1s, LTA = 70s

Figura 30: Curvas ROC para SMCA Grupo 1

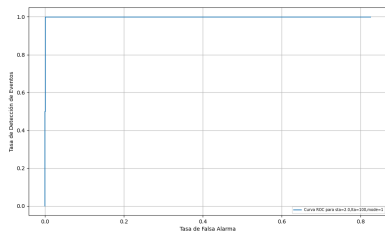


(a) STA = 2s, LTA = 100s

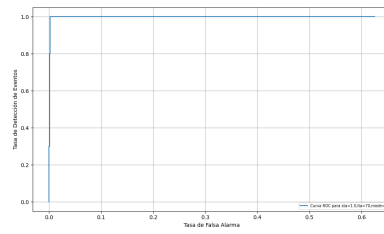


(b) STA = 1s, LTA = 70s

Figura 31: Curvas ROC para SMCA Grupo 2



(a) STA = 2s, LTA = 100s



(b) STA = 1s, LTA = 70s

Figura 32: Curvas ROC para SMCA Grupo 3

14.2. Análisis de histogramas para SMCA

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 33 Y 34.

Para la Figura 33a mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en SMCA, para ambos conjuntos de parámetros, y es posible evidenciar la alta similitud entre ambas; sin embargo, en la Figura 34 vemos que la respuesta para la combinación $\{STA = 1s \text{ y } LTA = 70s\}$ (Figura 34b) resulta ser un tanto mas impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor sera la respuesta de $T(x[n])$, esto nos permite distribuir de manera conveniente los valores STA/LTA pico, atípicos en eventos de los grupos 2 y 3.

También se hace necesario resaltar que en comparación con otras estaciones esta presenta una distribución de ruido menor, lo cual la hace más precisa y provee una mejor respuesta STA/LTA debido a esta misma característica.

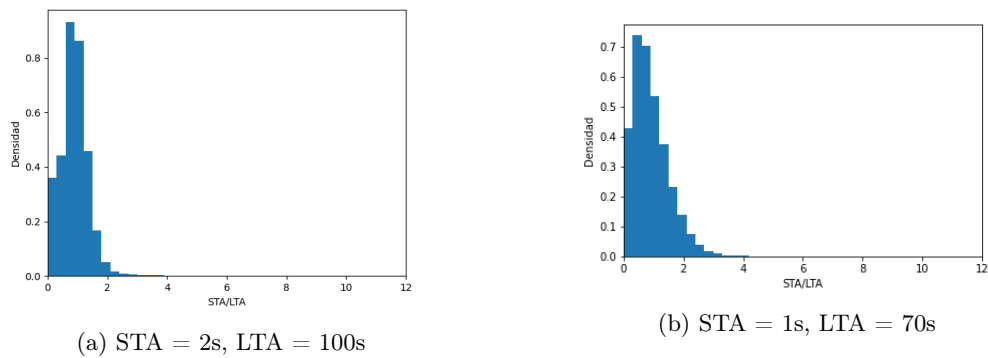


Figura 33: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en SMCA

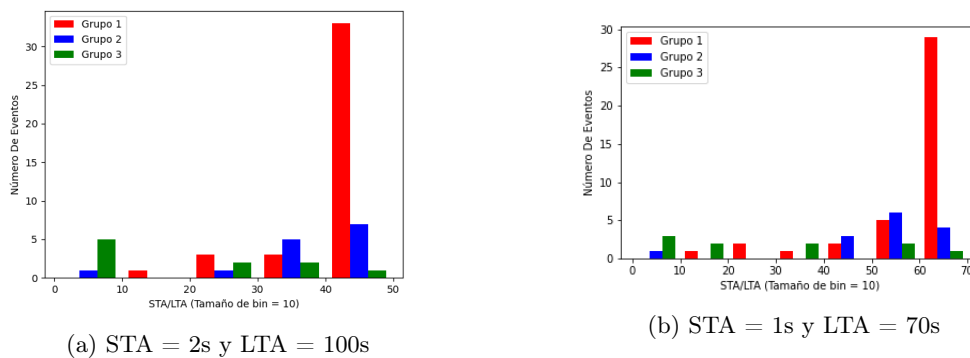


Figura 34: Histogramas de número de eventos para valores pico de función característica STA/LTA en SMCA

14.3. Análisis de probabilidades para SMCA

En el Cuadro 16 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 10.05 |
| 2 | 2.0 | 100 | 5.71 |
| 3 | 2.0 | 100 | 4.56 |
| 1 | 1.0 | 70 | 5.01 |
| 2 | 1.0 | 70 | 4.28 |
| 3 | 1.0 | 70 | 3.97 |

Cuadro 16: Valores mínimos para $T(S[n])$ en SMCA

El Cuadro 17 muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto STA = 1, LTA = 70 y $\eta_a = 5.6$ para los grupos 1, 2 y 3. Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 33b, se propone utilizar un nivel de umbral de desactivación del 65% del valor de η debido a que como se discutió anteriormente esta combinación tiene una respuesta más impulsiva.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en SMCA, este conjunto corresponde a $\{STA = 1s, LTA = 70s, \eta_a = 5.6, \eta_d = 4.0\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100.0 | 0.69 |
| 2 | 2.0 | 100 | 3.5 | 100.0 | 0.56 |
| 3 | 2.0 | 100 | 3.5 | 100.0 | 0.24 |
| 1 | 2.0 | 100 | 4.5 | 100.0 | 0.50 |
| 2 | 2.0 | 100 | 4.5 | 100.0 | 0.40 |
| 3 | 2.0 | 100 | 4.5 | 100.0 | 0.16 |
| 1 | 1.0 | 70 | 5.6 | 100.0 | 0.32 |
| 2 | 1.0 | 70 | 5.6 | 100.0 | 0.27 |
| 3 | 1.0 | 70 | 5.6 | 100.0 | 0.11 |

Cuadro 17: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para SMCA

 Optimización de parámetros STA/LTA estación APG

En el Cuadro 18 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación APG fue registrada para el 38 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 (Cuadro 1) la estación fue registrada en 17 % de los mismos, y a su vez fue registrada en 1 % de los eventos para el grupo 3 (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 24 | 38 % |
| 2 | 8 | 17 % |
| 3 | 4 | 1 % |

Cuadro 18: Cantidad de eventos analizados para APG.BHZ

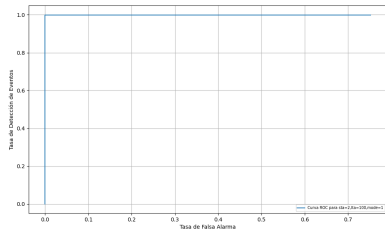
Los conjuntos de parámetros analizados para la estación APG corresponden a $\{\text{STA} = 2\text{s}, \text{LTA} = 100\text{s}\}$ y $\{\text{STA} = 5\text{s}, \text{LTA} = 100\text{s}\}$

15.1. Análisis de curvas ROC para APG

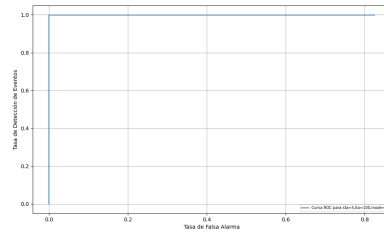
En la Figura 35a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 35b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Debido a que para ambos casos la curva se aproxima a la idealidad para los eventos de grupo 1 no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica solamente con esta observación.

La Figura 36 muestra la respuesta de la función característica para el grupo 2, tanto en (a) como en (b) la curva se aproxima a la idealidad para los eventos de grupo 2 no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica solamente con esta observación.,

Como definimos en el Cuadro 1 el grupo 3 nos servira para evidenciar la mejora entre falsos disparos y eventos declarados, por lo cual al analizar la Figura 37 podemos ver que (b) muestra se aleja de forma mínima del punto ideal, mientras que (a) se mantiene como el caso de los grupo 1 y 2, sin embargo, esta prueba resulta ser no concluyente en los tres casos.

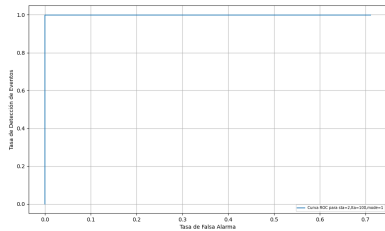


(a) STA = 2s, LTA = 100s

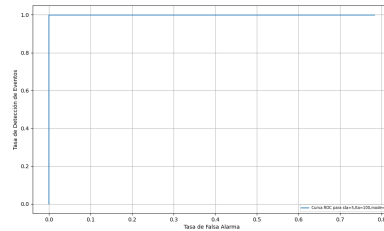


(b) STA = 5s, LTA = 100s

Figura 35: Curvas ROC para APG Grupo 1

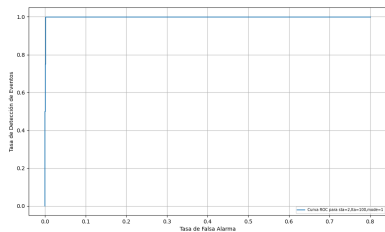


(a) STA = 2s, LTA = 100s

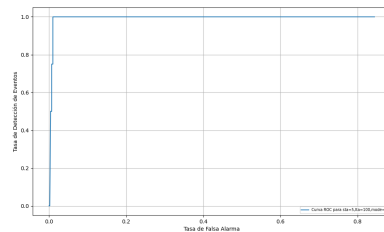


(b) STA = 5s, LTA = 100s

Figura 36: Curvas ROC para APG Grupo 2



(a) STA = 2s, LTA = 100s



(b) STA = 5s, LTA = 100s

Figura 37: Curvas ROC para APG Grupo 3

15.2. Análisis de histogramas para APG

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 38 Y 39.

Para la Figura 38 mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en APG. Para este caso es posible evidenciar en la Figura 39 los valores de STA = 2s y LTA = 100s resultan en una respuesta mas impulsiva, sin embargo, aunque la respuesta es menos impulsiva para el conjunto de parámetros STA = 5s y LTA = 100s aun siendo este el caso, este conjunto tambien resulta en valores altos, esto puede ser interpretado como una respuesta impulsiva en APG dadas las propiedades geológicas de su ubicación.

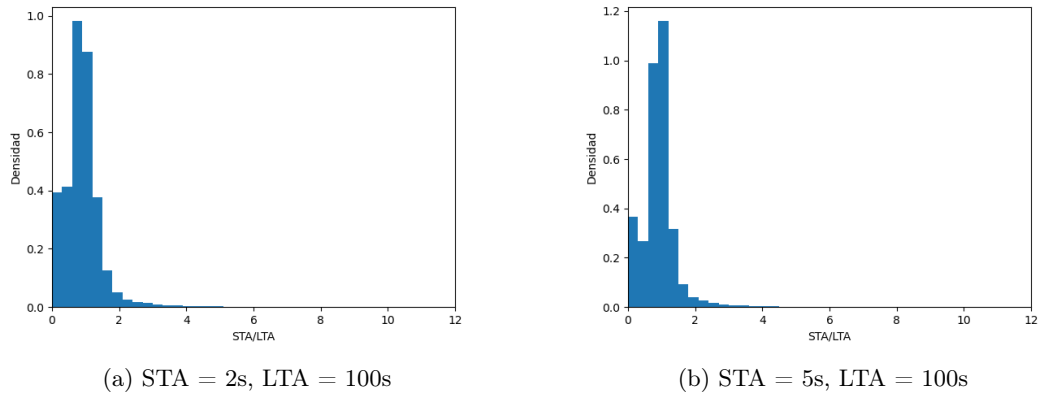


Figura 38: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en APG

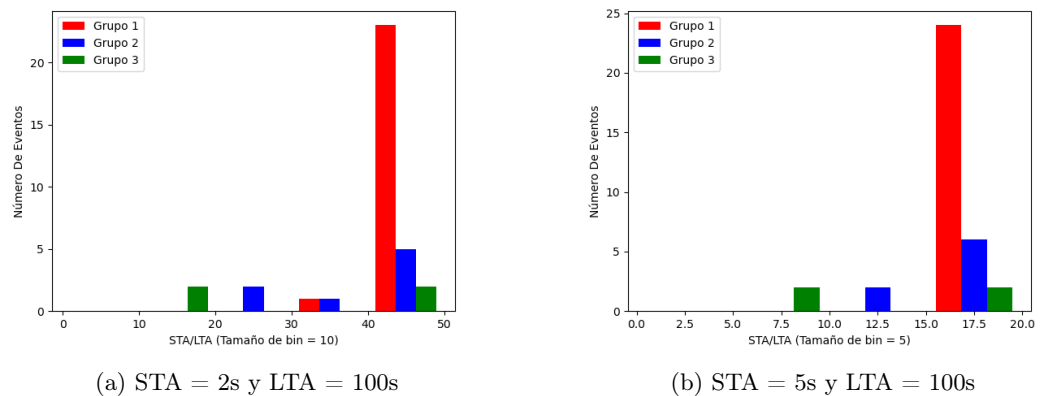


Figura 39: Histogramas de número de eventos para valores pico de función característica STA/LTA en estación APG

15.3. Análisis de probabilidades para APG

En el Cuadro 19 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

Los mínimos posibles se presentan en el Cuadro 19, como podemos ver en el Cuadro 19 los valores atípicos que hemos encontrado en análisis anteriores no están presentes en este caso, por lo cual consideramos las características propias de la estación al optimizar la parametrización de la misma. Si prevemos la probabilidad de ocurrencia de valores atípicos y consideramos $\eta_a \approx 11.2$ estaríamos probablemente discriminando eventos. Por lo cual, al considerar que el conjunto de parámetros STA = 5s y LTA = 100s convenientemente reduce la impulsividad de la estación podemos prevenir la discriminación de los mismos. Por lo cual se considera el valor de $\eta_a \approx 11.2$ como no conveniente.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 36.58 |
| 2 | 2.0 | 100 | 21.41 |
| 3 | 2.0 | 100 | 11.24 |
| 1 | 5.0 | 100 | 17.07 |
| 2 | 5.0 | 100 | 12.32 |
| 3 | 5.0 | 100 | 6.61 |

Cuadro 19: Valores mínimos para $T(S[n])$ en APG

El Cuadro 20 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto STA = 5, LTA = 100 y $\eta = 12$. también la probabilidad de detección para eventos de los grupos 1, 2 y 3 es de 100 %.

Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la distribución de ruido mostrada en la Figura 38b, se propone utilizar un nivel de umbral de desactivación del 55 % del valor de η_d a manera de maximizar el tiempo de registro de eventos y debido que se considera que esta estación no se ve fuertemente afectada por el ruido como sucede en otros casos.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en APG, este conjunto corresponde a $\{STA = 5s, LTA = 100s, \eta_a = 6.6, \eta_d = 3.7\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100.0 | 1.19 |
| 2 | 2.0 | 100 | 3.5 | 100.0 | 1.52 |
| 3 | 2.0 | 100 | 3.5 | 100.0 | 0.79 |
| 1 | 5.0 | 100 | 6.6 | 100.0 | 0.25 |
| 2 | 5.0 | 100 | 6.6 | 100.0 | 0.30 |
| 3 | 5.0 | 100 | 6.6 | 100.0 | 0.50 |

Cuadro 20: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para APG

 Optimización de parámetros STA/LTA estación RETA

En el Cuadro 21 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación RETA fue registrada para el 92 % de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 (Cuadro 1) la estación fue registrada en 38 % de los mismos, y a su vez fue registrada en 19 % de los eventos para el grupo 3 (Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 57 | 92 % |
| 2 | 18 | 38 % |
| 3 | 9 | 19 % |

Cuadro 21: Cantidad de eventos analizados para RETA.BHZ

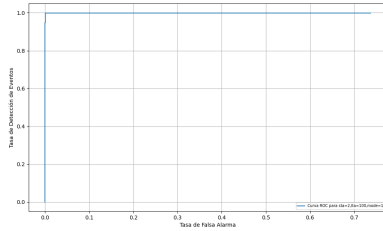
Los conjuntos de parámetros analizados para la estación RETA corresponden a **{STA = 2s, LTA = 100s}** y **{STA = 1.5s, LTA = 60s}**

16.1. Análisis de curvas ROC para RETA

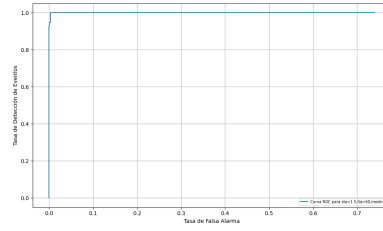
En la Figura 40a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 40b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Debido a que para ambos casos la curva se aproxima a la idealidad para los eventos de grupo 1 no podemos inferir en si la nueva combinación de parámetros mejora la respuesta de la función característica solamente con esta observación.

La Figura 41 muestra la respuesta de la función característica para el grupo 2, como podemos ver 41a tiene un decaimiento mientras se aproxima a la idealidad, contrario a la Figura 41b que muestra un curva ideal.

Como definimos en el Cuadro 1 el grupo 3 nos servira para evidenciar la mejora entre falsos disparos y eventos declarados, por lo cual al analizar la Figura 42 podemos ver que ambos conjuntos de parámetros proveen una curva ROC ideal para eventos de grupo 3.

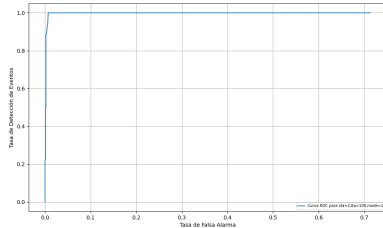


(a) STA = 2s, LTA = 100s

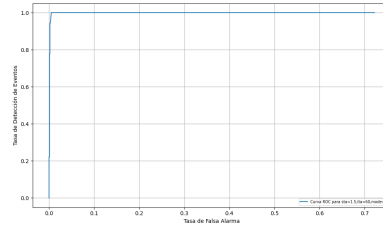


(b) STA = 1.5s, LTA = 60s

Figura 40: Curvas ROC para RETA Grupo 1

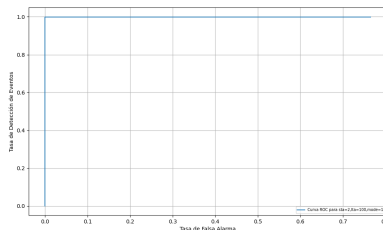


(a) STA = 2s, LTA = 100s

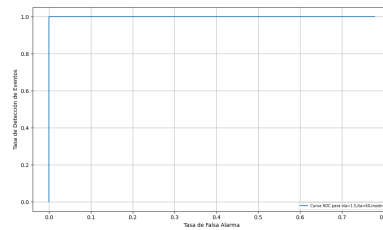


(b) STA = 1.5s, LTA = 60s

Figura 41: Curvas ROC para RETA Grupo 2



(a) STA = 2s, LTA = 100s



(b) STA = 1.5s, LTA = 60s

Figura 42: Curvas ROC para RETA Grupo 3

16.2. Análisis de histogramas para RETA

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia las Figuras 43 Y 44.

Para la Figura 43 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad, de esta es posible evidenciar la alta similitud entre ambas. Es importante resaltar que dado las características de esta estación lo que se buscó fue suavizar la respuesta impulsiva de la misma, debido a que valores muy altos para η_a suelen no ser la mejor opción en la parametrización de algoritmos STA/LTA, además, al suavizar la respuesta logramos disminuir, para este caso, los efectos del ruido. Es posible evidenciar este efecto en la Figura 44b, ya que hemos reorganizado convenientemente la respuesta de la función característica STA/LTA.

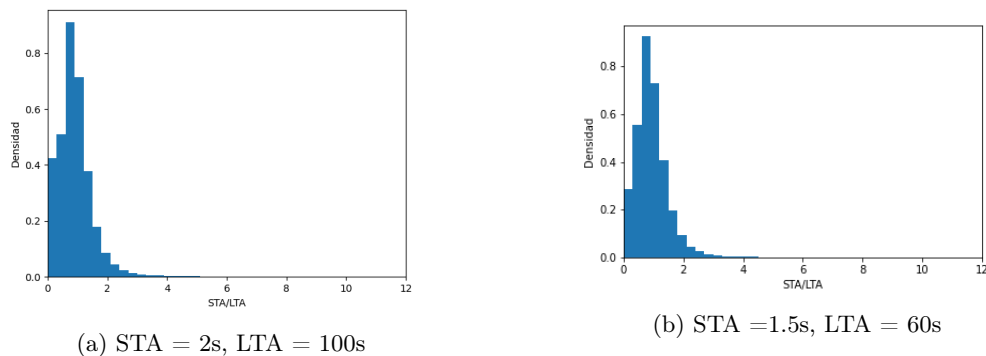


Figura 43: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en RETA

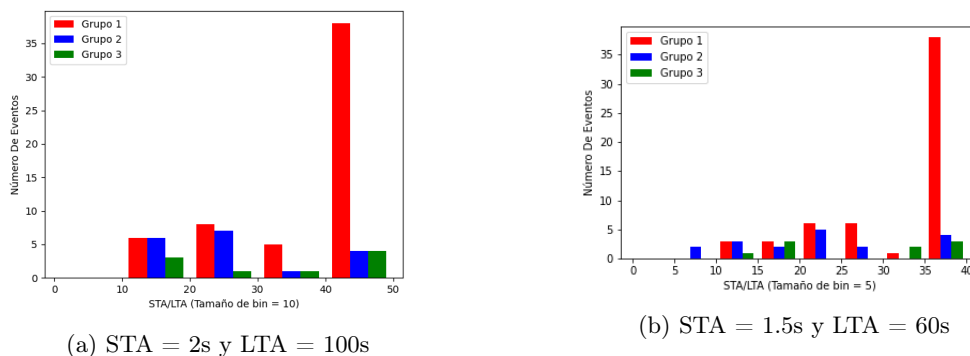


Figura 44: Histogramas de número de Eventos para valores pico de función característica STA/LTA en RETA

16.3. Análisis de probabilidades para RETA

En el Cuadro 22 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se decide tomar un valor η_a cercano y menor al menor de los mínimos entre los conjuntos de grupo 1, grupo 2 y grupo 3 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 11.68 |
| 2 | 2.0 | 100 | 10.07 |
| 3 | 2.0 | 100 | 14.62 |
| 1 | 1.5 | 60 | 10.04 |
| 2 | 1.5 | 60 | 8.11 |
| 3 | 1.5 | 60 | 11.82 |

Cuadro 22: Valores mínimos para $T(S[n])$ en RETA

El Cuadro 23 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores utilizados en la actualidad. Es posible evidenciar entonces que ambos conjuntos muestra un punto óptimo con sus respectivos valores para η_a , sin embargo se valora mejor el conjunto propuesto dado que este permite la utilización de un valor, entre lo común para η_a . Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 43b, se propone utilizar un nivel de umbral de desactivación del 65% del valor de η_a como se discutió anteriormente esta combinación tiene una respuesta más impulsiva.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en RETA, este conjunto corresponde a $\{STA = 1.5s, LTA = 70s, \eta_a = 8.0, \eta_d = 5.2\}$.

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 100 | 1.1 |
| 2 | 2.0 | 100 | 3.5 | 100 | 1.7 |
| 3 | 2.0 | 100 | 3.5 | 100 | 1.1 |
| 1 | 2.0 | 100 | 10.0 | 100 | 0.3 |
| 2 | 2.0 | 100 | 10.0 | 100 | 0.4 |
| 3 | 2.0 | 100 | 10.0 | 100 | 0.1 |
| 1 | 1.5 | 60 | 8.0 | 100 | 0.2 |
| 2 | 1.5 | 60 | 8.0 | 100 | 0.4 |
| 3 | 1.5 | 60 | 8.0 | 100 | 0.1 |

Cuadro 23: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA, LTA y η_a para RETA

 Optimización de parámetros STA/LTA estación CHIE

En el Cuadro 24 se puede observar el tamaño de la muestra por grupo utilizada para la optimización de la misma. Así mismo se detalla la representatividad dentro del tamaño total de la muestra. Esta nos indica que de los 62 eventos de subducción local con magnitud mayor a 4.5 registrados para el año 2019 la estación CHIE fue registrada para el 92% de los mismos. En su contra parte de los 48 eventos utilizados como muestra del grupo 2 la estación fue registrada en 10% de los mismos, y a su vez no hubo ocurrencias del grupo 3. (Ver Cuadro 1).

| Grupo | Muestra | Representatividad |
|-------|---------|-------------------|
| 1 | 37 | 60 % |
| 2 | 5 | 10 % |
| 3 | 0 | 0 % |

Cuadro 24: Cantidad de eventos analizados para CHIE.EHZ

Los conjuntos de parámetros analizados para la estación CHIE corresponden a $\{\mathbf{STA} = 2\mathbf{s}, \mathbf{LTA} = 100\mathbf{s}\}$ y $\{\mathbf{STA} = 1\mathbf{s}, \mathbf{LTA} = 120\mathbf{s}\}$

17.1. Análisis de curvas ROC para CHIE

En la Figura 45a se muestra la respuesta de la función característica STA/LTA dados los parámetros por defecto actualmente utilizados, mientras que en la Figura 45b se muestra la respuesta de acuerdo a la mejor combinación de parámetros de los utilizados en las diferentes simulaciones. Para ambas es posible evidenciar que es imposible una clasificación absoluta, sin embargo la Figura 45b es más escalonada y se aproxima de forma más consistente al punto (1, 1) (Ver Figura 10) lo cual denota una mejor relación de probabilidad de detección

versus probabilidad de falsa alarma.

La Figura 46 muestra que para ambas combinaciones de parámetros STA, LTA existe un valor de η_a capaz de proveer una completa clasificación entre ruido y evento de grupo 2.

Dado que para la muestra utilizada no hay ocurrencia de eventos de grupo 3 que hayan efectuado activación para la estación CHIE, se realiza el siguiente análisis considerando eventos de grupo 1 y 2 únicamente.

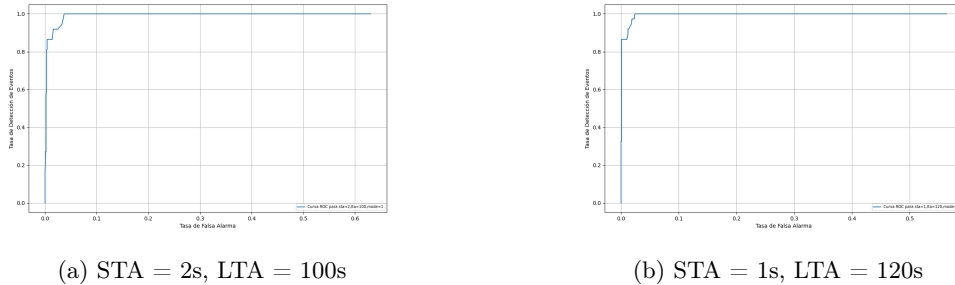


Figura 45: Curvas ROC para CHIE Grupo 1

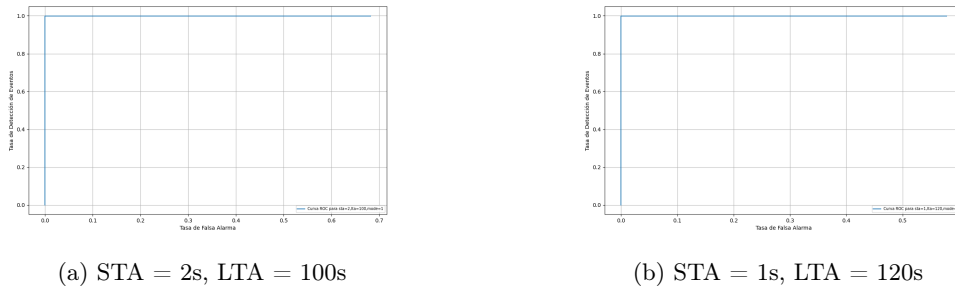


Figura 46: Curvas ROC para CHIE Grupo 2

17.2. Análisis de histogramas para CHIE

A continuación analizamos la diferencia en la respuesta bajo las combinaciones de parámetros STA y LTA, para esto utilizaremos como referencia la Figura 47 Y 48.

Para la Figura 47 mostramos la densidad de probabilidad de valores STA/LTA para la sección de ruido en CHIE para ambos conjuntos de parámetros, y es posible evidenciar la alta similitud entre ambas; sin embargo, en la Figura 48 vemos que la respuesta para la combinación $\{STA = 1s \text{ y } LTA = 120s\}$ (Figura 48b) resulta ser más impulsiva, es decir que a pequeños cambios de amplitud en $x[n]$, mayor será la respuesta de $T(x[n])$, esto nos permite distribuir de manera conveniente los valores STA/LTA pico, atípicos en eventos del Grupo 1.

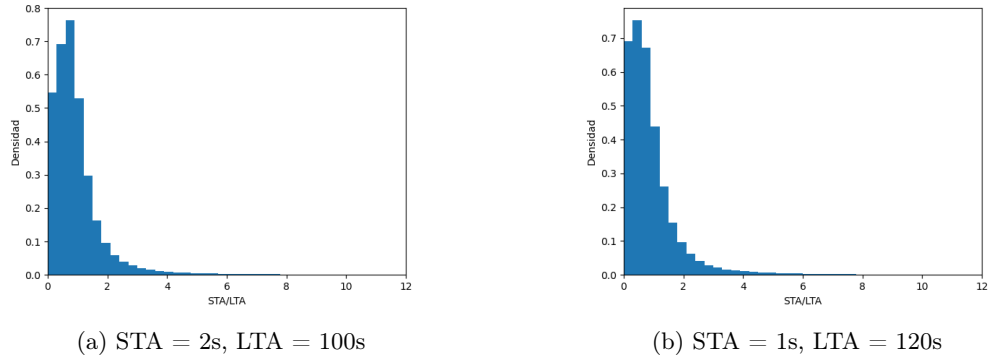


Figura 47: Histogramas de densidad de probabilidad de valores STA/LTA para la sección de ruido en CHIE

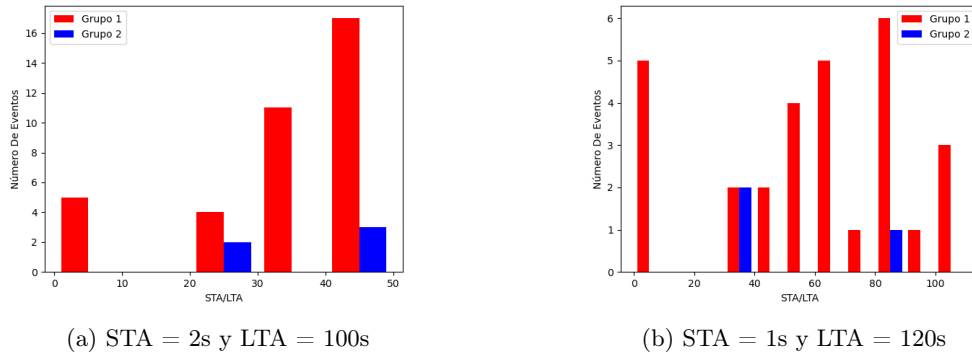


Figura 48: Histogramas de número de eventos para valores pico de función característica STA/LTA en CHIE

17.3. Análisis de probabilidades para CHIE

En el Cuadro 25 se muestra el valor mínimo para $T(S_k[n])$ dado un conjunto de parámetros. De la sección 9.1 obtenemos los criterios de selección para declaración de eventos, se decide tomar un valor $\eta - a$ cercano y menor al menor de los mínimos entre los conjuntos de grupo 1 y grupo 2 para ambos conjuntos STA, LTA.

| Grupo | STA (s) | LTA (s) | MIN |
|-------|---------|---------|-------|
| 1 | 2.0 | 100 | 3.11 |
| 2 | 2.0 | 100 | 21.63 |
| 1 | 1 | 120 | 4.3 |
| 2 | 1 | 120 | 35.8 |

Cuadro 25: Valores mínimos para $T(S[n])$ en CHIE

El Cuadro 26 nos muestra la probabilidad de falsa alarma y la probabilidad de detección para cada grupo dado un conjunto de parámetros. El primer conjunto representa los valores

utilizados en la actualidad. Es posible evidenciar entonces que la probabilidad de falsa alarma es menor al utilizar el conjunto $STA = 1s$, $LTA = 120s$, $\eta_a = 4.2$, se toma la decisión de utilizar aquel conjunto capaz de minimizar la probabilidad de falsa alarma y maximizar la probabilidad de detección para los 3 grupos de eventos. (Ver Cuadro 1).

Es además evidente la mejora de la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros y el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$).

Dada la densidad de probabilidad de valores STA/LTA para la sección de ruido mostrada en la Figura 47b, se propone utilizar un nivel de umbral de desactivación (η_d) del 65 % del valor de η_a debido a que se desea minimizar la probabilidad de desactivación temprana, así como la probabilidad de registro continuo.

El Cuadro 2 muestra los valores propuestos que optimizaran la captura de eventos en CHIE, este conjunto corresponde a $\{STA = 1s, LTA = 120s, \eta_a = 4.2, \eta_d = 2.7\}$

| Grupo | STA (s) | LTA (s) | η_a | P_D (%) | P_{FA} (%) |
|-------|---------|---------|----------|-----------|--------------|
| 1 | 2.0 | 100 | 3.5 | 94.6 | 2.72 |
| 2 | 2.0 | 100 | 3.5 | 100.0 | 2.29 |
| 1 | 2.0 | 100 | 3.0 | 100.0 | 3.67 |
| 2 | 2.0 | 100 | 3.0 | 100.0 | 3.04 |
| 1 | 1 | 120 | 4.2 | 100.0 | 2.20 |
| 2 | 1 | 120 | 4.2 | 100.0 | 1.84 |

Cuadro 26: Probabilidad de detección y probabilidad de falsa alarma dado un conjunto de parámetros STA , LTA y η_a para CHIE

- El algoritmo STA/LTA recursivo no es el mejor método para optimización debido a que su diseño mismo no es óptimo para implementaciones orientadas a detección de eventos locales. Este, además, en ninguno de los casos mostro una mejor combinación de parámetros.
- Se elige como mejor método de implementación para el algoritmo STA/LTA el método clásico, dada su similitud con la implementación en RTQUAKE, este fue además el que mostro en todos los casos la mejor combinación de parámetros.
- Fue posible para cada una de las estaciones optimizar la relación entre la probabilidad de detección y la probabilidad de falsa alarma para el nuevo conjunto de parámetros en comparación con el utilizado actualmente ($\{STA=2s, LTA=100s, \eta_a=3.5, \eta_d=1.5\}$), para eventos de subducción locales.
- Debido a la naturaleza de una implementación de un algoritmo STA/LTA óptimo para cada estación, el error en el tiempo de identificación del tiempo de llegada podrá ser potencialmente minimizado.
- Fue posible demostrar estadísticamente el método y parametrización que proveen los resultados óptimos para la identificación del tiempo de inicio de un evento sísmico.

- Para posteriores estudios se sugiere seleccionar la muestra en base a la estación que se desean optimizar, a manera de tener una muestra uniforme para cada grupo de eventos con características similares. Esto permitirá contar con una mayor cantidad de información para análisis.
- Debido a los recursos computacionales limitados con los que se realiza este estudio, la cantidad de valores que fueron procesados como banco de prueba en el tiempo de duración de este fue reducido, se recomienda, para posteriores estudios de este tipo, utilizar un ordenador con capacidad incrementada o bien más tiempo de estudio para poder así procesar las grandes cantidades de información sísmica.
- Dado que la optimización de filtros no forma parte del alcance de este estudio se recomienda realizar un análisis de frecuencias para cada una de las estaciones y de esta manera optimizar los valores superior e inferior de los filtros pasa bandas a manera de reducir, en lo mayor posible, los efectos del ruido.
- Se recomienda ponderar el voto de las estaciones FG12, STG8 y CHIE con una valoración menor dado que estas tienen una mayor probabilidad de falsa alarma, este mecanismo puede mejorar la declaración eventos de la red.

-
- [1] M. Withers, R. Aster, C. Young, J. B. and; Mark Harris, S. Moore y J. Trujillo, “A comparison of select trigger algorithms for automated global seismic phase and event detection,” *Bulletin of the seismological Society of America*, vol. 88, n.º 1, págs. 95-106, 1998.
 - [2] Y. Vaezi y M. V. der Baan, “Comparison of the STA/LTA and power spectral density methods for microseismic event detection,” *Geophysical Journal International*, vol. 203, n.º 3, págs. 1896-1908, dic. de 2015.
 - [3] B. K. Sharma, A. Kumar y V. M. Murthy, “Evaluation of seismic events detection algorithms,” *Journal of the Geological Society of India*, vol. 75, págs. 533-538, 2010.
 - [4] CEPREDENAC, *Política Centroamericana de Gestión Integral de Riesgo de Desastres*, 2011.
 - [5] I. N. de Sismología, Vulcanología, Meteorología e Hidrología, *Marco Tectónico Para Guatemala*, <https://insivumeh.gob.gt/destacados-portada/marco-tectonico-para-guatemala>.
 - [6] J. A. Á. Gómez, *Tectónica Activa y Geodinámica en el Norte de Centroamérica*. Departamento de Geodinámica de la Universidad Complutense de Madrid, 2009, pág. 20.
 - [7] I. N. de Sismología, Vulcanología, Meteorología e Hidrología, *Red Sismológica Nacional*, doi: 10.7914/SN/GI, 2013.
 - [8] H. Tavera, I. Bernal, C. Condori, M. Ordaz, A. Ceballos y O. Ishizawa, “Evaluación del Peligro Sísmico en Perú,” en. 2014, págs. 15-16.
 - [9] D. Boutelier, “On the relationships between slab dip, back-arc stress, upper plate absolute motion, and crustal nature in subduction zones,” *G3: An electronic Journal Of The Earth Sciences*, vol. 6, n.º 9, sep. de 2005, ISSN: 1525-2027.
 - [10] T. Mejia, *¿Qué es el movimiento convergente de placas?* <https://www.lifeder.com/movimiento-convergente-placas>.
 - [11] A. Trnkoczy, *Understanding and parameter setting of STA/LTA trigger algorithm*, sep. de 1999.

- [12] J. Evans y S. Allen, "A Teleseism-Specific Detection Algorithm For Single Short-Period Trace," *Bulletin of the seismological Society of America*, vol. 73, n.º 4, págs. 1173-11 860, ago. de 1983.
- [13] L. Krischer, J. Smith, W. Lei, M. Lefebvre, Y. Ruan, E. Sales de Andrade, N. Podhorszki, E. Bozdağ y J. Tromp, "An Adaptable Seismic Data Format," *Geophysical Journal International*, vol. 207, n.º 2, págs. 1003-1011, nov. de 2016, doi: 10.1093/gji/ggw319.
- [14] I. R. I. for Seismology, I. F. of Digital Seismograph Networks y U. S. G. Survey, "SEED Reference Manual," en. IRIS, 2012, págs. 1-10.
- [15] T. Utheim y J. Havskov, "RTQUAKE: A Real-Time Earthquake Detection System Integrates with SEISAN," Department of Earth Science: University of Earth Bergen, inf. téc., 2017.
- [16] K. Markham, *ROC curves and Area Under the Curve explained*, <https://www.dataschool.io/roc-curves-and-auc-explained>, nov. de 2014.

21.1. Implementación de STA/LTA clásico en Obspy

```

""" Titulo: obspy.signal.trigger
Autor: Beyruther, M. y Megies, T.
Fecha: 2012
Disponible en https://docs.obspy.org/\_modules/obspy/signal/trigger.html#classic\_sta\_lta """

def classic_sta_lta_py(a, nsta, nlta):
    """
    Computes the standard STA/LTA from a given input
    array a. The length of the STA is given by
    nsta in samples, respectively is the length of the
    LTA given by nlta in samples. Written in Python.

    .. note::

    There exists a faster version of this
    trigger wrapped in C called
    :func:`~obspy.signal.trigger.classic_sta_lta`
    in this module!

    :type a: NumPy :class:`~numpy.ndarray`
    :param a: Seismic Trace
    :type nsta: int
    :param nsta: Length of short time average window in samples
    :type nlta: int
    :param nlta: Length of long time average window in samples

```

```

:rtype: NumPy :class: '~numpy.ndarray '
:return: Characteristic function of classic STA/LTA
"""
# The cumulative sum can be exploited
# to calculate a moving average (the
# cumsum function is quite efficient)
sta = np.cumsum(a ** 2)

# Convert to float
sta = np.require(sta , dtype=np.float)

# Copy for LTA
lta = sta.copy()

# Compute the STA and the LTA
sta[nsta:] = sta[nsta:] - sta[:-nsta]
sta /= nsta
lta[nlta:] = lta[nlta:] - lta[:-nlta]
lta /= nlta

# Pad zeros
sta[:nlta - 1] = 0

# Avoid division by zero by setting
# zero values to tiny float
dtiny = np.finfo(0.0).tiny
idx = lta < dtiny
lta[idx] = dtiny

return sta / lta

```

21.2. Implementación de STA/LTA recursivo en Obspy

```

""" Titulo: obspy.signal.trigger
Autor: Beyruther, M. y Megies, T.
Fecha: 2012
Disponible en https://docs.obspy.org/\_modules/obspy/signal/trigger.html#classic\_sta\_lta """

```

```

def recursive_sta_lta_py(a, nsta, nlta):
    """

```

```

    Recursive STA/LTA written in Python.

```

```

    .. note::

```

```

    There exists a faster version of

```

```

this trigger wrapped in C called :func: ‘
~obspy.signal.trigger.recursive_sta_lta ‘
in this module!

:~type a: NumPy :class: ‘~numpy.ndarray ‘
:~param a: Seismic Trace
:~type nsta: int
:~param nsta: Length of short time average window in samples
:~type nlta: int
:~param nlta: Length of long time average window in samples
:~rtype: NumPy :class: ‘~numpy.ndarray ‘
:~return: Characteristic function of recursive STA/LTA

.. seealso:: [Withers1998]_ (p. 98) and [Trnkoczy2012]_
"""
try:
    a = a.tolist()
except Exception:
    pass
ndat = len(a)
# compute the short time average (STA)
# and long time average (LTA)
# given by Evans and Allen
csta = 1. / nsta
clta = 1. / nlta
sta = 0.
lta = 1e-99 # avoid zero division
charfct = [0.0] * len(a)
icsta = 1 - csta
iclta = 1 - clta
for i in range(1, ndat):
    sq = a[i] ** 2
    sta = csta * sq + icsta * sta
    lta = clta * sq + iclta * lta
    charfct[i] = sta / lta
    if i < nlta:
        charfct[i] = 0.
return np.array(charfct)

```

21.3. Banco de pruebas

```

"""
mode = 1    implementa el algoritmo STA/LTA recursivo de la
            libreria obspy
mode = 2    implementa el algoritmo STA/LTA clasico de la
            libreria obsoy

```

stas y ltas definen los valores que seran utilizados para la obtencion de la funcion caracteristica STA/LTA """

```
modes = ["1","2"]
stas = [0.5, 1.0, 1.5, 2.5, 3.0, 3.5, 4.0, 5.0]
ltas = [30, 50, 60, 70, 80, 90, 100, 110, 120]
```

```
for mode in modes:
    for sta in stas:
        for lta in ltas:
            -> Apply stalta(mode = mode,
                            sta = sta*fs
                            lta = lta*fs)
            -> Save max(T(s[n]))
            -> Save T(w[n])
```

21.4. Código para obtención de función característica STA/LTA

"""

Author: Diego Valdez

Title: characteristic_function_stalta.py

"""

```
from obspy.core import Stream, read, UTCDateTime, Trace
import stalta_defaultParameters as dp
import databaseConnect as db
import numpy as np
import csv
import os
```

```
station = "STG8"
groups = [1,2,3]
mode = "1"
```

```
# data[0] defines STA, data[1] defines LTA
data = [[1.5, 120]]
```

```
for val in data:
    sta = val[0]
    lta = val[1]
    for group in groups:
```

```
# Define save paths for all results
```

```

start_path = os.getcwd()
path = start_path + "/Grupos/grupo{}-ID/".format(group)
save_noise_path = start_path +
"/Resultados/{}/Grupo{}/Noises/".format(station, group)
save_stalta_path = start_path +
"/Resultados/{}/Grupo{}/Stalta/".format(station, group)

print(( "Doing: _STA=_{} , _LTA=_{}").format(sta , lta))

# Get Relational data from database fileids , start , etc...
data = db.retrieve_filenames(station , group)
ids = []
staltas = []
noises = []
coda_count = 0
noncoda_count = 0

# Process each event data for station
for file in data:
    coda = file [2]
    start = file [3]
    date = file [4]
    start = start.replace("_", "0")
    if int(coda) == 0:
        noncoda_count +=1
        coda = 100
    # basica sismic data allocation and trace processing
    st = read(path+file [0]+".MSEED")
    st = st.select(component = "Z")
    st2 = st.select(station = station)
    st_filter = st2.copy()

    st_filter = st_filter.filter ("bandpass",
                                freqmin = 2.0 ,
                                freqmax = 8.0)

    tr = st_filter.merge(fill_value='interpolate')[0]

    # obtain characteristic function
    cft = dp.stalta_trace(tr , length_sta=sta ,
                          length_lta=lta , alg=mode)

    # event start and end times
    station_start = UTCDateTime.strptime(date[0:11]+
                                         start[0:6] , '%Y-%m-%dT%H%M%S')
    station_end = station_start + int(coda)
    station_start = station_start - 30
    event_times = tr.times("utcdatetime")

```

```

# get noise and event related data separated
idx=(event_times>=station_start) &
      (event_times<station_end)
on_event = cft[idx]

if len(on_event) == 0:
    idx=(event_times>=station_start)
    on_event = cft[idx]

on_noise = cft[~idx]

on_event = on_event[~np.isnan(on_event)]
on_noise = on_noise[~np.isnan(on_noise)]

print("max_is", max(on_event))

staltas.append(max(on_event))
noises = np.concatenate([noises, on_noise])
ids.append(file[1])
coda_count += 1
print("Done:_", coda_count)

# save files in directory
information = "{}-{}-{}-{}.txt".format(station,
                                       sta, lta,
                                       mode)
save_noise = save_noise_path + information
save_stalta = save_stalta_path + information

with open(save_noise, "w") as txt_file:
    for line in noises:
        txt_file.write("{}\n".format(line))

with open(save_stalta, "w") as txt_file:
    for line in staltas:
        txt_file.write("{}\n".format(line))

```

21.5. Código para análisis estadístico

```

"""
Author: Diego Valdez
Title: plots_statistics.py
"""

```

```

from matplotlib import pyplot as plt
import numpy as np
from matplotlib.font_manager import FontProperties
import math
import os

station = "ESSG"
groups = [1,2,3]
sta = 1
lta = 120
mode = "1"
mins = 4.2
first = True

start_path = os.getcwd()

for group in groups:

    filename = ("{}-{}-{}-{}").format(station ,
                                     sta ,
                                     lta ,
                                     mode).replace(".txt","")

    data_arr = filename.split("-")
    sta = data_arr[1]
    lta = data_arr[2]
    mode = data_arr[3]
    save_noise_path = start_path +
        "/Resultados/{}"/Grupo{}"/Noises/".format(station ,group)
    save_noise = save_noise_path + filename + ".txt"
    noises = []
    with open(save_noise, "r") as txt_file:
        lines = txt_file.readlines()
        for line in lines:
            temp = line.replace("\n","")
            if not(math.isnan(float(temp))):
                noises.append(float(temp))

    save_stalta_path = start_path +
        "/Resultados/{}"/Grupo{}"/Stalta/".format(station ,group)
    save_stalta = save_stalta_path + filename + ".txt"
    staltas = []
    with open(save_stalta, "r") as txt_file:
        lines = txt_file.readlines()
        for line in lines:
            temp = line.replace("\n","")
            if not(math.isnan(float(temp))):
                staltas.append(float(temp))

```

```

limit = 5 * round(max(staltas)/5) + 5
bins = np.arange(0, limit, 10) # fixed bin size

n_event, ebins, patches1 = plt.hist(staltas, bins=bins)
plt.xlim([0, limit])
# plt.title(event_title)
plt.xlabel('STA/LTA')
plt.ylabel('Numero_De_Eventos')

save_path = start_path +
            "/Resultados/{}/Grupo{}/Event_Histograms/".format(
                station, group)
save_path = save_path + filename + "-E.png"
plt.savefig(save_path)

plt.clf()
plt.cla()
plt.close()

noise_bins = np.arange(0, limit, 0.3) # fixed bin size

n_noise, nbins, patches2 = plt.hist(noises,
                                     bins=noise_bins,
                                     density=True)

plt.xlim([0, 12])
# plt.title(noise_title)
plt.xlabel('STA/LTA')
plt.ylabel('Densidad')

save_path = start_path +
            "/Resultados/{}/Grupo{}/Noise_Histograms/".format(station,
                                                                group)
save_path = save_path + filename + "-N.png"
plt.savefig(save_path)

plt.clf()
plt.cla()
plt.close()

noise_s = np.random.choice(noises, size=1000)

""" EVENT DISTRIBUTION DETECTION PROBABILITY"""
eta = np.arange(0.5, limit, 0.5)
bins = np.arange(0, limit, eta[1]-eta[0]) # fixed bin size
n_event, ebins, patches1 = plt.hist(staltas, bins=bins)
n_noise, nbins, patches2 = plt.hist(noise_s, bins=noise_bins)
plt.clf()

```

```

plt.cla()
plt.close()

pd = []
pfa = []
for e in eta:
    bin_width = ebins[1] - ebins[0]
    nidx = nbins[0: len(n_noise)] >= e
    eidx = ebins[0: len(n_event)] >= e
    pd_s = sum(n_event[eidx]) / sum(n_event)
    pfa_s = sum(n_noise[nidx]) / sum(n_noise)
    pd.append(pd_s)
    pfa.append(pfa_s)

pd_pfa = np.transpose(np.array([pd, pfa]))

leg = "Curva_ROC_para_sta={},lta={},mode={}".format(sta,
                                                    lta,
                                                    mode)

fontP = FontProperties()
fontP.set_size('x-small')
fig2 = plt.figure(figsize=(12.0, 7.0))
l = fig2.add_subplot(111).plot(pfa, pd, label = leg)
fig2.add_subplot(111).grid()
fig2.add_subplot(111).set_ylabel("Tasa
.....de_Deteccion_de_Eventos")
fig2.add_subplot(111).set_xlabel("Tasa_de_Falsa_Alarma")
fig2.add_subplot(111).legend(loc='lower_right',
                             prop=fontP)

ppd = (sum(np.array(staltas) > mins) / len(staltas)) * 100
ppfa = (sum(np.array(noises) > mins) / len(noises)) * 100

print("Probabilidades:_grupo_", group)
print("eta_", mins)
print("Pfa:_", ppfa)
print("Pd:_", ppd)
print("min:_", min(staltas))

save_path = start_path +
            "/Resultados/{}/Grupo{}/ROC/".format(station,
            group)

save_path = save_path + station +
            "-{}-{}-{}R.png".format(sta,
            lta,
            group)

plt.savefig(save_path, dpi = 100)

```

```

plt.clf()
plt.cla()
plt.close()

```

21.6. Archivo de parámetros de RTQUAKE

SERVERS

S01 172.20.1.225:16501

ALLSUBNETS

NETWORK CHILE1 6

| NW STAT | LOC | CMP | FL | FH | STA | LTA | T-ON | T-OFF | SERVER |
|----------|-----|-----|-----|-----|-----|-------|------|-------|--------|
| GI APG | .. | BHZ | 2.0 | 8.0 | 5.0 | 100.0 | 6.6 | 4.2 | S01 |
| GI SMCA | 00 | EHZ | 2.0 | 8.0 | 1.0 | 70.0 | 5.6 | 3.6 | S01 |
| GI FG8 | 00 | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI FG10 | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI FG12 | 00 | BHZ | 2.0 | 8.0 | 0.5 | 70.0 | 4.0 | 2.6 | S01 |
| GI CHIE | .. | HHZ | 2.0 | 8.0 | 1.0 | 120.0 | 4.2 | 2.7 | S01 |
| GI HUEH | .. | HHZ | 2.0 | 8.0 | 1.0 | 70.0 | 4.2 | 2.7 | S01 |
| GI RETA | .. | HHZ | 2.0 | 8.0 | 1.5 | 70.0 | 8.0 | 5.2 | S01 |
| GI STG2 | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI STG5 | .. | HHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI STG8 | .. | HHZ | 2.0 | 8.0 | 1.5 | 120.0 | 6.2 | 4.0 | S01 |
| MX TGIG | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| MX THIG | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| MX SCIG | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| MX CCIG | .. | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| SV NUBE | .. | HHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| SV MTO3 | .. | HHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| SV JAYA | .. | HHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI QUIS | 00 | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI SOKI | 00 | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI GCG4 | 00 | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI JUAM | 00 | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI PCG5 | 00 | BHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI ZAFR2 | .. | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| HN SARH | .. | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| HN NISHN | .. | EHZ | 2.0 | 8.0 | 2.0 | 100.0 | 3.5 | 1.5 | S01 |
| GI ESSG | 00 | EHZ | 2.0 | 8.0 | 1.0 | 120.0 | 7.3 | 4.3 | S01 |

21.7. Código para selección de muestra

```

"""
Author: Diego Valdez
Title: muestra_eventos.py
"""

import csv
import numpy as np
import random as rn
import datetime as dt
import pandas as pd
from dateutil.relativedelta import relativedelta

# Index counter
i = 0;

# Numpy structure array for simplified data clasification
structdef = np.dtype([( 'date', 'U30'),
                        ( 'latitude', 'f8'),
                        ( 'longitude', 'f8'),
                        ( 'depth', 'f8'),
                        ( 'mc', 'f8'),
                        ( 'zone', 'U4'),
                        ( 'nst', 'i4'),
                        ( 'felt', 'U2')])

# Relocating space for data
catalogue = np.zeros(3446, structdef)

# Reading csv file and input data into structure array
with open('reporte2019.csv', newline='') as csvfile:
    spamreader = csv.reader(csvfile, delimiter=',')
    for row in spamreader:
        if (i != 0):
            catalogue[ 'date' ][i-1] = row[0]
            catalogue[ 'latitude' ][i-1] = row[1]
            catalogue[ 'longitude' ][i-1] = row[2]
            catalogue[ 'depth' ][i-1] = row[3]
            catalogue[ 'mc' ][i-1] = row[4]
            catalogue[ 'zone' ][i-1] = row[5]
            catalogue[ 'nst' ][i-1] = row[6]
            catalogue[ 'felt' ][i-1] = row[7]
        i += 1

"""
Group #1: Magnitude >= 4.5
For this group I am using the 62 events registered
"""

```

```

# group1 = catalogue[catalogue['mc']>=4.5][:]
# group1 = group1[group1['zone']=='SL']['date']

# a_file = open("group1.txt", "w")
# np.savetxt(a_file, group1, delimiter=" ", fmt="%s")
# a_file.close()

"""
Group #2: Magnitude >= 3.3, Magnitude <= 4.1
Select 4 events for each month of the year
"""

# save_file = "group2.txt"
# group2 = catalogue[catalogue['mc']>=3.3][:]
# group2 = group2[group2['mc']<=4.1][:]
# group2 = group2[group2['zone']=='SL'][::]

"""
Group #3: Magnitude < 3.3
Select 4 events for each month of the year
"""

save_file = "group3.txt"
group2 = catalogue[catalogue['mc']<3.3][:]
group2 = group2[group2['zone']=='SL'][::]

"""
MAIN ALGORITHM TO GET RANDOMIZED GROUP
ENSURING AT LEAST 4 EVENTS
PER MONTH TO HAVE A REPRESENTATIVE
SAMPLE OF 2019
"""

randomized_g2 = np.array(np.zeros(0), dtype = 'U30')

group2_dated = pd.to_datetime(group2['date'],
                              format='%Y-%m-%d_%H:%M')

from_date = dt.datetime(2019,1,1,0,0,0)
to_date = from_date+relativedelta(months=1)

idx=(group2_dated>=from_date) & (group2_dated<to_date)

while from_date < dt.datetime(2020,1,1,0,0,0):
    idx=(group2_dated>=from_date) & (group2_dated<to_date)
    temp = group2_dated[idx]
    print(len(temp), from_date)
    for x in range(0,4):
        i = rn.randint(0, len(temp)-1)
        event = np.array(temp[i], dtype = 'U30')

```

```

        randomized_g2 = np.append(randomized_g2, event)
        temp = np.delete(temp, i)
    from_date = from_date+relativedelta(months=1)
    to_date = to_date+relativedelta(months=1)

a_file = open(save_file, "w")
np.savetxt(a_file, randomized_g2,
           delimiter="_", fmt="%s")
a_file.close()

```

21.8. Código para solicitudes y administración de base de datos

```

"""
Created on Sun Aug 30 09:19:52 2020

@author: diego
"""
import pymongo
import numpy as np

"""
Function to update event end and start time on MongoDB
filename = filename of record to be updated (string),
station = station of record to be updated (string),
ts = event start time value in samples (integer)
te = event end time value in samples (integer)

Dictionary structure {"_id": ,
                      "file_id": ,
                      "date": ,
                      "station": ,
                      "group": ,
                      "event_start": ,
                      'coda': ,
                      "max_stalta": ,
                      }
"""

def update_times(doc_id, ts, te):

    client = "mongodb+srv://DiegoValdez:123@
    cluster0.bwutu.mongodb.net/<dbname>
    ?retryWrites=true&w=majority"

```

```

cluster = pymongo.MongoClient(client)
db = cluster["tesis"]
collection = db["eventos"]

collection.update_one({"_id":doc_id},
                      {"$set": {"event_start":ts}},
                      upsert=False)
collection.update_one({"_id":doc_id},
                      {"$set": {"event_end":te}},
                      upsert=False)

"""
Function to update max noise and stalta on MongoDB
filename = filename of record to be updated (string),
station = station of record to be updated (string),
max_e = maximun stalta ratio on CF of sismic event window (double)
max_n = maximun noise on CF of noise event window (double)
"""

def update_channel(ids, channel):

    client = "mongodb+srv://DiegoValdez:123@
    ~~~~~cluster0.bwutu.mongodb.net/<dbname>
    ~~~~~?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)
    db = cluster["tesis"]
    collection = db["eventos"]

    for i in range(0,len(ids)):

        update_rec = collection.find_one({ "_id": int(ids[i]) })
        new_val = { "$set": { 'channel': channel}}
        collection.update_one(update_rec, new_val)

"""
Function to retrieve all filenames of
events where coincidence trigger
exists for an specific station.
station = stations to analyze from files (string)
returns: string array of all the files
with the station of interest
"""

def retrieve_filenames(station, group):

    client = "mongodb+srv://DiegoValdez:123@

```

```

cluster0.bwutu.mongodb.net/<dbname>
?retryWrites=true&w=majority"
cluster = pymongo.MongoClient(client)
db = cluster["tesis"]
collection = db["eventos"]

results = collection.find({"station":station, "group": group})
filenames = []
ids = []
coda = []
event_start = []
date = []
for result in results:
    filenames.append(result["file_id"])
    ids.append(result["_id"])
    coda.append(result["coda"])
    event_start.append(result["event_start"])
    date.append(result["date"])
data = np.array([filenames,
                 ids,
                 coda,
                 event_start,
                 date])
data = np.transpose(data)
return data

def post_records(records):
    client = "mongodb+srv://DiegoValdez:123@
cluster0.bwutu.mongodb.net/<dbname>
?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)
    db = cluster["tesis"]
    collection = db["eventos"]

    for record in records:
        post = {"_id": int(record['id']),
               "file_id": str(record['file_id']),
               "date": str(record['date']),
               "station": str(record['station']),
               "group": int(record['group']),
               "event_start": str(record['event_start']),
               'coda': str(record['coda']).replace("_", ""),
               "max_stalta": list(np.zeros(15, dtype="f8")),
               }
        collection.insert_one(post)

def retrieve_times(doc_id):

```

```

        client = "mongodb+srv://DiegoValdez:123@
.....cluster0.bwutu.mongodb.net/<dbname>
.....?retryWrites=true&w=majority"
        cluster = pymongo.MongoClient(client)
        db = cluster["tesis"]
        collection = db["eventos"]

        result = collection.find_one({"_id":doc_id})

        coda = result["coda"]
        event_start = result["event_start"]
        date = result["date"]

        return coda, event_start, date

def available_stations():
    client = "mongodb+srv://DiegoValdez:123@
.....cluster0.bwutu.mongodb.net/<dbname>
.....?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)
    db = cluster["tesis"]
    collection = db["eventos"]

    return collection.distinct("station")

def get_stalta(station, group):
    client = "mongodb+srv://DiegoValdez:123@
.....cluster0.bwutu.mongodb.net/<dbname>
.....?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)
    db = cluster["tesis"]
    collection = db["eventos"]

    results = collection.find({"station":station, "group": group})
    stalta_list = []

    for result in results:
        stalta_list.append(result['max_stalta'])
        channel = result['channel']

    return stalta_list, channel

def retrieve_file(station, group):

    client = "mongodb+srv://DiegoValdez:123@
.....cluster0.bwutu.mongodb.net/<dbname>
.....?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)

```

```

db = cluster["tesis"]
collection = db["eventos"]

result = collection.find_one({"station": station,
                              "group": group})

return result['file_id']

def counter(station):
    client = "mongodb+srv://DiegoValdez:123@
    cluster0.bwutu.mongodb.net/<dbname>
    ?retryWrites=true&w=majority"
    cluster = pymongo.MongoClient(client)
    db = cluster["tesis"]
    collection = db["eventos"]

    results = collection.find({"station": station})
    count = [0,0,0]
    for result in results:
        if result['group'] == 1:
            count[0] += 1
        if result['group'] == 2:
            count[1] += 1
        if result['group'] == 3:
            count[2] += 1

    return count

```

