



**UNIVERSIDAD  
DE GRANADA**

---

MASTER EN GEOFÍSICA Y  
METEOROLOGÍA  
**GEOMET 2016/17**

**ANÁLISIS DE MÉTODOS DE  
IDENTIFICACIÓN AUTOMÁTICA DE  
LLEGADAS DE FASES P Y SU APLICACIÓN  
A LAS SEÑALES SISMO-VOLCÁNICAS DEL  
MISTI (PERÚ)**

**Riky Gustavo Centeno Quico**

Tutores:

**Dr. Gerardo Alguacil de la Blanca  
Dra. María Luz García Martínez**

---

Trabajo de Fin de Master  
Granada, Setiembre 2017

## Agradecimientos

En primer lugar, quiero agradecer a mi amigo el Dr. Gerardo Alguacil por invaluables enseñanzas, por su gran paciencia y por su amistad. Sin su ayuda y sus sabios consejos no hubiera podido cumplir mi propósito de crecer profesionalmente en el mundo de la Geofísica con lo que más me apasiona, la ciencia y los volcanes. A la Dra. Luz García, por sus enseñanzas, y compartir conmigo sus sabios conocimientos en el mundo de la teoría de la señal.

Por otro lado, quiero agradecer el apoyo del Dr. Orlando Macedo, director del Observatorio Vulcanológico del Sur(OVS) del instituto Geofísico del Perú (IGP), por permitirme trabajar con los datos del volcán Misti.

Por último, quiero agradecer el apoyo moral constante de mi familia, en especial a mis padres Juana y Natalio. El de mis nuevos amigos del GEOMET 2016/17 y mis amigos en Perú, en especial a mi amiga Go.

# Índice General

<b>Índice de figuras .....</b>	<b>3</b>
<b>Índice de tablas.....</b>	<b>4</b>
<b>Lista de acrónimos y símbolos.....</b>	<b>5</b>
<b>Resumen.....</b>	<b>6</b>
<b>Capítulo 1 Introducción.....</b>	<b>7</b>
1.1. Justificación del tema .....	7
1.2. Estado del arte .....	8
1.3. Objetivos .....	9
1.3.1 General .....	9
1.3.2 Específicos .....	9
1.4. Área de estudio.....	9
1.5. Alcances .....	10
1.6. Estructura de la memoria.....	10
<b>Capítulo 2 Fundamentos teóricos .....</b>	<b>11</b>
2.1. Fases sísmicas locales .....	11
2.2. Sismos asociados al volcán Misti.....	12
2.3. Identificación (manual) de la onda P.....	13
2.3.1 Detección ( <i>Triggering</i> ).....	13
2.3.2 Identificación y temporización de fases .....	13
2.3.3 Error de observación .....	15
2.3.4 Polaridad del primer movimiento.....	15
2.3.5 Filtrado .....	16
2.4. Realce del arribo de la onda P mediante la potencia específica .....	17
<b>Capítulo 3 Tratamiento de los datos.....</b>	<b>18</b>
3.1. Estaciones sísmicas .....	18
3.2. Datos .....	18
3.3. Pre procesamiento .....	20
3.3.1 Aplicación de la potencia instantánea .....	20
3.3.2 Elección de filtro óptimo de los datos .....	21
3.4. Aplicación de los métodos de picking automático de onda P .....	21
3.4.1 Método basado en la Energía .....	21

3.4.2 Método basado en procesos autorregresivos y el criterio de información Akaike(AIC) .	23
3.4.3 Método basado en estadísticos de orden superior .....	25
3.4.4 Método basado en el periodo predominante.....	26
3.4.5 Método AMPA.....	27
3.4.6 Método AR-K.....	30
3.5. Declaración de la onda P y filtrado de identificaciones falsas .....	33
<b>Capítulo 4 Resultados y su discusión.....</b>	<b>34</b>
4.1. Capacidad de detección: Desempeño .....	34
4.2. Calculo del error: precisión .....	35
4.3. Discusión.....	40
<b>Conclusiones .....</b>	<b>41</b>
<b>Referencias.....</b>	<b>42</b>
<b>Anexo A .....</b>	<b>46</b>

# Índice de figuras

Figura 1.1.-Ubicación la zona de estudio: el volcán Misti. ....	10
Figura 2.1.-Trayectoria de rayos de las principales fases sísmicas locales observadas en un modelo simple de dos capas planas de la corteza terrestre (Diehl et al., 2009). ....	12
Figura 2.2.-Ejemplar de un sismo VT del Misti, registrados en la componente vertical (Z) de la estación MISE, mostrando a) forma de onda, b) espectrograma y c) su lectura de fases P y S. ....	12
Figura 2.3.-Enfoque probabilístico de picking de onda P basado en un cambio de ASNR; donde la amplitud es utilizada para determinar $t_E$ y $t_L$ (Diehl et al., 2009). ....	14
Figura 2.4.-Enfoque probabilístico de picking de onda P basado en un cambio de FSNR; donde la frecuencia es utilizada para determinar $t_E$ y $t_L$ (Diehl et al., 2009). ....	14
Figura 2.5.- Estimación de la polaridad de la onda P y su incertidumbre. Se puede observar la dependencia de la polaridad respecto a la asignación de llegadas a y b. en estos casos es muy común utilizar como calidad intermedia ‘+’ o ‘-’ para una polaridad de tipo.....	15
Figura 2.6.- Efecto de un filtro paso-alto Butterworth de 2º orden (1 Hz) en un registro de período corto. En este ejemplo, el filtro provoca un cambio significativo en la forma de onda (Diehl et al., 2012). ....	16
Figura 3.1.- Distribución de las estaciones de la red Misti, así como los 150 sismos seleccionados para este estudio. ....	19
Figura 3.2.- Periodo de funcionamiento de las estaciones de la red Misti. Las barras en azul, corresponde al número de registros (por hora), utilizados. ....	19
Figura 3.3.- Ejemplo de aplicación de la potencia instantánea a dos señales de las estaciones a) MISE, con una SNR alta; y la b) MISA, con un valor de SNR baja. Para un evento VT, ocurrido el día 15 de enero 2015 a las 02:58Hrs. ....	20
Figura 3.4.- Aplicación de la SSNR a todos los datos y por estación. La parte sombreada en amarillo, corresponde a la banda espectral seleccionada por estación. ....	21
Figura 3.5.- Ejemplo de un sismograma, en la componente Z de la estación MISA, en su forma a) original filtrado en sus cinco bandas (FB1-FB5) y sus respectivas CF(CF <sub>1</sub> -CF <sub>5</sub> ); y b) el mismo evento, pero realzado con la potencia instantánea, sus 5 sub-bandas y su respectiva CF. ....	23
Figura 3.6.- Ejemplo cálculo de la CF, utilizando el método AIC, para un sismo local en la componente Z de la estación MISC;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea. ....	25
Figura 3.7.- Ejemplo cálculo de la CF, utilizando la curtosis (ventana de 1 segundo). Para un sismo local en la componente Z de la estación MISF;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea. ....	26
Figura 3.8.- Ejemplo cálculo de la CF, utilizando el periodo predominante amortiguado $T^{pd}$ . Para un sismo local en la componente Z de la estación MISG;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea. ....	27
Figura 3.9.- a) Diagrama de bloques de un proceso multibanda adaptativo (AMPA) (Alvarez et al., 2013); y b) un ejemplo de eliminación del ruido multibanda a un sismograma de la estación MISE, así como la mejora final de su envolvente, ‘lztot’.....	28
Figura 3.10.- a)Respuesta impulsional teórica para un detector optimo, para un evento impulsivo y de decaimiento suave, así como la compensación a la respuesta impulsional anterior; b) diagrama	

de bloques describiendo la mejora de la envolvente filtrada para diferentes ventanas de longitud L, y el cálculo de la CF (Alvarez et al., 2013).....	29
Figura 3.11.- Cálculo de la CF, utilizando el método AMPA, para un sismo local en la componente Z de la estación MISD.....	30
Figura 3.12.- Diagrama de bloques que ilustra (a) la estimación del error de predicción de un proceso autorregresivo (AR) y b) la segunda parte del proceso que corresponde al cálculo de la CF, utilizando la curtosis. ....	31
Figura 3.13.- Ejemplo cálculo de la CF, utilizando el método propuesto AR-K. Para un sismo local en la componente Z de la estación MISF;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea. ....	32
Figura 4.1.- Histogramas de la diferencia del picking automático y manual obtenidos con los seis métodos empleados, para los datos de las seis estaciones de la red Misti, en sus dos versiones (datos originales en azul y datos realzados en rojo).....	37
Figura 4.2.- Comparación del residuo entre los seis métodos de picking automático utilizados, en función de la SNR de cada lectura. ....	38
Figura 4.3.- Comparación de los seis métodos de picking automático de onda P frente al obtenido manualmente (línea de color negro), para un evento representativo registrado en la componente Z de todas las estaciones de la red Misti, en a) su versión original, y b) los mismos, pero utilizando la potencia específica instantánea. La línea en amarillo, corresponde a la onda P con el método de la energía; la verde, con la curtosis; la magenta, con AIC; la celeste, con AMPA; la azul, con $T^{pd}$ , y la roja con el método AR-K. ....	39
Figura 4.4.- Desempeño de los seis métodos de picking, en función del tiempo que se toman en procesar los 150 sismos seleccionados, tomando como referencia la estación MISD. ....	39

## Índice de tablas

Tabla 2.1.- Esquema que relaciona el peso y precisión de picking.....	15
Tabla 3.1.- Valores de la constante K utilizados, en función del valor de la SNR. ....	33
Tabla 4.1.- Resultados de detección y precisión por estación, para los seis métodos utilizando los datos originales.....	36
Tabla 4.2.- Resultados de detección y precisión por estación, para los seis métodos utilizando los datos realzados. ....	36

## Lista de acrónimos y símbolos

### Acrónimos:

AIC	:	Criterio de Información Akaike.
AMPA	:	Algoritmo adaptativo de temporización automática de onda P.
AR	:	Proceso Autorregresivo.
ASNR	:	Amplitud en términos de la relación señal-ruido.
BHZ	:	Canal vertical de una estación de banda ancha.
CF	:	Función característica.
CVZ	:	Zona Volcánica de los Andes Centrales.
Fig.	:	Figura.
FSNR	:	Frecuencia en términos de la relación señal ruido.
IF	:	Identificaciones falsas.
IGP	:	Instituto Geofísico del Perú.
K	:	Estadístico Curtosis.
OVS	:	Observatorio Vulcanológico del Sur.
p. ej.	:	Por ejemplo.
Pg	:	Onda P que viaja por la capa granítica.
SAC	:	Formato archivos de señales sísmicas.
SNR	:	Relación señal-ruido.
STA/LTA	:	Relación entre promedios de ventana de longitud corta y larga.
T <sub>auto</sub>	:	Tiempo de onda P automático.
T <sub>man</sub>	:	Tiempo de onda P manual.
T <sup>pd</sup>	:	Periodo predominante amortiguado.
UTC	:	Tiempo Coordinado Universal.
VEI	:	Índice de explosividad volcánica.
VT	:	Sismo tipo volcano-tectónico.

### Símbolos:

$\mu$	:	Media
$\sigma$	:	desviación estándar

## Resumen

Este trabajo describe el análisis de seis métodos utilizados en la identificación automática de fases P, aplicados a 150 eventos VT registrados por la red sísmica volcán Misti. Tres de ellos son convencionales y están basados en el análisis de la energía, la curtosis y el criterio de información Akaike (AIC). Dos fueron desarrollados a partir del periodo predominante amortiguado ( $T^{pd}$ ), y del análisis de envolvente multi-banda (AMPA). Así mismo, en este trabajo se propone un método novedoso, denominado “AR-K”. Este método consiste en combinar un análisis autorregresivo de la señal y la curtosis del error de predicción como función característica (CF). Además de este enfoque novedoso, también se utilizó la potencia específica instantánea, la cual permitió realzar la llegada de la onda P y comparar la precisión y exactitud de los métodos con dos tipos de datos: originales y realzados. Se buscó identificar de manera automática 655 fases P; los resultados muestran que el método AR-K es el que detectó el menor número de identificaciones falsas y el que mejor detecta la llegada de ondas P, con un 99% de aciertos utilizando los datos originales. Luego, destacan los métodos AMPA,  $T^{pd}$  y AIC ajustándose mejor a los datos realzados. El método AR-K, también demostró ser el más preciso con una diferencia mínima de  $0.02 \pm 0.02s$ , con los datos originales y  $0.01 \pm 0.02s$ , con los datos realzados. Los métodos AMPA y  $T^{pd}$ , además de su precisión, destacan por su rapidez. Estos tres métodos muestran un buen desempeño, y son propuestos para realizar sistemáticamente la identificación automática de fases P para los datos del volcán Misti.



# Capítulo 1

## Introducción

### 1.1. Justificación del tema

A diferencia de los terremotos que se originan en el límite de las placas tectónicas, los sismos generados en ambientes volcánicos son más difíciles de estudiar. Esto se debe, principalmente, a que la fuente implica movimientos dinámicos de gases, fluidos y ruptura del macizo rocoso, y los trayectos de propagación de onda suelen ser extremadamente heterogéneos, anisótropos y absorbentes, con topografías e interfaces irregulares, incluyendo grietas de todas las escalas y orientaciones (Aki, K., 1992).

El objetivo final de la sismología volcánica es la alerta temprana de erupciones volcánicas a través de la observación directa de los diversos tipos de sismos de origen volcánico. Hacia tal objetivo, en los últimos años, modernas redes de vigilancia sísmica digital se han incrementado en todo el mundo. En el Perú, dada su ubicación en la Zona Volcánica de los Andes Centrales (CVZ), la cual se extiende desde 16°S (sur del Perú) hasta 28°S (norte de Chile); existen antecedentes históricos de actividad volcánica muy explosiva en la zona sur del país. El Instituto Geofísico del Perú (IGP) opera cuatro modernas redes de vigilancia sísmo-volcánica sobre los volcanes Sabancaya, Ubinas Misti y Ticsani.

Otro de los retos más importantes de la sismología volcánica es determinar, con la máxima precisión, el arribo de las diferentes fases sísmicas, desde las ondas P o S directas a otro tipo de ondas, reflejadas, refractadas o superficiales. Un análisis de este tipo demanda mucho tiempo y recurso humano calificado. Sin embargo, un gran volumen de información puede hacer que este análisis manual sea subjetivo debido a la inconsistencia de lectura fases, asociadas a diferentes analistas.

En los últimos años, se han desarrollado diferentes algoritmos que tratan de complementar el trabajo humano con modernos sistemas de reconocimiento automático de fases sísmicas. La mayoría fueron desarrollados para un determinado conjunto de datos o para un problema en particular, tal como la alerta temprana de terremotos, localización de sismos en tiempo real, tomografía sísmica, etc.; no obstante, solo unos cuantos fueron establecidos ampliamente en la comunidad científica.

En este contexto, una detección rápida, precisa, automática y en tiempo real de la llegada de la onda P en sismos volcánicos, como herramienta útil al propósito de alerta temprana de erupciones volcánicas, sería de vital importancia en las tareas de vigilancia como parte de la gestión del riesgo volcánico, ya que haría posible la rápida localización, estimación de magnitud e incluso el mecanismo focal de estos sismos.

## 1.2. Estado del arte

Los primeros intentos de automatizar el proceso de identificación y marcado de fases datan de hace aproximadamente cuarenta años, tiempo en el cual numerosos autores han intentado abordar el problema desde diversas técnicas y perspectivas. Por ejemplo, Stewart (1977) desarrolló el primer procedimiento de identificación automática de onda P, con una precisión algo limitada.

Un paso fundamental hacia la identificación automática de fases sísmicas fue el algoritmo propuesto por Rex V. Allen (1978, 1982). Este autor introdujo por primera vez el concepto de “*función característica*” (CF), la cual resulta de una transformación ‘no lineal’ de la traza sísmica, que es sensible a cambios de amplitud y frecuencia. La relación entre un promedio de corto termino (STA) y uno de largo (LTA) de la amplitud de la señal fue utilizado ampliamente como algoritmo de detección de eventos en registros analógicos (Ambuter y Salomon, 1974). La comparación entre los promedios STA/LTA se aplica a la CF de Allen, en lugar de la señal directamente, lo que permite identificar el instante de llegada de la onda P.

Baer y Kradolfer (1987) mejoraron la función envolvente de Allen, incorporando un umbral de detección de señal dinámico. Este algoritmo marcó un hito en la identificación automática de fases sísmicas, la cual, si es combinada con una evaluación sofisticada de calidad, puede identificar tiempos de llegada de onda P con alta precisión. Recientemente, Lomax et al. (2012) y Vasallo et al. (2012) crearon un algoritmo de reconocimiento automático de alta precisión basado en los trabajos antes mencionados, diseñado para registros sísmicos de banda ancha en tiempo real.

En varios trabajos se propuso evaluar los momentos estadísticos de orden superior, tales como la curtosis y la asimetría (p. ej. Saragiotis et al., 2002; Saragiotis et al., 2004; Panagiotakis et al., 2008; Galiana-Merino et al., 2008; Küperkoch et al., 2010; Baillard et al., 2014; Hibert et al., 2014; Li et al., 2016). La idea subyacente es que el ruido de fondo presenta características diferentes a la de la señal de interés (fases sísmicas de los terremotos). Estos algoritmos proporcionan estimaciones muy fiables de la llegada de la onda P.

Otros métodos que buscan distinguir las diferentes propiedades estadísticas del ruido y la señal son los que usan modelos autorregresivos (AR) (p. ej. Takanami y Kitagawa, 1988; Leonard y Kennett, 1999; Sleeman y van Eck, 1999; y Zhang et al., 2003). Este método está basado en el Criterio de Información de Akaike (AIC) y la asunción de dos segmentos estacionarios en el sismograma, antes (ruido) y después del inicio de la onda P (señal), los cuales pueden ser modelados mediante un proceso autorregresivo (AR).

También fueron propuestos otros enfoques, como aquellos que buscan cambios de polarización (Vidale, 1986; Magotra et al., 1987; Kulesh et al., 2007) o del periodo predominantemente amortiguado ( $T^{pd}$ ) (Lockman y Allen, 2005; Hildyard et al., 2008; Hildyard y Rietbrock, 2010). Otros autores han aplicado la teoría de la lógica difusa (Chu y Mendel, 1994), métodos de máxima verosimilitud (Christofferson et al., 1988; Roberts et al., 1989), amortiguamiento basado en la energía (Kalkan, 2016), análisis de envolvente multibanda y detección de sus formas características (Alvarez et al., 2013; Romero et al., 2016; García et al. 2016), el entrenamiento de un modelo de red neuronal artificial (Dai y MacBeth, 1995; Wang y Teng, 1997; Zhao y Takano, 1999; Gentili y Michellini, 2006) y los basados en wavelets (Anant y Dowlá, 1997; Tibuleac et al., 1999, Gendron et al., 2000; Akansu et al., 2010; Hafez et al., 2013, Karamzadeh et al., 2013; Bogiatzis e Ishii, 2015).

La mayoría de estos métodos se adaptan a un tipo particular de datos (p.ej. telesismos) y son poco eficientes con otros. Esta es la razón por la cual la tendencia actual es optar por combinar varios enfoques, según el tipo de datos utilizados y sus necesidades específicas. Estos enfoques pueden ser iterativos, simultáneos o paralelos; a menudo, se aplican conjuntamente para combinar

refuerzos complementarios (p. ej. Nippres et al., 2010; Küperkoch et al., 2012; Álvarez et al., 2013; y Ross y Ben-Zion, 2014).

## **1.3. Objetivos**

### **1.3.1 General**

Determinar de manera automática, precisa y eficiente, el tiempo de llegada de la onda P para sismos asociados a ambientes volcánicos y, en particular, para el volcán Misti en Perú.

### **1.3.2 Específicos**

- Realzar al máximo posible la llegada de la onda P en cada una de las estaciones sísmicas utilizadas.
- Aplicar cinco de las técnicas más utilizadas para la identificación automática de ondas P, tales como el análisis de energía, procesos autorregresivos combinado con el criterio de información Akaike (AR-AIC), la curtosis, el periodo predominante amortiguado y el análisis de envolvente multibanda (AMPA); todos ellos al conjunto total de datos.
- Analizar el desempeño de cada uno de los métodos utilizados.
- Proponer un algoritmo para la identificación automática de la onda P.

## **1.4. Área de estudio**

El volcán Misti está ubicado en el sur del Perú ( $16^{\circ}18'$ ,  $71^{\circ}24'$ , 5822 m s.n.m.). Colinda al noroeste con el volcán Chachani y al sureste con el volcán Pichu Pichu. Por el sector Suroeste, tiene a sus pies a la ciudad de Arequipa, la cual aparece con un impresionante desnivel de 3500 metros con relación a la cima del volcán. Al pie del flanco Norte y Noroeste discurre el río Chili. Asimismo, algunos kilómetros aguas arriba, se ubican importantes reservorios (Aguada Blanca y El Frayle), los cuales constituyen la principal fuente de agua para los habitantes de la ciudad de Arequipa, así como la agricultura local (Fig. 1.1).

El Misti es un estratovolcán de forma cónica muy regular, formado por numerosas coladas de lava de composición andesítica a riolítica y de espesas tefras. En la cumbre se puede observar una estructura en forma de herradura de 1 x 1.5 km que contiene dos cráteres concéntricos: uno externo de 950 m de diámetro y uno interno de 550 m de diámetro. El interior de este último está ocupado por un “domo tapón” andesítico básico que presenta una intensa actividad fumarólica (Thouret et al., 1999).

Estudios recientes acerca de su evolución muestran que el edificio volcánico fue construido en los últimos 833 mil años, sobre secuencias volcánicas mucho más antiguas, emplazadas durante y después de las erupciones del volcán vecino Chachani (Paquereau-Lebti et al., 2006). El crecimiento del edificio volcánico se debe principalmente a una actividad efusiva que produjo flujos de lava de 20-60 m de espesor y hasta 12 km de longitud, con el aporte adicional de erupciones explosivas de diferente magnitud, con un Índice de Explosividad Volcánica (VEI, por sus siglas en inglés) máximo de 5 (Harpel et al. et al., 2011).

Es uno de los volcanes más activos de la región y es considerado el volcán de mayor riesgo del país, debido a su cercanía a la ciudad de Arequipa que, al día de hoy, cuenta con una población superior al millón de habitantes, así como una muy importante infraestructura (represas, hidroeléctricas, aeropuertos, centros mineros, etc.). Si se produjera una erupción importante, no hay duda de que la ciudad de Arequipa estaría en una posición crítica debido a su proximidad con

respecto al volcán: muy probablemente los materiales de la erupción caerían directamente sobre ella.

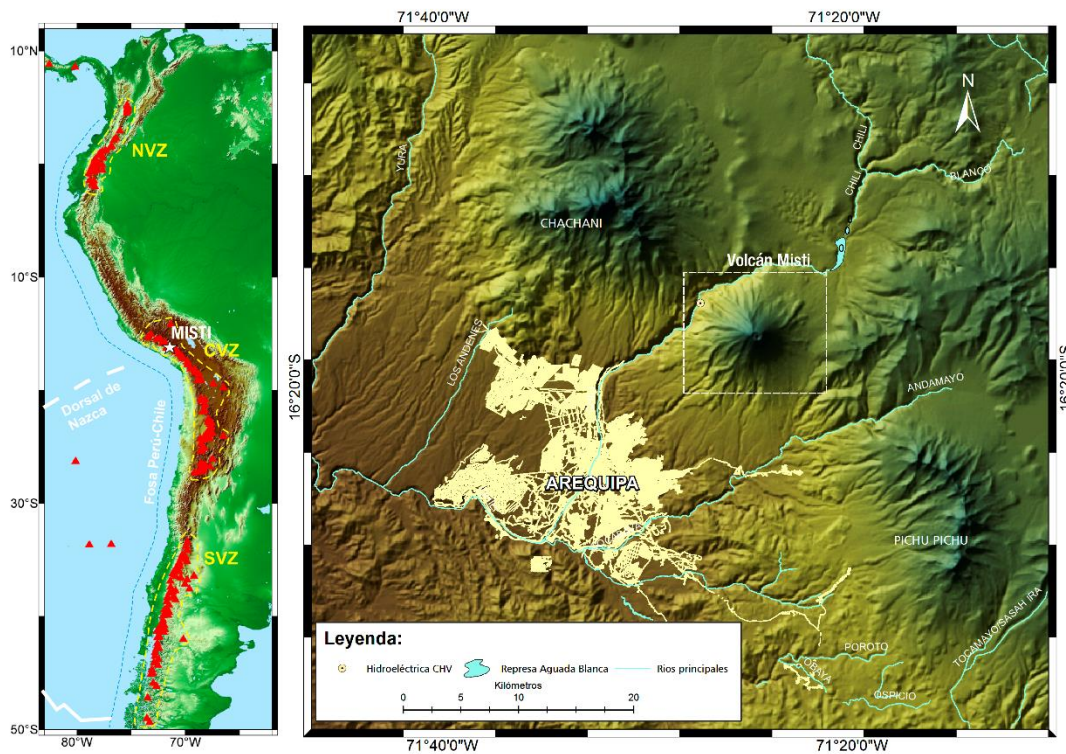


Figura 1.1.-Ubicación la zona de estudio: el volcán Misti.

## 1.5. Alcances

Este trabajo tendrá como propósito fundamental dar solución a uno de los problemas relacionados con la detección automática de señales sismo-volcánicas: la temporización precisa de la onda P, en este caso, para los sismos del volcán Misti (Perú). Para ello, se realizará un análisis comparativo de cinco enfoques utilizados en la identificación automática de la llegada de onda P.

Seguidamente, se hará un análisis comparativo del desempeño de cada método en función de las características de los datos, de manera que sirva de guía al momento de discernir qué método utilizar.

Finalmente, se propondrá un nuevo algoritmo robusto para la identificación automática de fases P, con el objetivo de consolidarlo como una herramienta más para el sistema de alerta temprana de erupciones volcánicas en el Misti. Además, aprovechando la consistencia y rapidez del método, podrá ser utilizado también para otro tipo de estudios, por ejemplo, en la tomografía sísmica.

## 1.6. Estructura de la memoria

La memoria está organizada en cinco capítulos: en el Capítulo 1 se enmarca el problema de investigación, se explica las motivaciones de la misma, el estado del arte y se define los objetivos de este trabajo. En el Capítulo 2 se mencionan algunas nociones básicas necesarias para la comprensión de este trabajo, así como una breve explicación de la potencia específica instantánea como herramienta en el realce de fases sísmicas. En el Capítulo 3 se detalla la base de datos utilizada, la descripción de los métodos empleados y los criterios de detección. En el Capítulo 4 se muestran los resultados obtenidos de la aplicación de los algoritmos de *picking* automático, mostrando distintas gráficas y cuadros comparativos entre ellos. Finalmente, en el Capítulo 5 se da cuenta de las conclusiones del trabajo.

# Capítulo 2

## Fundamentos teóricos

### 2.1. Fases sísmicas locales

Un sismo, es una perturbación o movimiento del suelo originado por un cambio en la distribución de energía elástica en un punto al que suele llamarse *fente sísmica* o *foco*. La fuente radia ondas elásticas internas (*body waves*) de tipo compresional (P) y de cizalla (S), que se propagan por el medio elástico constituido por las rocas del interior de la Tierra. En las capas superficiales, las condiciones de contorno dan lugar a ondas superficiales de tipo Love y Rayleigh, que se forman debido a la interferencia de diferentes ondas internas.

El registro del movimiento del suelo correspondiente es el sismograma. En la mayoría de los sismógrafos actuales, este registro es proporcional a la velocidad del suelo en una cierta banda de frecuencias. Así mismo, es una combinación de los efectos de la fuente, la distancia fuente-estación y el instrumento.

Cada onda (interna, superficial, etc.), que se pueda distinguir en la señal sísmica, se le denomina *fase sísmica*. El punto del sismograma donde comienza, *arribo* o *llegada de fase*; y el tiempo de la misma, *tiempo de arribo* y que, por regla general siempre estará referido al Tiempo Coordinado Universal (UTC). Cuando el arribo es súbito, es decir comienza con un movimiento grande y empinado, se le llama '*impulsivo*'. Por otro lado, cuando el arribo es gradual y resulta difícil determinar su comienzo, se le llama '*emergente*'. Toda esta terminología será utilizada de aquí en adelante.

Cuando se inicia un proceso de identificación de fases, lo primero es determinar el tipo de evento y estimar su distancia. En función de ésta, los sismos pueden ser divididos en tres tipos: locales (cercaos), regionales y telesismos.

Sin embargo, actualmente no existe una única definición estándar sobre un rango de distancia para referirse a cada tipo, debido a que las variaciones regionales de la corteza y la estructura del manto superior, hacen imposible definir una sola distancia en la cual se detenga la propagación de fases al interior de la Tierra (Havskov, 2010; Bormann, 2012). Las fases locales más comunes, son las directas (P y S), las que viajan por la capa granítica de la corteza (Pg y Sg), y las críticamente refractadas entre la corteza y el manto (Pn y Sn) (Fig. 2.1).

Se considera como una fuente local, si las fases Pg y Sg llegan a la estación, como ondas directas P y S.

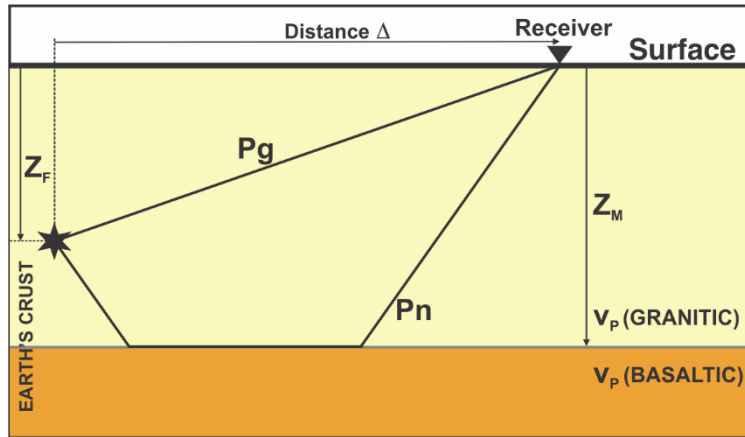


Figura 2.1.-Trayectoria de rayos de las principales fases sísmicas locales observadas en un modelo simple de dos capas planas de la corteza terrestre (Diehl et al., 2009).

## 2.2. Sismos asociados al volcán Misti

De acuerdo al catálogo de sismos del volcán Misti, existen tres tipos de tipos eventos volcánicos principales: Volcano-Tectónicos (VT), Largo Periodo (LP) y Tremor (TRE). Por otro lado, también se registran otro tipo eventos (OTR), sin ninguna relación con el volcán (Centeno R., 2012). El 75% de la actividad volcánica, está conformada por eventos VT, los cuales están asociados a ruptura de roca, inducida por un movimiento de magma y gas al interior del volcán (Zobin, 2012; McNutt, 2005). En su forma de onda, se resalta el carácter impulsivo de sus fases iniciales P y S (Fig. 2.2c), además de una duración relativamente corta entre 12 y 35 segundos (Fig. 2.2a), debido a sus bajas magnitudes. Espectralmente, su energía está concentrada en una amplia banda de frecuencias que puede llegar a superar los 20Hz (Fig. 2.2b). Suelen ocurrir en forma de enjambres sísmicos, con hipocentros localizados entre 1 y 6 km por debajo del cráter.

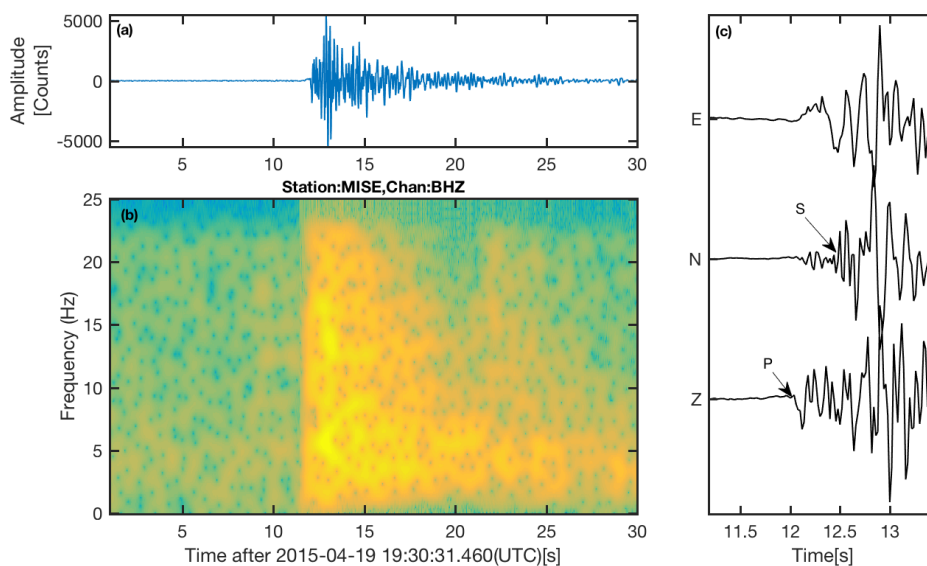


Figura 2.2.-Ejemplar de un sismo VT del Misti, registrados en la componente vertical (Z) de la estación MISE, mostrando a) forma de onda, b) espectrograma y c) su lectura de fases P y S.

## 2.3. Identificación (manual) de la onda P

Antes de cualquier análisis de tipo sismológico es muy importante, previamente, identificar con exactitud la llegada de las ondas sísmicas P y S, en especial la onda P debido a que provee una estimación rápida de la localización espacial de eventos, identificación de eventos, y un análisis de mecanismo fuente (Zhang et al., 2003), además por ser un dato esencial para estudios tomográficos de ondas P (p. ej., Di Stefano et al., 2006, Diehl et al., 2009).

Al proceso de identificar y temporizar la llegada de ondas o fases sísmicas se le denomina “*picking*”. Tradicionalmente es realizado por sismólogos expertos mediante inspección visual de los datos, esto debido a la gran precisión y fiabilidad que se necesita de esta medida.

Las etapas de un proceso de *picking* manual en el sismograma, son las siguientes:

### 2.3.1 Detección (*Triggering*)

En general, la llegada de una fase sísmica se determina mediante dos observaciones visuales (Diehl et al., 2012):

- Por un cambio súbito de amplitud. - En términos de SNR (la relación señal-ruido) la expresaremos como “ASNR”. Si la amplitud supera un nivel umbral, de por lo menos, un factor de 1.5 ( $ASNR \geq 1.5$ ) en más de un registro a la vez; se declara el inicio o llegada de fase (Fig. 2.3).
- Por cambios de frecuencia. - En términos de SNR la expresaremos como FSNR. Este observable, a diferencia de la ASNR, es mucho más difícil de analizar visualmente. Debido a que la frecuencia dominante del ruido y la señal, a veces pueden ser similares. Pero para ventanas de tiempo más largas, la FSNR puede ayudar a determinar aproximadamente la llegada de una fase (Fig. 2.4).

### 2.3.2 Identificación y temporización de fases

Una vez detectada la llegada de fase, en base a ASNR o por FSNR (o una combinación de ambas), seguidamente se define el tiempo absoluto.

Entonces, si consideramos la llegada de una fase sísmica como una función probabilística  $P_a(t)$ , el tiempo de arribo “más probable”  $t_A$  podría expresarse como  $P_a(t_A) = \text{Max}(P_a)$ . De otro lado podemos definir un tiempo de llegada de fase, lo más ‘anticipado’ posible  $t_E$ , donde la probabilidad que se inicie se aproxime a cero  $P_a(t_E) \geq 0$ . Equivalentemente, el tiempo de inicio más ‘tardío’ posible  $t_L$ , se define como  $P_a(t_L) \geq 0$ .

En la práctica, es el mismo analista el que define los límites de  $t_E$  y  $t_L$ . En el ejemplo de la figura 2.3,  $t_L$  se ubica en la intercepción de la señal y el nivel umbral de ruido de fondo (1.5 veces la amplitud del ruido). Ubicar  $t_E$  no es tan simple como  $t_L$ : primero se ajusta una tangente (línea ‘a’ entrecortada) a la pendiente suavizada del inicio de señal y, desde  $t_L$  ‘a priori’ donde la pendiente sea cero con la curva de la señal, se posiciona  $t_E$ . Seguidamente determinamos el tiempo absoluto de la fase (*picking*) en la posición más probable  $t_A$ .

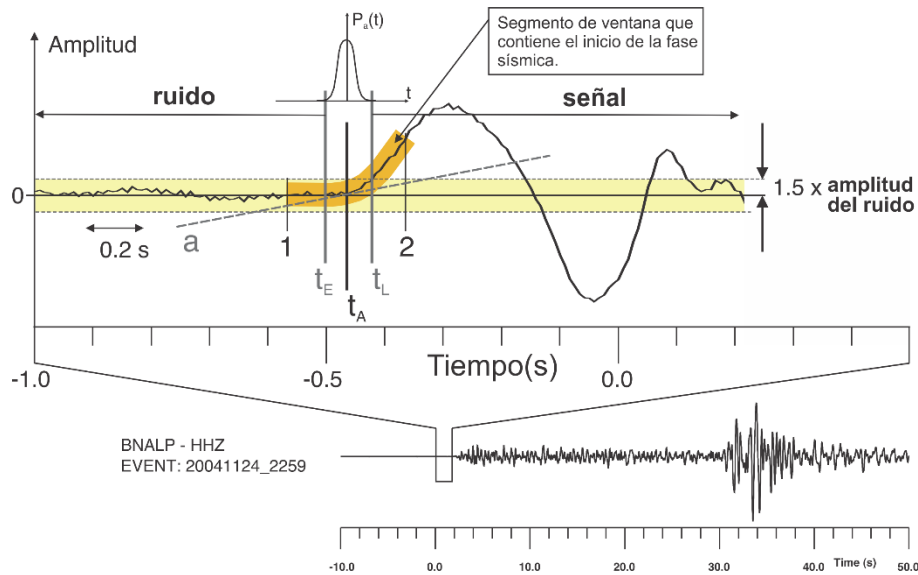


Figura 2.3.-Enfoque probabilístico de picking de onda P basado en un cambio de ASNR; donde la amplitud es utilizada para determinar  $t_E$  y  $t_L$  (Diehl et al., 2009).

En la figura 2.4, el cambio en frecuencia es más notorio posterior a la posición 'b' y, según este se propone ubicar  $t_L$  a un cuarto de periodo de la ondícula ( $T/4$ ) después de la posición 'b'. Luego considerando un tiempo de llegada anticipado de fase, al cual llamaremos 'a', ubicamos  $t_E$  a medio periodo de la ondícula ( $T/2$ ) antes de la posición 'a'. Una determinación consistente de  $t_A$ , en este caso, es mucho más difícil que el anterior. Aunque parezca que  $t_A$  este ubicada en el medio de  $t_E$  y  $t_L$ , no siempre es así (especialmente en registros de banda ancha).

Actualmente, no existe una definición universal para determinar  $t_A$ , pero su posición generalmente coincide con un cambio notable en la forma de onda que separa el ruido de la señal.

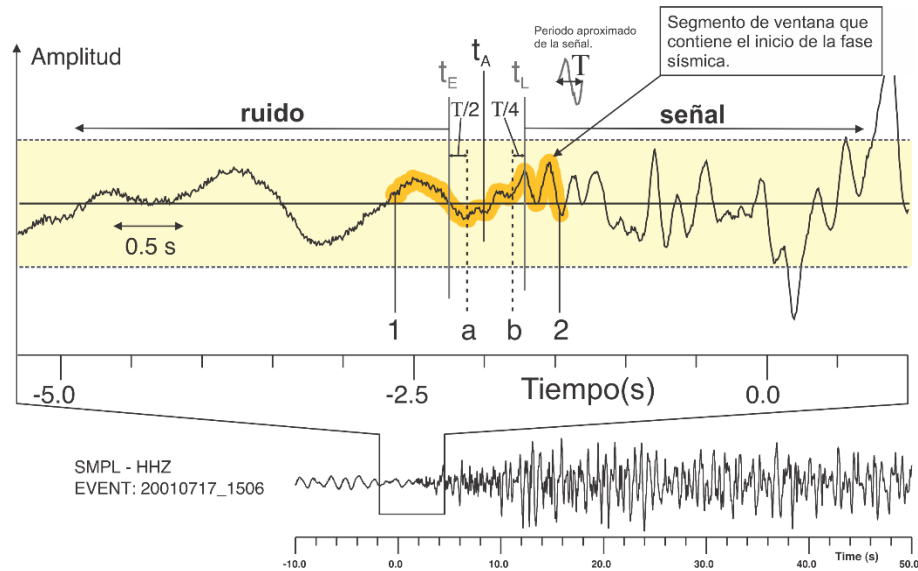


Figura 2.4.-Enfoque probabilístico de picking de onda P basado en un cambio de FSNR; donde la frecuencia es utilizada para determinar  $t_E$  y  $t_L$  (Diehl et al., 2009).



### 2.3.3 Error de observación

Después de identificar o temporizar una fase sísmica, es muy importante también determinar el error de observación y frecuentemente está asociado a la asignación de un peso o grado de calidad. El esquema de pesos fue introducido por Lee y Lahr (1975) en el primer programa estándar de localización hipocentral “HYPO71”. Por ejemplo, para el procesamiento de sismos en el observatorio vulcanológico de Alaska (Lahr, 1999), este puede variar de 0 a 4: 0 para un intervalo de error ( $t_L-t_E$ ) menor a 0.2 segundos; y 4 para lecturas de  $t_A$  “demasiado inciertas” y un intervalo de error mayor a 1 segundo (Tabla 2.1).

En ese sentido, la disponibilidad de intervalos de error permitirá evaluar el desempeño de los algoritmos de *picking* automático de forma cuantitativa (Diehl et al., 2012).

Tabla 2.1.- Esquema que relaciona el peso y precisión de picking.

Peso	Precision(s)
0	< 0.2
1	$\geq 0.2$ - <0.4
2	$\geq 0.4$ - <0.7
3	$\geq 0.7$ - <1.0
4	> 1.0

### 2.3.4 Polaridad del primer movimiento

Es muy útil para una rápida determinación del mecanismo focal. Normalmente se le asignan dos letras: 'U' o 'C', si el movimiento o polaridad de la onda P es para ‘arriba’ (corresponde a un primer movimiento compresivo) y 'D', si el movimiento es opuesto al anterior (dilatación). La figura 2.3, es un claro ejemplo del primer caso.

Por otra parte, hay casos como el de la figura 2.4 donde es imposible determinar la polaridad. Así como también existen casos, donde a veces es necesario definir una clase de polaridad de calidad intermedia. Por ejemplo, en la figura 2.6 se muestran dos posibles lecturas de tiempo de llegada, 'a' y 'b'. Y dependiendo cual se elija, 'U' o 'D', en ambos casos es relativamente incierto definir una polaridad real; para este tipo de escenarios se sugiere utilizar los símbolos '+' o '-', como clase intermedia de polaridad 'U' o 'D', respectivamente.

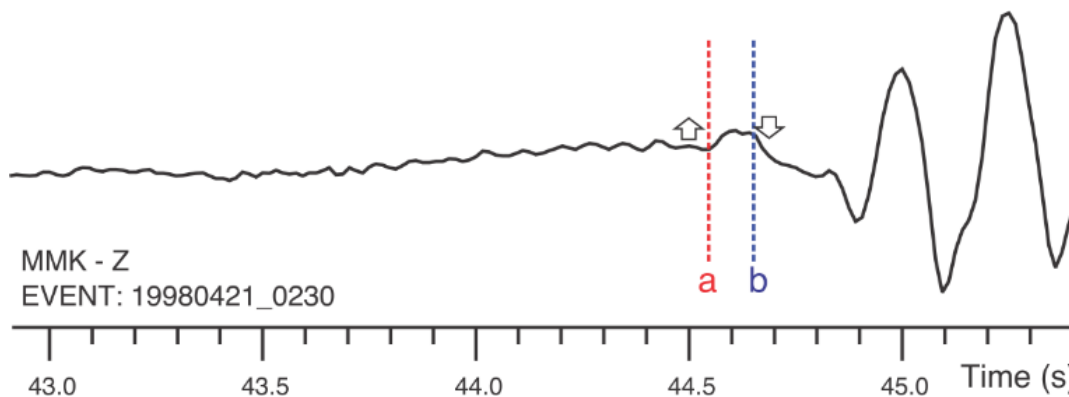


Figura 2.5.- Estimación de la polaridad de la onda P y su incertidumbre. Se puede observar la dependencia de la polaridad respecto a la asignación de llegadas a y b. en estos casos es muy común utilizar como calidad intermedia '+' o '-' para una polaridad de tipo

### 2.3.5 Filtrado

En un proceso de identificación de fases se recomienda (en lo posible), leer las señales sin filtrar. En el caso de registros de banda ancha (en la mayoría de casos), es necesario filtrar (p. ej. un filtro paso-alto, con una frecuencia de esquina 1 Hz) debido a la presencia de ruido de largo periodo presente en este tipo de registros. Además, esto simula la señal característica de un registro de período corto, minimizando la inconsistencia debida al uso de diferentes tipos de sensores.

El uso de filtros en el análisis sísmológico, tiene como finalidad reducir los niveles de ruido, que afecten el reconocimiento exacto de llegada de fases sísmicas. Por otro lado, los filtros comúnmente utilizados (pasa-banda, paso-alto) también distorsionan parte de la señal mientras intentan aumentar la relación señal-ruido (SNR). Por lo tanto, se debe tener especial cuidado al aplicar algún tipo de filtro en un proceso de *picking*.

En primer lugar, debe tenerse en consideración que el registro sísmico original, es por sí mismo la salida de un filtro (función de transferencia de sensor y digitalizador). Un desplazamiento de fase introducido por un filtro, depende de la respuesta del instrumento, de la respuesta del filtro y del contenido de frecuencias de la señal.

Si la frecuencia dominante de la señal está próxima a la frecuencia de esquina ( $f_c$ ) del filtro, más notorio será el efecto esperado sobre la señal. Además, cuanto más pronunciados estén los flancos del filtro (mayor orden), más fuerte será el efecto esperado en frecuencias próximas a la  $f_c$ .

La Figura 2.7, muestra el efecto resultante de la aplicación de un filtro arbitrario, a un registro de periodo corto proporcional a la velocidad. La frecuencia de esquina de un filtro pasa-alta de 2<sup>do</sup> orden ( $f=1\text{Hz}$ ), está cerca de la frecuencia dominante de la primera parte del grupo de ondas. Como resultado, el filtro afecta significativamente las bajas frecuencias de la señal.

La posición 'a' y 'b' representan dos posibles tiempos de llegada de onda P en el registro original y el filtrado (paso-alto), respectivamente. La diferencia de tiempo entre ambas es mayor a 0,2 s, un valor de incertidumbre alto para ambas llegadas. Por otro lado, el proceso de filtrado también puede afectar a la polaridad aparente de la onda P.

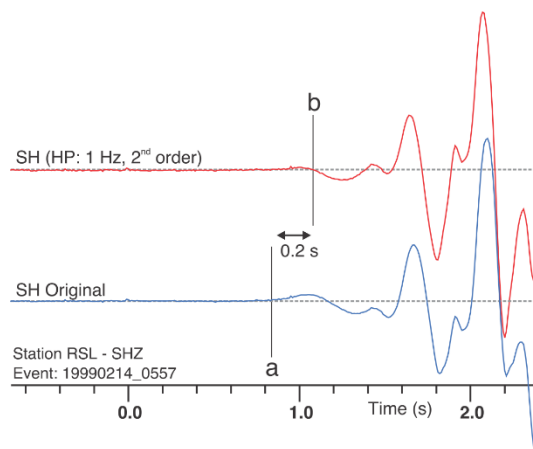


Figura 2.6.- Efecto de un filtro paso-alto Butterworth de 2º orden (1 Hz) en un registro de período corto. En este ejemplo, el filtro provoca un cambio significativo en la forma de onda (Diehl et al., 2012).

Una recomendación importante, si se utiliza algún tipo de filtro, es asegurarse que no provoque cambios de fase en la señal (p. ej. filtros de fase mínima).

## 2.4. Realce del arribo de la onda P mediante la potencia específica

Una manera de mejorar el potencial de los métodos antes descritos, es realzando las características de las primeras llegadas de la señal sísmica mediante la *potencia específica*  $p$ , introducida como un indicador de la severidad de la sacudida en movimientos fuertes (Olea et al, 2011; Alguacil y Vidal, 2012).

Supongamos una masa  $m$  rígida que es sometida a un movimiento del suelo, debido al paso de las ondas sísmicas. La fuerza total que actúa sobre ella debe ser

$$\vec{F}(t) = m\vec{a}(t) \quad (2.1)$$

donde  $\vec{a}(t)$  es la aceleración de su centro de masa y  $\vec{F}(t)$  es la resultante de todas las fuerzas que actúan sobre ella. Llamando  $\vec{r}$  y  $\vec{v}(t)$  al vector de posición y la velocidad instantánea de su centro de masa (c.m.), la potencia instantánea de traslación desarrollada por las fuerzas sobre ella vendrá dada por

$$\vec{F}(t) \cdot \frac{d\vec{r}}{dt} = m\vec{a}(t) \cdot \vec{v}(t) \quad (2.2)$$

Si se considera la potencia por unidad de masa, que llamamos *potencia específica*  $p$ , será

$$p = \vec{a}(t) \cdot \vec{v}(t) \quad (2.3)$$

Esta función es escalar y se puede calcular para una o las tres componentes del sismograma. Su interés como auxiliar del *picking* viene dado por dos características:

- Resalta los paquetes de energía de las ondas.
- Resalta los movimientos con alta rectilinealidad (ondas P o S), frente a los elípticos (ondas superficiales que predominan en el ruido local).

En efecto, para un movimiento puramente circular uniforme,  $\vec{a}$  es la aceleración normal y  $\vec{v}$  sería tangencial, por lo que el producto escalar  $\vec{a} \cdot \vec{v}$  sería nulo. En cambio, en un movimiento rectilíneo la dirección de ambos es la misma y su producto escalar es máximo en valor absoluto.

El teorema de la energía cinética nos dice que el trabajo realizado por estas fuerzas es la energía cinética de traslación de la masa  $m$ , es decir

$$e_c(t) = \int_0^t p(\tau) d\tau = \frac{v(t)^2}{2} \quad (2.4)$$

que obviamente es no-negativa.

Los registros obtenidos con un sismógrafo son proporcionales a la velocidad del suelo al menos en una cierta banda de frecuencias, por lo que el cálculo de  $p$  se reduce a derivar para estimar la aceleración y realizar el producto escalar de ambos. En registros de aceleración el procedimiento sería integrar para calcular la velocidad y de nuevo el producto escalar.

# Capítulo 3

## Tratamiento de los datos

### 3.1. Estaciones sísmicas

Para este estudio, se emplearon datos de la red de vigilancia sísmica del volcán Misti, para el periodo enero - abril 2015. Esta red pertenece al Instituto Geofísico del Perú (IGP) y es operada a través del Observatorio Vulcanológico del Sur (OVS) desde octubre del año 2005, como parte del proyecto de monitoreo sismo-volcánico en el sur del país.

Actualmente, la red Misti cuenta con seis estaciones sísmicas situadas en la cima y en los alrededores del volcán, entre los 4200 y 5700 msnm, transmiten las señales sísmicas vía radio telemetría hasta la estación central del OVS, ubicada en la ciudad de Arequipa-Perú. Allí se analizan los datos diariamente, con el fin de observar el estado de la actividad sismo-volcánica del Misti.

Cada estación está compuesta por los siguientes elementos: un mástil de concreto para el panel solar, GPS y las antenas de comunicación; una caseta de seguridad la que contiene una batería, un digitalizador RefTek RT130 y el sismómetro. Este último, puede ser de dos tipos: tres son de banda ancha (MISE, MISD y MISG) y están equipadas con sensores de la marca Guralp modelo GMG-3SPEC, de 30 segundos; y las tres siguientes, son de periodo corto (MISA, MISC y MISF) y tienen sensores Lennartz LE-3D Lite, de 1 segundo. Todas son de 3 componentes (Norte-Sur, Este-Oeste y Vertical), tienen una tasa de muestreo de 100 Hz (a excepción de la estación MISA, con 50 Hz), y operan en modo continuo, en tiempo real. La Figura 3.1, muestra la distribución de las seis estaciones sobre el edificio volcánico.

### 3.2. Datos

Los datos empleados, corresponden al registro de eventos tipo VT o volcano-tectónicos ocurridos durante el periodo de estudio (1877 eventos). Pero antes, era necesario hacer una selección de eventos representativos, considerando para ello, sismos con una 'buena' lectura de fases P y S, obtenidas de forma manual en un mínimo de tres estaciones de la red Misti.

Por otro lado, la disponibilidad de datos fue restringida a un determinado número de eventos usados para la elaboración de mapas de sismicidad, es decir, se trabajó con algunas horas de registros sísmico continuo por estación y en formato SAC. En la Figura 3.2, las barras en azul, representan la disponibilidad total de datos durante el periodo de estudio y por estación.

Bajo esta premisa, se seleccionaron (aleatoriamente) 150 eventos VT los cuales fueron localizados con 655 lecturas de fases P y 573 de S. Estas lecturas, sobre todo de la onda P, serán utilizadas en lo posterior como referencia al momento de evaluar similitudes con las obtenidas automáticamente. Del mismo modo, permitirá probar la robustez de los diferentes algoritmos de

*picking* automático utilizados. Los detalles de los 150 sismos seleccionados (fecha, hora, magnitud, hipocentro, etc.), podrán ser consultadas en el Anexo A.

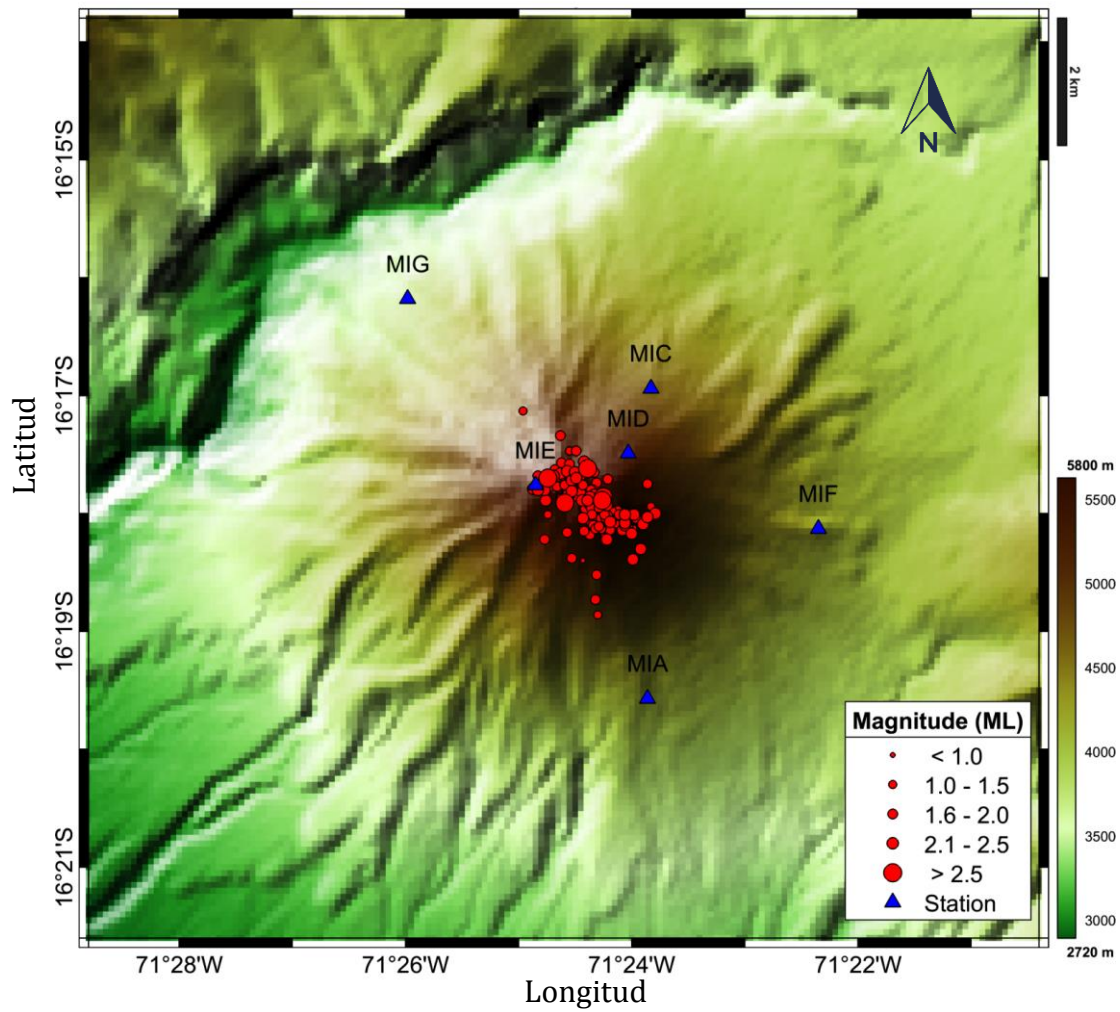


Figura 3.1.- Distribución de las estaciones de la red Misti, así como los 150 sismos seleccionados para este estudio.

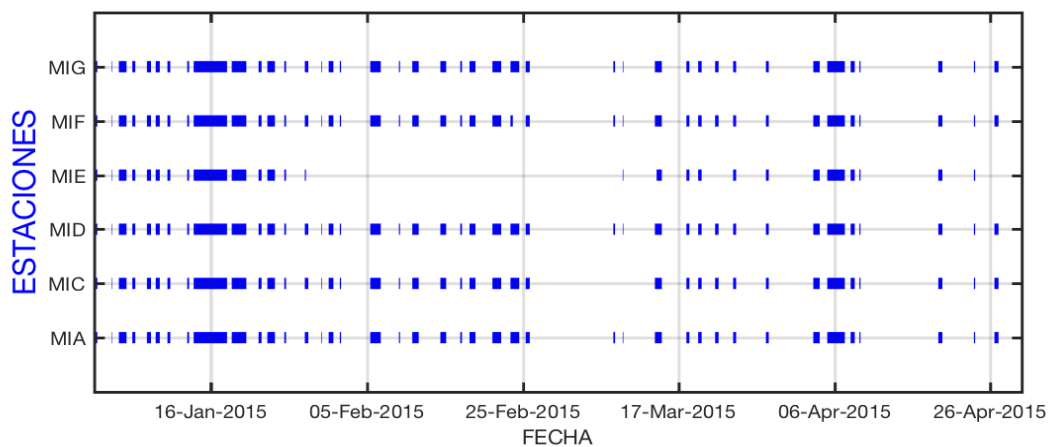


Figura 3.2.- Periodo de funcionamiento de las estaciones de la red Misti. Las barras en azul, corresponde al número de registros (por hora), utilizados.

### 3.3. Pre procesamiento

Cualquier algoritmo de *picking* automático, usualmente requiere de algunos parámetros antes de su ejecución. Debido a que son algoritmos computacionales y una mala elección de los mismos pueden hacer que los métodos fallen, y si los resultados fuesen independientes de tal elección, se deberá tener aún más cuidado en los resultados. Entonces, para una correcta elección de los mismos, se deberá tener en cuenta: el tiempo y amplitud de la señal; contenido espectral; y la SNR de cada estación.

Tomando como referencia el tiempo de ocurrencia de los 150 eventos antes mencionados, fueron segmentados en ventanas de 47 segundos, alrededor de la llegada de cada evento. La razón del tamaño de ventana, fue tratar de evitar la ocurrencia de más de un evento por traza.

#### 3.3.1 Aplicación de la potencia instantánea

Como se mencionó en el capítulo 2, una manera de realzar la llegada de fases P y S frente al ruido sísmico, es la aplicación de la potencia instantánea. En nuestro caso, como es de solo interés la onda P, solo fue necesario aplicarla a la componente vertical (Z) de los sismogramas (Fig. 3.3).

Además de realzar la llegada de la onda P en el tiempo, como en frecuencia (Fig. 3.3). Se pudo observar una mejora notable en los valores de SNR en la mayoría de los datos y las estaciones; por ejemplo, los valores de SNR mejoraron en más del 80% respecto a los datos originales, tal y como se muestra en la Figura 3.3, a continuación.

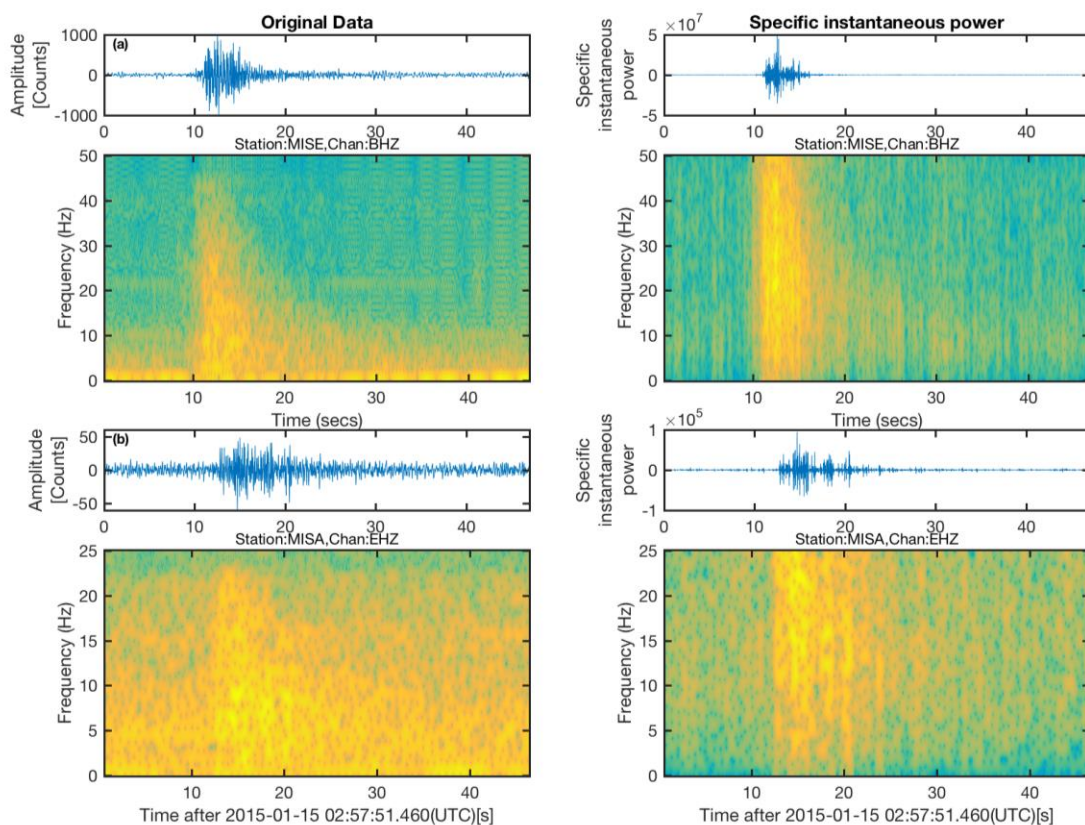


Figura 3.3.- Ejemplo de aplicación de la potencia instantánea a dos señales de las estaciones a) MISE, con una SNR alta; y la b) MISA, con un valor de SNR baja. Para un evento VT, ocurrido el día 15 de enero 2015 a las 02:58Hrs.

### 3.3.2 Elección de filtro óptimo de los datos

Lo siguiente consistió en determinar un filtro óptimo que resalte la llegada de la onda P. Una manera de hacerlo fue utilizando la relación espectral señal-ruido media (SSNR), para todos los eventos y por estación.

Para ello: primero, se tomaron dos ventanas del mismo tamaño (1 segundo) que contengan: la onda P y el ruido; segundo, se calcula la transformada de Fourier para ambos segmentos; tercero, se calcula el logaritmo de la relación de amplitudes en modulo; y para finalizar, se determina la media de todas las razones espectrales como

$$SSNR = \langle \log_{10} |S_s/S_r| \rangle$$

Donde:  $S_s$  es el espectro medio de la señal y  $S_r$  del ruido.

En la Figura 3.4, se muestra la aplicación de la SSNR a todos los datos seleccionados, en su forma original (línea azul) y con el cálculo de la potencia instantánea (línea roja). En general, se puede observar una SSNR media buena en la mayoría de estaciones, hasta frecuencias altas próximas a los 30Hz.

Entonces, si consideramos una banda para filtrar en donde la SSNR media este por encima del 70% del máximo, tenemos que para la estación: MISA, 1-17Hz; MISC, 1-23Hz; MISD, 1-27Hz; MISE, 1-22Hz; MISF, 1-12Hz; y MISG, 1-17Hz.

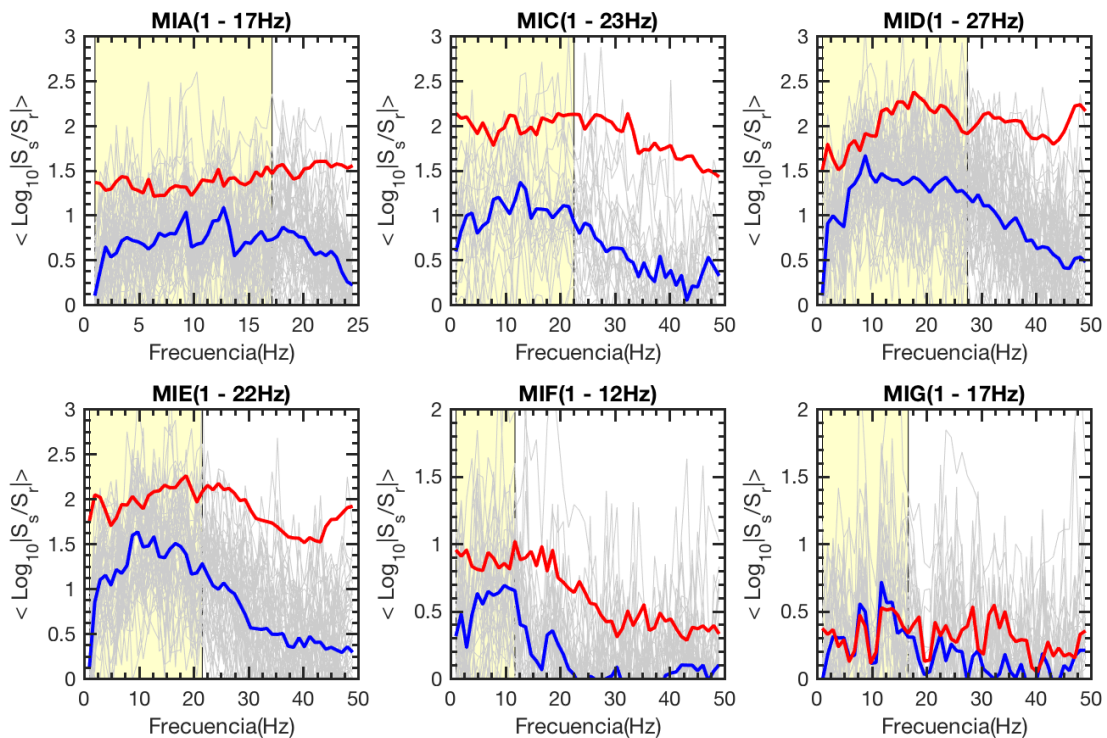


Figura 3.4.- Aplicación de la SSNR a todos los datos y por estación. La parte sombreada en amarillo, corresponde a la banda espectral seleccionada por estación.

## 3.4. Aplicación de los métodos de *picking* automático de onda P

### 3.4.1 Método basado en la Energía

Este método consiste en comparar muestra a muestra la señal, entre un valor actual y uno predicho a partir de muestras anteriores; si la relación entre ambas supera un valor umbral, entonces se declara una posible llegada de fase. Esta es pocas palabras la idea básica de la mayoría de

algoritmos convencionales de tipo STA/LTA, ampliamente utilizados para la identificación de la onda P (Allen, 1978,1982; Baer y Kradolfer, 1987; Lomax et al., 2012; Vassallo et al., 2012; Chen y Holland, 2016).

A medida que la potencia de cálculo ha mejorado, se ha hecho posible que algoritmos más complejos se ejecuten en tiempo corto o incluso en tiempo real. Una mejora reciente al método sugiere un procesamiento multibanda simple y por cada banda se calcula una CF (Lomax et al., 2012).

Este algoritmo, primero aplica filtro de ancho de banda una octava (relación de frecuencias, de una al doble) la cual genera algunas bandas de frecuencia por sismograma; estas se adaptan automáticamente a la frecuencia de muestreo de los datos y dependiendo de la elección de la frecuencia mínima central, puede incluir frecuencias bajas y altas, no llegando a superar la frecuencia de esquina (Nyquist).

Seguidamente, calcula la energía  $E_n$  de los datos filtrados para cada banda, como:

$$E_n[i] = BF_n[i]^2 \quad (3.1)$$

Donde  $BF_n[i]$ , es la amplitud de los datos filtrados (pasa-banda) de la  $i$ -ésima muestra de la  $n$ -ésima banda. Para calcular la CF de cada banda, se añadió el valor medio cuadrático (*rms*) y la desviación estándar (*st dev.*) en las ecuaciones (3.2) y (3.3).

$$CF_n^{rms}[i] = \frac{E_n[i]}{rms(E_n[i-1-l:i-1])} \quad (3.2)$$

$$CF_n^{st.dev.}[i] = \frac{E_n[i] - mean(E_n[i-1-l:i-1])}{st.dev.(E_n[i-1-l:i-1])} \quad (3.3)$$

Donde  $l$ , es la longitud de ventana de la muestra. Para el modo *rms*,  $CF_n^{rms}[i]$  es la relación entre la energía transitoria de la muestra  $i$  en la banda de frecuencias  $n$  y el valor *rms* de energía en la ventana móvil anterior; para el modo *st.dev.*,  $CF_n^{st.dev.}[i]$ , es la relación entre diferencias de energía transitoria de la muestra  $i$  y la media de su ventana anterior, para una banda de frecuencias  $n$  y la desviación estándar de la energía en su ventana anterior.

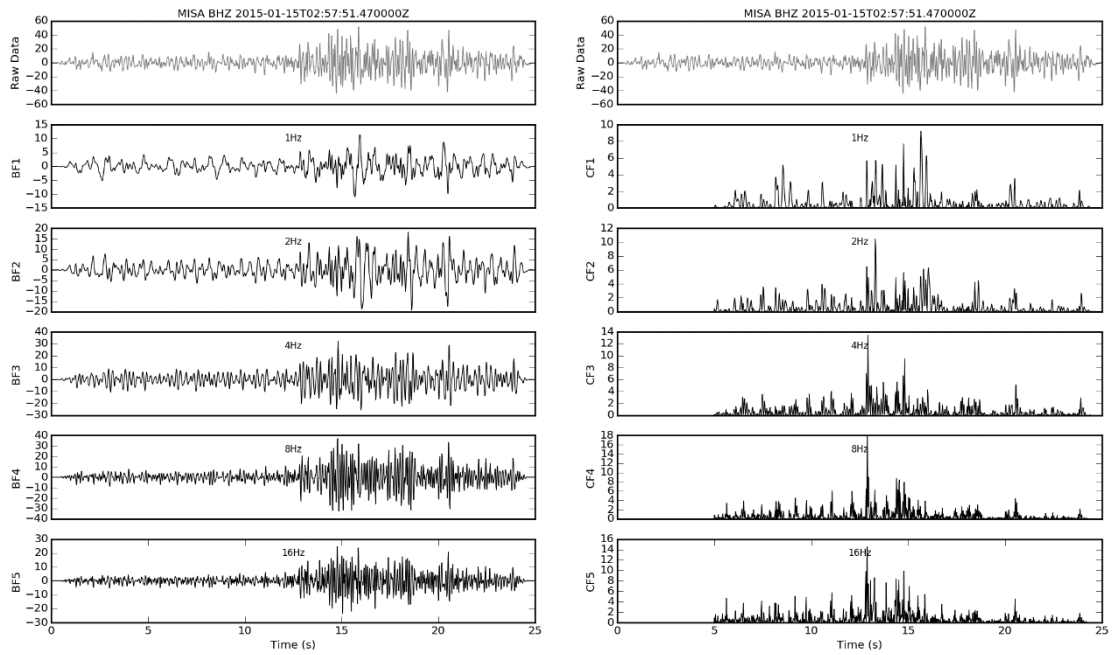
La CF cuantifica el cambio de energía en relación al nivel de energía de la ventana anterior. Por lo tanto, un alto valor de  $CF_n[i]$  indica cambios de energía de una muestra respecto a la ventana anterior. Finalmente, el algoritmo para obtener la CF, resume la  $CF_n[i]$  tomando el valor máximo sobre todas las  $n$  bandas para cada muestra  $i$  (Lomax et al., 2012). Además, el modo *rms* permite mejorar el tiempo de cálculo en aproximadamente un 30%, sin afectar significativamente el rendimiento del algoritmo. La razón es que se elimina la tendencia de los datos, en la fase de acondicionamiento de los mismos.

Según la Figura 3.4, los registros del Misti tienen una buena SSNR por encima de 1Hz. Entonces utilizando la misma, como frecuencia mínima central, se determinaron 6 bandas con una frecuencia central en 1,2,4,8,16,32 Hz (a excepción de la estación MISA, con 5), respectivamente (Fig. 3.5). Se empleó el modo *rms* y una longitud de ventana de muestra  $l$  de 5 segundos, para el cálculo de la CF.

Todo lo anterior fue aplicado tanto a los datos originales filtrados (Fig. 3.5a), como a los de la potencia instantánea (Fig. 3.5b).



a).



b).

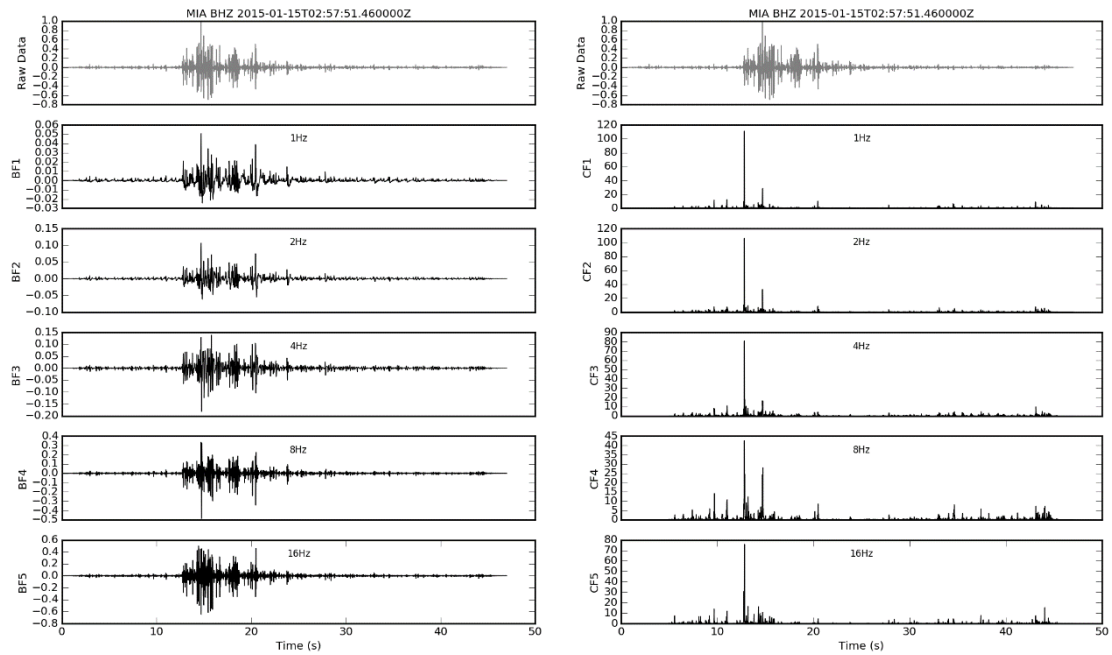


Figura 3.5.- Ejemplo de un sismograma, en la componente Z de la estación MISA, en su forma a) original filtrado en sus cinco bandas (BF1-FB5) y sus respectivas CF(CF<sub>1</sub>-CF<sub>5</sub>); y b) el mismo evento, pero realzado con la potencia instantánea, sus 5 sub-bandas y su respectiva CF.

### 3.4.2 Método basado en procesos autorregresivos y el criterio de información Akaike(AIC)

Un proceso autorregresivo consiste que cada muestra ' $x_t$ ' puede ser predicha a partir de las anteriores, con un cierto error (que se trata de minimizar) y matemáticamente se expresa como:

$$x_t = \sum_{m=1}^M a_m x_{t-m} + e_t \quad (3.4)$$

Donde:  $M$ , es el orden del modelo AR;  $a_m$  ( $m=1, \dots, M$ ) es el coeficiente de predicción; y  $e_t$  es el error (residuo) de predicción. Entonces este modelo AR permite caracterizar toda la serie únicamente multiplicando por el coeficiente  $a_m$ , excepto el error de predicción. No obstante, lo que se hace es minimizar  $e_t$  para que el modelo AR tenga el mejor ajuste posible. esto equivale a considerar la serie como una suma de una parte causal y otra aleatoria, que tiende a tener una distribución normal.

Para estimar el arribo de la onda P se asume que, dos segmentos del sismograma (antes y después de la llegada de la señal sísmica), son estacionarios; por lo tanto, pueden ser modelados mediante un proceso AR utilizando la ecuación 3.4. Sin embargo, la estimación del orden de un proceso AR es casi siempre muy imprecisa, de ahí que la mayoría de aplicaciones AR fijan una orden utilizando métodos de ensayo y error (Leonard y Kennett, 1999). Seguidamente se calculan los errores o residuos de predicción para ambos segmentos; y para terminar el mínimo global de la función AIC para los dos modelos, es el indicador del tiempo de arribo.

La función AIC de un sismograma, para dos segmentos de señal, pueden representarse como una función punto división (Sleeman y van Eck, 1999), tal como:

$$AIC(P) = (P - M) \log(\sigma_{1,max}^2) + (N - M - P) \log(\sigma_{2,max}^2) + Const. \quad (3.5)$$

Donde  $P$  es el punto de división,  $M$  es el orden del modelo AR,  $N$  es la longitud total de los datos,  $\sigma_{1,max}^2$  y  $\sigma_{2,max}^2$  indican la varianza del sismograma en los dos segmentos.

Otro enfoque que no utiliza un modelo AR, lo determina la función AIC directamente del sismograma (Maeda, 1985).

$$AIC(P) = P \log\{var(x[1, P])\} + (N - P - 1) \log\{var(x[P + 1, N])\} \quad (3.6)$$

Donde el punto de división  $P$ , se extiende sobre toda la muestra del sismograma ( $x$ ). Si lo que se busca en el sismograma es sólo la onda P, utilizando este último enfoque, se puede determinar fácilmente el inicio de fase buscando el mínimo global de la función AIC.

Sin embargo, si el sismograma también contiene la onda S, difícilmente identificará ambas fases a la vez si se decide tomar solo el mínimo valor global de la función AIC. Entonces como alternativa (más eficiente), para determinar la llegada de la onda P, es utilizar el valor absoluto de la primera derivada de la función AIC como CF, porque la primera derivada de la función AIC es más sensible a cualquier cambio (Chen y Holland, 2016).

Para nuestros datos, empleamos el segundo enfoque y criterio para estimar de manera más eficiente la CF. No fue necesario ningún parámetro de entrada, más que el mismo sismograma y fue aplicado tanto a los datos originales (filtrados), como a los de la potencia instantánea; como se muestra en la Figura 3.6 a continuación.

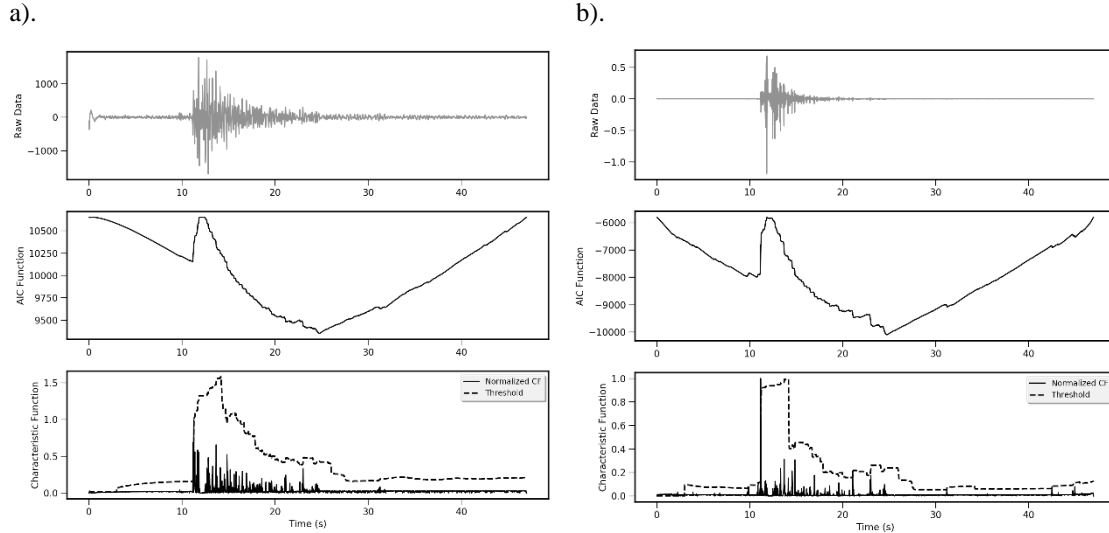


Figura 3.6.- Ejemplo cálculo de la CF, utilizando el método AIC, para un sismo local en la componente Z de la estación MISC;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea.

### 3.4.3 Método basado en estadísticos de orden superior

La curtosis  $K$  es un estadístico de orden superior de una variable y que, para este tipo de estudios, ha sido de gran utilidad para identificar automáticamente el inicio de la onda P y S (p. ej. Panagiotakis et al., 2008; Baillard et al., 2014; Hibert et al., 2014). Se define como:

$$K = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2} = \frac{\mu_4}{\sigma^4} \quad (3.7)$$

Donde:  $E$  es el operador de expectación,  $X$  son las mediciones,  $\mu$  es la media de las mediciones,  $\mu_4$  es el cuarto momento sobre la media y  $\sigma$  es la desviación estándar. La forma discreta de la ecuación 3.7 puede escribirse como

$$K = \frac{\frac{1}{n} \sum_{i=1}^{n+1} (X_i - \bar{X})^4}{\left(\frac{1}{n} \sum_{i=1}^{n+1} (X_i - \bar{X})^2\right)^2} - 3 \quad (3.8)$$

Donde:  $n$  es el número total de muestras,  $X_i$  es el valor de la  $i$ -ésima muestra y  $\bar{X}$  es la media de las  $n$  muestras.

El ruido blanco usualmente es una distribución normal (gaussiana), cuya curtosis es 3 (p. ej. De Carlo, 1997). Los valores de  $K > 3$  resulta una concentración máxima (apuntamiento) comparada con la distribución gaussiana, mientras que los valores  $< 3$  dan como resultado una concentración menor que la distribución gaussiana.

Por lo tanto, la curtosis caracteriza a la señal basándose en la forma o tipo de distribución, en lugar de la SNR. Entonces la curtosis puede funcionar muy bien para las señales con baja SNR (Chen y Holland, 2016).

En la práctica, la curtosis se estimó con una ventana móvil de 1 segundo, esta avanza muestra a muestra y se le asigna su valor, al final de la ventana. Y es en base a ello que

se construyó la CF para nuestros datos, en sus dos versiones antes mencionadas, como se muestra en la Figura 3.7.

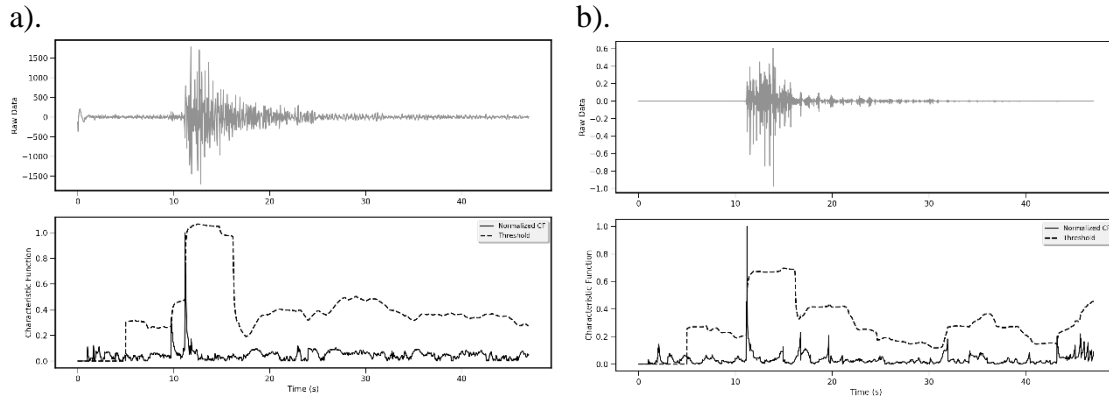


Figura 3.7.- Ejemplo cálculo de la CF, utilizando la curtosis (ventana de 1 segundo). Para un sismo local en la componente Z de la estación MISF; (a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea.

### 3.4.4 Método basado en el periodo predominante

Este método, en un primer momento, se introdujo para poder estimar el periodo predominante ( $T^p$ ) de la velocidad del suelo, como un indicador rápido de la para estimar magnitud de terremotos en sistemas de alerta temprana (p. ej. Allen y Kanamori, 2003; y Lockman y Allen, 2005). Con el tiempo se observaron cambios notorios de la función  $T^p$  coincidentemente con el inicio de la fase P en los terremotos, pero solo en algunos casos. Para mejorar la estabilidad de la estimación de la onda P, Hildyard et al. (2008) introducen el periodo predominante amortiguado  $T^{pd}$ , el cual se calcula de la siguiente manera:

$$T_i^{pd} = 2\pi \sqrt{\frac{X_i}{D_i + D_s}} \quad (3.9)$$

Las variables:  $X_i$  y  $D_i$  se definen como

$$X_i = \alpha D_{i-1} + x_i^2, \quad D_i = \alpha D_{i-1} + \left(\frac{dx}{dt}\right)_i^2 \quad (3.10)$$

También se pueden expresar en su forma no recursiva como

$$X_i = \sum_{j=0}^i \alpha^{i-j} x_j^2, \quad D_i = \sum_{j=0}^i \alpha^{i-j} \left(\frac{\partial x}{\partial t}\right)_j^2 \quad (3.11)$$

$D_s$  es una constante que inicialmente fue introducida para estabilizar la función  $T^p$ ,  $x$  es el segmento de señal sísmica filtrada y  $\alpha$  es una constante de amortiguación. El  $T^{pd}$  se calcula para toda la señal sísmica, entonces se entiende que el periodo dominante del ruido y la onda P deben ser diferentes.

Los parámetros que influyen en la sensibilidad del método  $T^{pd}$  son:  $\alpha$  y  $D_s$ . El primero, controla el amortiguamiento de las funciones  $X_i$  y  $D_i$  (Hildyard et al., 2008), en otras palabras, como influye la forma de onda en el cálculo del periodo predominante.

Hildyard et al. (2008) especifica una ventana de tiempo  $\tau_w$ , en el cual la función de amortiguación alcanza (arbitrariamente) un valor de 0.1, de tal manera que

$$\alpha = e^{\frac{\ln(0.1)}{\tau_w/\Delta t}} \quad (3.12)$$

Donde,  $1/\Delta t$  es la frecuencia de muestreo.

El segundo controla el amortiguamiento del periodo predominante y está relacionado con el nivel de SNR. Una alternativa para el cálculo de  $D_s$  basado en este último, es la especificación de un nivel o un mínimo periodo esperable  $\tau_{mx}$ , antes de la llegada de fase.

$$D_s = \frac{4\pi^2 \langle x_i^2 \rangle \tau_w}{\tau_{mx}^2} \quad (3.13)$$

Donde  $\langle x_i^2 \rangle$  expresa la media del nivel de ruido al cuadrado.

Para los datos del Misti, los parámetros óptimos para el cálculo de la CF ( $T^{pd}$ ) fueron: una ventana de tiempo  $\tau_w$  de 5 segundos y un periodo mínimo  $\tau_{mx}$  de 0.04. En la Figura 3.8, se muestra un ejemplo de aplicación del método  $T^{pd}$  a los datos originales filtrados y a los realzados con la potencia instantánea.

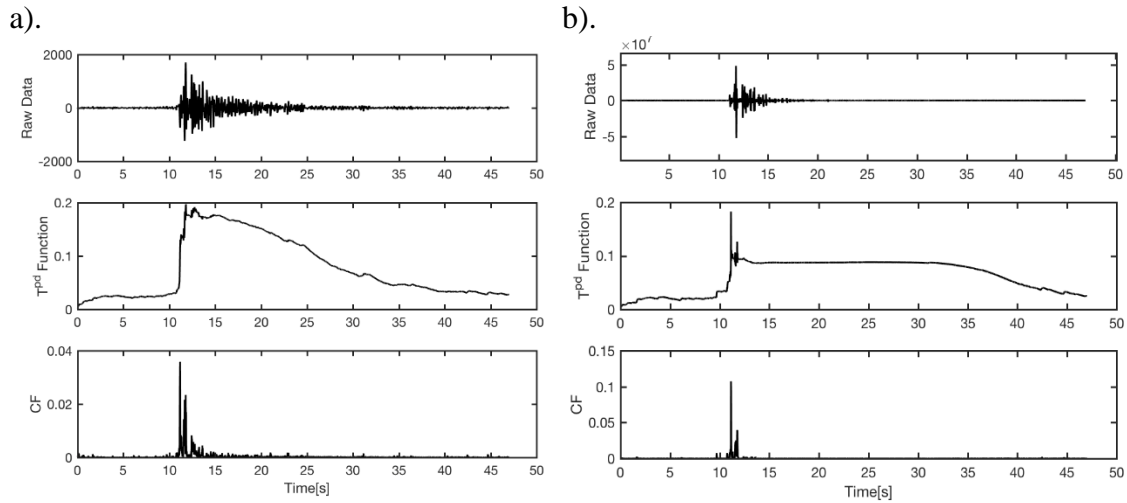


Figura 3.8.- Ejemplo cálculo de la CF, utilizando el periodo predominante amortiguado  $T^{pd}$ . Para un sismo local en la componente Z de la estación MISG; (a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea.

### 3.4.5 Método AMPA

AMPA por sus siglas en inglés “Adaptive Multiband Picking Algorithm” (Alvarez et al., 2013), es un algoritmo de *picking* automático reciente. Su estrategia está enfocada en determinar el tiempo de llegada de la fase P para señales fuertemente afectadas por ruido sísmico no-estacionario, debido, p. ej., a ráfagas de viento, actividad humana, etc. Para ello, AMPA trabaja en dos etapas.

En la primera etapa, se define una ventana de señal filtrada  $S_v(n)$ , en la que se espera una llegada de fase P.  $S_v(n)$  se divide en "k" bandas de análisis a través de filtros pasa-banda. Para cada sub-banda se calcula una envolvente como la amplitud de su función analítica (transformada de Hilbert) y se le aplica un umbral para reducción de ruido (Fig. 3.9a). En el panel derecho de la Figura 3.9b, se muestra un ejemplo de seis envolventes y su respectiva sub-banda ( $z_1, \dots, z_6$ ) después de aplicar el umbral de reducción de ruido. La señal final de esta primera etapa, se denominada ' $z_{tot}$ ' y es el logaritmo de la suma de todas las envolventes mejoradas y normalizadas (Fig. 3.9).

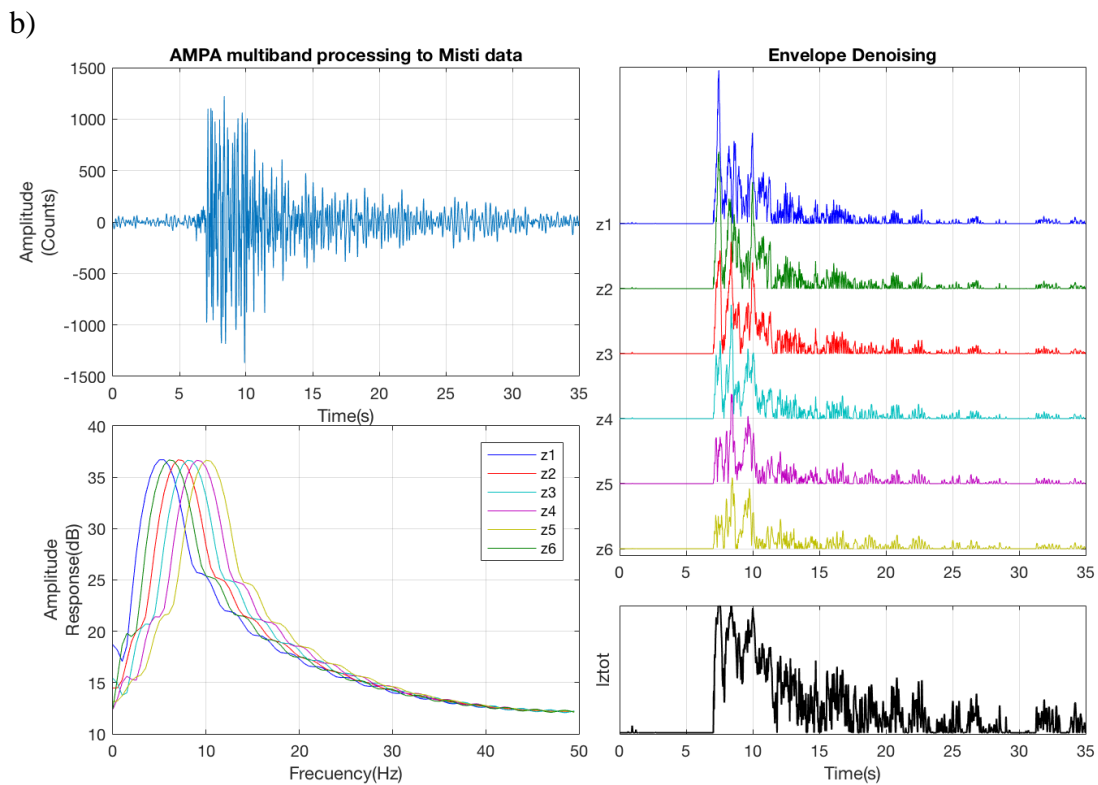
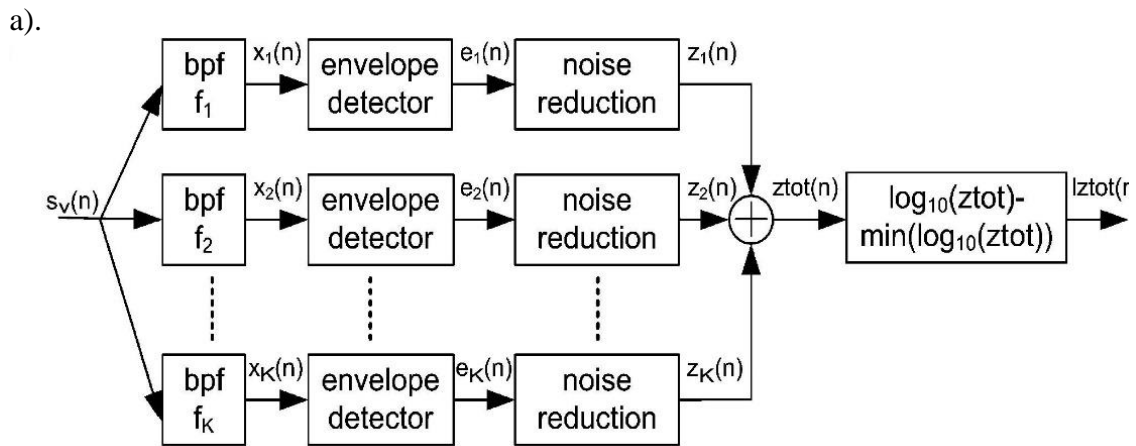


Figura 3.9.- a) Diagrama de bloques de un proceso multibanda adaptativo (AMPA) (Alvarez et al., 2013); y b) un ejemplo de eliminación del ruido multibanda a un sismograma de la estación MISE, así como la mejora final de su envolvente, ' $z_{tot}$ '.

En la segunda etapa de detección, se busca una forma de envolvente típica de fase P. Para ello, a la señal final ( $lztot$ ) se somete a un conjunto de filtros de realce, diseñados siguiendo los principios de la teoría de la comunicación, como detector óptimo. La respuesta impulsional resultante, debería ser la inversión temporal de la señal esperada ( $h_a(n)$ ). Para evitar la influencia de ruidos emergentes o impulsivos, la respuesta impulsional se compensa con un valor negativo, dando como resultado la respuesta impulsional final  $h_b(n)$  (Fig. 3.10a).

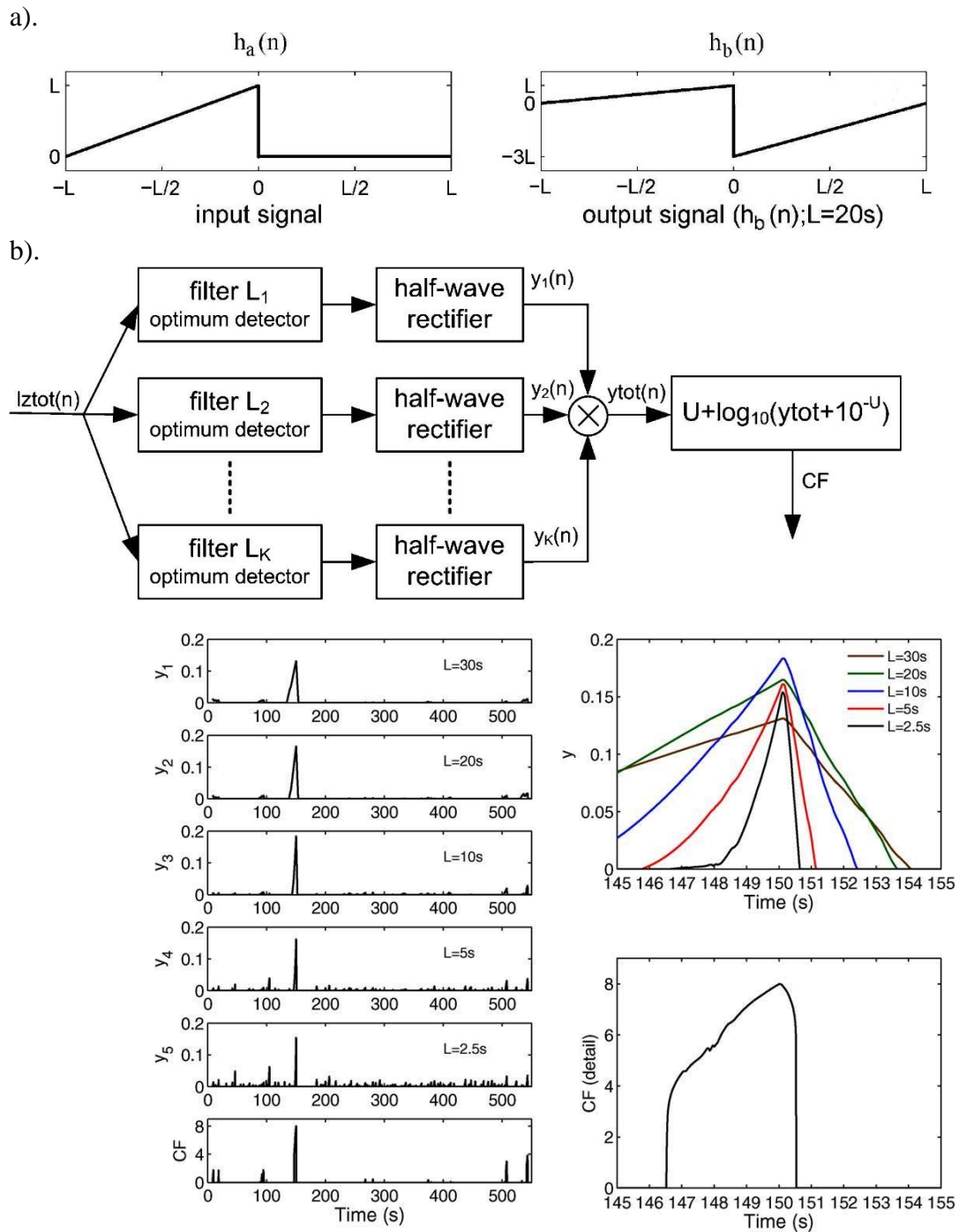


Figura 3.10.- a) Respuesta impulsional teórica para un detector óptimo, para un evento impulsivo y de decaimiento suave, así como la compensación a la respuesta impulsional anterior; b) diagrama de bloques describiendo la mejora de la envolvente filtrada para diferentes ventanas de longitud  $L$ , y el cálculo de la CF (Alvarez et al., 2013).

Para terminar, definimos la función característica final (CF) que resulta después de eliminar el ruido de las envolventes multibanda y mejorar las llegadas impulsivas con decaimientos suaves. Finalmente, el tiempo de llegada de la fase P se detectará como el instante de tiempo cuando CF toma la amplitud máxima (Fig. 3.10b).

Una ventaja de este método es que hace automáticamente dos evaluaciones de calidad para definir un rango de fiabilidad de los resultados. En primer lugar, la amplitud máxima de la función característica CF nos da una idea rápida de la impulsividad de la onda P. En segundo lugar, evalúa una SNR alrededor del tiempo de llegada de la onda P.

En comparación con otros métodos, AMPA (Álvarez et al. 2013, Romero et al. 2016) tiene un coste computacional relativamente bajo, sin configuración compleja y rendimiento efectivo en entornos ruidosos. Un ejemplo de aplicación del mismo, para los datos del Misti se muestra en la Figura 3.11. El cual no requiere más que la señal misma.

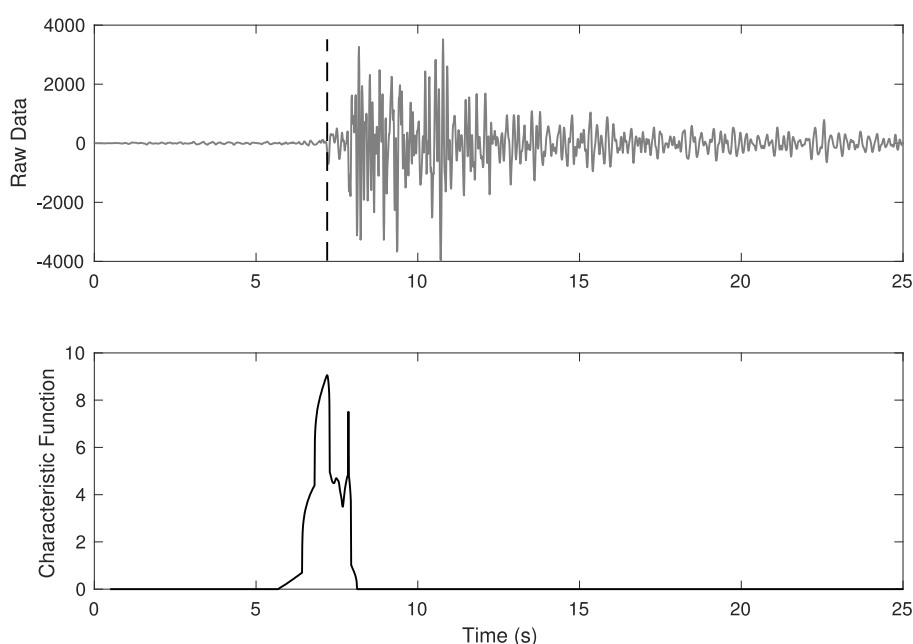


Figura 3.11.- Cálculo de la CF, utilizando el método AMPA, para un sismo local en la componente Z de la estación MISD.

### 3.4.6 Método AR-K

Este enfoque novedoso se basa en la idea de la estacionariedad del ruido sísmico, en ausencia de un terremoto (de cualquier origen) y que se puede modelar mediante un proceso autorregresivo (AR); se asume que el residuo o error de predicción estimado es aleatorio y tiene una distribución gaussiana. Entonces, si utilizamos la “curtosis” como CF de tal residuo, este será más sensible a la llegada de la onda P, que solo a través del sismograma filtrado.

Esto aprovechando la ventaja que tiene la curtosis de identificar transiciones entre distribuciones gaussianas y no gaussianas en una señal, sobre todo, en registros altamente afectados por ruido sísmico; incluso, es muy sensible incluso a llegada de ondas P emergentes. Entonces si este método se combina con un sofisticado algoritmo de *picking* (p. ej. Kuperkoch et al., 2010), se puede aprovechar aún más la CF y obtener excelentes resultados.

Muchos modelos AR han sido ampliamente utilizados en sismología, por ejemplo: en el análisis espectral de señales sísmicas, reducción del ruido sísmico de fondo y como se mencionó



anteriormente, en técnicas de *picking* automático. Algunos de estos métodos, sobre todo el propuesto por Takanami y Kitagawa, mostró resultados exitosos de *picking* automático para sistemas de localización de sismos (Takanami y Kitagawa, 1988;1991). Sin embargo, una desventaja de este es que, si a priori, la estimación del tiempo de inicio no es lo suficientemente precisa, obliga al método a buscar el mismo en una ventana de tiempo más amplia, haciendo de este un algoritmo lento debido su alto coste computacional (Alvarez et al., 2013).

Actualmente, herramientas de cálculo modernas, como Matlab o Python, hacen que algoritmos complejos como el Takanami y Kitagawa sean más rápidos, incluso, en tiempo real. Este nuevo enfoque trata de optimizar el cálculo de un proceso AR, para el cálculo del error de predicción de la ecuación 3.4, y sobre este, aplicar la curtosis como la variable de entrada en la ecuación 3.8.

El proceso se inicia con una ventana de señal inicial  $w_z(n)$ , se le aplica un filtro pasa-banda (según la banda espectral que corresponda por estación) y como resultado una señal filtrada  $x(n)$ . Luego, de esta última se toma una ventana de 6 segundos que solo contenga ruido y de ésta, se calculan los coeficientes de predicción lineal  $a_i$  ajustando un filtro de predicción lineal (LPC) de orden 5; como resultado se obtienen 6 coeficientes. Seguidamente, para modelar una señal  $w_m$ , el primero de los 6 coeficientes debe hacerse 0, para evitar el valor actual de la señal y saber cómo predice el modelo usando solo los valores anteriores. Y para finalizar esta primera etapa, de la diferencia entre la señal modelada  $w_m$  y la señal filtrada  $x(n)$ , se obtiene una señal  $e_j$  construida del error de predicción del modelo AR (Fig. 3.12a).

La segunda etapa del método, consiste en construir la CF utilizando la curtosis (ecuación 3.8) a partir de la señal  $e_j$ .

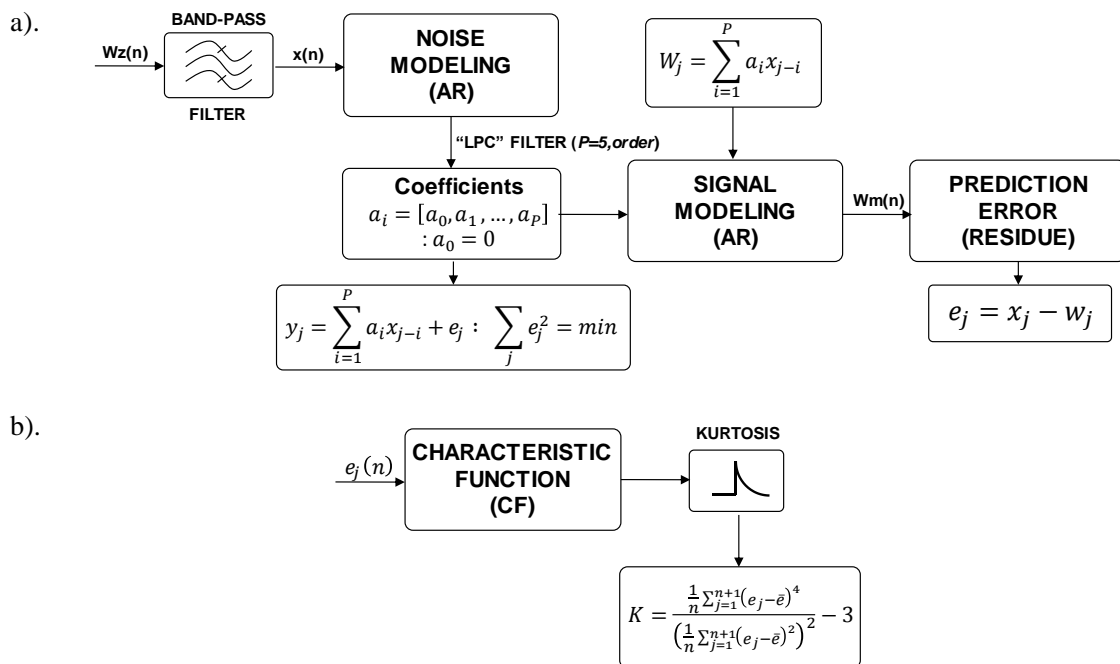
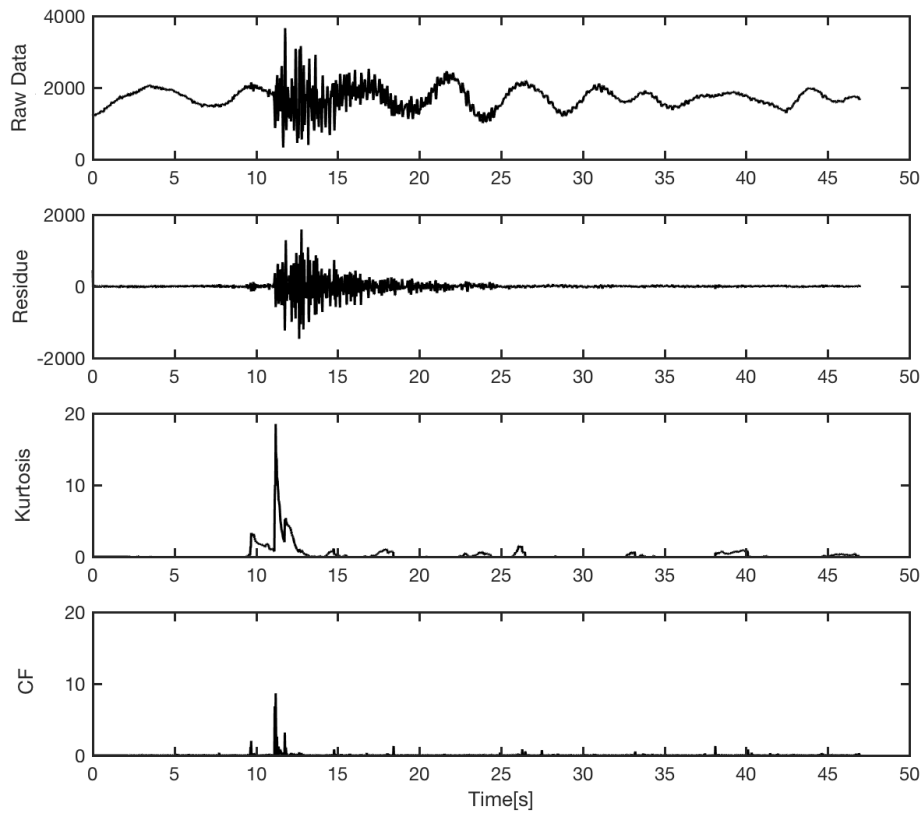


Figura 3.12.- Diagrama de bloques que ilustra (a) la estimación del error de predicción de un proceso autorregresivo (AR) y b) la segunda parte del proceso que corresponde al cálculo de la CF, utilizando la curtosis.

Una aplicación del método, se muestra en la Figura 3.13, con datos originales filtrados y los de la potencia instantánea.

a).



b).

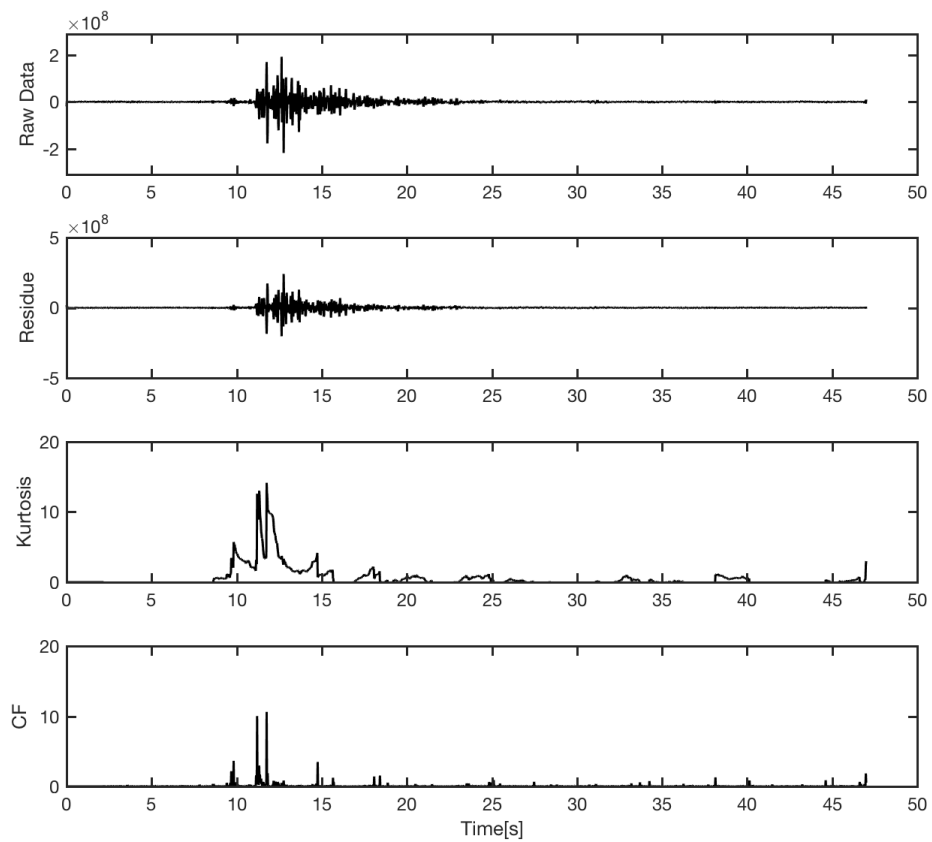


Figura 3.13.- Ejemplo cálculo de la CF, utilizando el método propuesto AR-K. Para un sismo local en la componente Z de la estación MISF;(a) en su forma original filtrado, y (b) su envolvente realzada, con la potencia instantánea.

### 3.5. Declaración de la onda P y filtrado de identificaciones falsas

Los métodos de energía, curtosis y AIC utilizan un *nivel umbral dinámico* para declarar la llegada de onda P, el cual se obtiene de la multiplicación del valor *rms* de la CF y una constante (*K*); la cual, es proporcional al número de desviaciones estándar de la CF. Entonces, el inicio de la onda P se declara cuando la CF excede dicho umbral. Este método, a menudo se aplica en procesos de control estadístico, como las cartas de control (Alwan y Roberts,1988). Esta técnica permite ajustar un nivel de umbral dinámico, al mismo tiempo que se calcula la CF.

Esta es una ventaja, frente a otros métodos recientes, como el de Lomax et al. (2012), donde se aplican umbrales estáticos y para ajustarlos, se necesita previamente una re-compilación del algoritmo (*script*). Por lo tanto, no permite ajustar la sensibilidad del método en función de los niveles cambiantes del ruido de los datos.

Una manera de elegir adecuadamente un nivel umbral dinámico óptimo, es a través del cálculo de SNR de la CF ( $SNR_{CF}$ ) y para ello se utilizó la siguiente expresión:

$$SNR_{CF} = 20 * \log\left(\frac{A}{A_0}\right) \quad (3.14)$$

Donde *A*, es el valor *rms* de la señal (ventana de 0.5segundos) y  $A_0$  el valor *rms* del ruido (ventana de 2 segundos). En la tabla 3.1, se muestran los valores de la constante *K* para cada rango de valores de SNR.

Tabla 3.1.- Valores de la constante *K* utilizados, en función del valor de la SNR.

$SNR_{CF}$	< 8	8 - 12	12-16	16 - 18	18 - 20	20 - 24	24 - 26	>26
<i>K</i>	2	3	4	5	6	7	8	12

El método AMPA declara la llegada la onda P, empleando solo el valor máximo de su CF; en cambio, los métodos  $T^{pd}$  y AR-K, consideran el valor máximo de la derivada de su CF.

Otro problema común entre los algoritmos de *picking* automático, son las “identificaciones falsas” (p. ej. detecciones que ocurren antes de la llegada de la onda P o mucho después de la coda). Todos los estudios intentan disminuir el número de identificaciones falsas, procurando siempre no perder sensibilidad en la identificación de fases reales. En nuestro caso, definimos una “identificación falsa” (IF) cuando la diferencia entre la detección automática y manual es mayor a un segundo. Esto debido a que los sismos del volcán Misti son muy locales y por lo general, están localizados entre 1km y 6km de la cima del volcán.

# Capítulo 4

## Resultados y su discusión

### 4.1. Capacidad de detección: Desempeño

Se aplicaron seis métodos de *picking* automático a una base de datos la cual tenía dos versiones: la primera, considerando los datos originales (primer conjunto de datos); y la segunda, los mismos solo que realizados mediante la potencia específica instantánea (segundo conjunto de datos).

Para evaluar los resultados de los métodos, se compararon 655 lecturas de onda P obtenidas automática y manualmente, en la componente vertical de las estaciones (MISA, MISC, MISD, MISE, MISF y MISG) de la red Misti.

Se evaluó, en primer lugar, la sensibilidad de métodos frente a su capacidad de detección, sin que se generen “identificaciones falsas”. Para el primer conjunto de datos (Tabla 4.1), el mayor número de identificaciones falsas corresponde a un 13% empleando el método de energía; en segundo lugar, el método de la curtosis, con un 8%; y para el resto de métodos, menos del 5%.

En general, la estación que mostro un mayor número de identificaciones falsas fue la MISG, con un máximo de 28 casos y utilizando el método de la curtosis; y 23, con el método de la energía. Por otro lado, las estaciones que no mostraron identificaciones falsas son: MISC, MISD y MISE, con el método AR-K; y MISA, con el método del  $T^{pd}$ . Las estaciones MISF y MISG mostraron pocas de identificaciones falsas, con el método AR-K.

Por lo general, el método AR-K es el que muestra el menor número identificaciones falsas (<1%), por lo tanto, es el que mejor detecta los eventos respecto a los demás métodos (99% de total). Esto demuestra que los parámetros del método se ajustan bien a los datos, incluso, en registros ruidosos. El segundo método en detectar la mayor cantidad de eventos es AIC, pero a diferencia del primero, es más lento debido su alto coste computacional. Los métodos computacionalmente más rápidos son AMPA y  $T^{pd}$ , ambos pueden procesar los datos en menos de 1 décima de segundo (Fig. 4.4).

Para el segundo conjunto de datos (Tabla 4.2), el número de identificaciones falsas se redujo ligeramente a un 12%, con el método de energía; mientras que, con el método de curtosis, estas aumentan considerablemente de un 8% al 22%; y con el resto de métodos, disminuyen ligeramente de un 5% a 3%.

El número de identificaciones falsas se acrecienta, notablemente, en las estaciones MISG, MISF y MISA utilizando la curtosis. Por otro lado, estas disminuyen notoriamente en las estaciones MISC, MISD, MISE y MISF, con los métodos AIC, AMPA y  $T^{pd}$ ; haciendo de estas cero, en algunos casos. Por lo tanto, se demuestra que, para este tipo de datos, el método de la curtosis es más sensible a cambios en la amplitud.

Los parámetros de los métodos AIC, AMPA y  $T^{pd}$  se ajustan mejor a este tipo de datos; mientras que el método AR-K, también, pero resulta ser más eficiente con los datos originales. Sin embargo, dependiendo de la estación y el método, las identificaciones falsas pueden aumentar, pero también pueden disminuir. Por ejemplo, las detecciones falsas con el método AMPA con los datos originales de la estación MISD, fueron 7 casos, pero con los datos realzados ninguno. Por otro lado, utilizando el mismo método y cambiando la estación a MISA, el número de identificaciones falsas aumentan ligeramente de 1 a 2 casos. Si analizamos los registros de ambas estaciones (Fig. 4.3), se puede observar que la llegada de la onda P en la estación MISD es mucho más impulsiva, que la observada en MISA. Esto sugiere que la potencia instantánea funciona mejor con llegadas de onda P de tipo impulsiva.

## 4.2. Calculo del error: precisión

Para estimar la exactitud y precisión de cada uno de los seis métodos aplicados, se calculó el valor absoluto de la diferencia entre las lecturas de llegada de onda P obtenidas automática y manualmente ( $|T_{auto} - T_{manual}|$ ), este valor en adelante lo definimos como “residuo”. Además de este, utilizaremos como indicadores de exactitud y precisión de los métodos: la media ( $\mu$ ) y la dispersión de los residuos ( $\sigma$ ).

En la figura 4.1, se muestran histogramas con la diferencia entre el *picking* automático y manual obtenido con los seis métodos, para los datos en sus dos versiones. En la misma, se puede observar que no existe ninguna tendencia respecto al valor central (cero) y, muy por el contrario (sobre todo en las estaciones MISA, MISD y MISE) los métodos muestran que aproximadamente más del 80% de los residuos, están muy cerca del valor central.

Con el nuevo método AR-K se identificó automáticamente la llegada de la P dentro de un margen de error  $\pm 0.2s$  en 90% mediante el conjunto de datos originales y un 92%, con los datos. Luego destaca el método AMPA, con una efectividad de 81%, con los datos originales; y un 84%, con los datos realzados. Y por último  $T^{pd}$ , con una efectividad del 71%, con los datos originales; y un 85%, con los datos realzados.

En las tablas 4.1 y 4.2, se hace un resumen de las medias de los errores y el intervalo ( $\mu \pm \sigma$ ), que ilustra el comportamiento estadístico de los errores acorde con el método empleado, la estación y el tipo de datos. Para el primer conjunto de datos (Tabla 4.1), se destaca como primer resultado, dos residuos medios óptimos de  $0.02 \pm 0.02s$  y  $0.02 \pm 0.03s$ , obtenidos en las estaciones MISD y MISE utilizando el método AR-K, y al 0% de identificaciones falsas (IF).

El segundo mejor resultado, tiene un residuo medio de  $0.06 \pm 0.12s$  obtenidos en las estaciones MISD y MIE con el método de la Curtosis, al 1% de IF. Después particularmente, resaltan residuos medios de  $0.07 \pm 0.14s$  en MISA con los métodos AMPA y AIC, al 1% de IF. Por otro lado, se muestran residuos medios ligeramente grandes de  $0.14 \pm 0.21s$ , en MISA y  $0.13 \pm 0.14s$ , en MISC; ambos obtenidos con el método  $T^{pd}$ , al 0% de IF. Los residuos más grandes se presentan en las estaciones MISC y MISF, con un residuo medio máximo de  $0.30 \pm 0.22s$  en aproximadamente un 10% de IF.

Estos resultados son bastante buenos, considerando la dificultad que se tiene al tratar de identificar manualmente la llegada de la onda P, para señales sísmicas en ambientes volcánicos, sobre todo las que son de tipo emergentes (Fig. 4.3). Por lo tanto, el método AR-K además de ser el que mejor detecta los eventos, también es el “más preciso” respecto a los demás métodos, en la mayoría de estaciones de la red Misti (MISA, MISC, MISD y MISE).

Para el segundo conjunto de datos (Tabla 4.2), se muestran residuos medios mucho más bajos, del orden de  $0.01 \pm 0.02s$  obtenidos en las estaciones MISD y MISE, con el método AR-K y al 0% de IF. El segundo mejor resultado, tiene un residuo de  $0.04 \pm 0.07s$  en la MISD, con el método

AMPA y al 0% de IF. Luego destacan residuos medios de  $0.06 \pm 0.11$  s y  $0.07 \pm 0.14$  s en MISA, con los métodos  $T^{pd}$  y AMPA, respectivamente y al 1% de IF. La mayoría de los métodos mostraron algunas mejoras en la precisión de sus resultados.

Los residuos medios más altos tienen residuos medios de  $0.22 \pm 0.31$  s en MISG, con el método de la energía y al 7% de IF; y  $0.35 \pm 0.34$  s en MISF, con el método AIC y al 6% de IF. Esto se debe principalmente a que ambas estaciones son las que más ruido muestran, con una SNR por debajo de los 20dB en las lecturas de onda P, para ambos conjuntos de datos (Fig. 4.2).

Tabla 4.1.- Resultados de detección y precisión por estación, para los seis métodos utilizando los datos originales.

		MISA	MISC	MISD	MISE	MISF	MISG	Total IF
<b>Energía</b>	N / IF	137/11	74/15	134/16	109/9	68/9	50/23	83(13%)
	$\mu \pm \sigma(s)$	$0.09 \pm 0.15$	$0.17 \pm 0.19$	$0.07 \pm 0.12$	$0.08 \pm 0.12$	$0.11 \pm 0.18$	$0.15 \pm 0.20$	
<b>Curtosis</b>	N / IF	140/8	88/1	148/2	115/3	69/8	45/28	50(8%)
	$\mu \pm \sigma(s)$	$0.10 \pm 0.17$	$0.30 \pm 0.22$	$0.06 \pm 0.12$	$0.06 \pm 0.12$	$0.29 \pm 0.30$	$0.15 \pm 0.24$	
<b>AIC</b>	N / IF	142/6	88/1	149/1	115/3	73/4	62/11	26(4%)
	$\mu \pm \sigma(s)$	$0.06 \pm 0.12$	$0.22 \pm 0.23$	$0.06 \pm 0.15$	$0.06 \pm 0.14$	$0.22 \pm 0.31$	$0.13 \pm 0.22$	
<b>AMPA</b>	N / IF	147/1	86/3	143/7	107/11	75/2	65/8	32(5%)
	$\mu \pm \sigma(s)$	$0.07 \pm 0.14$	$0.16 \pm 0.17$	$0.11 \pm 0.21$	$0.11 \pm 0.21$	$0.12 \pm 0.18$	$0.13 \pm 0.19$	
<b><math>T^{pd}</math></b>	N / IF	148/0	89/0	136/14	106/12	76/1	68/5	32(5%)
	$\mu \pm \sigma(s)$	$0.14 \pm 0.21$	$0.13 \pm 0.14$	$0.15 \pm 0.15$	$0.17 \pm 0.22$	$0.18 \pm 0.22$	$0.18 \pm 0.22$	
<b>AR-K</b>	N / IF	147/1	89/0	150/0	118/0	76/1	70/3	5(<1%)
	$\mu \pm \sigma(s)$	$0.06 \pm 0.15$	$0.09 \pm 0.13$	$0.02 \pm 0.03$	$0.02 \pm 0.02$	$0.12 \pm 0.19$	$0.15 \pm 0.24$	
N.º de lecturas Manuales		148	89	150	118	77	73	655(100%)

N=número de detecciones, IF=número de identificaciones falsas,  $\mu$ =media,  $\sigma$ =desviación estándar.

Tabla 4.2.- Resultados de detección y precisión por estación, para los seis métodos utilizando los datos realizados.

Método		MISA	MISC	MISD	MISE	MISF	MISG	Total IF
<b>Energía</b>	N / IF	137/11	80/9	131/19	103/15	62/15	61/12	81(12%)
	$\mu \pm \sigma(s)$	$0.10 \pm 0.15$	$0.19 \pm 0.22$	$0.10 \pm 0.15$	$0.12 \pm 0.17$	$0.24 \pm 0.28$	$0.22 \pm 0.31$	
<b>Curtosis</b>	N / IF	119/29	72/17	134/16	108/10	45/32	31/42	146(22%)
	$\mu \pm \sigma(s)$	$0.09 \pm 0.17$	$0.24 \pm 0.23$	$0.08 \pm 0.12$	$0.09 \pm 0.15$	$0.31 \pm 0.33$	$0.16 \pm 0.26$	
<b>AIC</b>	N / IF	143/5	89/0	149/1	115/3	72/5	60/13	27(4%)
	$\mu \pm \sigma(s)$	$0.09 \pm 0.18$	$0.17 \pm 0.19$	$0.11 \pm 0.16$	$0.10 \pm 0.13$	$0.35 \pm 0.34$	$0.19 \pm 0.30$	
<b>AMPA</b>	N / IF	146/2	88/1	150/0	115/3	74/3	65/8	17(3%)
	$\mu \pm \sigma(s)$	$0.07 \pm 0.14$	$0.14 \pm 0.16$	$0.04 \pm 0.07$	$0.07 \pm 0.16$	$0.20 \pm 0.26$	$0.15 \pm 0.24$	
<b><math>T^{pd}</math></b>	N / IF	147/1	89/0	142/8	116/2	77/0	65/8	19(3%)
	$\mu \pm \sigma(s)$	$0.06 \pm 0.11$	$0.14 \pm 0.21$	$0.07 \pm 0.15$	$0.05 \pm 0.10$	$0.19 \pm 0.25$	$0.18 \pm 0.25$	
<b>AR-K</b>	N / IF	148/0	88/1	150/0	118/0	75/2	67/6	9(1%)
	$\mu \pm \sigma(s)$	$0.05 \pm 0.10$	$0.09 \pm 0.15$	$0.01 \pm 0.02$	$0.02 \pm 0.02$	$0.09 \pm 0.17$	$0.12 \pm 0.22$	
N.º de lecturas Manuales		148	89	150	118	77	73	655(100%)

N=número de detecciones, IF=número de identificaciones falsas,  $\mu$ =media,  $\sigma$ =desviación estándar.

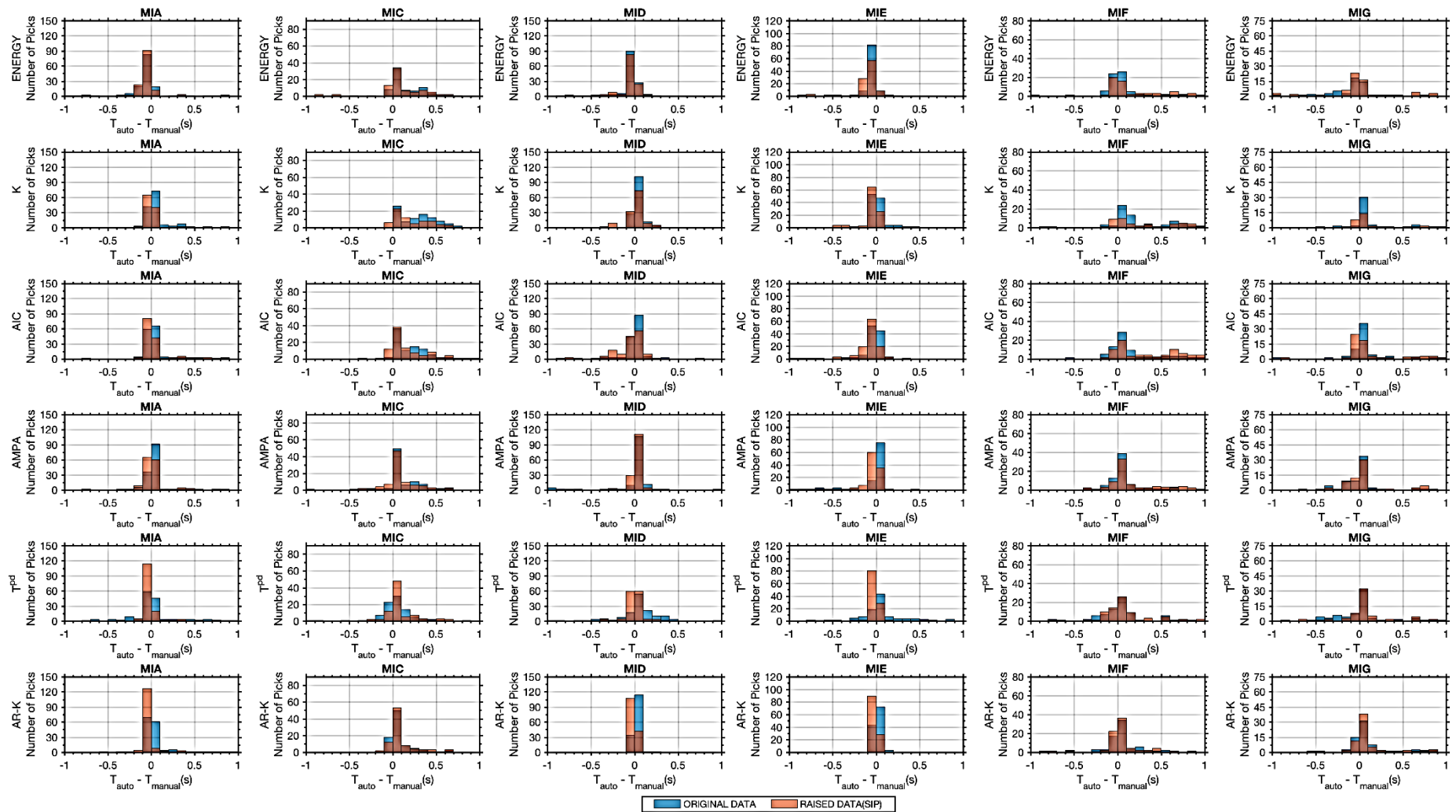


Figura 4.1.- Histogramas de la diferencia del picking automático y manual obtenidos con los seis métodos empleados, para los datos de las seis estaciones de la red Misti, en sus dos versiones (datos originales en azul y datos realizados en rojo).

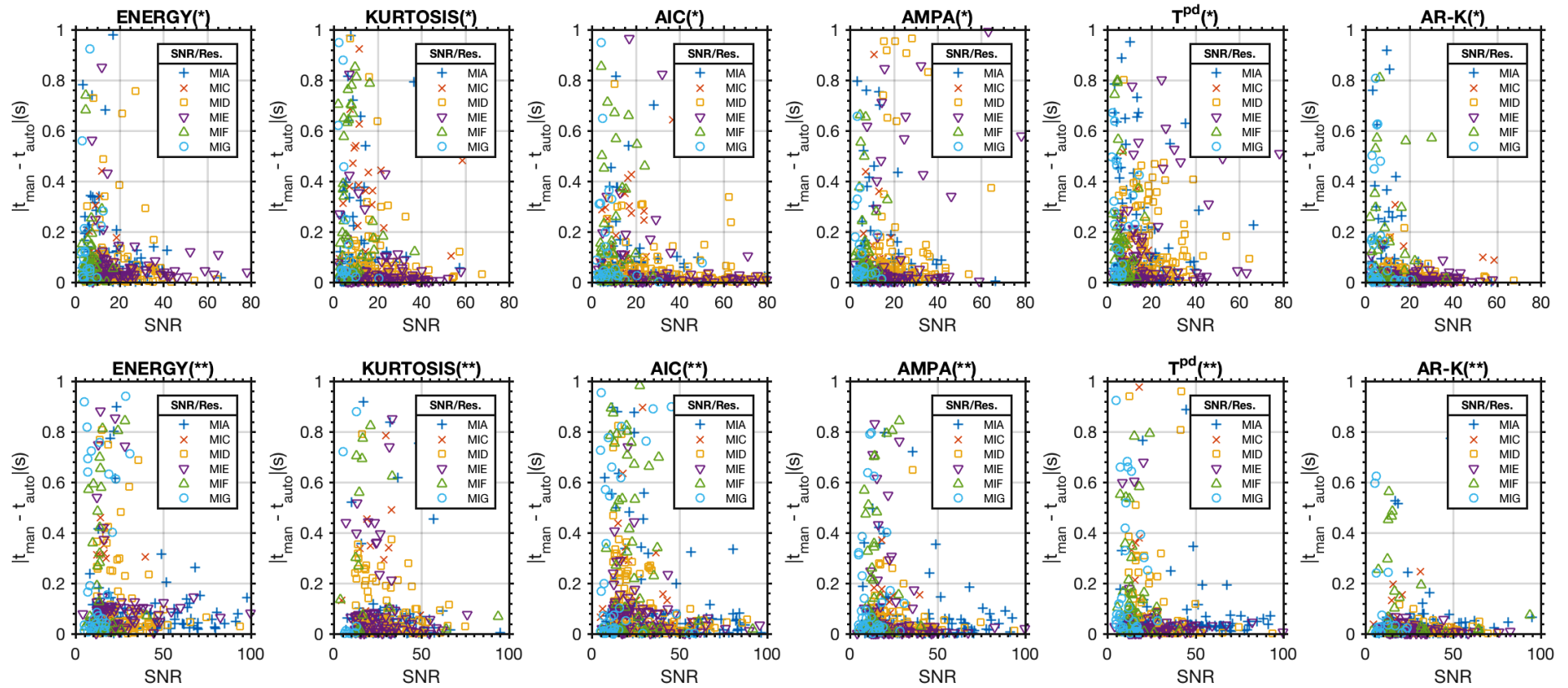


Figura 4.2.- Comparación del residuo entre los seis métodos de picking automático utilizados, en función de la SNR de cada lectura.



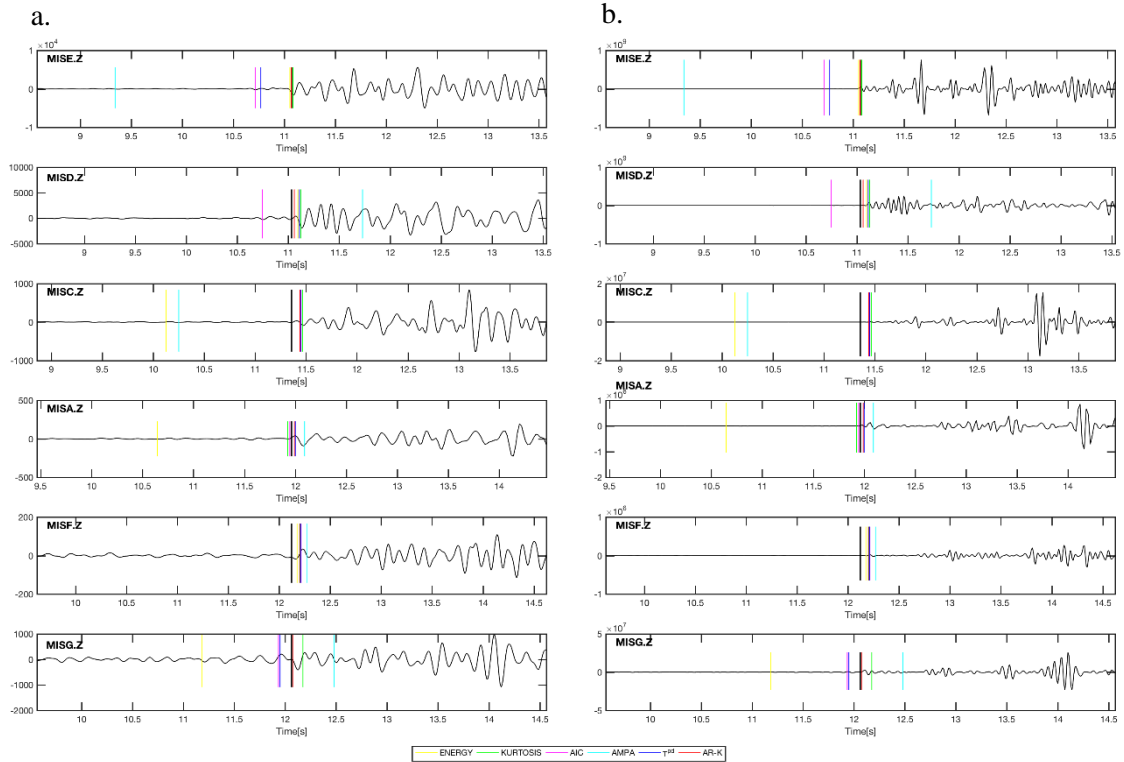


Figura 4.3.- Comparación de los seis métodos de picking automático de onda P frente al obtenido manualmente (línea de color negro), para un evento representativo registrado en la componente Z de todas las estaciones de la red Misti, en a) su versión original, y b) los mismos, pero utilizando la potencia específica instantánea. La línea en amarillo, corresponde a la onda P con el método de la energía; la verde, con la kurtosis; la magenta, con AIC; la celeste, con AMPA; la azul, con  $T^{pd}$ ; y la roja con el método AR-K.

Adicionalmente, mostramos en la figura 4.4 un análisis acerca del tiempo que tardan los métodos en procesar cada dato y tomando como referencia los datos de la estación MISD. Se resalta que el método más rápido de todos es AMPA y  $T^{pd}$ , con una duración menor a 1 décima de segundo en promedio, luego están los métodos de energía, AR-K, kurtosis y por último el método AIC. Con una duración intermedia esta AR-K, con menos de 3 décimas de segundo. Así mismo, los métodos de Energía, Kurtosis y AIC, fueron implementados en Python (Obspy); mientras, que los métodos AMPA,  $T^{pd}$  y AR-K, en Matlab.

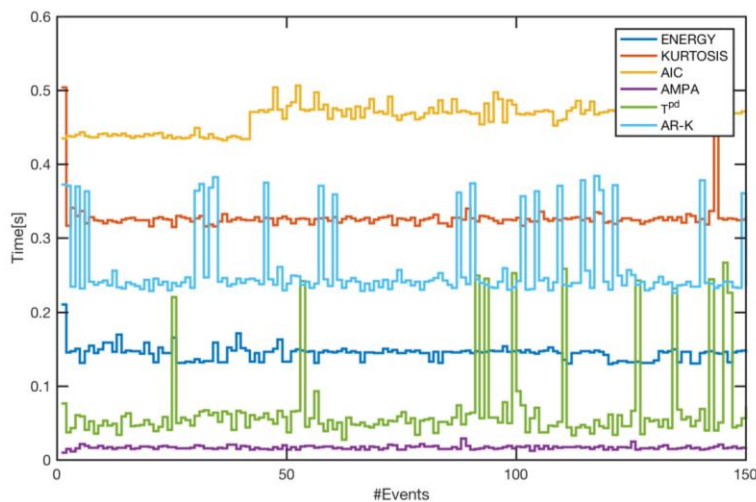


Figura 4.4.- Desempeño de los seis métodos de picking, en función del tiempo que se toman en procesar los 150 sismos seleccionados, tomando como referencia la estación MISD.

### 4.3. Discusión

En los últimos 30 años, los métodos de *picking* automático han ido mejorando sus enfoques en la detección y precisión de fases P y S, pero la mayoría solo están pensados para eventos sísmicos de magnitud importante. En realidad, son muy pocos los trabajos que tratan de detectar fases sísmicas en micro-terremotos debido a que, sencillamente, son muy difíciles de detectar, y más aún en sismos de origen volcánico. Por esta razón, que el primer reto planteado en este estudio, fue solo tratar de identificar la onda P de manera automática y precisa.

En el presente estudio, se resalta el buen desempeño del método propuesto AR-K, el cual demostró ser un método robusto en términos de detección y precisión, frente al resto de métodos convencionales utilizados, como: los basados en la energía, curtosis y AIC. Así, como otros métodos recientes basados en el periodo predominante amortiguado  $T^{pd}$  y en el análisis multibanda AMPA.

La implementación de AR-K es algo novedosa, y nació de la idea de combinar el error de predicción de un modelo autorregresivo (AR) al ruido de la señal, y de la señal resultante (residuo) aplicar el método de la curtosis, como CF. Este último, además de ser sensible a cambios de amplitud, demostró que también detecta llegadas débiles o emergentes.

Como en muchos trabajos de *picking* automático (p. ej. Nippres et al., 2010), se recomienda combinar el método AR-K con otros métodos como AMPA y  $T^{pd}$ , debido a que estos dos últimos son mucho más rápidos que AR-K. De esta manera, se puede conseguir dos cosas: rapidez y precisión.

Otra novedad en este trabajo, fue aplicar la potencia específica instantánea a la componente vertical de las estaciones de la red Misti, con el objetivo de resaltar la llegada de la onda P. En muchas ocasiones funciona muy bien, sobre todo con los métodos AMPA,  $T^{pd}$  y AIC, para señales de regular a buena SNR ( $>20\text{dB}$ ). Pero en otros casos, donde la llegada de la onda P era de tipo emergente y de baja frecuencia (MISF), la atenuaba por completo. Con el método AR-K, se obtuvo ligeramente un número menor de detecciones de onda P, que con los datos originales. Esto se debe principalmente, a que la potencia no necesariamente se ajusta a un modelo autorregresivo

## Conclusiones

- Se han analizado seis métodos para la identificación automática de fases P, aplicados a 150 eventos VT del volcán Misti, con 655 trazas sísmicas. Tres de ellos son convencionales y se basan, respectivamente en el análisis de la energía, la curtosis y el criterio de información Akaike (AIC). Dos métodos más recientes usan el periodo predominante amortiguado ( $T^{pd}$ ), y el análisis de envolvente multi-banda (AMPA).
- Se propone un método novedoso, al que se le denomina “AR-K”. Este método es combinación de un análisis autorregresivo de la señal y la curtosis del error de predicción como CF.
- Se han construido dos conjuntos de datos: uno con los registros originales y otro con los mismos realizados mediante la potencia instantánea. Los métodos de energía y curtosis aplicados al primer conjunto de datos muestran el máximo de identificaciones falsas (13% y 8%, respectivamente). Los 4 métodos restantes mostraron menos de 5% de identificaciones falsas.
- El nuevo método aquí propuesto (AR-K) presenta el mínimo de identificaciones falsas: 0% en dos estaciones (MISE y MISD). Este método es el que detecta mejor las llegadas de P, con un 99% de aciertos, un éxito sin precedentes en las referencias sobre el tema.
- Para el segundo conjunto de datos, el método de la curtosis muestra hasta un 22% de identificaciones falsas, principalmente en las estaciones MISF y MISG. Sin embargo, los métodos AIC, AMPA y  $T^{pd}$ , se ajustan mejor a este tipo de datos, disminuyendo ligeramente el número de identificaciones falsas.
- Con el nuevo método AR-K se identificó automáticamente la llegada de la P dentro de un margen de error  $\pm 0.2s$  en un 92% de las trazas realizadas con la potencia y un 90% con las trazas originales. En segundo lugar, destaca el método AMPA, con una efectividad del 81% con los datos originales y un 84%, con los datos realizados.  $T^{pd}$  alcanza una efectividad del 71%, con los datos originales y un 85%, con los realizados.
- El método AR-K, además de ser el que mejor detecta los datos, también demuestra ser el más preciso, respecto al *picking* manual, con un residuo medio  $\mu = 0.02 s$  y  $\sigma = 0.02s$ , con los datos originales; y  $\mu = 0.01s$  y  $\sigma = 0.02s$ , con los datos realizados; sobre todo en las estaciones MISD y MISE. Los residuos medios relativamente más altos ( $0.22 \pm 0.31s$ ) se obtienen en MISG, la más ruidosa, con el método de la energía y  $0.35 \pm 0.34s$  en MISF, con el método AIC.
- Se propone una combinación de los métodos AR-K, AMPA y  $T^{pd}$ , para la identificación automática de la onda P para los sismos del volcán Misti. Debido a que estos dos últimos son mucho más rápidos que AR-K y de esta manera, conseguir rapidez y precisión.

## Referencias

- Akansu, A. N., W. A. Serdijn, and I. W. Selesnick (2010). Emerging applications of wavelets: A review, *Phys. Comm.* 3, no. 1, 1–18.
- Aki, K. (1992). State of the art in volcanic seismology. In: *Gasparini, P., Scarpa, R., Aki, K. (Eds.), Volcanic Seismology. IAVCEI Proc. Volcanol.* 3, 3-10.
- Alguacil, G., Vidal, F. (2012). Medidas Instrumentales De La Intensidad Del Movimiento Del Suelo. Aplicación A Terremotos Europeos. En **Homenaje A La Profesora María Dolores Romacho Romero**, pp 15-32. Universidad De Almería.
- Allen, R. (1982). Automatic phase pickers: Their present use and future prospects, *Bull. Seismol. Soc. Am.* 72, S225–S242.
- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* 68, no. 5, 1521–1532.
- Allen, R. and Kanamori, H. (2003). The potential for earthquake early warning in southern California, *Science*, 300, 786–789.
- Allen, R.V. (1978). Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.* 68, 1521–1532.
- Alvarez, I., Garcia, L., Mota, S., Cortes, G., Benitez, C., de la Torre, A. (2013). An automatic P-phase *picking* algorithm based on adaptive multiband processing. *Geosci. Remote Sens. Lett.*, IEEE 10, 1488–1492.
- Alwan, L. C., and H. V. Roberts (1988). Time-series modeling for statistical process control, *J. Bus. Econ. Stat.* 6, no. 1, 87–95.
- Ambuter, B. P. and Solomon, S. C. (1974). An event-recording system for monitoring Small earthquakes. *Bull. Seis. Soc. Am.* 64, 1181-1188.
- Anant, K. S., and F. U. Dowla (1997). Wavelet transform methods for phase identification in three-component seismograms, *Bull. Seismol. Soc. Am.* 87, no. 6, 1598–1612.
- Baer, M., and U. Kradolfer (1987). An automatic phase picker for local and teleseismic events, *Bull. Seismol. Soc. Am.* 77, no. 4, 1437–1445.
- Baillard, C., W. C. Crawford, V. Ballu, C. Hibert, and A. Mangeney (2014). An automatic kurtosis-based P- and S-phase picker designed for local seismic networks, *Bull. Seismol. Soc. Am.* 104, no. 1, 394–409.
- Bogiatzis, P., and M. Ishii (2015). Continuous wavelet decomposition algorithm for automatic detection of compressional- and shear-wave arrival times, *Bull. Seismol. Soc. Am.* 105, no. 3, 1628–1641.
- Bormann, P., Engdahl, B., Kind, R. (2012): Seismic Wave Propagation and Earth models. - In: Bormann, P. (Ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP2)*, **Potsdam: Deutsches GeoForschungsZentrum GFZ**, pp. 1—105.
- Centeno, R. (2010): Análisis de la actividad sísmica del Volcán Misti entre octubre 2005 a diciembre 2008 y su dinámica interna actual. (**Tesis de grado**), **Universidad Nacional San Agustín de Arequipa**. 75 p.
- Chen, C., and A. A. Holland, (2016). PhasePAPy: A Robust Pure Python Package for Automatic Identification of Seismic Phases, *Seismological Research Letters*, 87(6).
- Christoffersson, A., Husebye, E.S., Ingate, S.F., 1988. Wavefield decomposition using ML-probabilities in modelling single-site 3-component records. *Geophys. J. Int.* 93, 197–213.

- Chu, C.-K.P., Mendel, J.M., 1994. First break refraction event *picking* using fuzzy logic systems. *Fuzzy Syst., IEEE Trans.* on 2, 255–266.
- Dai, H., and C. MacBeth (1995). Automatic *picking* of seismic arrivals in local earthquake data using an artificial neural network, *Geophys. J. Int.* 120, no. 3, 758–774.
- DeCarlo, L. T. (1997). On the meaning and use of kurtosis, *Psychol. Meth.* 2, no. 3, 292–307.
- Diehl, T., Kissling, E., Bormann, P. (2012): Tutorial for consistent phase *picking* at local to regional distances. - In: Bormann, P. (Ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, Potsdam : **Deutsches GeoForschungsZentrum GFZ**, pp. 1—21.
- Diehl, T., Kissling, E., Husen, S. and Aldersons, F. (2009). Consistent phase *picking* for regional tomography models: application to the greater Alpine region. *Geophys. J. Int.*, 176, 542-554.
- Galiana-Merino, J.J., Rosa-Herranz, J.L., Parolai, S., 2008. Seismic P phase *picking* using a Kurtosis-based criterion in the stationary wavelet domain. *Geosci. Remote Sens. IEEE Trans.* on 46, 3815–3826.
- García, L., Álvarez, I., Benítez, C., Titos, M., Bueno, Á., Mota, S., de la Torre, Á., Segura, J.C., Alguacil, G., Díaz-Moreno, A., Prudencio, J., García-Yeguas, A., Ibáñez, J.M., Zuccarello, L., Cocina, O., Patanè, D. (2016). Advances on the automatic estimation of the P-wave onset time. *Ann. Geophys.* 59, S0434.
- Gendron, P., J. Ebel, and D. Manolakis (2000). Rapid joint detection and classification with wavelet bases via Bayes theorem, *Bull. Seism. Soc. Am.* 90, 764–774.
- Gentili, S., and A. Michelini (2006). Automatic *picking* of P and S phases using a neural tree, *J. Seismol.* 10, no. 1, 39–63.
- Hafez, A.G., Khan, M.T.A., Kohda, T. (2010). Clear P-wave arrival of weak events and automatic onset determination using wavelet filter banks. *Digit. Signal Process.* 20, 715–723.
- Hafez, A.G., Rabie, M., Kohda, T. (2013). Seismic noise study for accurate P-wave arrival detection via MODWT. *Comput. Geosci.* 54, 148–159.
- Harpel, C., de Silva, S., Salas, G. (2011). The 2 ka Eruption of Misti Volcano, Southern Peru—The Most Recent Plinian Eruption of Arequipa's Iconic Volcano. *Geological Society of America Special papers* Vol. 484 (72 pp.).
- Havskov, J., and Ottemöller L. (2010). *Routine Data Processing in Earthquake Seismology: With Sample Data, Exercises and Software*, **Springer, New York**, 326 pp.
- Hibert, C., A. Mangeney, G. Grandjean, C. Baillard, D. Rivet, N. M. Shapiro, C. Satriano, A. Maggi, P. Boisser, V. Ferrazzini, and W. Crawford (2014). Automated identification, location, and volume estimation of rockfalls at Piton de la Fournaise volcano, *J. Geophys. Res.* 119, no. 5, 1082–1105.
- Hildyard, M.W., Nippress, S.E., Rietbrock, A. (2008). Event detection and phase *picking* using a time-domain estimate of predominate period Tpd. *Bull. Seismol. Soc. Am.* 98, 3025–3032.
- Hildyard, M.W., Rietbrock, A. (2010). Tpd, a damped predominant period function with improvements for magnitude estimation. *Bull. Seismol. Soc. Am.* 100, 684–698.
- Kalkan, E. (2016). An automatic P-phase arrival-time picker. *Bull. Seismol. Soc. Am.* 106, 971–986.

- Karamzadeh, N., Doloei, G.J., Reza, A.M. (2013). Automatic earthquake signal onset *picking* based on the continuous wavelet transform. *IEEE Trans. Geosci. Remote Sens.* 51, 2666–2674.
- Kulesh, M., Diallo, M.S., Holschneider, M., Kurennaya, K., Kruger, F., Ohrnberger, M., Scherbaum, E. (2007). Polarization analysis in the wavelet domain based on the adaptive covariance method. *Geophys. J. Int.* 170, 667–678.
- Küperkoch, L., Meier, T., Diehl, T. (2012): Automated Event and Phase Identification. - In: Bormann, P. (Ed.), *New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, Potsdam: **Deutsches GeoForschungsZentrum GFZ**, pp. 1—52.
- Küperkoch, L., Meier, T., Lee, J., Friederich, W. (2010). Automated determination of P-phase arrival times at regional and local distances using higher order statistics. *Geophys. J. Int.* 181, 1159–1170.
- Lahr, J.C., (1999). HYPOELLIPSE: a computer program for determining local earthquake hypocentral parameters, magnitude, and first-motion pattern: **U.S. Geological Survey Open-File Report 99–23**, version 1.1, 119 p.
- Lee, W. H. K. and J. C. Lahr (1975). HYP071 (Revised): A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, **U. S. Geological Survey Open File Report 75-311**, 113 pp.
- Leonard, M., and B. L. N. Kennett (1999). Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth. Planet. In.* 113, 247–263.
- Leonard, M., and B. T. N. Kennett (1999). Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth Planet. In.* 113, no. 1, 247–263.
- Li, X.B., Shang, X.Y., Wang, Z.W., Dong, L.J., Weng, L. (2016). Identifying P-phase arrivals with noise: an improved Kurtosis method based on DWT and STA/LTA. *J. Appl. Geophys.* 133, 50–61.
- Lockman, A.B., Allen, R.M. (2005). Single-station earthquake characterization for early warning. *Bull. Seismol. Soc. Am.* 95, 2029–2039.
- Lomax, A., C. Satriano, and M. Vassallo (2012). Automatic picker developments and optimization: FilterPicker: A robust, broadband picker for real-time seismic monitoring and earthquake early warning, *Seismol. Res. Lett.* 83, no. 3: 531–540.
- Maeda, N. (1985). A method for reading and checking phase times in autoprocesing system of seismic wave data, *Zisin* 38, no. 2, 365–379.
- McNutt, S. (2005). Volcanic seismology. *Annual Review Earth and Planetary Science.* 33, 461–491.
- Magotra, N., Ahmed, N., Chael, E. (1987). Seismic event detection and source location using single-station (three-component) data. *Bull. Seismol. Soc. Am.* 77, 958–971.
- Nippres, S. E. J., A. Rietbrock, and A. E. Heath (2010). Optimized automatic pickers: Application to the ANCORP data set, *Geophys. J. Int.* 181, 911–925.
- Olea, B., Alguacil, G., Vidal, F., Feriche, M. (2011). Parámetros de movimiento intenso y su relación con la intensidad macrosísmica en el área euro-mediterránea. IV Congreso Nacional de Ingeniería Sísmica. Granada, Mayo 2011. Ponencia 45. 8pp.
- Panagiotakis, C., E. Kokinou, and F. Vallianatos (2008). Automatic P-phase *picking* based on local-maxima distribution, *IEEE Trans. Geosci. Remote Sens.* 46, no. 8, 2280–2287.
- Paquereau-Lebti, P., Thouret, J.-C., Wörner, G., Fornari, M. (2006). Neogene and Quaternary ignimbrites in the area of Arequipa, southern Peru: stratigraphical and petrological correlations. *Journal of Volcanology and Geothermal Research* 154, 251–275.

- Roberts, R.G., Christoffersson, A., Cassidy, F. (1989). Real-time event detection, phase identification and source location estimation using single station three-component seismic data. *Geophys. J. Int.* 97, 471–480.
- Romero, J.E., Titos, M., Bueno, A., Alvarez, I., Garcia, L., de la Torre, A., Benitez, M.C. (2016). APASVO: a free software tool for automatic P-phase *picking* and event detection in seismic traces. *Comput. Geosci.* 90, 213–220.
- Ross, Z.E., Ben-Zion, Y. (2014). Automatic *picking* of direct P, S seismic phases and fault zone head waves. *Geophys. J. Int.* 199, 368–381.
- Saragiotis, C.D., Hadjileontiadis, L.J., Panas, S.M. (2002). PAI-S/K: a robust automatic seismic P phase arrival identification scheme. *Geosci. Remote Sens. IEEE Trans. on* 40, 1395–1404.
- Saragiotis, C.D., Hadjileontiadis, L.J., Rekanos, I.T., Panas, S.M. (2004). Automatic P phase *picking* using maximum kurtosis and k-statistics criteria. *Geosci. Remote Sens. Lett. IEEE* 1, 147–151.
- Sleeman, R., and Van Eck, T. (1999). Robust automatic P-phase *picking*: an on-line implementation in the analysis of broadband seismogram recordings. *Phys. Earth Planet. Inter.* 113, 265–275.
- Stewart, S. W. (1977). Real-time detection and location seismic events in central California, *Bull. Seism. Soc. Am.* 67, 433-452.
- Takanami, T., and Kitagawa, G. (1991). Estimation of the arrival times of seismic waves by multivariate time series model. *Ann. Inst. Stat. Math.*, 43, 407–433.
- Takanami, T., and Kitagawa, G. (1988). A new efficient procedure for the estimation of onset times of seismic waves. *J. Phys. Earth* 36, 267–290.
- Thouret, J.-C., Finizola, A., Fornari, M., Suni, J., Legeley-Padovani, A., Frechen, M. (2001). Geology of El Misti volcano nearby the city of Arequipa, Peru. *Geological Society of America Bulletin* 113 (12), 1593–1610.
- Tibuleac, I. M. and E.T. Herrin, (1999), An automatic method for determination of Lg arrival times using wavelet methods, *Seism. Res. Lett*, 70, 577–595.
- Vassallo, M., C. Santriano, and A. Lomax (2012). Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers, *Seismol. Res. Lett.* 83, no. 3, 541–554.
- Vidale, J.E. (1986). Complex polarization analysis of particle motion. *Bull. Seismol. Soc.*
- Wang, J., and T. Teng (1997). Identification and *picking* of S phase using an artificial neural network, *Bull. Seismol. Soc. Am.* 87, no. 5, 1140–1149.
- Zhang, H., C. Thurber, and C. Rowe (2003). Automatic P-wave arrival detection and *picking* with multiscale wavelet analysis for single-component recordings, *Bull. Seismol. Soc. Am.* 93, no. 5, 1904–1912.
- Zhao, Y., Takano, K. (1999). An artificial neural network approach for broadband seismic phase *picking*. *Bull. Seismol. Soc. Am.* 89, 670–680.
- Zobin, V.M. (2012). Introduction to Volcanic Seismology. 2nd edition, **Elsevier, Amsterdam-New York-Tokyo.** 482 p.

## Anexo A

Detalle de los 150 sismos VT seleccionados en este estudio:

N°	Fecha y Hora (UTC)	Coordenadas Geográficas (Datum:WGS84, Zona 19k)		Profundidad (km)	Magnitud (ML)
		Latitud(°)	Longitud(°)		
1	01-01-2015 06:26:21.78	-16.298 S	-71.406 W	1.3	1.9
2	03-01-2015 07:36:39.90	-16.302 S	-71.404 W	2.4	1.5
3	03-01-2015 07:50:46.86	-16.294 S	-71.408 W	2.0	1.8
4	03-01-2015 09:39:36.23	-16.295 S	-71.414 W	2.1	1.7
5	04-01-2015 15:42:32.71	-16.296 S	-71.407 W	1.9	1.9
6	05-01-2015 03:02:21.75	-16.299 S	-71.407 W	1.4	1.6
7	06-01-2015 02:48:37.89	-16.298 S	-71.406 W	2.4	1.6
8	06-01-2015 07:34:51.04	-16.294 S	-71.411 W	2.1	1.6
9	07-01-2015 23:18:41.97	-16.294 S	-71.406 W	2.1	1.6
10	08-01-2015 08:17:24.86	-16.298 S	-71.411 W	1.9	1.4
11	10-01-2015 09:33:26.09	-16.295 S	-71.409 W	2.0	1.8
12	13-01-2015 01:47:12.98	-16.296 S	-71.411 W	2.1	1.6
13	14-01-2015 08:07:05.11	-16.299 S	-71.406 W	1.5	2.0
14	15-01-2015 00:33:22.42	-16.307 S	-71.407 W	2.7	0.2
15	15-01-2015 02:58:01.88	-16.295 S	-71.404 W	1.5	1.7
16	16-01-2015 14:11:07.79	-16.293 S	-71.410 W	1.7	1.8
17	16-01-2015 17:36:45.44	-16.294 S	-71.411 W	1.8	2.4
18	17-01-2015 00:22:10.01	-16.298 S	-71.406 W	1.9	1.3
19	17-01-2015 06:03:23.61	-16.298 S	-71.405 W	1.2	1.5
20	17-01-2015 06:22:01.97	-16.300 S	-71.402 W	1.6	1.3
21	17-01-2015 06:58:10.12	-16.298 S	-71.408 W	1.4	1.2
22	17-01-2015 21:46:12.60	-16.301 S	-71.406 W	2.5	1.7
23	18-01-2015 01:00:33.56	-16.300 S	-71.404 W	2.4	1.7
24	18-01-2015 01:47:35.07	-16.301 S	-71.406 W	2.6	1.8
25	19-01-2015 04:39:04.43	-16.294 S	-71.410 W	1.8	2.4
26	19-01-2015 05:51:36.94	-16.302 S	-71.401 W	2.4	2.1
27	19-01-2015 05:52:17.49	-16.298 S	-71.404 W	1.2	2.6
28	19-01-2015 06:09:11.73	-16.295 S	-71.409 W	2.0	2.2
29	19-01-2015 07:01:14.45	-16.295 S	-71.410 W	1.9	1.7
30	19-01-2015 07:06:42.95	-16.298 S	-71.407 W	1.5	1.7
31	19-01-2015 07:13:48.80	-16.295 S	-71.410 W	1.8	1.5
32	19-01-2015 08:04:14.53	-16.300 S	-71.404 W	1.2	1.6
33	19-01-2015 08:33:11.70	-16.300 S	-71.406 W	1.3	1.6
34	19-01-2015 08:36:58.87	-16.295 S	-71.409 W	1.9	1.7
35	19-01-2015 08:43:23.99	-16.297 S	-71.406 W	1.7	1.1
36	19-01-2015 08:48:38.14	-16.299 S	-71.406 W	1.4	1.6



37	19-01-2015 09:05:15.68	-16.298 S	-71.406 W	1.3	1.8
38	19-01-2015 09:13:15.12	-16.297 S	-71.407 W	1.0	1.8
39	19-01-2015 09:30:54.36	-16.296 S	-71.405 W	1.2	1.8
40	19-01-2015 23:50:26.20	-16.296 S	-71.410 W	1.7	1.6
41	20-01-2015 00:02:24.31	-16.303 S	-71.406 W	2.1	1.8
42	20-01-2015 12:03:48.99	-16.301 S	-71.404 W	2.3	1.6
43	21-01-2015 05:38:15.87	-16.302 S	-71.405 W	2.1	1.3
44	21-01-2015 05:49:12.24	-16.293 S	-71.409 W	1.4	1.9
45	23-01-2015 04:50:36.49	-16.297 S	-71.409 W	1.6	2.1
46	23-01-2015 06:07:44.60	-16.301 S	-71.405 W	2.2	1.5
47	23-01-2015 07:40:29.30	-16.301 S	-71.401 W	1.8	1.4
48	23-01-2015 07:55:07.01	-16.301 S	-71.404 W	2.5	1.6
49	23-01-2015 08:01:02.78	-16.298 S	-71.407 W	1.0	1.5
50	24-01-2015 01:14:45.16	-16.303 S	-71.407 W	2.3	1.5
51	25-01-2015 09:06:19.72	-16.302 S	-71.404 W	2.1	1.1
52	25-01-2015 10:27:12.68	-16.302 S	-71.403 W	2.8	2.0
53	27-01-2015 04:44:29.85	-16.301 S	-71.400 W	1.2	1.5
54	28-01-2015 01:50:28.39	-16.301 S	-71.403 W	2.9	2.1
55	28-01-2015 05:44:06.17	-16.302 S	-71.405 W	2.1	1.5
56	31-01-2015 08:36:58.58	-16.297 S	-71.415 W	1.5	2.2
57	06-02-2015 03:10:22.39	-16.302 S	-71.401 W	2.4	1.6
58	06-02-2015 06:05:18.82	-16.295 S	-71.409 W	2.0	1.4
59	06-02-2015 06:29:46.14	-16.300 S	-71.403 W	1.1	1.5
60	06-02-2015 14:35:12.35	-16.294 S	-71.408 W	2.2	2.1
61	11-02-2015 02:56:55.45	-16.300 S	-71.404 W	2.7	1.8
62	11-02-2015 03:34:16.29	-16.300 S	-71.402 W	1.0	1.8
63	11-02-2015 04:53:02.81	-16.302 S	-71.402 W	1.0	1.1
64	11-02-2015 05:17:10.02	-16.302 S	-71.398 W	2.8	2.0
65	11-02-2015 05:24:16.41	-16.300 S	-71.401 W	1.0	1.8
66	11-02-2015 05:37:49.73	-16.300 S	-71.397 W	2.9	2.1
67	11-02-2015 05:43:25.68	-16.302 S	-71.400 W	2.1	1.2
68	11-02-2015 06:39:13.30	-16.295 S	-71.408 W	2.1	2.2
69	11-02-2015 07:15:34.35	-16.294 S	-71.407 W	2.2	2.8
70	11-02-2015 07:23:48.74	-16.295 S	-71.408 W	2.2	1.8
71	11-02-2015 09:04:03.30	-16.301 S	-71.398 W	2.8	2.1
72	11-02-2015 09:29:24.00	-16.302 S	-71.401 W	1.1	2.1
73	11-02-2015 10:49:32.41	-16.301 S	-71.401 W	2.9	2.0
74	11-02-2015 11:49:50.57	-16.296 S	-71.414 W	1.5	2.1
75	11-02-2015 13:19:14.53	-16.300 S	-71.400 W	3.0	2.0
76	17-02-2015 03:29:34.26	-16.301 S	-71.400 W	2.5	1.6
77	17-02-2015 04:10:09.41	-16.301 S	-71.400 W	2.7	1.9
78	18-02-2015 05:30:26.71	-16.299 S	-71.397 W	1.2	1.5
79	18-02-2015 08:13:10.83	-16.301 S	-71.400 W	2.4	1.6
80	18-02-2015 09:15:31.11	-16.293 S	-71.407 W	2.2	2.3
81	18-02-2015 10:53:31.61	-16.294 S	-71.408 W	1.9	1.9

82	22-02-2015 01:47:13.99	-16.301 S	-71.399 W	2.6	1.7
83	22-02-2015 02:22:46.85	-16.301 S	-71.401 W	2.4	1.6
84	23-02-2015 23:20:19.81	-16.297 S	-71.412 W	1.5	1.7
85	24-02-2015 00:07:58.65	-16.298 S	-71.413 W	1.5	2.4
86	24-02-2015 00:50:15.58	-16.297 S	-71.413 W	1.8	2.0
87	04-03-2015 04:25:28.60	-16.300 S	-71.400 W	1.0	2.2
88	09-03-2015 21:23:44.91	-16.299 S	-71.406 W	1.2	1.7
89	14-03-2015 03:08:07.97	-16.297 S	-71.406 W	1.5	2.0
90	14-03-2015 03:37:20.96	-16.298 S	-71.405 W	1.2	2.3
91	14-03-2015 06:02:27.47	-16.299 S	-71.404 W	1.3	2.1
92	14-03-2015 08:15:55.58	-16.298 S	-71.406 W	1.2	2.1
93	14-03-2015 09:45:20.39	-16.299 S	-71.406 W	1.3	1.4
94	14-03-2015 09:58:34.77	-16.298 S	-71.406 W	1.3	2.1
95	14-03-2015 10:30:48.53	-16.298 S	-71.405 W	1.3	1.8
96	14-03-2015 10:40:26.04	-16.295 S	-71.412 W	1.7	2.8
97	14-03-2015 11:27:44.73	-16.300 S	-71.404 W	2.4	1.9
98	14-03-2015 12:48:49.40	-16.298 S	-71.406 W	1.4	2.0
99	18-03-2015 01:49:44.87	-16.297 S	-71.406 W	1.0	1.6
100	19-03-2015 10:50:20.25	-16.298 S	-71.404 W	1.3	2.5
101	25-03-2015 03:19:07.32	-16.303 S	-71.400 W	2.5	2.1
102	28-03-2015 04:38:26.20	-16.298 S	-71.406 W	1.2	1.8
103	28-03-2015 07:27:01.64	-16.297 S	-71.406 W	1.1	2.0
104	30-03-2015 08:22:34.04	-16.297 S	-71.407 W	1.3	2.0
105	02-04-2015 09:07:59.79	-16.304 S	-71.413 W	2.6	1.7
106	03-04-2015 10:13:21.15	-16.299 S	-71.407 W	1.3	1.5
107	03-04-2015 21:56:13.85	-16.298 S	-71.409 W	1.9	1.5
108	03-04-2015 22:45:11.03	-16.295 S	-71.412 W	2.1	2.2
109	04-04-2015 00:08:56.21	-16.296 S	-71.398 W	1.9	1.6
110	04-04-2015 00:46:19.02	-16.297 S	-71.414 W	2.2	2.0
111	04-04-2015 01:34:00.70	-16.300 S	-71.412 W	1.5	1.2
112	04-04-2015 01:44:59.71	-16.295 S	-71.411 W	2.3	1.7
113	05-04-2015 12:28:23.47	-16.303 S	-71.410 W	1.8	1.5
114	06-04-2015 00:33:14.81	-16.286 S	-71.416 W	3.4	1.3
115	06-04-2015 18:50:50.34	-16.312 S	-71.405 W	2.1	1.7
116	06-04-2015 23:16:34.58	-16.297 S	-71.408 W	1.3	1.4
117	07-04-2015 00:50:01.56	-16.289 S	-71.411 W	1.9	1.6
118	07-04-2015 02:29:15.09	-16.309 S	-71.405 W	2.0	1.7
119	07-04-2015 03:24:17.73	-16.300 S	-71.406 W	2.7	1.8
120	07-04-2015 03:57:17.87	-16.291 S	-71.409 W	1.1	1.4
121	07-04-2015 05:22:20.91	-16.291 S	-71.408 W	1.8	1.6
122	07-04-2015 05:30:59.63	-16.306 S	-71.409 W	2.6	1.7
123	07-04-2015 06:27:34.15	-16.298 S	-71.407 W	1.3	2.3
124	07-04-2015 06:56:34.05	-16.305 S	-71.399 W	2.7	2.4
125	07-04-2015 08:10:18.83	-16.307 S	-71.400 W	2.5	2.2
126	07-04-2015 08:12:36.21	-16.300 S	-71.401 W	2.4	2.1

127	07-04-2015 08:36:55.41	-16.314 S	-71.405 W	1.8	1.5
128	07-04-2015 08:45:20.19	-16.300 S	-71.406 W	2.1	1.8
129	08-04-2015 01:56:12.23	-16.301 S	-71.407 W	1.3	2.1
130	09-04-2015 04:25:19.47	-16.298 S	-71.406 W	1.3	2.2
131	16-04-2015 02:41:19.23	-16.299 S	-71.407 W	1.4	1.9
132	16-04-2015 02:44:16.25	-16.297 S	-71.408 W	1.6	1.8
133	19-04-2015 06:08:14.78	-16.299 S	-71.404 W	1.2	1.7
134	19-04-2015 07:41:38.83	-16.297 S	-71.406 W	1.5	1.8
135	19-04-2015 07:49:34.35	-16.297 S	-71.409 W	2.1	2.3
136	19-04-2015 10:55:56.78	-16.298 S	-71.406 W	1.0	1.8
137	19-04-2015 15:50:16.33	-16.302 S	-71.406 W	2.3	1.6
138	19-04-2015 17:51:21.19	-16.299 S	-71.406 W	1.4	1.6
139	19-04-2015 19:30:42.93	-16.299 S	-71.410 W	1.5	2.6
140	21-04-2015 05:12:01.57	-16.302 S	-71.405 W	1.9	1.6
141	23-04-2015 19:10:51.98	-16.302 S	-71.405 W	1.3	1.7
142	23-04-2015 20:43:58.64	-16.299 S	-71.406 W	1.1	1.6
143	24-04-2015 00:19:01.32	-16.304 S	-71.404 W	1.1	2.1
144	26-04-2015 10:47:35.30	-16.300 S	-71.406 W	1.0	1.7
145	26-04-2015 12:35:59.97	-16.295 S	-71.409 W	1.8	1.7
146	28-04-2015 02:16:10.59	-16.299 S	-71.405 W	1.3	2.0
147	29-04-2015 21:38:38.70	-16.303 S	-71.404 W	1.2	1.5
148	30-04-2015 00:16:38.34	-16.296 S	-71.407 W	1.2	1.7
149	30-04-2015 07:11:05.29	-16.298 S	-71.407 W	1.4	2.2
150	30-04-2015 07:43:39.89	-16.297 S	-71.406 W	1.4	1.6