



UNIVERSIDAD DE CONCEPCIÓN  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

# VARIACIÓN DE VELOCIDAD DE ONDAS SÍSMICAS PRODUCTO DE MEGASEQUÍA DE CHILE MEDIANTE INTERFEROMETRÍA SÍSMICA PASIVA

**Por: Héctor Elías Carrasco Neira**

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la  
Universidad de Concepción para optar al grado académico de Título de  
Geofísico

Marzo 2023

Concepción, Chile

**Profesor Guía:**

**Dr. Matthew Miller**

**Comisión:**

**Dr. Matthew Miller, Dr. Jaime Belmonte, Dr. Rodrigo Abarca**



© 2022, Héctor E. Carrasco Neira

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

## AGRADECIMIENTOS

Gracias a mis padres que me han apoyado todos estos años y a mis amigos que me han ayudado a seguir adelante.

## Resumen

En esta tesis se explora la relación entre la variación relativa de ondas sísmicas y variables meteorológicas mediante la interferometría sísmica pasiva. Para esto utilizamos los datos obtenidos de un grupo de 4 estaciones sísmicas de la zona central de Chile: BO01, BO03, MT01 Y VA05. Se realiza la correlación cruzada en cada combinación posible y luego se calcula el factor de estrechamiento de cada serie de datos respecto a una señal de referencia específico para cada par, resultado de promediar todos los datos de dicha serie de datos. Luego de obtener los resultados en distintos formatos (diario con media móvil, mensual y anual), los comparamos con datos obtenidos de 3 cuencas que se encuentran dentro del área de estudio a lo largo de 4 años: 2016 a 2019, en el periodo de la megasequía de Chile. Estos datos son temperatura, caudal, precipitaciones y evapotranspiración potencial. De esta manera determinamos que la temperatura tiene un mayor impacto que el almacenamiento de agua, al menos a corto plazo, pero que esta última igual presenta efectos y es la principal diferencia entre los resultados de cada estación, puesto que las variaciones de temperatura son mucho más homogéneas dentro de una zona reducida.

# Índice general

<b>AGRADECIMIENTOS</b>	<b>I</b>
<b>Resumen</b>	<b>II</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Marco Teórico</b>	<b>3</b>
2.1. Función de Green . . . . .	3
2.2. Interferometría sísmica pasiva . . . . .	4
2.3. Cambios relativos en la velocidad de ondas sísmicas . . . . .	6
2.3.1. Representación del campo de ondas perturbado . . . . .	7
2.3.2. Coeficiente de correlación . . . . .	7
2.3.3. Perturbación del tiempo de viaje . . . . .	8
2.3.4. Velocidad de propagación perturbada . . . . .	9
<b>3. Contexto Hidrológico</b>	<b>10</b>
3.1. Mediciones y datos . . . . .	10
<b>4. Metodología</b>	<b>15</b>
4.1. Instrumentación . . . . .	16
4.2. Pre-procesamiento de datos . . . . .	17
4.2.1. Normalización espectral . . . . .	18
4.3. Extraer la función de Green . . . . .	18
4.3.1. Correlación cruzada . . . . .	21
4.4. Factor de estrechamiento . . . . .	21
4.4.1. Análisis . . . . .	25
<b>5. Resultados</b>	<b>29</b>
5.0.1. MT01-BO01 . . . . .	30
5.0.2. VA05-BO01 . . . . .	31
5.0.3. MT01-VA05 . . . . .	32
5.0.4. BO03-BO01 . . . . .	33
5.0.5. BO03-MT01 . . . . .	34
5.0.6. BO03-VA05 . . . . .	35
5.1. Promedios . . . . .	35

---

5.1.1. Promedios mensuales . . . . .	38
<b>6. Discusión</b>	<b>42</b>
6.1. Relación con variables meteorológicas . . . . .	45
6.1.1. Datos anuales . . . . .	48
<b>7. Conclusión</b>	<b>50</b>
<b>Referencias</b>	<b>52</b>
<b>Apéndices</b>	<b>54</b>
<b>A. Glosario</b>	<b>54</b>
A1. Operadores vectoriales . . . . .	54
A2. Tensores especiales . . . . .	55
A2.1. Delta de Kronecker . . . . .	55
A2.2. Tensor de Levi Civita . . . . .	55
A2.3. Delta de Dirac . . . . .	55
A3. Propiedades tensoriales . . . . .	55
<b>B. Tablas</b>	<b>56</b>
A1. Información de estaciones sismológicas . . . . .	56
A2. Polos y Zeros Trillium 120P . . . . .	56
<b>C. Código de Python</b>	<b>58</b>
A1. Correlación cruzada entre estaciones . . . . .	58
A2. Stretch factor y calculo de $dv/v$ . . . . .	64

# Índice de cuadros

3.1.1.Pendientes de regresión lineal. Se encuentran organizadas por cuenca en columnas y categoría en filas horizontales. . . . .	14
5.1.1.Pendientes de regresión lineal para $dv/v$ junto a la cantidad promedio de datos diarios para el resultado de promediar todas las estaciones y luego los resultados de utilizar el método de la navaja.	36
5.1.2.Pendientes de la regresión lineal para $dv/v$ para el resultado de promediar todas las estaciones y luego los resultados de utilizar el método de la navaja. . . . .	40
6.0.1.Pendientes de regresión lineal para $dv/v$ para los 3 pares de estaciones incluidos ahora además de la zona de depresión central, como complemento a la tabla 5.1.2 . . . . .	43
6.1.1.Datos anuales para la variación de velocidad, temperatura y almacenamiento de agua para los años 2016, 2017 y 2018. . . . .	49
A1.1.Información de las 6 estaciones pertenecientes a la red C1 que fueron utilizadas en el estudio. . . . .	56
A2.1.Polos y ceros para la componente vertical de sensor Trillium 120P.	57

# Índice de figuras

2.2.1.Representación de dos fuentes de ruido $N_1(t)$ y $N_2(t)$ , ubicados en las posiciones $x_s$ y $x'_s$ , actuando sobre receptores ubicados en las posiciones $x_a$ y $x_b$ . . . . .	5
3.0.1.Fotos para cada uno de los 3 esteros observados. . . . .	11
3.0.2.Mapa en donde se presentan, como triángulos azules, los 3 esteros a los que se le realizaron las mediciones. . . . .	12
3.1.1.Datos obtenidos de CAMELS-CL de $(CR)^2$ , para las 3 cuencas nombradas. 5746001 en magenta, 6043001 en verde y 6132001 en negro. . . . .	13
4.1.1.Mapa en donde se presentan, como diamantes purpuras, las 4 estaciones listadas utilizadas para este estudio. . . . .	16
4.2.1.Serie de tiempo de la estación MT01 el día 1 de enero de 2017. En 4.2.1a se ve antes de cualquier clase de preprocesamiento. En 4.2.1b ya se le removió la media y la inclinación lineal y en 4.2.1c por un filtro pasabanda. En 4.2.1d se le aplicó normalización espectral sobre todo lo anterior. . . . .	19
4.2.2.Transformada de Fourier para una serie de tiempo de la estación MT01 el día 1 de enero de 2017. En 4.2.2a se ve antes de cualquier clase de preprocesamiento. En 4.2.2b ya se le removió la media y la inclinación lineal, además de haber pasado por un filtro pasabanda. En 4.2.2c se le aplicó normalización espectral sobre todo lo anterior. . . . .	20
4.3.1.Comparativa para matrices de correlación de 3 pares de estaciones utilizando filtros pasabanda para dos intervalos de frecuencia: a la izquierda 0,4 y 2[Hz] y a la derecha entre 1 y 4[Hz]. Además a la izquierda de cada matriz se encuentra la traza de referencia de la matriz. . . . .	22
4.3.2.Matriz de correlación para los 3 pares de estaciones de la figura 4.3.1. Van desde el 1 de enero de 2016 hasta el 31 de diciembre de 2019, completando cuatro años de correlaciones diarias. En 4.3.2a se hace un acercamiento en el eje Y, dejando el time lag entre 0 y 40 segundos en vez de 80. . . . .	23

4.3.3. Matriz de correlación para los 3 pares de estaciones restantes, todas filtradas en frecuencias entre 0,4 y 2[Hz], los cuales también van desde el 11 de febrero de 2016 hasta el 31 de diciembre de 2019. . . . .	24
4.4.1. Ejemplo de obtención de factor de estrechamiento utilizando autocorrelación en la estación KZK <a href="#">Wegler et al. (2009)</a> . En (a) se observan 11 trazas que fueron expandidas o comprimidas y en gris, detrás de cada una, la traza de referencia con la que se compara. En (b) se observa la correlación entre las 11 trazas distintas con la traza de referencia. Se observa que la correlación alcanza su máximo con un factor de estrechamiento de $-0,4\%$ , llegando a una correlación máxima de $R = 0,88$ . . . . .	26
4.4.2. Variación del factor de estrechamiento diario para el par de estaciones VA05-BO01 entre los años 2016 y 2019. La matriz representa el nivel de correlación considerando cada factor de estrechamiento, mientras que los puntos rojos representan el máximo diario. Se utilizó una ventana de tiempo de 15 segundos para la coda . . . . .	27
4.4.3. En 4.4.3a se ve como varía el factor de estrechamiento a lo largo de los 4 años para el mismo par de estaciones de 4.4.2. Se le aplicó una media móvil de 60 días para suavizar el gráfico y hacerlo más claro de visualizar. En 4.4.3b se observa la correlación máxima para cada día respecto a la traza de referencia. La leyenda en cada figura corresponde a la ventana de tiempo de la coda utilizada, en segundos. . . . .	27
5.0.1. Media móvil a lo largo de 4 años para el el $dv/v$ entre MT01-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	30
5.0.2. Media móvil a lo largo de 4 años para el el $dv/v$ entre VA05-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	31
5.0.3. Media móvil a lo largo de 4 años para el el $dv/v$ entre MT01-VA05 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	32
5.0.4. Media móvil a lo largo de 4 años para el el $dv/v$ entre BO03-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	33
5.0.5. Media móvil a lo largo de 4 años para el el $dv/v$ entre BO03-MT01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	34
5.0.6. Media móvil a lo largo de 4 años para el el $dv/v$ entre BO03-VA05 como un gráfico rojo y tendencia lineal de los peaks como una recta negra. . . . .	35
5.1.1. En rojo gráfico que promedia la variación de velocidad de los 6 pares de estaciones. En negro la media móvil sin pesos que abarca 60 días y en azul la tendencia lineal. . . . .	36
5.1.2. Gráfico que representa la cantidad de datos para cada día a lo largo de los 4 años de datos. . . . .	37

5.1.3. Gráfico que compara la media móvil que incluye los 6 pares de estaciones en negro, presentada en 5.1.1, mientras que en colores se encuentran los resultados de aplicar el método de la navaja. . . .	37
5.1.4. Gráficos que comparten la misma estructura que 5.1.1 pero se diferencian en que se ignora una de las 4 estaciones diferentes en cada uno, es decir que cada gráfico es solo el promedio entre 3 pares de estaciones. . . . .	39
5.1.5. Gráfico que compara la media móvil que incluye los 6 pares de estaciones en negro con gráficos de distintos colores que representan los distintos resultados de aplicar el método de la navaja. Por otro lado la línea punteada representa la tendencia general a lo largo de los años. . . . .	40
5.1.6. Gráficos con promedios mensuales que ignoran una de las 4 estaciones diferentes en cada uno. . . . .	41
6.0.1. Variación mensual para el promedio de los 3 pares de estaciones utilizados para estudiar la zona de la depresión central. . . . .	43
6.0.2. Mapa con todas las estaciones, incluyendo las presentes en 4.1.1, representados nuevamente con diamantes morados, junto a las las 2 incluidas en esta sección, como diamantes azules. Además se incluyen los coeficiente observado en las tablas 6.0.1 simbolizados con líneas que unen ambas estaciones. Líneas rojas representan variaciones negativas en la velocidad y líneas azules variaciones positivas. . . . .	44
6.1.1. Datos de almacenamiento de agua a lo largo del tiempo para la zona estudiada, junto a su media móvil. . . . .	46
6.1.2. Correlación cruzada entre el promedio de todas las estaciones mostrado en el gráfico de la figura 5.1.1 y los datos de meteorológicos de cuencas presentados en las gráficos de la figura 3.1.1. . . . .	47
6.1.3. En a) tenemos la temperatura en °C modelada con una función seno. En b) la misma función luego de aplicarle el delay extra de 0.06 años (equivalente a 22 días aproximadamente) e invertirla. . .	48

# Capítulo 1

## Introducción

La interferometría sísmica pasiva es un método de análisis de datos sismológicos y/o volcánicos que no requiere conocer la estructura de la corteza terrestre que se desea estudiar (Sens-Schönfelder and Wegler, 2006). Para esto, hay que reconstruir la función de Green entre dos sensores utilizando la correlación del ruido sísmico ambiental captado por estos (Campillo, 2006), para así obtener la variación de la velocidad de ondas sísmicas en función del tiempo (Snieder, 2006).

Este método de análisis de datos sismológicos ya ha sido ocupado exitosamente en la estudio de aguas subterráneas (Lecocq et al., 2017), donde demuestran que la interferometría sísmica pasiva sirve para observar variaciones en la velocidad del ruido sísmico a largo plazo y que es posible enlazar dichas variaciones a variaciones climatológicas externas, encontrando su origen tanto en contribuciones termo-elásticas e hidrológicas a partes iguales, trabajando bajo la hipótesis de que el agua subterránea genera un aumento en la presión del suelo, por tanto aumentando la velocidad de las ondas sísmicas propagándose a través de este

Actualmente Chile sufre de una sequía que se viene extendiendo desde 2010, la cual es conocida como Mega Sequía (Garreaud et al., 2019). Con déficit en las precipitaciones medias de entre 20 % y 40 % y disminuyendo en gran manera el acceso a agua fresca, la mega sequía ha tenido fuertes impactos en el ámbito social y económico. Esta sequía también ha afectado enormemente a los cuerpos de agua y el flujo de los ríos (Aguayo et al., 2019). Se prevé que este fenómeno vaya en aumento con sequías anuales al nivel de la sufrida en 2016, que marcó el año más seco en 70 años. Este trabajo busca utilizar la interferometría sísmica pasiva para

observar el desarrollo de esta mega sequía y como ha afectado y podría afectar a las aguas subterráneas, las cuales representan a su vez una de las mayores reservas de agua fresca.

## Capítulo 2

### Marco Teórico

En esta sección se explicará la función de Green, para luego proponer un modelo de interferometría sísmica pasiva para dos estaciones a través del cual mostrar y explicar como se calcula la correlación cruzada entre estas y finalmente el factor de estrechamiento.

#### 2.1. Función de Green

La función de Green representa la forma en que un medio responde ante acciones que se ejercen de manera dinámica sobre él. (Sánchez Sesma et al., 2017)

Al aplicar fuerzas sobre un volumen y su superficie se genera un campo de desplazamiento en donde sus soluciones están restringidas al movimiento de una partícula en un medio continuo, descrito por la función  $\delta_i = f_i + \tau_{ij,j}$ . Sin embargo en sismología, el representar el desplazamiento de una partícula en un sistema lagrangiano sobre un medio elástico se torna algo más complejo al considerar las variables que pueden originar el movimiento (fuerzas de cuerpo y de superficie) Aki and Richards (1976).

Si consideramos un impulso en la coordenada  $\vec{x} = \vec{x}'$  aplicado en el instante de tiempo  $t = t'$  con dirección  $-n$  y magnitud  $A = 1$ , entonces la componente  $i$ -ésima del campo sería tal que

$$G_{in} = G_{in}(\vec{x}, t, \vec{x}', t') \quad (2.1.1)$$

Donde  $G_{in}$  es la función de Green. En esta,  $(\vec{x}, t)$  representan el campo de desplazamiento para cualquier punto del medio elástico mientras que  $(\vec{x}', t')$  son las coordenadas de la fuente del impulso. La función de Green tiene que cumplir la ecuación de movimiento considerando que se aplica una fuerza de cuerpo sobre una partícula cuyo impulso tiene dirección  $-n$  tal que  $f_i = \delta(\vec{x} - \vec{x}')\delta(t - t')\delta_{in}$ .

Si el desplazamiento y la tracción desaparecen en toda la superficie al considerar que  $G_{in}$  establece condiciones de borde homogéneas, podemos recurrir al Teorema de Betti en su forma integral, para obtener la siguiente relación entre las posiciones de la fuente y el receptor:

$$\begin{aligned}
u_n(\vec{x}, t) = & \int_{-\infty}^{\infty} dt' \int_V f_i(\vec{x}', t') G_{in}(\vec{x}', t', \vec{x}, t) dx'^3 \\
& + \int_{-\infty}^{\infty} dt' \oint_S T_i(u(\vec{x}', t'), \hat{n}) G_{in}(\vec{x}', t', \vec{x}, t) dx'^2 \\
& - \int_{-\infty}^{\infty} dt' \oint_S C_{ijkl} n_j G_{kn,l}(\vec{x}', t', \vec{x}, t) u_i(\vec{x}', t') dx'^2
\end{aligned} \tag{2.1.2}$$

El termino  $u_n(\vec{x}, t)$  en la ecuación 2.1.2 representa el desplazamiento en dirección  $n$ -ésima del receptor producto de: El desplazamiento de la fuente  $(\vec{x}', t')$ , discontinuidades en el medio elástico y tracciones sobre la superficie de la fuente  $(T_i(\vec{x}', t'))$ . En  $f_i(\vec{x}', t')$  encontramos la ponderación del desplazamiento de todos los puntos del volumen evaluados en la fuente, mientras que  $G_{in}(\vec{x}', t', \vec{x}, t)$  representa la forma que tiene la onda de desplazamiento generada por todos los puntos del volumen de la fuente.

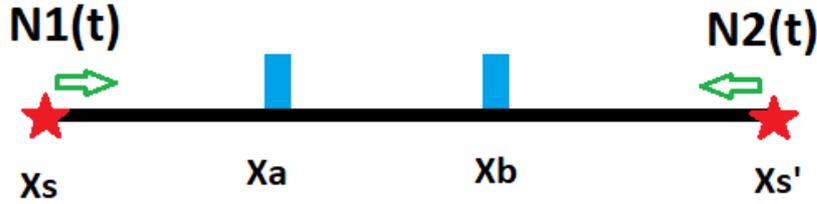
## 2.2. Interferometría sísmica pasiva

La interferometría sísmica es observar una fuente que tendrá respuestas sísmicas en distintos sensores para luego, haciendo uso de la correlación cruzada entre entre dichas respuestas, se obtiene una nueva respuesta de un sensor respecto a otro, donde este actúa como una fuente virtual. La interferometría sísmica pasiva consiste en realizar este proceso pero utilizando las respuestas de las estaciones frente al ruido sísmico constante. Al realizar esto se asume que la fuente presenta una distribución homogénea alrededor de los sensores.

Para realizar un modelo al respecto, consideremos que existen dos fuentes de ruido:  $N_1(t)$  y  $N_2(t)$  actuando al mismo tiempo en las posiciones  $x_s$  y  $x'_s$  como muestra la figura 4.4.1, la que obtengamos en  $x_a$  y  $x_b$  producto de ambas fuentes de ruido será:

$$u(x_a, t) = \sum_{i=1}^2 G(x_a, x_s^{(i)}, t) * N_i(t) \quad (2.2.1)$$

$$u(x_b, t) = \sum_{j=1}^2 G(x_b, x_s^{(j)}, t) * N_j(t) \quad (2.2.2)$$



**Figura 2.2.1:** Representación de dos fuentes de ruido  $N_1(t)$  y  $N_2(t)$ , ubicados en las posiciones  $x_s$  y  $x'_s$ , actuando sobre receptores ubicados en las posiciones  $x_a$  y  $x_b$

Esto demuestra que la respuesta de los sensores ubicados en  $x_a$  y  $x_b$  es la superposición de las ondas que provienen de ambas fuentes. Especialmente al tratar con fuentes ruidosas, la correlación cruzada entre ambas respuestas,  $u(x_a, t)$  y  $u(x_b, t)$ , no será una versión desplazada de alguna de las fuentes, ya que las fuentes no son correlacionables entre si. Si vemos la autocorrelación de las fuentes, esta es:

$$\langle N_j(t) * N_i(-t) \rangle = \delta_{ij} C_N(t) \quad (2.2.3)$$

En esta ecuación  $\delta_{ij} C_N(t)$  representa la autocorrelación de la fuente  $N(t)$  mientras que el operador  $\langle \cdot \rangle$  es un promedio integrado sobre el tiempo. La correlación cruzada de la respuestas de ambos sensores está dada por:

$$\langle u(x_b, t) * u(x_a, -t) \rangle = \left\langle \sum_{j=1}^2 G(x_b, x_s^{(j)}, t) * N_j(t) * \sum_{i=1}^2 G(x_a, x_s^{(i)}, t) * N_i(t) \right\rangle \quad (2.2.4)$$

Luego, utilizando las ecuaciones 2.2.3 y 2.2.4 tenemos que

$$\langle u(x_b, t) * u(x_a, -t) \rangle = [G(x_b, x_a, t) + G(x_b, x_a, -t)] * C_N(t) \quad (2.2.5)$$

La ecuación 2.2.5 muestra que la correlación cruzada de las respuestas de ambos sensores, producto de la superposición del ruido de las fuentes, como se ve en la figura 4.4.1, donde el ruido de las fuentes  $N_1(t)$  y  $N_2(t)$  actúan en conjunto sobre los receptores, da como resultado la función de Green entre ambos sensores sumado a su versión invertida, convolucionando con la autocorrelación de la función de ruido. Los términos cruzados de la correlación  $t'_b - t_a$  y  $t_b - t'_a$  carecen de importancia física pues las fuentes en  $x_s$  y  $x'_s$  no son correlacionables (Wapenaar et al., 2010), tal como muestra la ecuación 2.2.3.

Con esto el modelo ya está planteado en una dimensión, pero puede expandirse a 2 y 3 dimensiones.

## 2.3. Cambios relativos en la velocidad de ondas sísmicas

Tras realizar la correlación cruzada entre las señales sísmicas, se pueden observar uno o dos peaks principales. Estos peaks están dominados por ondas que siguieron la trayectoria más rápida entre ambas estaciones. Por otro lado la coda está compuesta por ondas que fueron dispersadas por el medio, las cuales son muy sensibles a cualquier cambio en este (Snieder, 2006) y por tanto es posible utilizar la interferometría sísmica pasiva para detectarlos. Esta teoría se basa en 4 puntos:

1. El campo de onda total es el resultado de la superposición de las ondas que se propagan por el medio en los distintos caminos posibles.
2. El coeficiente de correlación entre la onda perturbada y la no perturbada es un reflejo de los cambios en la coda de la onda.
3. El coeficiente de correlación está relacionado a la variación en los tiempos de viaje de las ondas, producto de las perturbaciones.
4. Perturbaciones o cambios en el medio se reflejan como perturbaciones en la onda sísmica.

### 2.3.1. Representación del campo de ondas perturbado

Es posible representar el campo de onda perturbado  $\hat{u}$ , producto de cambios de velocidad o posición de los scatterer, dependiendo del tiempo  $t$  como

$$\hat{u}(t) = \sum_p S_p(t - \tau_p) \quad (2.3.1)$$

Las perturbaciones en el tiempo  $\tau_p$  generan cambios en la onda  $S_p$ , las cuales dependen de la trayectoria  $p$ . La ecuación 2.3.1 solo es valida cuando la onda de referencia no perturbada y la onda perturbada provienen de la misma fuente. El cambio de fase es el de mayor relevancia al observar los cambios de velocidad. Además, suponiendo que las perturbaciones son no dispersivas, la variación de velocidad se ve reflejada directamente en el tiempo de llegada de la onda al receptor.

### 2.3.2. Coeficiente de correlación

Tenemos que definir el coeficiente de correlación entre un campo de ondas perturbadas y no perturbadas, tomando en cuenta una ventana de tiempo de largo  $2T$  y centrado en  $t$ , tal que:

$$R(t_S) \equiv \frac{\int_{t-T}^{t+T} u(t') \tilde{u}(t' + t_S) dt'}{\sqrt{\int_{t-T}^{t+T} u^2(t') dt' \int_{t-T}^{t+T} \tilde{u}^2(t') dt'}} \quad (2.3.2)$$

El tiempo de desplazamiento entre la onda perturbada y la onda no perturbada en la correlación cruzada está representado por  $t_S$ . Esta ecuación queda:

$$R(t_S) = \frac{\sum_p C_p(t_S - \tau_p)}{\sum_p C_p(0)} \quad (2.3.3)$$

La autocorrelación es la Transformada de Fourier del espectro de potencia. Si suponemos que el espectro de potencia de las ondas tiene la misma forma en cada trayectoria  $p$  por la que viajan, implicamos que también la autocorrelación es la misma para cada trayectoria  $p$ .

$$C_p(t) = I_p C(t) \quad (2.3.4)$$

donde  $C(t)$  es la autocorrelación de  $S_p(t)$  normalizada en su máximo cuando  $t = 0$ .  $I_p$  representa la intensidad de la onda propagándose a través de la trayectoria  $p$ . Así  $C(0) = 1$ . De esta manera la ecuación 2.3.3 queda:

$$R(t_S) = \frac{\sum_p I_p C(t_S - \tau_p)}{\sum_p I_p} \quad (2.3.5)$$

Las ondas deben viajar siguiendo las mismas trayectorias  $p$  para poder comparar los cambios en las velocidades entre dos periodos de tiempo. Por este motivo las ventanas de tiempo que se escogen al observar la coda no deben ser muy largas.

### 2.3.3. Perturbación del tiempo de viaje

Si asumimos que la perturbación que la perturbación  $\tau_p$  no varía considerablemente en las diferentes trayectorias  $p$  donde viajan las ondas superficiales para la ventana de tiempo seleccionada, es posible extraer la perturbación en el tiempo de viaje utilizando la correlación entre la onda superficial perturbada y no perturbada.

Si la autocorrelación  $C(t)$  es una función para donde  $C(0) = 1$ , entonces la expansión en series de Taylor de segundo orden está dada por:

$$C(t) = C(0) + \frac{1}{2} t^2 \ddot{C}(0) \quad (2.3.6)$$

No tomamos en cuenta el término de primer orden ya que queremos que la función de autocorrelación sea par. De esta manera, obtendremos el mismo resultado independientemente de la dirección de origen de la fuente. Utilizando el tiempo  $t = t_S - \tau_p$  tendremos que

$$C(t) = \frac{\int_{t-T}^{t+T} u(t') u(t' + t) dt'}{\int_{t-T}^{t+T} u^2(t') dt'} \quad (2.3.7)$$

y el coeficiente de correlación queda representado por:

$$R(t_S) = 1 - \frac{1}{2} \bar{\omega}^2 \left[ \frac{\sum_p I_p (t_S - \tau_p)^2}{\sum_p I_p} \right] \quad (2.3.8)$$

Cuando  $\frac{dR(t_S)}{dt_s}$  encontramos el punto máximo en la correlación, lo que implica que:

$$t_S = \frac{\sum_p I_p \tau_p}{\sum_p I_p} \quad (2.3.9)$$

Estadísticamente  $t_{max} \equiv \langle \tau \rangle$ , donde  $\langle \tau \rangle$  es el promedio de las perturbaciones de las ondas que viajan en cada trayectoria  $p$ .

### 2.3.4. Velocidad de propagación perturbada

Suponiendo que la velocidad  $v$  es perturbada con  $\delta v$  por una cambio en el medio. La velocidad relativa estará dada por el factor de estrechamiento  $-\frac{\delta v}{v}$  y será constante para cualquier lugar en el espacio. El tiempo de viaje perturbado será:

$$t_p + \tau_p = \int_p \frac{1}{v + \delta v} dx \quad (2.3.10)$$

Si aproximamos  $(1 + \alpha)^n = 1 + n\alpha$  obtenemos

$$\tau_p = - \left( \frac{\delta v}{v} \right) t_p \quad (2.3.11)$$

Reemplazamos el tiempo de viaje sin perturbar  $t_p$  por el tiempo central  $t$  de la ventana de tiempo. Así consideramos los promedios de los tiempos de llegada perturbados como:

$$\langle \tau \rangle = - \left( \frac{\delta v}{v} \right) t \quad (2.3.12)$$

el cambio de velocidad relativa es igual al inverso del estrechamiento de la correlación cruzada entre una onda perturbada y una no perturbada tomando en cuenta pequeñas ventanas en la coda. Esto se debe a cambios en el medio por el que viajan las ondas sísmicas con múltiples trayectorias  $p$ .

$$-\frac{\delta v}{v} = \frac{t_{max}}{t} \quad (2.3.13)$$

Esta ecuación se conoce como "factor de estrechamiento".

# Capítulo 3

## Contexto Hidrológico

Para observar de forma más clara los efectos que ha tenido la gran sequía observamos datos obtenidos de 3 cuencas. Una cuenca es un lugar en el que muchos flujos de agua convergen, en este caso serán 3 esteros. Estos esteros son Puangue, Nilahue y Alhué. Un estero es un curso o torrente de agua similar a un río pero de menor tamaño.

Normalmente los puntos de convergencia de las cuencas son observados constantemente para medir variaciones en temperatura, agua transportada y otras variables que entreguen información sobre las condiciones hidrológicas de la zona y son identificadas por códigos en los catálogos de datos. Las 3 cuencas observadas son:

- 5746001: Converge en el estero Puangue en latitud -33.66 y longitud -71.34.
- 6043001: Converge en el estero Alhué en latitud -34.06 y longitud -71.24.
- 6132001: Converge en el estero Nilahue en latitud -34.57 y longitud -71.79.

### 3.1. Mediciones y datos

Obteniendo los datos necesarios desde el [Catálogo CAMELS-CL \(2022\)](#) que proporciona Center of Climate and Resilience Research (*CR*)<sup>2</sup>, el cual cuenta con series de tiempo para temperatura, caudal, precipitaciones y evapotranspiración potencial. El caudal representa la cantidad de agua que fluye por la cuenca hacia el punto de observación en el estero y la evapotranspiración potencial es la cantidad



(a) Estero Puangue

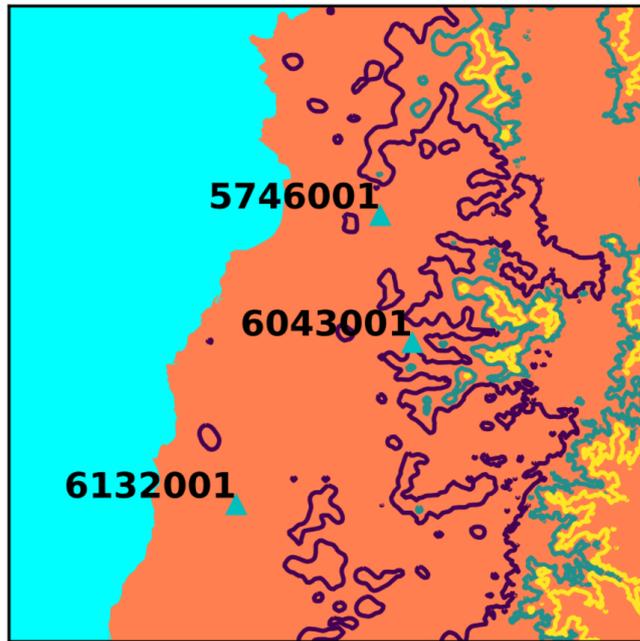


(b) Estero Nilahue



(c) Estero Alhué

**Figura 3.0.1:** Fotos para cada uno de los 3 esteros observados.

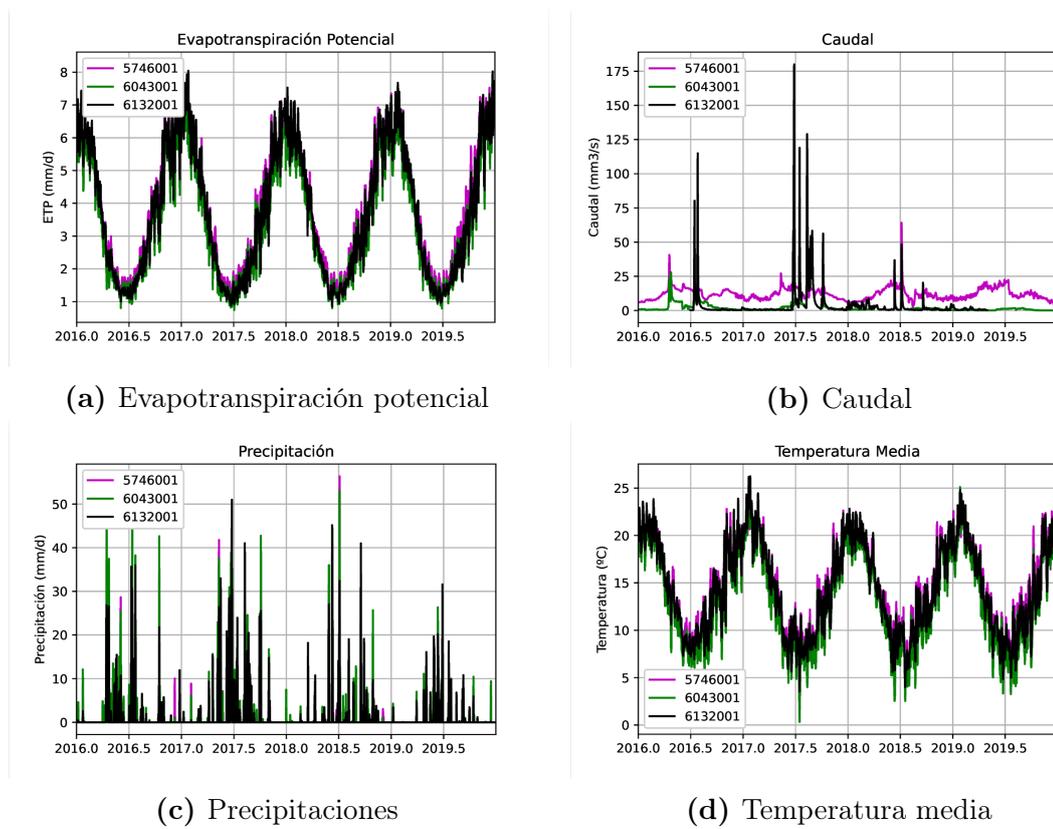


**Figura 3.0.2:** Mapa en donde se presentan, como triángulos azules, los 3 esteros a los que se le realizaron las mediciones.

máxima conjunta de evaporación de cualquier superficie, ya sea agua o suelo, y la transpiración de las plantas. Esta variable es difícil de medir y por tanto se calcula teniendo muchas variables en cuenta. Estos datos son graficados en conjunto en la figura 3.1.1 en el intervalo de tiempo que va desde el año 2016 hasta 2019, incluyendo ambos. Se observan comportamientos similares en las 3 cuencas, con la excepción del caudal de la cuenca 5746001 que presenta un comportamiento diferente.

Para una mejor visión de la situación hay que observar la tendencia lineal de estas series de datos, cuyas pendientes se encuentran en la tabla 3.1.1. En esta se observa que tanto las precipitaciones, el caudal de los esteros y la temperatura media presentan una pendiente negativa, es decir han disminuido durante los años, mientras que la evapotranspiración potencial presenta distintos comportamientos en cada cuenca.

Teniendo en cuenta principalmente la disminución constante de las precipitaciones es posible darse cuenta de como afecta la sequía al flujo de agua y como afecta a estos 3 esteros, sin embargo existen otras reservas de aguas muy importantes, las subterráneas. Estas no son fáciles de medir a grandes escalas y para esto es posible utilizar la interferometría sísmica pasiva (Lecocq et al., 2017).



**Figura 3.1.1:** Datos obtenidos de CAMELS-CL de  $(CR)^2$ , para las 3 cuencas nombradas. 5746001 en magenta, 6043001 en verde y 6132001 en negro.

---

	5746001	6043001	6132001
Evapotranspiración Potencial	0.00010651	4.75227829e-05	-5.72787255e-06
Precipitaciones	-0.00074316	-0.00101349	-0.00042872
Caudal	-0.00046489	-0.00217843	-0.00431499
Temperatura Media	-0.00074445	-0.00085444	-0.00094417

**Cuadro 3.1.1:** Pendientes de regresión lineal. Se encuentran organizadas por cuenca en columnas y categoría en filas horizontales.

# Capítulo 4

## Metodología

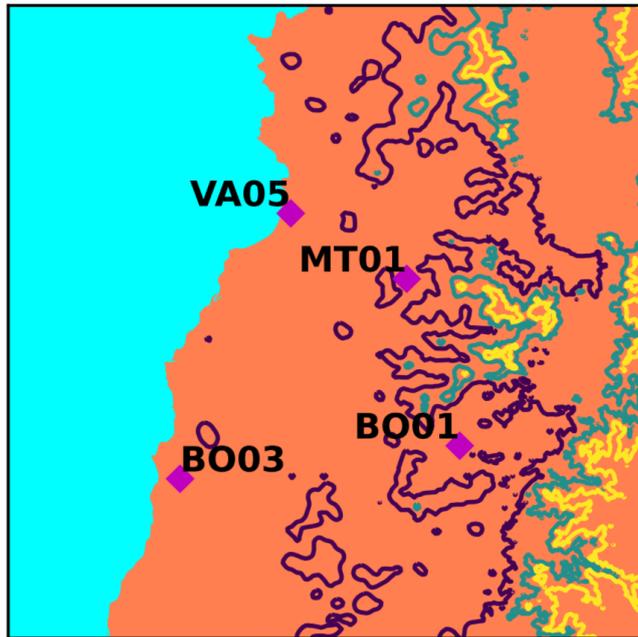
Para realizar un análisis de ruido ambiental debemos contar con un registro continuo de datos de una red sísmológica. Los datos utilizados van desde 1 de enero de 2016 hasta 31 de diciembre de 2019. Las series de tiempo utilizadas corresponden a los datos de la componente vertical de 4 estaciones de la red sísmológica C1.

Procesar los datos de ruido sísmico ambiental consiste en obtener la función de Green entre dos estaciones mediante la correlación de sus respuestas frente a la señal ruidosa proveniente de una fuente, para así extraer los cambios en la velocidad relativa en base al tiempo. (Bensen et al., 2007)

Para el procesamiento de los datos se utilizaron varios paquetes de funciones de Python, siendo los dos más destacables ObsPy y MIIC:

- ObsPy es un proyecto de acceso libre que tiene el objetivo de proveer herramientas para el análisis de datos sísmológicos en Python, siendo capaz de leer los formatos de archivos más comunes utilizados por centros sísmológicos y entregando rutinas de procesamiento que permiten la manipulación de series de tiempo sísmológicas.
- MIIC (Monitoring and Imaging based on Interferometric Concepts) es un proyecto que tiene el objetivo de desarrollar herramientas de monitoreo mediante interferometría en el ámbito de la investigación sísmológica. Incluye principalmente rutinas para analizar señales de ruido ambiental.

Este procedimiento se basa en cuatro fases principales:



**Figura 4.1.1:** Mapa en donde se presentan, como diamantes púrpuras, las 4 estaciones listadas utilizadas para este estudio.

1. Pre-procesamiento de los datos extraídos de cada estación.
2. Matriz de correlación cruzada y extracción de la función de Green.
3. Obtención del factor de estrechamiento.
4. Selección de resultados aceptables.

## 4.1. Instrumentación

Las estaciones escogidas de la red C1 son las siguientes:

- Estación BO01: Latitud -34.3917 - Longitud -71.0848
- Estación BO03: Latitud -34.49605 - Longitud -71.9612
- Estación MT01: Latitud -33.8641 - Longitud -71.2509
- Estación VA05: Latitud -33.65711 - Longitud -71.61428

Se estudiaron todos los pares de estaciones que se podían formar entre estas:

- MT01-BO01: Estas estaciones se encuentran a una distancia de 60.68 kilómetros

- MT01-B003: Estas estaciones se encuentran a una distancia de 96.05 kilómetros
- MT01-VA05: Estas estaciones se encuentran a una distancia de 40.76 kilómetros
- BO01-BO03: Estas estaciones se encuentran a una distancia de 81.28 kilómetros
- BO01-VA05: Estas estaciones se encuentran a una distancia de 95.24 kilómetros
- BO03-VA05: Estas estaciones se encuentran a una distancia de 98.70 kilómetros

Todas estas estaciones utilizan el mismo instrumento para medir: un sensor Trillium 120P, el cual responde ante señales que varían entre periodos de 120[s] y 175[Hz] y que presentan una intensidad superior a -3[dB]. Además presenta muy baja sensibilidad a variaciones en temperatura e inclinación, por lo que rara vez requieren de calibración en el centro de masa.

## 4.2. Pre-procesamiento de datos

Los datos son descargados en series de largo de un día entero. Estas series son cortadas en segmentos de largo de 1 hora cada 15 minutos, teniendo varias series que irán de 0:00 a 1:00, 0:15 a 1:15, 0:30 a 1:30, etc, dando un total de 93 segmentos para un día completo de datos. Cada uno de estos segmentos es procesado de manera independiente.

El primer paso es la preparación o pre-procesamiento. Esto consiste en remover el promedio y la tendencia lineal y aplicar un taper sobre la serie de datos, lo que se aprecia en la figura 4.2.1 comparando 4.2.1a con 4.2.1b. Luego hay que filtrar la serie en una banda de frecuencia determinada, para lo que utilizamos un filtro pasabanda, el que se observa en 4.2.1c y 4.2.2b. Posteriormente se intenta disminuir la influencia de terremotos y micro-sismicidad utilizando normalización espectral.

### 4.2.1. Normalización espectral

La normalización espectral o spectral whitening consiste en ampliar la banda de frecuencia del ruido ambiental suavizando su espectro de Fourier. De esta forma se logra disminuir la interferencia producto de microsismos en la correlación del ruido ambiental [Bensen et al. \(2007\)](#).

El ruido ambiental presenta grandes variaciones en su frecuencia, tenemos micro-sismicidad primaria con periodos de 15 segundos, micro-sismicidad secundaria con periodos de 7.5 segundos, zumbidos de periodos mayores a 50 segundos y ruido monocromático con periodos de 26 segundos, el cual produce grandes problemas al realizar correlaciones de ruido ambiental [Bensen et al. \(2007\)](#).

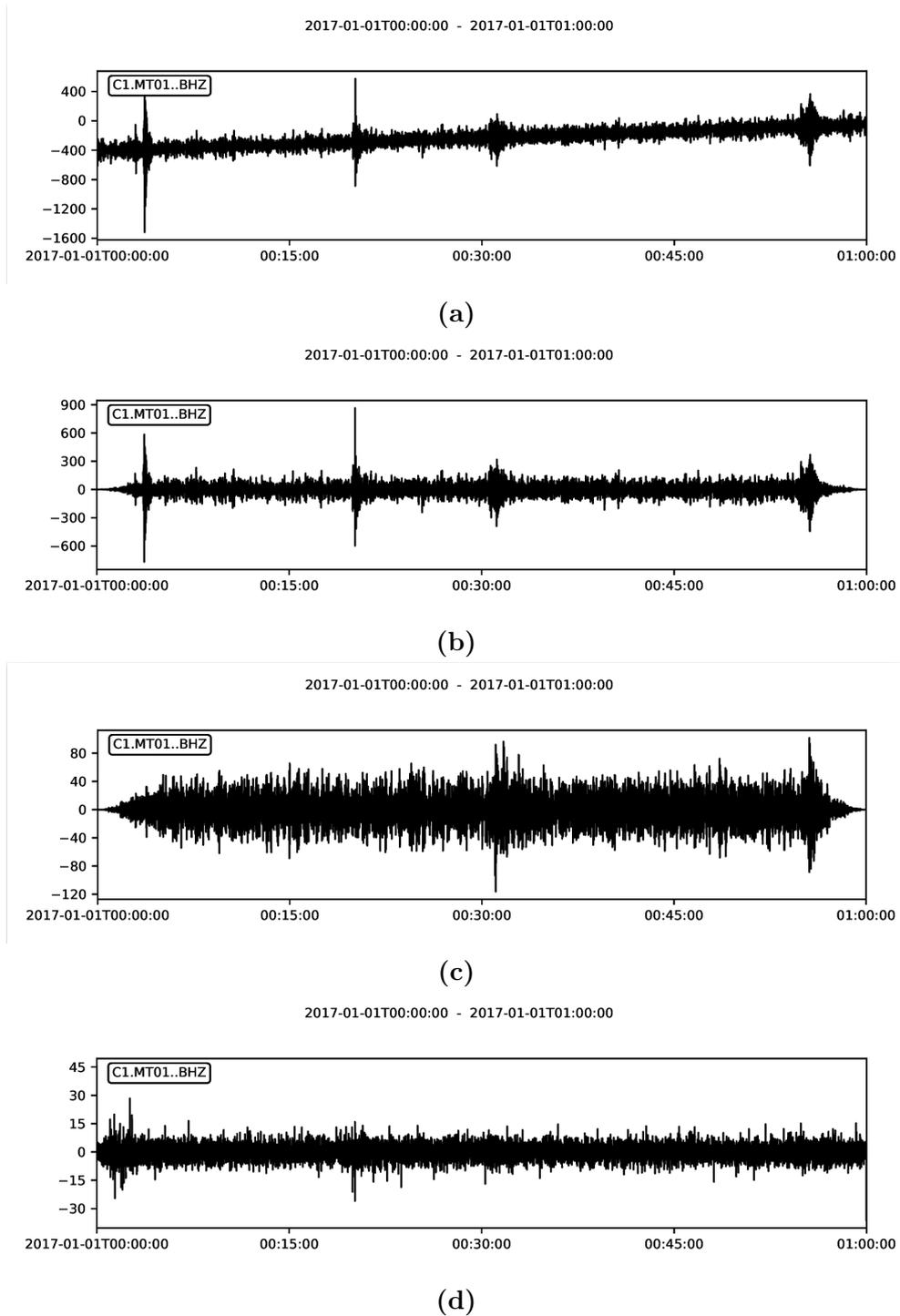
Para eliminar estas contaminaciones hay que aplicar una normalización espectral que excluya ruido ambiental con periodo superior a 7 segundos, lo cual equivale a una frecuencia de  $0,143[Hz]$ , cuyos resultados se aprecian en [4.2.1d](#) y [4.2.2c](#).

## 4.3. Extraer la función de Green

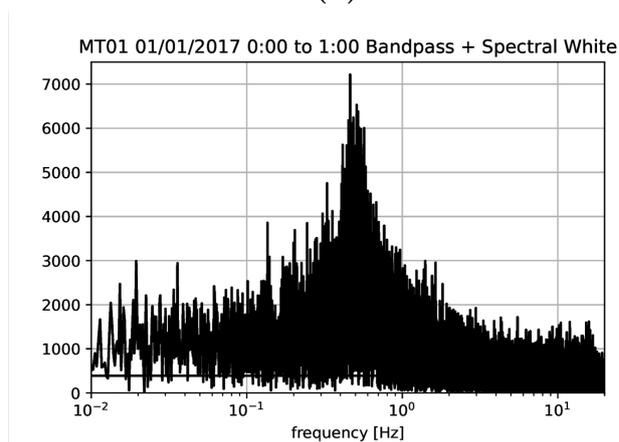
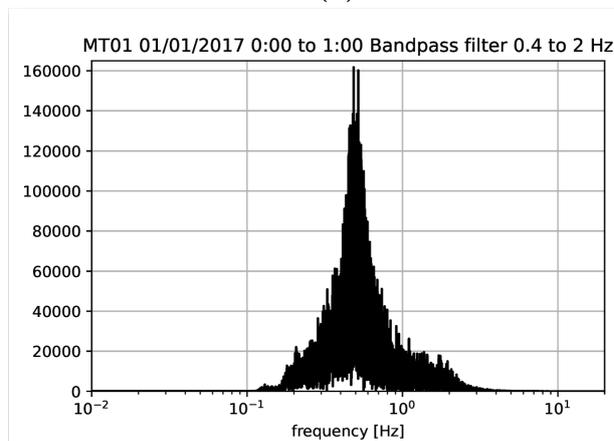
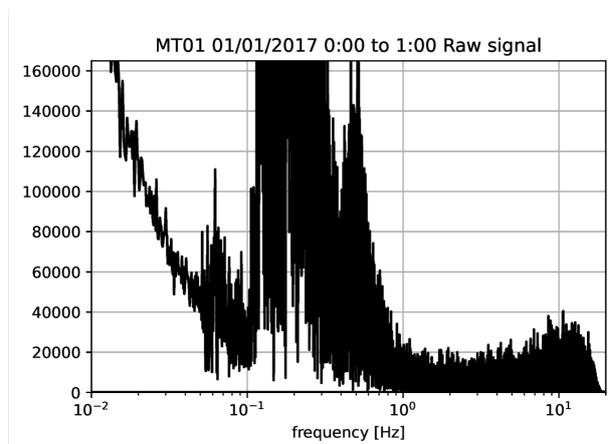
Ahora hay que obtener la función de Green entre un par de estaciones. Calculamos la función de correlación para cada segmento de una hora y los promediamos todos para así obtener la función de correlación de un día. Luego estos se agrupan cronológicamente en una matriz llamada “matriz de correlación”. En este proceso, consideramos una estación como “fuente virtual” del ruido ambiental mientras que la otra estación será considerada el “sensor virtual”.

En la correlación cruzada existe un parte causal y una parte acausal, divididas por “time lag” igual a 0. La parte causal corresponde a un time lag positivo, es decir ondas que van desde la fuente virtual hacia el sensor virtual, mientras que la parte acausal corresponde a la mitad opuesta en la matriz y sentido opuesto en el viaje de las ondas. Aunque ambas secciones difieren en magnitudes, sus peaks principales son simétricos respecto a 0. Por este motivo promediaremos la correlación cruzada como si la plegáramos por el eje de 0, para así resaltar las señales significativas y reducir la influencia del ruido ambiental que no aporta.

La coda es lo que sigue al peak principal, alejándose de time lag  $t_l = 0$  mientras reduce su amplitud. Esta coda está dominada por ondas superficiales que fueron dispersadas por el medio y adquieren información del medio, gracias a esto, con



**Figura 4.2.1:** Serie de tiempo de la estación MT01 el día 1 de enero de 2017. En 4.2.1a se ve antes de cualquier clase de preprocesamiento. En 4.2.1b ya se le removió la media y la inclinación lineal y en 4.2.1c por un filtro pasabanda. En 4.2.1d se le aplicó normalización espectral sobre todo lo anterior.



**Figura 4.2.2:** Transformada de Fourier para una serie de tiempo de la estación MT01 el día 1 de enero de 2017. En 4.2.2a se ve antes de cualquier clase de preprocesamiento. En 4.2.2b ya se le removió la media y la inclinación lineal, además de haber pasado por un filtro pasabanda. En 4.2.2c se le aplicó normalización espectral sobre todo lo anterior.

la función de correlación podemos calcular variaciones de velocidad relativa en función del tiempo, asociadas a el medio alrededor de las estaciones por el que viajan las ondas sísmicas.

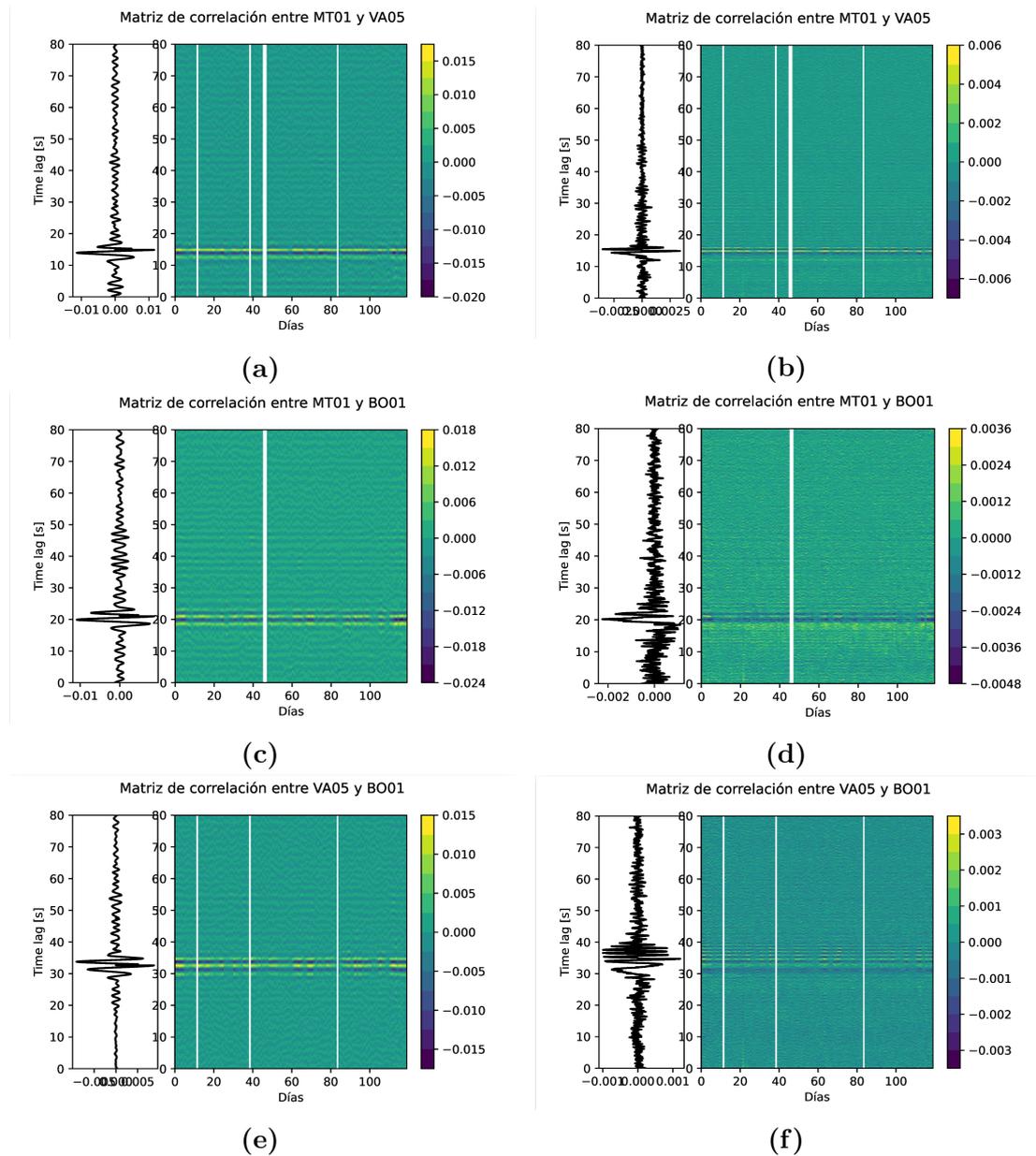
### 4.3.1. Correlación cruzada

Una vez completado el pre-procesamiento, se debe calcular la correlación cruzada entre estaciones para cada día y juntarlas formando una matriz. El rango de frecuencias a utilizar puede variar dependiendo del objeto que se está estudiando y de la distancia entre ambas estaciones. El segundo motivo es el que importa en este caso, lo que significa que mientras más alejadas se encuentren las estaciones, es mejor utilizar frecuencias menores, esto se debe a que para las ondas de menor frecuencia es más fácil alcanzar mayores distancias, mientras que las de frecuencias más altas se disipan con mayor facilidad y al no alcanzar ambas estaciones ensucian los datos, empeorando la calidad de la correlación cruzada. En la figura 4.3.1 se aprecia el uso de dos intervalos de frecuencia para 3 pares de estaciones, a la izquierda usando filtros entre 0,4 y 2[Hz] y a la derecha entre 1 y 4[Hz]. Como se dijo anteriormente, cada par tiene distintas distancias entre estaciones y de entre estas el par MT01-VA05 es el más cercano. De esta forma se aprecia en 4.3.1b que aún mantiene una forma clara mientras que los otros dos pares de estaciones presentan una señal sucia y con ruido, puesto que las altas frecuencias de dispersan a menor distancia. Además en 4.3.1a se observa como esta pierde mucha más información al utilizar frecuencias menores como límites del filtro pasabanda. De manera opuesta, los otros dos pares de estaciones se ven más definidos al utilizar el filtro pasabanda entre 0,4 y 2[Hz].

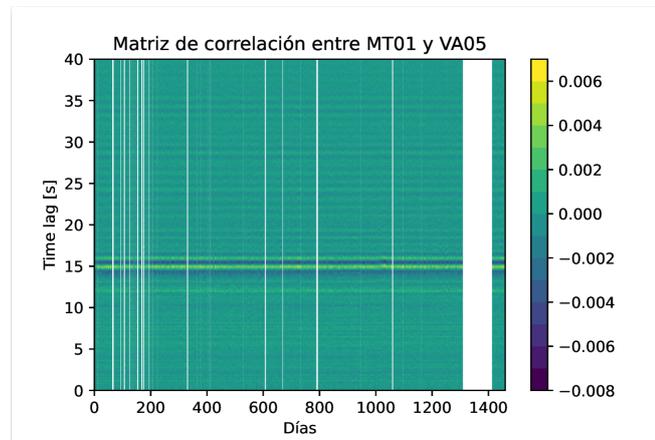
Hay días en los que la calidad de la función de correlación o la cantidad de datos son muy bajas y por tanto fueron dejados como espacios en blanco en la matriz de correlación. Consideramos validos entonces solo los días que tengan datos representativo de al menos dos tercios del día, es decir equivalentes a 16 horas.

## 4.4. Factor de estrechamiento

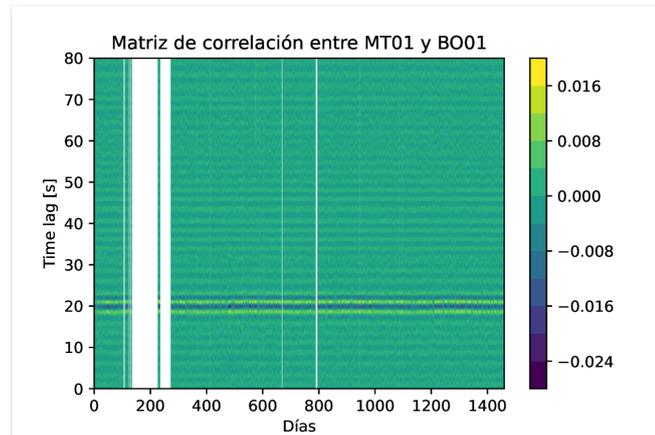
Las matrices de correlación pueden ser simétricas o asimétricas, las cuales se definen según la distribución de energía de la onda en la matriz de correlación. La matriz es simétrica si la onda presenta la misma energía en los peaks de la parte



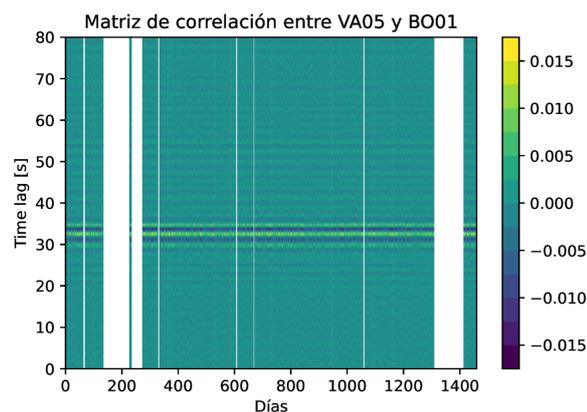
**Figura 4.3.1:** Comparativa para matrices de correlación de 3 pares de estaciones utilizando filtros pasabanda para dos intervalos de frecuencia: a la izquierda 0,4 y 2[Hz] y a la derecha entre 1 y 4[Hz]. Además a la izquierda de cada matriz se encuentra la traza de referencia de la matriz.



(a) MT01 - VA05: Bandpass entre 1 y 4[Hz].

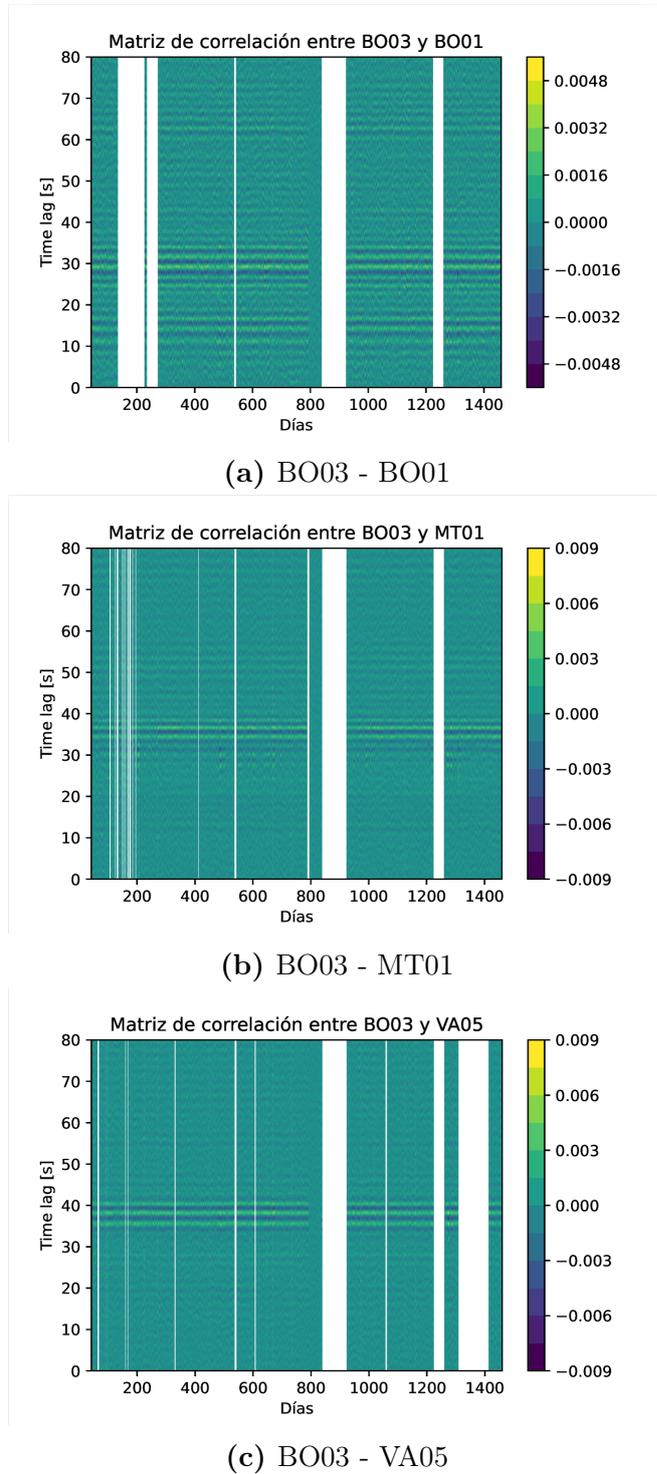


(b) MT01 - BO01: Bandpass entre 0,4 y 2[Hz].



(c) VA05 - BO01: Bandpass entre 0,4 y 2[Hz].

**Figura 4.3.2:** Matriz de correlación para los 3 pares de estaciones de la figura 4.3.1. Van desde el 1 de enero de 2016 hasta el 31 de diciembre de 2019, completando cuatro años de correlaciones diarias. En 4.3.2a se hace un acercamiento en el eje Y, dejando el time lag entre 0 y 40 segundos en vez de 80.



**Figura 4.3.3:** Matriz de correlación para los 3 pares de estaciones restantes, todas filtradas en frecuencias entre  $0,4$  y  $2[Hz]$ , los cuales también van desde el 11 de febrero de 2016 hasta el 31 de diciembre de 2019.

causal y acausal de la matriz. Por otro lado, la matriz es asimétrica cuando el peak de energía de una de las mitades es mucho mayor que el otro. Normalmente se promedia la mitad causal con la acausal solo en matrices simétricas, pero en este caso ese procedimiento se realizará con ambos tipos de matrices.

Procedemos a calcular la traza de referencia. La traza de referencia es el promedio de las funciones de correlación de la matriz para cada día y la utilizaremos para compararla con cada día y calcular su factor de estrechamiento.

Luego se selecciona la ventana de tiempo en la que se estudiará la coda y se compara con el mismo segmento en la traza de referencia. Así, a medida que se estira o contrae la función de correlación, se calcula el valor de correlación entre ambas trazas. Cuando esta correlación alcance su mayor valor, lo más cercano a 1 posible, el factor de estrechamiento asociado es el que nos entrega la información del cambio de velocidad de las ondas superficiales. La relación entre el factor de estrechamiento y el cambio de velocidad es:

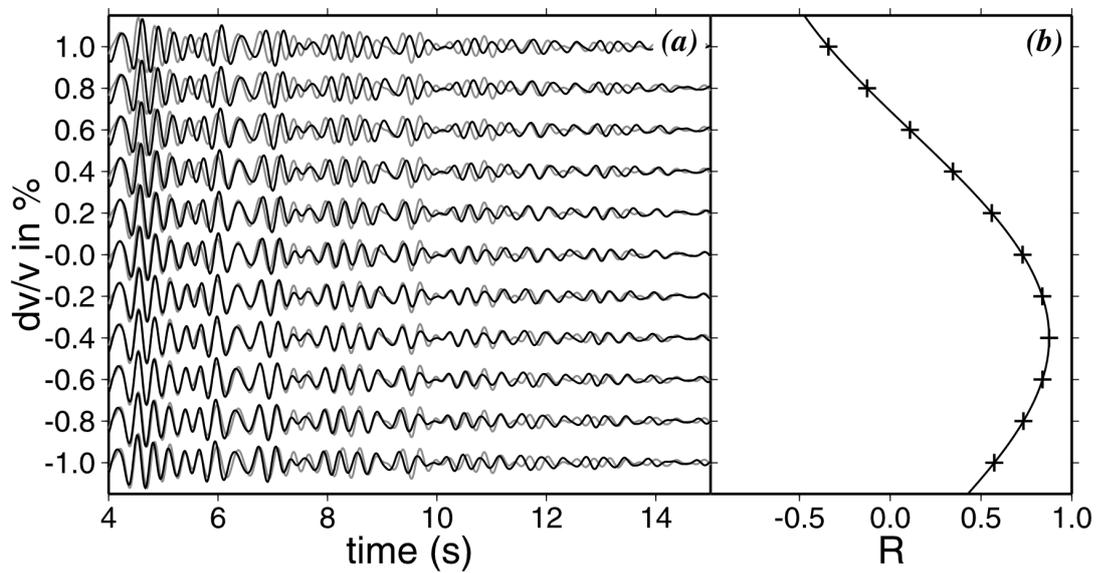
$$\frac{dt}{t} = -\frac{dv}{v} \quad (4.4.1)$$

Donde  $\frac{dt}{t}$  es el estrechamiento y representa la variación en el tiempo de viaje. Según [Schober et al. \(2018\)](#), se puede considerar como significativa una correlación con un coeficiente de 0.65, sin embargo para redondear, consideraremos que el procedimiento anterior se considera válido y relevante para cada día que la correlación máxima entre la traza de referencia y la traza del día sea superior a 0.6 como mínimo.

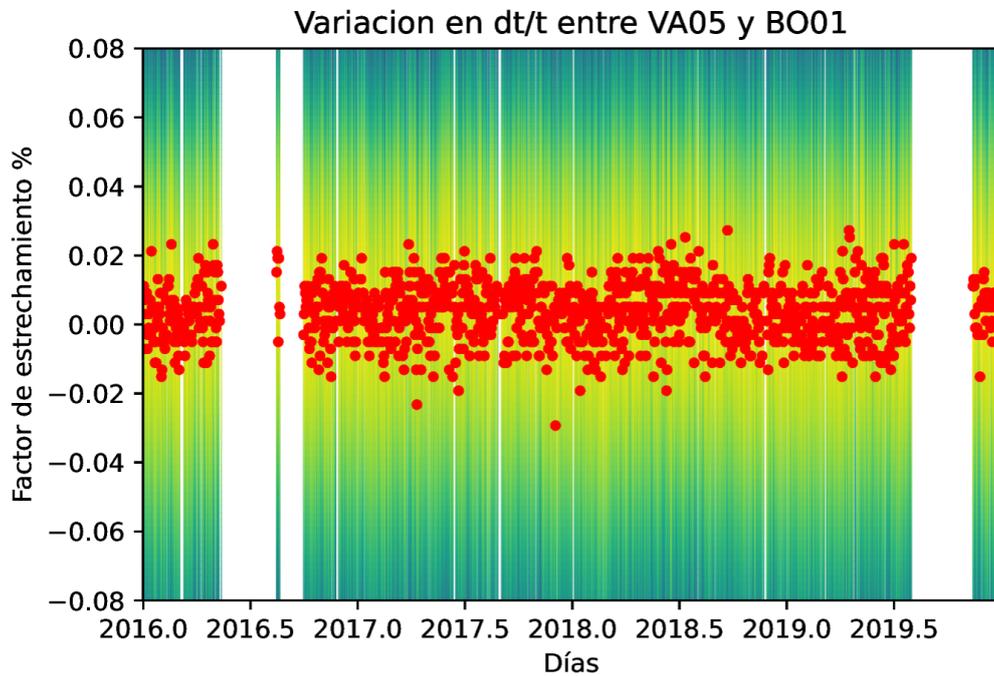
#### 4.4.1. Análisis

Para determinar la ventana de tiempo más apropiada para la coda hay que probar a calcular el factor de estrechamiento y la correlación para distintas ventanas de tiempo en la coda. La que muestre las mejores correlaciones es la ventana de tiempo más apropiada. Sin embargo es necesario tener en cuenta que de ser muy larga el ruido disperso por fuentes externas puede ensuciar la señal, y si es muy corta es posible que se pierda información importante de la coda.

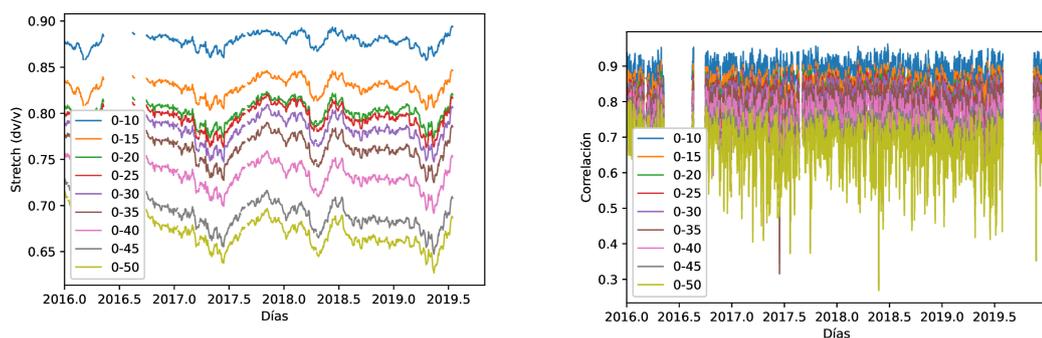
En la figura [4.4.3b](#) se aprecia que la correlación es mayor con ventanas más pequeñas de tiempo, además en las imágenes de la figura [4.3.1](#) podemos observar



**Figura 4.4.1:** Ejemplo de obtención de factor de estrechamiento utilizando autocorrelación en la estación KZK [Wegler et al. \(2009\)](#). En (a) se observan 11 trazas que fueron expandidas o comprimidas y en gris, detrás de cada una, la traza de referencia con la que se compara. En (b) se observa la correlación entre las 11 trazas distintas con la traza de referencia. Se observa que la correlación alcanza su máximo con un factor de estrechamiento de  $-0,4\%$ , llegando a una correlación máxima de  $R = 0,88$ .



**Figura 4.4.2:** Variación del factor de estrechamiento diario para el par de estaciones VA05-BO01 entre los años 2016 y 2019. La matriz representa el nivel de correlación considerando cada factor de estrechamiento, mientras que los puntos rojos representan el máximo diario. Se utilizó una ventana de tiempo de 15 segundos para la coda



(a) Variación del factor de estrechamiento a través del tiempo

(b) Correlación máxima a través del tiempo.

**Figura 4.4.3:** En 4.4.3a se ve como varía el factor de estrechamiento a lo largo de los 4 años para el mismo par de estaciones de 4.4.2. Se le aplicó una media móvil de 60 días para suavizar el gráfico y hacerlo más claro de visualizar. En 4.4.3b se observa la correlación máxima para cada día respecto a la traza de referencia. La leyenda en cada figura corresponde a la ventana de tiempo de la coda utilizada, en segundos.

que las coda tienen en su mayoría una duración de aproximadamente 10 segundos, por estos motivos esta será la ventana de tiempo que utilizaremos para calcular el factor de estrechamiento.

# Capítulo 5

## Resultados

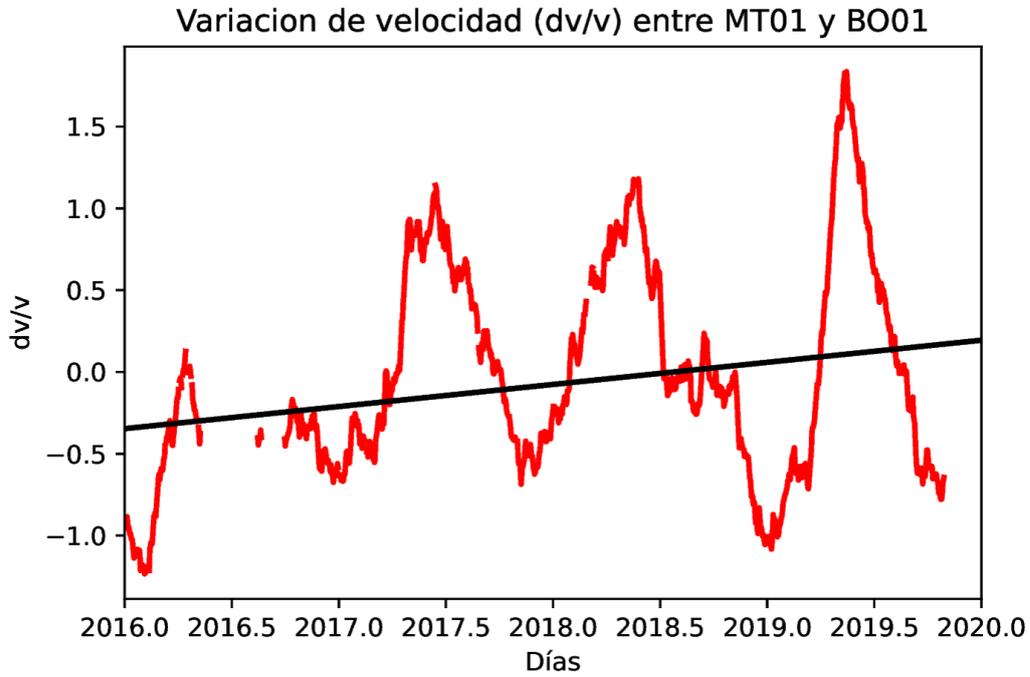
En este capítulo se calculará la variación relativa de velocidad de las ondas sísmica  $dv/v$  a partir de  $dt/t$ , utilizando la relación planteada en la ecuación 2.3.13, la cual es:

$$\frac{dt}{t} = -\frac{dv}{v}$$

El término  $dt/t$  corresponde al factor de estrechamiento calculado en la sección 4.4 y mostrado en la figura 4.4.3a. Desde este punto expresaremos  $dv/v$  en porcentaje, para esto la multiplicaremos por 100. Además centramos la curva en 0 restando le media de todos los datos diarios.

También calculamos la regresión lineal, la cual representa le tendencia de crecimiento o decrecimiento de la velocidad a lo largo del tiempo.

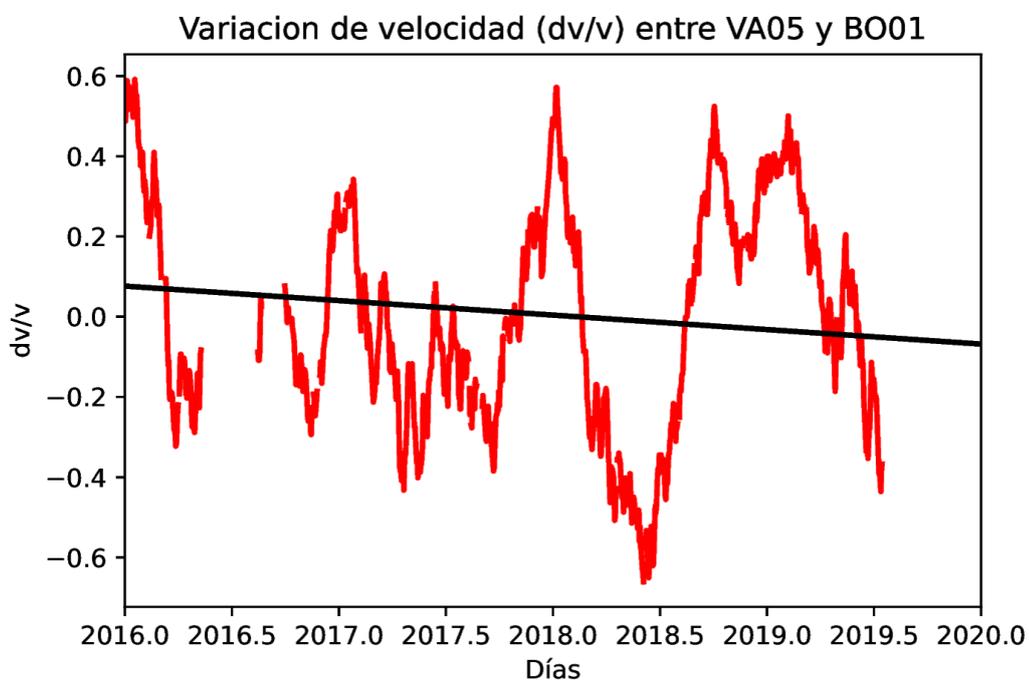
### 5.0.1. MT01-BO01



**Figura 5.0.1:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre MT01-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra.

En la figura 5.0.1 se observan tres picos en los que aumenta la velocidad de las ondas, el primero se encuentra a mediados de 2017, el segundo a mediados de 2018 y el último a mediados de 2019. El coeficiente de la regresión lineal indica un aumento del 0,135 % por año, una tendencia bastante pronunciada.

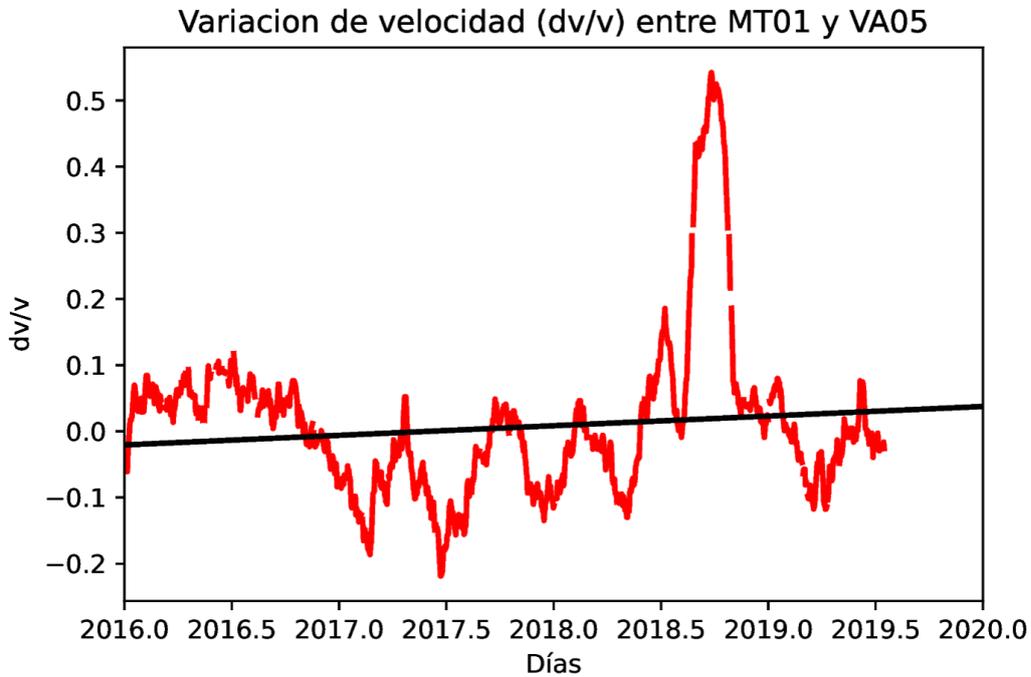
### 5.0.2. VA05-BO01



**Figura 5.0.2:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre VA05-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra..

Para el par de estaciones VA05-BO01, los datos presentan peaks menos claros que en el primer par, además estos se encuentran en distintas secciones del año, alcanzando su máximo poco después del inicio de año en vez de a mitad de año. La regresión lineal tiene un coeficiente de  $-0,036\%$  de variación por año, es decir un orden de magnitud menor respecto al anterior par de estaciones, esto significa que que la variación a lo largo de los años fue mucho menor en comparación.

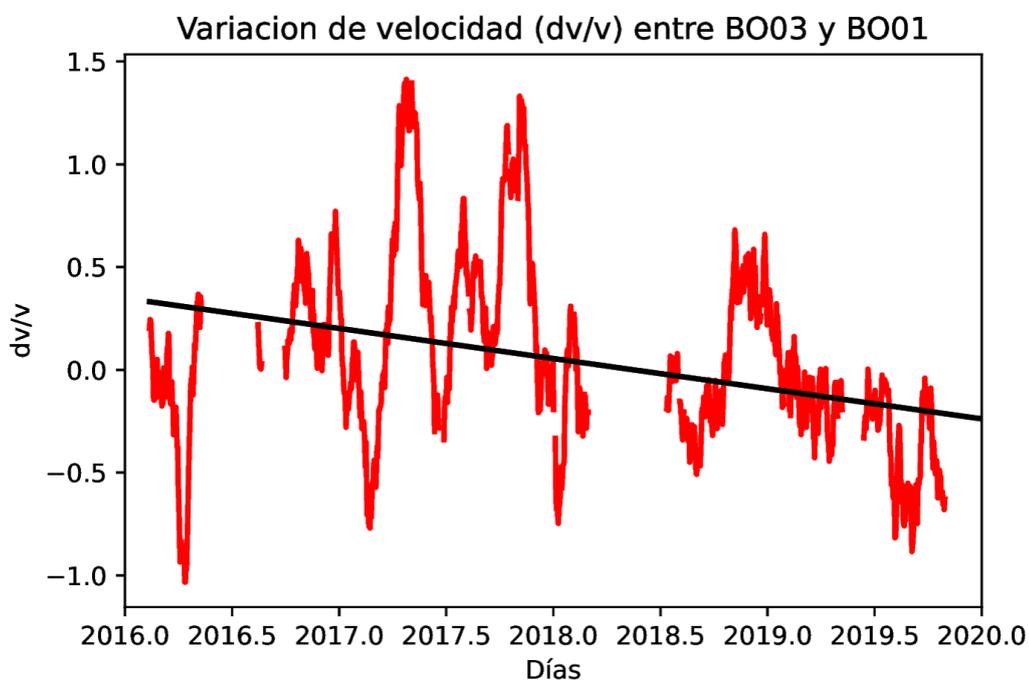
### 5.0.3. MT01-VA05



**Figura 5.0.3:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre MT01-VA05 como un gráfico rojo y tendencia lineal de los peaks como una recta negra.

En este par de estaciones vemos un comportamiento donde los datos se mantienen muy cerca de del 0 durante la mayor parte del tiempo, teniendo solo un peak muy notorio durante la segunda mitad del 2018. Esto puede deberse a que es el par de estaciones con menor distancia entre ellas y por tanto hay menos posibilidades para que las ondas se disipen. Con respecto a la regresión lineal, esta tiene un coeficiente de 0,014%, siendo aún menor que el anterior pero aún permaneciendo dentro del mismo orden de magnitudes.

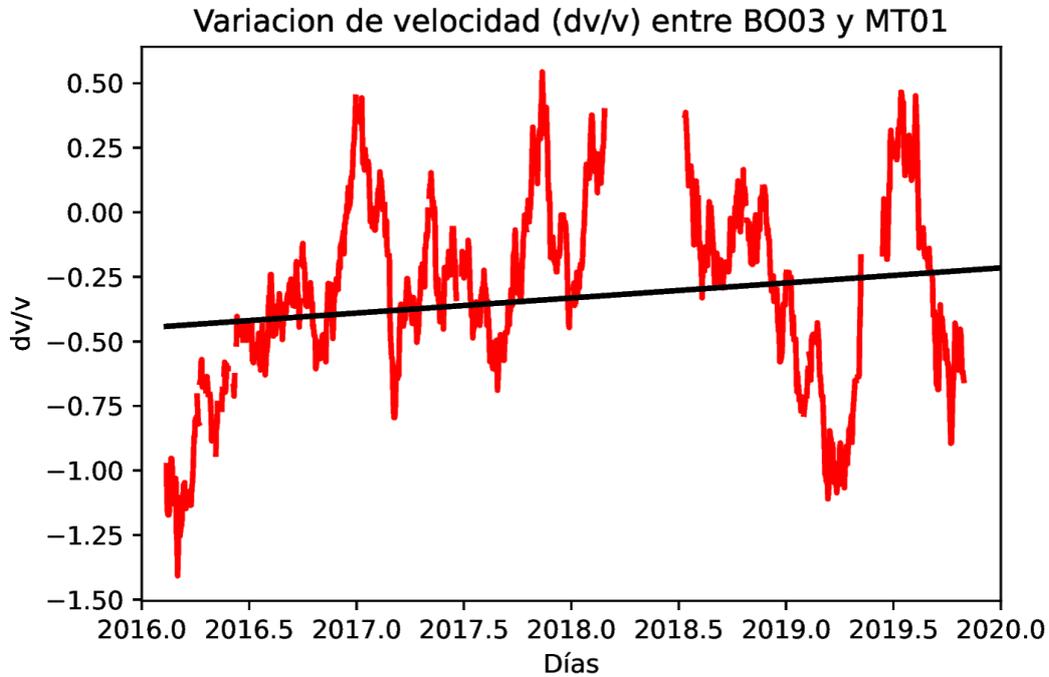
#### 5.0.4. BO03-BO01



**Figura 5.0.4:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre BO03-BO01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra.

Empezando con las parejas que incluyen BO03 tenemos BO03-BO01. En esta, vemos que los peaks son bastante grandes pero, a diferencia de MT01-BO01 y VA05-BO01, estas no parecen guiarse por un patrón anual. Además la regresión lineal tiene una pendiente muy marcada, con un coeficiente de  $-0,147\%$  por año.

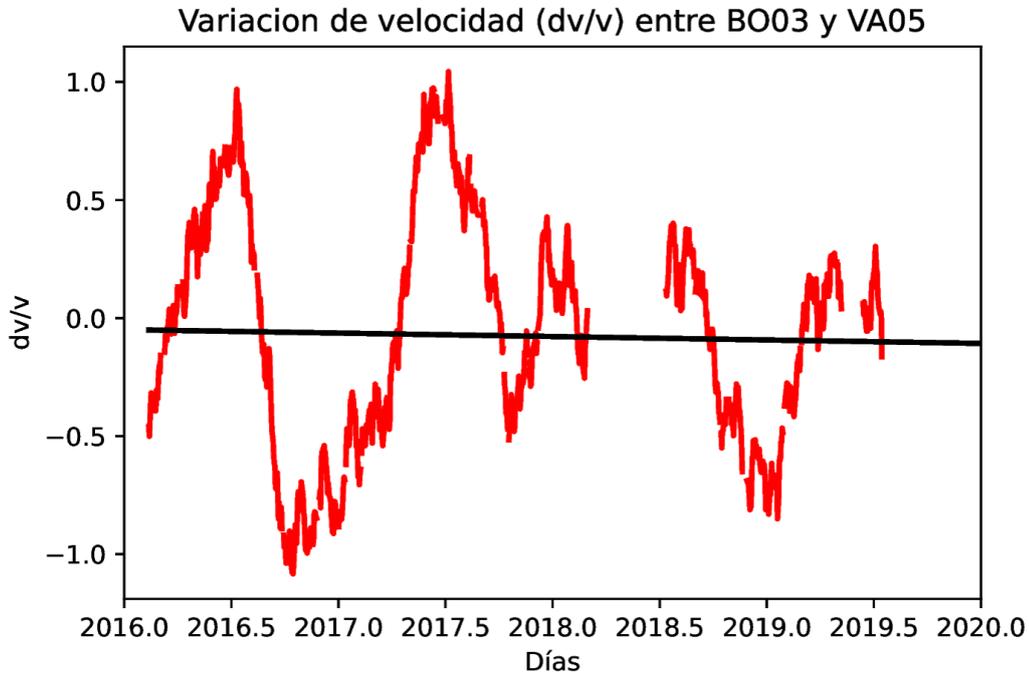
### 5.0.5. BO03-MT01



**Figura 5.0.5:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre BO03-MT01 como un gráfico rojo y tendencia lineal de los peaks como una recta negra..

En este par de estaciones, BO03-MT01, se observan datos muy dispersos similar al par anterior, y de la misma forma los peaks parecen no estar relacionados con el ciclo anual. La regresión lineal tiene un coeficiente de 0,058 %, lo que quiere decir que tiende a aumentar ligeramente en los años.

### 5.0.6. BO03-VA05



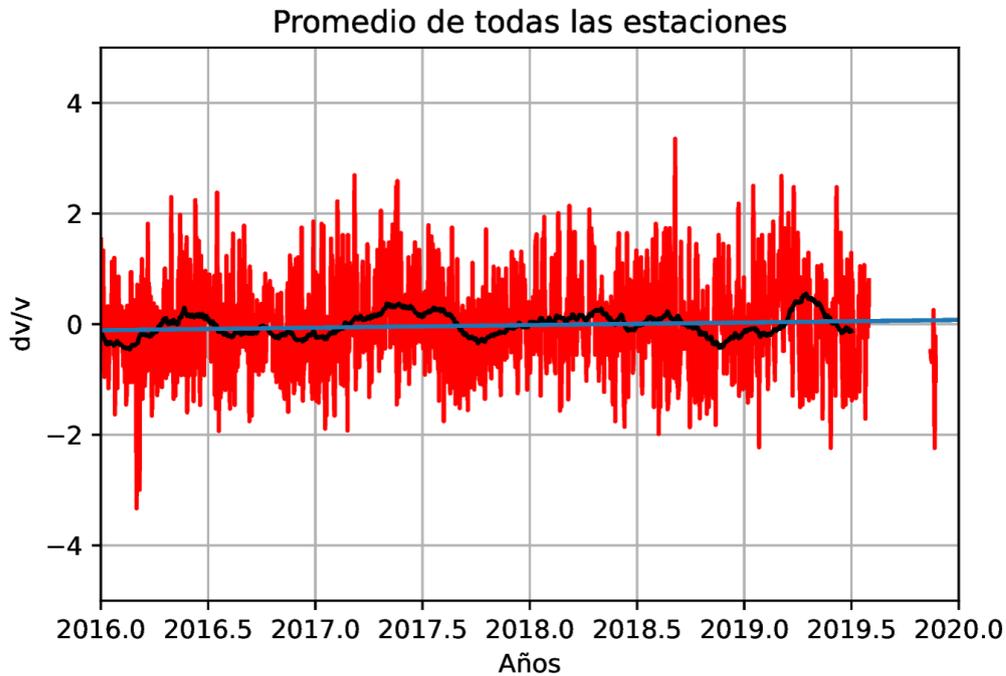
**Figura 5.0.6:** Media móvil a lo largo de 4 años para el el  $dv/v$  entre BO03-VA05 como un gráfico rojo y tendencia lineal de los peaks como una recta negra.

En el último par de estaciones estudiado, BO03-VA05, vemos que los datos presentan un comportamiento anual con peaks a mediados de año, tal como es el caso de MT01-BO01, al menos para los años 2016 y 2017, mientras que en 2018 y 2019 no es posible observar estos peaks por falta de datos, si es posible observar los valles entre ellos. La regresión lineal tiene un coeficiente de  $-0,015\%$ , encontrándose en el mismo orden de magnitud que los otros dos pares de estaciones que incluían la estación VA05.

## 5.1. Promedios

Teniendo los resultados para los seis pares de estaciones, ahora podemos promediarlos. El resultado de esto da información de la zona comprendida entre las cuatro estaciones utilizadas para la prueba y sus alrededores, pues las ondas dispersadas desde fuera también afectan (Lecocq et al., 2017). Solo consideramos validos días que tengan datos de 3 estaciones o más para asegurar que no se

produzcan picos repentinos producto de demasiada influencia de un solo par de estaciones.

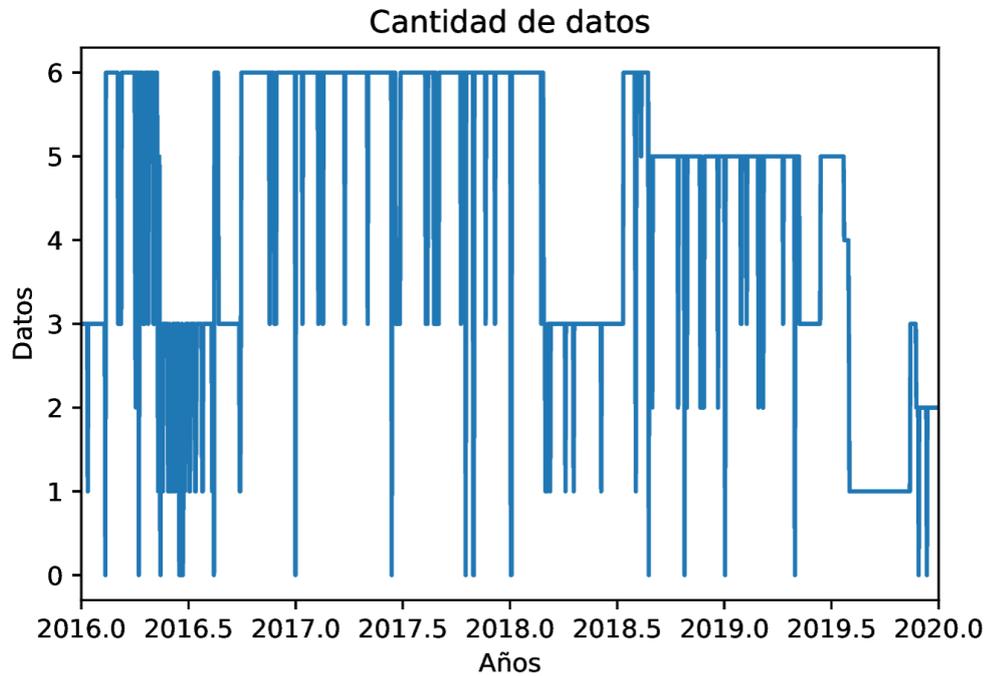


**Figura 5.1.1:** En rojo gráfico que promedia la variación de velocidad de los 6 pares de estaciones. En negro la media móvil sin pesos que abarca 60 días y en azul la tendencia lineal.

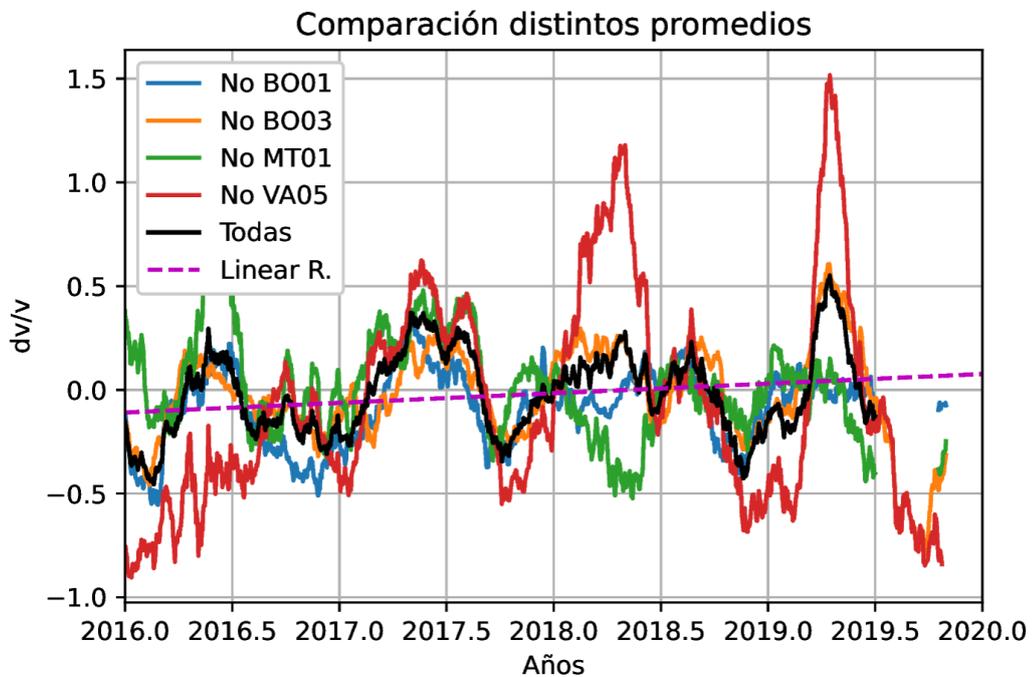
Luego, para realizar una prueba de estabilidad de los resultados, utilizamos el método de dejar fuera uno o de la navaja (Efron and Gong, 1983). Este consiste en dejar fuera una estación, reduciendo así los 6 pares de estaciones a tan solo 3, quedando tan solo con el 50 % del espacio muestral.

	Coefficiente	Prom. cantidad de datos
Todos los pares	0.04648008	4.33
Excluyendo BO01	0.04615481	2.05
Excluyendo BO03	-0.00034987	2.51
Excluyendo MT01	-0.09634996	2.13
Excluyendo VA05	0.10179704	1.98

**Cuadro 5.1.1:** Pendientes de regresión lineal para  $dv/v$  junto a la cantidad promedio de datos diarios para el resultado de promediar todas las estaciones y luego los resultados de utilizar el método de la navaja.



**Figura 5.1.2:** Gráfico que representa la cantidad de datos para cada día a lo largo de los 4 años de datos.



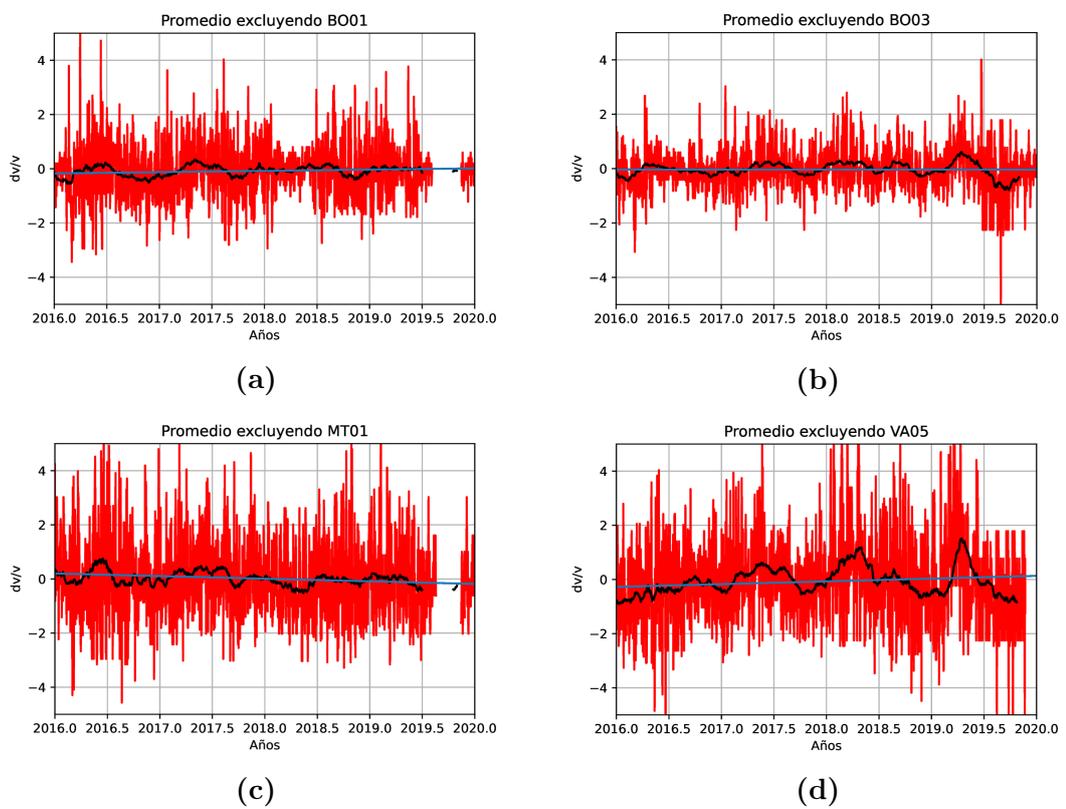
**Figura 5.1.3:** Gráfico que compara la media móvil que incluye los 6 pares de estaciones en negro, presentada en 5.1.1, mientras que en colores se encuentran los resultados de aplicar el método de la navaja.

Tras aplicar el método de la navaja observamos que las variaciones son bastante consistentes entre el promedio de todas las estaciones y las variaciones sin una estación, con excepción del periodo que va desde mediados de 2018 a mediados de 2019, en donde la variación es mucho mucho mayor. En cada caso se pueden observar cosas distintas:

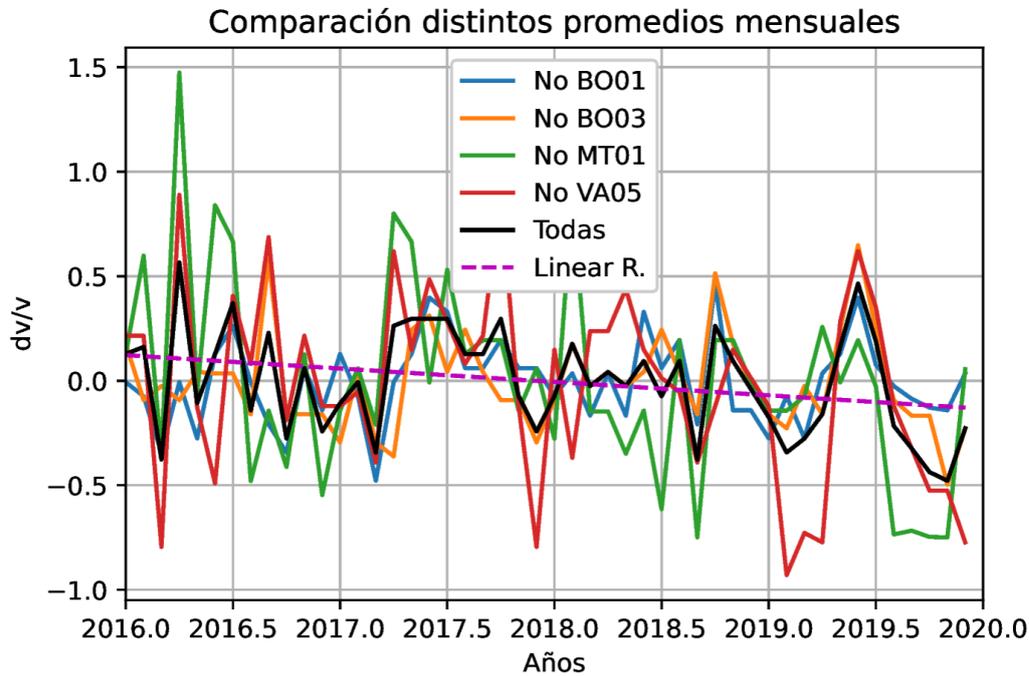
- Al extraer BO03 el gráfico no varía mucho. Esto se debe probablemente a que es la estación que menos datos aportaba y por tanto tendrá menos peso al calcular el promedio.
- Al extraer VA05 las curvas ocurren aproximadamente en los mismos intervalos de tiempo, sin embargo las variaciones son mucho más pronunciadas. Esto era un resultado esperable puesto que los 3 pares de estaciones que la incluían presentaban poca variación en la velocidad, especialmente MT01-VA05.
- Tanto al extraer BO01 o MT01 es donde más discrepancias se producen en el periodo de tiempo entre mediados de 2018 y 2019 anteriormente mencionado.

### 5.1.1. Promedios mensuales

Además de trabajar con resultados diarios y media móvil, también realicé un estudio para promedios mensuales de las correlación cruzada, es decir previo al calculo del factor de estiramiento, a diferencia de la media móvil. Esto sigue la misma lógica que el promediar series de tiempo de una hora para reforzar la información útil y reducir el ruido que no aporta nada. De esta forma en la figura 5.1.5 observamos de manera general el comportamiento general, teniendo el promedio de las 6 estaciones. Por otro lado, en el cuadro 5.1.2 y las figuras en 5.1.6 podemos ver de manera algo más detallada los datos para cada caso del método de la navaja.



**Figura 5.1.4:** Gráficos que comparten la misma estructura que 5.1.1 pero se diferencian en que se ignora una de las 4 estaciones diferentes en cada uno, es decir que cada gráfico es solo el promedio entre 3 pares de estaciones.

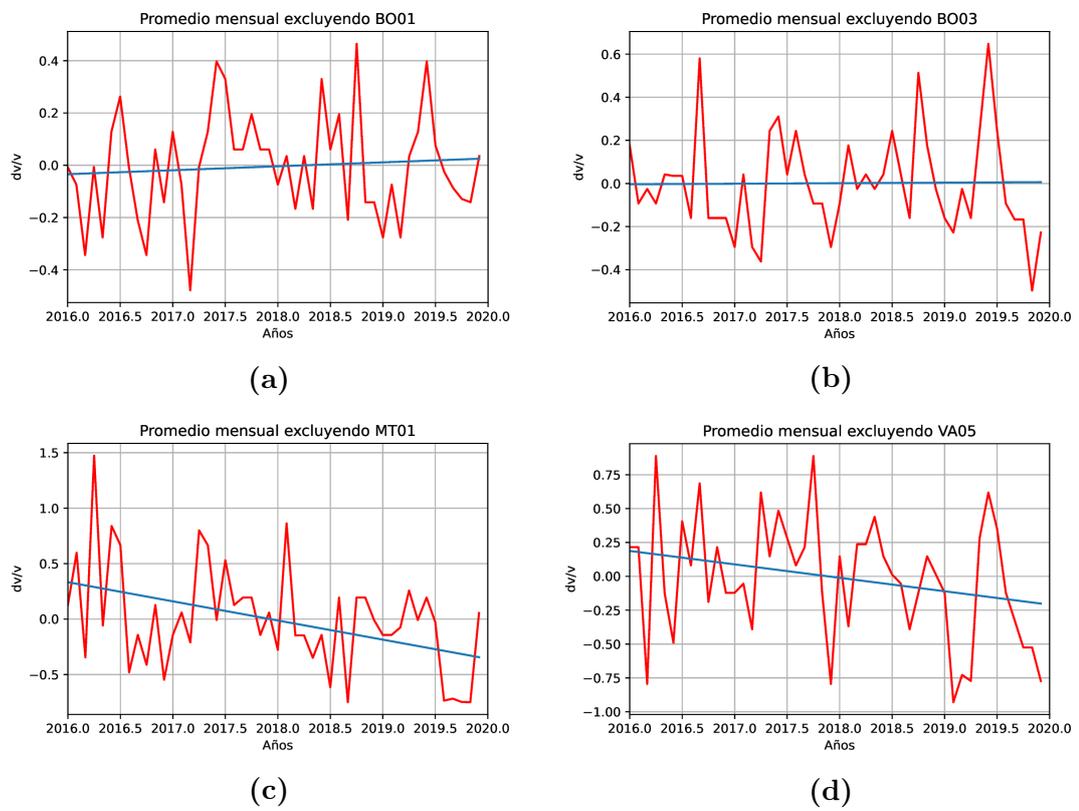


**Figura 5.1.5:** Gráfico que compara la media móvil que incluye los 6 pares de estaciones en negro con gráficos de distintos colores que representan los distintos resultados de aplicar el método de la navaja. Por otro lado la línea punteada representa la tendencia general a lo largo de los años.

	Pendiente
Todos los pares	-0.06406397
BO01	0.01508888
BO03	0.00254249
MT01	-0.17264247
VA05	-0.09926326

**Cuadro 5.1.2:** Pendientes de la regresión lineal para  $dv/v$  para el resultado de promediar todas las estaciones y luego los resultados de utilizar el método de la navaja.

Mientras que los datos diarios con media móvil nos presentan un comportamiento mucho más claro de los procesos anuales, los promedios mensuales nos entregan una mejor visión de la tendencia a largo plazo.



**Figura 5.1.6:** Gráficos con promedios mensuales que ignoran una de las 4 estaciones diferentes en cada uno.

# Capítulo 6

## Discusión

Si vemos la figura 4.1.1 podemos observar que la estación BO01 es la más lejana a la costa, encontrándose en lo que llamamos depresión central, mientras que las otras 3 estaciones se encuentran dentro de la cordillera de la costa. Además, BO01 es también la estación que presenta los resultados con mayor diferencia al resto en lo que respecta a la tendencia en la variación de velocidad a lo largo del tiempo, según la tabla 5.1.2. Esto significa que las cercanías sufrieron un mayor cambio a lo largo de los años.

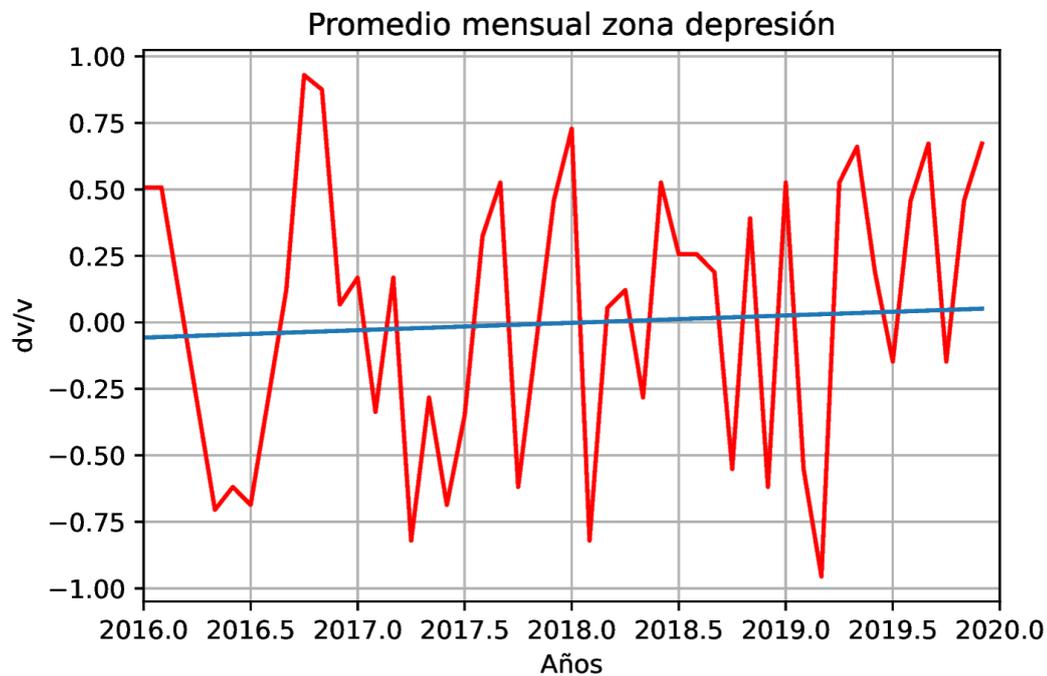
Juntando ambos hechos, tanto la posición de la estación BO01 como la gran variación en la tendencia al excluirla de los promedios, podemos intuir que la variación en la velocidad de propagación de las ondas sísmicas es mucho más notoria en la depresión central, mientras que en la zona costera y de la cordillera de la costa, se ha mantenido relativamente estable a lo largo de los años. Esto puede deberse a diferencias en la geología de ambas zonas, como un suelo relleno de sedimentos que permita una mayor infiltración y almacenamiento de aguas subterráneas o una mayor penetración de los cambios de temperatura.

Para comprobar esto realizamos el mismo estudio, añadiendo ahora la estación BO04 y BO02, la cual se encuentra más al noreste y sudeste de la estación BO01, respectivamente, con el objetivo de tener más mediciones en el espacio perteneciente a la depresión central.

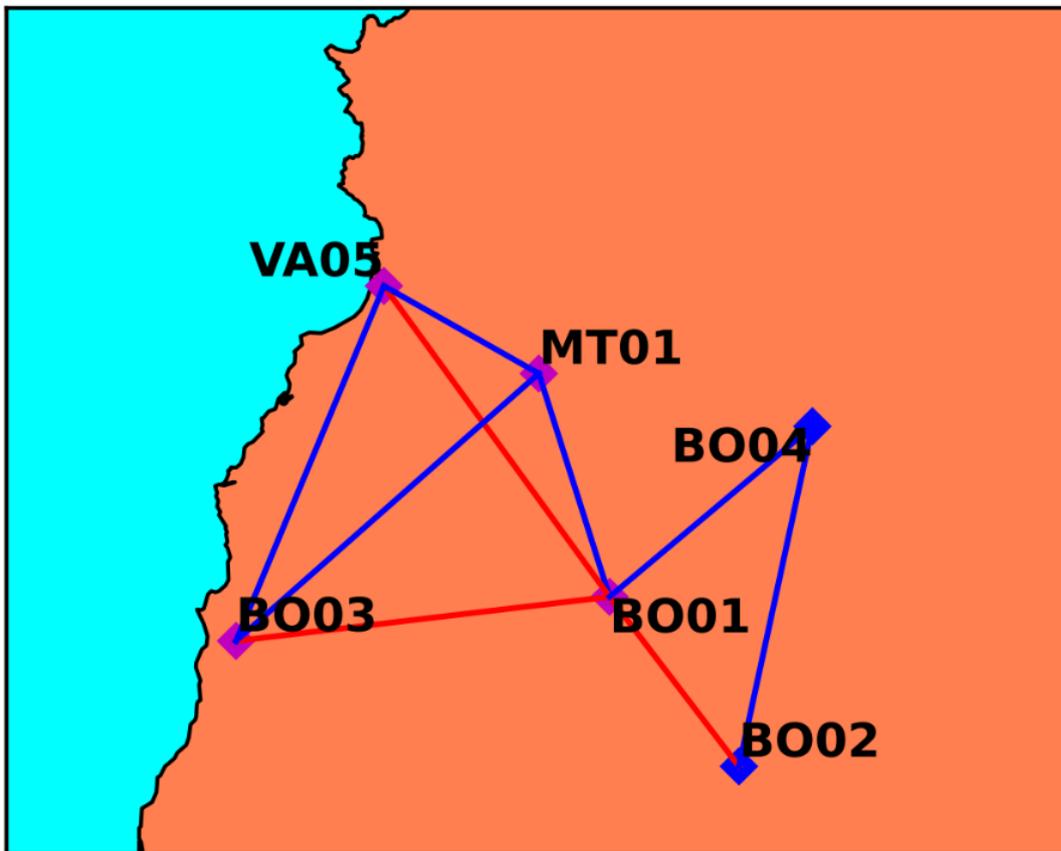
Tras observar los resultados en la tabla 6.0.1 podemos notar que, en realidad, este no es el caso, ya que los resultados son muy variados, y lo que es más, el promedio de la variación de velocidad general de la zona es positivo, de la misma forma que

	Coefficiente
Depresión central	0.0277352
BO01-BO02	-0.00257821
BO01-BO04	0.05888848
BO04-BO02	0.17668443

**Cuadro 6.0.1:** Pendientes de regresión lineal para  $dv/v$  para los 3 pares de estaciones incluidos ahora además de la zona de depresión central, como complemento a la tabla 5.1.2



**Figura 6.0.1:** Variación mensual para el promedio de los 3 pares de estaciones utilizados para estudiar la zona de la depresión central.



**Figura 6.0.2:** Mapa con todas las estaciones, incluyendo las presentes en 4.1.1, representados nuevamente con diamantes morados, junto a las las 2 incluidas en esta sección, como diamantes azules. Además se incluyen los coeficiente observado en las tablas 6.0.1 simbolizados con líneas que unen ambas estaciones. Líneas rojas representan variaciones negativas en la velocidad y líneas azules variaciones positivas.

el promedio que excluye a BO01 en la tabla 5.1.2, el cual representaría a la zona costera.

Sin embargo, tras observar el mapa de la figura 6.0.2, pareciera que en realidad la tendencia negativa en la variación de velocidad parece estar principalmente más ubicada hacia el sur de la zona estudiada.

## 6.1. Relación con variables meteorológicas

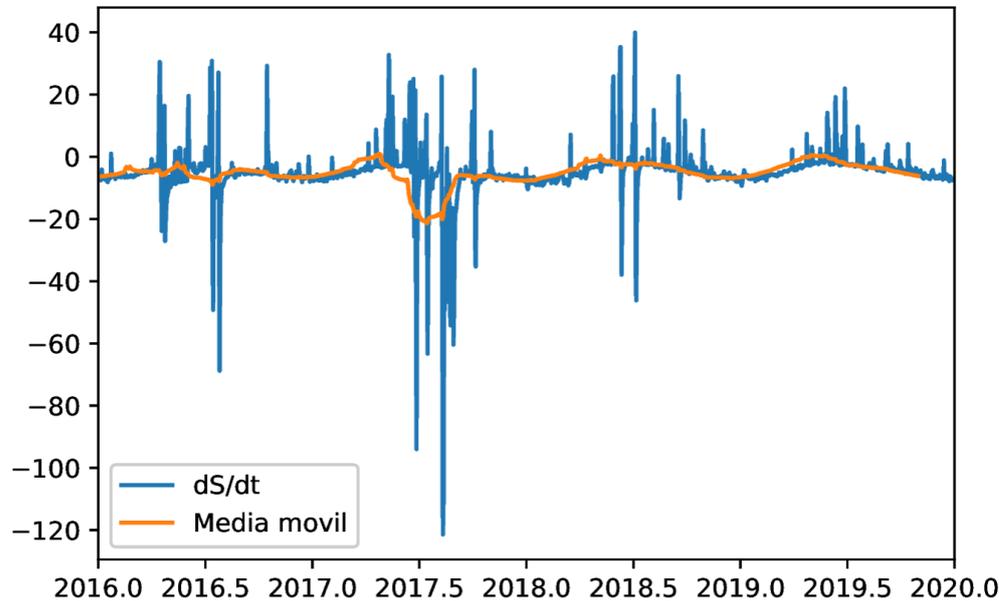
El resultado de promediar la serie de datos de todas las estaciones representa el comportamiento general de la zona. De esta forma nos es posible comparar y relacionar las variaciones de velocidad producidas a otras causas, en este caso con los datos climatológicos e hidrológicos (Hillers et al., 2014) expuestos en la sección 3.1. La temperatura afecta de manera directa en la presión termo-elástica (Turcotte and Schubert, 2014), por lo que es posible realizar la correlación sin mayor procesamiento siempre y cuando se utilice el delay adecuado. Sin embargo para datos hidrológicos debemos utilizar la siguiente ecuación (Lecocq et al., 2017):

$$\frac{dS}{dt} = P - E - R \quad (6.1.1)$$

donde  $S$  corresponde al almacenamiento de agua,  $P$  a las precipitaciones,  $E$  a la evotranspiración y  $R$  al caudal registrado. Al realizar el calculo con un promedio de las 3 mediciones en cuencas expuestas en 3.1 obtenemos los resultados expuestos en la figura 6.1.1. Tras realizar una correlación entre los datos de temperatura y los datos de almacenamiento de agua con respecto a los datos de variación de velocidad tenemos los siguientes resultados. Tras hacer una correlación observamos los siguientes resultados:

1. Correlación temperatura: -0.663
2. Correlación almacenamiento de agua: 0.281

Observamos que tanto almacenamiento de agua presenta una correlación positiva, lo que quiere decir que es directamente proporcional a la variación de velocidad, sin embargo la variación de temperatura presenta una correlación negativa, lo que significa que la variación de velocidad es directamente proporcional al inverso aditivo de la temperatura. Tras expandir el proceso a una correlación cruzada, la cual se encuentra graficada en las figuras 6.1.2, observamos la temperatura,



**Figura 6.1.1:** Datos de almacenamiento de agua a lo largo del tiempo para la zona estudiada, junto a su media móvil.

además de ser inversamente proporcional como ya se había mencionado, necesita un pequeño delay para alcanzar la correlación negativa máxima, mientras que los datos de de almacenamiento de agua necesitan un delay mucho mayor, de casi un año, para alcanzar su valor máximo.

Empezando con la temperatura, para remediar esto la reemplazamos por una función sinusoidal, la cual tiene periodo de un año y un delay de 0,3 años.

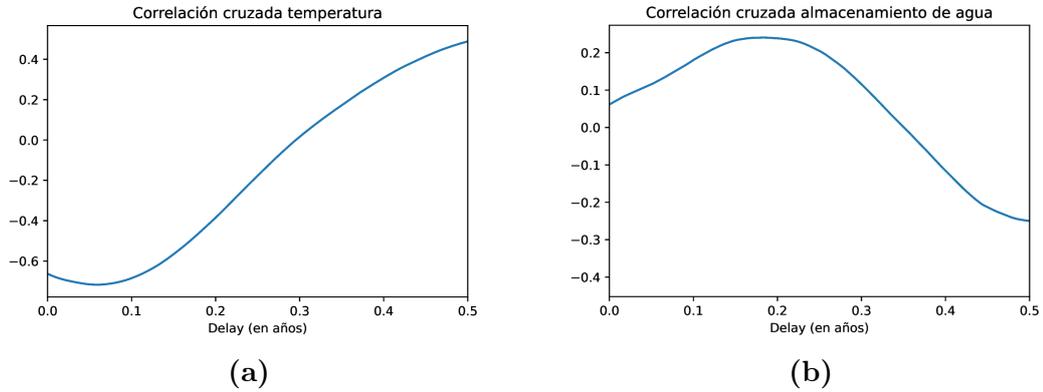
```
np.sin((time+0.3)*np.pi*2)
```

La correlación entre esta función sinusoidal y los datos de temperatura ya registrados es de 0.985, por lo que es perfectamente válido utilizarla como reemplazo al tratar con la variación de velocidad.

Luego, para alcanzar la correlación negativa máxima respecto a la variación de velocidad observada en 6.1.2a aplicamos un delay extra de 0,06.

```
np.sin((time+0.36)*np.pi*2)
```

Ahora lo que buscamos es explicar la variación de velocidad y para esto necesitamos que la correlación sea positiva. Podemos conseguir esto al invertir la función,



**Figura 6.1.2:** Correlación cruzada entre el promedio de todas las estaciones mostrado en el gráfico de la figura 5.1.1 y los datos de meteorológicos de cuencas presentados en las gráficos de la figura 3.1.1.

multiplicándola por  $-1$ , de forma que:

$$\begin{aligned}
 -\sin((t + 0,36) * 2\pi) &= \sin((t + 0,36) * 2\pi + \pi) \\
 &= \sin((t + 0,36 + 0,5) * 2\pi) \\
 &= \sin((t + 0,86) * 2\pi)
 \end{aligned}$$

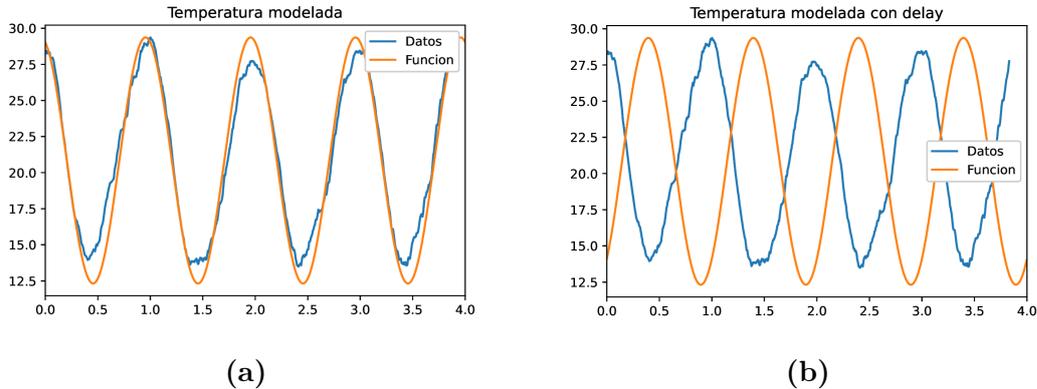
Es decir, para alcanzar la máxima correlación entre temperatura y variación de velocidad usaríamos la función

$$\text{np.sin}((\text{time}+0.86)*\text{np.pi}*2)$$

Utilizando esta función podemos representar de mejor manera la relación entre variación de temperatura y velocidad a lo largo del tiempo, con una correlación de 0.681.

Por otro lado, con el almacenamiento agua tan solo es necesario aplicar un delay de 68 días, y con esto ya alcanza su máximo. Esto implica que las variaciones de temperatura permean de forma más rápida hacia las profundidades que más afectan este experimento, incluso las temperaturas más bajas tienen la capacidad de permear más profundo y tienen mayor permanencia a lo largo de los años (Ben-Zion and Allam, 2013). Por otro lado la absorción de agua es más lenta.

De esta forma podemos ver que la temperatura tiene gran importancia en la variación de velocidad del ruido sísmico a través del suelo, y aunque el almacenamiento de agua influye en esta, es una correlación menor.



**Figura 6.1.3:** En a) tenemos la temperatura en  $^{\circ}C$  modelada con una función seno. En b) la misma función luego de aplicarle el delay extra de 0.06 años (equivalente a 22 días aproximadamente) e invertirla.

Cabe destacar la correlación entre dichas variables meteorológicas y la variación de velocidad de las ondas sísmicas no necesariamente implica causalidad, al menos no a corto plazo, por este motivo es necesario estudiar datos anuales a largo plazo para observar tendencias.

### 6.1.1. Datos anuales

Al trabajar con datos anuales observamos en la tabla 6.1.1 que la velocidad presenta un aumento tanto de 2016 a 2017 como de 2017 a 2018, sin embargo este es menor en comparación, el año 2019 presenta grandes periodos sin datos como se observa en la figura 5.1.2 por lo que podría ser poco representativo y por tanto fue excluido, en caso de que pueda crear confusión al interpretar los datos. Por otro lado la temperatura presenta primero un descenso, lo cual se corresponde con la correlación negativa ya observada, sin embargo luego presenta un pequeño ascenso. Esto sea probablemente compensado por el ascenso que ocurre en el almacenamiento de agua durante el año 2018 de más de  $3 \text{ mm/día}$ .

Además, es posible que los cambios de temperatura a largo plazo tengan efectos tendencias multi-anales (Ben-Zion and Allam, 2013).

	$dv/v$	Temperatura ( $^{\circ}C$ )	Almacenamiento Agua ( $mm/dia$ )
2016	-0.141	21.446	-5.399
2017	0.014	20.792	-7.557
2018	0.024	20.841	-4.284

**Cuadro 6.1.1:** Datos anuales para la variación de velocidad, temperatura y almacenamiento de agua para los años 2016, 2017 y 2018.

## Capítulo 7

### Conclusión

Se ha observado a partir del método de interferometría sísmica pasiva un aumento de  $dv/v$  a lo largo de los años dentro de la zona estudiada, que se encuentra comprendida entre las estaciones y sus alrededores. Hemos comprobado que la variación relativa de velocidad de ondas sísmicas está estrechamente relacionada a las variaciones temperatura en lo que a corto plazo se refiere. Esto se debe a la rápida propagación que tiene esta en el suelo y su efecto directo en la presión termo-elásticas.

Por otro lado, si bien evidentemente la variación de agua subterránea tiene importancia en el proceso, parece ser mucho menor al menos a corto plazo. Además, cabe recalcar que las variaciones de temperaturas experimentadas entre cada par de estaciones son muy similares, y por tanto las discrepancias entre distintos pares debería deberse principalmente a las diferencias en condiciones hidrológicas, ya sea por la capacidad del suelo de almacenar agua o en diferentes condiciones meteorológicas de cada zona.

Lamentablemente, no contamos con una gran historia de datos, puesto que la red de estaciones C1 es bastante reciente, sin embargo si esta metodología se aplica de aquí al futuro, este estudio sirve como línea de base para comparar resultados futuros y ver tendencias a largo plazo. En el caso de la red C tiene las estaciones muy alejadas entre ellas (exceptuando la zona de Santiago), por lo que no es apta para el método de interferometría sísmica pasiva. Además, sería bueno ampliar la red de datos para estudiar de mejor manera estas variaciones a lo largo del territorio chileno, o por otro lado, concentrar estaciones sismológicas con menor

separación entre ellas para enfocar una zona y estudiarla en mayor detalle.

La interferometría sísmica pasiva es una herramienta que puede ser muy útil a largo plazo, pero requiere una gran cantidad de datos para probarse efectiva. Incluso podría ser posible modelar como los cambios de temperaturas afectan  $dv/v$  en una corteza que posea una estructura similar a la de Chile, para así remover esta tendencia y observar solo como se ve afectada por efectos secundarios como el balance de aguas.

## Bibliografía

- Aguayo, R., León Muñoz, J., Vargas Baecheler, J., Montesinos, A., Garreaud, R., Urbina, M., Soto, D., and Iriarte, J. L. (2019). The glass half-empty: climate change drives lower freshwater input in the coastal system of the chilean northern patagonia.
- Aki, K. and Richards, P. G. (1976). *Quantitative Seismology*.
- Ben-Zion, Y. and Allam, A. A. (2013). Seasonal thermoelastic strain and postseismic effects in parkfield borehole dilatometers.
- Bensen, G. D., Ritzwoller, M. H. and Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements.
- Campillo, M. (2006). Phase and correlation in random seismic fields and reconstruction of the green function.
- Catálogo CAMELS-CL (2022). <https://camels.cr2.cl>.
- Efron, B. and Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation.
- Garreaud, R. D., Boisir, J. P., Rondanelli, R., Montecinos, A., Sepulveda, H. H., and Veloso Aguilar, D. (2019). The central chile mega drought (2010-2018): A climate dynamics perspective.
- Hillers, G., Retailleau, L., Campillo, M., Inbal, A., Ampuero, J.-P., and Nishimura, T. (2014). In situ observations of velocity changes in response to tidal deformation from analysis of the high-frequency ambient wavefield.
- Lecocq, T., Longuevergne, L., Pedersen, H. A., Brenguier, F., and Stammer, K. (2017). Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling.
- Schober, P., Boer, C., and Schwarte, L. (2018). Correlation coefficients: Appropriate use and interpretation.
- Sens-Schönfelder, C. and Wegler, U. (2006). Passive image interferometry and seasonal variations of seismic velocities at merapi volcano, indonesia.

- Snieder, R. (2006). The theory of coda wave interferometry.
- Sánchez Sesma, F. J., Carbajal Romero, M., Rodríguez Sánchez, J. E., Ávila Carrera, R., and Rodríguez Castellanos, A. (2017). Recuperación de función de green mediante ruido sísmico.
- Turcotte, D. L. and Schubert, G. (2014). *Geodynamics*.
- Wapenaar, K., Slob, E., Snieder, R., and Curtis, A. (2010). Tutorial on seismic interferometry: Part 2 - underlying theory and new advances.
- Wegler, U., Nakahara, H., Sens-Schönfelder, C., Korn, M., and Shiome, K. (2009). Sudden drop of seismic velocity after the 2005 mw 6.6 mid-niigata earthquake, japan, observed with passive image interferometry.

# Apéndice A

## Glosario

### A1. Operadores vectoriales

$$\vec{u}(x_1, x_2, x_3) = u_i \quad (\text{A1.1})$$

$$u_{i,j} = \frac{\partial u_i}{\partial x_j} \quad (\text{A1.2})$$

$$a_i b_j = \vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \quad (\text{A1.3})$$

$$\vec{a} \cdot \vec{b} = \vec{a}^T \vec{b} \quad (\text{A1.4})$$

$$\vec{a} \otimes \vec{b} = \vec{a} \vec{b}^T \quad (\text{A1.5})$$

$$\nabla \phi = \partial_i \hat{x}_i \phi = \left[ \frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}, \frac{\partial \phi}{\partial x_3} \right]^T \quad (\text{A1.6})$$

$$\nabla \cdot \vec{u} = \partial_i u_i = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \quad (\text{A1.7})$$

$$\nabla \times \vec{u} = \epsilon_{ijk} \partial_{jk} u_k \quad (\text{A1.8})$$

$$\nabla^2 \phi = \nabla \cdot (\nabla \phi) \quad (\text{A1.9})$$

$$\nabla \vec{u} = (\nabla u_{x_1}, \nabla u_{x_2}, \nabla u_{x_3}) \quad (\text{A1.10})$$

## A2. Tensores especiales

### A2.1. Delta de Kronecker

$$\delta_{ij} : \begin{cases} \delta_{ij} = 1, & i = j \\ \delta_{ij} = 0, & i \neq j \end{cases} \quad (\text{A2.1})$$

### A2.2. Tensor de Levi Civita

$$\epsilon_{ijk} : \begin{cases} \epsilon_{ijk} = 0, & i = j \vee i = k \vee j = k \\ \epsilon_{ijk} = +1, & ijk \text{ en orden} \\ \epsilon_{ijk} = -1, & ijk \text{ alternados} \end{cases} \quad (\text{A2.2})$$

### A2.3. Delta de Dirac

$$\delta(t - \tau) : \begin{cases} \delta(t - \tau) = \infty, & t = \tau \\ \delta(t - \tau) = 0, & t \neq \tau \end{cases} \quad (\text{A2.3})$$

## A3. Propiedades tensoriales

$$\delta_{ij}\phi_j = \phi_i \quad (\text{A3.1})$$

$$\epsilon_{ijk}a_jb_k = \vec{a} \times \vec{b} \quad (\text{A3.2})$$

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{jl}\delta_{im} \quad (\text{A3.3})$$

$$\epsilon_{ijk}U_{jk} = 0 \Rightarrow U_{jk} \text{ simetrico} \quad (\text{A3.4})$$

$$\sum_j A_j \delta_{ij} = A_i \quad (\text{A3.5})$$

$$A_{ij} = A_{ji} \Rightarrow \text{simetría} \quad (\text{A3.6})$$

$$A_{ij} = -A_{ji} \Rightarrow \text{asimetría} \quad (\text{A3.7})$$

$$\int_{-\infty}^{+\infty} \vec{f}(\vec{x})\delta(\vec{x} - \vec{a}) = \vec{f}(\vec{a}) \quad (\text{A3.8})$$

# Apéndice B

## Tablas

### A1. Información de estaciones sismológicas

Nombre de la estación	Modelo Sensor	Respuesta del instrumento	Frecuencia de Muestreo(Hz)
BO01	Trillium 120P	120s a 50Hz	100
BO02	Trillium 120P	120s a 50Hz	100
BO03	Trillium 120P	120s a 50Hz	100
BO04	Trillium 120P	120s a 50Hz	100
MT01	Trillium 120P	120s a 50Hz	100
VA05	Trillium 120P	120s a 50Hz	100

**Cuadro A1.1:** Información de las 6 estaciones pertenecientes a la red C1 que fueron utilizadas en el estudio.

### A2. Polos y Zeros Trillium 120P

ZEROS	4
0	0
0	0
-106	0
-158	0
POLES	7
-0.03859	0.03649
-0.03859	-0.03649
-190	0
-158	193
-158	-193
-639	1418
-639	-1418
CONSTANT	1.201e+09

**Cuadro A2.1:** Polos y ceros para la componente vertical de sensor Trillium 120P.

# Apéndice C

## Código de Python

### A1. Correlación cruzada entre estaciones

```
from obspy.core import read, UTCDateTime
import os
from obspy.signal.cross_correlation import xcorr_3c as xcorr
import matplotlib.pyplot as plt
import numpy as np
from numpy.fft import fft, ifft, fftfreq
from miic.core.normalization import temp_norm

def smooth(y, box_pts):
    box = np.ones(box_pts)/box_pts
    #box = np.hanning(box_pts)/box_pts
    y_smooth = np.convolve(y, box, mode='same')
    return y_smooth

def doPreprocessing(st_segment,minfreq,maxfreq):

    #####
    #demean
    st_segment.detrend('demean')
    #detrend
    st_segment.detrend('linear')
```

```
#taper
st_segment.taper(0.1)
#merge
st_segment.merge(method=1, fill_value=0, interpolation_samples=0)
#filter
st_segment.filter('bandpass',freqmin=minfreq,freqmax=maxfreq)
#Por default es 0.4 a 6

return st_segment

def spectralWhite(st_segment):
#####
# compute waveform' FFT for each station
for tr in st_segment:
    raw_data = tr.copy().data

    windowLengthNormalization = 0.4 # in Hz
    Nfft = len(raw_data)
    freqs = fftfreq(Nfft, d=1/tr.stats.sampling_rate)
    freqs = freqs[range(int(Nfft/2))]
    yfft = fft(raw_data, n=Nfft)
    yfft = yfft / Nfft*2
    spec_ampl = np.sqrt(abs(np.multiply(yfft,
        np.conjugate(yfft))))
    spec_ampl = smooth(spec_ampl, box_pts=
        int(tr.stats.sampling_rate*windowLengthNormalization))
    white_fft = yfft / spec_ampl
    wdata = ifft(white_fft)
    wdata = np.multiply(wdata, np.conjugate(wdata))
    wdata = wdata.real*np.sign(raw_data)
    tr.data = wdata

    windowLengthNormalization = 2.5 # in seconds
    tempnormdata = temp_norm(tr.data,
        int(windowLengthNormalization*tr.stats.sampling_rate))
```

```
#tempnorm_fft = fft(tempnormdata, Nfft)
tr.data = tempnormdata

return st_segment

"""
Parametros iniciales
"""
red1='C1' #Red de estaciones para la primera estacion
estacion1='B004' #Nombre de la primera estacion
red2='C1' #Red de estaciones de la segunda estacion
estacion2='B002' #Nombre de la segunda estacion
banda='BH' #Banda de frecuencia
stday=UTCDateTime(2017,2,1,0,0,0) #Fecha inicial
ndays=28 #Cantidad de dias desde la fecha inicial
minfreq=0.4 #Frecuencia minima del filtro pasabanda
maxfreq=2 #Frecuencia maxima del filtro pasabanda
directory='./Test042-' #Prefijo de la carpeta en el que se
    guardaran los datos
time_lag=80 #Intervalos de tiempo para la CC

"""
Carga y procesamiento de datos
"""

days=np.arange(0,ndays,1).tolist() #se genera una lista con la
    cantidad de dias
all_xcorr_mean=np.empty((6401,ndays)) #generamos la matriz de CC
    vacia
serietiempo=np.empty((1,ndays)) #generamos un vector vacío para
    los días
j=0 #contador de dias
b=0 #contador de dias con datos validos

for j in days:
```

```
#El nombre de los archivos a cargar
stdate=stday+86400*j
archivo1=''.join([str(elem) for elem in ['./',red1,
    '-',estacion1,'/',estacion1,banda,stdate,'.sac']])
archivo2=''.join([str(elem) for elem in ['./',red2,
    '-',estacion2,'/',estacion2,banda,stdate,'.sac']])

#Comprobamos si estos archivos existen o no
logic1=os.path.exists(archivo1)
logic2=os.path.exists(archivo2)
if logic1 == False:
    print([stdate.year,stdate.julday,estacion1,'omega bruh'])
elif logic2 == False:
    print([stdate.year,stdate.julday,estacion2,'omega bruh'])
else:

    #en caso de que existan los cargamos
    ST1 = read(archivo1)
    ST2 = read(archivo2)

    #comprobamos si es que tiene la cantidad suficiente de
    datos
    aaa=ST1[0].stats.npts
    bbb=ST2[0].stats.npts
    ccc=2304000
    if aaa < ccc:
        print([stdate.year,stdate.julday,estacion1,'bruh',aaa])
    elif bbb < ccc:
        print([stdate.year,stdate.julday,estacion2,'bruh',bbb])
    else:
        i=0 #Contador de segmentos
        a=0 #Contador de segmentos sin nan
        #Vemos cuantos tiempo de datos hay y por tanto en cuantos
        intervalos hay que dividir la serie
```

```
if aaa < bbb:
    missmeans=((3456001-aaa)/36000)//1
else:
    missmeans=((3456001-bbb)/36000)//1
missminutes=missmeans*15
nominutes=23*60+1-missminutes
nomeans=int(93-missmeans)
hours=np.arange(0,nominutes,15).tolist()

#Extraemos la informacion de la serie
stt=ST1[0].stats.starttime
sampling_rate=ST1[0].stats.sampling_rate
shift_len=int(sampling_rate*time_lag)
t=np.arange(-time_lag,time_lag+time_lag/(sampling_rate
    *time_lag),time_lag/(sampling_rate*time_lag))
#Matriz vacia donde iran las correlaciones de horas
all_xcorr=np.empty((len(t),nomeans))

#Realizamos CC para intervalos de 1 hora que avanzan
    por 15 mins
for i in hours:
    #Definimos los limites del intervalo
    t1=stt+60*i
    t2=t1+3600
    #Creamos copias de las series de datos diarias
    tr1=ST1.copy()
    tr2=ST2.copy()
    #Cortamos las series
    tr1.trim(t1,t2)
    tr2.trim(t1,t2)
    #Realizamos pre-procesamiento
    doPreprocessing(tr1,minfreq,maxfreq)
    doPreprocessing(tr2,minfreq,maxfreq)
    spectralWhite(tr1)
    spectralWhite(tr2)
```

```
#Realizamos la CC

if len(tr1[0].data)==len(tr2[0].data):
    index,value,fct=xcorr(tr1,tr2,shift_len,
        components="Z",full_xcorr=True)
    #print(max(np.abs(fct)))
    #Verificamos si el segmento tiene nan
    istherenan=np.isnan(np.sum(fct))
    if istherenan==False:
        all_xcorr[:,a]=fct
        #print(max(np.abs(all_xcorr[:,a])))
        a=a+1
#Comprobamos si luego de descartar los nan sigue teniendo
    datos suficientes
ddd=(a+3)*36000
if ddd < ccc:
    print([stdate.year,stdate.julday,estacion1,
        'bruh',ddd])
else:
    #Promediamos la CC de cada segmento para obtener
        una CC diaria
    all_xcorr_mean[:,b]=np.mean(all_xcorr[:,0:a-1],
        axis=1)
    serietiempo[0,b]=stdate.julday-1
    b=b+1
    print([stdate.year,stdate.julday,'N I C E'])

#Se corta lo que quede vacio en caso de dias sin datos suficientes
all_xcorr_final=all_xcorr_mean[:,0:b]
serietiemnocut=np.empty((b+1))
serietiemnocut[0:b]=serietiempo[0,0:b]
serietiemnocut[b]=stdate.julday

#Ploteamos la matriz de CC
plt.figure(b+2)
```

```
plt.pcolormesh(serietiemnocut[0:len(serietiemnocut)-1],t,  
              all_xcorr_final)
```

## A2. Stretch factor y calculo de $dv/v$

```
from obspy.core import UTCDateTime  
import os  
import matplotlib.pyplot as plt  
import numpy as np  
from miic.core.stretch_mod import time_stretch_estimate  
from matplotlib.colors import BoundaryNorm  
from matplotlib.ticker import MaxNLocator  
from sklearn.linear_model import LinearRegression  
  
"""  
Funciones  
"""  
def moving_average(maxcut,serie,w): #Media movil que funciona a  
    través de NaN  
    maximo=int(len(serie)-(w+1))  
    mediamovil=np.empty(maximo)  
    m=0  
    while m < maximo:  
        n=0  
        suma=0  
        numero=0  
        while n < w:  
            istherenan=np.isnan(maxcut[m+n])  
            if istherenan==True:  
                n=n+1  
            else:  
                suma=suma+maxcut[m+n]  
                numero=numero+1  
                n=n+1  
        prom=suma/numero
```

```
        mediamovil[m]=prom
        m=m+1
    return mediamovil

def nan_filler(matcut,serie): #Rellena los dias sin datos con NaN
    (para metricas 2D)
    seriearange=np.arange(min(serie),max(serie),1)
    m=0
    n=0
    matcutarange=np.empty((len(matcut[:,0]),len(seriearange)-1))
    while m < len(seriearange)-1:
        if serie[n]==seriearange[m]:
            matcutarange[:,m]=matcut[:,n]
            m=m+1
            n=n+1
        else:
            matcutarange[:,m]=np.nan
            m=m+1
    return matcutarange,seriearange

def vect_filler(vectcut,serie): #Rellena los dias sin datos con
    NaN (para vectores)
    seriearange=np.arange(min(serie),max(serie),1)
    m=0
    n=0
    vectcutarange=np.empty((len(seriearange)-1))
    while m < len(seriearange)-1:
        if serie[n]==seriearange[m]:
            vectcutarange[m]=vectcut[n]
            m=m+1
            n=n+1
        else:
            vectcutarange[m]=np.nan
            m=m+1
    return vectcutarange,seriearange
```

```

def nanremove(serietiempo,seriedatos):
    a=0
    b=0
    serietiemlocopy=serietiempo
    seriedatoscopy=seriedatos
    while a < len(serietiempo):
        isthisnan=np.isnan(seriedatos[a])
        if isthisnan == False:
            serietiemlocopy[b]=serietiempo[a]
            seriedatoscopy[b]=seriedatos[a]
            b=b+1
        a=a+1
    serietiemlocopyfinal=serietiemlocopy[0:b-1]
    seriedatoscopyfinal=seriedatoscopy[0:b-1]
    return serietiemlocopyfinal, seriedatoscopyfinal

"""
Carga de datos
"""
#Entregar información de las estaciones que hay que cargar
red1='C1'
estacion1='VA05'
red2='C1'
estacion2='B001'
prefijo='Test042' #Utilicé distintos prefijos para separar pruebas
                con distintos filtros
#Ventana de la coda. Recomendado dejar el minimo en 0 (peak mas
alto de la onda)
minlimcoda=0
maxlimcoda=10
#Fechas que comprenden la serie de datos
startdate=UTCDateTime(2016,1,1,0,0,0)
enddate=UTCDateTime(2019,12,31,0,0,0)
#Generando las direcciones para cargar los archivos

```

```

archivoload='' .join([str(elem) for elem in ['./',prefijo,'-',red1,
    estacion1,'-',red2,estacion2,'/daily_xcorr',stddate,'_',eddate,
    '.txt']])
serieload='' .join([str(elem) for elem in ['./',prefijo,'-',red1,
    estacion1,'-',red2,estacion2,'/serietiempo',stddate,'_',eddate,
    '.txt']])
lagtimeload='' .join([str(elem) for elem in ['./',prefijo,'-',red1,
    estacion1,'-',red2,estacion2,'/lagtime.txt']])
#Carfando archivos de datos
archivo=np.loadtxt(archivoload)
serie=np.loadtxt(serieload)
lagtime=np.loadtxt(lagtimeload)

"""
Promediar
"""
#Promedio de la parte positiva y negativa de la matriz de CC
dunno=int(((len(archivo[:,0])-1)/2)+1)
archivoprom=np.empty((dunno,len(archivo[0,:])))
archivoprom[0,:]=archivo[dunno-1,:]
i=1
while i<dunno:
    archivoprom[i,:]=(archivo[dunno-1+i,:]+archivo[dunno-1-i,:])/2
    i=i+1
#Rellenar los dias faltantes con NaN, esto con el objetivo de no
    ver datos estirados
[archivofill,seriefill]=nan_filler(archivoprom,serie)
#Dejar solo el lado positivo del time lag para encajar medidas
lagtimeprom=lagtime[dunno-1:len(lagtime)]

"""
Plotting
"""
plt.figure(1)
#Ploteo de la matriz de CC + barra de color

```

```

levels = MaxNLocator(nbins=15).tick_values(archivoprom.min(),
      archivoprom.max())
cmap = plt.get_cmap('viridis')
norm = BoundaryNorm(levels, ncolors=cmap.N, clip=True)
im = plt.pcolormesh(seriefill[0:len(seriefill)-1],lagtimeprom,
      archivofill, cmap=cmap, norm=norm)
plt.colorbar(im)
#Generar titulo
CCtitle=''.join([str(elem) for elem in ['Matriz de correlación
      entre ',estacion1,' y ',estacion2]])
plt.title(CCtitle)
#Nombrar ejes
plt.xlabel("Días")
plt.ylabel("Time lag [s]")
#plt.ylim([0,40])
#Comrpobamos si existe el directorio, si no existe es creado
savedirectory=''.join([str(elem) for elem in ['./Illust']])
logic=os.path.exists(savedirectory)
if logic==False:
    os.mkdir(''.join([str(elem) for elem in ['./Illust']]))

"""
Stretching
"""
dunno2=int(len(archivo[0,:]))
j=0
corr=np.empty(dunno2)
mat=np.empty((100,dunno2))
mat_max=np.empty(dunno2)
k=0
archivoprom2=np.empty(dunno)
while k < dunno:
    archivoprom2[k]=np.mean(archivoprom[k,:])
    k=k+1
maxindex=np.argmax(archivoprom2)

```

```

maxindexp=maxindex+40*maxlimcoda
maxindexm=maxindex+40*minlimcoda
#Corremos la funcion de stretching dia por dia
while j<dunno2:
    stretch_dict = time_stretch_estimate(archivoprom[maxindexm:
        maxindexp,j],archivoprom2[maxindexm:maxindexp]) #Funcion de stretch
    #Correlacion maxima del día:
    a=stretch_dict['corr']
    corr[j] = a
    #Vector de stretch de -0.1 a 0.1:
    stretch_factor=stretch_dict['second_axis']
    #Matriz de correlación:
    a=stretch_dict['sim_mat']
    mat[:,j]=a[:]
    #Valor de stretch correspondiente a la correlacion maxima del
        día:
    a=stretch_dict['value']
    mat_max[j] = a
    j=j+1

"""
Precesamiento extra
"""
l=0
#Cortar la matriz de stretching ente limites designados
posimin25a=np.where(stretch_factor > -0.1+np.nanmean(mat_max))
posimin25b=posimin25a[0]
posimin25c=posimin25b[0]
posimax25a=np.where(stretch_factor < 0.1+np.nanmean(mat_max))
posimax25b=posimax25a[0]
posimax25c=posimax25b[len(posimax25b)-1]
matcut=mat[posimin25c:posimax25c,:]
stretch_factorb=stretch_factor-np.nanmean(mat_max)
stretchcut=stretch_factorb[posimin25c:posimax25c]
maxcut=mat_max-np.nanmean(mat_max)

```

```

#Cambiar por nan los que se alejen mucho del centro o que tengan
    correlación maxima menor a lo buscado
while l < len(serie)-1:
    #Filtrar por correlación maxima:
    if corr[l]<0.6:
        matcut[:,l]=np.nan
        maxcut[l]=np.nan
    #Filtrar por distancia de stretch respecto a 0:
    elif mat_max[l] > stretchcut[len(stretchcut)-1]:
        matcut[:,l]=np.nan
        maxcut[l]=np.nan
    elif mat_max[l] < stretchcut[0]:
        matcut[:,l]=np.nan
        maxcut[l]=np.nan
    l=l+1
#Pasamos los datos de dt/t a dv/v en %
maxcut=maxcut*(-100)
stretchcut=stretchcut*(-100)
#Rellenar con NaN los días sin datos la matriz de stretching
[matcutfill,seriefill2]=nan_filler(matcut,serie)
#Calcular media movil
w=60
mediamovil=moving_average(maxcut,serie,w)
#Rellenar con NaN los dias sin datos de la media movil
[mediamovilfill,seriecutfill]=vect_filler(mediamovil,serie[0:
    len(serie)-(w)])
#Regresion lineal
seriereg,datosreg=nanremove(serie[0:len(mat_max)]/365.25, maxcut)
X = seriereg[:, np.newaxis]
model = LinearRegression().fit(X, datosreg)
y_pred = model.predict(seriefill2[:, np.newaxis]/365.25)
print('coef=',model.coef_)

"""

```

Plotting

```
"""
```

```
#Ploteo de la matriz de stretch junto a sus peaks y una media movil
```

```
starterday=stddate.year #Fecha inicial
```

```
plt.figure(2)
```

```
#Matriz de stretching:
```

```
#plt.pcolormesh(starterday+seriefill2[0:len(seriefill2)-1]/365.25,  
                stretchcut,matcutfill)
```

```
#Maximo diario:
```

```
#plt.plot(starterday+serie[0:len(mat_max)]/365.25,maxcut,'.r')
```

```
#Media movil
```

```
plt.plot(starterday+seriecutfill[0:len(seriecutfill)-1]/365.25,  
         mediamovilfill,'r',linewidth=2)
```

```
plt.plot(starterday+seriefill2/365.25,y_pred,'k',linewidth=2)
```

```
#plt.ylim([-8,8])
```

```
plt.xlim([2016,2020])
```

```
#Etiquetas para cada eje
```

```
plt.xlabel("Días")
```

```
plt.ylabel("Factor de estrechamiento %")
```

```
CCtitle3=''.join([str(elem) for elem in ['Variacion de velocidad  
( $dv/v$ ) entre ',estacion1,' y ',estacion2]])
```

```
plt.title(CCtitle3)
```