

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE CIENCIAS**

### **PROCESOS ESTOCÁSTICOS ARMA PARA EL ESTUDIO DE SERIES TEMPORALES SÍSMICAS DEL ECUADOR.**

**TRABAJO PREVIO A LA OBTENCIÓN DEL TÍTULO DE FÍSICO**

#### **PROYECTO DE INVESTIGACIÓN**

**WILLIAM JOHEL CASTILLO ARMIJO**

`william.castillo@epn.edu.ec`

**DIRECTOR: RAMÓN XULVI-BRUNET, PhD.**

`ramon.xulvi@epn.edu.ec`

**Quito, febrero 2021**

## **DECLARACIÓN**

Yo, WILLIAM JOHEL CASTILLO ARMIJO, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

**William Johel Castillo Armijo**

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por WILLIAM JOHEL CASTILLO ARMIJO, bajo mi supervisión.

---

**Ramón Xulvi-Brunet, PhD.**

**Director del Proyecto**

## AGRADECIMIENTOS

Es muy corto este espacio y mi dominio del idioma es muy limitado para poder plasmar cuan agradecido estoy con las personas nombradas a continuación:

A mi familia, que son un pilar fundamental de mi vida y quienes, a pesar de no estar en total acuerdo con mis decisiones siempre me han brindado su amor y apoyo incondicional.

Al Dr. Ramon Xulvi-Brunet, mi tutor y amigo, que con su infinita paciencia a dirigiendo este trabajo y me ha enseñado, de la manera más sincera posible, a ser un mejor profesional y ser humano.

A todos los profesores de los departamentos de física y matemática, quienes desinteresadamente compartieron sus conocimientos conmigo. En especial, a los doctores Esteban Iribarra, Christian Vásquez, Álvaro Dueñas y Luis Lascano quienes me han ayudado más allá de lo que sus obligaciones establecen.

A mis amigos, dentro y fuera de la universidad, que siempre me han ayudado y me alegran la vida.

A todos los ecuatorianos, quienes con su trabajo hacen posible la educación pública de calidad.

Los sismogramas, datos, metadatos y/o productos derivados utilizados en este estudio fueron proporcionados por el "Instituto Geofísico de la Escuela Politécnica Nacional" (IG-EPN).

El software "Seismic Analysis Code" (SAC) fue proporcionado por las Instituciones de Investigación Incorporadas para Sismología (IRIS). Las instalaciones del Consorcio IRIS cuentan con el apoyo del premio "Seismological Facilities for the Advancement of Geoscience" (SAGE) de la Fundación Nacional de Ciencias en virtud del Acuerdo de Apoyo Cooperativo EAR-1851048.

Los cálculos numéricos, con los que se obtuvieron los resultados de este estudio, se ejecutaron en el supercomputador del Laboratorio Nacional de Cálculo Científico, gracias al beneplácito del "Centro de Modelización Matemática MODEMAT".

## DEDICATORIA

*A tu ñaña.*

# CONTENIDO

<b>1</b>	<b>INTRODUCCIÓN</b>	<b>3</b>
1.1	Sismología . . . . .	3
1.2	Marco Sismológico del Ecuador . . . . .	5
1.3	Ondas Sísmicas y su Registro . . . . .	6
1.4	Análisis de Series Temporales . . . . .	9
1.4.1	Procesos Estocásticos . . . . .	9
1.4.2	Medidas Características . . . . .	9
1.5	Series Temporales . . . . .	11
1.6	Motivación y Objetivo de este Trabajo . . . . .	12
1.7	Outline . . . . .	13
<b>2</b>	<b>MARCO TEÓRICO</b>	<b>14</b>
2.1	Proceso ARMA . . . . .	14
2.1.1	Proceso Auto-Regresivo . . . . .	14
2.1.2	Proceso de Media Móvil . . . . .	20
2.1.3	Proceso Auto-Regresivo de Media Móvil (ARMA) . . . . .	21
2.2	Estimación de Máxima Verosimilitud . . . . .	23
2.2.1	Definición . . . . .	23
2.2.2	Función Likelihood para un Proceso ARMA(p,q) Gaussiano . . . . .	24
2.3	Algoritmo Davidon-Fletcher-Powell . . . . .	26
2.4	Métodos de Verificación . . . . .	29
2.4.1	Transformada de Fourier . . . . .	29
2.4.2	Criterio de Información de Akaike . . . . .	30
2.4.3	Coeficiente de Eficiencia del Modelo Nash-Sutcliffe . . . . .	31
2.4.4	Índice de Acuerdo de Willmott . . . . .	32
<b>3</b>	<b>METODOLOGÍA</b>	<b>33</b>
3.1	Tratamiento de los Datos . . . . .	33
3.2	Ajuste de los Parámetros del Proceso ARMA . . . . .	35
3.3	Construcción del Proceso ARMA y Pronósticos de Sismogramas . . . . .	39
<b>4</b>	<b>RESULTADOS Y DISCUSIÓN</b>	<b>42</b>

4.1	Validación del Modelo ARMA . . . . .	42
<b>5</b>	<b>CONCLUSIONES</b>	<b>50</b>
<b>6</b>	<b>REFERENCIAS BIBLIOGRÁFICAS</b>	<b>51</b>
<b>7</b>	<b>ANEXOS</b>	<b>I</b>
7.1	Apéndice: Cálculo de las derivadas parciales de residuos respecto a parámetros de media Móvil . . . . .	I
7.2	Códigos Fuente . . . . .	IV

## ÍNDICE DE FIGURAS

1.1	Proceso de convección en el cual el magma caliente sube a la superficie, se extiende y comienza a enfriarse. Luego se hunde de nuevo al fondo donde se recalienta y se eleva nuevamente. Dorsal: Lugar donde el Magma emerge y empuja a las placas; Fosa: Extremo de la placa que se hunde [1]. . . . .	3
1.2	Almacenamiento de energía elástica a medida que las placas tectónicas se deforman. Cuando la tensión revasa la fuerza de rozamiento, la falla se desliza liberando energía en forma de ondas sísmicas [4]. . . . .	4
1.3	Mapa tectónico activo del Ecuador. Los segmentos de fallas mayores se ilustran en líneas rojas continuas. Los restos de las zonas de subducción se dibujan en líneas de puntos negras [18]. . . . .	6
1.4	Un terremoto consta de una combinación de varios tipos de ondas sísmicas y sus movimientos característicos [26]. . . . .	7
1.5	Sismógrafo sensible a movimientos verticales. Puede ser visualizado como una pesa suspendida de un resorte que a su vez están suspendidos sobre una base que se mueve con los movimientos de la superficie de la Tierra. Para añadir un sistema de registro se coloca un tambor que gira en la base y un marcador sujetado a la masa [30]. . . . .	8
2.1	Función a optimizar de una sola variable real en la cual, el algoritmo empieza en un punto inicial arbitrario y después de tres iteraciones se localiza lo suficientemente cerca al valor que maximiza dicha función. . . . .	26
3.1	Series temporales sísmicas originales de 10,798,990 datos (125 datos por segundo) cada una. También, una línea horizontal en el instante 43,195.96 (5,399,495 datos) que divide estas series en dos partes iguales. La primera se utiliza para ajustar los parámetros de proceso <i>ARMA</i> y la segunda para verificar los resultados del mismo. Las series se las trata adimensionalmente ya que las unidades son irrelevantes para este análisis. . . . .	36
4.1	Dependencia del AIC a los parámetros $p$ y $q$ (toman valores enteros entre 1 y 7) para las cuatro series estudiadas. El punto de color rojo representa el AIC de menor valor y la superficie de color morado es una extrapolación gaussiana de los datos que se utiliza para visualizar de mejor manera la localización de los puntos en el espacio representado en cada figura. . . . .	43

4.2	Comparación de la segunda mitad de las cuatro series temporales estudiadas con el respectivo proceso <i>ARMA</i> que pretende describir dichas series. . . . .	44
4.3	Ampliación de aproximadamente 5 segundos de las series de la gráfica (4.2). La figura superior registra ruido mientras que las tres últimas figuras registran un sismo. 45	45
4.4	Sismogramas y pronósticos del proceso <i>ARMA</i> en el espacio de frecuencia con unidades de [ $x10^3Hz$ ]. . . . .	46
4.5	Comparación entre la distribución de residuos con su supuesta distribución de teórica. La distribución teórica se supone gaussiana según el planteamiento tradicional del Proceso <i>ARMA</i> para ajustar los parámetros que caracterizan este proceso. .	48

## ÍNDICE DE TABLAS

3.1	Valores iniciales del vector de parámetros $\Theta$ en notación de doble precisión en Fortran 90. . . . .	40
4.1	Media y desviación estándar de la serie producto del Procesos <i>ARMA</i> , la parte de ajuste y la parte de verificación de los cuatro sismogramas estudiados. . . . .	45
4.2	Índices que establecen la bondad del modelo <i>ARMA</i> en comparación con los datos empíricos. . . . .	47

## ÍNDICE DE CÓDIGOS FUENTE

7.1	Encontrar el mínimo de una función. . . . .	IV
7.2	Variación de la longitud del paso entre iteraciones. . . . .	VI
7.3	Cálculo de la función Log Likelihood. . . . .	VIII
7.4	Construcción del vector error. . . . .	IX
7.5	Gradiente de la función Log Likelihood. . . . .	IX
7.6	Script que administrar los recursos del HPC-MODEMAT. . . . .	XI
7.7	Números aleatorios independientes con distribución gaussiana. . . . .	XII
7.8	Construcción de un modelo en base a una serie temporal. . . . .	XII

## RESUMEN

Un sismo es el movimiento de la superficie terrestre producto de la súbita liberación de energía que se propaga en forma de ondas. Este fenómeno natural guarda una estrecha relación con los miembros de una sociedad y su desarrollo en conjunto. Dependiendo de la cantidad de energía liberada durante un sismo, las secuelas sociales y económicas pueden tomar años e incluso décadas para ser superadas. Si bien comprendemos bastante sobre este tema, no existe un modelo que describa satisfactoriamente el mecanismo subyacente que produce sismos. Hoy en día, es imposible conocer exactamente cuándo o dónde van a ocurrir. Razón por la cual, el conocimiento científico de este fenómeno natural es un componente esencial para la toma de decisiones con respecto a la inversión público/privada, la planificación del desarrollo y la seguridad nacional.

En este trabajo de titulación pretendemos indagar en la naturaleza de la sismicidad empleando los registros sísmicos del Instituto Geofísico de la Escuela Politécnica Nacional. Para este fin, desarrollamos un algoritmo que emplea la teoría del proceso ARMA (de su siglas en inglés *Auto Regressive Moving Average*) para recrear e incluso pronosticar la información que guardan estos registros. Los resultados muestran como la complejidad de este fenómeno natural rebasan las suposiciones que hace el modelo empleado.

## ABSTRACT

An earthquake is the movement of the Earth's surface due to the sudden release of energy that spreads in form of waves. This natural phenomenon is closely related to the society's members and their development. Depending on the amount of energy released during an earthquake, social and economic consequences can take years and even decades to be overcome. While we understand quite about this topic, there no exist any model that describes the underlying mechanism that produces earthquakes. Today, it is impossible to know exactly when or where they will occur. Thus, scientific knowledge of this natural phenomenon is an essential component for decision-making regarding public / private investment, development and national security.

In this project, we will investigate the nature of earthquakes using the seismic records of "Instituto Geofísico de la Escuela Politécnica Nacional". To do so, we developed an algorithm that uses the *Auto Regressive Moving Average* (ARMA) process theory to recreate and even forecast the information stored in these records. The results show how the complexity of this natural phenomenon exceeds the assumptions made by the model.

# 1 INTRODUCCIÓN

En este capítulo se discuten conceptos relevantes al estudio y pronóstico de sismos. Se revisa algunos conceptos básicos de sismología en especial el registro y análisis de ondas sísmicas. También, el entorno sismológico del país y sus implicaciones sociales que motivan este trabajo. Y finalmente, se establece el objetivo que también es motivado por razones pedagógicas.

## 1.1 SISMOLOGÍA

La **sismología** (del griego, *seismós*, que significa 'sismo', y *logía*, que significa ' estudio de') es la rama de la geofísica que se encarga del estudio de terremotos y la propagación de ondas sísmicas a través de la Tierra y otros cuerpos planetarios. Asimismo, este campo también se ocupa de la gestión de los riesgos y peligros sísmicos [2, 3].

Los **sismos** (también llamados seísmos, temblores, movimientos telúricos y, cuando su duración es relativamente larga o causa graves daños, terremotos) se caracterizan principalmente por su intensidad y magnitud. La intensidad es la descripción cualitativa de los efectos de los sismos (considera la percepción de las personas, daños materiales y económicos),

mientras que la magnitud es la descripción objetiva de los mismos (estima la energía liberada

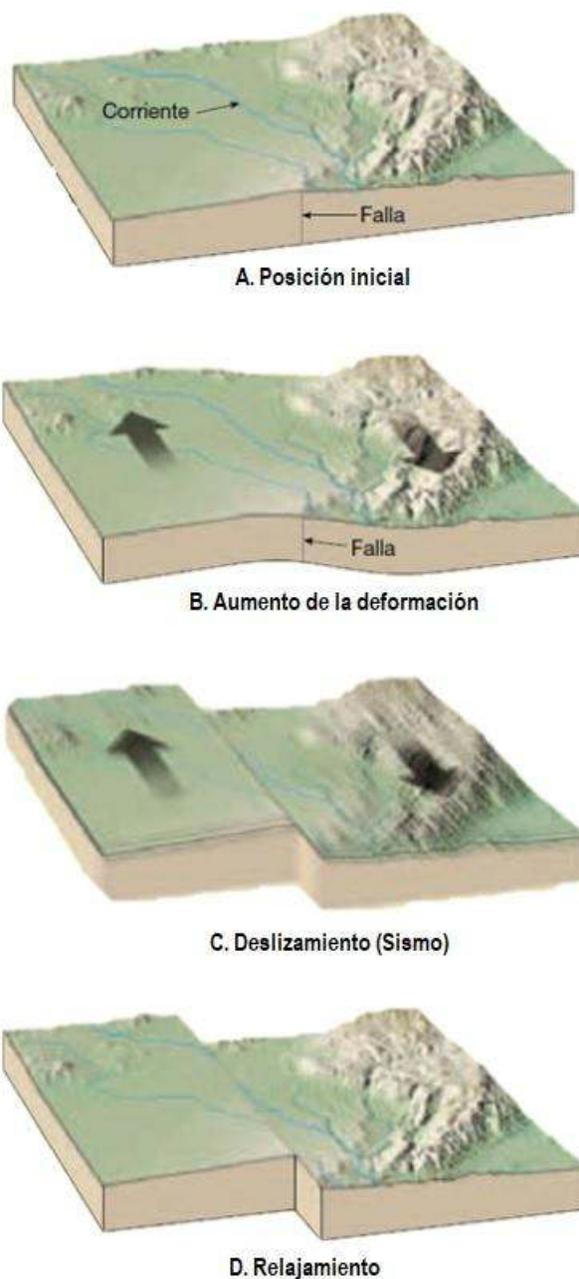


**Figura 1.1:** Proceso de convección en el cual el magma caliente sube a la superficie, se extiende y comienza a enfriarse. Luego se hunde de nuevo al fondo donde se recalienta y se eleva nuevamente. Dorsal: Lugar donde el Magma emerge y empuja a las placas; Fosa: Extremo de la placa que se hunde [1].

durante el sismo) [1, 3]. Estos fenómenos naturales pueden ser causados por fallas geológicas, erupciones volcánicas, o incluso por la explosión de bombas atómicas (aunque estos dos últimos casos, son relativamente débiles y ocurren con muy poca frecuencia) [1].

La existencia de sismos se debe principalmente a que el planeta es un sistema dinámico que está compuesto por cuatro partes principales: el núcleo interno, el núcleo externo, el manto y la corteza (Fig. 1.1). La parte superior del manto y la corteza forman una delgada capa en la superficie de nuestro planeta denominada **litosfera** [5]. Esta delgada capa se compone de muchas piezas, conocidas como **placas tectónicas**, que cubren toda la superficie de la tierra (como un rompecabezas). Debido a la convección del magma en la Tierra (Fig. 1.1), las placas se encuentran en constante movimiento chocando entre sí, generando un complejo conjunto de fracturas en el interior y en los bordes de dichas placas, denominadas **fallas geológicas** [1, 6].

En la Figura (1.2) se ilustra como el movimiento de las placas tectónicas produce sismos en dichas fallas. En la parte **A** de esta figura se observa una falla preexistente en la corteza. En **B**, el perpetuo movimiento de las placas tectónicas va deformando las rocas de la corteza a ambos lados de la falla. Debido a la fricción, estas rocas empiezan a almacenar energía elástica. Eventualmente, la tensión acumulada en las rocas es mayor que la fuerza de fricción entre los lados de la falla y ésta se desliza liberando la energía almacenada produciendo un sismo (Fig. 1.2.C). Este desli-



**Figura 1.2:** Almacenamiento de energía elástica a medida que las placas tectónicas se deforman. Cuando la tensión revasa la fuerza de rozamiento, la falla se desliza liberando energía en forma de ondas sísmicas [4].

zamiento permite que se disipe la tensión alrededor de la falla (Fig. 1.2.D). Este ciclo se reinicia debido al movimiento continuo de las placas tectónicas [4].

Las fallas geológicas se dividen en activas e inactivas. Las fallas activas pueden deslizarse súbitamente generando sismos, o pueden liberar energía gradualmente a través de "*sismos lentos*", caracterizados por un movimiento lento más o menos continuo que ocurre debido a una deformación tectónica continua denominada reptación. Por otro lado, las fallas inactivas son vestigios de actividad geológica en tiempos pasados y constituyen la mayoría de las fallas que se han registrado [4, 7].

Los procesos de deformación que ocurren en las fallas se dan de forma interconectada, en escalas que van desde milímetros a miles de kilómetros [8]. Por lo que una manera de interpretar la dinámica de las fallas geológicas es mediante la teoría de los sistemas que se autoorganizan de forma crítica [9-11]. Esta teoría describe el comportamiento de algunos sistemas dinámicos que a partir de la interacción natural entre las partes que los componen evolucionan a un estado '*crítico*' y equilibrado, donde perturbaciones aparentemente aleatorias conducen a eventos catastróficos de todos los tamaños. Esto produce una organización totalmente descentralizada e independiente de la escala espacial y/o temporal [12-15].

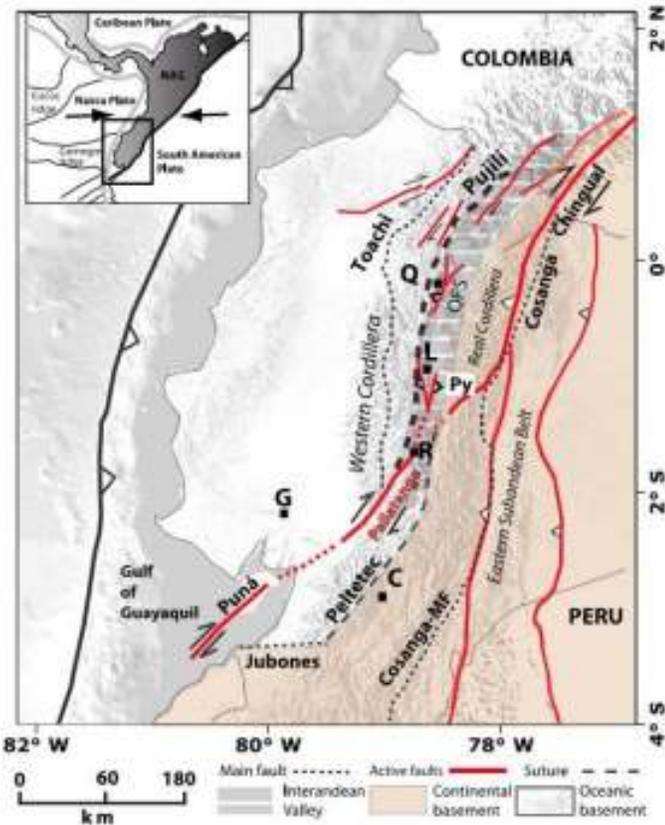
## 1.2 MARCO SISMOLÓGICO DEL ECUADOR

La ubicación del Ecuador en la costa oriental del Océano Pacífico hace al país parte del *Cinturón de Fuego del Pacífico*. Esta zona del planeta ha tomado esta denominación debido a que es la región más susceptible de experimentar los efectos de terremotos y erupciones volcánicas [16]. Frente a las costas del país, la placa oceánica Nazca se hunde (o se subduce) bajo la placa continental Sudamericana en dirección  $N83^{\circ}E$ , a una velocidad relativa aproximada de  $55 - 58\text{mm/a}$  con una inclinación de entre  $25^{\circ} - 30^{\circ}$ . La convergencia asociada con este proceso de subducción es responsable de la elevación de las montañas de los andes y de la cadena volcánica activa presente en el país [17].

En el Ecuador, debido a la subducción de la placa de Nazca, existen fallas geológicas en todo el territorio nacional (Fig. 1.3). Entre los sistemas de fallas más importantes se encuentran el Pallatanga-Chingual, que empieza en el golfo de Guayaquil, pasa por la Troncal y llega hasta Chingual; También, el sistema de fallas de Quito, que inicia en el sector de Tambillo y continua al norte por las lomas de Puengasí, etc [1].

### 1.3. ONDAS SÍSMICAS Y SU REGISTRO

La información proveniente del *Servicio Nacional de Sismología y Vulcanología* establece que anualmente el país experimenta un promedio de 2,600 eventos sísmicos [19]. Desde 1541 han ocurrido 38 terremotos de intensidad mayor o igual a VIII en la escala de intensidad MSK (viene de las iniciales Medvedev-Sponheuer-Karnik), intensidad desde la cual los efectos son importantes [1, 20]. Y si se toma en cuenta la magnitud de los eventos, existen 16 sismos mayores a 5 en la escala de Richter (equivalente a la energía liberada por 199 toneladas de explosivo TNT) [21]. De los cuales destacan el terremoto de 1906 ocurrido en Esmeraldas, cuya magnitud de 8.8 es el mayor registrado en Ecuador y Colombia [22]; y el terremoto de 2016 de magnitud 7.8 ocurrido en Pedernales [23]. A día de hoy, se calcula que directa o indirectamente alrededor de 80,000 seres humanos perdieron la vida a causa de este fenómeno natural y, en lo referente a perjuicios materiales, resulta imposible cuantificar una cantidad exacta [19].



**Figura 1.3:** Mapa tectónico activo del Ecuador. Los segmentos de fallas mayores se ilustran en líneas rojas continuas. Los restos de las zonas de subducción se dibujan en líneas de puntos negras [18].

### 1.3 ONDAS SÍSMICAS Y SU REGISTRO

Una **onda elástica** es una oscilación que se propaga sobre un medio material, el cual, cuando está sujeto a la acción de fuerzas externas, experimenta una resistencia interna a la deformación y recupera el estado original del material cuando estas fuerzas<sup>1</sup> ya no se aplican [24, 25]. Las **ondas sísmicas** se modelan como ondas elásticas dado que la manera y la velocidad con la que éstas viajan a través de la litosfera están descritas, aproximadamente, por propiedades elásticas

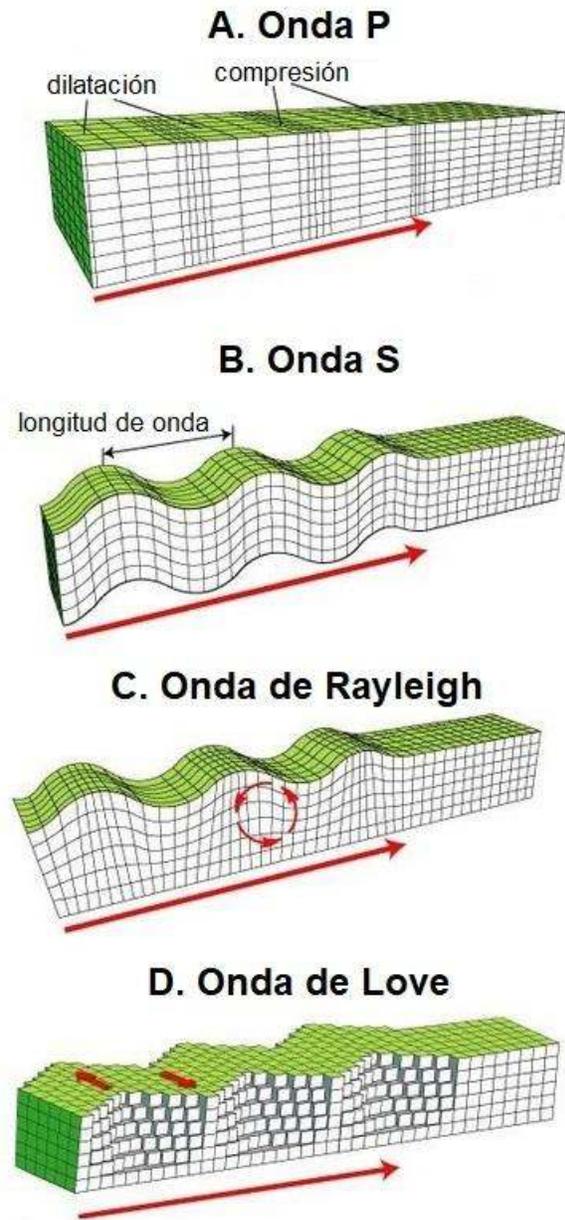
<sup>1</sup>Cuya magnitud no debe ser tan grande para comprometer la estructura del material.

[9]. Estas ondas son el producto de la liberación súbita de energía, que se irradia en todas las direcciones cuando una falla se desliza [7].

Existen dos grupos principales de ondas sísmicas. Las ondas de cuerpo, que viajan a través del interior de la Tierra y, las ondas superficiales, que viajan sobre la parte externa de la Tierra [4].

Las ondas de cuerpo se dividen a su vez en dos tipos, que se denominan **ondas primarias** o **P** y **ondas secundarias** o **S**. Las ondas P (Fig. 1.4.A) comprimen y dilatan el suelo en la dirección de propagación de la perturbación. A su vez, las ondas S (Fig. 1.4.B) se desplazan transversalmente a la dirección de propagación [27]. Las ondas primarias viajan generalmente a una velocidad 1.73 veces mayor que la de las ondas S y pueden viajar a través de cualquier tipo de material líquido o sólido, mientras que las ondas secundarias se trasladan únicamente a través de elementos líquidos y son las que producen la mayor cantidad de los daños [28].

La combinación de las ondas P y S cuando llegan a la superficie terrestre generan lo que se conoce como ondas superficiales. Existen dos tipos de ondas superficiales, las **ondas de Rayleigh** y las **ondas de Love**. Las ondas de Rayleigh (Fig. 1.4.C) se propagan generando que la superficie terrestre se mueva en forma ovalada, similar a las olas en la superficie del agua, a una velocidad menor a las ondas de cuerpo (aproximadamente el 90 % de la velocidad de las ondas S en medios elásticos homogéneos) [27]. Mientras que las ondas de Love (Fig. 1.4.D), también conocidas como ondas Q (Quer:

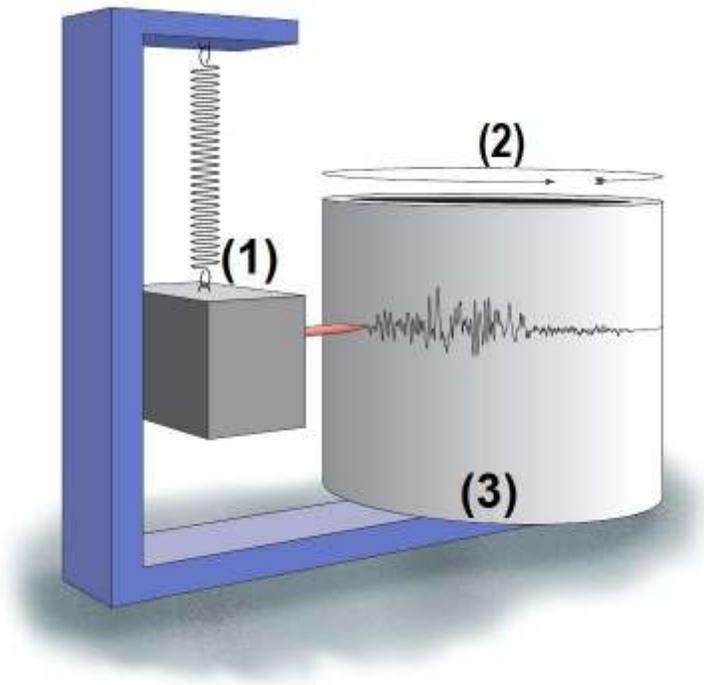


**Figura 1.4:** Un terremoto consta de una combinación de varios tipos de ondas sísmicas y sus movimientos característicos [26].

alemán para lateral), hacen que los materiales de la superficie terrestre se muevan de lado a lado en un ángulo de 90 grados con respecto a la dirección de propagación de la onda. Este movimiento es particularmente perjudicial para los cimientos de las estructuras [29].

Las ondas sísmicas son detectadas y registradas por un dispositivo denominado **sismógrafo**. Este instrumento consta de tres componentes (Fig. 1.5): **(1)** un sismómetro, que responde al movimiento del suelo y produce una señal proporcional a la aceleración o velocidad relativa a la tierra; **(2)** un dispositivo de temporización; y **(3)** un dispositivo de grabación que registra la actividad sísmica, dependiente del tiempo (denominado sismograma), en papeles o en medios de almacenamiento electrónico [31].

A pesar de que su diseño ha ido evolucionando a la par con la tecnología, los sismógrafos se basan en el principio de la inercia. Esta es la propiedad que tienen los cuerpos de mantener su estado de reposo o de movimiento a velocidad constante. Este dispositivo consta de una masa suspendida libremente de un soporte que se fija al terreno (Fig. 1.5). Cuando la vibración de un sismo alcanza el instrumento, la inercia de la masa suspendida la mantiene relativamente estacionaria mientras la tierra, junto al soporte, se mueven. El movimiento de la Tierra con respecto a la masa estacionaria se registra produciendo un sismograma [4]. En los sismómetros modernos el movimiento relativo entre la masa y la base generan un voltaje eléctrico que se registra en dispositivos de almacenamiento [30].



**Figura 1.5:** Sismógrafo sensible a movimientos verticales. Puede ser visualizado como una pesa suspendida de un resorte que a su vez están suspendidos sobre una base que se mueve con los movimientos de la superficie de la Tierra. Para añadir un sistema de registro se coloca un tambor que gira en la base y un marcador sujetado a la masa [30].

## 1.4 ANÁLISIS DE SERIES TEMPORALES

### 1.4.1 Procesos Estocásticos

Una **variable aleatoria**,  $Y$ , es una cantidad cuyo valor es el resultado de un proceso que siempre que se ejecuta, aparentemente bajo el mismo conjunto de condiciones iniciales, presenta resultados diferentes. Así pues, un **proceso estocástico** se define como una colección de variables aleatorias  $Y_t$ , donde  $t$  es un índice que determina el orden cronológico de dichas variables [32].

Una variable aleatoria que toma un número finito, o infinitamente contable, de valores  $R_Y = (y_1, y_2, y_3, \dots)^2$ , se denomina **variable aleatoria discreta**. En la mayoría de los casos, este tipo de variables suele representar datos que se cuentan, como el número de veces que se suscita un acontecimiento. Por el contrario, si la variable puede tomar cualquier valor numérico en uno o más intervalos reales,  $R_Y \subseteq \mathbb{R}$ , se denomina **variable aleatoria continua**. Generalmente, este tipo de variables aleatorias representa números reales como la duración de un suceso, o el peso de una persona [33].

Puesto que una variable aleatoria puede tomar cualquier valor dentro de un conjunto de posibles resultados, es necesario determinar una función matemática que defina la probabilidad relativa según la cual dicha variable aleatoria tome un determinado valor. A esta función se la denomina **función de densidad de probabilidad**  $f(y_t)$  (o función de probabilidad) [33]. Dependiendo del tipo de variable aleatoria, la función de probabilidad puede ser continua o discreta. Por ejemplo, la función continua de probabilidad normal o gaussiana tiene la forma [34]:

$$f(Y = y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(y - \mu)^2}{2\sigma^2}\right)}. \quad (1.1)$$

### 1.4.2 Medidas Características

El **valor esperado**, también llamado valor medio, de una variable aleatoria se define, en el caso discreto, como [32, 35]:

<sup>2</sup> $R_Y$  es el rango de posibles valores de la variable aleatoria  $Y$ .

$$E[Y] = \mu = \sum_{y_i \in R_Y} y_i f(y_i), \quad (1.2)$$

mientras que para una variable aleatoria continua en un determinado intervalo  $(a, b)$ , se define como:

$$E[Y] = \mu = \int_a^b y f(y) dy .$$

Este operador es la principal medida de centralización de una distribución de probabilidad y tiene las siguientes propiedades [33, 35]:

$$\begin{aligned} E[E[X]] &= E[X] \\ E[aX] &= aE[X] \\ E[X + Y] &= E[X] + E[Y] \\ X \geq 0 &\rightarrow E[X] \geq 0 \\ X \leq Y &\rightarrow E[X] \leq E[Y] \end{aligned}$$

donde  $X$  y  $Y$  son dos variables aleatorias y  $a$  es una constante.

La medida que determina cuan extendida o dispersa es una distribución de probabilidad con respecto al valor medio se denomina **varianza** [33, 34], y se define matemáticamente como:

$$E[(Y - \mu)^2] = \sigma^2 = \sum_{y_i \in R_Y} (y_i - \mu)^2 f(y_i), \quad (1.3)$$

para el caso discreto, y como:

$$E[(Y - \mu)^2] = \sigma^2 = \int_a^b (y - \mu)^2 f(y) dy ,$$

para el caso continuo.

La **autocovarianza** es una cantidad que mide la variabilidad conjunta de dos variables aleatorias de un mismo proceso estocástico en diferentes instantes, es decir, la autocovarianza es una generalización de la varianza ( $\gamma_{0t} = \sigma^2$ ). Su representación matemática para una variable aleatoria discreta es [34, 36]:

$$E [(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})] = \gamma_{jt} = \sum_{y_i \in \mathbb{R}_Y} (y_i - \mu_t) (y_i - \mu_{t-j}) f(y_i), \quad (1.4)$$

mientras que para una variable aleatoria continua:

$$E [(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})] = \gamma_{jt} = \iiint \cdots \int (y_t - \mu_t) (y_{t-j} - \mu_{t-j}) f_{Y_t, Y_{t-1}, \dots, Y_{t-j}}(y_t, y_{t-1}, \dots, y_{t-j}) dy_t dy_{t-1} \cdots dy_{t-j}.$$

Un proceso estocástico se denomina estacionario si la media (Ec. 1.2) y la autocovarianza (Ec. 1.4) no dependen explícitamente del tiempo [34, 37]:

$$\forall t \wedge j \begin{cases} E [Y_t] = \mu \\ E [(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})] = \gamma \end{cases} \quad (1.5)$$

## 1.5 SERIES TEMPORALES

Una **serie temporal** (Ec. 1.6) es una secuencia de datos, observaciones o valores medidos que están ordenados cronológicamente en intervalos igualmente espaciados en el tiempo [34]. También, se la puede considerar como una sucesión de variables aleatorias indexadas según un parámetro creciente con el tiempo [33]. En sismología, estas series se usan para obtener una comprensión de las fuerzas subyacentes y la estructura que produjo los datos observados. En consecuencia, un sismograma es un caso especial de una serie temporal.

$$\{y_t\}_{t=1}^T = (y_1, y_2, \dots, y_{T-1}, y_T) \quad (1.6)$$

Los ejemplos de series temporales abundan en diversos campos: en economía (una secuencia mensual de la cantidad de bienes enviados desde una fábrica), en ingeniería (observaciones por hora, realizadas sobre el rendimiento de un proceso químico), en ciencias sociales (una serie semanal del número de accidentes de tráfico), en las ciencias naturales (especialmente en geofísica y meteorología, con medidas como cantidades diarias de lluvia), etc [37].

Una serie temporal en la que todos los elementos se derivan de la misma distribución de probabilidad y son mutuamente independientes (i.i.d.),  $\{\epsilon_t\}_{t=-\infty}^{\infty}$ , se denomina *ruido blanco* si además

satisface las siguientes condiciones [32, 34, 38]:

$$\begin{cases} E[\epsilon_t] = 0 \\ E[\epsilon_t^2] = \sigma^2 \\ E[\epsilon_t \epsilon_\tau] = 0 \quad \forall t \neq \tau \end{cases} \quad (1.7)$$

El **análisis de series temporales** pretende encontrar un modelo matemático que proporcione una buena representación de las observaciones empíricas [37, 39]. Este análisis puede hacerse mediante métodos paramétricos o no paramétricos. Los planteamientos paramétricos suponen que el proceso estocástico estacionario, relacionado con la serie analizada, tiene una determinada estructura que puede describirse mediante una función que depende de varios parámetros. Cuando lo que se busca es un modelo paramétrico, la tarea es estimar los parámetros del modelo mediante diferentes técnicas estadísticas. Por otro lado, el análisis no paramétrico de las series temporales utiliza técnicas que no suponen función alguna, ya sea por el planteamiento del problema o porque ni siquiera se sabe describir analíticamente una estructura adecuada para los datos medidos. [36, 38]. En consecuencia, este método busca relaciones implícitas entre las observaciones de la serie temporal a través de transformaciones, medidas de dispersión, etc.

## 1.6 MOTIVACIÓN Y OBJETIVO DE ESTE TRABAJO

Tomando en consideración:

1. que es innegable la vulnerabilidad de la sociedad ecuatoriana ante los potenciales riesgos que representan los terremotos,
2. que el conocimiento científico de este fenómeno natural es un componente esencial para la toma de decisiones con respecto a la inversión público/privada, la planificación del desarrollo y la seguridad nacional en un país, y
3. que hoy en día no existe un modelo que describa satisfactoriamente la sismicidad de una región,

la motivación para llevar a cabo este trabajo de titulación es extraer información significativa de series temporales sísmicas, registradas por los sismógrafos del Instituto Geofísico de la Escuela

Politécnica Nacional (IG-EPN) y así contribuir en lo posible al desarrollo de modelos sísmicos para, en un futuro, prevenir los desastres asociados a los terremotos.

El objetivo de esta tesis es, sin embargo, mucho más modesto: Determinar si un sismograma puede ser descrito adecuadamente por un modelo estocástico paramétrico ARMA (en su formulación general que supone distribución de residuos gaussiana), modelo que se presenta en el capítulo 2.

### **1.7 OUTLINE**

Este trabajo de titulación está organizado de la siguiente manera. En el Capítulo 2 presentamos las técnicas ya conocidas que van a ser utilizadas para reconstruir y pronosticar series temporales. En el Capítulo 3, tras describir el procedimiento para convertir la información extraída de los sismógrafos a un formato adecuado para su análisis, presentamos la parte central de este trabajo, que consiste en diseñar el algoritmo que define la estructura del Proceso ARMA y ajustar este modelo (usando Estimación de Máxima Verosimilitud) a una serie temporal sísmica específica. En el Capítulo 4, comparamos parte de los datos empíricos con los resultados del algoritmo propuesto para analizar si el modelo propuesto tiene validez o carece de sentido. Finalmente, en el Capítulo 5, se discuten las conclusiones de este estudio.

## 2 MARCO TEÓRICO

Este capítulo presenta las técnicas básicas que utilizaremos para analizar series temporales sísmicas. Se introduce el modelo estocástico paramétrico denominado Proceso ARMA y la técnica estadística llamada Estimación de Máxima Verosimilitud. También describimos el algoritmo Davidon-Fletcher-Powell que usaremos para encontrar máximos y mínimos locales de funciones multivariantes. Finalmente, se presenta las técnicas que se utilizarán para validar o refutar los resultados del Proceso ARMA: Transformación de Fourier, el Criterio de Información de Akaike, el Coeficiente de Eficiencia del Modelo Nash-Sutcliffe y el Índice de Acuerdo de Willmott. Todas estas técnicas y métodos van a resultar esenciales para el trabajo que se hará en esta tesis.

### 2.1 PROCESO ARMA

#### 2.1.1 Proceso Auto-Regresivo

Un proceso Auto-regresivo de orden  $p$ , denotado  $AR(p)$ , es un proceso estocástico,  $Y_t$ , en el cual cada término es una combinación lineal de sus  $p$  valores anteriores más un término,  $\epsilon_t$ , de ruido blanco (Ec. 1.7) [34]. Matemáticamente se representa como:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad , \quad p \in \mathbb{Z}^+ ; \quad (2.1)$$

donde  $c$  es una constante que puede tomar cualquier valor real y el conjunto de parámetros  $\{\phi_1, \phi_2, \dots, \phi_p\}$  determina la influencia de cada uno de los valores anteriores en el nuevo término del proceso auto-regresivo [37]. Este modelo se usa para describir ciertos procesos que varían en el tiempo. En la igualdad anterior, el ruido blanco sirve para introducir en el modelo cierto tipo de perturbaciones estocásticas.

El modelo auto-regresivo asume que los procesos que modela son estacionarios (Ec. 1.5), así

## 2.1. PROCESO ARMA

---

la forma en la que el proceso cambia no varía en el tiempo. Dicho de otra manera, este modelo supone que la media y la autocovarianza de la serie temporal estudiada no cambian con el tiempo. En consecuencia, es necesario establecer condiciones en la familia de parámetros  $\{\phi_i\}_{i=1}^p$  para que el proceso paramétrico  $AR(p)$  sea estacionario [34, 37].

Empecemos con el proceso auto-regresivo de orden 1 (representado como  $AR(1)$ ), que en base a la ecuación (2.1) tiene la forma:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t . \quad (2.2)$$

Este proceso es estacionario si su media y autocovarianza no dependen explícitamente del tiempo. Para determinar el valor medio del proceso  $AR(1)$ , consideremos tomar el valor esperado en ambos lados de la expresión (2.2),

$$\begin{aligned} E[Y_t] &= E[c + \phi Y_{t-1} + \epsilon_t] \\ &= E[c] + \phi E[Y_{t-1}] + E[\epsilon_t] , \end{aligned}$$

aplicando las condiciones de un proceso estacionario (1.5) y de ruido blanco (1.7) la expresión anterior toma la siguiente forma:

$$E[Y_t] = \mu = c + \phi \mu + 0 ,$$

por lo que el valor medio,  $\mu$ , es:

$$\mu = \frac{c}{1 - \phi} , \quad (2.3)$$

el cual está bien definido y no depende del tiempo  $t$  si  $\phi \neq 1$ .

Más aún, reemplazando la igualdad (2.3) en (2.2),

$$\begin{aligned} Y_t &= \mu(1 - \phi) + \phi Y_{t-1} + \epsilon_t \\ Y_t - \mu &= \phi(Y_{t-1} - \mu) + \epsilon_t , \end{aligned}$$

y multiplicando ambos lados de la igualdad anterior por el término  $(Y_{t-j} - \mu)$  y tomando el valor esperado de ese producto, la autocovarianza del proceso  $AR(1)$  es:

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= \phi E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + E[(Y_{t-j} - \mu)\epsilon_t] \\ &= \phi E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + E[Y_{t-j}\epsilon_t] - \mu E[\epsilon_t] . \end{aligned}$$

En la definición de un proceso auto-regresivo de orden 1 (Ec. 2.2)  $Y_{t-j}$  depende linealmente de  $\epsilon_{t-j}, \epsilon_{t-j-1}, \epsilon_{t-j-2}, \dots$  a través de  $Y_{t-j}, Y_{t-j-1}, Y_{t-j-2}, \dots$  respectivamente, es decir:

$$\begin{aligned} Y_{t-j} &= c + \phi Y_{t-j-1} + \epsilon_{t-j} \\ &= c + \phi(c + \phi Y_{t-j-2} + \epsilon_{t-j-1}) + \epsilon_{t-j} \\ &= c + \phi(c + \phi(c + \phi(\dots) + \epsilon_{t-j-2}) + \epsilon_{t-j-1}) + \epsilon_{t-j} , \end{aligned}$$

que al reordenar los términos se convierte en la siguiente igualdad:

$$Y_{t-j} = c(1 + \phi + \phi^2 + \dots) + \epsilon_{t-j} + \epsilon_{t-j-1}\phi + \epsilon_{t-j-2}\phi^2 + \dots , \quad (2.4)$$

y al multiplicarla por  $\epsilon_t$  para tomar el valor esperado se tiene el siguiente resultado:

$$\begin{aligned} E[Y_{t-j}\epsilon_t] &= E[(c(1 + \phi + \phi^2 + \dots) + \epsilon_{t-j} + \epsilon_{t-j-1}\phi + \epsilon_{t-j-2}\phi^2 + \dots) \epsilon_t] \\ &= c(1 + \phi + \phi^2 + \dots)E[\epsilon_t] + E[\epsilon_{t-j}\epsilon_t] + \phi E[\epsilon_{t-j-1}\epsilon_t] + \phi^2 E[\epsilon_{t-j-2}\epsilon_t] + \dots . \end{aligned}$$

Por definición de ruido blanco (Ec. 1.7)  $\epsilon_t$  no está correlacionado con  $\epsilon_{t-j}$  (para  $j = 1, 2, \dots$ ), y la igualdad anterior se reduce a la siguiente expresión:

$$E[Y_{t-j}\epsilon_t] = 0 ,$$

con lo cual, la autocovarianza se la puede expresar de la siguiente forma:

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= \phi E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] + 0 - \mu 0 \\ E[(Y_t - \mu)(Y_{t-j} - \mu)] &= \phi E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] , \end{aligned}$$

que no depende explícitamente del tiempo. Con lo cual, siempre y cuando  $\phi \neq 1$  (dado que así la media y la autocovarianza están bien definidas y no dependen del tiempo) el proceso auto-regresivo de orden 1 es estacionario.

Consideremos ahora el proceso auto-regresivo de orden  $p$ . Para determinar qué condiciones,

## 2.1. PROCESO ARMA

---

sobre la familia de parámetros  $\{\phi_i\}_{i=1}^p$ , hacen que el proceso  $AR(p)$  (con  $p \in \mathbb{Z}^+$ ) sea estacionario podemos empezar por calcular su valor esperado. Así pues, recordando las condiciones (1.5) y (1.7) se tiene:

$$\begin{aligned} E[Y_t] &= E\left[c + \epsilon_t + \sum_{i=1}^p \phi_i Y_{t-i}\right] \\ E[Y_t] &= E[c] + E[\epsilon_t] + \phi_1 E[Y_{t-1}] + \phi_2 E[Y_{t-2}] + \cdots + \phi_p E[Y_{t-p}] \\ \mu &= c + 0 + \phi_1 \mu + \phi_2 \mu + \cdots + \phi_p \mu \end{aligned}$$

y despejando  $\mu$  se obtiene la siguiente expresión:

$$\mu = \frac{c}{1 - \sum_{i=1}^p \phi_i} . \quad (2.5)$$

También, si se reemplaza la ecuación anterior en la igualdad (2.1) y se toma el valor esperado del producto  $(Y_t - \mu)(Y_{t-j} - \mu)$  se tiene la autocovarianza,

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \sum_{i=1}^p \phi_i E[(Y_{t-i} - \mu)(Y_{t-j} - \mu)] \quad \text{para } j \neq i; j = 1, 2, \dots \quad (2.6)$$

Es evidente que la media (Ec. 2.5) y la autocovarianza (Ec. 2.6) no dependen explícitamente del tiempo; por lo que el proceso  $AR(p)$  es estacionario. Sin embargo, estos resultados dependen de que la sumatoria del denominador del lado derecho de la ecuación (2.5) converja, incluso para el caso cuando  $p$  tiende al infinito:

$$\mu = \frac{c}{1 - \sum_{j=1}^{\infty} \phi_j} , \quad (2.7)$$

donde la igualdad (2.7) es el valor medio del proceso auto-regresivo de orden infinito:

$$Y_t = c + \epsilon_t + \sum_{j=1}^{\infty} \phi_j Y_{t-j} . \quad (2.8)$$

Por lo tanto, para determinar bajo qué condiciones los parámetros auto-regresivos,  $\{\phi_j\}_{j \in \mathbb{Z}^+}$ , hacen que el valor medio esté bien definido utilizamos el criterio de Cauchy estocástico. Este criterio establece que una secuencia de números, en el caso del proceso auto-regresivo (Ec. 2.8):

$$\sum_{j=1}^{\infty} \phi_j Y_{t-j},$$

converge si y solo si, para cualquier  $\delta > 0$ , existe un número  $N$  adecuadamente grande tal que para  $M > N$ , una vez que se han sumado  $N$  términos, la diferencia entre esa suma y la obtenida sumando hasta  $M$  es una variable aleatoria ( $\phi_j Y_{t-j}$ ) cuya media y varianza son arbitrariamente cercanas a cero [33, 34]

$$E \left[ \sum_{j=1}^M \phi_j Y_{t-j} - \sum_{j=1}^N \phi_j Y_{t-j} \right]^2 < \delta. \quad (2.9)$$

El lado izquierdo de la condición (2.9) se puede reescribir como:

$$\begin{aligned} E \left[ \sum_{j=1}^M \phi_j Y_{t-j} - \sum_{j=1}^N \phi_j Y_{t-j} \right]^2 &= E \left[ \phi_M Y_{t-M} + \phi_{M-1} Y_{t-M+1} + \dots + \phi_{N+1} Y_{t-N-1} \right]^2 \\ &= E \left[ \sum_{i=N+1}^M \phi_i^2 Y_{t-i}^2 + 2 \sum_{j=N+1; j \neq k}^M \phi_j Y_{t-j} \left( \sum_{k=N+1}^M \phi_k Y_{t-k} \right) \right] \\ &= E \left[ \sum_{i=N+1}^M \phi_i^2 Y_{t-i}^2 \right] + 2E \left[ \sum_{j=N+1; j \neq k}^M \phi_j Y_{t-j} \left( \sum_{k=N+1}^M \phi_k Y_{t-k} \right) \right] \\ &= E \left[ \phi_M^2 Y_{t-M}^2 + \phi_{M-1}^2 Y_{t-M+1}^2 + \dots + \phi_{N+1}^2 Y_{t-N-1}^2 \right] \\ &\quad + 2E \left[ \sum_{j=N+1; j \neq k}^M \sum_{k=N+1}^M \phi_j Y_{t-j} \phi_k Y_{t-k} \right] \\ &= \phi_M^2 E \left[ Y_{t-M}^2 \right] + \dots + \phi_{N+1}^2 E \left[ Y_{t-N-1}^2 \right] \\ &\quad + 2E \left[ \sum_{j=N+1; j \neq k}^M \sum_{k=N+1}^M \phi_j \phi_k E \left[ Y_{t-j} Y_{t-k} \right] \right] \end{aligned}$$

donde  $Y_t$  depende linealmente de  $\epsilon_{t-1}, \epsilon_{t-2}, \dots$ , a través de  $Y_{t-1}, Y_{t-2}, \dots$  respectivamente (de manera similar a la ecuación (2.4) para el proceso  $AR(1)$ ) y como  $\epsilon_{t-j}$  no es correlacionado con  $\epsilon_{t-k}$  para  $j \neq k$ ,

$$E[\epsilon_{t-j} \epsilon_{t-k}] = 0,$$

con lo cual,

$$E[Y_{t-j} Y_{t-k}] = 0,$$

el lado izquierdo de la expresión (2.9) toma la siguiente forma [37]:

$$\begin{aligned} E \left[ \sum_{j=1}^M \phi_j Y_{t-j} - \sum_{j=1}^N \phi_j Y_{t-j} \right]^2 &= \phi_M^2 \sigma^2 + \phi_{M-1}^2 \sigma^2 + \dots + \phi_{N+1}^2 \sigma^2 + 0 \\ &= \left[ \sum_{j=1}^M \phi_j^2 - \sum_{j=1}^N \phi_j^2 \right] \sigma^2, \end{aligned}$$

y por consiguiente, el criterio de Cauchy (Ec. 2.9) se transforma en:

$$\left[ \sum_{j=1}^M \phi_j^2 - \sum_{j=1}^N \phi_j^2 \right] \sigma^2 < \delta. \quad (2.10)$$

En la igualdad anterior, si a la familia de parámetros auto-regresivos,  $\{\phi_j\}_{j \in \mathbb{Z}^+}$ , se impone la condición:

$$|\phi_j| < 1 \quad \forall j \in \mathbb{Z}^+,$$

las series convergen ya que son sumables al cuadrado, es decir [40, 41]:

$$\sum_{j=1}^{\infty} \phi_j^2 < \infty.$$

Por esta razón, el criterio de Cauchy (Ec. 2.9) se cumple dado que el lado izquierdo de la igualdad (2.10) puede ser tan arbitrariamente pequeño como se desee mediante la elección de un  $N$  adecuadamente grande. En consecuencia:

1. La secuencia de valores,  $\{\phi_i Y_{t-i}\}_{i \in \mathbb{Z}^+}$ , que caracteriza el proceso auto-regresivo (Ec. 2.1) va a permanecer con media y varianza arbitrariamente cercanas a cero (Ec. 2.9), incluso en el caso cuando  $p$  tiende al infinito (2.8).
2. El valor medio del proceso  $AR(p)$  (Ec. 2.5) está bien definido, dado que la sumatoria del denominador de esta expresión, como establece la condición (2.10), converge en una vecindad cercana a cero. Por lo que el proceso auto-regresivo es estacionario para cualquier  $p \in \mathbb{Z}^+$  [42].

### 2.1.2 Proceso de Media Móvil

Otro tipo de modelo es el proceso de media móvil (abreviado como  $MA(q)$  del inglés Moving Average), donde cada nuevo valor de este proceso es una constante más ruido aleatorio independiente y una fracción del ruido aleatorio de instantes anteriores [37]. Este modelo describe como valores aleatorios influyen en la evolución de un proceso invariable. Su representación matemática es:

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad , \quad q \in \mathbb{Z}^+ ; \quad (2.11)$$

en la cual  $\{\epsilon_t\}_{t=1}^q$  es ruido blanco (Ec. 1.7),  $\mu$  es el valor medio y los parámetros  $(\theta_1, \theta_2, \dots, \theta_q)$ , a diferencia de lo que pasa en el proceso auto-regresivo, son números reales sin ninguna restricción [34]. Esto se debe a que, por su definición (Ec. 2.11), este proceso ya es estacionario. En otras palabras, la media y la autocovarianza del proceso  $MA(q)$  no varían con el tiempo [34, 37].

Ver esto es sencillo si tomamos el valor esperado a ambos lados de la igualdad (2.11) y, recordando la definición de ruido blanco (Ec. 1.7):

$$\begin{aligned} E[Y_t] &= E[\mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}] \\ &= E[\mu] + E[\epsilon_t] + \theta_1 E[\epsilon_{t-1}] + \theta_2 E[\epsilon_{t-2}] + \cdots + \theta_q E[\epsilon_{t-q}] \\ &= \mu + 0 + \theta_1 0 + \theta_2 0 + \cdots + 0 , \end{aligned}$$

el valor medio del proceso  $MA(q)$  es igual a una constante:

$$E[Y_t] = \mu . \quad (2.12)$$

Para el caso de la autocovarianza, en la ecuación (2.11) se aplica el valor esperado a la diferencia entre  $Y_t$  y  $\mu$  de la siguiente manera:

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= E[(\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}) \\ &\quad (\epsilon_{t-j} + \theta_1 \epsilon_{t-j-1} + \theta_2 \epsilon_{t-j-2} + \cdots + \theta_q \epsilon_{t-j-q})] . \end{aligned}$$

Como, por definición, los términos de ruido blanco en diferentes instantes no se correlacionan, el valor esperado de su producto es igual a cero:

$$E[\epsilon_t \epsilon_\tau] = 0 \quad ; \forall t \neq \tau ,$$

con lo cual,

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= E\left[\theta_j \epsilon_{t-j}^2 + \theta_{j+1} \theta_1 \epsilon_{t-j-1}^2 + \theta_{j+2} \theta_2 \epsilon_{t-j-2}^2 + \cdots + \theta_q \theta_{q-j} \epsilon_{t-q}^2\right] \\ &= \theta_j E\left[\epsilon_{t-j}^2\right] + \theta_{j+1} \theta_1 E\left[\epsilon_{t-j-1}^2\right] + \theta_{j+2} \theta_2 E\left[\epsilon_{t-j-2}^2\right] + \\ &\quad \cdots + \theta_q \theta_{q-j} E\left[\epsilon_{t-q}^2\right], \end{aligned}$$

la autocovarianza para  $j = 1, 2, \dots, q$ , no depende del tiempo:

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = (\theta_j + \theta_{j+1} \theta_1 + \theta_{j+2} \theta_2 + \cdots + \theta_q \theta_{q-j}) \sigma^2. \quad (2.13)$$

En los valores,  $Y_t$ , del modelo de media móvil (Ec. 2.11), el ruido blanco tiene influencia tanto en el instante actual como en los  $q$  períodos anteriores. Esto le otorga a este proceso una memoria a corto plazo. Por el contrario, en el modelo auto-regresivo (Ec. 2.1), el ruido blanco afecta a valores de  $Y_t$  infinitamente lejanos en el futuro. Esto se debe a que, debido a la estructura que define este modelo,  $\epsilon_t$  afecta  $Y_t$ , que afecta a  $Y_{t+1}$ , que afecta a  $Y_{t+2}$ , y así sucesivamente [43]. Es así como la combinación de estos dos modelos proporciona una mejor descripción de sistemas donde la interacción entre todas las partes que los componen.

### 2.1.3 Proceso Auto-Regresivo de Media Móvil (ARMA)

En su tesis doctoral de 1951 [44], Peter Whittle estableció que la representación de un proceso estocástico tiene una mayor flexibilidad al incluir términos de media móvil y auto-regresivos en el modelo. Así que combinó el modelo Auto-regresivo de orden  $p$  (Ec. 2.1) con el proceso de Media Móvil de orden  $q$  (Ec. 2.11) en un solo modelo denominado proceso  $ARMA(p, q)$  (Del inglés Autoregressive Moving Average). Matemáticamente este proceso se representa como:

$$Y_t = c + \overbrace{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}}^{AR(p)} + \overbrace{\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}}^{MA(q)} \quad (2.14)$$

donde los  $p+q+2$  parámetros  $(\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, c, \sigma^2)$  ajustan el modelo a un determinado caso. Dicho de otra manera, un proceso Auto-Regresivo de Media Móvil define una familia de series temporales que se caracteriza por tener errores aleatorios  $\{\epsilon_t\}$  descritos por una determinada densidad de probabilidad conjunta.

## 2.1. PROCESO ARMA

---

Para que la estructura matemática del proceso  $ARMA(p, q)$  produzca series temporales que varíen en el tiempo de una manera establecida, al igual que en el proceso  $AR(p)$ , este modelo debe ser estacionario; es decir, la familia de parámetros  $\{\phi_i\}$  debe tener como restricción que:

$$|\phi_i| < 1 \text{ para } i = 1, 2, \dots, p. \quad (2.15)$$

Como prueba de esto considere aplicar el valor esperado en ambos lados de la ecuación (2.14),

$$\begin{aligned} E[Y_t] &= E[c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}] \\ &= E[c] + E[\phi_1 Y_{t-1}] + E[\phi_2 Y_{t-2}] + \dots + E[\phi_p Y_{t-p}] + \\ &\quad E[\epsilon_t] + E[\theta_1 \epsilon_{t-1}] + E[\theta_2 \epsilon_{t-2}] + \dots + E[\theta_q \epsilon_{t-q}] \\ &= E[c] + \phi_1 E[Y_{t-1}] + \phi_2 E[Y_{t-2}] + \dots + \phi_p E[Y_{t-p}] + \\ &\quad E[\epsilon_t] + \theta_1 E[\epsilon_{t-1}] + \theta_2 E[\epsilon_{t-2}] + \dots + \theta_q E[\epsilon_{t-q}] \end{aligned}$$

y recordando las condiciones de un proceso estacionario (1.5) y ruido blanco (1.7),

$$\mu = c + \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu + 0 + \theta_1 0 + \theta_2 0 + \dots + \theta_q 0,$$

se tiene el siguiente resultado [45]:

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p}, \quad (2.16)$$

que establece restricciones solo a los parámetros auto-regresivos, más no en la familia de parámetros de media móvil [34, 36].

Nuevamente, reemplazando la ecuación (2.16) en la igualdad (2.14) y despejando la desviación de la media  $(Y_t - \mu)$  para multiplicarlo por  $(Y_{t-j} - \mu)$  y tomando el valor esperado, la autocovarianza para  $j = q + 1, q + 2, \dots$  tiene la forma:

$$\begin{aligned} E[(Y_t - \mu)(Y_{t-j} - \mu)] &= \phi_1 E[(Y_{t-1} - \mu)(Y_{t-j} - \mu)] \\ &\quad + \phi_2 E[(Y_{t-2} - \mu)(Y_{t-j} - \mu)] + \dots + \phi_p E[(Y_{t-p} - \mu)(Y_{t-j} - \mu)] \quad (2.17) \end{aligned}$$

Tanto la expresión (2.16) como la expresión (2.17) no dependen explícitamente del tiempo, y por tanto el proceso  $ARMA(p, q)$  es estacionario siempre y cuando los parámetros auto-regresivos cumplan la condición (2.15).

## 2.2 ESTIMACIÓN DE MÁXIMA VEROSIMILITUD

### 2.2.1 Definición

Descrito oficialmente entre 1912 y 1922 por R. A. Fisher [46, 47], la Estimación de Máxima Verosimilitud (o MLE del inglés Maximum Likelihood Estimation) es un método utilizado ampliamente para estimar los parámetros de un modelo que describe un conjunto de observaciones. El MLE determina que el valor óptimo de los parámetros es aquel que maximiza la probabilidad de que el modelo produzca dichas observaciones [48].

En un proceso estocástico como medir la altura de los estudiantes de una clase, la probabilidad de que un estudiante mida  $x$  centímetros está descrito por una determinada distribución de probabilidad. Supongamos que dicha distribución es gaussiana (Ec. 1.1) y está caracterizada por los parámetros  $\mu$  y  $\sigma^2$ . Al variar estos parámetros, la función de distribución será una u otra. El método de estimación de máxima verosimilitud calcula, de entre todos los posibles valores que pueden tomar los parámetros del modelo, qué cantidades hacen que el modelo describa los datos empíricos de forma adecuada.

El ejemplo anterior bosqueja la idea que hay detrás de la estimación de máxima verosimilitud [32]. Detallado formalmente, supongamos un proceso estocástico,  $Y_t$  con  $t \in \mathbb{Z}^+$ , descrito por una determinada distribución de probabilidad que, digamos, depende de  $N$  parámetros (generalmente representados en el vector  $\Theta = (\theta_1, \theta_2, \dots, \theta_N)'$ ). Y supongamos que este proceso ha arrojado los siguientes resultados:  $y_1, y_2, \dots, y_t$ . Entonces, la probabilidad de que el proceso estocástico tenga como resultado dichos valores es:

$$f(y_t, y_{t-1}, \dots, y_1; \Theta) = P(Y_1 = y_1, Y_2 = y_2, \dots, Y_t = y_t; \Theta) . \quad (2.18)$$

Al contrario de una función de probabilidad, que determina la posibilidad de tener una observación como resultado de un proceso estocástico; la igualdad (2.18) ajusta el valor numérico del vector de parámetros  $\Theta$  en base los resultados (de los que depende dicha función) de un proceso

estocástico. Por ende, la ecuación (2.18) se la denomina *función likelihood* [34, 49]. Adicionalmente, como el logaritmo de una función alcanza su valor máximo en el mismo punto que dicha función<sup>1</sup>, por lo general se aplica el logaritmo a la función likelihood. Así, esta función se vuelve matemáticamente más simple y es más sencillo determinar su punto máximo [50].

$$\mathcal{L}(\Theta) = \ln [f(y_t, y_{t-1}, \dots, y_1; \Theta)]$$

Al maximizar la función likelihood en base al vector de parámetros  $\Theta$ , se establece el valor de las componentes de dicho vector que hacen más probable la aparición de los resultados empíricos,  $y_1, y_2, \dots, y_t$ , del proceso estocástico. El punto del espacio N-dimensional de parámetros que maximiza la función likelihood se denomina estimación de máxima verosimilitud [37]:

$$\Theta_{MLE} = \arg \max_{\Theta \in \mathbb{R}^t} f(y_t, y_{t-1}, \dots, y_1; \Theta) . \quad (2.19)$$

Calcular los parámetros de un modelo mediante la estimación de máxima verosimilitud involucra dos pasos. Primero debe obtenerse la función likelihood (Ec. 2.18). En el caso general del proceso ARMA (Ec. 2.14), la función likelihood que describe los residuos i.i.d. se supone gaussiana (Ec. 1.1) [34, 37, 49]. Segundo, debe calcularse el valor del conjunto de parámetros que maximizan la función likelihood gaussiana. El vector de parámetros a optimizar referente a dicha función está compuesto por los parámetros auto regresivos, los parámetros de media móvil, una constante y la varianza:

$$\Theta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, c, \sigma^2)'$$

### 2.2.2 Función Likelihood para un Proceso ARMA(p,q) Gaussiano

Consideremos el proceso ARMA(p, q) (Ec. 2.14):

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} ,$$

donde el término de *ruido blanco*,  $\epsilon_t$ , representa el residuo que aparece al aproximar una observación empírica mediante un determinado modelo:

<sup>1</sup>Propiedad denominada monótonamente creciente.

$$\epsilon_t = \underbrace{Y_t}_{Obs.} - \underbrace{(c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q})}_{Modelo}. \quad (2.20)$$

Todos los elementos de la familia de errores,  $\{\epsilon_t\}_{t=1}^T$ , son mutuamente excluyentes (Ec. 1.7) y tienen densidad de probabilidad gaussiana (Ec. 1.1). Por lo que la función likelihood conjunta para el proceso *ARMA* gaussiano es igual al producto de las funciones de probabilidad de cada elemento de dicha familia de ruido blanco [32, 33]:

$$f_{Y_T, Y_{T-1}, \dots, Y_1}(y_T, y_{T-1}, \dots, y_1 | \Theta) = \prod_{t=1}^T f_{Y_t}(y_t | y_{t-1}, \dots, y_{t-p}, \Theta). \quad (2.21)$$

Asimismo, Box y Jenkins [37] recomiendan empezar el cálculo de la función likelihood a partir del instante  $t = p + 1$ , con  $Y_1, Y_2, \dots, Y_p$  iguales a sus respectivos valores reales  $y_1, y_2, \dots, y_p$ . Con el valor de los errores aleatorios en estos instantes:

$$\epsilon_p = \epsilon_{p-1} = \dots = \epsilon_{p-(q+1)} = 0. \quad (2.22)$$

Así pues, empleando la igualdad (2.20), la distribución de probabilidad para  $\epsilon_t$  (con  $t = p + 1, p + 2, \dots, T$ ) es:

$$f_{Y_t}(y_t | y_{t-1}, \dots, y_{t-p}, \Theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-\epsilon_t^2}{2\sigma^2}\right]. \quad (2.23)$$

Por lo que la ecuación (2.21) toma la siguiente forma:

$$f(y_T, \dots, y_{p+1} | y_p, \dots, y_1, \epsilon_p = 0, \dots, \epsilon_{p-(q+1)} = 0) = \left[\frac{1}{\sqrt{2\pi\sigma^2}}\right]^{T-p} \prod_{t=p+1}^T \exp\left[\frac{-\epsilon_t^2}{2\sigma^2}\right] \quad (2.24)$$

Como es más fácil trabajar una suma de funciones cuadráticas en lugar del producto de funciones exponenciales, aplicamos el logaritmo en ambos lados de la expresión (2.24), teniendo como

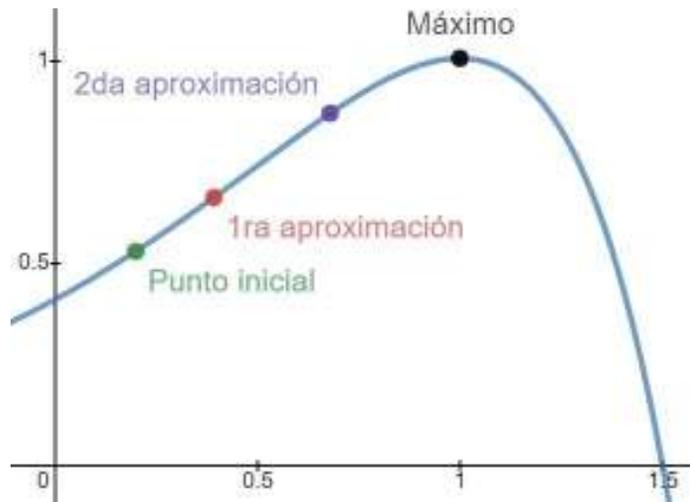
resultado:

$$\mathcal{L}(\Theta) = \ln \left[ (y_T, \dots, y_{p+1} | y_p, \dots, y_1, \epsilon_p = 0, \dots, \epsilon_{p-(q+1)} = 0) \right] = -\frac{T-p}{2} \ln(2\pi) - \frac{T-p}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=p+1}^T \epsilon_t^2 \quad (2.25)$$

En virtud de la complejidad de la expresión (2.25), es necesario un método numérico adecuado para poder encontrar el conjunto de parámetros,  $\Theta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, c, \sigma^2)'$ , que maximice esta función. El algoritmo utilizado para llevar a cabo esta tarea se describe a continuación.

### 2.3 ALGORITMO DAVIDON-FLETCHER-POWELL

El método de Davidon-Fletcher-Powell (también referido como DFP) es un algoritmo de optimización que busca el máximo o el mínimo de una función  $f(x)$  denominada 'función objetivo' [51]. Esta técnica, a partir de un punto inicial arbitrario, ejecuta operaciones (numéricas o lógicas) en un ciclo que se repite varias veces, con el resultado del ciclo anterior como valor inicial del nuevo ciclo (Fig. 2.1) [52]. Este procedimiento se denomina *algoritmo iterativo* y pertenece a los métodos *Quasi-Newton*. Estos métodos se basan en



**Figura 2.1:** Función a optimizar de una sola variable real en la cual, el algoritmo empieza en un punto inicial arbitrario y después de tres iteraciones se localiza lo suficientemente cerca al valor que maximiza dicha función.

una aproximación en serie de Taylor de segundo grado de la función objetivo [53]. El término *quasi* se debe a que estas técnicas se derivan del *Método de Newton* que se desarrolló originalmente para encontrar las raíces de funciones diferenciables [51, 54].

Al igual que todos los algoritmos de optimización, el método DFP busca el valor extremo (máximo o mínimo) de la función objetivo en una vecindad cercana al punto inicial. Así mismo, estos algoritmos no encuentran un valor exacto, sino que lo aproximan con cualquier precisión especi-

ficada mediante un "criterio de convergencia" que define cuando se está lo suficientemente cerca del punto mínimo o máximo [51]. Más aún, estos métodos pueden fallar si el punto inicial se encuentra muy distante del valor extremo de la función, las variables están mal escaladas o si la función objetivo no cumple con las condiciones que cada algoritmo específica para trabajar adecuadamente [55].

En el caso del algoritmo DFP, la función debe tener segunda derivada y debe ser cóncava en una vecindad cercana al punto N-dimensional que optimice la función objetivo en el espacio de parámetros. Por cóncava nos referimos a que -1 veces la matriz de segundas derivadas debe ser simétrica, y el producto escalar con cualquier vector tiene como resultado un número positivo [55].

Supongamos que  $\Theta$  es un vector de dimensión  $N$  que se va a optimizar. Sea  $g(\Theta_0)$  el vector gradiente de la función objetivo  $f(\Theta)$  evaluada en un punto N-dimensional inicial arbitrario  $\Theta_0$ :

$$g(\Theta_0) = \left. \frac{\partial f(\Theta)}{\partial \Theta} \right|_{\Theta=\Theta_0}$$

y  $H(\Theta_0)$  el negativo de la matriz hessiana de la función  $f(\Theta)$ :

$$H(\Theta_0) = - \left. \frac{\partial^2 f(\Theta)}{\partial \Theta \partial \Theta'} \right|_{\Theta=\Theta_0}.$$

Desarrollando  $f(\Theta)$  en serie de Taylor hasta segundo orden alrededor de  $\Theta_0$  [50, 53]:

$$f(\Theta) \cong f(\Theta_0) + g(\Theta_0)' [\Theta - \Theta_0] - 1/2 [\Theta - \Theta_0]' H(\Theta_0) [\Theta - \Theta_0] \quad (2.26)$$

El método de Newton-Raphson consiste en escoger  $\Theta$  de tal manera que la expresión (2.26) alcance su valor máximo, es decir, su derivada en este punto es igual a cero [52]:

$$g(\Theta_0) - H(\Theta_0) [\Theta - \Theta_0] = 0$$

Suponiendo que la inversa de la matriz hessiana  $[H(\Theta_0)]^{-1}$  existe, se define la siguiente expresión,

$$\Theta - \Theta_0 = [H(\Theta_0)]^{-1} g(\Theta_0).$$

que en notación iterativa toma la forma:

$$\Theta_{m+1} - \Theta_m = [H(\Theta_m)]^{-1} g(\Theta_m), \quad (2.27)$$

donde el m-ésimo paso de la iteración actualiza el valor estimado de  $\Theta$  hasta que algún criterio de convergencia sea cumplido. Nótese que en el lado derecho de la expresión (2.27), el vector  $[H(\Theta_m)]^{-1} g(\Theta_m)$  define la longitud y la dirección entre los vectores  $\Theta_{m+1}$  y  $\Theta_m$  conocida como la *dirección de Newton* ( $N_m$ ):

$$N_m = [H(\Theta_m)]^{-1} g(\Theta_m) . \quad (2.28)$$

Existe la posibilidad de que la inversa de la matriz hessiana  $[H(\Theta)]^{-1}$  no exista, y por esta razón las expresiones (2.27) y (2.28) no tengan sentido alguno. Asimismo, es necesario que la matriz hessiana sea definida positiva en los puntos que recorre el algoritmo hasta optimizar la función. Sin embargo, cuando se está lejos del extremo de la función, esto no ocurre necesariamente. Para evitar estos inconvenientes, y acelerar el algoritmo al no calcular explícitamente la matriz hessiana y posteriormente su inversa, Davidon [56] y Fletcher and Powell [57] considerando que la matriz hessiana ( $-H(\Theta)$ ), al ser la primera derivada del vector gradiente ( $g(\Theta)$ ), define la variación de  $g(\Theta)$  en función de la variación de  $\Theta$ . Con lo cual, propusieron reemplazar la inversa de la matriz hessiana, en la dirección de Newton (Ec. 2.28), por una secuencia de matrices que se actualizan en cada iteración [55]:

$$\lim_{i \rightarrow \infty} A_i = [H]^{-1} , \quad (2.29)$$

es decir,

$$N_m = A_m g(\Theta_m) , \quad (2.30)$$

con la igualdad (2.27) tomando la forma:

$$\Theta_{m+1} = \Theta_m + A_m(\Theta_m) g(\Theta_m) . \quad (2.31)$$

Una vez que  $\Theta_{m+1}$  y el gradiente  $g(\Theta_{m+1})$  se calculan, el valor de  $A_m$ , detallado en [58], se determina según la siguiente expresión:

$$A_{m+1} = A_m - \frac{A_m [g(\Theta_{m+1}) - g(\Theta_m)] [g(\Theta_{m+1}) - g(\Theta_m)]' A_m}{[g(\Theta_{m+1}) - g(\Theta_m)]' A_m [g(\Theta_{m+1}) - g(\Theta_m)]} - \frac{[\Theta_{m+1} - \Theta_m] [\Theta_{m+1} - \Theta_m]'}{[g(\Theta_{m+1}) - g(\Theta_m)]' [\Theta_{m+1} - \Theta_m]} \quad (2.32)$$

Puesto que la dirección de Newton (Ec. 2.30) depende del gradiente de la función objetivo (Ec. 2.30), el algoritmo DFP garantiza que dicha función varía más rápidamente en ésta dirección [53]. Pero también, siguiendo dicha dirección  $\Theta_{m+1}$  no necesariamente se acerca al valor que optimiza la función objetivo. Por lo que una manera de asegurar la convergencia del algoritmo es introducir un escalar positivo  $s$  en la dirección de Newton (Ec. 2.31):

$$\Theta_{m+1} = \Theta_m + s A_m (\Theta_m) g (\Theta_m) , \quad (2.33)$$

de tal forma que si la distancia entre los vectores  $\Theta_{m+1}$  y  $\Theta_m$  no satisface un determinado criterio (cuando  $s = 1$ ), se retrocede a lo largo de la dirección de Newton disminuyendo el valor del escalar  $s$  [34].

## 2.4 MÉTODOS DE VERIFICACIÓN

### 2.4.1 Transformada de Fourier

La *Transformación de Fourier* es una teoría desarrollada entre 1807 y 1822 cuando el matemático francés Jean-Baptiste Joseph Fourier estudiaba la transferencia de calor [59]. Esta teoría permite construir la aproximación de una función como la suma de ondas representadas matemáticamente por senos y cosenos [60, 61]. En esta aproximación la función, que originalmente se describe en unidades de tiempo (o espacio), pasa a ser representada en unidades de frecuencia. Más aún, esta transformación siempre se puede invertir [61].

Además de aproximar funciones, es posible analizar series temporales mediante este proceso. Por lo general, a estas series se las representa como la suma de una secuencia de datos,  $\{\psi_t\}_{t=-\infty}^{\infty}$ , que tiene la forma:

$$Y_t = \mu + \sum_{i=0}^{\infty} \theta_i \psi_{t-i} .$$

Al transformar esta representación al espacio de frecuencia, se describe el valor de  $Y_t$  como la suma ponderada de funciones periódicas,

$$Y_t = \mu + \int_0^{\pi} \alpha(\omega) \cos(\omega t) d\omega + \int_0^{\pi} \delta(\omega) \sin(\omega t) d\omega ,$$

donde  $\delta(\omega)$  y  $\alpha(\omega)$  representan la ponderación de las funciones sinusoidales [34]. En este dominio se puede determinar la importancia que tiene cada frecuencia en el comportamiento de  $Y_t$ , lo cual facilita la detección o eliminación de características asociadas a dichas frecuencias [55, 62]. En ocasiones, la descripción en el dominio del tiempo es más simple para determinadas características mientras que la descripción en el dominio de la frecuencia resulta más sencilla para estudiar otras características [63, 64].

### 2.4.2 Criterio de Información de Akaike

El Criterio de información de Akaike, también conocido como *AIC* (del inglés Akaike Information Criterion) [65], es un criterio estadístico que emplea información<sup>2</sup> para determinar la calidad relativa de un modelo en comparación a otros. Ningún modelo puede representar completamente la realidad. Por lo que al describir un proceso -de naturaleza desconocida- a través de un modelo siempre se omite cierta cantidad de información. El criterio de información de Akaike define la bondad de un modelo determinando la cantidad de información que este ignora en relación con otros modelos. Cuanta menos información omite un modelo "*más cercano*" es este a la realidad desconocida [66].

El principio de parsimonia (también llamado Navaja de Ockham) establece que cuando dos teorías en igualdad de condiciones conllevan a las mismas conclusiones, la explicación más sencilla suele ser la más probable [67]. Esto se debe a que mientras más parámetros tiene un modelo, este aumenta su complejidad. Por lo que, en el desarrollo de dicho modelo, es más fácil cometer errores en comparación con un modelo relativamente simple con pocos parámetros. Por tanto, el AIC determina el número de parámetros idóneo de un modelo buscando un balance entre simplicidad y eficiencia [66, 68].

En la práctica, se calcula el AIC para cada uno de los modelos candidatos y selecciona el modelo con el valor más pequeño de AIC. Matemáticamente, se representa de la siguiente forma [36]:

$$AIC = -2 \ln (f (\Theta_{MLE})) + 2K , \quad (2.34)$$

donde  $f (\Theta_{MLE})$  es el valor numérico de la función Likelihood (Ec. 2.18) evaluada en el punto que la maximiza (i.e. en el estimador de máxima verosimilitud, Ec. 2.19), y  $K$  es el número de parámetros que definen la estructura del modelo [68].

<sup>2</sup>Datos que una vez percibidos y procesados cambian el estado de conocimiento de un individuo o sistema.

Este criterio determina la bondad relativa de un modelo en comparación a otros. Por lo tanto, si todos los modelos describen pobremente el proceso, AIC seleccionará el mejor modelo de entre los propuestos a pesar de que el modelo seleccionado continuará siendo un modelo que describe mal los datos [68].

### 2.4.3 Coeficiente de Eficiencia del Modelo Nash-Sutcliffe

El *Coeficiente de Eficiencia del Modelo Nash-Sutcliffe* (ó *NSE* del inglés Nash–Sutcliffe Efficiency) es un índice desarrollado originalmente para determinar cuantitativamente la precisión de los resultados de modelos hidrológicos [69]. Este indicador es una generalización para modelos no lineales del Coeficiente de Determinación ( $R^2$ ) [32], y tiene un gran número de aplicaciones [70].

Este coeficiente utiliza datos empíricos para comprobar cual alejados de la realidad están los resultados del modelo analizado. Matemáticamente se representa de la siguiente forma [69]:

$$E_f = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - E[y_i])^2}, \quad (2.35)$$

donde  $y_i$  es el dato observado en el instante  $i$ , mientras que  $\hat{y}_i$  es la predicción realizada por el modelo en el mismo instante.

La expresión en el numerador del segundo término del lado derecho de la igualdad (2.35) es la suma de residuos,  $\epsilon_i$ , al cuadrado [34],

$$\epsilon_i = y_i - \hat{y}_i. \quad (2.36)$$

La suma cuadrática de estos residuos mide la variación en los datos que no han sido explicados por el modelo. Si el numerador se dividiera entre los grados de libertad y se tomara la raíz cuadrada de este cociente, el valor resultante sería igual al error estándar de estimación [71].

El coeficiente NSE puede tomar valores menores o iguales a uno ( $E_f \leq 1$ ). Si las predicciones del modelo coinciden perfectamente con la información observada, este coeficiente tomará el valor de uno ( $E_f = 1$ ). Si este coeficiente es igual a cero ( $E_f = 0$ ), los resultados del modelo son tan precisos como la media de los datos observados. Mientras que una eficiencia menor que cero

( $E_f < 0$ ) ocurre cuando la media de los datos originales es un mejor predictor que el modelo, es decir, cuando la varianza residual (el numerador de la segunda expresión a la derecha de la igualdad (2.35)), es mayor que la varianza de datos (descrita por el denominador) [71].

La mayor desventaja de la eficiencia de Nash-Sutcliffe es el hecho de que las diferencias entre los valores observados y predichos se calculan como cantidades al cuadrado. En consecuencia, los valores más grandes en una serie temporal se sobreestiman fuertemente mientras que los valores más bajos no tiene un gran impacto en el cálculo de este indicador [71, 72].

#### 2.4.4 Índice de Acuerdo de Willmott

El *Índice de Acuerdo de Willmott*, presentado por primera vez en 1981 [73], pretendía superar la insensibilidad de  $E_f$  y  $R^2$  a las diferencias entre las observaciones empíricas y las predicciones de medias y varianzas. Pero al igual que  $E_f$ , este índice también es muy sensible a valores grandes e insensible a valores bajos [70, 74]:

$$d = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (|\hat{y}_i - E[y_i]| + |y_i - E[y_i]|)^2} . \quad (2.37)$$

Así que en 1985 Willmott presentó una nueva versión de este criterio [75],

$$d_1 = 1 - \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{\sum_{i=1}^N (|\hat{y}_i - E[y_i]| + |y_i - E[y_i]|)} , \quad (2.38)$$

cuyo valor oscila entre 0 y 1, donde 1 implica un ajuste perfecto y 0 representa un ajuste muy pobre. Este indicador, al utilizar valores absolutos, no potencializa los residuos más grandes producto de valores extremos como en el caso de las igualdades (2.35) y (2.37) [74]. También es mucho menos sensible a la forma de la distribución de frecuencia de error y, como consecuencia, a errores concentrados en valores atípicos [70].

## 3 METODOLOGÍA

En este capítulo detallamos el procedimiento para pronosticar series temporales sísmicas. Empezamos con datos empíricos procedentes de un sismógrafo. Esta información no está diseñada para su análisis, con lo cual es necesario convertir las series a un formato adecuado para su estudio mediante un procedimiento que resulta no trivial. A continuación, las series temporales sísmicas se van a dividir en dos partes iguales. Con la primera mitad de las series, y empleando la técnica de MLE, determinaremos el valor de los parámetros que ajustan el Proceso *ARMA* a dichas series. Para esto, usamos un algoritmo que maximice la función log likelihood en base al vector de parámetros en el espacio  $(p + q + 1)$ -dimensional. En nuestro problema, el algoritmo se basa en la información del gradiente de la función objetivo para definir la dirección a seguir en la búsqueda del máximo. Posteriormente, con los parámetros optimizados, procedemos a pronosticar las series objeto de este estudio en base a la estructura del modelo *ARMA*( $p, q$ ). Por medio del AIC escogemos el modelo que mejor se ajuste a los datos según los parámetros  $p$  y  $q$ . Finalmente, para determinar si el proceso *ARMA* puede describir los sismogramas estudiados, comparamos la segunda mitad de las series temporales sísmicas con su respectiva predicción. Esta comparación la haremos por medio análisis espectral, el índice de eficiencia de Nash-Sutcliffe y el índice de acuerdo de Willmott.

### 3.1 TRATAMIENTO DE LOS DATOS

Las series temporales sísmicas que utilizamos en este trabajo provienen del sismógrafo *Trillium 120p*, ubicado en el cantón Rioverde de la provincia de Esmeraldas (latitud 1.07; longitud -79.39), propiedad del Instituto Geofísico de la Escuela Politécnica Nacional [76]. En vista de la remota localización y de la gran cantidad de información que registra este instrumento, el formato en el que se almacenan los datos es un formato especial asociado al aparato. Para poder analizar los sismogramas es necesario pasar los registros a formato de texto [77]. Los archivos originalmente se registran en formato MiniSEED, una versión reducida del formato SEED (del inglés Standard

for the Exchange of Earthquake Data) [78]. Estos deben ser convertidos previamente a un formato intermedio denominado SAC [79], para posteriormente pasarlos a archivos de texto.

El proceso de conversión se realiza mediante un software de código abierto llamado PASSCAL PASSOFT, desarrollado y distribuido por el consorcio IRIS [80]. Este software sirve como base para descargar datos sísmicos, evaluar su calidad y convertir el formato de los datos. PASSOFT está actualmente desarrollado, en el caso especial de Linux (con el que nosotros trabajamos), exclusivamente para Fedora 10 i386 (y versiones posteriores) [81]. También existe la posibilidad de trabajar con herramientas de conversión<sup>1</sup> o con el código fuente (para instalarlo en cualquier otra distribución de Linux). Para reproducir los resultados a partir de los mismos datos empíricos, a continuación describimos el arduo proceso de instalación al igual que el uso del software.

En la base de datos del consorcio IRIS [82], los paquetes están etiquetados de la siguiente forma: PASSOFT-MAJORVERSION(Año)-MINORVERSION(Día del año).i386.rpm. Para instalar PASSOFT desde el terminal ejecutamos el siguiente comando: `rpm -ivh -nodeps **Archivo.rpm**`. Dado que todas las librerías de PASSOFT dependen de otras librerías, es muy probable que se deba instalar paquetes adicionales para que el comando mencionado anteriormente funcione.

Para convertir ficheros en formato miniSEED a ficheros en formato SAC se utiliza el paquete `mseed2sac v.2.1`. [83]. Al instalar este archivo es recomendable mover la descarga a la carpeta `/usr/local` para evitar problemas cuando se actualice el software del sistema. A continuación, se ejecuta el siguiente comando: `make install makefile`. Para correr el programa se utiliza el comando: `mseed2sac [opciones] *archivo*`. Para este caso no es necesario utilizar [opciones], pero éstas se especifican en [84].

Finalmente, para pasar del formato SAC a formato texto, es necesario instalar un software capaz de tratar con ficheros en formato SAC (del inglés Seismic Analysis Code), que lleva el mismo nombre. Por causa de las restricciones de la empresa, se debe llenar una solicitud que el personal de IRIS debe revisar [85]. Este proceso puede llevar varias semanas.

Desde el terminal, los pasos para instalar este programa son:

1. Cambie la dirección a: `sac-101.6a-linux_x86_64.tar.gz`
2. Para crear la carpeta `sac101.6a`: `sudo tar -zxvf sac-101.6a-linux_x86_64.tar.gz -C /usr/local/`
3. Coloque en la carpeta: `cd sac11.6a`

<sup>1</sup>En el caso de Ubuntu 'alien'.

4. Actualice sus librerías (opcional):

```
> sudo apt-get update
> sudo apt-get upgrade
> sudo apt-get install ncurses-dev
> sudo apt-get install libx11-dev
```

5. Construya e instale el programa:

```
> chmod +x configure
> chmod +x config/install-sh
> ./configure
> make
> sudo make install
```

Para arrancar el programa de conversión en el terminal, escriba `sac`, lo que iniciará la interfaz del programa [79]. A continuación, para transformar a formato de texto se usa:

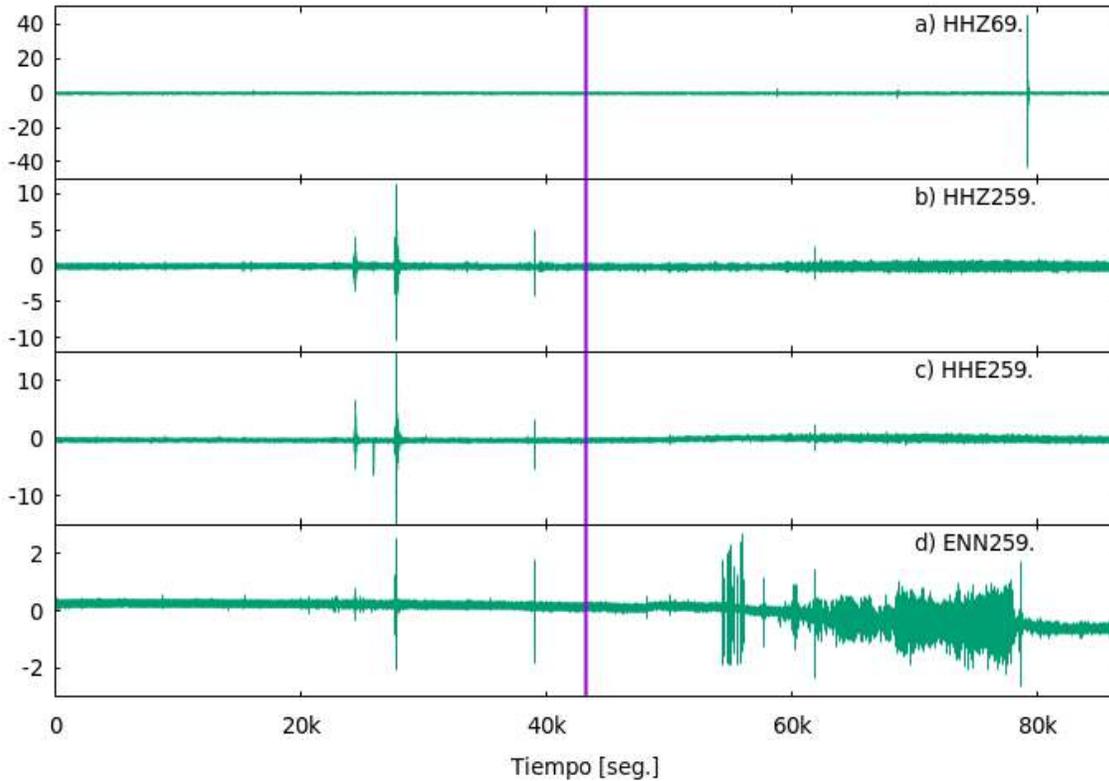
```
SAC> read binary <archivo.sac>
SAC> write alpha <archivo.dat>
```

El sismógrafo registra cantidades que, en promedio, son del orden de  $10^4$ . Con lo cual, al realizar operaciones con estas cifras, las cantidades obtenidas exceden la cantidad de bits del programa que las maneja. Por esta razón, para emplear variables de doble precisión en el análisis, es recomendable reescalar todo el sismograma multiplicándolo por  $10^{-4}$  para que dichas cantidades se encuentren en un rango de 0 a 1.

## 3.2 AJUSTE DE LOS PARÁMETROS DEL PROCESO ARMA

Al utilizar el proceso *ARMA* para pronosticar series temporales sísmicas (ya convertidas en ficheros de texto), se debe garantizar que estas series sean estacionarias (Ec. 1.5). Para determinar hasta qué punto un proceso *ARMA* puede modelar adecuadamente sismogramas, en este estudio utilizaremos 4 series temporales sísmicas (Fig. 3.1) con propiedades estadísticas diferentes (referente al proceso estacionario).

El proceso *ARMA* define una estructura matemática cuyos parámetros se deben ajustar para tratar de describir una serie temporal dada. La técnica empleada para ajustar estos parámetros es



**Figura 3.1:** Series temporales sísmicas originales de 10,798,990 datos (125 datos por segundo) cada una. También, una línea horizontal en el instante 43,195.96 (5,399,495 datos) que divide estas series en dos partes iguales. La primera se utiliza para ajustar los parámetros de proceso *ARMA* y la segunda para verificar los resultados del mismo. Las series se las trata adimensionalmente ya que las unidades son irrelevantes para este análisis.

la llamada "*estimación de máxima verosimilitud*" (Sección 2.2). En este planteamiento empleamos la primera mitad de cada serie temporal de la figura (3.1) para maximizar la función log likelihood en base a dichos parámetros.

El algoritmo que utilizamos para optimizar la función log likelihood, producido según la teoría de Davidon-Fletcher-Powell (Sección 2.3), es una adaptación del código descrito en la Sección 10.7 del libro *Numerical Recipes* [55]. Este proceso está compuesto por las subrutinas MINDFP (Cod. 7.1) y LNSRCH (Cod. 7.2). Este es un algoritmo de minimización, por lo que al minimizar la función  $-f$  se maximiza  $f$ , donde  $f$  será nuestra función likelihood [51]. Por ello, la función de doble precisión FUNC (Cod. 7.3) calcula el valor numérico de menos una vez la función log likelihood (Ec. 2.25), es decir:

$$-\mathcal{L}(\Theta) = \frac{T-p}{2} \ln(2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{t=p+1}^T \epsilon_t^2. \quad (3.1)$$

Para calcular la sumatoria del segundo término del lado derecho de la igualdad anterior, em-

### 3.2. AJUSTE DE LOS PARÁMETROS DEL PROCESO ARMA

---

pleamos la subrutina ERRVECT (Cod. 7.4). Esta calcula de forma iterativa el valor de estos errores en base a la ecuación (2.20),

$$\epsilon_t = y_t - c - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots - \phi_p y_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q},$$

y los guarda en un vector (ya que cada  $\epsilon_t$  depende de  $q$  errores anteriores). Por el planteamiento de Box y Jenkins [37], los primeros  $p$  valores de la familia de errores,  $\{\epsilon_i\}_{i=1}^T$ , son iguales a cero y las variables aleatorias,  $\{Y_t\}_{t=1}^T$ , toman el valor de la serie temporal sísmica a estudiar.

La función (3.1) se puede simplificar a la siguiente expresión:

$$\mathcal{L}(\Theta) = \sum_{t=p+1}^T \epsilon_t^2. \quad (3.2)$$

Esto se debe a que, como plantea el AIC (Sección 2.4.2), el modelo más simple con las mismas implicaciones debe ser escogido. Es así como MLE se reduce a mínimos cuadrados si los residuos se distribuyen de forma gaussiana y  $\sigma^2$  se determina a partir de la distribución de residuos con los demás parámetros optimizados.

El algoritmo DFP utiliza la información suministrada por el gradiente para determinar, en cada iteración, la dirección (Ec. 2.33) en la que se encuentra el punto que maximiza la función log likelihood. Por consiguiente, la subrutina DFUNC (Cod. 7.5), basándose en los resultados de la subrutina ERRVECT, calcula el gradiente de la ecuación (3.2), cuya forma general es:

$$g_i(\Theta) = \frac{\partial \mathcal{L}(\Theta)}{\partial x_i} = 2 \sum_{t=p+1}^T \epsilon_t \frac{\partial \epsilon_t}{\partial x_i}, \quad (3.3)$$

donde  $x_i$  es el  $i$ -ésimo elemento del vector de parámetros  $\Theta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, c)'$  y, basándose en la igualdad (2.20),  $\epsilon_t$  tiene la forma:

$$\epsilon_t = y_t - c - \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \theta_j \epsilon_{t-j}. \quad (3.4)$$

En la ecuación (3.3) se aprecia como los elementos del vector gradiente  $g_i(\Theta)$ , difieren entre sí

### 3.2. AJUSTE DE LOS PARÁMETROS DEL PROCESO ARMA

---

debido a la derivada parcial de los residuos (Ec. 3.4) en función de cada parámetro del vector  $\Theta$ :

$$\frac{\partial \epsilon_t}{\partial c} = -1 - \sum_{j=1}^q \theta_j \frac{\partial \epsilon_{t-j}}{\partial c} \quad (3.5)$$

$$\frac{\partial \epsilon_t}{\partial \phi_i} = -y_{t-i} - \sum_{j=1}^q \theta_j \frac{\partial \epsilon_{t-j}}{\partial \phi_i} \quad i = 1, 2, \dots, p \quad (3.6)$$

$$\frac{\partial \epsilon_t}{\partial \theta_k} = -\epsilon_{t-k} - \sum_{j=1}^q \theta_j \frac{\partial \epsilon_{t-j}}{\partial \theta_k} \quad k = 1, 2, \dots, q \quad (3.7)$$

El algoritmo determina todas las  $p + q + 1$  derivadas parciales referentes a las tres ecuaciones anteriores de forma similar: El cálculo empieza en el instante  $p + 1$ , esto se debe a que los errores anteriores a este instante son iguales a cero (Ec. 2.22). Con lo cual, usando el resultado del instante  $p + 1$  se calcula la derivada del instante  $p + 2$ . Asimismo, con estos dos resultados se calcula la derivada parcial en el instante  $p + 3$ . Y de forma general, la derivada parcial en el instante  $t$  se calcula con las  $q$  derivadas anteriores a este instante.

Cada derivada parcial depende de las derivadas de instantes anteriores. Debido a lo cual, todos los resultados se guardan en un vector, para el caso de las derivadas referentes a la constante  $c$ , y en matrices tanto para los  $p$  parámetros auto-regresivos,

$$DPH = \begin{bmatrix} \frac{\partial \epsilon_{p+1}}{\partial \phi_1} & \frac{\partial \epsilon_{p+1}}{\partial \phi_2} & \cdots & \frac{\partial \epsilon_{p+1}}{\partial \phi_p} \\ \frac{\partial \epsilon_{p+2}}{\partial \phi_1} & \frac{\partial \epsilon_{p+2}}{\partial \phi_2} & \cdots & \frac{\partial \epsilon_{p+2}}{\partial \phi_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \epsilon_T}{\partial \phi_1} & \frac{\partial \epsilon_T}{\partial \phi_2} & \cdots & \frac{\partial \epsilon_T}{\partial \phi_p} \end{bmatrix},$$

como para los  $q$  parámetros de media móvil,

$$DTH = \begin{bmatrix} \frac{\partial \epsilon_{p+1}}{\partial \theta_1} & \frac{\partial \epsilon_{p+1}}{\partial \theta_2} & \cdots & \frac{\partial \epsilon_{p+1}}{\partial \theta_q} \\ \frac{\partial \epsilon_{p+2}}{\partial \theta_1} & \frac{\partial \epsilon_{p+2}}{\partial \theta_2} & \cdots & \frac{\partial \epsilon_{p+2}}{\partial \theta_q} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \epsilon_T}{\partial \theta_1} & \frac{\partial \epsilon_T}{\partial \theta_2} & \cdots & \frac{\partial \epsilon_T}{\partial \theta_q} \end{bmatrix}. \quad (3.8)$$

Si se calcula explícitamente los elementos de la matriz anterior se obtiene el siguiente resultado

(Sección 7.1):

$$DTH = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ \frac{\partial \epsilon_{p+2}}{\partial \theta_1} & 0 & 0 & 0 & \dots & 0 \\ \frac{\partial \epsilon_{p+3}}{\partial \theta_1} & \frac{\partial \epsilon_{p+2}}{\partial \theta_1} & 0 & 0 & \dots & 0 \\ \frac{\partial \epsilon_{p+4}}{\partial \theta_1} & \frac{\partial \epsilon_{p+3}}{\partial \theta_1} & \frac{\partial \epsilon_{p+2}}{\partial \theta_1} & 0 & \dots & 0 \\ \frac{\partial \epsilon_{p+5}}{\partial \theta_1} & \frac{\partial \epsilon_{p+4}}{\partial \theta_1} & \frac{\partial \epsilon_{p+3}}{\partial \theta_1} & \frac{\partial \epsilon_{p+2}}{\partial \theta_1} & \dots & 0 \\ \frac{\partial \epsilon_{p+6}}{\partial \theta_1} & \frac{\partial \epsilon_{p+5}}{\partial \theta_1} & \frac{\partial \epsilon_{p+4}}{\partial \theta_1} & \frac{\partial \epsilon_{p+3}}{\partial \theta_1} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \epsilon_{q+1}}{\partial \theta_1} & \frac{\partial \epsilon_q}{\partial \theta_1} & \frac{\partial \epsilon_{q-1}}{\partial \theta_1} & \frac{\partial \epsilon_{q-2}}{\partial \theta_1} & \dots & \frac{\partial \epsilon_{p+2}}{\partial \theta_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \epsilon_T}{\partial \theta_1} & \frac{\partial \epsilon_{T-1}}{\partial \theta_1} & \frac{\partial \epsilon_{T-2}}{\partial \theta_1} & \frac{\partial \epsilon_{T-3}}{\partial \theta_1} & \dots & \frac{\partial \epsilon_{T-q}}{\partial \theta_1} \end{bmatrix}. \quad (3.9)$$

Por lo que el algoritmo puede utilizar solo la primera columna de la matriz anterior para determinar las derivadas parciales referentes a todos los  $q$  parámetros de media móvil:

$$DTH = \left[ \frac{\partial \epsilon_{p+2}}{\partial \theta_1}, \frac{\partial \epsilon_{p+3}}{\partial \theta_1}, \frac{\partial \epsilon_{p+4}}{\partial \theta_1}, \dots, \frac{\partial \epsilon_T}{\partial \theta_1} \right]'$$

### 3.3 CONSTRUCCIÓN DEL PROCESO ARMA Y PRONÓSTICOS DE SISMOGRAMAS

Con el fin de determinar si el proceso  $ARMA(p, q)$  describe apropiadamente un sismograma, los parámetros que caracterizan este proceso se van a ajustar, para cada serie temporal (Fig. 3.1), con diferentes valores de los parámetros  $p$  y  $q$ . Cada parámetro tomará valores enteros entre 1 y 7 para un total de 49 simulaciones cubriendo todas las combinaciones entre este par de parámetros. En consecuencia, contando a la constante  $c$ , el vector de parámetros,  $\Theta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, c)'$ , tendrá desde 3 componentes (cuando  $p$  y  $q$  son iguales a 1), hasta 15 componentes (cuando  $p$  y  $q$  son iguales a 7).

Los valores iniciales de las componentes del vector de parámetros con los que el algoritmo de optimización empieza la búsqueda del punto  $p + q + 1$ -dimensional que maximizan la función log likelihood son:

### 3.3. CONSTRUCCIÓN DEL PROCESO ARMA Y PRONÓSTICOS DE SISMOGRAMAS

Nº	Valor	Nº	Valor	Nº	Valor
1	1.268D-2	2	2.78D-3	3	1.718D-2
4	-1.8129D-3	5	1.6802D-2	6	-1.11D-2
7	-1.904D-2	8	1.268D-2	9	1.718D-2
10	-1.8129D-2	11	1.56802D-2	12	2.78D-2
13	-2.29D-2	14	5.6802D-3	15	7.18D0-3

**Tabla 3.1:** Valores iniciales del vector de parámetros  $\Theta$  en notación de doble precisión en Fortran 90.

Para el caso  $ARMA(1, 1)$  el vector de parámetros utilizará los tres primeros valores de la tabla (3.1):

$$\Theta_3 = \left( \underbrace{x_1}_{\phi}, \underbrace{x_2}_{\theta}, \underbrace{x_3}_c \right)',$$

donde  $x_i$  representa uno de los 15 valores de dicha tabla ordenados según el subíndice  $i$ . A continuación, para el caso  $ARMA(1, 2)$ :

$$\Theta_4 = \left( \underbrace{x_1}_{\phi}, \underbrace{x_2}_{\theta_1}, \underbrace{x_3}_{\theta_2}, \underbrace{x_4}_c \right)',$$

y en general, para cualquier combinación de  $p$  y  $q$  se utiliza los valores iniciales según el orden establecido hasta llegar al caso donde se ocupan todos estos valores:

$$\Theta_{p+q+1} = \left( \underbrace{x_1}_{\phi_1}, \underbrace{x_2}_{\phi_2}, \dots, \underbrace{x_p}_{\phi_p}, \underbrace{x_{p+1}}_{\theta_1}, \underbrace{x_{p+2}}_{\theta_2}, \dots, \underbrace{x_{p+q}}_{\theta_q}, \underbrace{x_{p+q+1}}_c \right)'$$

A causa de la gran cantidad de información que se va a procesar, optamos por usar el supercomputador 'QUINDE' propiedad del Laboratorio de Cálculo Científico 'MODEMAT' [86]. Para administrar los recursos de este supercomputador, al correr las simulaciones, es necesario un script.slurm (Cod.7.6) cuyo funcionamiento se detalla en [87].

Con el punto en el espacio de parámetros que maximiza la función log likelihood, empleando la subrutina ERRVECT, generamos un conjunto de errores aleatorios según la igualdad (2.20). Así, con estos errores podemos determinar la varianza que maximiza la función log likelihood original (2.25). A continuación, con la subrutina NDRN (Cod. 7.7) generamos ruido blanco con distribución gaussiana a partir de la media y varianza que maximizan la función (2.25). Finalmente, con la subrutina TSERIES (Cod. 7.8) de forma iterativa procedemos a pronosticar la serie temporal sísmica

### 3.3. CONSTRUCCIÓN DEL PROCESO ARMA Y PRONÓSTICOS DE SISMOGRAMAS

---

mediante el modelo  $ARMA(p, q)$  (Ec. 2.14),

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} ,$$

donde  $\{\epsilon_i\}_{i \in \mathbb{Z}^+}$  es el ruido blanco generado por la subrutina NDRN y los primeros  $p$  valores de esta nueva serie  $(Y_1, Y_2, \dots, Y_p)$  son iguales a los últimos  $p$  valores de la primera mitad de las series temporales a predecir.

De los 49 procesos  $ARMA(p, q)$ , con diferentes valores de  $p$  y  $q$ , el modelo que mejor se ajusta a cada sismograma estudiado se determina mediante el Criterio de información de Akaike (Sección 2.4.2). Es decir, calculamos este indicador (Ec. 2.34) para cada proceso y aquel que tenga el menor valor numérico se considera el modelo más cercano a la realidad descrita por los datos del sismograma.

Para validar las series temporales producto del modelo  $ARMA(p, q)$ , utilizamos la segunda parte de los sismogramas originales de la figura (3.1). Estos sismogramas se van a comparar con sus respectivos pronósticos tanto en el dominio del tiempo como en el dominio de la frecuencia. Es así como la implementación de la transformación de Fourier (Sección 2.4.1) de estas series la realizaremos mediante el algoritmo de la Transformada Rápida de Fourier (o FFT del inglés Fast Fourier transform). El código de este procedimiento está descrito en la sección 12.3 del libro Numerical Recipes [55]. Asimismo, empleamos el Coeficiente de Eficiencia del Modelo Nash-Sutcliffe (Sección 2.4.3) y el Índice de Acuerdo de Willmott (Sección 2.4.4) para determinar la bondad de los ajustes realizados por el proceso  $ARMA$ .

## 4 RESULTADOS Y DISCUSIÓN

En este capítulo utilizamos el Criterio de Información de Akaike para seleccionar el mejor modelo *ARMA* de entre los propuestos. Asimismo, analizamos si los resultados presentados por este proceso capturan las características más importantes para el estudio de series temporales sísmicas.

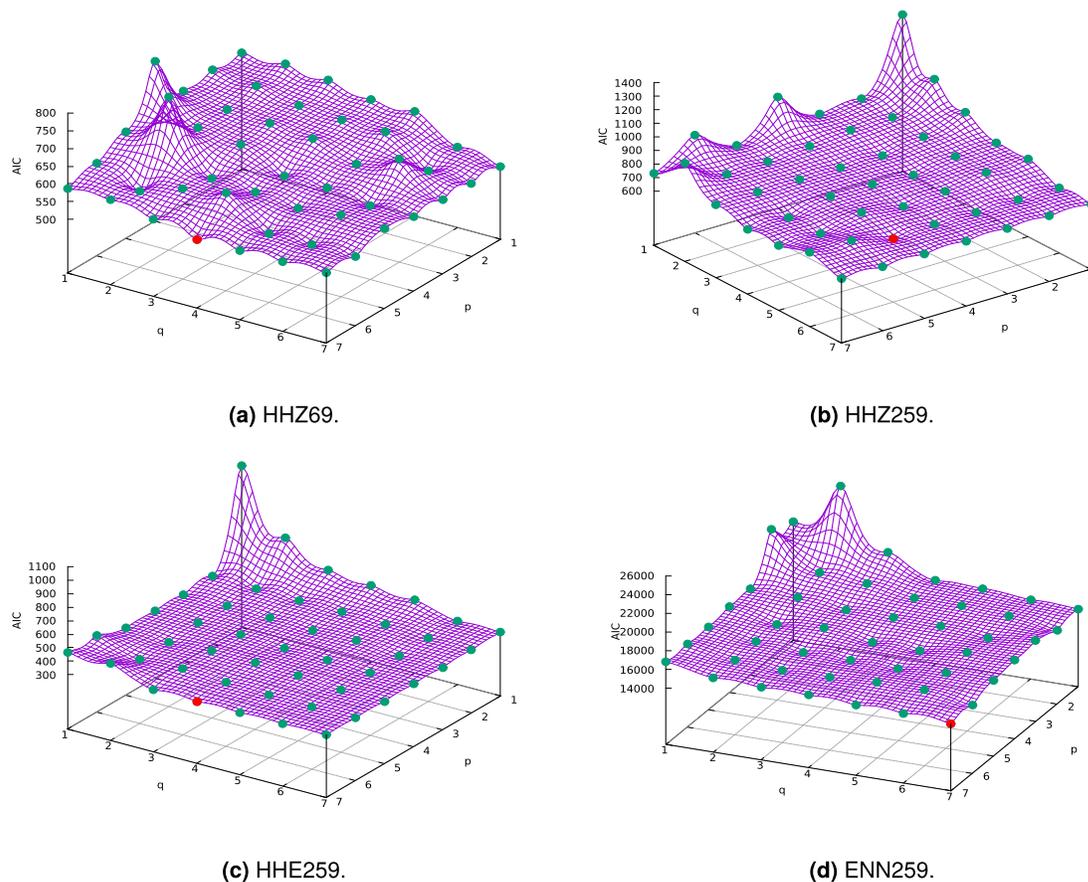
### 4.1 VALIDACIÓN DEL MODELO ARMA

Para cada serie temporal analizada (Fig. 3.1) realizamos 49 predicciones con valores enteros entre 1 y 7 para los parámetros  $p$  y  $q$ . Cada una de estas 49 predicciones representa un modelo diferente de la familia de procesos  $ARMA(p, q)$ . El modelo, de entre los 49 propuestos, que mejor ajusta cada sismograma se puede determinar empleando el Criterio de Información de Akaike (Ec. 2.34). Donde el número de parámetros es  $p+q+1$  y la subrutina FUNC calcula el valor negativo de la función log likelihood (Ec. 3.2), con lo cual, se debe multiplicar por -1 en el cálculo de AIC:

$$\begin{aligned} AIC &= -2 \ln (f (\Theta_{MLE})) + 2K \\ &= -2(-\mathcal{L} (\Theta_{MLE})) + 2(p + q + 1) \end{aligned}$$

Los resultados de aplicar este indicador a las series temporales sísmicas se presentan en la figura (4.1). El índice AIC más bajo para las figuras **(a)**, **(b)**, **(c)** y **(d)** corresponde a los procesos  $ARMA(7, 4)$ ,  $ARMA(5, 6)$ ,  $ARMA(7, 4)$  y  $ARMA(7, 7)$  respectivamente. Cabe destacar que el hecho de que el mismo proceso  $ARMA(7, 4)$  describa tanto a la serie HHZ69 como a la serie HHE259, es mera coincidencia y no se debe a relación alguna entre estas dos series. Por otro lado, las figuras (4.1b) y (4.1c) presentan índices AIC con valores similares. Esto se debe a que los datos empíricos de cada serie corresponden a componentes de un mismo sismograma.

## 4.1. VALIDACIÓN DEL MODELO ARMA

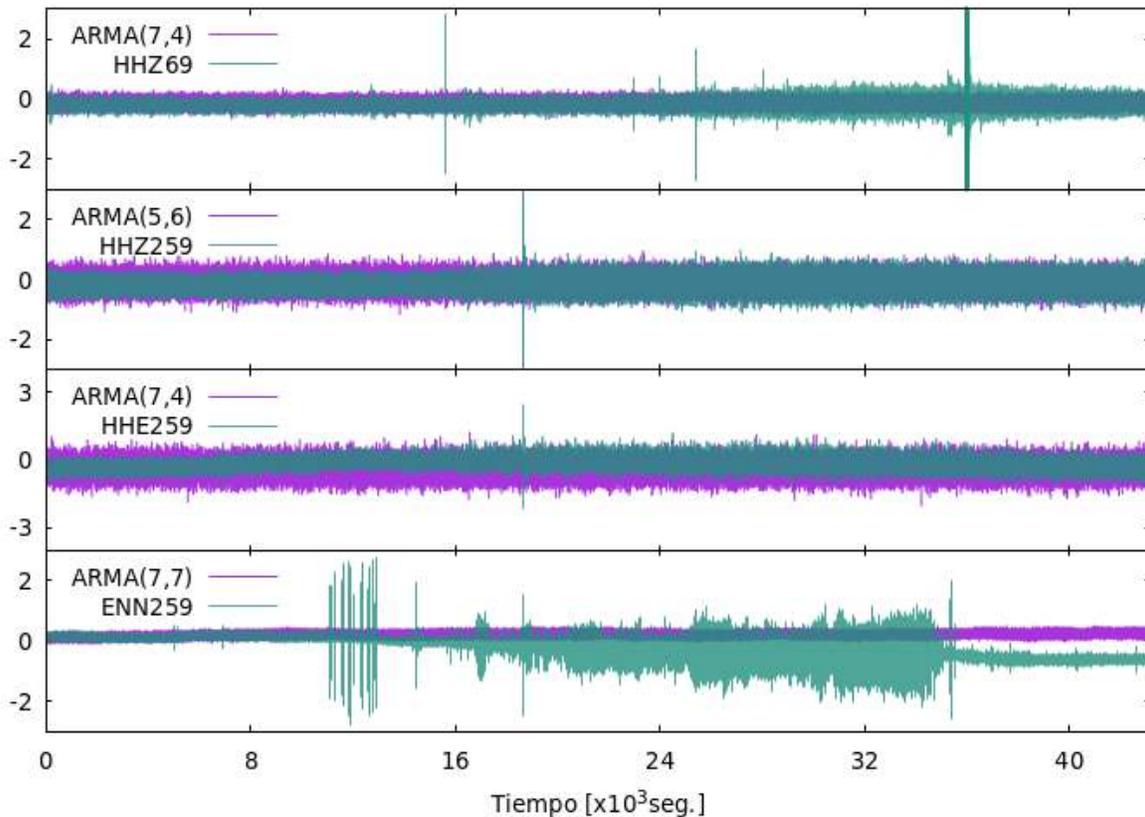


**Figura 4.1:** Dependencia del AIC a los parámetros  $p$  y  $q$  (toman valores enteros entre 1 y 7) para las cuatro series estudiadas. El punto de color rojo representa el AIC de menor valor y la superficie de color morado es una extrapolación gaussiana de los datos que se utiliza para visualizar de mejor manera la localización de los puntos en el espacio representado en cada figura.

Con la primera mitad de las series temporales estudiadas (parte de ajuste), se determinó el valor de los parámetros que ajustan los procesos *ARMA* (Fig. 3.1). Ahora, para comprobar la efectividad de los pronósticos realizados por estos modelos, utilizamos la segunda mitad de estas series temporales (parte de verificación). En consecuencia, la figura (4.2) presenta a la segunda mitad de los sismogramas originales y las series producidas mediante el proceso *ARMA*. Para todos los casos, no observamos un comportamiento exactamente igual entre el pronóstico y el sismograma original. El modelo *ARMA* supone que tanto la media como la autocovarianza permanecen constantes. Pero los sismogramas de interés, al presentar eventos sísmicos como terremotos, necesariamente no son estacionarios (Tabla 4.1).

La serie de la figura (3.1a), registrada el 03 de marzo de 2009, es la representación del eje  $z$  (perpendicular a la superficie terrestre) de un velocímetro. Fijémonos que esta es una serie aproximadamente estacionaria que presenta un aumento de la variación -ensanchamiento- al final de la misma y exhibe eventos sísmicos aislados (representados como el súbito aumento de la am-

#### 4.1. VALIDACIÓN DEL MODELO ARMA



**Figura 4.2:** Comparación de la segunda mitad de las cuatro series temporales estudiadas con el respectivo proceso *ARMA* que pretende describir dichas series.

plitud en dicha serie). De igual modo, las series de las figuras (3.1b) y (3.1c) (registradas el 16 de septiembre de 2009) son componentes del mismo sismograma en el eje  $z$  y en el eje  $x$  respectivamente. En la serie HHZ259 existen varios sismos (de diferentes intensidades en instantes aleatorios) y un aumento de la variación al final de la misma, lo cual implica que su varianza no necesariamente es constante. A la par, en la serie HHE259 la varianza aumenta de manera similar a la serie anterior y, adicionalmente, presenta una media variable con una desviación a partir de los 50,000 segundos. Finalmente, la figura (3.1d) es la componente en el eje  $y$  (dirección Sur-Norte) de un velocímetro registrado el mismo día que los dos casos anteriores. Esta serie presenta una media que decae con el transcurso del tiempo. También, una varianza que aumenta a partir de los 60,000 segundos. Claramente, este sismograma no es estacionario.

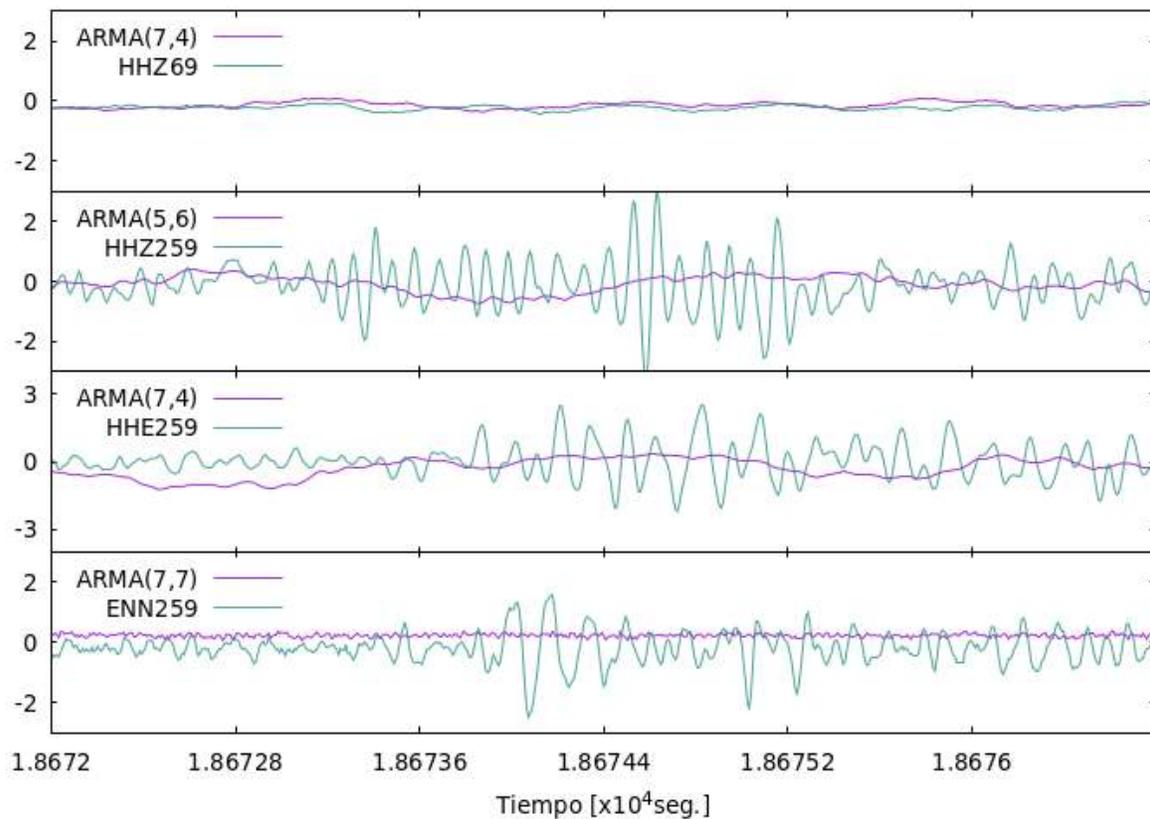
La serie producto del proceso *ARMA*(7,4) no presenta un aumento de la variación, y menos aún, los sismos que aparecen en la serie HHZ69 que pretende pronosticar (considerada como "*la más estacionaria*"). De igual modo, las series producto de los procesos *ARMA*(5,6) y *ARMA*(7,4) no pueden reproducir la evolución de las series HHZ259 y HHE259 respectiva-

#### 4.1. VALIDACIÓN DEL MODELO ARMA

Series		Ajuste	ARMA	Verificación
HHZ69	$\mu$	$-0.132 \pm 3E-4$	$-0.132 \pm 5E-4$	$-0.1799 \pm 4E-4$
	$\sigma$	$0.1153 \pm 7E-4$	$0.1156 \pm 5E-4$	$0.3894 \pm 9E-4$
HHZ259	$\mu$	$-9.58E-2 \pm 7E-4$	$-9.62E-2 \pm 4E-4$	$-0.1433 \pm 4E-4$
	$\sigma$	$0.2188 \pm 5E-4$	$0.2184 \pm 4E-4$	$0.1894 \pm 4E-4$
HHE259	$\mu$	$-0.352 \pm 6E-3$	$-0.355 \pm 3E-3$	$-0.108 \pm 2E-3$
	$\sigma$	$0.3206 \pm 7E-4$	$0.3201 \pm 9E-4$	$0.2432 \pm 8E-4$
ENN259	$\mu$	$0.222 \pm 9E-3$	$0.209 \pm 8E-3$	$-0.222 \pm 5E-3$
	$\sigma$	$6.04E-2 \pm 5E-4$	$6.39E-2 \pm 5E-4$	$0.2906 \pm 4E-4$

**Tabla 4.1:** Media y desviación estándar de la serie producto del Procesos *ARMA*, la parte de ajuste y la parte de verificación de los cuatro sismogramas estudiados.

mente. Por último, la serie producto del modelo *ARMA(7,7)* presenta la varianza de menor valor (Tabla 4.1). Pero debido a las limitaciones del proceso *ARMA* y al hecho de que la serie experimental ENN259 es la que más variaciones producto de eventos sísmicos tiene, la discrepancia entre estas dos series es la mayor de todos los casos estudiados.

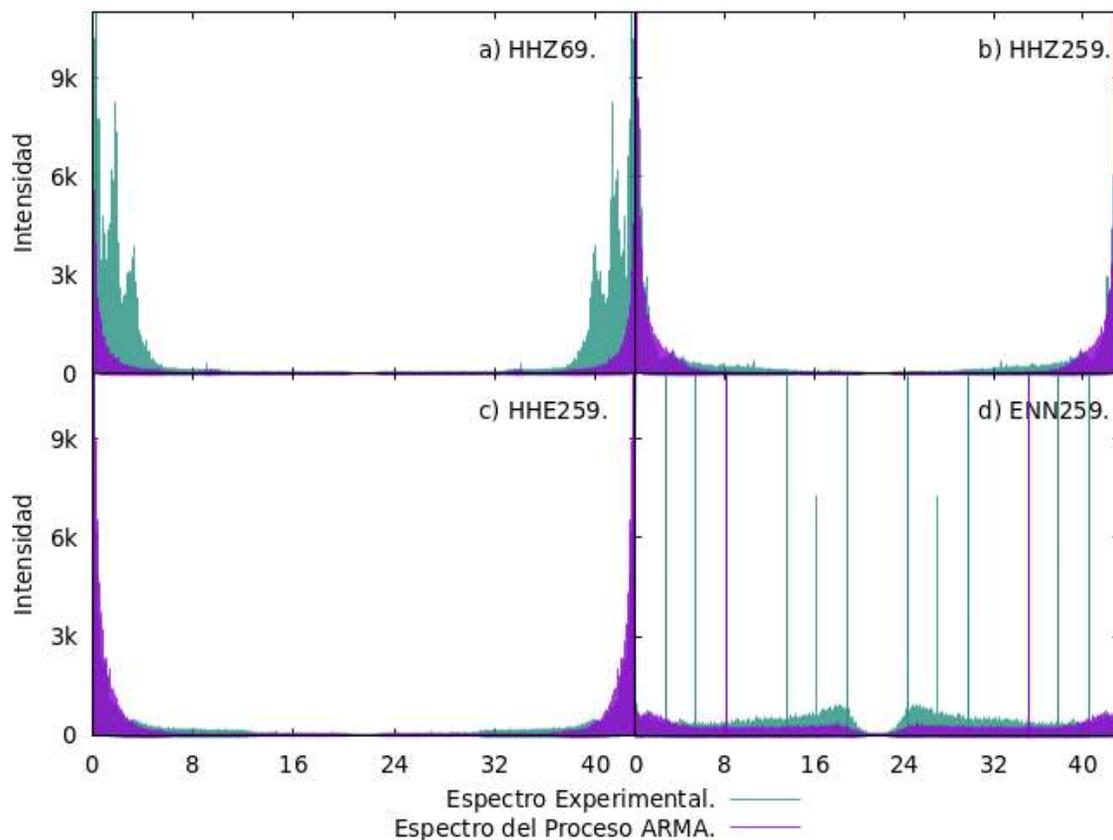


**Figura 4.3:** Ampliación de aproximadamente 5 segundos de las series de la gráfica (4.2). La figura superior registra ruido mientras que las tres últimas figuras registran un sismo.

Al ampliar cada gráfica en el instante donde se registra un sismo (a excepción de la serie HHZ69 que funciona como control) observamos el comportamiento de estas series más a detalle. Para

#### 4.1. VALIDACIÓN DEL MODELO ARMA

cada caso, apreciamos como a pesar de que no coinciden exactamente los procesos *ARMA* con los datos empíricos, sí existe cierta similitud en la forma que tiene cada serie con su predicción (Fig. 4.3). A notar, la predicción del proceso *ARMA*(7,4) presenta una frecuencia muy similar a la de los datos empíricos de la serie HHZ69. El valor de las predicciones de las series HHZ259 y HHE259 crecen y decrecen conforme las series originales tienen sus picos en la parte superior o inferior a la media. La serie EEN259, debido al sismo, presenta una forma muy accidentada que el modelo *ARMA*(7,7) también presenta a pesar de que esta serie no es estacionaria.



**Figura 4.4:** Sismogramas y pronósticos del proceso ARMA en el espacio de frecuencia con unidades de  $[x10^3Hz]$ .

Posiblemente existen relaciones implícitas entre las series experimentales y las series producto del modelo *ARMA* que no son evidentes en el dominio temporal. Con lo cual, empleamos la transformación de Fourier para descomponer las series temporales en sus diversas componentes dentro del dominio frecuencial. Así pues, al comparar el espectro de los sismogramas con sus respectivos pronósticos (Fig. 4.4), observamos que no existen frecuencias en las que estas series coincidan exactamente: El espectro de los sismogramas HHZ69 y ENN259, presentan un mayor número de picos y, en general son mucho más complejos que los espectros de los pronósticos

#### 4.1. VALIDACIÓN DEL MODELO ARMA

---

del modelo  $ARMA(7, 4)$  y  $ARMA(7, 7)$  respectivamente. Al contrario, a pesar de la complicada naturaleza de los sismogramas, los espectros de las series HHZ259 y HHE259 presentan formas similares a los espectros de sus respectivos pronósticos. En cada gráfica observamos picos pronunciados en los extremos de las figuras que alcanzan intensidades semejantes.

Para establecer cuantitativamente la bondad del ajuste del proceso  $ARMA$ , en comparación con los datos empíricos, utilizamos el coeficiente de eficiencia Nash-Sutcliffe (Ec. 2.35) y el índice de acuerdo de Willmott (Ec. 2.38). El coeficiente NSE ( $E_f$ ) toma valores menores o iguales a uno. Uno, para un ajuste perfecto; cero cuando los pronósticos son tan precisos como la media de los datos empíricos y valores negativos cuando el valor medio de los datos es un mejor predictor que el modelo analizado. De forma similar, el índice de acuerdo de Willmott ( $d_1$ ) toma valores que oscilan entre 0 y 1. Donde uno implica una coincidencia perfecta entre el modelo y los datos observados, mientras que cero representa un ajuste muy pobre por parte del modelo.

	$E_f$	$d_1$
<b>HHZ69</b>	-0.1033	0.2653
<b>HHZ259</b>	-1.3841	0.2865
<b>HHE259</b>	-2.7668	0.2769
<b>ENN259</b>	-2.3897	0.3627

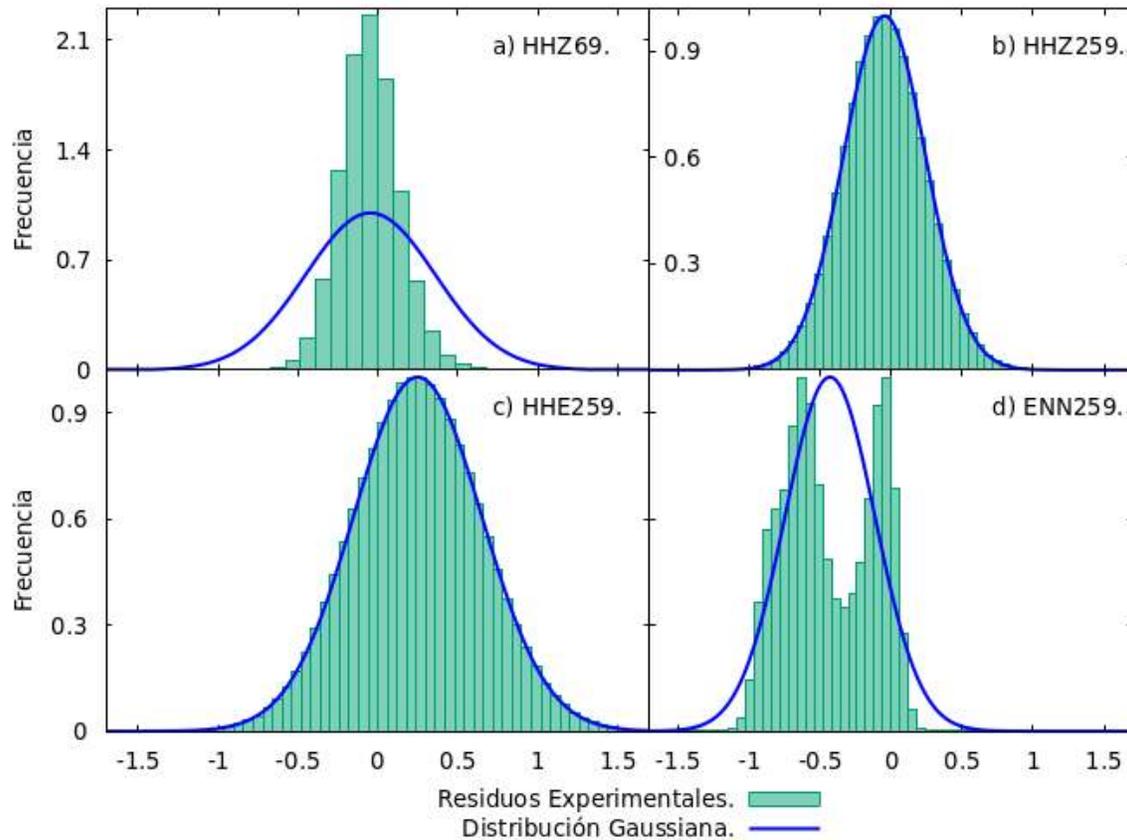
**Tabla 4.2:** Índices que establecen la bondad del modelo  $ARMA$  en comparación con los datos empíricos.

En la tabla (4.2) observamos que el índice NSE, para todas las series, presenta valores negativos. Teniendo en cuenta que un modelo es de calidad a partir de valores entre  $0.5 \leq NSE \leq 0.65$  [88], estos resultados sugieren que los modelos no describen aceptablemente la realidad representada en los sismogramas. Cabe mencionar que, como se aprecia en la tabla (4.2), el valor de NSE es descendiente desde la serie HHZ69 hasta la serie ENN529. Este resultado confirma la suposición de que las características estacionarias disminuyen acorde se cambia de serie temporal de manera descendiente en la figura (3.1).

Los resultados que obtuvimos al aplicar el índice de acuerdo de Willmott, al encontrarse más cercanos a cero que a uno, sugieren también que las predicciones realizadas por el proceso  $ARMA$  no describen apropiadamente series temporales sísmicas. Este índice, al utilizar valores absolutos, no potencializa los residuos más grandes producto de valores extremos como en el caso del indicador NSE -que los eleva al cuadrado-. Por esta razón, observamos que, contradictoriamente, el índice  $d_1$  crece conforme se analizan series menos estacionarias (orden descendiente de la tabla (4.2)).

Antes de concluir categóricamente cualquier resultado, debemos comprobar si las suposiciones

#### 4.1. VALIDACIÓN DEL MODELO ARMA



**Figura 4.5:** Comparación entre la distribución de residuos con su supuesta distribución de teórica. La distribución teórica se supone gaussiana según el planteamiento tradicional del Proceso *ARMA* para ajustar los parámetros que caracterizan este proceso.

de nuestro estudio han sido razonables. ¿están los residuos  $\epsilon_t$  distribuidos de forma gaussiana? En base a esta suposición determinamos en su momento cuál es la función log likelihood que se iba a optimizar para determinar los parámetros de cada proceso *ARMA* (Sección 2.2.2). En consecuencia, para comprobar si la distribución experimental es gaussiana, determinamos los residuos según la igualdad (2.36) y contamos la frecuencia del valor de estos términos para determinar su distribución. A continuación, determinamos la media y varianza de estos residuos experimentales. Con estos parámetros, calculamos la supuesta función likelihood gaussiana que deberían tener los residuos según la ecuación (1.1). Y finalmente, comparamos la distribución experimental con la teórica.

En la figura (4.5) se aprecia claramente que los residuos, entre la predicción de los procesos *ARMA* y la segunda mitad de los sismogramas originales, se distribuyen de forma diferente para cada serie. Para el caso de la serie HHZ69 observamos que la distribución experimental aparenta una forma gaussiana con una varianza menor a la varianza de la distribución teórica. Pero como

la distribución teórica se calculó con la media y la varianza de la distribución experimental, no es posible que la distribución experimental sea gaussiana. Por otro lado, la distribución experimental de las series temporales sísmicas HHZ259 y HHE259 es gaussiana. Dichas series coinciden exactamente con sus respectivas distribuciones teóricas. A continuación, notamos como la distribución de residuos de la serie ENN259 es bimodal y, por tanto, no concuerda con su supuesta distribución teórica.

Recapitulando, la tabla (4.1) muestra como el valor medio y la desviación estándar del proceso *ARMA* y la primera parte de los datos empíricos (parte de ajuste), en todos los casos, son muy semejantes. Más aún, en la figura (4.3) apreciamos como este proceso se desempeña de manera excelente dentro de su rango de acción. Esto demuestra como el método de estimación de máxima verosimilitud (Sección 2.2) determinó estos valores con gran precisión a pesar de que la distribución de residuos está bien caracterizada solo para la mitad de los casos (Fig. 4.4). Por otro lado, en la figura (4.2) está claro que la complejidad de los sismogramas rebasa la capacidad del proceso *ARMA* para predecir este tipo series temporales. Por lo que la correcta elección de la función log likelihood (Ec. 2.25) no garantiza mejores resultados. También, en la figura (4.4) apreciamos como en el dominio de la frecuencia los pronósticos del proceso *ARMA* no coinciden exactamente con los datos empíricos. Cabe notar que en las series donde la distribución experimental es gaussiana existe gran similitud entre los espectros teóricos y experimentales (Fig. 4.4b y 4.4c). A la par, la eficiencia de Nash-Sutcliffe y el índice de acuerdo de Willmott confirman cuantitativamente que este modelo no es ideal para pronosticar sismogramas (Tabla 4.2).

## 5 CONCLUSIONES

En este trabajo de titulación se analizó la posibilidad de describir sismogramas empleando el modelo estocástico paramétrico *ARMA* (en su formulación general que supone distribución de residuos gaussianas) y el método para calcular parámetros denominado Estimación de Máxima Verosimilitud. A partir de ajustar los parámetros que caracterizan la estructura matemática que el proceso *ARMA* establece, es posible pronosticar futuros valores de un sismograma y así indagar en la naturaleza de la sismicidad.

En base a las figuras (4.2) y (4.4) queda claro que, tanto en el dominio del tiempo como en el dominio de la frecuencia, los pronósticos del proceso *ARMA* no coinciden exactamente con los datos empíricos. A pesar de que la función log likelihood que ajusta los parámetros que caracterizan este modelo sea correcta, el coeficiente NSE y el índice de acuerdo de Willmott confirman cuantitativamente el pobre desempeño de este proceso.

El proceso *ARMA* supone que las series temporales sísmicas que analiza son estacionarias. Pero en este estudio fue evidente que los sismogramas de interés (que presentan eventos sísmicos) no son estacionarios. Con lo cual, el proceso *ARMA* se ve rebasado por la complejidad de los sismogramas, obteniendo como resultado predicciones poco precisas que no reproducen toda la información contenida en dichas series. Esto imposibilita el desarrollo de un posible modelo y la comprensión de este fenómeno natural empleando, únicamente, procesos *ARMA*.

Debido a la versatilidad de la teoría que respalda las técnicas empleadas, en futuros trabajos se puede aplicar un filtro que analice la veracidad de una determinada distribución de residuos. Además, es posible utilizar estructuras matemáticas más sofisticadas que se basan en el proceso *ARMA* como son el modelo no lineal auto-regresivo de media móvil (*NARMA*) y el modelo de media móvil integrado auto-regresivo (*ARIMA*). También se podrían intentar otros enfoques como el método generalizado de momentos, la transformación de Hilbert-Huang (más conocida como empirical mode decomposition), dimensión Fractal, entre otros.

## 6 REFERENCIAS BIBLIOGRÁFICAS

- [1] Francisco Rivadeneira y col. *Breves fundamentos sobre los terremotos en el Ecuador*. Ed. por Jorge Ortega. Instituto. Quito: Corporación Editora Nacional, 2007, pág. 101. URL: <https://www.igepon.edu.ec/publicaciones-para-la-comunidad/comunidad-espanol/35-breves-fundamentos-sobre-los-terremotos-en-el-ecuador/file>.
- [2] C. Frohlich. «Fundamentals of geophysics». En: *Eos, Transactions American Geophysical Union* 79.15 (1998), págs. 187-187. ISSN: 0096-3941. DOI: 10.1029/98eo00138.
- [3] *Seismology - Latest research and news | Nature*. URL: <https://www.nature.com/subjects/seismology> (visitado 10-01-2020).
- [4] Edward J. Tarbuck y col. *Ciencias de la tierra : una introducción a la geología física*. Pearson Educación, 2013. ISBN: 9788490353943.
- [5] USGS Earthquake Hazards Program. *The Science of Earthquakes*. URL: <https://earthquake.usgs.gov/learn/kids/eqscience.php> (visitado 02-04-2019).
- [6] Donald L. Turcotte y Bruce D. Malamud. «14 Earthquakes as a complex system». En: *International Geophysics*. Vol. 81. PART A. Academic Press, 2002. DOI: 10.1016/S0074-6142(02)80217-0.
- [7] Clifford Thurber. «An Introduction to Seismology, Earthquakes, and Earth Structure». En: *Eos, Transactions American Geophysical Union* 84.22 (2003), pág. 209. ISSN: 0096-3941. DOI: 10.1029/2003eo220007.
- [8] Hiroo Kanamori. «72 Earthquake prediction: An overview». En: *International Geophysics*. Vol. 81. PART B. Academic Press, 2003, págs. 1205-1216. DOI: 10.1016/S0074-6142(03)80186-9.
- [9] R Burridge, L Knopoff - Bulletin of the Seismological y undefined 1967. «Model and theoretical seismicity». En: *pubs.geoscienceworld.org* (). URL: <https://pubs.geoscienceworld.org/ssa/bssa/article-abstract/57/3/341/116471>.



- 
- [19] Subsecretaría General de Gestión de Riesgos. *PLAN NACIONAL DE RESPUESTA ANTE DESASTRES*. Inf. téc. Subsecretaría de Preparación y Respuesta ante Eventos Adversos, 2018. URL: [www.gestionderiesgos.gob.ec](http://www.gestionderiesgos.gob.ec).
- [20] Alexandra Alvarado y col. «Seismic, volcanic, and geodetic networks in Ecuador: building capacity for monitoring and research». En: *Seismological Research Letters* 89.2A (mar. de 2018), págs. 432-439. ISSN: 19382057. DOI: 10.1785/0220170229. URL: <https://pubs.geoscienceworld.org/ssa/srl/article/89/2A/432/528855/Seismic-Volcanic-and-Geodetic-Networks-in-Ecuador>.
- [21] Thomas C. Hanks e Hiroo Kanamori. «A moment magnitude scale». En: *Journal of Geophysical Research* 84.B5 (mayo de 1979), pág. 2348. ISSN: 0148-0227. DOI: 10.1029/JB084iB05p02348. URL: <http://doi.wiley.com/10.1029/JB084iB05p02348>.
- [22] Carlos Mendoza y James W Dewey. «Seismicity associated with the great Colombia-Ecuador earthquakes of 1942, 1958, and 1979: Implications for barrier models of earthquake rupture». En: *Bulletin of the Seismological Society of America* 74.2 (abr. de 1984), págs. 577-593. ISSN: 0037-1106.
- [23] Céline Beauval y col. «Comparison of observed ground-motion attenuation for the 16 April 2016 Mw 7.8 Ecuador Megathrust earthquake and its two largest aftershocks with existing ground-motion prediction equations». En: *Seismological Research Letters* 88.2 (mar. de 2017), págs. 287-299. ISSN: 19382057. DOI: 10.1785/0220160150.
- [24] A. Bedford y D. S. Drumheller. «Introduction to elastic wave propagation». En: *Introduction to elastic wave propagation* (1994). ISSN: 01410296. DOI: 10.1016/0141-0296(94)90093-0.
- [25] P. G. Appleby, R. J. Atkin y N. Fox. «An Introduction to the Theory of Elasticity». En: *The Mathematical Gazette* 65.433 (1981), pág. 230. ISSN: 00255572. DOI: 10.2307/3617161.
- [26] Pokapu Akoranga Putaiao. *Earth waves — Science Learning Hub*. 2007. URL: <https://www.sciencelearn.org.nz/images/353-earth-waves> (visitado 27-02-2020).
- [27] W. Jacquelyne Kious y Robert I. Tilling. «This dynamic earth: the story of plate tectonics». En: *General Interest Publication* (1996). URL: <https://pubs.er.usgs.gov/publication/7000097>.
- [28] Markus Båth. «Introduction to Seismology». En: *Geophysical Journal of the Royal Astronomical Society* 40.1 (1975), págs. 143-143. ISSN: 1365246X. DOI: 10.1111/j.1365-246X.1975.tb01612.x.
- [29] Edward J Tarbuck, Frederick K Lutgens y Dennis Tasa. *Earth Science*, pág. 787. ISBN: 9781292061313.

- 
- [30] IRIS. «Como Trabaja un Sismógrafo ?» En: *New York* 7 (2004), págs. 20005-20005.
- [31] W. H.K. Lee. «17 Challenges in observational seismology». En: *International Geophysics*. Vol. 81. PART A. Academic Press, 2002, págs. 269-281. DOI: 10.1016/S0074-6142(02)80220-0.
- [32] Hossein Pishro-Nik. *Introduction to probability, statistics, and random processes*, pág. 732. ISBN: 0990637204.
- [33] Javier Gorgas, Nicolás Cardiel y Jaime Zamorano. *Estadística Basica Para Estudiantes De Ciencias*. 2009, pág. 205. ISBN: 978-84-691-8981-8.
- [34] James D. (James Douglas) Hamilton. *Time series analysis*. Princeton University Press, 1994, pág. 799. ISBN: 9780691042893.
- [35] Anders Hald. *History of Probability and Statistics and Their Applications before 1750*. Wiley Series in Probability and Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., ene. de 1990. ISBN: 9780471725169. DOI: 10.1002/0471725161. URL: <http://doi.wiley.com/10.1002/0471725161>.
- [36] Holger Kantz y Thomas Schreiber. *Nonlinear Time Series Analysis*. 2nd. Vol. 47. 3. Cambridge university press, 2005, págs. 381-381. ISBN: 0521529026. DOI: 10.1198/tech.2005.s306. arXiv: 0911.0331. URL: <http://pubs.amstat.org/doi/abs/10.1198/tech.2005.s306>.
- [37] George E.P. Box, Gwilym M. Jenkins y Gregory C. Reinsel. *Time series analysis: Forecasting and control: Fourth edition*. Vol. 20. October 1984. Princeton University Press, 2013, págs. 1-746. ISBN: 9781118619193. DOI: 10.1002/9781118619193. arXiv: arXiv:1011.1669v3. URL: <https://press.princeton.edu/titles/5386.html>.
- [38] Terence C. Mills. *Applied time series analysis : a practical guide to modeling and forecasting*. ISBN: 0128131179.
- [39] P. J. Brockwell. «Time series analysis». En: *International Encyclopedia of Education*. Elsevier Ltd, ene. de 2010, págs. 474-481. ISBN: 9780080448947. DOI: 10.1016/B978-0-08-044894-7.01372-5.
- [40] James Rovnyak. «Ideals of Square Summable Power Series». En: *Proceedings of the American Mathematical Society* 13.3 (jun. de 1962), pág. 360. ISSN: 00029939. DOI: 10.2307/2034938.
- [41] Louis De Branges. *SQUARE SUMMABLE POWER SERIES*. Inf. téc.
- [42] Ionuț Florescu. *Probability and stochastic processes*. ISBN: 9780470624555.

- 
- [43] MW y Walter Enders. «Applied Econometric Time Series.» En: *Journal of the American Statistical Association* 90.431 (1995), pág. 1135. ISSN: 01621459. DOI: 10.2307/2291367.
- [44] Peter Whittle. *Hypothesis testing in time series analysis*. Uppsala: Almqvist & Wiksells boktr., 1951.
- [45] Zihua Zhang y John C. Moore. «Autoregressive Moving Average Models». En: *Mathematical and Physical Fundamentals of Climate Change*. Elsevier, ene. de 2015, págs. 239-290. DOI: 10.1016/b978-0-12-800066-3.00008-5.
- [46] R. A. Fisher. «On the mathematical foundations of theoretical statistics». En: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 222.594-604 (ene. de 1922), págs. 309-368. ISSN: 0264-3952. DOI: 10.1098/rsta.1922.0009.
- [47] R. A. Fisher. «Theory of Statistical Estimation». En: *Mathematical Proceedings of the Cambridge Philosophical Society* 22.5 (1925), págs. 700-725. ISSN: 14698064. DOI: 10.1017/S0305004100009580.
- [48] Richard J. Rossi. *Mathematical Statistics : An Introduction to Likelihood Based Inference*. New York: Wiley, 2018, pág. 227. ISBN: 978-1-118-77104-4.
- [49] Halbert White. «Maximum Likelihood Estimation of Misspecified Models». En: *Econometrica* 50.1 (ene. de 1982), pág. 1. ISSN: 00129682. DOI: 10.2307/1912526. URL: <https://www.jstor.org/stable/1912526?origin=crossref>.
- [50] Germán Rojas, Juan Carlos Trujillo y Fabián Barba. *Cálculo Diferencial en una Variable*. Ed. por Juan Carlos Trujillo. Vol. 1. Quito: Facultad de Ciencias - Escuela Politécnica Nacional, 2009. URL: <http://blog.espol.edu.ec/cesdamos/files/2011/11/calculo-diferencial-en-una-variable-juan-carlos-trujillo.pdf>.
- [51] Alfio Quarteroni y col. *Matematica Numerica*. 2008, pág. 509. ISBN: 9788847007826. DOI: 10.1007/978-88-470-0818-2.
- [52] Helmut Neunzert y Abul Hasan Siddiqi. «Algorithms for Optimization». En: 2000, págs. 53-77. ISBN: 0262039427. DOI: 10.1007/978-1-4757-3222-1\_2.
- [53] Jerrold E. Marsden y Anthony Tromba. *Cálculo vectorial*. ISBN: 9788490355787.
- [54] J. E. (John E.) Dennis y Robert B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. Society for Industrial y Applied Mathematics, 1996, pág. 378. ISBN: 9780898713640.

- 
- [55] William H Press y col. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Vol. 1. 2007, pág. 1262. ISBN: 0521880688. DOI: 10.1137/1031025. URL: <http://www.amazon.com/Numerical-Recipes-3rd-Edition-Scientific/dp/0521880688?SubscriptionId=0JYN1NVW651KCA56C102%7B%5C%26%7Dtag=techkie-20%7B%5C%26%7DlinkCode=xm2%7B%5C%26%7Dcamp=2025%7B%5C%26%7Dcreative=165953%7B%5C%26%7DcreativeASIN=0521880688>.
- [56] WC Davidon - Labs., ANL-5990 Rev y undefined 1959. «Variable metric method for minimization, Argonne Natl». En: () .
- [57] R. Fletcher y M. J. D. Powell. «A Rapidly Convergent Descent Method for Minimization». En: *The Computer Journal* 6.2 (ago. de 1963), págs. 163-168. ISSN: 0010-4620. DOI: 10.1093/comjnl/6.2.163. URL: <https://academic.oup.com/comjnl/article-lookup/doi/10.1093/comjnl/6.2.163>.
- [58] E. (Elijah) Polak. *Computational methods in optimization : a unified approach*. Academic Press, 1971, pág. 329. ISBN: 9780080960913.
- [59] J.B. Joseph Fourier. *Théorie analytique de la chaleur*. Paris: Firmin Didot, père et fils, 1822. URL: <https://books.google.com/?id=TDQJAAAAIAAJ%7B%5C%26%7Dpg=PA525%7B%5C%26%7Ddq=%7B%5C%26%7D22c%7B%5C%26%7D27est-%7B%5C%60%7Ba%7D%7D-dire+qu%7B%5C%26%7D27on+a+l%7B%5C%26%7D27%7B%5C%27%7Be%7D%7Dquation%7B%5C%26%7D22%7B%5C%26%7Dv=onepage%7B%5C%26%7Dq=%7B%5C%26%7D22c%7B%5C%26%7D27est-%7B%5C%60%7Ba%7D%7D-dire%20qu%7B%5C%26%7D27on%20a%20l%7B%5C%26%7D27%7B%5C%27%7Be%7D%7Dquation%7B%5C%26%7D22>.
- [60] Elias Stein y Rami Shakarchi. *Fourier Analysis: An introduction*. Princeton University Press, 2003. ISBN: 978-0-691-11384-5. URL: <https://books.google.com/?id=FAOc24bTfGkC%7B%5C%26%7Dpg=PA158%7B%5C%26%7Ddq=%7B%5C%26%7D22The+mathematical+thrust+of+the+principle%7B%5C%26%7D22%7B%5C%26%7Dv=onepage%7B%5C%26%7Dq=%7B%5C%26%7D22The%20mathematical%20thrust%20of%20the%20principle%7B%5C%26%7D22>.
- [61] Gerald Kaiser y Lonnie H. Hudgins. «A Friendly Guide to Wavelets». En: *Physics Today* 48.7 (1995), págs. 57-58. ISSN: 19450699. DOI: 10.1063/1.2808105. URL: <https://books.google.com/?id=rfRnrhJwoloC%7B%5C%26%7Dpg=PA29%7B%5C%26%7Ddq=%7B%5C%26%7D22becomes+the+Fourier+%7B%5C%26%7D28integral%7B%5C%26%7D29+transform%7B%5C%26%7D22%7B%5C%26%7Dv=onepage%7B%5C%26%7Dq=%7B%5C%26%7D22becomes%20the%20Fourier%20%7B%5C%26%7D28integral%7B%5C%26%7D29%20transform%7B%5C%26%7D22>.

- 
- [62] Anders Vretblad. *Fourier Analysis and its Applications*. Graduate Texts in Mathematics. New York: Springer, 2000. ISBN: 978-0-387-00836-3.
- [63] Boualem Boashash. *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Oxford: Elsevier Science, 2015, págs. 1-1020. ISBN: 9780123984999. DOI: 10.1016/C2012-0-00024-5.
- [64] David Kammler. *A First Course in Fourier Analysis*. Prentice Hall, 2000. ISBN: 978-0-13-578782-3.
- [65] Hirotugu Akaike. «A New Look at the Statistical Model Identification». En: *IEEE Transactions on Automatic Control* 19.6 (1974), págs. 716-723. ISSN: 15582523. DOI: 10.1109/TAC.1974.1100705.
- [66] Leon Brillouin y R. W. Hellwarth. «Science and Information Theory». En: *Physics Today* 9.12 (1956), págs. 39-40. ISSN: 0031-9228. DOI: 10.1063/1.3059856.
- [67] Amy Courtney y Michael Courtney. «Comments Regarding "On the Nature of Science"». En: *Physics in Canada* 64.3 (dic. de 2008), págs. 7-8. arXiv: 0812.4932. URL: <http://arxiv.org/abs/0812.4932>.
- [68] Kenneth P. Burnham, David R. Anderson y Kenneth P. Burnham. *Model selection and multimodel inference : a practical information-theoretic approach*, pág. 208. ISBN: 9781441929730.
- [69] J. E. Nash y J. V. Sutcliffe. «River flow forecasting through conceptual models part I - A discussion of principles». En: *Journal of Hydrology* 10.3 (abr. de 1970), págs. 282-290. ISSN: 00221694. DOI: 10.1016/0022-1694(70)90255-6.
- [70] P Krause, D P Boyle y F Bäse. «Comparison of different efficiency criteria for hydrological model assessment». En: *Advances in Geosciences* 5 (dic. de 2005), págs. 89-97. ISSN: 16807359. DOI: 10.5194/adgeo-5-89-2005. URL: <https://hal.archives-ouvertes.fr/hal-00296842>.
- [71] Richard H. McCuen, Zachary Knight y A. Gillian Cutter. «Evaluation of the Nash–Sutcliffe Efficiency Index». En: *Journal of Hydrologic Engineering* 11.6 (nov. de 2006), págs. 597-602. ISSN: 1084-0699. DOI: 10.1061/(ASCE)1084-0699(2006)11:6(597). URL: <http://ascelibrary.org/doi/10.1061/%7B%5C%%7D28ASCE%7B%5C%%7D291084-0699%7B%5C%%7D282006%7B%5C%%7D2911%7B%5C%%7D3A6%7B%5C%%7D28597%7B%5C%%7D29>.

- 
- [72] C. J. Willmott y K. Matsuura. «On the use of dimensioned measures of error to evaluate the performance of spatial interpolators». En: *International Journal of Geographical Information Science* 20.1 (ene. de 2006), págs. 89-102. ISSN: 13658816. DOI: 10.1080/13658810500286976.
- [73] Cort J. Willmott. «On the validation of models». En: *Physical Geography* 2.2 (1981), págs. 184-194. ISSN: 02723646. DOI: 10.1080/02723646.1981.10642213. URL: <https://www.tandfonline.com/doi/abs/10.1080/02723646.1981.10642213>.
- [74] Cort J. Willmott, Scott M. Robeson y Kenji Matsuura. «A refined index of model performance». En: *International Journal of Climatology* 32.13 (nov. de 2012), págs. 2088-2094. ISSN: 08998418. DOI: 10.1002/joc.2419.
- [75] Cort J. Willmott y col. «Statistics for the evaluation and comparison of models». En: *Journal of Geophysical Research* 90.C5 (sep. de 1985), pág. 8995. ISSN: 0148-0227. DOI: 10.1029/jc090ic05p08995.
- [76] Instituto Geofísico - EPN. *Red Nacional de Sismógrafos (RENSIG)*. 2018. URL: <https://www.igepn.edu.ec/red-nacional-de-sismografos> (visitado 20-07-2018).
- [77] Jeanne C Adams y col. *Fortran 90 Handbook*. 1992, pág. 740. ISBN: 0070004064.
- [78] IRIS. *MiniSEED*. 2018. URL: <https://ds.iris.edu/ds/nodes/dmc/data/formats/miniseed/> (visitado 25-07-2018).
- [79] George Helffrich, James Wookey y Ian Bastow. *The Seismic Analysis Code: A Primer and User's Guide*. 2013, pág. 183. ISBN: 9781107045453. DOI: <http://dx.doi.org/10.1108/ET-01-2016-0019>. URL: <http://www.cambridge.org/ec/academic/subjects/earth-and-environmental-science/solid-earth-geophysics/seismic-analysis-code-primer-and-users-guide?format=PB%7B%5C%20&%7Disbn=9781107613195%7B%5C%20%7DAdam7kdg2GoJQ0B.97>.
- [80] Incorporated Research Institutions for Seismology. *IRIS*. 2020. URL: <https://www.iris.edu/hq/> (visitado 12-03-2020).
- [81] IRIS PASSCAL Instrument Center y New Mexico Tech. *How do I install PASSOFT? | PASSCAL Instrument Center*. URL: <https://www.passcal.nmt.edu/content/how-do-i-install-passoft> (visitado 28-05-2018).
- [82] IRIS PASSCAL Instrument Center y New Mexico Tech. *PASSCAL FTP Server: Index of /ftp/software/passoft/linux*. URL: <https://www.passcal.nmt.edu/ftp/software/passoft/linux/> (visitado 12-03-2020).

- 
- [83] Incorporated Research Institutions for Seismology. *IRIS: Data Services: Nodes: DMC: Software: mseed2sac v. 2.1*. URL: <https://ds.iris.edu/ds/nodes/dmc/software/downloads/mseed2sac/2-1/> (visitado 12-03-2020).
- [84] Chad Trabant e IRIS Data Management Center. *mseed2sac/mseed2sac.md at master · iris-edu/mseed2sac · GitHub*. 2017. URL: <https://github.com/iris-edu/mseed2sac/blob/master/doc/mseed2sac.md> (visitado 13-03-2020).
- [85] IRIS. *IRIS: SAC Software Request*. 2013. URL: <http://ds.iris.edu/ds/nodes/dmc/forms/sac/> (visitado 11-08-2020).
- [86] Laboratorio de Cálculo Científico MODEMAT. *...:HPC-MODEMAT:..*. URL: <https://hpcmodemat.epn.edu.ec/> (visitado 14-10-2019).
- [87] SchedMD. *Slurm Workload Manager - Documentation*. 2019. URL: <https://slurm.schedmd.com/> (visitado 14-06-2020).
- [88] D. N. Moriasi y col. «Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations». En: *Transactions of the ASABE* 50.3 (2007), págs. 885-900. ISSN: 2151-0032. DOI: 10.13031/2013.23153. URL: <http://elibrary.asabe.org/abstract.asp?JID=3%7B%5C%7DAID=23153%7B%5C%7DCID=t2007%7B%5C%7Dv=50%7B%5C%7Di=3%7B%5C%7DT=1>.

## 7 ANEXOS

### 7.1 APÉNDICE: CÁLCULO DE LAS DERIVADAS PARCIALES DE RESIDUOS RESPECTO A PARÁMETROS DE MEDIA MÓVIL

Este apéndice demuestra que las matrices (3.8) y (3.9) son equivalentes.

El algoritmo empieza el cálculo de las derivadas parciales de la ecuación (3.4) respecto a los parámetros de media móvil en el instante  $p + 1$ :

$$\epsilon_{p+1} = y_{p+1} - c - \phi_1 y_{p+1-1} - \dots - \phi_p y_{p+1-p} - \theta_1 \epsilon_{p+1-1} - \dots - \theta_p \epsilon_{p+1-p} .$$

Pero como los errores anteriores a este instante son iguales a cero (Ec. 2.22) la ecuación anterior toma la siguiente forma:

$$\epsilon_{p+1} = y_{p+1} - c - \phi_1 y_{p+1-1} - \dots - \phi_p y_{p+1-p} ,$$

es decir,

$$\epsilon_{p+1} \neq \epsilon_{p+1}(\theta_i) \quad i = 1, 2, \dots, q .$$

Lo que implica que la derivada parcial de este error referente a cualquier parámetro de media móvil es:

$$\frac{\partial \epsilon_{p+1}}{\partial \theta_i} = 0 \quad ; i = 1, 2, \dots, q \quad (7.1)$$

Para el instante  $p + 2$  se tiene:

7.1. APÉNDICE: CÁLCULO DE LAS DERIVADAS PARCIALES DE RESIDUOS RESPECTO A PARÁMETROS DE MEDIA MÓVIL

---

$$\begin{aligned}
 \epsilon_{p+2} &= y_{p+2} - c - \phi_1 y_{p+2-1} - \dots - \phi_p y_{p+2-p} - \theta_1 \epsilon_{p+2-1} - \dots - \theta_q \epsilon_{p+1-q} \\
 &= y_{p+2} - c - \phi_1 y_{p+1} - \dots - \phi_p y_2 - \theta_1 \epsilon_{p+1} \\
 &= \epsilon_{p+2}(\theta_1)
 \end{aligned}$$

con lo cual su derivada parcial es:

$$\frac{\partial \epsilon_{p+2}}{\partial \theta_i} = \begin{cases} -\epsilon_{p+1} & i = 1 \\ 0 & i = 2, 3, \dots, q \end{cases} \quad (7.2)$$

Y en general para cualquier instante posterior a  $p + 1$ :

$$\begin{aligned}
 \epsilon_{p+j} &= y_{p+j} - c - \phi_1 y_{p+j-1} - \dots - \phi_p y_{p+j-p} - \theta_1 \epsilon_{p+j-1} - \dots - \theta_q \epsilon_{p+j-q} \\
 &= \epsilon_{p+j}(\theta_1, \theta_2, \dots, \theta_k) \quad k = 1, 2, \dots, q
 \end{aligned}$$

la derivada parcial de la ecuación anterior respecto al parámetro  $\theta_i$  (Ec. 3.7) se divide en dos casos:

■  $i \leq j$ :

$$\frac{\partial \epsilon_{p+j}(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_i} = 0 \quad \text{con} \quad k = 1, 2, \dots, q \quad (7.3)$$

■  $j < i$ :

$$\begin{aligned}
 \frac{\partial \epsilon_{p+j}}{\partial \theta_i} &= -\epsilon_{p+j-i} - \theta_1 \frac{\partial \epsilon_{p+j-1}}{\partial \theta_i} - \theta_2 \frac{\partial \epsilon_{p+j-2}}{\partial \theta_i} - \dots - \theta_q \frac{\partial \epsilon_{p+j-q}}{\partial \theta_i} \\
 &= -\epsilon_{p+(j-i)} - \theta_1 \frac{\partial \epsilon_{p+i-1+(j-i)}}{\partial \theta_i} - \theta_2 \frac{\partial \epsilon_{p+i-2+(j-i)}}{\partial \theta_i} - \dots - \theta_q \frac{\partial \epsilon_{p+i-q+(j-i)}}{\partial \theta_i}
 \end{aligned}$$

donde la diferencia entre los índices  $i$  y  $j$  definen el valor de la derivada parcial.

Como ejemplo de lo mencionado anteriormente consideremos que  $j = i$  (es decir  $j - i = 0$ ):

$$\frac{\partial \epsilon_{p+j}}{\partial \theta_i} = -\epsilon_p - \theta_1 \frac{\partial \epsilon_{p+i-1}}{\partial \theta_i} - \theta_2 \frac{\partial \epsilon_{p+i-2}}{\partial \theta_i} - \dots - \theta_q \frac{\partial \epsilon_{p+i-q}}{\partial \theta_i}$$

en la ecuación anterior, recordando la igualdad (7.3), todos los términos del lado derecho de la

7.1. APÉNDICE: CÁLCULO DE LAS DERIVADAS PARCIALES DE RESIDUOS RESPECTO A PARÁMETROS DE MEDIA MÓVIL

---

ecuación anterior desaparecen. Esto lleva a una generalización de la igualdad (7.1),

$$\frac{\partial \epsilon_{p+i+1}}{\partial \theta_i} = 0.$$

Ahora, si la diferencia entre  $i$  y  $j$  es igual a uno ( $j - i = 1$ ) se obtiene el mismo resultado de la igualdad (7.2),

$$\begin{aligned} \frac{\partial \epsilon_{p+j}}{\partial \theta_i} &= -\epsilon_{p+1} - \theta_1 \frac{\partial \epsilon_{p+i-1+1}}{\partial \theta_i} - \theta_2 \frac{\partial \epsilon_{p+i-2+1}}{\partial \theta_i} - \dots - \theta_q \frac{\partial \epsilon_{p+i-q+1}}{\partial \theta_i} \\ \frac{\partial \epsilon_{p+i+1}}{\partial \theta_i} &= -\epsilon_{p+1} - \theta_1 \frac{\partial \epsilon_{p+i}}{\partial \theta_i} - \theta_2 \frac{\partial \epsilon_{p+i-1}}{\partial \theta_i} - \dots - \theta_q \frac{\partial \epsilon_{p+i-q+1}}{\partial \theta_i} \\ \frac{\partial \epsilon_{p+i+1}}{\partial \theta_i} &= -\epsilon_{p+1} \end{aligned}$$

Y siguiendo esta lógica se puede determinar todas las derivadas parciales de la matriz (3.8), que dependen de la diferencia entre  $i$  y  $j$ . Por tanto, es evidente que la matriz (3.9) se deriva de la matriz (3.8).

## 7.2 CÓDIGOS FUENTE

**Código 7.1:** Encontrar el mínimo de una función.

```

1  SUBROUTINE MINDFP(P0,N,P,Q,Y,GTOL,ITER,FRET)
-  IMPLICIT NONE
-  INTEGER, INTENT(IN) :: N, P, Q           !N: TAMAÑO DE LA MUESTRA TOMADA DE LA SERIE; P
      : ORDEN AUTO-REGRESIVO; Q: ORDEN DE MEDIA MÓVIL.
-  REAL*8,INTENT(INOUT) :: P0(P+Q+1)      !PUNTO ARBITRARIO INICIAL Y PUNTO MÍNIMO QUE
      RETORNA EL ALGORITMO.
5  INTEGER,INTENT(OUT) :: ITER              !NÚMERO DE ITERACIONES.
-  REAL*8, INTENT(OUT) :: FRET              !VALOR MÍNIMO DE LA FUNCIÓN.
-  REAL*8,INTENT(IN) :: GTOL, Y(1:N)      !GTOL: REQUERIMIENTO DE CONVERGENCIA DEL GRADIENTE
      . Y: SERIE TEMPORAL SÍSMICA ORIGINAL.
-  INTEGER, PARAMETER :: NMAX=50,ITMAX=200,STPMX=100 !NMAX: NÚMERO MÁXIMO DE PARÁMETROS
      A OPTIMIZAR. ITMAX: NÚMERO MÁXIMO DE ITERACIONES. STPMAX: LONGITUD MÁXIMA DEL
      PASO ESCALADO PERMITIDO.
-  REAL*8, PARAMETER :: EPS=3.D-8,TOLX=4.D0*EPS !CRITERIOS DE CONVERGENCIA DE LA FUNCIÓN.
10  INTEGER :: I, ITS, J
-  LOGICAL :: CHECK
-  REAL *8 :: DEN,FAC,FAD,FAE,FP,STPMAX,SUMA,SUMDG,SUMXI,TEMP,TEST, DG(NMAX),G(NMAX),HDG(
      NMAX),HESSIN(NMAX,NMAX),PNEW(NMAX),XI(NMAX)
-  FP=FUNC(N,P,Q,Y,P0)                    !FUNCIÓN EVALUADA EN EL PUNTO INICIAL.
-  CALL DFUNC(N,P,Q,Y,P0,G)               !GRADIENTE EN EL PUNTO INICIAL.
15  SUMA=0D0
-  DO I=1, P+Q+1 !12 !CÁLCULO MATRIZ UNIDAD.
-      DO J=1, P+Q+1 !11
-          HESSIN(I,J)=0.D0
-      END DO !11
20  HESSIN(I,I)=1.D0
-  XI(I)=-G(I) !DIRECCIÓN INICIAL.
-  SUMA=SUMA+P0(I)**2
-  END DO !12
-  STPMAX=STPMX*DMAX1(DSQRT(SUMA),FLOAT(P+Q+1))
25  DO ITS=1,ITMAX !27 !LAZO PRINCIPAL SOBRE LAS ITERACIONES.
-      ITER=ITS
-      !LA EVALUACIÓN DE FUNCIONES OCURRE EN LNSRCH; EL VALOR DE LA FUNCIÓN SE GUARDA EN
      FP PARA VOLVER A LLAMAR A LNSRCH.
-      !LNSRCH(N,P,Q,Y,XOLD,FOLD,G,P0,X,F,STPMAX,CHECK)
-      CALL LNSRCH(N,P,Q,Y,P0,FP,G,XI,PNEW,FRET,STPMAX,CHECK)
30  FP=FRET

```

## 7.2. CÓDIGOS FUENTE

```
- DO I=1,P+Q+1 !13
-   XI(I)=PNEW(I)-P0(I) !ACTUALIZACIÓN DE LA DIRECCIÓN,
-   P0(I)=PNEW(I) !Y DEL PUNTO EN EL ESPACIO DE PARÁMETROS.
- ENDDO !13
35 TEST=0.D0 !TEST DE LA CONVERGENCIA DE DELTA X.
- DO I=1,P+Q+1 !14
-   TEMP=DABS(XI(I))/DMAX1(DABS(P0(I)),1.D0)
-   IF(TEMP .GT. TEST) TEST=TEMP
- ENDDO !14
40 IF(TEST .LT. TOLX) RETURN
- DO I=1,P+Q+1 !15 !GUARDAR EL GRADIENTE ANTERIOR,
-   DG(I)=G(I)
- ENDDO !15
- CALL DFUNC(N,P,Q,Y,P0,G) !Y EL NUEVO GRADIENTE.
45 TEST=0.D0 !TEST DE LA CONVERGENCIA DEL GRADIENTE.
- DEN=DMAX1(FRET,1.D0)
- DO I=1,P+Q+1 !16
-   TEMP=DABS(G(I))*DMAX1(DABS(P0(I)),1.D0)/DEN
-   IF(TEMP .GT. TEST) TEST=TEMP
50 ENDDO !16
- IF(TEST .LT. GTOL) RETURN !INICIO CORRECCIÓN BFGS.
- DO I=1,P+Q+1 !17 !DIFERENCIA ENTRE GRADIENTES,
-   DG(I)=G(I)-DG(I)
- ENDDO !17
55 DO I=1,P+Q+1 !19 !Y DIFERENCIA ENTRE MATRICES.
-   HDG(I)=0.D0
-   DO J=1,P+Q+1 !18
-     HDG(I)=HDG(I)+HESSIN(I,J)*DG(J)
-   ENDDO !18
60 ENDDO !19
- FAC=0D0 !PRODUCTO DE LOS DENOMINADORES.
- FAE=0D0
- SUMDG=0D0
- SUMXI=0D0
65 DO I=1,P+Q+1 !21
-   FAC=FAC+DG(I)*XI(I)
-   FAE=FAE+DG(I)*HDG(I)
-   SUMDG=SUMDG+DG(I)**2
-   SUMXI=SUMXI+XI(I)**2
70 ENDDO !21
- IF(FAC .GT. DSQRT(EPS*SUMDG*SUMXI)) THEN !SI FAC NO ES SUFICIENTEMENTE POSITIVO,
-   SALTAR ACTUALIZACIÓN.
```

## 7.2. CÓDIGOS FUENTE

```

-      FAC=1D0/FAC
-      FAD=1D0/FAE
-      DO I=1,P+Q+1 !22          !VECTOR QUE CORRIGE EL PLANTEAMIENTO DE DFP.
75      DG(I)=FAC*X1(I)-FAD*HDG(I)
-      END DO !22
-      DO I=1,P+Q+1 !24          !FÓRMULA DE BFGS:
-      DO J=I,P+Q+1 !23
-      HESSIN(I,J)=HESSIN(I,J)+FAC*X1(I)*X1(J)-FAD * HDG(I)*HDG(J)+FAE*DG(I)*
-      DG(J)
80      HESSIN(J,I)=HESSIN(I,J)
-      END DO !23
-      END DO !24
-      ENDIF
-      DO I=1,P+Q+1 !26          !CÁLCULO DE LA NUEVA DIRECCIÓN A SEGUIR.
85      X1(I)=0D0
-      DO J=1,P+Q+1 !25
-      X1(I)=X1(I)-HESSIN(I,J)*G(J)
-      END DO !25
-      END DO !26
90  END DO !27          !REGRESO A EMPEZAR OTRA ITERACIÓN.
-      PRINT*, 'NÚMERO MÁXIMO DE ITERACIONES ALCANZADO'
-      RETURN
-      END SUBROUTINE MINDFP

```

**Código 7.2:** Variación de la longitud del paso entre iteraciones.

```

1  SUBROUTINE LNSRCH(N,P,Q,Y,XOLD,FOLD,G,P0,X,F,STPMAX,CHECK)
-  IMPLICIT NONE
-  INTEGER,INTENT(IN) :: N, P, Q      !TAMAÑO DE LA SERIE TEMPORAL, ORDEN AUTO-REGRESIVO Y
-  ORDEN DE MEDIA MÓVIL RESPECTIVAMENTE.
-  LOGICAL,INTENT(OUT) :: CHECK
5  REAL*8,INTENT(IN) :: XOLD(P+Q+1),FOLD,G(P+Q+1),Y(1:N)      !XOLD ES EL PUNTO INICIAL;
-  FOLD (FUNCIÓN) Y G (GRADIENTE) EVALUADOS EN XOLD; Y(N) ES LA SERIE TEMPORAL.
-  REAL*8,INTENT(INOUT) :: P0(1:P+Q+1) !DIRECCIÓN DE NEWTON.
-  REAL*8,INTENT(IN) :: STPMAX      !CANTIDAD QUE LIMITA LA LONGITUD DE LOS PASOS.
-  REAL*8,INTENT(OUT) :: X(P+Q+1),F      !X NUEVO PUNTO A LO LARGO DE LA DIRECCIÓN; F VALOR
-  DE LA FUNCIÓN QUE RETORNA.
-  REAL*8,PARAMETER :: ALF=1D-4,TOLX=1D-7      !ALF ASEGURA UNA DISMINUCIÓN SUFICIENTE EN EL
-  VALOR DE LA FUNCIÓN; TOLX ES LA CONVERGENCIA CRITERIO SOBRE DELTA X.
10  INTEGER :: I
-  REAL*8::A, ALAM, ALAM2, ALAMIN, B, DISC, F2, RHS1, RHS2, SLOPE, SUMA, TEMP, TEST,
-  TEMPLAM

```

## 7.2. CÓDIGOS FUENTE

```
- CHECK=.FALSE.
- SUMA=0D0
- DO I=1,P+Q+1 !11
15   SUMA=SUMA+P0(I)*P0(I)
- ENDDO !11
- SUMA=DSQRT(SUMA)
- IF(SUMA .GT. STPMAX) THEN      !REESCALAMIENTO SI EL PASO ES MUY LARGO.
-   DO I=1,P+Q+1 !12
20   P0(I)=P0(I)*STPMAX/SUMA
-   ENDDO !12
- ENDIF
- SLOPE=0D0
- DO I=1,P+Q+1 !13
25   SLOPE=SLOPE+G(I)*P0(I)
- ENDDO !13
- IF(SLOPE .GE. 0D0) PRINT*, 'LNSRCH: LA DIRECCIÓN SEGUIDA NO MINIMIZA LA FUNCIÓN'
- TEST=0D0      !CÁLCULO DEL VALOR MÍNIMO DEL ENTERO POSITIVO S.
- DO I=1,P+Q+1 !14
30   TEMP=DABS(P0(I))/DMAX1(DABS(XOLD(I)),1D0)
-   IF(TEMP .GT. TEST) TEST=TEMP
- ENDDO !14
- ALAMIN=TOLX/TEST
- ALAM=1D0      !PRIMERO SE INTENTA EL PASO DE NEWTON COMPLETO.
35 1 CONTINUE      !INICIO DE LAZO DE ITERACIÓN.
-   DO I=1,P+Q+1 !15
-     X(I)=XOLD(I)+ALAM*P0(I)
-   ENDDO !15
-   F=FUNC(N,P,Q,Y,X)
40   IF(ALAM .LT. ALAMIN) THEN      !CONVERGENCIA EN DELTA X.
-     DO I=1,P+Q+1 !16
-       X(I)=XOLD(I)
-     ENDDO !16
-     CHECK=.TRUE.
45   RETURN
-   ELSE IF(F .LE. FOLD+ALF*ALAM*SLOPE) THEN      !CRITERIO DE DECRECIMIENTO DE LA FUNCIÓN, (P0=X-XOLD).
-     RETURN
-   ELSE      !RETROCESO.
-     IF(ALAM .EQ. 1D0) THEN      !PRIMERA VEZ QUE SE RETROCEDE.
50       TEMPLAM=-SLOPE/(2D0*(F-FOLD-SLOPE))
-     ELSE      !SIGUIENTES RETROCESOS.
-       RHS1=F-FOLD-ALAM*SLOPE
```

## 7.2. CÓDIGOS FUENTE

```

-      RHS2=F2-FOLD-ALAM2*SLOPE      ! F2=F(XOLD+ALAM2*P0)
-      A=(RHS1/(ALAM**2)-RHS2/(ALAM2**2))/(ALAM-ALAM2)
55     B=(-ALAM2*RHS1/(ALAM**2)+ALAM*RHS2/(ALAM2**2))/(ALAM-ALAM2)
-      IF(A .EQ. 0D0) THEN
-          TEMPLAM=-SLOPE/(2D0*B)
-      ELSE
-          DISC=B*B-3D0*A*SLOPE      ! DISCRIMINANTE.
60     IF(DISC .LT. 0.D0) THEN
-          TEMPLAM=0.5D0*ALAM
-          ELSE IF(B .LE. 0D0) THEN
-          TEMPLAM=(-B+DSQRT(DISC))/(3D0*A)
-          ELSE
65     TEMPLAM=-SLOPE/(B+DSQRT(DISC))
-          ENDIF
-      ENDIF
-      IF(TEMPLAM .GT. 0.5D0*ALAM) TEMPLAM=0.5D0*ALAM      !S <= 0.5*S_1.
-      ENDIF
70     ENDIF
-      ALAM2=ALAM
-      F2=F
-      ALAM=DMAX1(TEMPLAM,0.1D0*ALAM)      !S >= 0.1 S_1 .
-      GOTO 1                          ! INICIO DEL NUEVO LOOP.
75 END SUBROUTINE LNSRCH

```

**Código 7.3:** Cálculo de la función Log Likelihood.

```

1  DOUBLE PRECISION FUNCTION FUNC(T,P,Q,Y,TH)
-  IMPLICIT NONE
-  !TH=(PHI_1, ... , PHI_P, THETA_1,... ,THETA_Q,C)
-  INTEGER, INTENT(IN) :: T, P, Q      !TAMAÑO DE LA SERIE TEMPORAL, ORDEN AUTO-REGRESIVO Y
-      ORDEN DE MEDIA MÓVIL RESPECTIVAMENTE.
5  REAL*8 :: EP(T)      !VECTOR DE ERRORES ALEATORIOS.
-  INTEGER :: I
-  REAL*8 :: SUMAT
-  REAL*8, INTENT(IN) :: TH(P+Q+1), Y(1:T)      !Y(T) SERIE TEMPORAL.
-  SUMAT=0D0
10 CALL ERRVECT(T,P,Q,Y,TH,EP)
-  DO I=P+1,T
-      SUMAT=SUMAT+EP(I)*EP(I)
-  END DO
-  FUNC=SUMAT
15 END FUNCTION FUNC

```

## 7.2. CÓDIGOS FUENTE

**Código 7.4:** Construcción del vector error.

```
1 SUBROUTINE ERRVECT(T,P,Q,Y,TH,ER)
- IMPLICIT NONE
- !ER_T=Y(T)-TH(P+Q+1)-AR-MA
- !TH=(PHI_1, ..., PHI_P, THETA_1,...,THETA_Q,C) .
5 INTEGER, INTENT(IN) :: T, P, Q
- REAL*8,INTENT(IN) :: Y(1:T), TH(1:P+Q+1) !CONSTANTE=TH(P+Q+1) .
- REAL*8, INTENT(OUT) :: ER(1:T)
- INTEGER :: I, J, K
- REAL*8 :: AR, MA
10 ER=0D0
- DO I=P+1,T !1 !LASO PRINCIPAL .
- AR=0.D0; MA=0.D0
- DO J=1, P !PARTE AUTO-REGRESIVA DEL ERROR.
- AR=AR+TH(J)*Y(I-J)
15 END DO
- IF (Q .LE. P) THEN !PARTE DE MEDIA MÓVIL .
- DO K=1,Q
- MA=MA+TH(K+P)*ER(I-K)
- END DO
20 ELSE
- DO K=1,P
- MA=MA+TH(K+P)*ER(I-K)
- END DO
- DO K=P+1,Q
25 IF (I-K .LT. 1) EXIT
- MA=MA+TH(K+P)*ER(I-K)
- END DO
- END IF
- ER(I)=Y(I)-TH(P+Q+1)-AR-MA
30 END DO !1
- END SUBROUTINE ERRVECT
```

**Código 7.5:** Gradiente de la función Log Likelihood.

```
1 SUBROUTINE DFUNC(T,P,Q,Y,P0,G)
- IMPLICIT NONE
- !!!P0=(PHI_1, ..., PHI_P, THETA_1,...,THETA_Q,C)
- !!!G(I)=(d FUNC)/(d P0(I)); I=1,...,P+Q+1
5 INTEGER, INTENT(IN) :: T, P, Q !TAMAÑO DE LA SERIE TEMPORAL, ORDEN AUTO-REGRESIVO Y
ORDEN DE MEDIA MÓVIL RESPECTIVAMENTE.
- REAL*8,INTENT(IN) :: Y(T) !VECTOR DE SERIE TEMPORAL.
```

## 7.2. CÓDIGOS FUENTE

```
- REAL*8, INTENT(IN), DIMENSION(P+Q+1):: P0
- DOUBLE PRECISION, INTENT(OUT), DIMENSION(1:P+Q+1):: G
- REAL*8:: EP(1:T), DC(P+1:T), DPH(P+1:T,1:P), DTH(P+1:T) !EP = VECTOR DE ERRORES
  ALEATORIOS; DC = DERIVADAS PARCIALES REFERENTES A C; DPH = DERIVADAS PARCIALES EN
  FUNCIÓN DE PARÁMETROS AUTO-REGRESIVOS; DTH = DERIVADAS PARCIALES REFERENTES A PAR
  ÁMETROS DE MEDIA MÓVIL.
10 INTEGER:: I,J,K
- G=0.D0
- CALL ERRVECT(T,P,Q,Y,P0,EP)
- !1. G(P+Q+1)=(d FUNC)/(d CONSTANT).
- DC=-1D0
15 DC(P+2)=-1D0+P0(P+1)
- G(P+Q+1)=EP(P+1)*DC(P+1)+EP(P+2)*DC(P+2) !PRIMER COMPONENTE DE G(P+Q+1).
- DO I=P+3,T !2 !ELEMENTOS DE G(P+Q+1).
  IF (I .LE. P+Q) THEN
    DO J=1, I-P-1 !2.1
      DC(I)=DC(I)-P0(P+J)*DC(I-J)
20    END DO !2.1
  ELSE
    DO J=1,Q !2.2
      DC(I)=DC(I)-P0(P+J)*DC(I-J)
25    END DO !2.2
  END IF
  G(P+Q+1)=G(P+Q+1)+EP(I)*DC(I)
- END DO !2
- !2. G(I)=(d FUNC)/(d P0(I)); I=1,...,P.
30 DO I=P+1,T !3
  DO J=1,P !3.1
    DPH(I,J)=-Y(I-J)
  END DO !3.1
- END DO !3
35 DO J=1,P !4
  DO I=P+2,T !4.1
    IF (I .LE. P+Q) THEN
      DO K=1,I-P-1 !4.1.1
        DPH(I,J)=DPH(I,J)-P0(K+P)*DPH(I-K,J)
40      END DO !4.1.1
    ELSE
      DO K=1,Q !4.1.2
        DPH(I,J)=DPH(I,J)-P0(P+K)*DPH(I-K,J)
      END DO !4.1.2
45    END IF
```

## 7.2. CÓDIGOS FUENTE

```
-      END DO          !4.1
- END DO          !4
- DO I=1,P !1
-     DO J=P+1,T !1.1
50      G(I)=G(I)+EP(J)*DPH(J,I)
-     END DO          !1.1
- END DO          !1
- !3. G(I)=(d FUNC)/(d P0(I));    I=P+1,...,P+Q
- DTH=0D0
55 DTH(P+2)=-EP(P+1)
- DTH(P+3)=-EP(P+2)+P0(P+1)*EP(P+1)
- DO I=P+4,T !6
-     DTH(I)=-EP(I-1)
-     IF (I .LE. P+Q) THEN
60         DO K=1,I-P-1 !6.1
-             DTH(I)=DTH(I)-P0(P+K)*DTH(I-K)
-         END DO          !6.1
-     ELSE
-         DO K=1,Q          !6.2
65         DTH(I)=DTH(I)-P0(P+K)*DTH(I-K)
-         END DO          !6.2
-     END IF
- END DO          !6
- DO I=1,Q !1
70     DO J=P+I,T !1.1
-         G(P+I)=G(P+I)+EP(J)*DTH(J-I+1)
-     END DO          !1.1
- END DO          !1
- !4. G(I)
75 G=2D0*G !G=-2*G, PARA MAXIMIZAR SE MULTIPLICA POR -1.
- END SUBROUTINE DFUNC
```

**Código 7.6:** Script que administrar los recursos del HPC-MODEMAT.

```
1 #!/bin/bash
-
- #SBATCH --qos=extended
- #SBATCH --time=24:00:00 # walltime
5 #SBATCH --ntasks=1 # number of processor cores (i.e. tasks)
- #SBATCH --nodes=1 # number of nodes
- #SBATCH --mem-per-cpu=1024M # memory per CPU core
- #SBATCH -J "seismictimeseries" # job name
```

## 7.2. CÓDIGOS FUENTE

```
- #SBATCH --mail-user=william.castillo@epn.edu.ec # email address
10 #SBATCH --mail-type=BEGIN
- #SBATCH --mail-type=END
- #SBATCH --mail-type=FAIL
-
- module load intel/2017      #Fortran compiler.
15 cd /home/wcastillo
- ifort -o result.out 1gauss.f90 2mindfp.f90 3whiten.f90 4main.f90
- ./result.out
```

**Código 7.7:** Números aleatorios independientes con distribución gaussiana.

```
1 SUBROUTINE NDRN(N,M,SD,WN)
- IMPLICIT NONE
- USE IFPORT
- REAL*8, PARAMETER:: PI=DACOS(-1D0)
5 INTEGER, INTENT(IN):: N !TAMAÑO DE LA MUESTRA.
- REAL*8, INTENT(IN):: M, SD !MEDIA Y DESVIACIÓN ESTANDAR RESPECTIVAMENTE.
- REAL*8, INTENT(OUT):: WN(1:N)
- INTEGER:: I
- REAL*8:: TEMP
10 CALL RANDOM_NUMBER(WN) ! DISTRIBUCIÓN UNIFORME.
- DO I = 1, N-1, 2 ! CONVERSIÓN A DISTRIBUCIÓN GAUSSIANA SEGÚN EL MÉTODO DE BOX-
  MULLER.
-   TEMP = SD*DSQRT(-2D0*DLOG(WN(I))) *DCOS(2D0*PI*WN(I+1))+M
-   WN(I+1) = SD*DSQRT(-2D0*DLOG(WN(I))) *DSIN(2D0*PI*WN(I+1))+M
-   WN(I) = TEMP
15 END DO
- END SUBROUTINE NDRN
```

**Código 7.8:** Construcción de un modelo en base a una serie temporal.

```
1 SUBROUTINE TSERIES(T,TOUT,P,Q,Y,TH,ITER,X)
- IMPLICIT NONE
- INTEGER, INTENT(IN):: T, TOUT, P, Q !T DEFINE EL NÚMERO DE DATOS PARA DETERMINAR EL
  VALOR DE LOS PARÁMETROS, MIENTRAS QUE, TOUT ES NÚMERO DE DATOS A PRONOSTICAR.
- REAL*8, INTENT(IN):: Y(1:T)
5 REAL*8, INTENT(INOUT):: TH(1:P+Q+1)
- REAL*8, INTENT(OUT):: X(1:TOUT)
- INTEGER, INTENT(OUT):: ITER
- REAL*8:: ER(1:T), MA, AR, GTOL, FRET, WN(1:TOUT) !WN ES EL VECTOR DE ERRORES
  ALEATORIOS GENERADO CON LA SUBRUTINA NDRN.
- REAL*8:: MED, STD
```

## 7.2. CÓDIGOS FUENTE

```
10 INTEGER:: I, J, K
- GTOL=1D-7          ! CRITERIO DE CONVERGENCIA.
- ITER=0D0; FRET=0D0    ! NÚMERO DE ITERACIONES Y VALOR MÍNIMO DE LA FUNCIÓN.
- CALL MINDFP(TH, T, P, Q, Y, GTOL, ITER, FRET)
- CALL ERRVECT(T, P, Q, Y, TH, ER)
15 MED=0D0; STD=0D0
- DO I=1, T          !3 ! MEDIA DE LOS ERRORES ALEATORIOS.
-   MED=MED+ER(I)
- ENDDO              !3
- MED=MED/DBLE(T)
20 DO I=1, T          !4 ! DESVIACIÓN ESTADAR, VARIANZA.
-   STD=STD+(ER(I)-MED)**2
- END DO             !4
- STD=DSQRT(STD/DBLE(T))
- CALL NDRN(TOUT, MED, STD, WN)  ! I.I.D. CON DISTRIBUCIÓN NORMAL.
25 X=0D0
- DO I=0, P-1        !1
-   X(P-I)=Y(T-I)    ! P PRIMEROS DATOS DEL MODELO, SON LOS P ÚLTIMOS DE LA SERIE
-   ORIGINAL.
- END DO             !1
- DO I=P+1, T        !2 ! CONSTRUCCIÓN DEL MODELO.
30   AR=0.D0; MA=0.D0    ! PARTE AUTO REGRESIVA Y MEDIA MÓVIL RESPECTIVAMENTE.
-   DO J=1, P          !2.1
-     AR=AR+TH(J)*X(I-J)
-   END DO            !2.1
-   IF (Q .LE. P) THEN
35     DO K=1, Q        !2.2
-       MA=MA+TH(K+P)*WN(I-K)
-     END DO          !2.2
-   ELSE
-     DO K=1, P        !2.3
40     MA=MA+TH(K+P)*WN(I-K)
-     END DO          !2.3
-     DO K=P+1, Q      !2.4
-       IF (I-K .LT. 1) EXIT
-       MA=MA+TH(K+P)*WN(I-K)
45     END DO          !2.4
-   END IF
-   X(I)=TH(P+Q+1)+WN(I)+AR+MA
- END DO              !2
- END SUBROUTINE TSERIES
```