



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO

Grado en Matemáticas

**ANÁLISIS MEDIANTE *WAVELETS* DEL ÍNDICE DE PRECIPITACIÓN
ESTANDARIZADO EN LA PENÍNSULA IBÉRICA**

Autor: Lydia Moreno Martín

Tutor/a/es: José Ignacio Farrán Martín y María Luisa Martín Pérez

Año: 2024

ANÁLISIS MEDIANTE *WAVELETS* DEL ÍNDICE DE PRECIPITACIÓN ESTANDARIZADO EN LA PENÍNSULA IBÉRICA

Lydia Moreno Martín

Agradecimientos

Quiero expresar mi agradecimiento a todas aquellas personas que me han acompañado en toda esta etapa académica llena de aprendizaje, algunos desafíos y por supuesto, momentos inolvidables.

En primer lugar, quiero dar las gracias a mis dos tutores D. José Ignacio Farrán Martín y Dra. María Luisa Martín Pérez por su ayuda, paciencia, consejos y ánimos durante el desarrollo del proyecto. También agradecer a D. Javier Díaz Fernández por el interés de seguir desde el inicio el proyecto.

Agradezco de corazón a mis padres, Agustín y Pilar y a mi hermana Nuria por apoyarme, darme ánimos y confiar en mí desde la distancia, sin ellos no habría sido posible.

A mi compañero de vida, Alonso, por acompañarme en los momentos buenos y en los no tan buenos. Por hacer que no me rinda nunca y confiar en mí, haciendo cumplir mi sueño. Gracias por celebrar mis logros como si fueran los tuyos.

A mis amigas de siempre, Esther y Tina, gracias por estar siempre ahí, incluso en los momentos más complicados haciendo que cada situación se convierta siempre en una locura de tres. A mis amigos de la universidad, que serán para siempre, especialmente a Paula y Miguel por dejarme aprender a vuestro lado y apoyarme a lo largo de estos años.

Y para finalizar, pero no menos importante, agradecer al resto de familiares, los que están y los que siguen viviendo en mi corazón por su amor y ser una fuente de palabras que sustentan mi motivación.

A todos ellos, GRACIAS, sin vosotros no lo habría conseguido.

Resumen

Este Trabajo Fin de Grado (TFG) se centra en el estudio de las *wavelets*, una herramienta matemática de gran utilidad para el análisis de señales y series temporales, con un enfoque especial para el análisis de datos no estacionarios. A diferencia de la transformada de Fourier, las *wavelets* permiten representar simultáneamente dichas señales en el tiempo y la frecuencia, lo que las hace especialmente útiles para descomponer y estudiar señales que varían a lo largo del tiempo.

Dado que actualmente el cambio climático es uno de los aspectos más preocupantes dentro de la sociedad, se ha decidido aplicar las *wavelets* a series temporales de datos meteorológicos, concretamente al análisis del Índice de Precipitación Estandarizado (SPI) en distintos puntos de la península ibérica. Este índice permite identificar periodos de déficit y/o superávit de precipitaciones, es decir, periodos secos y húmedos.

El estudio se enfoca en el análisis del SPI a diferentes escalas temporales mensuales (SPI3, SPI6, SPI9 y SPI12), utilizando *k-means* y el método del codo para disminuir el volumen de datos y para agrupar los datos en distintos *clusters* optimizando la representación de las series temporales, obteniendo un total de 6 *clusters*. Posteriormente, se emplea la transformada de *wavelets* para analizar la media de cada serie temporal en los distintos grupos, lo cual facilita la identificación de patrones de anomalías de precipitación recurrentes a largo plazo.

Los resultados obtenidos muestran la utilidad de las *wavelets* para identificar componentes periódicas y analizar fenómenos meteorológicos. Dichos resultados se han obtenido del análisis del SPI12 debido a que uno de los objetivos de este trabajo es estudiar los grandes periodos de anomalías de precipitación desde 1962 hasta 2024. Además, se seleccionó la *wavelet* de sombrero mexicano ya que son más adecuadas para altas frecuencias frente a la *wavelet* de Morlet la cual es más adecuada para datos de baja frecuencia.

Palabras claves: *wavelets*, SPI, sequías, precipitación, Morlet, sombrero mexicano, series temporales.

Abstract

This paper will focus on the study of wavelets, a Mathematical tool widely used for non-stationary data. Opposing to Fourier Transform, wavelets allows simultaneous representations of time signals and frequency, making them specially useful for tasks such as decomposing and studying signals that changes over time.

Taking into consideration that, nowadays, climate change is one of the most alarming topics withing society, it has been decided to apply the wavelets to a a temporal series of meteorologic data, specifically in the analysis of the Standardized Precipitation Index (SPI) in different places of the Iberian Peninsula. This index allows to identify periods of deficit and/or surplus of precipitations, which means dry and humid periods.

The paper also focuses on the analysis of the SPI on different monthly-timed scales (SPI3, SPI6, SPI9 y SPI12) using k-means and the Elbow method to shrink the volume of data and group the different data in clusters, optimizing the representation of the temporal series, resulting a total of 6 clusters. Later, it uses the wavelet transformation to analyze the mean of each temporal series in the different groups, which makes it easier to to identify anomaly precipitation patterns recurring over time.

The results obtained shows the utility of wavelets to identify periodic components and analyze meteorologic phenomenons. These results had been obtained from the analysis of the SPI12 because one of their objectives is to study the great precipitation anomalies periods from 1962 to 2024. In addition, the Mexican Hat Wavelet selected the ones that showed to be more adequate for high frequencies opposing to the Morlet wavelet which is more adequate for low frequency data.

Keywords: wavelets, SPI, drought, precipitation, Morlet, Mexican Hat, time series.

Índice general

Agradecimientos

Resumen

Abstract

Lista de figuras v

Lista de tablas vii

I Fundamentos teóricos 1

1. Introducción 3

1.1. Objetivos 4

1.2. Conceptos previos 5

1.3. Espacios L^p 7

1.4. Estructura de la memoria 9

2. Serie temporal 11

2.1. Clasificación de series temporales 11

2.2. Componentes de las series temporales 12

3. K-Means 13

3.1. Regla del codo 13

3.2. Algoritmo 14

3.3. Criterio de parada 15

4. Teoría de Fourier 17

4.1. Series trigonométricas 17

4.2. Funciones ortogonales 18

4.3. Serie de Fourier generalizada 18

4.4. Coeficientes y series de Fourier 19

4.4.1. Serie de Fourier compleja 20

| | | |
|-----------|--|-----------|
| 4.5. | Convergencia de la serie | 21 |
| 4.6. | Transformada de Fourier | 22 |
| 4.6.1. | Propiedades | 22 |
| 4.6.2. | Transformada inversa de Fourier | 22 |
| 4.6.3. | Transformada Discreta de Fourier (DFT) | 24 |
| 4.6.4. | Transformada rápida de Fourier (FFT) | 25 |
| 4.6.5. | Transformada corta de Fourier (STFT) | 26 |
| 4.6.6. | Aplicaciones | 26 |
| 5. | Wavelets | 29 |
| 5.1. | La transformada <i>Wavelet</i> | 29 |
| 5.2. | Transformada de Fourier vs <i>Wavelet</i> | 30 |
| 5.3. | Algunas <i>wavelet</i> | 30 |
| 5.3.1. | <i>Wavelet</i> de Daubechies | 30 |
| 5.3.2. | <i>Wavelet</i> Symmlets | 31 |
| 5.3.3. | Familia Coiflets | 31 |
| 5.3.4. | Familia de las Biortogonales | 32 |
| 5.3.5. | <i>Wavelet</i> Morlet | 33 |
| 5.3.6. | <i>Wavelet</i> Sombrero Mexicano | 34 |
| 5.3.7. | <i>Wavelet</i> hermíticas | 35 |
| 5.4. | <i>Wavelets</i> en Python | 35 |
| 5.4.1. | <i>Wavelets</i> utilizadas en el estudio | 37 |
| 5.5. | Transformada Wavelet Continua (CWT) | 37 |
| 5.6. | Análisis multiresolución (MRA) | 40 |
| 5.7. | Transformada Wavelet Discreta (DWT) | 44 |
| 5.7.1. | Función escala | 44 |
| 5.7.2. | Función wavelet | 45 |
| 5.7.3. | Aplicaciones | 45 |
| II | Experimentación Práctica | 47 |
| 6. | Datos utilizados | 49 |
| 7. | Herramientas tecnológicas y procedimiento metodológico | 55 |
| 7.1. | Herramientas tecnológicas | 55 |
| 7.2. | Procedimiento metodológico | 56 |
| 8. | Resultados | 59 |
| 8.1. | Selección índice de agregación del SPI | 59 |
| 8.2. | Selección del tipo de <i>wavelet</i> : Sombrero Mexicano vs Morlet | 61 |
| 8.3. | Análisis del SPI y <i>wavelets</i> | 63 |
| 9. | Conclusiones | 75 |

| | |
|--------------------------|-----------|
| III Apéndices | 77 |
| A. Acrónimos | 79 |
| B. Documentación adjunta | 81 |
| Bibliografía y Webgrafía | 95 |

Índice de figuras

| | |
|---|----|
| 3.1. Ejemplo del método del codo | 14 |
| 4.1. Esquema de la transformación mediante la transformada de Fourier por intervalos STFT | 26 |
| 5.1. <i>Wavelet</i> Daubechies: dbN | 31 |
| 5.2. <i>Wavelet</i> symmlets: symN | 31 |
| 5.3. <i>Wavelet</i> coiflets: coifN | 32 |
| 5.4. <i>Wavelet</i> biortogonales: bior | 32 |
| 5.5. Representación gráfica de la <i>wavelet</i> de Morlet | 34 |
| 5.6. Representación gráfica de la <i>wavelet</i> sombrero mexicano | 35 |
| 5.7. Esquema de la transformación mediante la transformada <i>wavelet</i> | 38 |
| 5.8. Esquema de la transformación mediante la transformada de Fourier. | 38 |
| 5.9. Paso 1 para obtener la Transformada <i>Wavelet</i> Continua | 39 |
| 5.10. Paso 2 para obtener la Transformada <i>Wavelet</i> Continua | 39 |
| 5.11. Paso 3 para obtener la Transformada <i>Wavelet</i> Continua | 40 |
| 5.12. Representaciones de la Transformada <i>Wavelet</i> Continua | 40 |
| 5.13. (a) Función original y (b) Representación mediante resolución unitaria | 41 |
| 5.14. Tres ejemplos de aproximación a f : (a) Resolución unitaria f_0 , (b) Resolución doble f_1 y (c) Resolución cuádruple f_2 | 41 |
| 6.1. Esquema de un evento de sequía [23]. | 51 |
| 7.1. Método del codo práctico: (a) Método del codo SPI12 y (b) Variabilidad explicada en cada cluster SPI12. | 57 |
| 7.2. Mapa con la visualización de los datos agrupados en 6 clusters | 57 |
| 7.3. Precipitación anual de la península ibérica de 1961-2022 [33]. | 58 |
| 8.1. Representación de (a) SPI3, (b) SPI6, (c) SPI9 y (d) SPI12 de la ciudad de Valladolid | 60 |
| 8.2. SPI3 para Valladolid: (a) Mexh y (b) Morlet | 62 |
| 8.3. SPI6 para Valladolid: (a) Mexh y (b) Morlet | 62 |
| 8.4. SPI9 para Valladolid: (a) Mexh y (b) Morlet | 63 |
| 8.5. SPI12 para Valladolid: (a) Mexh y (b) Morlet | 63 |

| | |
|--|----|
| 8.6. Para el <i>Cluster</i> 1: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 65 |
| 8.7. Para el <i>Cluster</i> 2: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 67 |
| 8.8. Para el <i>Cluster</i> 3: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 68 |
| 8.9. Para el <i>Cluster</i> 4: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 70 |
| 8.10. Para el <i>Cluster</i> 5: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 71 |
| 8.11. Para el <i>Cluster</i> 6: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano | 73 |

Índice de tablas

| | |
|--|----|
| 5.1. Diferencias entre wavelet Sombrero Mexicano y wavelet de Morlet | 37 |
| 6.1. Categorías de los valores SPI | 50 |
| 6.2. Códigos de las distintas regiones y sus coordenadas | 53 |

Parte I

Fundamentos teóricos

Capítulo 1

Introducción

En la historia de las matemáticas, el análisis que utiliza la herramienta de las *wavelets* tiene diferentes orígenes. Gran parte de su desarrollo comenzó alrededor de 1930 pero en aquella época se desconocía que todo pertenecía a una misma teoría.

Cada uno de los orígenes del análisis de wavelets corresponde a un punto de vista distinto y a una técnica concreta que actualmente tiene un significado científico común.

Esta teoría comienza gracias a Joseph Fourier con su análisis en el dominio-frecuencia, conocido como Análisis de Fourier, comentado con detalle en el Capítulo 4 de este proyecto. Las afirmaciones llevadas a cabo por Fourier tuvieron un valor esencial en la evolución de las ideas acerca de las funciones e integrales. El cambio de poder representar funciones mediante polinomios a hacerlo según afirmaba Fourier abrió un nuevo camino en la teoría de funciones. Posteriormente, dicha teoría fue verificada y ampliada gracias a las contribuciones matemáticas de Henri Lebesgue. El estudio de convergencia de las series de Fourier y el significado de los sistemas ortogonales provocó un avance gradual desde el análisis en el dominio tiempo-frecuencia al análisis en el dominio tiempo-escala, dando lugar más tarde a la definición de wavelets.

Se puede decir que la primera aparición de wavelets es gracias a Haar al preguntarse si existiría otro sistema ortonormal de funciones $h_0(x), h_1(x), \dots, h_n(x)$ definidas en el intervalo $[0, 1]$ de manera que, para cualquier función, $f(x)$, continua en el intervalo $[0, 1]$, la serie $\sum_{n=0}^{\infty} \langle f, h_n \rangle$ convergiera uniformemente a $f(x)$ en dicho intervalo, siendo $\langle f, h_n \rangle$ el producto escalar entre dichas funciones. Este problema no tiene solución única, pero, Haar en 1909 descubrió la solución más simple, abriendo camino a las wavelets. Las funciones Haar fueron criticadas, principalmente por utilizar funciones discontinuas para la construcción de funciones continuas, lo que resultaba incoherente para ciertas aplicaciones. Dicha limitación y la idea de aproximar el gráfico de una función mediante funciones poligonales, llevó a Faber y Schauder a sustituir las funciones de Haar por sus primitivas, iniciando la investigación entre 1910 y 1920 que dio lugar a las bases de Schauder [1][2].

Hacia 1930, diversos investigadores exploraron la representación de funciones utilizando bases de funciones de escala variable. Por ejemplo, el físico Paul Lévy demostró que las bases de Schauder son superiores a las de Fourier para analizar la regularidad de movimiento browniano (señal random). Este trabajo, junto con otros avances, reveló que los

sistemas trigonométricos no eran tan efectivos para estudiar las regularidades locales de las señales. Se sabe que la energía total de una función viene dada por la suma de los cuadrados de los coeficientes de Fourier siempre que la función se pueda desarrollar en una base de Fourier, ahora bien, en física a veces es conveniente saber si dicha energía se encuentra concentrada en algunos puntos o distribuida a lo largo del eje real, y para ello, simultáneamente, Littlewood y Paley desarrollaron métodos para analizar la distribución de la energía de una función en la serie de Fourier, contribuyendo a la definición de las wavelets. Este enfoque fue ampliado por el grupo de Zygmund en la Universidad de Chicago, que introdujo la wavelet madre y las funciones Littlewood-Paley-Stein, fundamentales para el análisis de escalas variables. Estas ideas sirvieron de base para los algoritmos de procesamiento de imágenes desarrolladas por Marr y Mallat durante 1980.

En 1927, Philip Franklin creó una base ortonormal derivada de la base de Schauder gracias al método de Gram-Schmidt, aunque su uso fue limitado debido a problemas en su construcción mediante traslaciones y dilataciones. Posteriormente, entre 1960 y 1980, Weiss y Coifman estudiaron la descomposición atómica de los espacios funcionales, es decir, el objetivo es hallar para los espacios funcionales, los átomos (elementos más simples de un espacio de funciones) y las reglas que permiten reconstruir todos los elementos de ese espacio funcional utilizando dichos átomos, mientras que Grossman y Morlet redescubrieron la identidad de Calderón, denominando wavelet analizadora a la función correspondiente [1][2].

Por último, en 1985, las wavelets alcanzaron un impacto significativo gracias a Stephane Mallat, quien estableció conexiones claras entre filtros, algoritmos piramidales y bases ortonormales de wavelets, con aplicaciones en procesamiento de imágenes. Inspirado por estos avances, Yves Meyer construyó la primera base diferenciable de wavelets, y poco después, Ingrid Daubechies desarrolló wavelets de soporte compacto que generan bases ortonormales de funciones, las cuales se han convertido en un pilar fundamental en las aplicaciones de hoy en día de esta teoría [1].

1.1. Objetivos

El objetivo principal de este TFG es conocer el concepto *wavelet* y su utilidad práctica, en este caso, su utilidad para el análisis de déficit o superávit de precipitaciones en la península ibérica siguiendo el Índice de Precipitación Estandarizado (SPI).

Durante el estudio se obtendrán distintos *clusters* de las distintas ciudades de la península ibérica para los cuales se evaluarán los periodos de superávit y/o déficit de precipitaciones. Además, se comparará cada *cluster* con la precipitación total registrada para la España peninsular.

Todo este estudio es de utilidad para poder tomar medidas anticipadas ante posibles periodos secos o húmedos.

De manera esquemática, los objetivos quedarían definidos de la siguiente manera:

- **Objetivo-1:** Estudio teórico del concepto *wavelet*, sus familias y los distintos tipos.

- **Objetivo-2:** Estudio teórico de las diferencias entre la teoría de Fourier y las *wavelets*.
- **Objetivo-3:** Estudio teórico del SPI y del fenómeno de El Niño.
- **Objetivo-4:** Desarrollo de un código en lenguaje Python capaz de realizar los distintos tipos de *wavelets* para conocer los periodos secos y húmedos de la península ibérica desde el año 1962 hasta el año 2024. Este objetivo se divide en los siguientes:
 - **Objetivo 4.1:** Seleccionar el SPI y el tipo de *wavelet* para realizar el estudio mediante el análisis de una ciudad española.
 - **Objetivo 4.2:** Obtener los distintos *clusters* de las ciudades de la península ibérica mediante el método *k-means*.
 - **Objetivo 4.3:** Obtener las distintas gráficas de la serie temporal media de cada *cluster*.
 - **Objetivo 4.4:** Obtener los gráficos correspondientes al tipo de *wavelet* seleccionado de la serie temporal de cada *cluster*.

1.2. Conceptos previos

En esta sección se recuerdan algunos conceptos básicos que usan continuamente a lo largo del proyecto.

Definición 1. *Un espacio vectorial es una terna $(V, +, \cdot)$, donde V es un conjunto no vacío y $+$, \cdot son dos operaciones definidas como $+$: $V \times V \rightarrow V$ y \cdot : $\mathbb{R} \times V \rightarrow V$, denominadas suma de vectores y producto por escalares respectivamente. Además, cuentan con las siguientes propiedades:*

1. **Asociativa:** $u + (v + w) = (u + v) + w, \forall u, v, w \in V$
2. **Conmutativa:** $u + v = v + u, \forall u, v \in V$
3. **Elemento neutro:** $\exists e \in V$ tal que $e + v = v + e = v, \forall v \in V$
4. **Elemento opuesto:** Para cada $v \in V$ existe w tal que $v + w = w + v = e$
5. **Seudo-asociativa:** $\lambda(\mu v) = (\lambda\mu)v, \forall v \in V, \forall \lambda, \mu \in \mathbb{R}$
6. **Distributiva:**
 - $\lambda(u + v) = \lambda u + \lambda v \forall u, v \in V \forall \lambda \in \mathbb{R}$
 - $(\lambda + \mu)v = \lambda v + \mu v, \forall v \in V \forall \lambda, \mu \in \mathbb{R}$
7. **Unimodular:** $1v = v, \forall v \in V$

Se dice que un espacio vectorial es de **dimensión finita** si tiene un conjunto generador con una cantidad finita de elementos.

Definición 2. Para dos vectores $u = (u_1, u_2, \dots, u_n)$ y $v = (v_1, v_2, \dots, v_n)$ en \mathbb{R}^n el **producto escalar** se define como

$$\langle u, v \rangle = \sum_{i=1}^n u_i v_i$$

Sus propiedades son:

1. **Conmutatividad:** $\langle u, v \rangle = \langle v, u \rangle$
2. **Linealidad:** $\langle au + bw, v \rangle = a\langle u, v \rangle + b\langle w, v \rangle$ para los escalares a, b .
3. **Positividad:** $\langle u, u \rangle \geq 0$ y $\langle u, u \rangle = 0$ solo si $u = 0$.

Definición 3. Sea V un espacio vectorial sobre \mathbb{R} y S un conjunto de vectores de V . Se dice que S es **ortogonal** si cualquier par de vectores distintos y no nulos de S es ortogonal, es decir, si para todo $v, w \in S$ $v, w \neq 0$, con $v \neq w$ se tiene que $\langle v, w \rangle = 0$

Definición 4. Una base, B , de un espacio vectorial V es una **base ortogonal** si sus vectores son ortogonales dos a dos.

Definición 5. Un subconjunto de la recta, $A \subset \mathbb{R}$, se denomina **conjunto abierto** si para todo $x \in A$ existe $r > 0$ tal que $(x - r, x + r) \subset A$.

Definición 6. Un subconjunto de la recta, $C \subset \mathbb{R}$, se denomina **conjunto cerrado** si el conjunto $\mathbb{R} \setminus C = \{x \in \mathbb{R} / x \notin C\}$ es abierto.

Definición 7. Un subconjunto de la recta real, $B \subset \mathbb{R}$, se denomina **conjunto acotado** si existe $M > 0$ tal que $|x| < M$ para todo $x \in B$.

Definición 8. Se dice que un subconjunto de la recta, $K \subset \mathbb{R}$ es **compacto** si es a la vez cerrado y acotado.

Definición 9. Se dice que una función tiene **soporte compacto** si el conjunto donde no se anula es cerrado y acotado. Por tanto, si $f(x)$ es una función cualquiera, su soporte es $Sop(f) = \{x \in X / f(x) \neq 0\}$. Si el conjunto $Sop(f)$ es cerrado y acotado entonces es compacto.

Definición 10. Se dice que una función tiene **soporte infinito** si el conjunto donde no se anula la función no es ni acotado ni finito, es decir, el conjunto donde no se anula la función no es acotado.

1.3. Espacios L^p

En lo que sigue, I denota un intervalo de la recta real.

Definición 11. Una función continua a trozos $f(x)$ definida en un intervalo I está acotada (o es L^∞) en I si existe $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in I$.

La L^∞ -norma de $f(x)$ se define como:

$$\|f\|_\infty = \sup\{|f(x)| : x \in I\} \quad (1.1)$$

Teorema 1. Si I es cerrado y acotado entonces cualquier función continua, $f(x)$, en I es también L^∞ en I .

Demostración. Sea $f(x)$ una función continua en un intervalo I cerrado y acotado. Entonces $f(x)$ es compacto por ser la imagen de un intervalo compacto por una aplicación continua, en particular está acotada en I , es decir, existe $M > 0$ tal que $|f(x)| \leq M$ para todo $x \in I$ \square

Ejemplos:

1. La función $f(x) = \frac{1}{x}$ es continua y tiene un valor finito en cada punto del intervalo $(0, 1]$ pero no está acotada en $(0, 1]$
2. Las funciones $\text{sen}(x)$ y $\text{cos}(x)$ son L^∞ en \mathbb{R} . Además, la función compleja e^{ix} es L^∞ en \mathbb{R} . De hecho, $\|\text{sen}(x)\|_\infty = \|\text{cos}(x)\|_\infty = \|e^{ix}\|_\infty = 1$

Definición 12. Una función continua a trozos, $f(x)$, definida en un intervalo I es integrable (o de clase L^1 o simplemente L^1) en I si la integral

$$\int_I |f(x)| dx$$

es finita.

La L^1 -norma de una función $f(x)$ se define como:

$$\|f\|_1 = \int_I |f(x)| dx \quad (1.2)$$

Observación Si $f(x)$ es L^∞ en un intervalo finito I , entonces $f(x)$ es L^1 en I .

Demostración. Sea $f(x) \in L^\infty$ entonces por definición existen $M > 0$ tal que $|f(x)| \leq M$. Para ver si $f(x)$ es L^1 necesitamos ver que $\int_I |f(x)| dx$ es finita.

$$\int_I |f(x)| dx \leq \int_I M dx < \infty \text{ Por tanto, } f(x) \text{ es } L^1 \text{ en } I. \quad \square$$

Teorema 2. Sea $f(x) \in L^1$ en \mathbb{R} y sea $\epsilon > 0$. Entonces existe un número R tal que si

$$g(x) = f(x)\chi_{[-R,R]}(x) = \begin{cases} f(x) & \text{si } x \in [-R, R], \\ 0 & \text{si } x \notin [-R, R] \end{cases}$$

entonces

$$\int_{-\infty}^{\infty} |f(x) - g(x)| dx = \|f - g\|_1 < \epsilon$$

Demostración. Como $f(x)$ es integrable, la definición de integral impropia de Riemann implica que existe v tal que

$$\lim_{r \rightarrow \infty} \int_{-r}^r |f(x)| dx = v = \int_{-\infty}^{\infty} |f(x)| dx$$

Por tanto, dado $\epsilon > 0$, existe $r_0 > 0$ tal que si $r \geq r_0$, entonces

$$\left| \int_{-\infty}^{\infty} |f(x)| dx - \int_{-r}^r |f(x)| dx \right| = \left| \int_{-r}^r |f(x)| dx - v \right| < \epsilon \quad (1.3)$$

Tomando $R \geq r_0$ y definiendo $g(x) = f(x)\chi_{[-R,R]}(x)$, entonces,

$$\begin{aligned} \int_{-\infty}^{\infty} |f(x) - g(x)| dx &= \lim_{r \rightarrow \infty} \int_{-r}^r |f(x) - g(x)| dx = \lim_{r \rightarrow \infty} \int_{R < |x| < r} |f(x)| dx \\ &= \left| \int_{-\infty}^{\infty} |f(x)| dx - \int_{-R}^R |f(x)| dx \right| < \epsilon \end{aligned}$$

siguiendo la Ecuación 1.3 □

Definición 13. Una función continua a trozos, $f(x)$, definida en un intervalo I es de cuadrado integrable (o de clase L^2 o simplemente L^2) en I si la integral

$$\int_I |f(x)|^2 dx$$

es finita.

La L^2 -norma de la función $f(x)$ se define como

$$\|f\|_2 = \left(\int_I |f(x)|^2 dx \right)^{1/2}$$

Teorema 3. (Desigualdad de Cauchy-Schwarz) Sean $f(x), g(x) \in L^2$ en un intervalo I . Entonces

$$\left| \int_I f(x)g(x) dx \right| \leq \left(\int_I |f(x)|^2 dx \right)^{1/2} \left(\int_I |g(x)|^2 dx \right)^{1/2} \quad (1.4)$$

Demostración. Supongamos que $f(x), g(x) \in \mathbb{R}$ y sea $t \in \mathbb{R}$ arbitrario. Entonces,

$$\begin{aligned} 0 &\leq \int_I (f(x) + tg(x))^2 dx = \int_I |f(x)|^2 + 2tf(x)g(x) + t^2|g(x)|^2 dx \\ &= t^2 \int_I |g(x)|^2 dx + 2t \int_I f(x)g(x) dx + \int_I |f(x)|^2 dx = At^2 + Bt + C \end{aligned}$$

Esta expresión representa una función cuadrática no negativa para todo $t \in \mathbb{R}$, es decir, o tiene una raíz doble o no tiene solución, por tanto, su discriminante debe cumplir, $B^2 - 4AC \leq 0$, es decir,

$$4\left(\int_I f(x)g(x) dx\right)^2 \leq 4\left(\int_I |g(x)|^2 dx\right)\left(\int_I |f(x)|^2 dx\right)$$

dividiendo a ambos lados por 4 y haciendo la raíz cuadrada se obtiene la desigualdad 1.4

Por otro lado, si $f(x), g(x) \notin \mathbb{R}$, entonces sabemos que si una función $h(x)$ toma valores complejos en un intervalo I y si su integral de Riemann existe como un número finito en I entonces,

$$\left|\int_I h(x) dx\right| \leq \int_I |h(x)| dx$$

aplicando esto a la función $h(x) = f(x)g(x)$ se tiene

$$\left|\int_I f(x)g(x) dx\right| \leq \int_I |f(x)||g(x)| dx$$

y ahora es igual que el caso anterior. □

1.4. Estructura de la memoria

Este documento cuenta con la siguiente estructura:

- **Parte I. Fundamentos teóricos:** Se detalla toda la teoría matemática necesaria para la realización de la experimentación práctica.
 - **Capítulo 1. Introducción:** En este capítulo se presenta la historia de las *wavelet* en distintos campos y los principales objetivos del proyecto. Además, se recuerdan algunos conceptos.
 - **Capítulo 2. Series temporales:** Se define el concepto de serie temporal así como los distintos tipos y campos en las que son utilizadas.
 - **Capítulo 3. *K-means*:** Se presenta la teoría del método de agrupación de datos utilizado.
 - **Capítulo 4. Teoría de Fourier:** Se define los coeficientes, las series y la transformada de Fourier.

- **Capítulo 5. Wavelet:** Se define el concepto *wavelet*, algunas de sus características y los diferentes tipos. Además, cuenta con un apartado de comparación con la transformada de Fourier. por otro lado, se presentan los distintos tipos de *wavelets* implementadas en Python así como las utilizadas en la Parte II. Finalmente, se presentan CWT y DWT.
- **Parte II. Experimentación Práctica:** Se detalla los resultados obtenidos del caso práctico estudiado.
 - **Capítulo 6. Datos utilizados:** Se define el Índice de Precipitación Estandarizado (SPI) utilizado para el caso práctico y la base de datos necesaria.
 - **Capítulo 7. Herramientas tecnológicas y procedimiento metodológico:** Se describen las distintas herramientas tecnológicas utilizadas. También contiene la explicación del procedimiento de análisis llevado a cabo para realizar el experimento práctico.
 - **Capítulo 8. Resultados:** Se detallan los resultados obtenidos del análisis realizado. En la primera Sección 8.1 se explica la elección del índice de agregación mediante los datos de una ciudad de la península ibérica y en la segunda sección 8.2 se explica la elección del tipo de *wavelet* entre *wavelet* sombrero mexicano y Morlet, por último, se tiene el apartado correspondiente a los resultados obtenidos en cada *cluster* de provincias analizado.
 - **Capítulo 9. Conclusiones:** Consta de las conclusiones obtenidas del estudio práctico asociado.
- **Parte III Apéndices:**
 - **Apéndice A. Acrónimos:** Lista de los distintos acrónimos utilizados en la memoria del proyecto.
 - **Apéndice B. Documentación adjunta:** Se presenta el código desarrollado en lenguaje Python para la obtención de los resultados descritos en el Capítulo 8.

Capítulo 2

Serie temporal

Las series temporales son una herramienta fundamental para el análisis meteorológico, ya que permite conocer como varían a lo largo del tiempo algunos fenómenos, como pueden ser la temperatura o la precipitación. En este capítulo se presenta la teoría básica sobre series temporales proporcionando un marco teórico para el análisis práctico (Parte II) realizado.

Definición 14. *Una serie temporal X_t $t = 1, 2, \dots, n$ es una variable estadística cuyas observaciones están ordenadas temporalmente, es decir, es una colección de observaciones recogidas secuencialmente en el tiempo generalmente en intervalos regulares.*

Este tipo de datos aparece en diversos campos, por ejemplo:

1. Economía y marketing: variación mensual del precio de alquiler de inmuebles, beneficios netos anuales de alguna empresa, etc.
2. Demografía: número de habitantes anuales en una cierta región, tasa de mortalidad o natalidad, etc.
3. Medioambiente: temperatura media mensual, precipitaciones mensual en una localidad, etc.

La herramienta descriptiva más básica es el gráfico temporal, los cuales se utilizan en este estudio para la representación de la serie temporal media obtenida.

El eje de ordenadas del gráfico representa los distintos valores de la serie y el eje de abscisas los instantes temporales considerados en la serie temporal [3][4].

2.1. Clasificación de series temporales

Una serie temporal puede ser de dos tipos dependiendo de cómo sean las observaciones recogidas: continua o discreta.

Por otro lado, existe otro tipo de clasificación en función del poder de predicción del futuro gracias a las observaciones de la serie temporal [4][5][6]:

1. Determinísticas: Se puede predecir los valores exactos.
2. Estocásticas: No se puede predecir los valores exactos y solo se puede determinar parcialmente el futuro a partir de las observaciones pasadas. Se considera que los futuros valores tienen una distribución de probabilidad que está condicionada a los valores pasados.

Además, existe la clasificación descriptiva de las series temporales:

- Estacionarias: Una serie temporal es estacionaria si es estable, es decir, la media y la variabilidad son constantes a lo largo del tiempo. Gráficamente, los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también es constante en el tiempo.
- No estacionarias: La media y/o la variabilidad cambian a lo largo del tiempo. Los cambios sobre la media indican una tendencia a crecer o decrecer a lo largo del tiempo por lo que la serie no oscila alrededor de un valor constante.

2.2. Componentes de las series temporales

- Tendencia: Dirección a la que van evolucionando los distintos datos a lo largo del tiempo. Dicha dirección puede ser ascendente, descendente o seguir un patrón más complejo.
- Estacionalidad: Patrones recurrentes en los datos con frecuencia regular o dicho de otra manera los datos presentan cierta periodicidad.
- Componente aleatoria: Variabilidad no sistemática en los datos que no puede atribuirse a ninguno de los componentes descritos anteriormente. Se pretende estudiar qué tipo de comportamiento aleatorio reside en los datos mediante modelos probabilísticos.

De estas tres componentes, las dos primeras (tendencia y estacionalidad) son componentes determinísticas mientras que la última (componente aleatoria) es, como su nombre indica, aleatoria. La serie temporal se puede denotar como

$$X_t = T_t + E_t + I_t$$

donde T_t es la tendencia y E_t e I_t es la componente estacional y la componente aleatoria respectivamente [4][5][6].

Capítulo 3

K-Means

K-means es un algoritmo de clasificación no supervisada encargado de agrupar datos en k grupos de tal manera que todos los datos o individuos de un mismo grupo o *cluster* cuentan con características similares mientras que los de los otros grupos son diferentes. Por tanto, el objetivo principal del algoritmo es agrupar observaciones similares para describir patrones que a simple vista se desconocen, para ello el algoritmo busca un número fijo (k) de *clusters*. Cada grupo está representado por el promedio de los datos que lo componen. El representante de cada grupo se denomina centroide. La cantidad de grupos, k , es un parámetro que debe establecerse a priori.

El método empieza con k centroides ubicados de manera aleatoria y después asigna cada observación al centroide más cercano. A continuación, los centroides son modificadas al promedio de los datos asignados inicialmente y se vuelve a resignar las observaciones al correspondiente centroide.

3.1. Regla del codo

Uno de los problemas que se encuentran en los métodos de *Clustering* es la elección del número de *clusters*. Es importante saber elegir el número óptimo ya que una mala elección de estos puede dar lugar a agrupaciones de datos muy heterogéneos, es decir, obtener pocos *clusters*, o datos que siendo similares se agrupan en distintos grupos, es decir, obtener muchos *clusters* [7].

En nuestro caso práctico, se utiliza el método o regla del codo para la elección del número óptimo de grupos (k).

El método del codo utiliza los valores de la inercia que se obtienen tras aplicar *k-means* sobre diferentes números de *clusters*, desde 1 a N *clusters*. Siendo la inercia la suma de las distancias al cuadrado de cada objeto del *cluster* a su centroide:

$$Inercia = \sum_{i=0}^N ||\mathbf{x}_i - \mu||^2 \quad (3.1)$$

Una vez obtenidas las inercias, se representa en una gráfica lineal la inercia respecto del número de *clusters*. En dicha gráfica se debe apreciar un cambio brusco (codo) en la

evolución de la inercia. El punto donde se observa ese cambio brusco de la inercia indica el número óptimo de *clusters* a seleccionar, es decir, el punto que representa el “codo” es el número óptimo de *clusters* [7].

En la Figura 3.1, se muestra un ejemplo del método del codo, mediante la flecha de la Figura 3.1(a) se indica donde se encuentra el número óptimo de *cluster*. Por otro lado, en la Figura 3.1(b) se muestran las distintas agrupaciones obtenidas tras la elección del número de *cluster*.

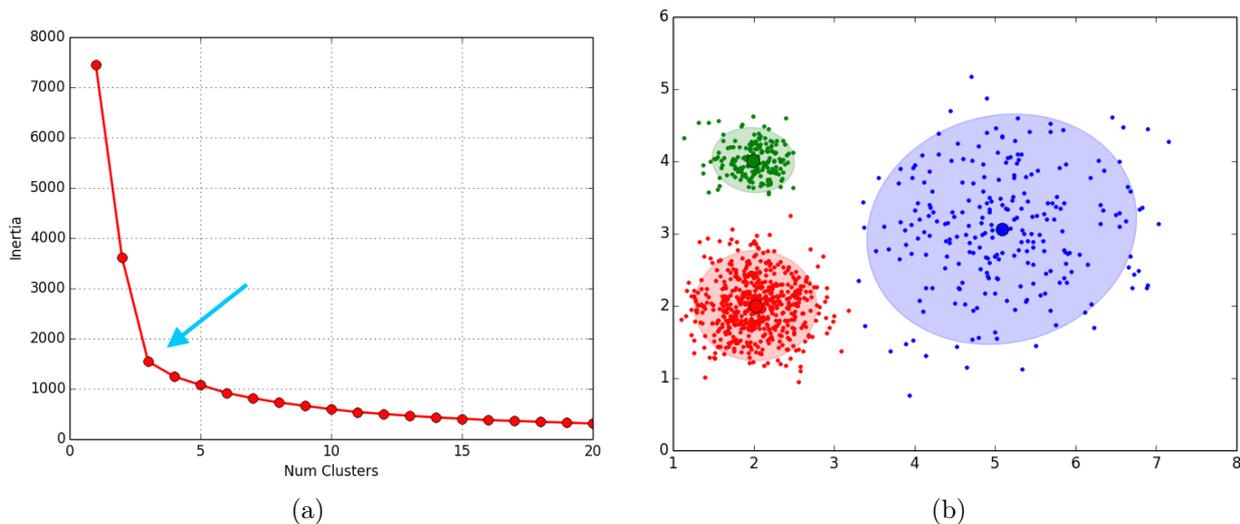


Figura 3.1: Ejemplo del método del codo

3.2. Algoritmo

El algoritmo consta de los siguientes pasos [8]:

1. **Elección del número de *clusters* k :** Elegir cuantas agrupaciones se quieren obtener, calculado mediante alguno de los métodos para tener el número óptimo.
2. **Inicializar las coordenadas de los centroides:** Los centroides se inicializan de manera aleatoria a unas coordenadas. Por ejemplo, si se tiene $k=2$ se inicializan dos centroides en puntos aleatorios de los datos.
3. **Asignar cada punto a un *cluster*:** Calcular la distancia de cada punto a cada centroide y se agrupa con aquel más próximo.
4. **Recalcular los centroides de cada *cluster*:** Una vez se asignan todos los datos a un grupo se recalculan los centroides de manera que vuelven a ser los centros de cada grupo.
5. Se repiten los pasos c) y d) hasta cumplir el criterio de parada (Sección 3.3)

El algoritmo *k-means* permite resolver un problema de optimización, siendo la función a minimizar la suma de las distancias cuadráticas de cada objeto al centroide de su *cluster*.

Los objetos son representados como vectores reales de d dimensiones $(x_1, x_2, x_3, \dots, x_n)$ de manera que el algoritmo *k-means* construye k grupos donde se minimiza la suma de distancias de los objetos, dentro de cada grupo $S = \{S_1, S_2, S_3, \dots, S_k\}$ a su centroide, es decir, se tiene el siguiente problema:

$$\min_s E(\mu_i) = \min_s \sum_{i=1}^k \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \mu_i\|^2 \quad (3.2)$$

siendo S el conjunto de datos cuyos elementos son los objetos representados como vectores. Se tendrán k grupos o *clusters* con su correspondiente centroide μ_i [8].

Por otro lado, en cada actualización de los centroides se impone la condición necesaria de extremo a la función $E(\mu_i)$, siguiendo la función cuadrática 3.2 se tiene:

$$\frac{\partial E}{\partial \mu_i} = 0 \Rightarrow \mu_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j \quad (3.3)$$

3.3. Criterio de parada

Existen tres criterios de parada:

1. **Los centroides no cambian:** Después de varias iteraciones, los centroides de cada *cluster* no varían. Se asume que el algoritmo ha convergido.
2. **Los datos dejan de cambiar de *cluster*:** Se observa que no datos que modifiquen su asignación a un *cluster*.
3. **Límite de iteraciones:** Se puede fijar un número máximo de iteraciones para la ejecución del algoritmo.

Capítulo 4

Teoría de Fourier

Jean B. Fourier demostró en 1807 que una función puede describirse mediante series trigonométricas, es decir por una serie de funciones periódicas. Los coeficientes se obtienen por integración.

4.1. Series trigonométricas

Definición 15. Una función $f(x) \in \mathbb{R}$ tiene periodo $p > 0$ si $f(x + p) = f(x) \forall x \in \mathbb{R}$. En este caso, se dice que la función es **periódica**.

Comentarios:

- Las funciones trigonométricas $\text{sen}(x)$ y $\text{cos}(x)$ tiene periodo 2π .
- Las funciones $\text{sen}(ax)$ y $\text{cos}(ax)$ tiene periodo $\frac{2\pi}{a}$
- Si una función tiene periodo $p > 0$, entonces también tiene periodo kp , para $k \in \mathbb{R}$. por tanto, una función periódica tiene muchos periodos, normalmente, se toma el periodo más pequeño para definir el periodo de la función.

Definición 16. Sea $f(x) \in \mathbb{R}$ y $p > 0$, la periodicidad de $f(x)$ se define como la función

$$f_p(x) = \sum_{n \in \mathbb{Z}} f(x + np) \quad (4.1)$$

La función 4.1 tiene periodo p puesto que.

$$f_p(x + p) = \sum_{n \in \mathbb{Z}} f(x + p + np) = \sum_{n \in \mathbb{Z}} f(x + (n + 1)p) = \sum_{r \in \mathbb{Z}} f(x + rp)$$

Definición 17. Dado $a > 0$, se define sistema trigonométrico a la colección de funciones

$$\left\{ e^{\frac{2\pi i n x}{a}} \right\}_{n \in \mathbb{Z}} \quad (4.2)$$

También puede ser representado debido a la función de Euler, $e^{i\alpha} = \text{cos}(\alpha) + i\text{sen}(\alpha)$, por

$$\left\{ 1, \text{cos}\left(\frac{2\pi n x}{a}\right), \text{sen}\left(\frac{2\pi n x}{a}\right) \right\}_{n \in \mathbb{N}} \quad (4.3)$$

Definición 18. Una función $f(x)$ definida en un intervalo I es diferenciable a trozos en I si cumple:

- a) $f(x)$ es continua por partes en I con discontinuidades de salto (si existen).
- b) $f'(x)$ existe en todos los puntos excepto para un número finito de puntos del intervalo I .
- c) $f'(x)$ es continua por partes en I con discontinuidades de salto (si existen).

4.2. Funciones ortogonales

Definición 19. El **producto escalar** o **producto interno** de dos funciones, f_1 y f_2 , pertenecientes a $L^2([a, b])$

$$\langle f_1, f_2 \rangle = \int_a^b f_1(x)f_2(x)dx$$

La norma de una función f definida en $[a, b]$ inducida por dicho producto escalar es

$$\|f\| = \left(\int_a^b f^2(x)dx \right)^{\frac{1}{2}}$$

Definición 20. Dos funciones f_1 y f_2 son **ortogonales** en el intervalo $[a, b]$ si

$$\langle f_1, f_2 \rangle = \int_a^b f_1(x)f_2(x)dx = 0$$

Definición 21. Un conjunto de funciones $\{\phi_n\}_{n=0}^{\infty}$ es **ortogonal** en el intervalo $[a, b]$ si

$$\langle \phi_m, \phi_n \rangle = \int_a^b \phi_m(x)\phi_n(x)dx = 0, m \neq n$$

Si además, se cumple la propiedad $\|\phi_n\| = 1$ para cualquier n , entonces se dice que $\{\phi_n\}_{n=0}^{\infty}$ es un **conjunto ortonormal** en el intervalo $[a, b]$

4.3. Serie de Fourier generalizada

Sea $\{\phi_n\}_{n=0}^{\infty}$ un conjunto ortogonal de funciones en el intervalo $[a, b]$ y sea f una función definida en dicho intervalo. De manera que,

$$f(x) = c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_s\phi_s(x) + \dots \quad (4.4)$$

Si f puede escribirse como en 4.4 y todas las operaciones necesarias se pueden realizar, es decir, se cumplen las condiciones suficientes de convergencia [9] entonces los coeficientes $c_m, m = 0, 1, \dots$ se calculan multiplicando la expresión 4.4 por ϕ_m e integrando en el intervalo $[a, b]$

$$\int_a^b f(x)\phi_m(x)dx = c_0 \int_a^b \phi_0(x)\phi_m(x)dx + c_1 \int_a^b \phi_1(x)\phi_m(x)dx + \dots + c_s \int_a^b \phi_s(x)\phi_m(x)dx + \dots \quad (4.5)$$

Por ortogonalidad, cada término de la derecha en la expresión 4.5 es nulo, excepto para $s = m$, es decir, se tiene

$$\int_a^b \phi_s(x)\phi_m(x)dx = 0, \quad \text{si } s \neq m$$

$$\int_a^b \phi_s(x)\phi_m(x)dx \neq 0, \quad \text{si } s = m$$

Por tanto,

$$\int_a^b f(x)\phi_m(x)dx = c_s \int_a^b \phi_s(x)\phi_s(x)dx = c_s \int_a^b \phi_s(x)^2 dx \Rightarrow f(x) = \sum_{n=0}^{\infty} c_n \phi_n(x)$$

donde los coeficientes c_n vienen dados por

$$c_n = \frac{\int_a^b f(t)\phi_n(t)dt}{\int_a^b \phi_n(t)^2 dt} = \frac{\langle f, \phi_n \rangle}{\|\phi_n\|^2}$$

En conclusión, la **serie de Fourier generalizada** o **serie ortogonal de f** es

$$f(x) = \sum_{n=0}^{\infty} \frac{\langle f, \phi_n \rangle}{\|\phi_n\|^2} \phi_n \quad (4.6)$$

4.4. Coeficientes y series de Fourier

Como se ha comentado anteriormente, las series de Fourier son series de términos coseno y seno, por las propiedades de estas funciones, las series de Fourier tienen la principal función de representar funciones periódicas. Otra de las utilidades de las series es para la resolución de ecuaciones diferenciales ordinarias y en derivadas parciales.

Supongamos que la función $f(t)$ es una función 2π -periódica en intervalo $[-\pi, \pi]$ de manera que se puede representar de la siguiente manera.

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} (a_n \cos(nt) + b_n \sin(nt)) \quad (4.7)$$

Es decir, la serie converge y su suma es $f(t)$. Los coeficientes, a_n y b_n , de las series 4.7 se

determinan a partir de la función $f(t)$ mediante las fórmulas de Euler (4.8)

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) dt \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt, \forall n > 0 \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \operatorname{sen}(nt) dt, \forall n \geq 1 \end{aligned} \quad (4.8)$$

Para un intervalo cualquiera $[-p, p]$ el conjunto $\{1, \cos(\frac{\pi}{p}x), \cos(\frac{2\pi}{p}x), \dots, \operatorname{sen}(\frac{\pi}{p}x), \operatorname{sen}(\frac{2\pi}{p}x)\}$ es ortogonal. Sea f una función que admite un desarrollo a partir del conjunto anterior (Ecuación 4.9)

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(\frac{n\pi}{p}x) + b_n \operatorname{sen}(\frac{n\pi}{p}x)) \quad (4.9)$$

donde los coeficientes $a_0, a_1, a_2, \dots, b_1, b_2, \dots$ se determinan como se indican en 4.10

$$\begin{aligned} a_0 &= \frac{1}{p} \int_{-p}^p f(t) dt \\ a_n &= \frac{1}{p} \int_{-p}^p f(t) \cos(\frac{n\pi}{p}t) dt, \forall n > 0 \\ b_n &= \frac{1}{p} \int_{-p}^p f(t) \operatorname{sen}(\frac{n\pi}{p}t) dt, \forall n \geq 1 \end{aligned} \quad (4.10)$$

Entonces la serie descrita en 4.9 recibe el nombre de **serie de Fourier** y los valores reales a_0, a_n y $b_n \forall n \geq 1$ reciben el nombre de **coeficientes de Fourier**.

4.4.1. Serie de Fourier compleja

En lo anterior se ha utilizado el sistema ortogonal formado por las funciones reales $\cos(\frac{n\pi}{p}x)$ y $\operatorname{sen}(\frac{n\pi}{p}x)$, pero también se puede utilizar un sistema ortogonal de funciones con valores complejos. Es este caso el producto interno queda definido de la manera siguiente:

$$\langle f_1, f_2 \rangle_{\mathbb{C}} = \int_p^{-p} f_1(x) \overline{f_2(x)} dx$$

El conjunto de funciones exponenciales, $\{e^{i\frac{n\pi}{p}x}\}$ es un sistema ortogonal con respecto al producto interno definido anteriormente.

Por otro lado, debido a la relación de Euler, $e^{it} = \cos(t) + i\operatorname{sen}(t)$, $i \in \mathbb{C}$, $t \in \mathbb{R}$ se tienen las siguientes relaciones para cada $t \in \mathbb{R}$,

$$\begin{aligned} \cos(t) &= \frac{e^{it} + e^{-it}}{2} \\ \operatorname{sen}(t) &= \frac{e^{it} - e^{-it}}{2i} \end{aligned}$$

La serie de Fourier compleja de f es

$$\sum_{-\infty}^{\infty} c_n e^{i \frac{n\pi}{p} x}$$

donde los coeficientes c_n vienen dados por

$$c_n = \frac{1}{2p} \int_{-p}^p f(s) e^{-i \frac{n\pi}{p} s} ds$$

La “dos” series de Fourier realmente es la misma escrita de dos maneras distintas pues los coeficientes se relacionan de la siguiente manera. Para cada $n > 0$,

$$c_n = \frac{a_n - ib_n}{2}$$

$$c_{-n} = \frac{a_n + ib_n}{2}$$

y para $n = 0$ se tiene $c_0 = \frac{a_0}{2}$

4.5. Convergencia de la serie

Sea $f(t)$ cualquier función dada de periodo 2π para la cual existen las integrales de 4.8. Entonces se pueden calcular los coeficientes de Fourier descritos en la Ecuación 4.8 de $f(t)$ y formar la serie de Fourier descrita en la Ecuación 4.7. Es interesante que la serie que se obtiene converge y su suma sea $f(t)$.

Teorema 4. Condición suficiente de convergencia puntual de una serie de Fourier.

Sea $f(t)$ una función 2π -periódica, continua a trozos en el intervalo $[-\pi, \pi]$ y que tiene derivada por la izquierda y por la derecha en todo punto de dicho intervalo. Entonces la serie de Fourier de $f(t)$ converge y su suma es

$$\frac{a_0}{2} + \sum_{n=1}^{+\infty} (a_n \cos(nt) + b_n \sin(nt)) = \frac{f(t^+) + f(t^-)}{2}$$

Observación: Si $f(t)$ es continua en t_0 , entonces se tiene que $f(t_0^-) = f(t_0^+) = f(t_0)$ y la serie de Fourier converge a $f(t_0)$ ya que

$$\frac{f(t_0^+) + f(t_0^-)}{2} = \frac{f(t_0) + f(t_0)}{2} = \frac{2f(t_0)}{2} = f(t_0)$$

4.6. Transformada de Fourier

Dada una función $f : \mathbb{R} \rightarrow \mathbb{C}$ se denomina transformada de Fourier a la función compleja dada por:

$$\mathcal{F}[f(t)](z) = \int_{-\infty}^{+\infty} f(t)e^{-itz} dt. \quad (4.11)$$

para todo $z \in \mathbb{R}$. Siempre que la expresión 4.11 tenga, sentido, es decir, sea convergente la integral impropia. Supongamos que $t, z \in \mathbb{R}$, entonces, $e^{itz} = \cos(tz) - i\text{sen}(tz)$. Si $f(t) \in \mathbb{R}$, para asegurar la convergencia absoluta basta ver que

$$\lim_{t \rightarrow \pm\infty} |f(t)e^{-itz}| = \lim_{t \rightarrow \pm\infty} |f(t)| = 0$$

pues $|e^{-itz}| = 1$.

4.6.1. Propiedades

La transformada de Fourier cuenta con dos propiedades principales, son las siguientes:

- a) **Existencia:** Dada una función, $f(t)$ definida en \mathbb{R} , es suficiente que cumpla las dos condiciones siguientes para existir la transformada de Fourier:
- $f(t)$ es continua a trozos en \mathbb{R} .
 - $f(t)$ es absolutamente integrable en \mathbb{R} .
- b) **Linealidad:** Para cualesquiera $f(t)$ y $g(t)$ que existe la transformada de Fourier y para a y b constantes,

$$\mathcal{F}[af(t) + bg(t)](z) = a\mathcal{F}[f(t)](z) + b\mathcal{F}[g(t)](z) \quad (4.12)$$

La demostración de 4.12 se debe a la linealidad de la integral y por definición de la transformada 4.11, como se puede comprobar a continuación,

$$\begin{aligned} \mathcal{F}[af(t) + bg(t)](z) &= \int_{-\infty}^{+\infty} (af(t) + bg(t)) \exp^{-itz} dt = \\ &= a \int_{-\infty}^{+\infty} f(t) \exp^{-itz} dt + b \int_{-\infty}^{+\infty} g(t) \exp^{-itz} dt = a\mathcal{F}[f(t)](z) + b\mathcal{F}[g(t)](z) \end{aligned} \quad (4.13)$$

4.6.2. Transformada inversa de Fourier

Sea $f(t)$ una función compleja de variable real que verifica:

- a) Continua o continua a trozos en cualquier intervalo finito de \mathbb{R} . Si es continua a trozos, en los puntos de discontinuidad viene dada por:

$$f(t) = \frac{1}{2}(f(t^-) + f(t^+))$$

Si no se cumple esta condición se dice que la fórmula integral de Fourier es válida o se cumple casi por todas partes. En las discontinuidades que son de salto finito la fórmula converge a:

$$f(t) = \frac{1}{2}(f(t^-) + f(t^+))$$

b) Contiene un número finito de extremos relativos.

c) Existe

$$\int_{-\infty}^{+\infty} |f(t)| dt < +\infty$$

Entonces se verifica la fórmula de la integral de Fourier

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{izt} \left(\int_{-\infty}^{+\infty} f(t) e^{-izt} dt \right) dz \quad (4.14)$$

Sabemos que la transformada de Fourier de la función f es

$$\mathcal{F}[f(t)](z) = F(z) = \int_{-\infty}^{+\infty} f(t) e^{-izt} dt.$$

por lo tanto la fórmula de la integral de Fourier (Ecuación 4.14) puede escribirse como

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{izt} F(z) dz$$

y escribimos

$$\mathcal{F}^{-1}[F(z)](t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(z) e^{izt} dz \quad (4.15)$$

que es la expresión de la **transformada inversa de Fourier**.

A modo de esquema se tiene lo siguiente:

$$f(t) \begin{array}{c} \xrightarrow{\mathcal{F}} \\ \xleftarrow{\mathcal{F}^{-1}} \end{array} F(z)$$

Si $f(t)$ es una función dada, entonces

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(z) e^{izt} dz$$

es una ecuación integral de función incógnita $F(z)$, de tipo singular por ser la integral impropia. La solución de esta ecuación integral es

$$F(z) = \int_{-\infty}^{+\infty} f(t) e^{izt} dt$$

A $F(z)$ se le llama también **función espectral** de la función $f(t)$.

4.6.3. Transformada Discreta de Fourier (DFT)

Debido a que en el ámbito computacional se trabaja con datos discretos, el cálculo numérico de la Transformada de Fourier de $f(t)$ requiere valores discretos o muestreos de $f(t)$, es decir, valores de la forma f_k con $k = 0, 1, 2, \dots$. De esta manera mediante un mecanismo computacional es posible calcular la transformada $F(w)$ sólo para valores discretos de w . Por tanto, se obtienen valores de la forma F_n con $n = 0, 1, 2, \dots$.

Supongamos que f es una señal periódica de periodo T y sólo se conocen los valores en N puntos equispaciados en el tiempo y se define como el número de muestras o la longitud de la señal. Entonces, si $f(kT_s)$ corresponde a la k -ésima muestra de $f(t)$ y $F(nw_s)$ corresponde al n -ésima muestra de $F(w)$ siendo $w_s = 2_s$ y f_s es la frecuencia con la que se obtienen las muestras. Entonces se puede escribir la Transformada de Fourier de una señal de periodo T en su forma discreta como:

$$F_n = \sum_{k=0}^{N-1} f_k e^{\frac{i2\pi kn}{N}} \quad n = 0, 1, 2, \dots, N-1 \quad (4.16)$$

Proposición 1. F_n tiene periodo N al igual que f_k .

Demostración. Por definición de periodicidad, se quiere probar que $F_n = F_{n+N}$. Por 4.16 se tiene que

$$F_{n+N} = \sum_{k=0}^{N-1} f_k e^{\frac{i2\pi k(n+N)}{N}} = \sum_{k=0}^{N-1} f_k e^{\frac{i2\pi kn}{N}} e^{\frac{i2\pi kN}{N}} = \sum_{k=0}^{N-1} f_k e^{\frac{i2\pi kn}{N}} e^{i2\pi k} = \sum_{k=0}^{N-1} f_k e^{\frac{i2\pi kn}{N}}$$

debido a la relación de Euler $e^{i2\pi k} = \cos(2\pi k) + i\sin(2\pi k) = 1 + i0 = 1 \quad \forall k = 0, 1, 2, \dots$ \square

Por tanto, se puede decir que el conjunto de coeficientes (F_n) con $n = 0, 1, 2, \dots, N-1$ es denominado la **Transformada Discreta de Fourier (DFT)** de los valores muestreados (f_k) con $k = 0, 1, 2, \dots, N-1$.

Mediante los coeficientes F_n también se puede obtener los valores f_k , es decir, se puede calcular la transformada discreta de Fourier inversa de la siguiente manera:

$$f_k = \frac{1}{N} \sum_{n=0}^{N-1} F_n e^{-\frac{i2\pi kn}{N}} \quad k = 0, 1, \dots, N-1$$

De forma matricial si $W_N^{kn} = e^{\frac{i2\pi kn}{N}}$ se puede representar de la siguiente manera la DFT (4.17) y su inversa (4.18):

$$\begin{bmatrix} F_0 \\ F_1 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N & \cdots & W_N^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix} \quad (4.17)$$

y

$$\begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N^{-1} & \cdots & W_N^{-(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{-(N-1)} & \cdots & W_N^{-(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} F_0 \\ F_1 \\ \vdots \\ F_n \end{bmatrix} \quad (4.18)$$

La matriz obtenida en 4.17 se denomina matriz de Fourier. De forma equivalente se tiene:

$$F_n = \sum_{k=0}^{N-1} f_k W_N^{kn}$$

$$f_k = \frac{1}{n} \sum_{n=0}^{N-1} F_n W_N^{-kn}$$

La matriz de Fourier es densa (todos sus elementos son no nulos).

4.6.4. Transformada rápida de Fourier (FFT)

Debido al alto costo computacional del cálculo de DFT, a mediados de la década de los sesenta desarrollan un algoritmo denominado Transformada rápida de Fourier (FFT) [10]. La FFT utiliza las propiedades de simetría y periodicidad del factor W_N para eliminar información redundante que existe en DFT, dichas propiedades son:

- **Simetría:**

$$W_N^{k+\frac{N}{2}} = e^{\frac{i2\pi(k+\frac{N}{2})}{N}} = e^{\frac{i2\pi k}{N}} e^{\frac{i2\pi \frac{N}{2}}{N}} = e^{\frac{i2\pi k}{N}} e^{\frac{i\pi N}{N}} = e^{\frac{i2\pi k}{N}} e^{i\pi} = -e^{\frac{i2\pi k}{N}} = -W_N^k$$

- **Periodicidad:**

$$W_N^{k+N} = e^{\frac{i2\pi(k+N)}{N}} = e^{\frac{i2\pi k}{N}} e^{\frac{i2\pi N}{N}} = e^{\frac{i2\pi k}{N}} e^{i2\pi} = e^{\frac{i2\pi k}{N}} = W_N^k$$

Existen dos tipos de algoritmos FFT:

- a) **Diezmado de tiempo:** Toma la totalidad de los datos de entrada f_k y los separa en dos muestras, una de ellas contiene las muestras pares y la otra las muestras impares. Cada una de ellas tiene como longitud la mitad de la original.
- b) **Diezmado en frecuencia:** Al igual que el anterior separa la señal original de longitud N en dos secuencias con longitud $\frac{N}{2}$ cada una de ellas. La diferencia está en que una secuencia contiene la primera mitad de las muestras, es decir, $k = 0, 1, 2, \dots, \frac{N}{2} - 1$ y la otra secuencia la otra mitad, $k = \frac{N}{2}, \frac{N}{2} + 1, \dots, N$

Una de las principales propiedades de FFT es su eficiencia cuando se aplica sobre señales donde el número de muestras es una potencia de 2. Además, el principio de FFT es “divide y vencerás”, ya que divide la señal en dos secuencias de datos de longitud la mitad [11].

4.6.5. Transformada corta de Fourier (STFT)

En 1946, Denis Gabor, corrige una de las desventajas de la Transformada de Fourier, cuenta con una pobre resolución en tiempo para ello, adaptó la transformada para poder analizar una pequeña sección de la señal en un determinado tiempo. Dicha adaptación recibe el nombre de Transformada corta de Fourier (STFT), la cual lleva una señal del plano del tiempo al plano bidimensional de tiempo y frecuencia (Figura 4.1)

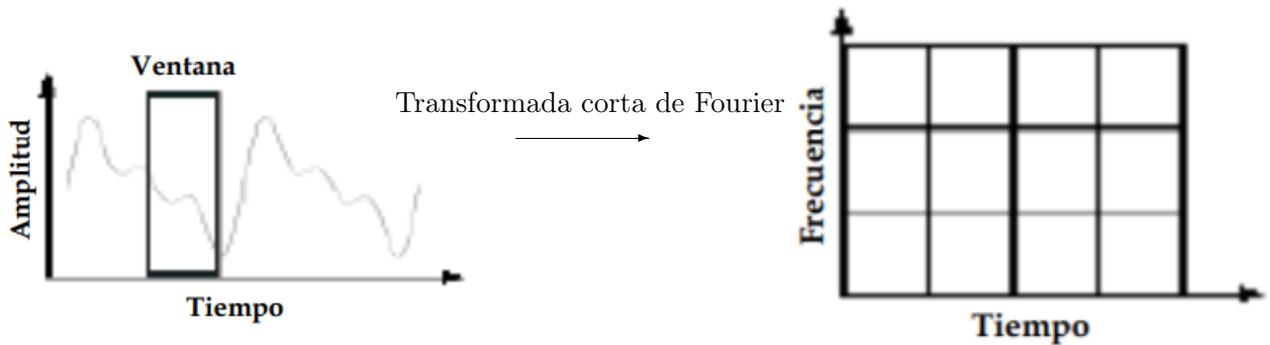


Figura 4.1: Esquema de la transformación mediante la transformada de Fourier por intervalos STFT

La división de la señal se lleva a cabo mediante una función denominada tiempo-ventana o función ventana, $h(t)$, cuyo ancho o soporte corresponde a la longitud de cada partición de la señal. Gracias a la función ventana se encuadra la señal alrededor de un instante de tiempo, τ , y se calcula la transformada de Fourier. A continuación, se traslada la función ventana hasta que no se superpone con la anterior, cubriendo de esta manera una nueva porción de la señal a la cual se vuelve a calcular la transformada de Fourier. Dicho proceso se repite hasta cubrir la totalidad de la señal.

De forma matemática:

$$STFT(t, w) = \int_{-\infty}^{+\infty} x(t)h^*(\tau - t)e^{-iwt} dt$$

siendo $h^*(\tau - t)$ el conjugado de la función ventana $h(\tau - t)$.

Si se considera $h(t) \in \mathbb{R}$ para todo t una función ventana de tal manera que $h(-t) = h^*(t)$ entonces

$$STFT(t, \xi) = \int_{-\infty}^{+\infty} x(t)h(\tau - t)e^{-i\xi t} dt \tag{4.19}$$

La Ecuación 4.19 calcula el producto interno entre la señal y la función tiempo-ventana trasladada y modulada.

4.6.6. Aplicaciones

La transformada de Fourier se encarga de cambiar el espacio temporal por el de frecuencias, y trata de localizar las frecuencias de una señal. Por ello es utilizado en diversos

campos: medicina, telecomunicaciones, circuitos eléctricos, etc. Donde más se utiliza es en la transformación de señales de potencia y energía para poder enviar información por cables u ondas. Entre las aplicaciones de telecomunicaciones también se utiliza para convertir señales analógicas en digitales. Al descomponer una señal en sus frecuencias constituyentes, se puede optimizar la codificación y transmisión, ajustando la modulación de señales para reducir interferencias y mejorar la calidad de la transmisión [12][13].

Otro ejemplo donde es utilizada la transformada de Fourier es en tratamiento de distintos archivos gráficos o de audio, en este caso suelen utilizarse las transformadas seno y coseno [14], como por ejemplo, comprimir archivos de audio, imágenes o vídeos sin perder demasiada calidad. Otra utilidad es la identificación y filtración del ruido en distintos archivos, en el campo de la medicina es de gran utilidad para imágenes de resonancia magnética. Actualmente, para el área de la inteligencia artificial se utiliza principalmente para el reconocimiento de voz se utiliza FFT o detectar patrones en imágenes [15].

Por otro lado, en ingeniería civil se utilizan para calcular el comportamiento de grandes estructuras como pueden ser el análisis de vibraciones en edificios ante terremotos.

Capítulo 5

Wavelets

En este capítulo se introduce el concepto *wavelet* y transformada de *wavelet*, la diferencia con la transformada de Fourier y la utilidad que tienen, como es el caso práctico del trabajo (Parte II).

5.1. La transformada *Wavelet*

El término “wavelet”, en inglés, viene del francés, “ondelette”, que se traduce al español como ondícula u onditas. Oscilación situada en el tiempo, cuenta con varias familias de funciones entre las que destacan: *wavelet* de Haar, en honor a Alfred Haar, la *wavelet* sombrero mexicano, onda de Morlet de Jean Morlet que cuenta también con una versión compleja, etc. Las distintas familias se pueden ver en la sección 5.4 de forma más detallada.

Definición 22. Una *wavelet* $\psi(x)$ es una función limitada en un cierto intervalo o de extensión limitada. La función *wavelet* puede ser cualquier función que cumpla las siguientes tres condiciones:

a) **Normalizada:** $\|\psi\| = 1$

b) **Energía finita**, es decir, $\psi(x) \in L^2(\mathbb{R})$:

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty \quad (5.1)$$

c) **Oscilatoria:**

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \quad (5.2)$$

Las *wavelets* tiene dos propiedades básicas:

a) **Escala o dilatación** (s): como de estirada o no se encuentra. Propiedad que tiene relación con la frecuencia.

b) **Ubicación** (t): dónde se ubica en el tiempo.

Definición 23. Una función $\psi \in L^2(\mathbb{R})$ es una **wavelet ortonormal** si el sistema $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ proporciona una base ortonormal, donde $\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k)$

La transformada *wavelet* es una función matemática que describe una señal, puede considerarse que representa tiempo-frecuencia, por lo que tiene relación con el análisis armónico. Además, se tiene dos grupos de transformadas: continua y discreta, este trabajo se centrará en las transformadas continuas explicadas en la sección 5.5.

5.2. Transformada de Fourier vs *Wavelet*

La principal diferencia entre las dos transformadas se encuentra en la descomposición de la señal, la transformada de Fourier descompone en oscilaciones que no cambian durante la secuencia, sin embargo, la transformada de *Wavelet* descompone la señal en oscilaciones que se encuentran en el espacio y tiempo, es decir, se obtiene información espectral y temporal simultáneamente de forma local.

Una segunda ventaja de la transformada *wavelet* es la existencia de varios tipos, pues si se sabe la forma de la señal se cuenta con una amplia gama para elegir la que mejor se adapte.

En cuanto al soporte de las funciones, una función *wavelet* tiene soporte compacto a diferencia de la transformada de Fourier que tiene soporte infinito [16].

5.3. Algunas *wavelet*

En esta sección se presentan algunos tipos de *wavelets*.

5.3.1. *Wavelet* de Daubechies

La familia de *wavelet* Daubechies suele denotarse, dbN, donde $N \in \mathbb{Z}$ indica el orden. Esta *wavelet* tiene soporte compacto, puede ser ortogona, biortogonal y no tiene simetría.

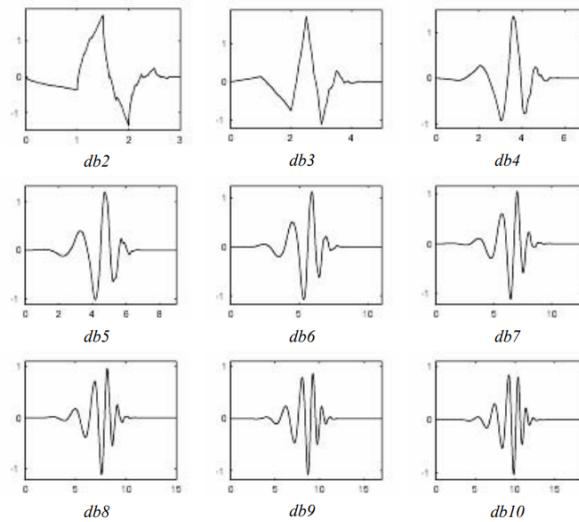
Estas *wavelets* no tienen una expresión determinada, salvo la *wavelet* de Haar o db1, detallada en la Ecuación 5.3, la cual es la primera y la más sencilla de las *wavelets*.

La *wavelet* de Haar es una *wavelet* ortonormal. Se define la función en $L^2(\mathbb{R})$

$$\psi(t) = \begin{cases} 1 & \text{si } t \in [0, \frac{1}{2}) \\ -1 & \text{si } t \in [\frac{1}{2}, 1) \\ 0 & \text{en otro caso} \end{cases} \quad (5.3)$$

A pesar de tener soporte compacto no tiene buena localización tiempo-frecuencia, además no es continua y por tanto es no diferenciable.

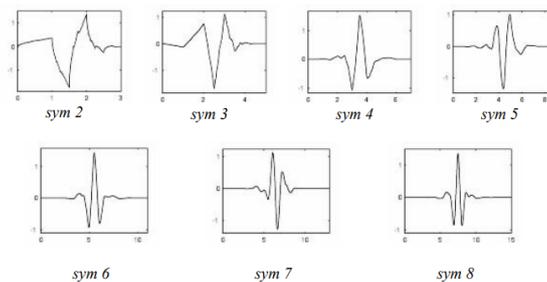
Las gráficas de las distintas *wavelets* de Daubechies se muestran en la Figura 5.1.

Figura 5.1: *Wavelet* Daubechies: dbN

5.3.2. *Wavelet* Symmlets

Esta familia de *wavelets* tiene características muy similares a las Daubechies. Se suelen denotar, symN.

Mayoritariamente son funciones asimétricas. Todas cumplen las condiciones de ortogonalidad y biortogonalidad permitiendo realizar tanto transformadas continuas como discretas. La anchura de la ventana es de $2N - 1$ donde N es el orden.

Figura 5.2: *Wavelet* symmlets: symN

5.3.3. Familia Coiflets

La anchura de estas funciones es de $6N - 1$ y van siendo más simétricas que las anteriores. son ortogonales y biortogonales y se pueden utilizar tanto para la representación de CWT como DWT.

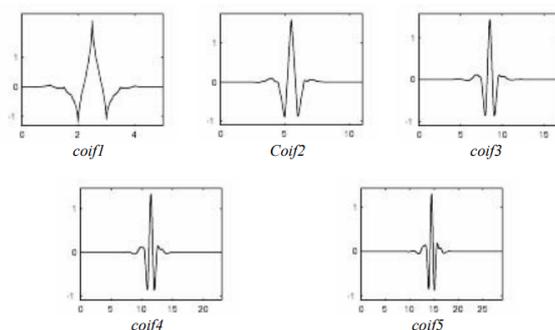


Figura 5.3: Wavelet coiflets: coifN

5.3.4. Familia de las Biortogonales

Se caracterizan por tener soporte compacto y simetría. Además, como su nombre indica son biortogonales.

Por otro lado, la reconstrucción exacta de la señal es posible gracias a los filtros *finite Impulse Response*, lo cual en las *wavelets* ortogonales no es posible de realizar a excepción de la *wavelet* de Haar. Por ello, cuentan con dos representaciones, una para la descomposición de la señal y otra para la reconstrucción. Este tipo de *wavelets* admiten representaciones con DWT y CWT

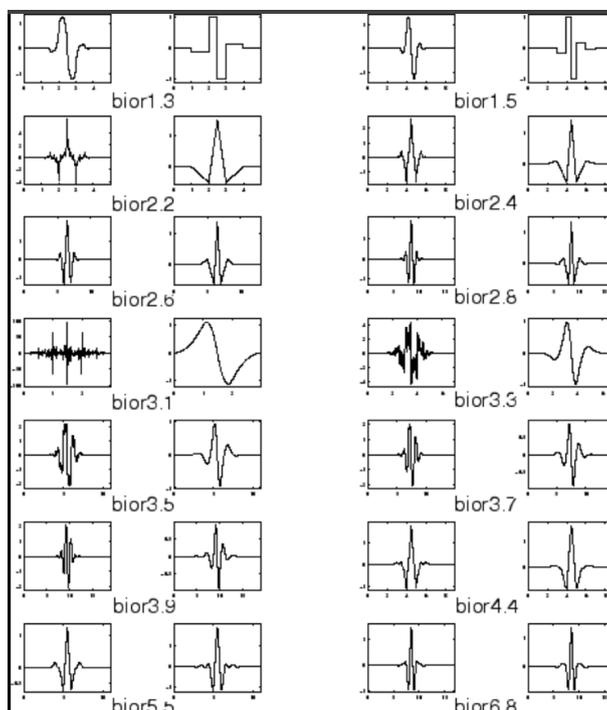


Figura 5.4: Wavelet biortogonales: bior

5.3.5. *Wavelet* Morlet

La *wavelet* de Morlet es introducida por Jean Morlet alrededor del año 1980 y es una de las *wavelets* más utiliza en el análisis de señales. Su expresión se puede reducir a la siguiente fórmula:

$$\psi(t) = C e^{\frac{-t}{2}} \cos(w_0 t) \quad (5.4)$$

donde la constante C se utiliza para la normalización de la gráfica y w_0 es la frecuencia base. La *wavelet* de Morlet no es ni ortogonal ni biortogonal, pero sí es simétrica. Por otro lado, no admite representación con la DWT pero sí con CWT.

Además, existe la *wavelet* de Morlet compleja dada por

$$\psi(t) = \frac{1}{\sqrt{B\pi}} e^{-\frac{t^2}{B}} e^{i2Ct\pi} \quad (5.5)$$

siendo B el ancho de banda y C la frecuencia central. La cual combina una onda sinusoidal compleja y una envolvente gaussiana y se representa tal y como se muestra en la Figura 5.5.

- **Onda sinusoidal compleja:**

$$e^{i2Ct\pi}$$

Esta parte de la ecuación introduce la oscilación sinusoidal compleja en la *wavelet*. La variable C controla la frecuencia de la oscilación, es decir, moviendo los valores de C se puede centrar la *wavelet* en diferentes frecuencias.

- **Envolvente gaussiana:**

$$\frac{1}{\sqrt{B\pi}} e^{-\frac{t^2}{B}}$$

Esta parte de la ecuación es una gaussiana normalizada que define cómo la *wavelet* está localizada en el tiempo. La variable B sirve para controlar la anchura de la gaussiana, es decir, la dispersión temporal de la *wavelet*. Para valores grandes de B implica una mayor dispersión temporal, es decir, mayor resolución en frecuencia pero menor en tiempo, mientras que para valores pequeños de B se consigue una mejor resolución temporal.

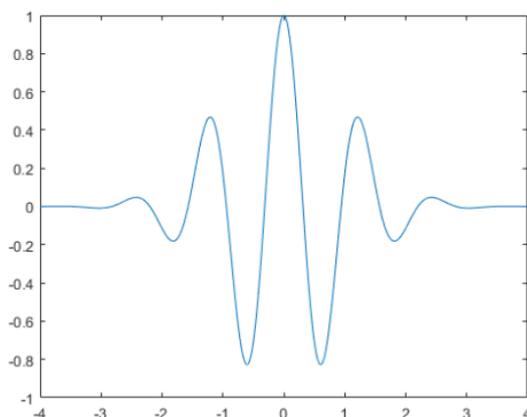


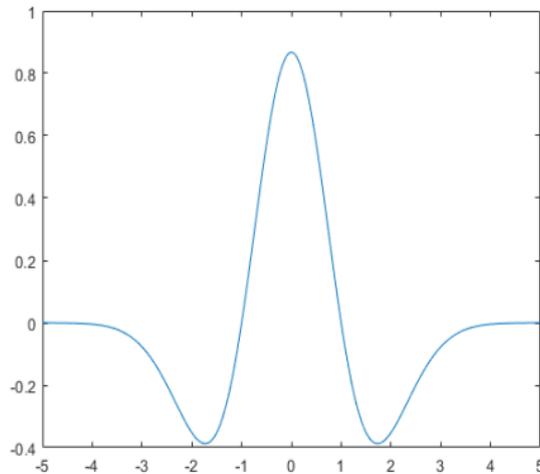
Figura 5.5: Representación gráfica de la *wavelet* de Morlet

5.3.6. *Wavelet* Sombrero Mexicano

La *wavelet* sombrero mexicano, también denominada *wavelet* de Ricker, es una función de gran utilidad para el análisis de señales, especialmente para la sismología y la electrodinámica computacional. Esta función se define como la segunda derivada normalizada negativa de una función gaussiana (Ecuación 5.8), por eso tiene una forma característica de campana invertida, con un pico central y oscilaciones que decrecen a medida que se alejan del centro (Figura 5.6). Al ser una derivada de una función gaussiana está vinculada con la segunda función de Hermite siendo un caso especial de *wavelets* continuas o *wavelets* hermíticas (Sección 5.3.7). La normalización permite asegurar que su energía es constante, independientemente de la escala utilizada, lo cual es importantísimo para poder comparar señales en diferentes resoluciones o frecuencias sin que la escala afecte al análisis.

Su forma particular la hace especialmente útil para la detectar picos y transiciones en los datos, permitiendo una localización precisa de eventos o características específicas en una señal. Por ejemplo, para la sismología, la *wavelet* sombrero mexicano se utiliza principalmente para modelar datos sísmicos, puesto que su forma se asimila a la respuesta sísmica generada por una explosión puntual, por tanto, es una herramienta de gran utilidad para la simulación de la propagación de ondas en el subsuelo. Además, tiene la capacidad de capturar tanto los aspectos de alta como de baja frecuencia de una señal lo cual es un factor muy importante para el análisis e interpretación de datos sísmicos.

Por otra parte, al ser un caso especial de *wavelet* continua su uso es común en el análisis de multirresolución, técnica que permite descomponer señales en diferentes escalas, especialmente de utilidad para el procesamiento de imágenes y la detección de bordes.

Figura 5.6: Representación gráfica de la *wavelet* sombrero mexicano

5.3.7. Wavelet hermíticas

Las *wavelet* hermíticas son una familia de *wavelets* discretas y continuas utilizadas en la transformada *wavelet* de Hermite continua y discreta. La ondícula hermítica se define como la n -ésima derivada de una distribución gaussiana para cada n positivo.

$$\Psi_n(x) = (2n)^{-\frac{n}{2}} c_n H_n(x) e^{-\frac{1}{2}x^2} \quad (5.6)$$

donde $H_n(x)$ es el polinomio de Hermite y $\Psi \in L_{\rho,\mu}(-\infty, \infty)$.

El coeficiente de normalización c_n viene dado por:

$$c_n = (n^{\frac{1}{2}-n} \Gamma(n + \frac{1}{2}))^{-\frac{1}{2}} = (n^{\frac{1}{2}-n} \sqrt{\pi} 2^{-n} (2n-1)!)^{-\frac{1}{2}} \quad (5.7)$$

5.4. Wavelets en Python

A continuación se enumeraran y se proporciona información sobre las distintas familias de *wavelets* que se recogen en Python.

a) Sombrero Mexicano: Dada por

$$\psi(t) = \frac{2}{\sqrt{3}\sqrt[4]{\pi}} e^{-\frac{t^2}{2}} (1-t^2) \quad (5.8)$$

b) Morlet: Dada por

$$\psi(t) = \cos(5t) e^{-\frac{t^2}{2}} \quad (5.9)$$

c) Morlet compleja: Dada por

$$\psi(t) = \frac{1}{\sqrt{B\pi}} e^{-\frac{t^2}{B}} e^{i2Ct\pi} \quad (5.10)$$

siendo B el ancho de banda y C la frecuencia central.

d) Onda derivadas gaussianas: Son las derivadas de orden P de la función

$$\psi(t) = Ce^{-t^2} \quad (5.11)$$

siendo C una constante de normalización depiente del orden.

e) Onda derivados gaussianos complejos: Son las derivadas de orden P de la función

$$\psi(t) = Ce^{-it}e^{-t^2} \quad (5.12)$$

siendo C una constante de normalización depiente del orden.

f) Ondas de Shannon: Dadas por

$$\psi(t) = \sqrt{B} \frac{\text{sen}(Bt\pi)}{Bt\pi} e^{i2Ct\pi} \quad (5.13)$$

siendo B el ancho de banda y C la frecuencia central.

g) Frecuencia *Wavelets* B-Spline: Dada por

$$\psi(t) = \sqrt{B} \left(\frac{\text{sen}(B\pi \frac{t}{M})}{B\pi \frac{t}{M}} \right)^M e^{i2Ct\pi} \quad (5.14)$$

siendo M el orden de los spline, B el ancho de banda y C la frecuencia central.

La elección de una u otra dependerá del tipo de señal, así como el análisis o las conclusiones que se quiere realizar. Se pueden seguir dos criterios de elección:

- Buscar varias con forma similar a la señal a analizar.
- Realizar distintas pruebas eligiendo la que mejor resultados nos aporte.

Para entender de forma sencilla como funciona, pensaremos de la siguiente forma, una vez seleccionada una *wavelet* con una escala particular, la deslizamos por toda la señal, es decir, cambiamos de forma continua la ubicación. En cada paso se multiplica la señal y la *wavelet*, el resultado es un coeficiente para dicha escala en ese paso de tiempo particular. Después se va aumentando la escala de las ondas y se repite el proceso.

La *wavelet* elegida recibe el nombre de *wavelet* madre, la cual se expandirá o comprimirá y se trasladará por la señal original.

Definición 24. Una *wavelet madre* es una función localizada, perteneciente al espacio $L^2(\mathbb{R})$ de manera que tiene energía infinita y el cuadrado es integrable, es decir,

$$f \in L^2(\mathbb{R}) \Rightarrow \int |f(t)|^2 = E < \infty \quad (5.15)$$

5.4.1. Wavelets utilizadas en el estudio

Los dos tipos de wavelet utilizados en este trabajo son la *wavelet* sombrero mexicano y la *wavelet* de Morlet, cuyas funciones matemáticas están descritas en la Sección 5.4. Las principales diferencias se encuentran en la siguiente Tabla 5.1.

Tabla 5.1: Diferencias entre wavelet Sombrero Mexicano y wavelet de Morlet

| | Sombrero Mexicano | Morlet |
|-------------------------------------|---|---|
| Frecuencia | <ul style="list-style-type: none"> ▪ Banda ancha ▪ Banda central no claramente definida | <ul style="list-style-type: none"> ▪ Buena resolución en frecuencia (sinusoidal modulada) ▪ Banda central clara |
| Resolución Tiempo-Frecuencia | Mayor concentración en el dominio del tiempo pero menos en el dominio de frecuencia | Equilibrio entre tiempo y frecuencia |
| Simetría | Real y simétrica | Compleja y no necesariamente simétrica |

5.5. Transformada Wavelet Continua (CWT)

La Transformada Wavelet Continua (CWT) expresa una señal $x(t)$ continua en el tiempo, a través de una expansión de términos o coeficientes proporcionales al producto interno entre la señal y las diferentes versiones escaladas y transformadas de la *wavelet* madre, $\psi(t)$.

Debido a que esta función es infinita, la CWT puede expresarse de la siguiente manera:

$$CWT(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (5.16)$$

Expresada en términos de la transformada de Fourier, Ecuación 5.17

$$CWT(a, b) = \frac{1}{2\pi\sqrt{|a|}} \int_{-\infty}^{\infty} X(w) \psi(aw) e^{-iwt} dt \quad (5.17)$$

La variable a da la escala, es decir, domina el ancho de banda y la variable b corresponde a la traslación por lo que da el dominio del tiempo. La variable de escala permite comprimir o dilatar la señal e indica el grado de resolución utilizado para el análisis de la señal. Para escalas pequeñas se tendrá una buena resolución en el dominio del tiempo mientras que

para escalas grandes se tendrá una buena resolución en el dominio de frecuencia. Esta es una de las principales diferencias entre la Transformada Wavelet Continua (CWT) y la Transformada de Fourier, ya que la CWT tiene ventanas de corta (largas) duración para altas (bajas) frecuencias (Figura 5.7) mientras que la Transformada de Fourier tiene una única ventana con la misma duración independientemente de la frecuencia (Figura 5.8) [17].

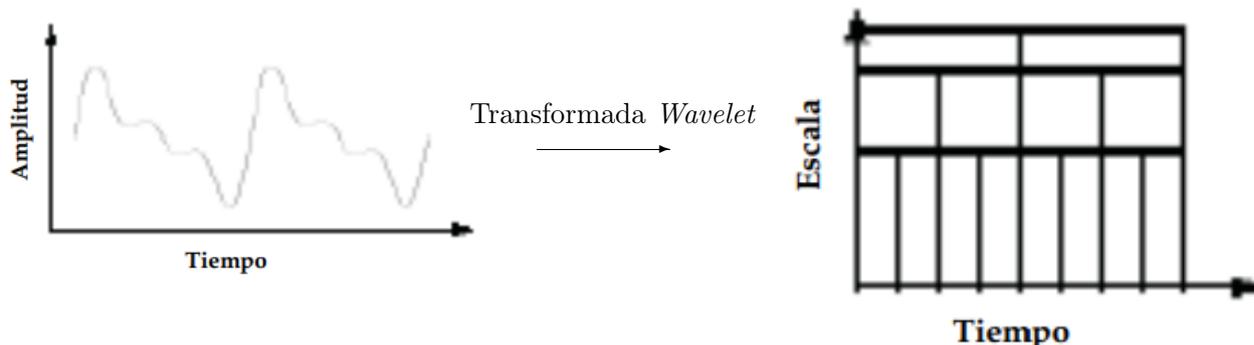


Figura 5.7: Esquema de la transformación mediante la transformada *wavelet*.

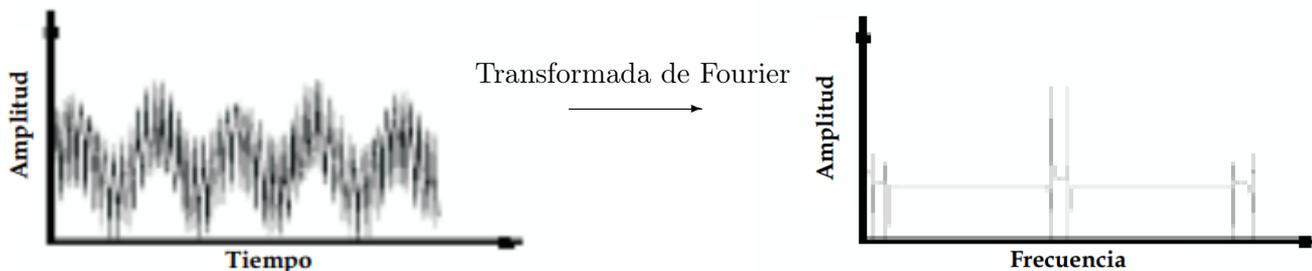


Figura 5.8: Esquema de la transformación mediante la transformada de Fourier.

Además, la función *wavelet* madre debe cumplir la siguiente condición, $\psi(0) = 0$, lo que nos lleva a pensar que se trata de una onda.

Para obtener la transformada *wavelet* continua de una determinada señal se deben seguir los siguientes pasos de manera cualitativa [18]:

Paso 1: Inicializamos en escala y ubicación, $a = 1$ y $t = 0$. A continuación, se multiplican entre sí ambas señales y el resultado se integra sobre todo el espacio de tiempo. El resultado de la integral se multiplica por $\frac{1}{\sqrt{a}}$, con el objetivo de normalizar la energía y de este modo obtener una función transformada con la misma energía a cualquier

escala. El resultado final es la Transformada *Wavelet* en tiempo cero, $t = 0$, y $a = 1$, indica la correlación entre la *wavelet* y la señal original. En la Figura 5.9 observamos el resultado del primer paso.

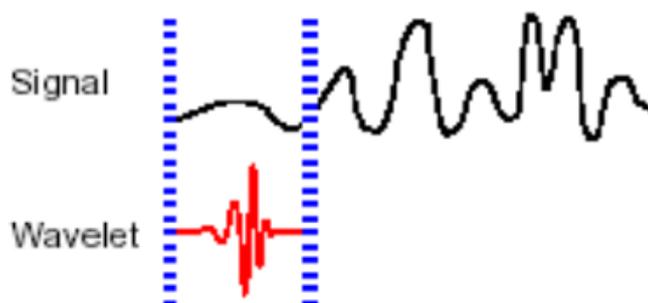


Figura 5.9: Paso 1 para obtener la Transformada *Wavelet* Continua

Paso 2: A continuación, la función *wavelet* se traslada en el tiempo (hacia la derecha) y se repite el procedimiento del Paso 1 hasta llegar al final de la señal, tal y como se muestra en la Figura 5.10

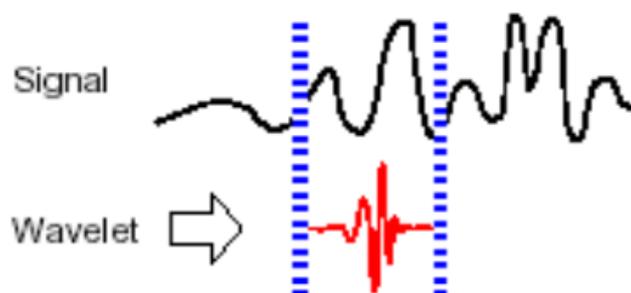


Figura 5.10: Paso 2 para obtener la Transformada *Wavelet* Continua

Paso 3: Ahora se cambia el valor de la escala, a , y se realizan los dos pasos anteriores hasta completar el rango de frecuencias (ver Figura 5.11). Debido a ser una transformación continua, tanto la transición en tiempo como las variaciones de escalas deben realizarse de manera continua. Sin embargo, si es necesario obtener la transformada *wavelet* mediante medios computacionales entonces la condición anterior se reduce a considerar un paso suficientemente pequeño.

Una vez se completen todos los valores de escala se obtendrá la Transformada *Wavelet* continua de la señal. Con respecto a la representación hay diversas opciones, por ejemplo tridimensional como en la Figura 5.12(a) o representación en un plano, en el eje x se representa el tiempo y en el eje y se representa la escala, el color de cada punto indica la magnitud de los coeficientes *Wavelet*, tal como se muestra en la Figura 5.12(b). En la Parte II utilizaremos la representación bidimensional.

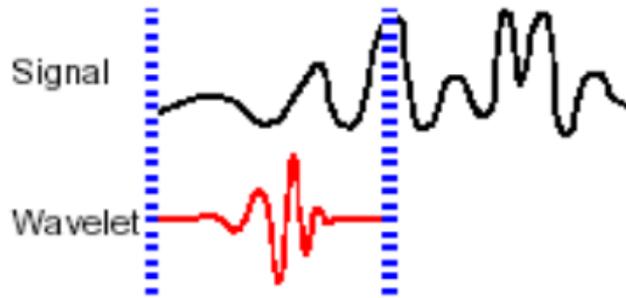
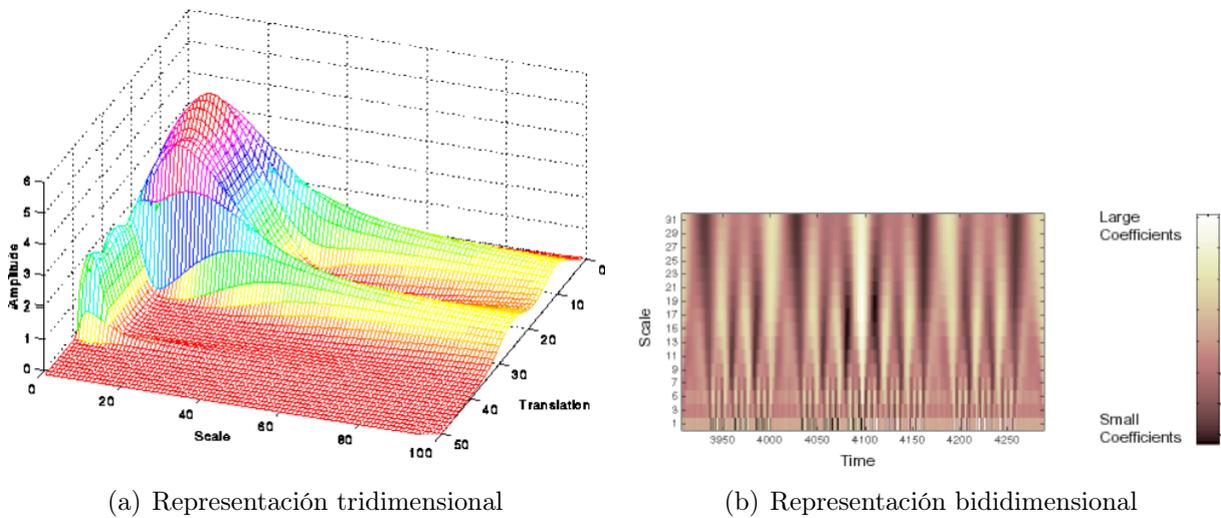


Figura 5.11: Paso 3 para obtener la Transformada *Wavelet* Continua



(a) Representación tridimensional

(b) Representación bidimensional

Figura 5.12: Representaciones de la Transformada *Wavelet* Continua

5.6. Análisis multiresolución (MRA)

Para poder realizar un Análisis multiresolución (MRA) es necesario tener una función base, $\phi = \phi(t)$ localizada tanto en tiempo como en frecuencia. Una de las más comunes es la función indicadora, $\phi = 1_{[0,1]}$ definida como:

$$1_{[0,1]} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } 0 \leq t < 1 \\ 0 & \text{si } \textit{resto} \end{cases} \quad (5.18)$$

La función 5.18 está localizada en 0, tiene escala unitaria debido a que es diferente de 0 en un intervalo de longitud 1. Dicha función puede ser trasladada, por ejemplo, al intervalo $[k, k + 1)$, mediante la siguiente fórmula:

$$\phi(t) = 1_{[0,1]} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } k \leq t < k + 1 \\ 0 & \text{si } \textit{resto} \end{cases} \quad k = \pm 1, \pm 2, \dots$$

Cualquier función, f , puede describirse con resolución unitaria:

$$f_0(t) = \sum_k a_k \phi_k(t) \quad (5.19)$$

La función, f_0 , es constante en el intervalo $[k, k+1)$ para cualquier k , donde toma el valor $a_k = \int f(t)\phi_k(t)dt$, el valor promedio de f en dicho intervalo. Si f no varía mucho en dicho intervalo entonces f_0 se aproxima a f (Figura 5.13).

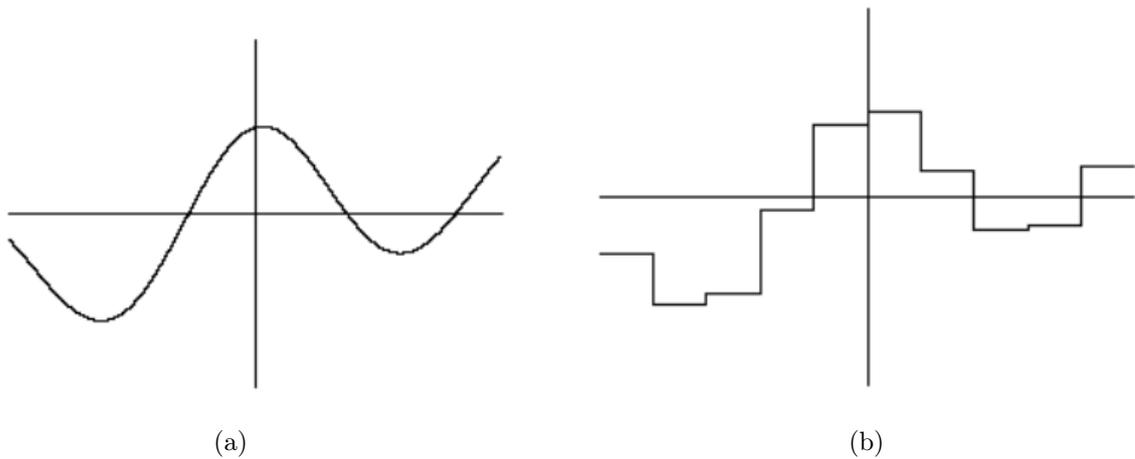


Figura 5.13: (a) Función original y (b) Representación mediante resolución unitaria

Para obtener una mejor aproximación de f se aumenta la resolución en la Ecuación 5.19 mediante una versión más pequeña de ϕ_k dilatando la variable t por un factor de 2:

$$\phi_{j,k}(t) \stackrel{\text{def}}{=} \phi(2^j t - k) \quad j = 1, 2, \dots \quad k = \pm 1, \pm 2, \dots$$

Con esta nueva notación se obtiene $\phi_{0,0} = \phi(2^0 t - 0) = \phi(t) = \phi$ y $\phi_{0,k} = \phi(2^0 t - k) = \phi(t - k) = \phi_k$. En la Figura 5.14 se muestran tres ejemplos de aproximaciones de f con $j = 0, 1, 2$.

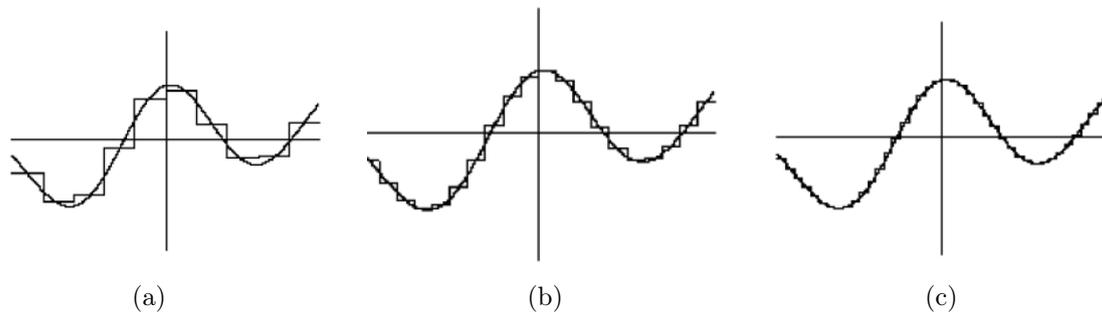


Figura 5.14: Tres ejemplos de aproximación a f : (a) Resolución unitaria f_0 , (b) Resolución doble f_1 y (c) Resolución cuádruple f_2 .

Para tener un MRA correcto f debe cumplir

$$\lim_{j \rightarrow \infty} |f - f_j| = 0 \quad (5.20)$$

Por tanto, esta propiedad depende de la elección de ϕ , debido a que ϕ determina a $f_j \quad \forall j = 1, 2, \dots$. La función indicador cumple dicha propiedad, sin embargo, hay otras opciones para elegir ϕ que cuentan con mayor velocidad de aproximación. Una forma de generar dichas funciones es obtener la solución, usualmente denominada **función escala**, de :

$$\phi(t) = \sum_k h_k \phi(2t - k) \quad (5.21)$$

$$\int \phi(t) dt = 1 \quad (5.22)$$

La solución viene determinada por los números $\{h_k\}$ denominados coeficientes del filtro pasabajo. Se obtiene una solución única para ϕ si se satisfacen las siguientes condiciones:

$\{h_k\}$ tiene un número finito de coeficientes diferentes de cero.

$$\begin{aligned} \sum_k h_k &= 2 \\ \sum_k h_k h_{k+2n} &= 0 \quad n \neq 0 \\ \sum_k h_k^2 &= 2 \end{aligned} \quad (5.23)$$

Ejemplo: La función indicador satisface la Ecuación 5.22 y cumple la Ecuación 5.21 para ello analicemos como se comporta la función indicadora al aplicar transformaciones de escalado y traslación.

a) Para $k = 0$, la función

$$1_{[0,1)} = \begin{cases} 1 & \text{si } 0 \leq t < \frac{1}{2} \\ 0 & \text{si } \textit{resto} \end{cases}$$

b) Para $k = 1$, la función

$$1_{[0,1)} = \begin{cases} 1 & \text{si } \frac{1}{2} \leq t < 1 \\ 0 & \text{si } \textit{resto} \end{cases}$$

Dado que $1_{[0,1)}$ cubre el intervalo $[0, 1)$ y la combinación de $1_{[0,1)}(2t)$ y $1_{[0,1)}(2t-1)$ cubre el mismo intervalo y además para otros valores de k , la función indicadora no contribuye porque las funciones $1_{[0,1)}(2t-k)$ se desplazan fuera del intervalo $[0, 1)$. Por tanto, se tiene lo siguiente:

$$1_{[0,1)}(t) = h_0 1_{[0,1)}(2t) + h_1 1_{[0,1)}(2t-1) + \sum_{k=2} h_k 1_{[0,1)}(2t-k) = \begin{cases} h_0 & \text{si } 0 \leq t < \frac{1}{2} \\ h_1 & \text{si } \frac{1}{2} \leq t < 1 \\ \sum_{k=2} h_k 1_{[0,1)}(2t-k) & \text{si } \textit{resto} \end{cases}$$

Dando lugar a $h_0 = h_1 = 1$ y $h_k = 0 \quad \forall k \geq 2$ que cumplen las condiciones 5.23

Cualquier secuencias que cumplan las tres últimas condiciones de 5.23 da lugar a una función ϕ que es ortogonal a sus traslaciones enteras:

$$\int \phi(t)\phi(t-k)dt = 0 \text{ con } k \in \mathbb{Z}, k \neq 0$$

$$\int \phi^2(t)dt = 1$$

En estos casos, la fórmula que expande la función f para una resolución j es similar a la fórmula en series de Fourier

$$f_j(t) = \sum_k a_k \phi_{j,k}(t)$$

con $a_j = 2^j \int \phi_{j,k}(t)dt$. El factor 2^j es necesario debido a la reducción de la función ϕ en la resolución j . Sin embargo, dicha expansión difiere de la lista de muestras de f en puntos separados 2^{-j} , lo que resulta en una baja eficiencia de compresión. Hay una forma de codificar solamente las diferencias entre las aproximaciones sucesivas de f .

$$f = f_0 + (f_1 - f_0) + (f_2 - f_1) + (f_3 - f_2) + \dots$$

Esta codificación funciona debido a 5.20 de forma que los coeficientes necesarios para codificar $f_j - f_{j-1}$ deben ir siendo más pequeños a medida que aumenta j . Además, el número de coeficientes necesarios para codificar $f_j - f_{j-1}$ es proporcional a 2^j .

El “truco” es expandir las diferencias como superposición de otra función:

$$f(t) = \sum_k c_{0,k} \phi_{0,k}(t) + \sum_k d_{0,k} \psi_{0,k}(t) + \sum_k d_{1,k} \psi_{1,k}(t) + \sum_k d_{2,k} \psi_{2,k}(t) + \dots \quad (5.24)$$

donde $c_{0,k}$ y $d_{j,k}$, $j = 0, 1, 2, \dots$ son amplitudes que se determinan para cada función f . Dicha fórmula es válida gracias a que ϕ cumple las Ecuaciones 5.21 y 5.22. Como resultado de esto aparece la función *wavlet* madre, que genera las componentes de la función 5.24 mediante dilataciones y desplazamientos:

$$\psi(t) = \sum_k g_k \psi(2t - k)$$

con $g_k = (-1)^k h_{1-k}(t)$, $k = 0, \pm 1, \pm 2, \dots$

Las dos secuencias $\{h_k\}$ y $\{g_k\}$ se denominan filtros de cuadratura conjugados mediante los cuales es posible encontrar eficientemente los coeficientes $c_{0,k}$ y $d_{j,k}$ de la ecuación 5.24

5.7. Transformada Wavelet Discreta (DWT)

La DWT se obtiene al discretizar los parámetros de desplazamiento y escalamiento dentro de la transformada wavelet continua. Normalmente, los valores que se implementa para realizarlo son [19][20]:

$$\begin{aligned} a &= 2^{-j} \\ b &= k2^{-j} \end{aligned}$$

siendo a el escalamiento, b el desplazamiento y $j, k \in \mathbb{Z}$

Una vez se han definido los parámetros de escala y desplazamiento con valores discretos, la wavelet madre es:

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k) \quad (5.25)$$

con $j, k \in \mathbb{Z}$

Dando lugar a la Transformada Wavelet Discreta (DWT) como:

$$DWT_{j,k} = 2^{\frac{j}{2}} \int_{-\infty}^{\infty} x(t) \psi(2^j t - k) dt \quad (5.26)$$

donde $x(t)$ es una señal discreta.

La DWT permite reconstruir la señal una vez se han calculado los coeficientes wavelet. Para realizar dicha reconstrucción se necesitan dos funciones:

- **Función escala:** $\phi(t)$
- **Función wavelet:** $\psi(t)$

y se obtiene siguiendo la siguiente ecuación:

$$x(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t) \quad (5.27)$$

con $j, k \in \mathbb{Z}$, $c_{j,k}$ son los coeficientes de escala o de aproximación, asociados con la función escala y $d_{j,k}$ los coeficientes wavelet o de detalle, relacionados con la función wavelet.

5.7.1. Función escala

Sea $\phi(t) \in L^2(\mathbb{R})$ una función escala, que trasladada y escalada genera una familia de funciones definidas como [21][20]:

$$\phi_{j,k}(t) = 2^{\frac{j}{2}} \phi(2^j t - k) \quad j, k \in \mathbb{Z}$$

Entonces una función $f(t)$ se puede escribir como

$$f(t) = \sum_{k \in \mathbb{Z}} c_{j,k} \phi_{j,k}(t) \quad (5.28)$$

Multiplicando por $\phi_{j,k}(t)$ en la ecuación 5.28 se obtiene:

$$\begin{aligned} \langle f(t), \phi_{j,k}(t) \rangle = & \langle \sum_{k \in \mathbb{Z}} c_{j,k} \phi_{j,k}(t), \phi_{j,k}(t) \rangle = c_{j,-\infty} \langle \phi_{j,-\infty}(t), \phi_{j,k}(t) \rangle + \dots + \\ & c_{j,k} \langle \phi_{j,k}(t), \phi_{j,k}(t) \rangle + \dots + c_{j,\infty} \langle \phi_{j,\infty}(t), \phi_{j,k}(t) \rangle \end{aligned}$$

Puesto que la función escala debe ser ortonormal, todos los productos anteriores son nulos salvo uno, dando lugar a.

$$\int_{-\infty}^{+\infty} |f(t)\phi_{j,k}(t)| dt = \langle f(t), \phi_{j,k}(t) \rangle = c_{j,k} \langle \phi_{j,k}, \phi_{j,k} \rangle = c_{j,k}$$

Por tanto, se tiene lo siguiente:

- a) Para cada $j \in \mathbb{Z}$, $\{\phi_{j,k}/k \in \mathbb{Z}\}$ forma una base ortonormal para el subespacio $V_j \subset L^2(\mathbb{R})$, siendo V_j los espacios generados por las funciones escalas.
- b) Los subespacios V_j están anidados, es decir, $V_j \subset V_{j+1} \forall j \in \mathbb{Z}$
- c) La función $\phi(t)$ tiene soporte compacto. Es decir, existe un subconjunto del dominio de $\phi(t)$ donde esta no es nula.

5.7.2. Función wavelet

Se define $W_j = \overline{\text{Span}_{k \in \mathbb{Z}}\{\psi_{j,k}(t)\}}$ como el complementario ortogonal de V_j en V_{j+1} , es decir, todos los elementos de V_j son ortogonales a todos los elementos de W_j . Donde Span indica el espacio vectorial generado por.

$$\begin{aligned} \langle \phi_{j,k}(t), \psi_{j,k}(t) \rangle = & \int_{-\infty}^{+\infty} \phi_{j,k}(t)\psi_{j,k}(t) dt = 0 \quad j, k, l \in \mathbb{Z} \\ \psi_{j,k}(t) = & 2^{\frac{j}{2}}\psi(2^j t - k) \quad j, k \in \mathbb{Z} \\ & V_1 = V_0 \oplus W_0 \\ f(t) = & \sum_{k \in \mathbb{Z}} d_{j,k}\psi_{j,k}(t) \end{aligned}$$

de este modo la función *wavelet madre*, $\psi(t)$, se puede representar como

$$\psi(t) = \sum_{k \in \mathbb{Z}} \sqrt{2}h(n)\phi(2t - k)$$

5.7.3. Aplicaciones

La DWT se aplica en ciencia, ingeniería, matemáticas e informática. Particularmente, para la codificación de señales, compresión de datos o para algunas aplicaciones prácticas en el procesamiento de señales de aceleraciones para el análisis de marcha, procesamiento de imágenes, en comunicaciones digitales, entre otras cosas.

Además, se usa como banco de filtros analógicos en el procesamiento de señales bio-médicas para el diseño de marcapasos de bajo consumo.

Parte II

Experimentación Práctica

Capítulo 6

Datos utilizados

Los datos utilizados se basan en el Índice de Precipitación Estandarizado (SPI), el cual representa el número de desviaciones estándar de las precipitaciones a lo largo de un determinado periodo respecto de la media, antes se debe transformar los datos originales de precipitación a una distribución normal. Por tanto, el SPI sirve para cuantificar la intensidad de sequía o la escasez de precipitaciones. Su historia empieza en el año 1993 gracias al estadounidense MC Kee con el fin de poder evaluar el impacto de la escasez de precipitaciones en diferentes tiempos y tomar medidas en los distintos recursos hídricos [22].

Interpretar la magnitud de precipitaciones es un desafío puesto que la climatología varía según la zona geográfica y la época del año. Para evaluar el SPI se define una escala de valores, los valores positivos indican un periodo húmedo mientras que valores negativos indican un periodo seco, las distintas categorías se presentan en la Tabla 6.1. Además, se suele indicar un límite para poder expresar épocas extremadamente húmedas o secas. El SPI compara el total de precipitación para un intervalo determinado con la distribución histórica de precipitación para dicho intervalo.

Cada valor de SPI se calcula separado para una selección de escalas de tiempo como las siguientes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 18, 24, 30, 36, 48, 60 y 72 meses. Las escalas de 1-3 meses reciben el nombre de escalas temporales cortas, donde el SPI está ligado a la humedad del suelo, para escalas mayores a 6 meses el índice se asocia con el agua subterránea o el nivel de agua embalsada. Los índices para cada escala temporal se basan en las series de precipitaciones acumuladas para esa escala. El SPI de X meses muestra una comparación de la precipitación a lo largo de un periodo específico de X meses con las precipitaciones totales del mismo periodo de X meses de todos los años incluidos en el registro. Por ejemplo, para calcular un índice de 3 meses para abril de 2023, se considera la suma de los valores mensuales de precipitación desde febrero de 2023 y se compara con los totales de precipitación entre febrero y abril de todos los años presentes en el registro de la localidad que se quiere estudiar. De la misma manera, si la escala es de 6 meses para abril de 2023, se considera la suma de los valores mensuales desde octubre de 2022 y se compara con todos los totales de precipitación registrados en ese mismo periodo. Por tanto, para un índice de 12 meses se considera la suma de los

valores mensuales de precipitación de un año completo y se compara con los totales de precipitación del resto de años del registro [23] [24].

Para estimar el valor del índice en cada estación meteorológica se siguen los siguientes pasos:

- a) Calcular el total de precipitaciones para todas las escalas temporales elegidas.
- b) Ajustar una distribución de probabilidad paramétrica o no paramétrica a los totales calculados anteriormente para la escala correspondiente. Habitualmente los datos tienen una distribución asimétrica, es decir, la mayor parte de los totales son valores bajos o intermedios y existe una probabilidad decreciente de mayores precipitaciones, por ello la distribución gamma o la de Pearson III son las más utilizadas ya que cuentan con un sesgo positivo, es decir, cuentan con una gran cola en la parte derecha. Dichas distribuciones se definen a partir de dos o tres parámetros estimados en el proceso de ajuste. Guttman (1999) recomienda un mínimo de 50 años para alcanzar precisión computacional.
- c) Una vez se tiene la distribución ajustada se transforma a una distribución normal o gaussiana.
- d) Por último, se interpreta los valores obtenidos. Normalmente los valores de SPI se encuentran entre -3 y 3, ya que este rango incluye el 99.7% de la densidad de una distribución normal con media 0 y desviación estándar 1.

Tabla 6.1: Categorías de los valores SPI

| Categoría | SPI |
|-----------------------|------------------------|
| Sequía extrema | $SPI \leq -1,5$ |
| Sequía severa | $-1,5 < SPI \leq -1,0$ |
| Sequía moderada | $-1,0 < SPI \leq -0,5$ |
| Normal | $-0,5 < SPI < 0,5$ |
| Moderadamente húmedo | $0,5 \leq SPI < 1,5$ |
| Severamente húmedo | $1,0 \leq SPI < 1,5$ |
| Extremadamente húmedo | $SPI \geq 1,5$ |

Las principales ventajas del SPI son:

- a) Basado únicamente en las precipitaciones, es decir, no se requiere ningún otro dato como puede ser la localización o la temperatura de esta.
- b) Calculado en diferentes escalas temporales, mencionadas anteriormente, asociadas a distintos tipos de recursos hídricos como son la humedad del suelo, caudal de ríos o aguas subterráneas.

- c) Permite comparar índices de distintas regiones con características climáticas muy diferentes. En el caso de la península ibérica pueden ser la comparación del norte y el sur.
- d) Cálculo más sencillo que otros índices como puede ser el Índice Palmer de Severidad de Sequía (PDSI).

y sus principales desventajas son:

- a) Al ser un indicador únicamente de precipitación limita su capacidad de analizar el efecto en temperaturas elevadas asociadas al cambio climático.
- b) Los valores pueden verse afectados debido a la longitud y la confiabilidad de las series de precipitaciones necesarias para el ajuste de la distribución en el cálculo, por ello se recomienda usar series de 30 a 50 años.

En la Figura 6.1 se muestra en línea naranja una serie SPI hipotética. La línea de puntos representa el umbral por el cual se considera sequía moderada. También se representa un área rayada entre el inicio y el fin del evento, la magnitud que es el producto entre la duración y la intensidad.

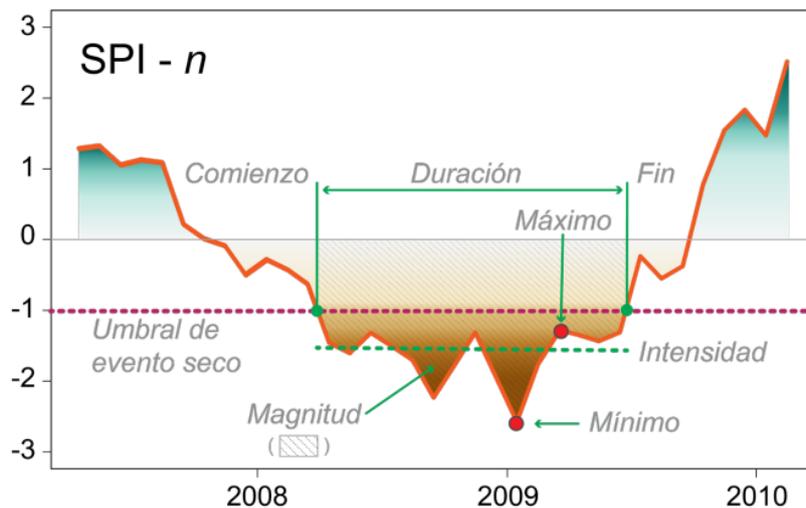


Figura 6.1: Esquema de un evento de sequía [23].

La base de datos de la cual se extraen los datos para este estudio contiene valores SPI que abarcan desde 1 mes hasta 36 meses, en el periodo comprendido entre enero de 1961 y enero de 2024 para 75 puntos distintos de España entre ciudades, comunidades autónomas y ríos teniendo un total de $757 \text{ (filas)} \times 36 \text{ (columnas con valores SPI)} \times 75 \text{ (número de tablas)} = 2043900$ datos. Para cada punto de España, se cuenta con una tabla (757 filas y 38 columnas) que recoge todos estos valores, cada una identificada con

un código numérico. Además, se dispone de una lista que relaciona dichos códigos con su respectiva región.

De toda esta información, se extrae la base de datos que se utilizará en este trabajo, la cual incluye únicamente los valores SPI de las 52 ciudades de la península ibérica (Tabla 6.2) desde enero de 1961 hasta enero de 2024. El tratamiento de estos datos se explica en la Sección 7.2.

Tabla 6.2: Códigos de las distintas regiones y sus coordenadas

| Código | Región | Latitud | Longitud |
|---------------|---------------|----------------|-----------------|
| 1 | A Coruña | 43,37135 | -8,396 |
| 2 | Albacete | 38,99424 | -1,85643 |
| 3 | Alicante | 38,34517 | -0,48149 |
| 4 | Almería | 36,83814 | -2,45974 |
| 5 | Álava | 42,84998 | -2,67268 |
| 6 | Asturias | 43,36029 | -5,84476 |
| 7 | Ávila | 40,65724 | -4,69951 |
| 8 | Badajoz | 38,87789 | -6,97061 |
| 10 | Barcelona | 41,38879 | 2,15899 |
| 11 | Bizkaia | 43,26271 | -2,92528 |
| 12 | Burgos | 42,34106 | -3,70184 |
| 13 | Cáceres | 39,47649 | -6,37224 |
| 14 | Cádiz | 36,52672 | -6,2891 |
| 15 | Cantabria | 43,46472 | -3,80444 |
| 16 | Castellón | 39,98567 | -0,04935 |
| 18 | Ciudad Real | 38,98626 | -3,92907 |
| 19 | Córdoba | 37,89155 | -4,77275 |
| 20 | Cuenca | 40,06667 | -2,13333 |
| 21 | Gipuzkoa | 43,31283 | -1,97499 |
| 22 | Girona | 41,98311 | 2,82493 |
| 23 | Granada | 37,18817 | -3,60667 |
| 24 | Guadalajara | 40,62862 | -3,16185 |
| 25 | Huelva | 37,26638 | -6,94004 |
| 26 | Huesca | 42,13615 | -0,4087 |
| 27 | Jaen | 31,76922 | -3,79028 |
| 28 | La Rioja | 42,46667 | -2,45 |
| 30 | León | 42,60003 | -5,57032 |
| 31 | Lleida | 41,61674 | 0,62218 |
| 32 | Lugo | 43,00992 | -7,55602 |
| 33 | Madrid | 40,4165 | -3,70256 |
| 34 | Málaga | 36,72016 | -4,42034 |
| 35 | Melilla | 35,29369 | -2,93833 |
| 36 | Murcia | 37,98704 | -1,13004 |
| 37 | Navarra | 42,81687 | -1,64323 |
| 38 | Ourense | 42,33669 | -7,86407 |
| 39 | Palencia | 42,00955 | -4,52406 |
| 40 | Pontevedra | 42,431 | -8,64435 |
| 41 | Salamanca | 40,96882 | -5,66388 |
| 43 | Segovia | 40,94808 | -4,11839 |
| 44 | Sevilla | 37,38283 | -5,97317 |
| 45 | Soria | 41,76401 | -2,46883 |
| 46 | Tarragona | 41,11905 | 1,24544 |
| 47 | Teruel | 40,3456 | -1,10646 |
| 48 | Toledo | 39,8581 | -4,02263 |
| 49 | Valencia | 39,47391 | -0,37966 |
| 50 | Valladolid | 41,65518 | -4,72372 |
| 51 | Zamora | 41,50633 | -5,74456 |
| 52 | Zaragoza | 41,65606 | -0,87734 |

Capítulo 7

Herramientas tecnológicas y procedimiento metodológico

En este capítulo se describen las distintas herramientas tecnológicas utilizadas y el procedimiento metodológico llevado a cabo para el estudio práctico.

7.1. Herramientas tecnológicas

Para llevar a cabo la experimentación práctica se han utilizado las siguientes herramientas computacionales:

- Visual Studio Code: Aplicación disponible para Windows, Linux o macOS. Permite editar código fuente para casi cualquier tipo de lenguaje. En nuestro caso, será utilizado para crear un cuaderno en Jupyter Notebook en el lenguaje Python [25].
- Jupyter Notebook: Aplicación web utilizada para crear y compartir proyectos computacionales. Se basa en varias celdas ordenadas con entradas y salidas en formato JSON. Estas celdas pueden contener gráficos, textos y escribir en distintos lenguajes de programación. Se creó en el año 2014 con el fin de permitir crear software libre. Entre sus usos destaca principalmente la creación de modelos de aprendizaje automático y la modelización estadística. Durante el proyecto se han utilizado celdas de Python, para la programación, y Markdown, para escribir algunos títulos y organizar el código [26].

Además de las herramientas mencionadas anteriormente ha sido necesaria la siguiente tecnología.

- Python: Lenguaje de programación interpretado orientado a objetos de alto nivel y de código abierto, nació a finales de los años 80 creado por Guido van Rossum en los Países Bajos. En la actualidad, es considerado uno de los principales lenguajes de programación por su versatilidad. Se usa principalmente en análisis de datos y la

creación de modelos de aprendizaje automático gracias a la diversidad de bibliotecas, librerías y módulos que contiene [27][28].

Las bibliotecas utilizadas han sido:

- *NumPy*: Abreviatura de “Numerical Python”, biblioteca que incorpora funciones matemáticas de alto nivel para el procesamiento numérico y científico de datos. Principalmente es utilizado para crear matrices, trabajar con distintas formas lógicas y realizar operaciones algebraicas con mayor facilidad [29].
- *Pandas*: Estructura de datos de alto rendimiento y herramientas de análisis de datos, se basa en *NumPy*. Es fundamental para trabajar con datos de tipo *DataFrame*, pues facilita las operaciones de manipulación y análisis. Generalmente es utilizado para ciencia y análisis de datos [30].
- *Matplotlib*: Librería de Python utilizada para la representación de diversos gráficos a partir de listas o *arrays*. Entre sus gráficos los más comunes son histogramas, diagramas de barras, diagramas de color o mapas de color. Permite también realizar distintas combinaciones entre ellos [31][32].

7.2. Procedimiento metodológico

Siguiendo la base de datos proporcionada y descrita en el Capítulo 6 se crearán cuatro ficheros, cada uno correspondiente a un índice de agregación 3, 6, 9 y 12 meses. Dichos ficheros constan de 758 filas y 54 columnas. La primera fila son los nombres de las distintas columnas, de la segunda fila en adelante son los datos correspondientes al SPI seleccionado desde enero de 1961 hasta enero de 2024. Por otro lado, la primera y segunda columnas cuentan con el año y el mes de estudio. De la tercera columna hasta la 54 son las distintas ciudades de la península ibérica en orden alfabético.

Para seleccionar el índice de agregación de estudio, se visualizan las cuatro series temporales (SPI3, SPI6, SPI9 y SPI12) desde enero de 1962 hasta enero de 2024 en la ciudad de Valladolid (Figura 8.1). Permitiendo así comparar el comportamiento de los distintos índices de agregación y seleccionar el más adecuado. A su vez, para las distintas series temporales de la ciudad de Valladolid, según el índice de agregación, se visualizan dos tipos de *wavelet*: Morlet y sombrero mexicano (Figuras 8.2-8.5). Ambos tipos de *wavelet* serán comparadas y analizadas entre sí con el objetivo de seleccionar la más óptima.

Una vez se tiene el índice de agregación seleccionado, SPI12, se procede a reducir la dimensionalidad de los datos, para ello se utiliza el algoritmo *k-means* (Capítulo 3). Después, se aplicará el método del codo sobre todos los datos del SPI12 con el objetivo de conocer el número óptimo de *clusters* con los que trabajar. En la Figura 7.1(a) se muestra dónde se produce el cambio brusco en la inercia (codo), el cual se obtiene aproximadamente en 2 *clusters*. Sin embargo, en este trabajo se han seleccionado 6 *clusters* ya que optar por 2 *clusters* puede llevar a una pérdida significativa de información y una excesiva simplificación pues son insuficientes para capturar la diversidad regional de la

precipitación en la península ibérica. Por ello, en este trabajo se utiliza la variabilidad explicada acumulada (Figura 7.1(b)) en los distintos números de *clusters*. De esta manera se obtiene que el número óptimo es de 6 *clusters* con una variabilidad de 56,70 %, es decir, proporcionan una mayor representación y diferenciación de las características propias a las diferentes regiones de la península ibérica.

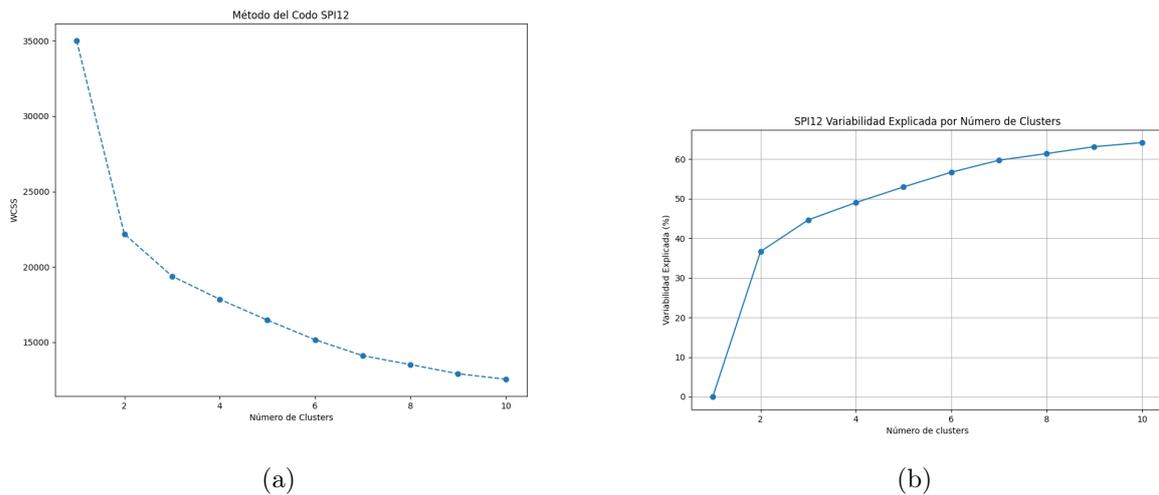


Figura 7.1: Método del codo práctico: (a) Método del codo SPI12 y (b) Variabilidad explicada en cada cluster SPI12.

La distribución de los datos en 6 *clusters* se muestran en la Figura 7.2.

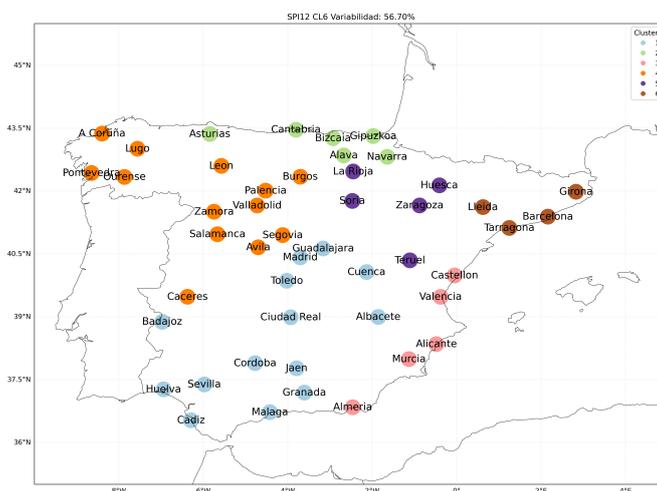


Figura 7.2: Mapa con la visualización de los datos agrupados en 6 clusters

Una vez se tienen los grupos, se pasa a calcular la media de cada grupo, obteniendo una serie temporal del SPI12 correspondiente a cada *cluster*. De cada una de las series

temporales se visualizarán la representación del SPI y la *wavelet* correspondiente, como se muestra en el Capítulo 8. Analizando los periodos que presentan déficit de precipitaciones (periodo seco) o superávit de precipitaciones (periodo húmedo).

Para completar el análisis de resultados de cada *cluster* se observan aquellos años de mayores déficit y superávit de precipitaciones registrados en la península ibérica. En la Figura 7.3 se representa la media anual de las precipitaciones totales (mm) en la España peninsular desde 1961 hasta 2022. La línea de color morado indica la media de precipitaciones sobre el periodo de referencia 1981-2010 [33]. Se puede observar que los periodos 1980-1983 y 1990-1996 son los periodos más largos por debajo de la media, que coinciden con los años más secos de España donde la sequía de 1980-1983 afectó especialmente al sur [34]. Destaca el año 2005 como el más seco de toda la serie temporal (450 mm) y el año 1963 como el más húmedo (900 mm).

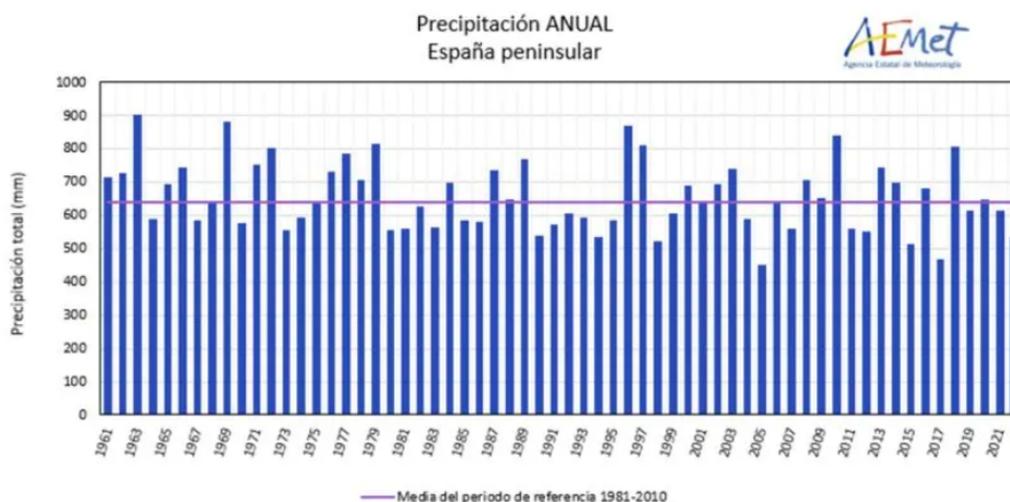


Figura 7.3: Precipitación anual de la península ibérica de 1961-2022 [33].

Capítulo 8

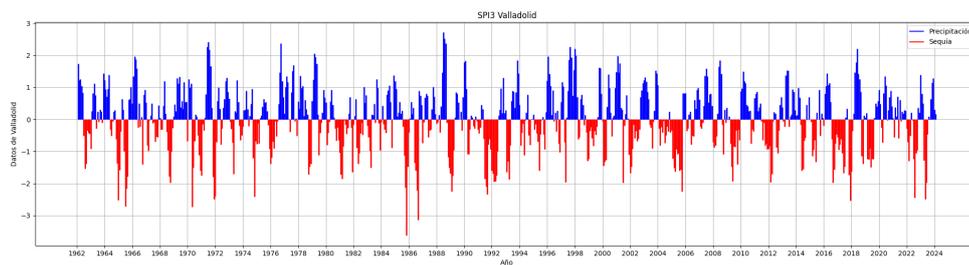
Resultados

En el siguiente capítulo se exponen los resultados obtenidos y sus interpretaciones, detallando la elección del índice de agregación del SPI y del tipo de wavelet.

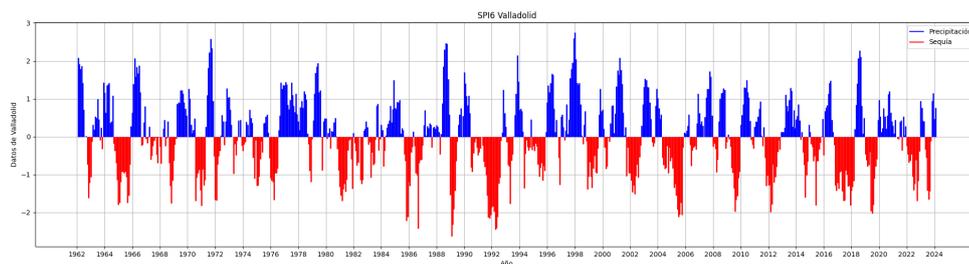
8.1. Selección índice de agregación del SPI

Para seleccionar el índice de agregación óptimo para este estudio, se representan gráficamente los distintos valores del SPI (SPI3, SPI6, SPI9 y SPI12) para una determinada ciudad, en este caso se ha escogido a modo de ejemplo Valladolid (Figura 8.1). En dicha figura se representa en color azul el superávit de precipitaciones y en color rojo el déficit de precipitaciones para los diferentes valores del SPI.

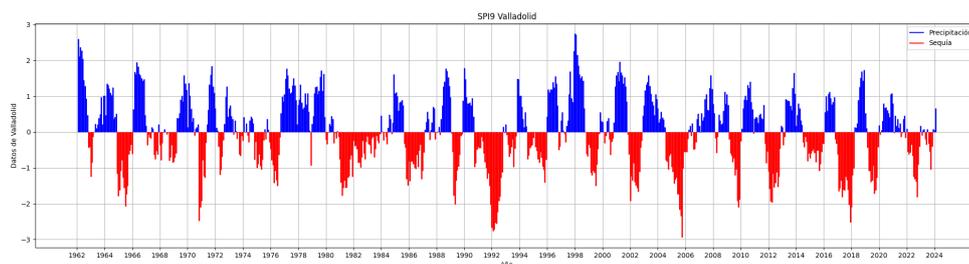
Se puede observar que cuanto mayor es el índice de agregación hay una mayor “suavidad” en la gráfica, es decir, se tiene menos “ruido” en los datos de mayor índice de agregación. Por ejemplo, si en un periodo seco (índice negativo) se registran dos meses consecutivos con precipitaciones abundantes, el SPI en las agregaciones más bajas (SPI3 y SPI6) pueden tomar valores positivos, en dichos meses, mientras que el SPI en las agregaciones más altas (SPI9 y SPI12) continúan reflejando condiciones de sequía, tal y como se puede ver en la Figura 8.1 en los años 1990-1993. Dado que uno de los objetivos de dicho trabajo es estudiar los grandes periodos de anomalías de precipitación desde 1962 hasta 2024, se han seleccionado los datos del SPI12, ya que el análisis con agregaciones más altas puede evitar conclusiones erróneas de que una sequía ha finalizado, cuando en realidad sólo ha ocurrido un periodo húmedo transitorio o viceversa [23].



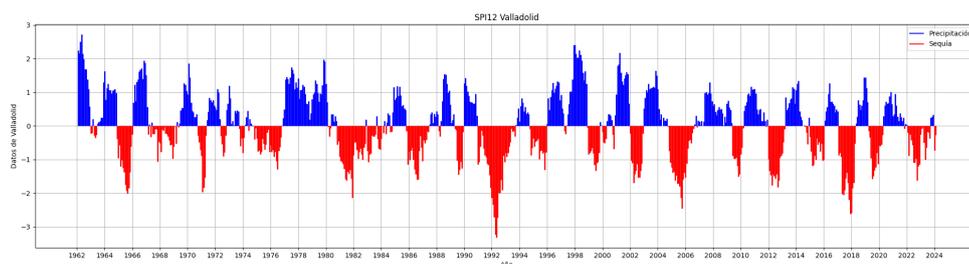
(a)



(b)



(c)



(d)

Figura 8.1: Representación de (a) SPI3, (b) SPI6, (c) SPI9 y (d) SPI12 de la ciudad de Valladolid

8.2. Selección del tipo de *wavelet*: Sombrero Mexicano vs Morlet

A modo de ejemplo, se ha representado las wavelets del tipo Mexicano (Mexh) y Morlet (Morl) para la ciudad de Valladolid en los distintos índices de agregación (Figura 8.2-8.5). El color rojo corresponde a los periodos con déficit de precipitaciones y el azul a periodos húmedos (superávit de precipitación).

Para la interpretación de las *wavelets* (Figuras 8.6(b)-8.11(b)) se debe tener en cuenta las siguientes características:

- a) **Patrón temporal:** El eje X identifica el registro temporal utilizado en la base de datos seleccionada.
- b) **Patrón escala:** El eje Y representa el periodo (años) de la señal. Valores bajos en el eje Y son indicativos de periodos pequeños, es decir, eventos poco comunes que se repiten cada poco tiempo y por el contrario valores altos indican periodos altos, es decir, eventos extraordinarios que se repiten cada largo periodo de tiempo.
- c) **Intensidad:** La intensidad del color indica la magnitud del comportamiento de las sequías (rojos) o de los periodos más húmedos (azules) del SPI.

Al igual que ocurría anteriormente, las gráficas son más representativas a medida que se aumenta el índice de agregación de manera que se observa mejor los periodos de anomalía de precipitación. En las Figuras 8.2-8.5 se aprecia una concordancia entre los periodos de déficit y superávit de precipitaciones con los mostrados en la Figura 8.1. Por ejemplo, en la Figura 8.1(d), se aprecia durante los periodos 1981-1984 y 1991-1994 épocas de grandes déficit de precipitación, mientras que en los años 1977-1980, se aprecia una época de periodos más húmedos. Observando esos mismos periodos en la Figura 8.5(a), se observa que durante el año 1982 se produce una sequía de gran intensidad entre los 5-10 años, además en 1992 se observa el valor mínimo en torno a los 10-15 años, es decir, la mayor sequía en el periodo analizado. Sin embargo, en el año 1979 se produce el valor máximo (superávit de precipitaciones) en torno a los 10-15 años, indicando el periodo de mayor intensidad de precipitaciones. Por otro lado, el año 1998 es el más húmedo para la ciudad de Valladolid, con un máximo entorno a los 25-30 años de periodo, aunque con menor intensidad que las sequías.

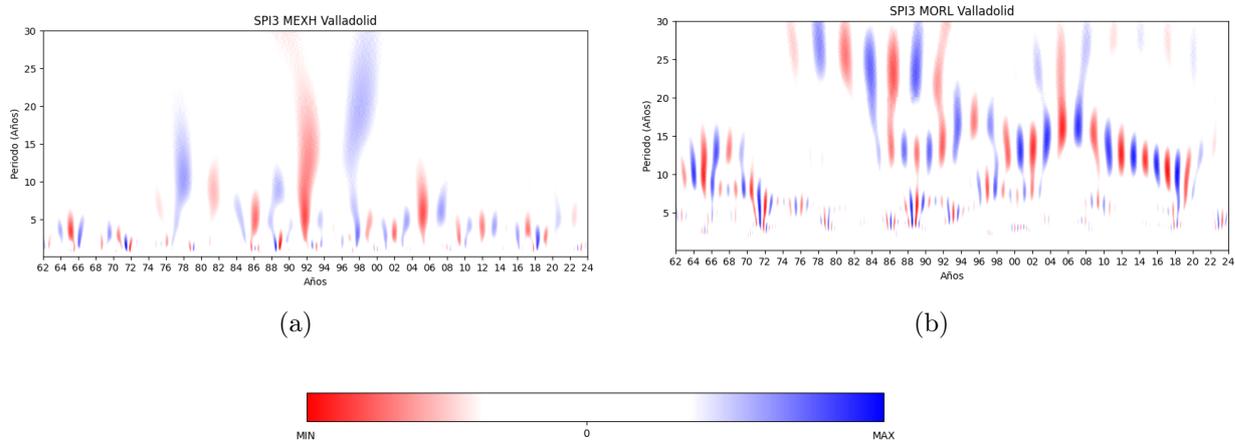


Figura 8.2: SPI3 para Valladolid: (a) Mexh y (b) Morlet

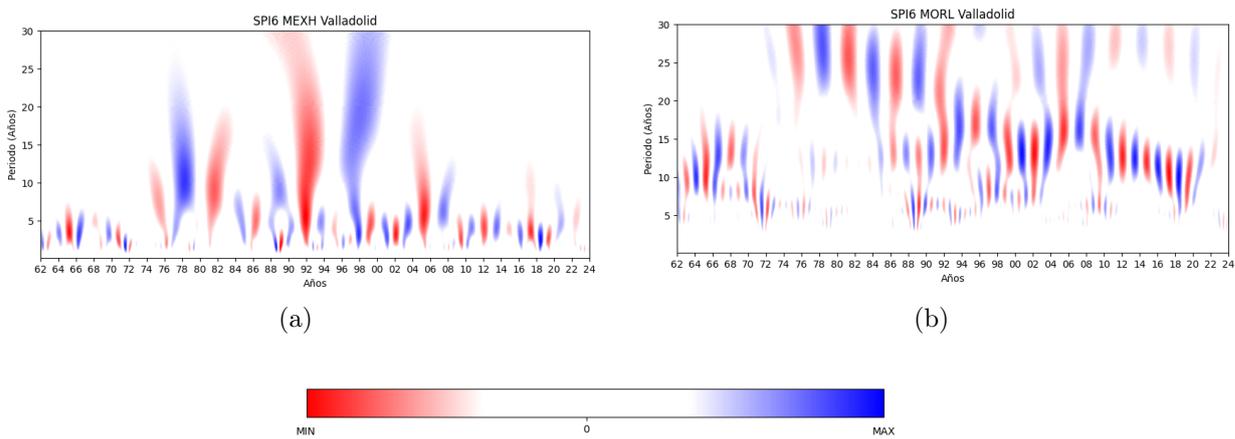


Figura 8.3: SPI6 para Valladolid: (a) Mexh y (b) Morlet

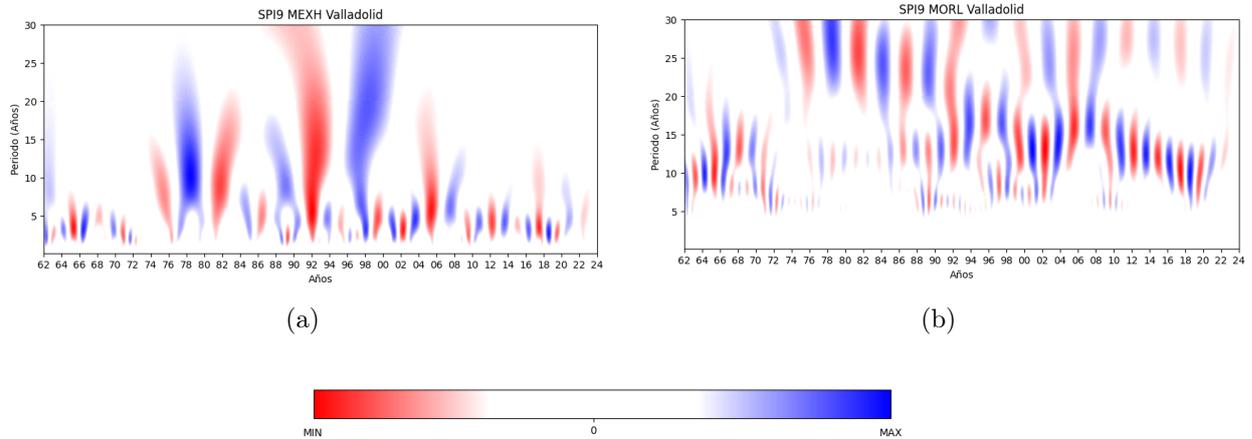


Figura 8.4: SPI9 para Valladolid: (a) Mexh y (b) Morlet

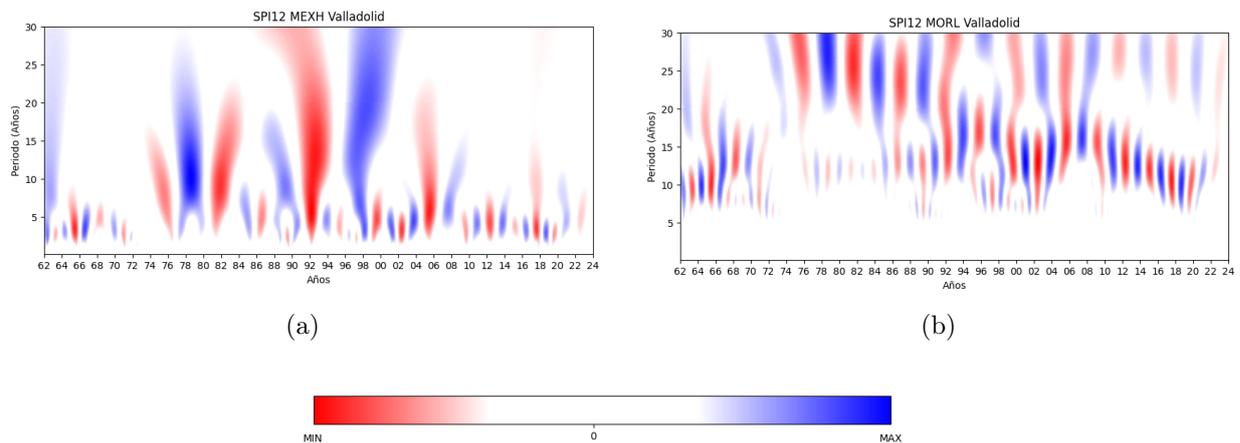


Figura 8.5: SPI12 para Valladolid: (a) Mexh y (b) Morlet

Siguiendo las Figuras 8.2-8.5 se puede apreciar que para las *wavelets* de Morlet (Figuras 8.2(b)-8.5(b)), la señal es excesiva entre $[-2, 2]$, lo que refleja la alta sensibilidad de esta *wavelet* a las frecuencias bajas. Esta representación de Morlet es más adecuada para encontrar señales de baja frecuencia, (por ejemplo, datos paleoclimáticos), ya que estas frecuencias son útiles para entender tendencias y ciclos a largo plazo. Sin embargo, las *wavelets* de Mexh (Figuras 8.2(a)-8.5(a)) son más adecuadas para altas y medias frecuencias. Por este motivo, se decide trabajar con la *wavelet* Mexh y no con la de Morlet.

8.3. Análisis del SPI y *wavelets*

Como se ha mencionado anteriormente en la Sección 7.2 se han obtenido 6 *clusters* (Figura 7.2). A continuación, se muestran la representación de cada serie temporal media

de cada *cluster* y la *wavelet* Mexh correspondiente al índice de agregación seleccionado (SPI12).

En cada figura se representa la serie temporal de la media de cada *cluster* para el SPI12 y la *wavelet* correspondiente. En las siguientes figuras (Figura 8.6-8.11) se representa en color azul el superávit de precipitaciones y en color rojo el déficit de precipitaciones o sequías.

Resultados en *cluster* 1

El *cluster* 1 es el grupo más numeroso y amplio espacialmente. Está formado por las siguientes regiones (por orden alfabético): Albacete, Badajoz, Cádiz, Ciudad Real, Córdoba, Cuenca, Granada, Guadalajara, Huelva, Jaén, Madrid, Málaga, Sevilla y Toledo.

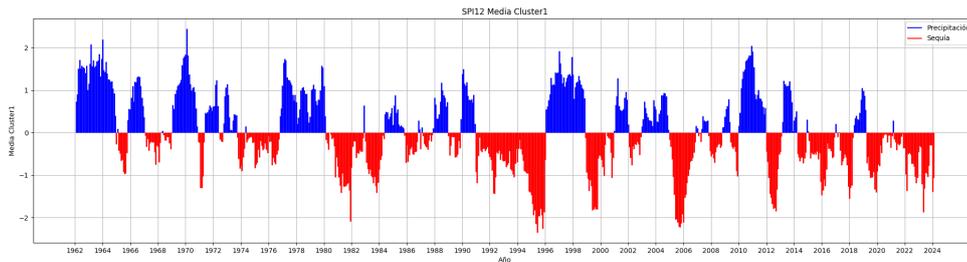
En la Figura 8.6(a) se muestran las variaciones del SPI12 a lo largo de los años, donde los valores positivos, en azul, indican periodos húmedos y los valores negativos, en rojo, representan periodos con déficit de precipitación y en la Figura 8.6(b) se proporciona un análisis de la frecuencia-temporal que permite observar cómo las características de los periodos de déficit y/o superávit de precipitaciones cambian a diferentes escalas temporales. La presencia de ciclos alternos de intensidad variable, sugiere la existencia de patrones cíclicos en el comportamiento de las precipitaciones. Este comportamiento está relacionado con los denominados patrones atmosféricos de teleconexión (Oscilación del Atlántico Norte (NAO), El Niño, etc.) [35][36][37][38][39][40][41].

Entre los eventos más destacados de la Figura 8.6(a) se puede resaltar un periodo húmedo alrededor de los años 1962-1964 y 1977-1980. En algunos meses de estos periodos, los valores del SPI llegan a superar los +2, es decir, según la Tabla 6.1 un periodo extremadamente húmedo. Estas anomalías positivas coinciden con los datos registrados de precipitación total para la península ibérica (Figura 7.3) donde también se reflejan años húmedos, por ejemplo, durante 1962 y 1963 se llegaron a alcanzar los 900 mm de precipitación y durante 1977-1979 los registros de la precipitación total superan la media de precipitaciones.

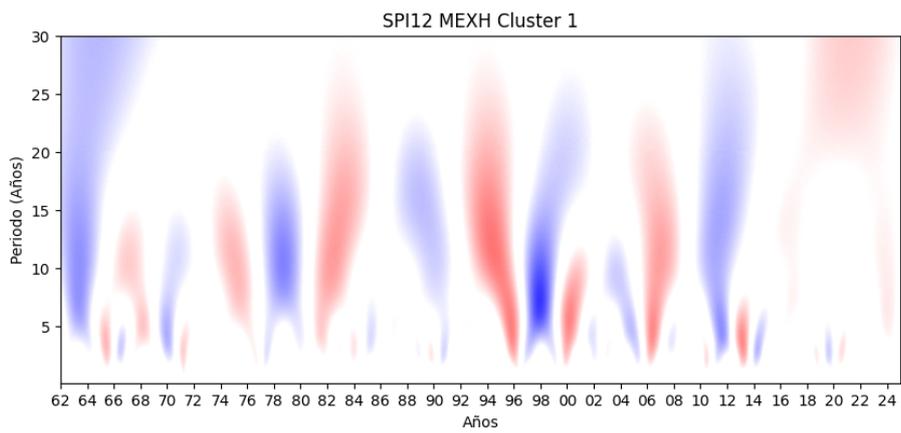
En contraste, también se identifican periodos de sequía extrema. Uno de los más pronunciados se encuentra a principio de la década de los 90, entre 1991-1996, donde los valores del SPI son inferiores a -2 , lo que denota una sequía extrema de acuerdo con la misma clasificación. Esta sequía también es coherente con los datos de precipitación total, los cuales muestran que los valores para estos años fueron inferiores a la media, excepto en 1996. Por otro lado, durante el periodo 1996-1998 en la Figura 8.6(a) se muestra un periodo húmedo pero sin mucha relevancia, sin embargo, en la Figura 8.6(b) se encuentra el valor máximo de superávit de precipitaciones (señal en bandas azules) en una escala entre 5-10 años, es decir, hubo un patrón de precipitación con dicha periodicidad, lo que puede sugerir que en este intervalo temporal existió algún evento atmosférico importante, como puede ser una NAO fuerte o un evento El Niño o La Niña extremo [35][36][37][38][39][40][41]. Además, es llamativo que dicho acontecimiento se produce tras un periodo muy seco, durante los años 1991-1996 (banda roja intensa) en una escala de 10-15 años como también se ha comentado que se mostraba en la Figura 8.6(a). Posteriormente, otro periodo de

déficit de precipitación se encuentra alrededor de 2005-2009, periodo en el cual 2005 destaca como el año con menor precipitación total para la España peninsular y hasta el año 2008 no se vuelve a superar la media.

Finalmente, a partir de 2014 hasta 2024, se observan más periodos de déficit de precipitaciones que periodos húmedos, lo cual es respaldado en los registros de precipitación total ya que muestra varios años a partir del 2014 con precipitaciones inferiores a la media.



(a)



(b)

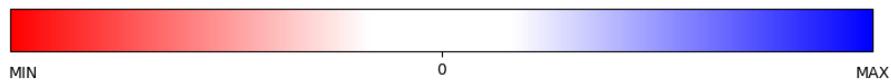


Figura 8.6: Para el *Cluster 1*: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

La comparación entre la serie temporal y el análisis de la wavelet muestra una clara concordancia entre los periodos secos y húmedos observados y los patrones de oscilación capturados por la wavelet, lo que es indicativo de la gran utilidad de esta herramienta para el estudio de este tipo de fenómenos geofísicos.

Resultados en *cluster 2*

El *cluster 2* está formado por las siguientes ciudades, por orden alfabético: Álava,

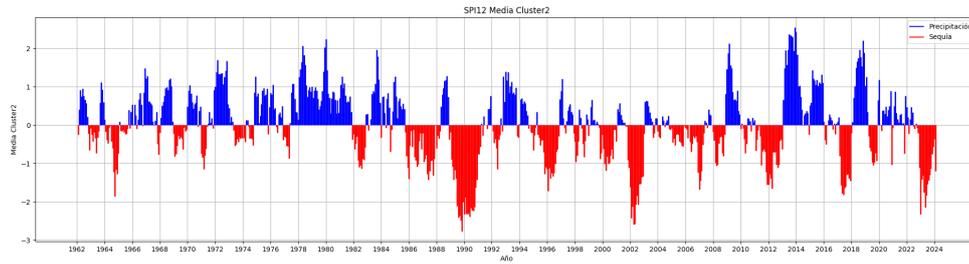
Asturias, Bizkaia, Cantabria, Gipuzkoa y Navarra.

En la Figura 8.7 se muestra la gráfica de la serie temporal media (Figura 8.7(a)) de las regiones que conforman el *cluster* 2 y en la Figura 8.7(b) se muestra la wavelet Mexh correspondiente a dicha serie temporal.

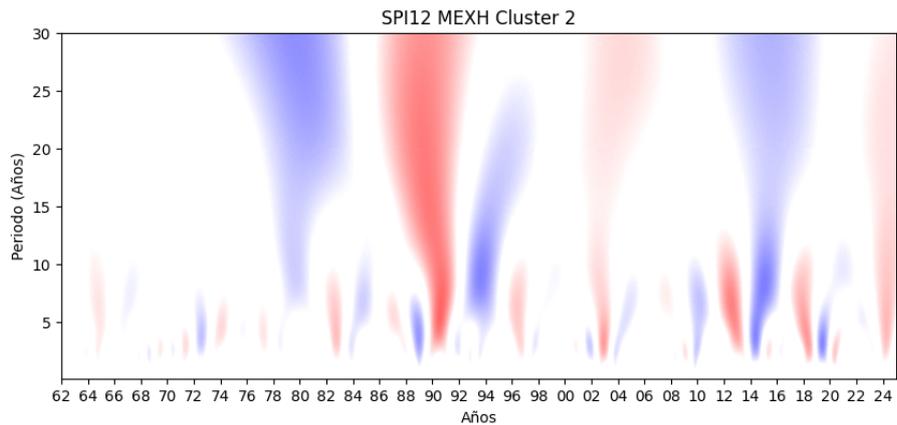
Entre los eventos más significativos que muestra la Figura 8.7(a) se observa un periodo extremadamente húmedo entre los años 1977 y 1981 ya que los valores del SPI12 (Figura 8.7(a)) superan +2 según la clasificación establecida en la Tabla 6.1. Para este mismo periodo, la Figura 8.7(b) muestra una fuerte oscilación, coincidiendo con lo observado en la serie temporal, además dicha señal es una de las más intensas. Otro periodo destacable se registra durante los años 2013-2014 donde se alcanzan los valores más altos del SPI12 (Figura 8.7(a)) para el *cluster* 2, indicando nuevamente condiciones de superávit de precipitaciones. Comparando con la Figura 7.3 (Registro de precipitaciones totales de la España peninsular), ambos periodos húmedos coinciden con los registros de precipitación, los cuales superan la media para dichos años. Por otro lado, el periodo de déficit de precipitaciones más largo se produce entre 1989-1991, con picos muy cercanos a -3 en el SPI12 (Figura 8.7(a)), lo cual se considera un periodo de sequía extrema, sin embargo los registros de la precipitación total de la España peninsular (Figura 7.3) muestran que en 1989 se registraron precipitaciones superiores a la media. Para este mismo periodo se puede observar que la figura 8.7(b) muestra una señal intensa, reflejando el periodo seco extremo que se visualiza en la serie temporal del SPI12 para el *cluster* 2 (Figura 8.7(a)).

En años más recientes, a partir de 2014 se pueden observar periodos principalmente húmedos (Figura 8.7(a)), en contra de lo obtenido para el *cluster* anterior. Sin embargo, también se produce algún periodo de sequía importante desde finales de 2022 hasta principios de 2024 (Figura 8.7(a)). Por otro lado, en la Figura 8.7(b) para los últimos años se observa una alternancia de colores a baja escala, lo que indica una variabilidad en las precipitaciones más rápida teniendo episodios de superávit y de déficit con mayor frecuencia y en periodos más cortos.

Cabe destacar que los periodos de mayor intensidad que se muestran en la Figura 8.7(b), tanto de déficit como de superávit de precipitaciones, se concentran en escalas de 5 a 10 años.



(a)



(b)

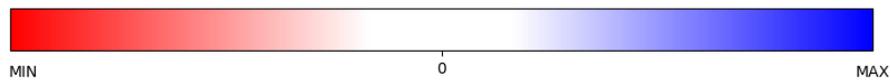


Figura 8.7: Para el *Cluster 2*: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

En conclusión, los años entre 1980-1982 y 1990-1995 destacan por eventos meteorológicos extremos teniendo alternancias entre precipitaciones y sequías. Además, dichas alternancias son más frecuentes durante los últimos años. De nuevo se observa la concordancia entre la información ofrecida por la serie temporal y la *wavelet* resultante.

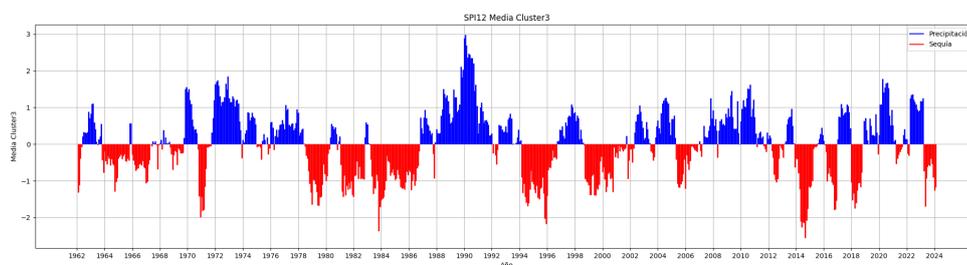
Resultados en *cluster 3*

En este caso el *cluster 3* está formado por las regiones del este peninsular, por orden alfabético: Alicante, Almería, Murcia y Valencia.

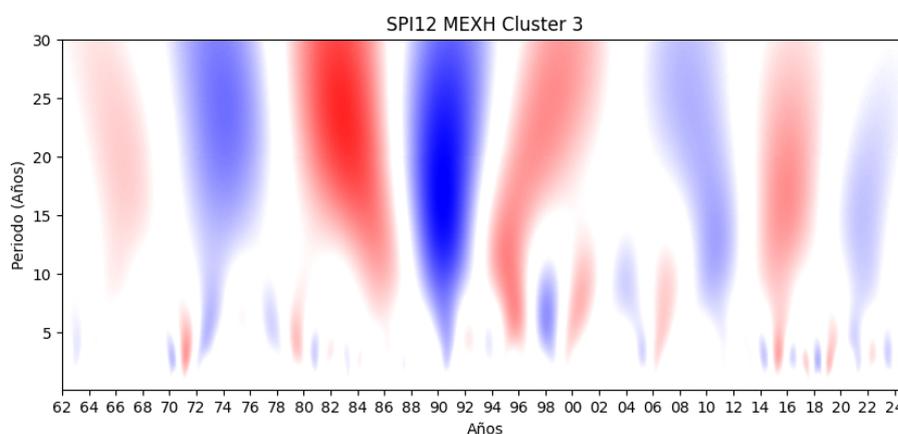
En la Figura 8.8 se muestran los resultados obtenidos de la serie temporal media del SPI12 en las regiones citadas anteriormente.

En la Figura 8.8(a) se presenta el mayor periodo de superávit de precipitación durante los años 1987-1994 llegando a tener un extremo de +3 de SPI12, es decir, es un periodo extremadamente húmedo, dentro de este periodo durante los años 1989-1993 se presenta el máximo (superávit de precipitación) de la representación de la *wavelet* en la escala de

10-15 años. El resto de valores de superávit de precipitaciones no superan el valor $+2$. Sin embargo, en los registros de precipitaciones totales de la España peninsular (Figura 7.3) este intervalo se divide en dos periodos, el primero de ellos es el periodo 1987-1989 para el cual las precipitaciones registradas superan la media mientras que, para el segundo intervalo, 1990-1994 son valores inferiores a la media. Por el contrario, el periodo de déficit de precipitación más largo para este *cluster* (Figura 8.8(a)) se presenta durante mediados de 1978 hasta principios de 1987, alcanzando en algún mes del periodo un valor inferior a -2 del SPI12, periodo en el que se produjo una de las peores sequías en la España peninsular especialmente en el sur, como se indica en la Sección 7.2. Además, durante el periodo 1981-1985 se observa durante la escala 15-20 años el valor mínimo (banda roja intensa) de la representación de la *wavelet* (Figura 8.8(b)). Por tanto, en orden cronológico primero se produce mayoritariamente una sequía severa y le sigue un periodo severamente húmedo con algún mes extremadamente húmedo.



(a)



(b)

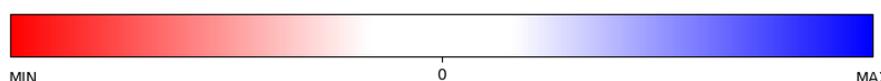


Figura 8.8: Para el *Cluster 3*: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

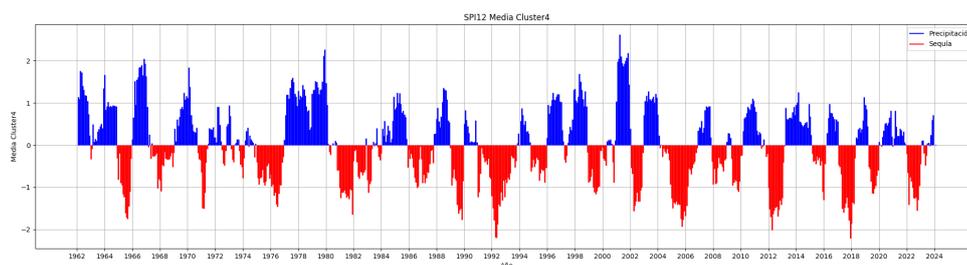
En conclusión, en las regiones del *cluster* 3 es notable el cambio que se produce durante el periodo 1980 y 1995 entre un déficit de precipitación prolongado a superávit de precipitación de gran intensidad.

Resultados en *cluster* 4

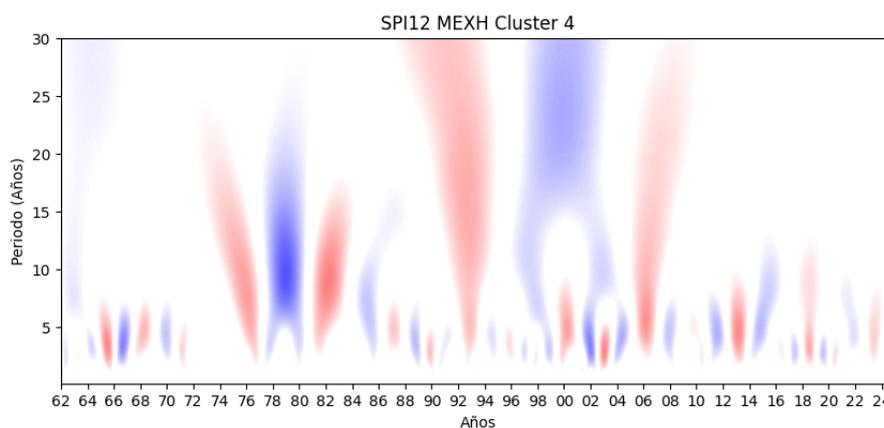
El *cluster* 4 está formado por las siguientes regiones: A Coruña, Ávila, Burgos, León, Lugo, Ourense, Palencia, Pontevedra, Salamanca, Segovia, Valladolid y Zamora.

En la Figura 8.9 se muestra la serie temporal media (Figura 8.9(a)) y la wavelet sombrero mexicano asociada a dicha serie temporal (Figura 8.9(b)).

Siguiendo la Figura 8.9(a) se identifican varios eventos destacables. Entre 1977 y 1980 se observa un periodo húmedo con picos que alcanzan valores de +2 en el SPI12, es decir, periodos extremadamente húmedos. Además, durante este mismo periodo los registros de precipitación total para la España peninsular (Figura 7.3) muestran que se superó la media de precipitaciones totales. Por otro lado, este periodo 1977-1980 en la Figura 8.9(b) cuenta con el máximo de intensidad para el evento de superávit de precipitación en periodos de 10-15 años, el cual está entre dos periodos secos, principios de 1977 y durante el año 1981, en la misma escala de tiempo. Un comportamiento similar se muestra en la Figura 8.9(a), durante los años 2001-2002 con también picos con valores cercanos a +2. Por otro lado, entre 1991 y 1993 se presenta un periodo seco, alcanzando valores de -2 en el SPI12, lo que indica un periodo extremadamente seco, sin embargo, en la Figura 8.9(b) se muestra que durante este periodo se tiene una intensidad media en la escala de 10-15 años. En los años más recientes, como 2005-2007, 2012-2013 y 2017-2018, se producen sequías importantes aunque de corta duración (Figura 8.9(a)) y de intensidad media en la escala de 5 años (Figura 8.9(b)), sin embargo, en los registros que se muestran en la Figura 7.3 para el periodo 2005-2007 también se muestra un periodo de precipitaciones inferiores a la media mientras que para los años 2012 y 2018 se tiene precipitaciones totales superiores a la media y para los años 2013 y 2017 se tiene precipitaciones totales inferiores a la media. Estos periodos secos se alternan con periodos húmedos, aunque con valores que apenas superan el valor +1 de SPI12, es decir, periodos moderadamente húmedos.



(a)



(b)

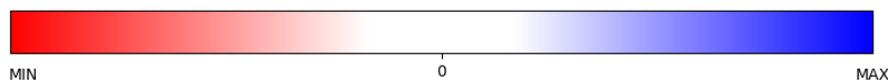


Figura 8.9: Para el *Cluster 4*: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

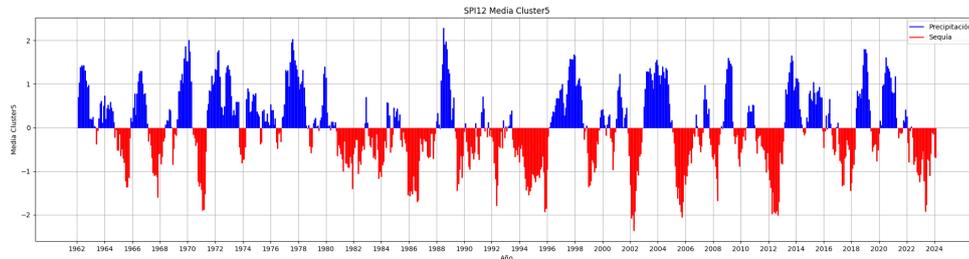
Resultados en *cluster 5*

El *cluster 5* está formado por las siguientes ciudades: Huesca, La Rioja, Soria, Teruel y Zaragoza.

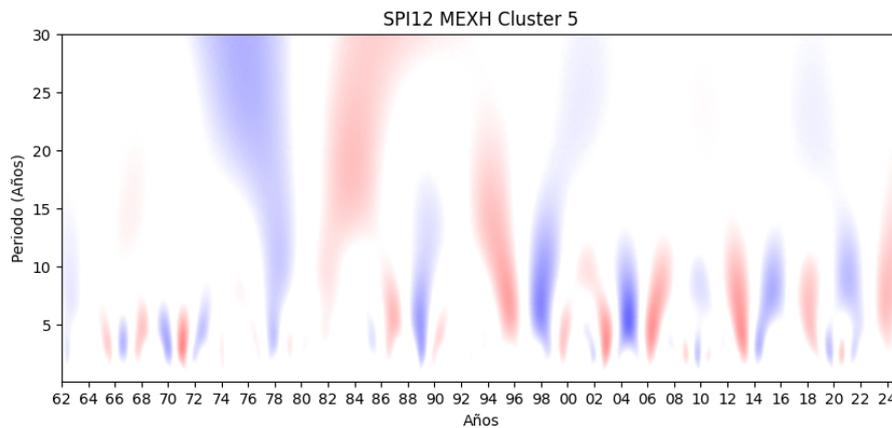
La Figura 8.10 consta de dos subfiguras, la primera de ellas (Figura 8.10(a)) es la representación gráfica de la serie temporal media del SPI12 asociado a las regiones del *cluster 5*, la segunda (Figura 8.10(b)) se trata de la wavelet sombrero mexicano asociada a dicha serie temporal.

Desde finales de 1971 hasta principios de 1981 se observa un superávit de precipitaciones con valores del SPI12 que superan los +2 (Figura 8.10(a)), este mismo periodo en la Figura 7.3 (registro de precipitación total para la España peninsular) también muestra mayoritariamente valores de precipitación superiores a la media. Posteriormente, durante 1981-1987 y 1989-1996, se presentan periodos de déficit de precipitación de larga duración, con valores del SPI12 inferiores a -2 (Figura 8.10(a)) pero observando la Figura 8.10(b)

dicho periodo de déficit de precipitación tiene muy baja intensidad en la escala de 15-20 años, sin embargo, durante 1988, año intermedio de ambos periodos, se produce un periodo extremadamente húmedo, según la clasificación de la Tabla 7.3. En la Figura 8.10(a) durante el periodo 2004-2005 se tiene un periodo húmedo no especialmente llamativo, sin embargo, en la Figura 8.10(b) es el periodo que cuenta con la mayor intensidad de superávit de precipitaciones en la escala de los 5 años. En los últimos años en la Figura 8.10(a), los periodos con déficit de precipitación son más abundantes y caracterizados como extremadamente secos mientras que los periodos de superávit de precipitación son menos abundantes y severamente húmedos, los años 2006 y 2013 son los de mayor intensidad (Figura 8.10(b)) en una escala de 5 años. Por otro lado, los periodos secos destacados en este *cluster* coinciden con periodos de precipitación inferior a la media en la Figura 7.3.



(a)



(b)



Figura 8.10: Para el *Cluster* 5: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

Resultados en *cluster* 6

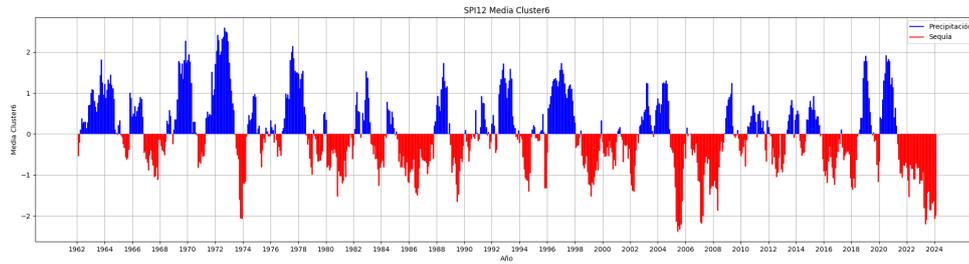
Por último, el *cluster* 6 está formado por las regiones que conforman Cataluña: Barcelona, Girona, Lleida y Tarragona.

En la Figura 8.11 se presenta la serie temporal media del SPI12 del *cluster* 6 (Figura 8.11(a)) y la wavelet sombrero mexicano asociado a dicha serie temporal (Figura 8.11(b)).

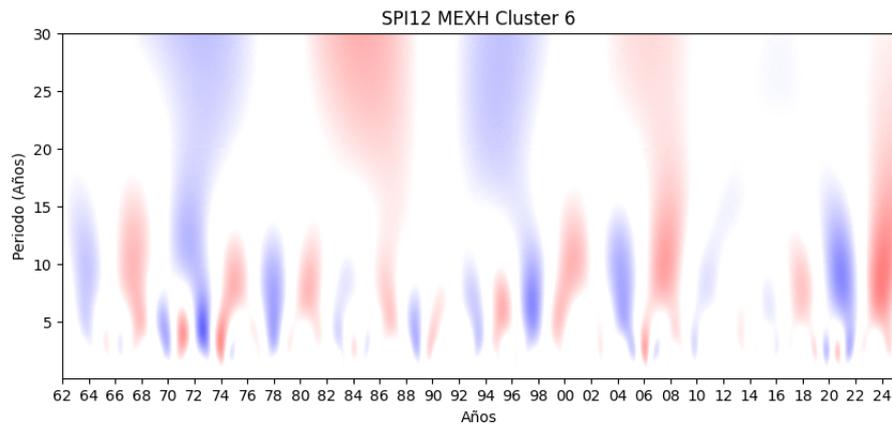
En la Figura 8.11(a) se observa un periodo extremadamente húmedo entre 1969 y 1973, con valores medios del SPI12 superiores a +2, a diferencia del registro de precipitaciones totales de la España peninsular (Figura 7.3) donde se observa que para parte de dicho periodo, 1972-1973, las precipitaciones totales son inferiores a la media. Para este mismo periodo, 1972-1973, la Figura 8.11(b) muestra señal en el periodo más intenso de superávit de precipitaciones en la escala de 5 años.

Posteriormente, entre 2005-2009 y 2021-2024 se registran los mayores periodos secos, es decir, periodos con déficit de precipitación, con valores del SPI12 inferiores a -2 . Además, alrededor de dicho periodos entre 2006-2016 y 2023-2024, en la Figura 8.11(b) también se presenta los periodos de déficit de precipitaciones más intensos en escalas entre 7-10 años.

Además, cabe destacar que en la Figura 8.11(b) no hay periodos de gran intensidad ni de déficit ni de superávit de precipitaciones por encima de los 5 años.



(a)



(b)

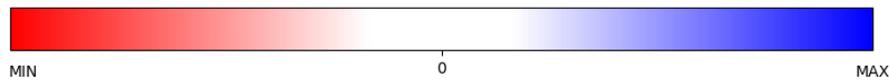


Figura 8.11: Para el *Cluster* 6: (a) SPI12 de la media de la serie temporal y (b) Wavelet sombrero mexicano

Los resultados obtenidos analizando la serie temporal media de cada uno de los *cluster* resultantes de aplicar las *k-means* a las diferentes series de SPI han revelado un conjunto de periodos secos y húmedos que, analizándolos con las *wavelets*, han puesto de manifiesto la gran utilidad que presenta esta herramienta matemática en el estudio de fenómenos geofísicos. Para obtener mayor información de los resultados obtenidos sería deseable desarrollar análisis adicionales relacionando estos resultados con otro tipo de procesos atmosféricos que son conexos con la frecuencia de aparición de la precipitación en la península ibérica.

Capítulo 9

Conclusiones

En este TFG se ha analizado el índice SPI, es decir, el índice que presenta periodos de déficit y superávit de precipitación en las distintas localidades de la península ibérica. La base de datos SPI utilizada comprende un periodo desde enero de 1962 hasta enero de 2024. Esta información se ha agrupado en diferentes regiones para una mayor generalidad a la hora de la explicación de los resultados.

Para analizar la aparición de eventos de precipitación extremos se han utilizado *wavelets* continuas. Se ha seleccionado un índice de agregación y tipo de *wavelet* concretos basándose en los resultados expuestos para la ciudad de Valladolid, es decir, de todos los índices de agregación utilizados se ha seleccionado el correspondiente a los 12 meses: por tanto, se ha utilizado el SPI12 con el objetivo de identificar grandes periodos temporales de anomalías de precipitación.

Con objeto de la reducir la dimensionalidad de la gran base de datos con los que se ha trabajado, se ha aplicado un método de regionalización. En concreto se ha utilizado *k-means* obteniéndose un total de 6 *clusters* formados por distintas regiones de la península ibérica que presentan un comportamiento similar en el régimen pluviométrico. Una vez agrupada la información, se procedió a calcular las series medias de cada *cluster* a las que se les aplicó las *wavelets* Morlet y sombrero mexicano. Se ha seleccionado finalmente la *wavelet* sombrero mexicano debido a la gran capacidad de obtener información de señales en alta y baja frecuencia frente a la Morlet, que ha resultado más adecuada para muy bajas frecuencias. El análisis se completa con la comparación de cada *cluster* con los datos anuales de precipitación media registrados en la España peninsular.

Los distintos grupos permiten observar patrones en la distribución de la precipitación en cada región.

En vista de todos los resultados obtenidos se puede concluir los siguientes puntos:

- La elección del SPI12 frente a SPI3, SPI6 y SPI9 para la realización del estudio se debe a la capacidad para ofrecer representaciones más estables y claras de los patrones de precipitación minimizando la posibilidad de una interpretación errónea. A diferencia de los índices a corto plazo, el SPI12 permitiría identificar los efectos que sobre las series utilizadas podrían ejercer fenómenos meteorológicos, como po-

drían ser patrones atmosféricos de teleconexión (NAO, El Niño, etc.). Sería deseable realizar un análisis más exhaustivo de las relaciones entre este tipo de patrones y el SPI.

- La elección de la *wavelet* tipo sombrero mexicano frente a la *wavelet* de Morlet se debe a la capacidad para detectar señales con mayor precisión en alta y media frecuencia en las series temporales estudiadas, mientras que la *wavelet* de Morlet es más adecuada para identificar señales de muy baja frecuencia.
- Los periodos más húmedos han sido entre 1975-1981 mientras que los periodos más secos han sido entre 1991-1996, sin embargo, siguiendo el registro de precipitaciones totales en la España peninsular tal y como se muestra en la Figura 7.3 se puede observar que el año más húmedo se registró en el año 1963 y el año más seco fue el año 2005.
- Se ha puesto de manifiesto la falta de uniformidad de los patrones de precipitación a lo largo del territorio estudiado, con resultados dispares entre los diferentes *clusters* seleccionados así como en sus *wavelets*. Comparando entre las distintas zonas geográficas se observa que en el norte y el oeste se tiene mayor frecuencia de periodos húmedos mientras que el sur y el este son zonas con sequías prolongadas e intensas. Por otro lado, el centro peninsular cuenta con una gran variabilidad de precipitación, alternando periodos extremadamente húmedos y secos.
- La gran coherencia encontrada en los resultados de las *wavelets* con las series temporales utilizadas ponen de manifiesto la gran utilidad de las *wavelets*, y en particular la *wavelet* sombrero mexicano, para el estudio de eventos pluviométricos así como sus frecuencias de aparición.

El estudio revela una gran variabilidad en los patrones de precipitación a lo largo de la península ibérica, con variaciones entre los diferentes *clusters*. Mientras que algunas regiones han experimentado periodos prolongados de déficit y superávit de precipitación, otras han mostrado variabilidad de las precipitaciones, con alternancia rápida entre condiciones extremas, es decir, periodos cortos de déficit y superávit de precipitación. Este estudio subraya la importancia de considerar las características regionales al analizar la distribución de las precipitaciones de los últimos años, es decir, se necesitan estrategias regionalizadas para la gestión del agua y de planificación adaptativa para mitigar los impactos de estos eventos extremos en el futuro, dado que cada zona enfrenta desafíos específicos en términos de disponibilidad de agua y resiliencia a eventos extremos. El cambio climático global también sugiere la posibilidad de que estas tendencias extremas se vuelvan más frecuentes, lo que subraya la urgencia de medidas de adaptación a eventos de precipitación extrema y/o sequías y por tanto implementar medidas de mitigación y adaptación efectivas.

Parte III

Apéndices

Apéndice A

Acrónimos

CWT Transformada Wavelet Continua

DFT Transformada Discreta de Fourier

DWT Transformada Wavelet Discreta

FFT Transformada rápida de Fourier

MRA Análisis multiresolución

NAO Oscilación del Atlántico Norte

PDSI Índice Palmer de Severidad de Sequía

SPI Índice de Precipitación Estandarizado

STFT Transformada corta de Fourier

TFG Trabajo Fin de Grado

Apéndice B

Documentación adjunta

En este Anexo se detalla parte del código fuente en el lenguaje de programación *Python* necesarios para el desarrollo del proyecto.

El código se ha desarrallado mediante *Jupyter Notebook* en el lenguaje *Python* como se ha comentado en la Sección 7.1 de esta memoria. Gracias a la estructura de *Jupyter Notebook* el código está dividido por “bloques” o celdas.

```
1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 import pywt
5 import csv
6 from sklearn.decomposition import PCA
7 from sklearn.cluster import KMeans
8 import geopandas as gpd
9 import cartopy.feature as cfeature
10 import cartopy.crs as ccrs
11 from cartopy.mpl.gridliner import LONGITUDE_FORMATTER,
    LATITUDE_FORMATTER
12 from sklearn.preprocessing import StandardScaler
13 from sklearn.datasets import make_blobs
```

```
1 # Cargar el archivo CSV
2 df_3meses = pd.read_csv('DatosSPI3meses.csv')
3 df_6meses = pd.read_csv('DatosSPI6meses.csv')
4 df_9meses = pd.read_csv('DatosSPI9meses.csv')
5 df_12meses = pd.read_csv('DatosSPI12meses.csv')
6
7 # Seleccionar las columnas de interés (desde la columna 3 en
    adelante)
8 data_3meses = df_3meses.iloc[:, 2:]
9 data_6meses = df_6meses.iloc[:, 2:]
10 data_9meses = df_9meses.iloc[:, 2:]
```

```

11 data_12meses = df_12meses.iloc[:, 2:]
12
13 df_3meses_filtrado = df_3meses[df_3meses.iloc[:, 0] != 1961]
14 df_6meses_filtrado = df_6meses[df_6meses.iloc[:, 0] != 1961]
15 df_9meses_filtrado = df_9meses[df_9meses.iloc[:, 0] != 1961]
16 df_12meses_filtrado = df_12meses[df_12meses.iloc[:, 0] != 1961]
17
18 data_3mesesPeninsula = df_3meses_filtrado.drop(columns=['Santa Cruz
19 de Tenerife', 'Las Palmas', 'Ceuta', 'Melilla', 'Balears'])
20 data_6mesesPeninsula = df_6meses_filtrado.drop(columns=['Santa Cruz
21 de Tenerife', 'Las Palmas', 'Ceuta', 'Melilla', 'Balears'])
22 data_9mesesPeninsula = df_9meses_filtrado.drop(columns=['Santa Cruz
23 de Tenerife', 'Las Palmas', 'Ceuta', 'Melilla', 'Balears'])
24 data_12mesesPeninsula = df_12meses_filtrado.drop(columns=['Santa
25 Cruz de Tenerife', 'Las Palmas', 'Ceuta', 'Melilla', 'Balears'])

```

SPI3, SPI6, SPI9 Y SPI12 DE VALLADOLID

```

1 # Seleccionar las dos columnas deseadas
2 años = data_3mesesPeninsula.iloc[:, 0] # Primera columna
3 mes = data_3mesesPeninsula.iloc[:, 1]
4 datos_valladolid_3 = data_3mesesPeninsula['Valladolid']
5
6 # Crear un nuevo DataFrame con las columnas seleccionadas
7 df_nuevo_3meses = pd.DataFrame({
8     'Años': años,
9     'Mes': mes,
10    'Datos_Valladolid': datos_valladolid_3
11 })
12
13 años_unicos = df_nuevo_3meses['Años'].unique()
14 # Graficar los datos
15 plt.figure(figsize=(25, 6))
16 for año in años_unicos:
17     df_año = df_nuevo_3meses[df_nuevo_3meses['Años'] == año]
18     for index, row in df_año.iterrows():
19         mes = row['Mes']
20         color = 'blue' if row['Datos_Valladolid'] > 0 else 'red'
21         plt.vlines(x=(año + mes / 12), ymin=0, ymax=row['
22             Datos_Valladolid'], color=color, linewidth=2)
23
24 # Crear líneas "falsas" para la leyenda
25 plt.plot([], [], color='blue', label='Precipitación')
26 plt.plot([], [], color='red', label='Sequía')
27
28 # Añadir la leyenda
29 plt.legend()

```

```

29
30 min_año = df_nuevo_3meses['Años'].min()
31 max_año = df_nuevo_3meses['Años'].max()
32 plt.xticks(np.arange(min_año, max_año + 1, 2))
33 plt.xlabel('Año')
34 plt.ylabel('Datos de Valladolid')
35 plt.title('SPI3 Valladolid')
36 plt.grid(True)
37 plt.savefig('SPI3/SPI3Valladolid')
38 plt.show()

```

```

1 # Seleccionar las dos columnas deseadas
2 años = data_6mesesPeninsula.iloc[:, 0] # Primera columna
3 mes = data_6mesesPeninsula.iloc[:, 1]
4 datos_valladolid_6 = data_6mesesPeninsula['Valladolid']
5
6 # Crear un nuevo DataFrame con las columnas seleccionadas
7 df_nuevo_6meses = pd.DataFrame({
8     'Años': años,
9     'Mes': mes,
10    'Datos_Valladolid': datos_valladolid_6
11 })
12
13 años_unicos = df_nuevo_6meses['Años'].unique()
14 # Graficar los datos
15 plt.figure(figsize=(25, 6))
16 for año in años_unicos:
17     df_año = df_nuevo_6meses[df_nuevo_6meses['Años'] == año]
18     for index, row in df_año.iterrows():
19         mes = row['Mes']
20         color = 'blue' if row['Datos_Valladolid'] > 0 else 'red'
21         plt.vlines(x=(año + mes / 12), ymin=0, ymax=row['
22             Datos_Valladolid'], color=color, linewidth=2)
23
24 # Crear líneas "falsas" para la leyenda
25 plt.plot([], [], color='blue', label='Precipitación')
26 plt.plot([], [], color='red', label='Sequía')
27
28 # Añadir la leyenda
29 plt.legend()
30 min_año = df_nuevo_6meses['Años'].min()
31 max_año = df_nuevo_6meses['Años'].max()
32 plt.xticks(np.arange(min_año, max_año + 1, 2))
33 plt.xlabel('Año')
34 plt.ylabel('Datos de Valladolid')
35 plt.title('SPI6 Valladolid')

```

```

35 plt.grid(True)
36 plt.savefig('SPI6/SPI6Valladolid')
37 plt.show()

1 # Seleccionar las dos columnas deseadas
2 años = data_9mesesPeninsula.iloc[:, 0] # Primera columna
3 mes = data_9mesesPeninsula.iloc[:, 1]
4 datos_valladolid_9 = data_9mesesPeninsula['Valladolid']
5
6 # Crear un nuevo DataFrame con las columnas seleccionadas
7 df_nuevo_9meses = pd.DataFrame({
8     'Años': años,
9     'Mes': mes,
10    'Datos_Valladolid': datos_valladolid_9
11 })
12
13 años_unicos = df_nuevo_9meses['Años'].unique()
14 # Graficar los datos
15 plt.figure(figsize=(25, 6))
16 for año in años_unicos:
17     df_año = df_nuevo_9meses[df_nuevo_9meses['Años'] == año]
18     for index, row in df_año.iterrows():
19         mes = row['Mes']
20         color = 'blue' if row['Datos_Valladolid'] > 0 else 'red'
21         plt.vlines(x=(año + mes / 12), ymin=0, ymax=row['
22             Datos_Valladolid'], color=color, linewidth=2)
23
24 # Crear líneas "falsas" para la leyenda
25 plt.plot([], [], color='blue', label='Precipitación')
26 plt.plot([], [], color='red', label='Sequía')
27
28 # Añadir la leyenda
29 plt.legend()
30
31 min_año = df_nuevo_9meses['Años'].min()
32 max_año = df_nuevo_9meses['Años'].max()
33 plt.xticks(np.arange(min_año, max_año + 1, 2))
34 plt.xlabel('Año')
35 plt.ylabel('Datos de Valladolid')
36 plt.title('SPI9 Valladolid')
37 plt.grid(True)
38 plt.savefig('SPI9/SPI9Valladolid')
39 plt.show()

1 # Seleccionar las dos columnas deseadas
2 años = data_12mesesPeninsula.iloc[:, 0] # Primera columna

```

```

3 mes = data_12mesesPeninsula.iloc[:, 1]
4 datos_valladolid_12 = data_12mesesPeninsula['Valladolid']
5
6 # Crear un nuevo DataFrame con las columnas seleccionadas
7 df_nuevo_12meses = pd.DataFrame({
8     'Años': años,
9     'Mes': mes,
10    'Datos_Valladolid': datos_valladolid_12
11 })
12
13 años_unicos = df_nuevo_12meses['Años'].unique()
14 # Graficar los datos
15 plt.figure(figsize=(25, 6))
16 for año in años_unicos:
17     df_año = df_nuevo_12meses[df_nuevo_12meses['Años'] == año]
18     for index, row in df_año.iterrows():
19         mes = row['Mes']
20         color = 'blue' if row['Datos_Valladolid'] > 0 else 'red'
21         plt.vlines(x=(año + mes / 12), ymin=0, ymax=row['
22             Datos_Valladolid'], color=color, linewidth=2)
23         # Crear líneas "falsas" para la leyenda
24 plt.plot([], [], color='blue', label='Precipitación')
25 plt.plot([], [], color='red', label='Sequía')
26
27 # Añadir la leyenda
28 plt.legend()
29
30 min_año = df_nuevo_12meses['Años'].min()
31 max_año = df_nuevo_12meses['Años'].max()
32 plt.xticks(np.arange(min_año, max_año + 1, 2))
33 plt.xlabel('Año')
34 plt.ylabel('Datos de Valladolid')
35 plt.title('SPI12 Valladolid')
36 plt.grid(True)
37 plt.savefig('SPI12/SPI12Valladolid')
38 plt.show()

```

WAVELET VALLADOLID

```

1     tipos_wavelet = ['mexh', 'morl']
2     escalas = np.arange(1, 768)
3
4 # Convertir las escalas a años
5 escalas_en_años = escalas / 12.0 # 12 meses por año
6
7 coeficientes_totales = []
8 frecuencias_totales = []

```

```

9 years = np.arange(1962, 2025, 2)
10
11 for wavelet in tipos_wavelet:
12     # Calcular la CWT
13     coeficientes, frecuencias = pywt.cwt(datos_valladolid_3,
14         escalas_en_anios, wavelet)
15     coeficientes_totales.append(coeficientes)
16     frecuencias_totales.append(frecuencias)
17
18     # Convertir a arrays para facilitar el manejo
19     coeficientes_totales = np.array(coeficientes_totales)
20     frecuencias_totales = np.array(frecuencias_totales)
21
22     vmax = np.max(coeficientes_totales)
23     vmin = np.min(coeficientes_totales)
24     # Creamos el colormap personalizado
25     cmap = create_custom_colormap(vmin, vmax)
26     # Graficar los coeficientes de la transformada para cada
27     # columna
28     plt.figure(figsize=(10, 6))
29     plt.imshow(coeficientes_totales.mean(axis=0), extent=[1962,
30         2024, min(escalas_en_anios), 30], aspect='auto', cmap=cmap,
31         origin='lower', vmax=vmax, vmin=vmin)
32     cbar = plt.colorbar(orientation='horizontal')
33
34     # Ocultar los ticks numéricos de la barra de color
35     cbar.set_ticks([0])
36
37     # Agregar texto personalizado para MIN y MAX
38     cbar.ax.text(-0.0015, -0.5, 'MIN', color='black', ha='center',
39         va='center', transform=cbar.ax.transAxes)
40     cbar.ax.text(1, -0.5, 'MAX', color='black', ha='center', va='
41         center', transform=cbar.ax.transAxes)
42     plt.xticks(years, [str(year)[-2:] for year in years])
43     plt.xlabel('Años')
44     plt.ylabel('Periodo (Años)')
45     plt.title(f'SPI3 {wavelet.upper()} Valladolid')
46     plt.savefig(f'SPI3/WAVELET/SPI3{wavelet.upper()}Valladolid')
47     plt.show()
48     coeficientes_totales = []
49     frecuencias_totales = []
50
51 tipos_wavelet = ['mexh', 'morl']
52 escalas = np.arange(1, 768)
53
54 # Convertir las escalas a años

```

```

5 escalas_en_anios = escalas / 12.0 # 12 meses por año
6
7 coeficientes_totales = []
8 frecuencias_totales = []
9 years = np.arange(1962, 2025, 2)
10
11 for wavelet in tipos_wavelet:
12     # Calcular la CWT
13     coeficientes, frecuencias = pywt.cwt(datos_valladolid_6,
14         escalas_en_anios, wavelet)
15     coeficientes_totales.append(coeficientes)
16     frecuencias_totales.append(frecuencias)
17
18     # Convertir a arrays para facilitar el manejo
19     coeficientes_totales = np.array(coeficientes_totales)
20     frecuencias_totales = np.array(frecuencias_totales)
21
22     vmax = np.max(coeficientes_totales)
23     vmin = np.min(coeficientes_totales)
24     # Creamos el colormap personalizado
25     cmap = create_custom_colormap(vmin, vmax)
26     # Graficar los coeficientes de la transformada para cada
27     # columna
28     plt.figure(figsize=(10, 6))
29     plt.imshow(coeficientes_totales.mean(axis=0), extent=[1962,
30         2024, min(escalas_en_anios), 30], aspect='auto', cmap=cmap,
31         origin='lower', vmax=vmax, vmin=vmin)
32     cbar = plt.colorbar(orientation='horizontal')
33
34     # Ocultar los ticks numéricos de la barra de color
35     cbar.set_ticks([0])
36
37     # Agregar texto personalizado para MIN y MAX
38     cbar.ax.text(-0.015, -0.5, 'MIN', color='black', ha='center',
39         va='center', transform=cbar.ax.transAxes)
40     cbar.ax.text(1, -0.5, 'MAX', color='black', ha='center', va='
41         center', transform=cbar.ax.transAxes)
42     plt.xticks(years, [str(year)[-2:] for year in years])
43     plt.xlabel('Años')
44     plt.ylabel('Periodo (Años)')
45     plt.title(f'SPI6 {wavelet.upper()} Valladolid')
46     plt.savefig(f'SPI6/WAVELET/SPI6{wavelet.upper()}Valladolid')
47     plt.show()
48     coeficientes_totales = []
49     frecuencias_totales = []

```

```

1     tipos_wavelet = ['mexh', 'morl']
2     escalas = np.arange(1, 768)
3
4     # Convertir las escalas a años
5     escalas_en_anios = escalas / 12.0 # 12 meses por año
6
7     coeficientes_totales = []
8     frecuencias_totales = []
9     years = np.arange(1962, 2025, 2)
10
11    for wavelet in tipos_wavelet:
12        # Calcular la CWT
13        coeficientes, frecuencias = pywt.cwt(datos_valladolid_9,
14            escalas_en_anios, wavelet)
15        coeficientes_totales.append(coeficientes)
16        frecuencias_totales.append(frecuencias)
17
18        # Convertir a arrays para facilitar el manejo
19        coeficientes_totales = np.array(coeficientes_totales)
20        frecuencias_totales = np.array(frecuencias_totales)
21
22        vmax = np.max(coeficientes_totales)
23        vmin = np.min(coeficientes_totales)
24        # Creamos el colormap personalizado
25        cmap = create_custom_colormap(vmin, vmax)
26        # Graficar los coeficientes de la transformada para cada
27        # columna
28        plt.figure(figsize=(10, 6))
29        plt.imshow(coeficientes_totales.mean(axis=0), extent=[1962,
30            2024, min(escalas_en_anios), 30], aspect='auto', cmap=cmap,
31            origin='lower', vmax=vmax, vmin=vmin)
32        cbar = plt.colorbar(orientation='horizontal')
33
34        # Ocultar los ticks numéricos de la barra de color
35        cbar.set_ticks([0])
36
37        # Agregar texto personalizado para MIN y MAX
38        cbar.ax.text(-0.015, -0.5, 'MIN', color='black', ha='center',
39            va='center', transform=cbar.ax.transAxes)
40        cbar.ax.text(1, -0.5, 'MAX', color='black', ha='center', va='
41            center', transform=cbar.ax.transAxes)
42        plt.xticks(years, [str(year)[-2:] for year in years])
43        plt.xlabel('Años')
44        plt.ylabel('Periodo (Años)')
45        plt.title(f'SPI9 {wavelet.upper()} Valladolid')

```

```

40 plt.savefig(f'SPI9/WAVELET/SPI9{wavelet.upper()}Valladolid')
41 plt.show()
42 coeficientes_totales = []
43 frecuencias_totales = []

```

```

1     tipos_wavelet = ['mexh', 'morl']
2 escalas = np.arange(1, 768)
3
4 # Convertir las escalas a años
5 escalas_en_anios = escalas / 12.0 # 12 meses por año
6
7 coeficientes_totales = []
8 frecuencias_totales = []
9 years = np.arange(1962, 2025, 2)
10
11 for wavelet in tipos_wavelet:
12     # Calcular la CWT
13     coeficientes, frecuencias = pywt.cwt(datos_valladolid_12,
14     escalas_en_anios, wavelet)
15     coeficientes_totales.append(coeficientes)
16     frecuencias_totales.append(frecuencias)
17
18 # Convertir a arrays para facilitar el manejo
19 coeficientes_totales = np.array(coeficientes_totales)
20 frecuencias_totales = np.array(frecuencias_totales)
21
22 vmax = np.max(coeficientes_totales)
23 vmin = np.min(coeficientes_totales)
24 # Creamos el colormap personalizado
25 cmap = create_custom_colormap(vmin, vmax)
26 # Graficar los coeficientes de la transformada para cada
27     columna
28 plt.figure(figsize=(10, 6))
29 plt.imshow(coeficientes_totales.mean(axis=0), extent=[1962,
30     2024, min(escalas_en_anios), 30], aspect='auto', cmap=cmap,
31     origin='lower', vmax=vmax, vmin=vmin)
32 cbar = plt.colorbar(orientation='horizontal')
33
34 # Ocultar los ticks numéricos de la barra de color
35 cbar.set_ticks([0])
36
37 # Agregar texto personalizado para MIN y MAX
38 cbar.ax.text(0.015, -0.5, 'MIN', color='black', ha='center', va=
39     'center', transform=cbar.ax.transAxes)
40 cbar.ax.text(1, -0.5, 'MAX', color='black', ha='center', va='
41     center', transform=cbar.ax.transAxes)

```

```

36 plt.xticks(years, [str(year)[-2:] for year in years])
37 plt.xlabel('Años')
38 plt.ylabel('Periodo (Años)')
39 plt.title(f'SPI12 {wavelet.upper()} Valladolid')
40 plt.savefig(f'SPI12/WAVELET/SPI12{wavelet.upper()}Valladolid')
41 plt.show()
42 coeficientes_totales = []
43 frecuencias_totales = []

```

MÉTODO DEL CODO

```

1 # Estandarizar los datos
2 scaler = StandardScaler()
3 data_scaled_12mesesPeninsula = scaler.fit_transform(
4     data_12mesesPeninsula.iloc[:,2:])
5 # Calcular el centroide global
6 global_centroid = np.mean(data_scaled_12mesesPeninsula, axis=0)
7
8 # Calcular la variabilidad total (suma de las distancias al
9     cuadrado al centroide global)
10 total_variability = np.sum((data_scaled_12mesesPeninsula -
11     global_centroid) ** 2)
12
13 # Lista para almacenar la variabilidad explicada por diferentes
14     valores de k
15 explained_variability = []
16
17 # Calcular WCSS y la variabilidad explicada para diferentes valores
18     de k
19 wcss = []
20 wcss_iter = []
21 for i in range(1, 11):
22     kmeans = KMeans(n_clusters=i, init='k-means++', max_iter=300,
23         n_init=10, random_state=42)
24     kmeans.fit(data_scaled_12mesesPeninsula)
25     wcss = kmeans.inertia_
26     explained_variability.append((total_variability - wcss) /
27         total_variability)
28     wcss_iter.append(kmeans.inertia_)
29
30 # Convertir la variabilidad explicada a porcentaje
31 explained_variability_percentage_12meses = [ev * 100 for ev in
32     explained_variability]
33
34 # Plotear el gráfico del codo
35 plt.figure(figsize=(10, 8))
36 plt.plot(range(1, 11), wcss_iter, marker='o', linestyle='--')

```

```

29 plt.title('Método del Codo SPI12')
30 plt.xlabel('Número de Clusters')
31 plt.ylabel('WCSS')
32 plt.savefig('SPI12/MetodoCodoSPI12')
33 plt.show()
34
35 # Graficar la variabilidad explicada en porcentaje
36 plt.figure(figsize=(10, 6))
37 plt.plot(range(1, 11), explained_variability_percentage_12meses,
38          marker='o')
39 plt.title('SPI12 Variabilidad Explicada por Número de Clusters')
40 plt.xlabel('Número de clusters')
41 plt.ylabel('Variabilidad Explicada (%)')
42 plt.grid(True)
43 plt.savefig('SPI12/VariabilidadClusterSPI12')
44 plt.show()
45
46 # Mostrar los porcentajes de variabilidad explicada
47 for i, ev in enumerate(explained_variability_percentage_12meses, 1)
48     :
49     print(f'Número de Clusters: {i}, Variabilidad Explicada: {ev:.2
50           f}%')

```

IDENTIFICAR LOS CLUSTERS

```

1     meses = 12
2
3 # Cargar las coordenadas con el delimitador correcto
4 coords = pd.read_csv('Coordenadas.csv', delimiter=';')
5
6 # Reemplazar comas por puntos y convertir a float
7 coords['LATITUD'] = coords['LATITUD'].str.replace(',', '.').astype(
8     float)
9 coords['LONGITUD'] = coords['LONGITUD'].str.replace(',', '.').
10    astype(float)
11
12 # Eliminar filas basadas en la columna 'NOMBRE'
13 coords = coords[~coords['NOMBRE'].isin(['Santa Cruz de Tenerife',
14    'Las Palmas', 'Ceuta', 'Melilla', 'Baleares'])]
15
16 # Seleccionar el número óptimo de clusters
17 optimal_clusters = 6
18 kmeans = KMeans(n_clusters=optimal_clusters, random_state=42)
19 coords['cluster12meses'] = kmeans.fit_predict(
20     data_scaled_12mesesPeninsula.T)
21 fig, ax = plt.subplots(1,1, figsize=(15, 10), subplot_kw={'
22     projection': ccrs.PlateCarree()})

```

```

18 # Definir el área de visualización y agregar características del
    mapa
19 ax.set_extent([-10, 5, 35, 46])
20 ax.add_feature(cfeature.BORDERS, linewidth=0.5)
21 ax.add_feature(cfeature.COASTLINE, linewidth=0.5)
22
23 # Graficar las provincias con los clusters en diferentes colores
24 scatter = ax.scatter(coords['LONGITUD'], coords['LATITUD'], c=
    coords['cluster12meses'], cmap='Paired', s=500)
25 # Obtener elementos y etiquetas de la leyenda
26 handles, labels = scatter.legend_elements()
27 # Crear las etiquetas para la leyenda (del 1 al 5)
28 new_labels = [str(i + 1) for i in range(optimal_clusters)]
29 legend1 = ax.legend(handles, new_labels, title="Clusters")
30 ax.add_artist(legend1)
31 # Añadir etiquetas de nombres a los puntos
32 for i, row in coords.iterrows():
33     ax.annotate(row['NOMBRE'], (row['LONGITUD'], row['LATITUD']),
34                 fontsize=15, ha='center', va='center', color='black')
35
36 gl = ax.gridlines(crs=ccrs.PlateCarree(), draw_labels=True,
37                 linewidth=1, color='black', alpha=0.1, linestyle='dotted')
38 gl.top_labels = False
39 gl.right_labels = False
40 gl.xformatter = LONGITUDE_FORMATTER
41 gl.yformatter = LATITUDE_FORMATTER
42
43 # Ajustar diseño y mostrar
44 plt.tight_layout()
45 # Guardar la figura
46 plt.title('SPI'+str(meses)+' CL6'+ ' Variabilidad: '+str("{:.2f}".
47     format(explained_variability_percentage_12meses[5]))+'%')
48 plt.savefig('SPI12/SPI'+str(meses)+'CL6'+'.png', dpi=300,
49     bbox_inches='tight')
50 plt.show()

```

WAVELET DE CADA CLUSTER

```

1     # Agrupar las ciudades por cluster
2 clusters_ciudades = coords[['NOMBRE', 'cluster12meses']].groupby('
    cluster12meses')['NOMBRE'].apply(list)
3
4 tipos_wavelet = ['mexh', 'morl']
5 escalas = np.arange(1, 768)
6 media = []
7 # Convertir las escalas a años
8 escalas_en_anios = escalas / 12.0 # 12 meses por año

```

```

9 years = np.arange(1962, 2025, 2)
10 # Mostrar las ciudades en cada cluster
11 for cluster, ciudades in clusters_ciudades.items():
12     print(f"Cluster {cluster+1}: {ciudades}")
13     datos_cluster = data_12mesesPeninsula[ciudades]
14     datos_cluster_filtro = datos_cluster.loc[2:,:]
15     datos_cluster_filtro['Media'+str(cluster+1)] =
16         datos_cluster_filtro.mean(axis=1)
17     media.append(datos_cluster_filtro['Media'+str(cluster+1)])
18     print(datos_cluster_filtro)
19     coeficientes_totales = []
20     frecuencias_totales = []
21     # Obtener nombres de ciudades como cadena separada por comas
22     nombres_ciudades = ', '.join(ciudades)
23
24     for wavelet in tipos_wavelet:
25         # Calcular la CWT
26         coeficientes, frecuencias = pywt.cwt(datos_cluster_filtro['
27             Media'+str(cluster+1)], escalas_en_anios, wavelet)
28         coeficientes_totales.append(coeficientes)
29         frecuencias_totales.append(frecuencias)
30
31     # Convertir a arrays para facilitar el manejo
32     coeficientes_totales = np.array(coeficientes_totales)
33     frecuencias_totales = np.array(frecuencias_totales)
34     vmax = np.max(coeficientes_totales)
35     vmin = np.min(coeficientes_totales)
36     # Creamos el colormap personalizado
37     cmap = create_custom_colormap(-12,12)
38     # Graficar los coeficientes de la transformada para cada
39     columna
40     plt.figure(figsize=(10, 6))
41     plt.imshow(coeficientes_totales.mean(axis=0), extent=[1962,
42         2025,min(escalas_en_anios),30], aspect='auto', cmap=
43         cmap, origin='lower',vmax=12, vmin=-12)
44     cbar = plt.colorbar(orientation = 'horizontal')
45     # cbar = plt.colorbar(orientation='horizontal', pad=0.2)
46
47     # Ocultar los ticks numéricos de la barra de color
48     cbar.set_ticks([0])
49
50     # Agregar texto personalizado para MIN y MAX
51     cbar.ax.text(0.015, -0.5, 'MIN', color='black', ha='center',
52         va='center', transform=cbar.ax.transAxes)
53     cbar.ax.text(1, -0.5, 'MAX', color='black', ha='center', va
54         ='center', transform=cbar.ax.transAxes)

```

```
48     plt.xticks(years, [str(year)[-2:] for year in years])
49     plt.xlabel('Años')
50     plt.ylabel('Periodo (Años)')
51     plt.title(f'SPI12 {wavelet.upper()} Cluster {cluster+1}')
52     plt.savefig(f'SPI12/WAVELET/SPI12{wavelet.upper()}Cluster{
53         cluster+1}')
54     plt.show()
55     coeficientes_totales = []
56     frecuencias_totales = []
```

Bibliografía y Webgrafía

- [1] *Capítulo 3: Revisión de la teoría de wavelets*. (Última visita el: 19/08/2024). 2010. URL: http://catarina.udlap.mx/u_dl_a/tales/documentos/mel/gonzalez_g_ra/capitulo3.pdf.
- [2] Liliana Raquel Castro y Silvia Mabel Castro. “Wavelets y sus Aplicaciones”. En: *I Congreso Argentino de Ciencias de la Computación*. 1995.
- [3] José Alberto Mauricio. “Análisis de series temporales”. En: *Universidad Complutense de Madrid* (2007).
- [4] Universidad de Valladolid. *Series temporales*. (Última visita el: 02/09/2024). URL: <https://www5.uva.es/estadmed/datos/series/series.htm>.
- [5] Esic. “¿Qué son y cómo se analizan las series temporales?” En: (2024). URL: <https://www.esic.edu/rethink/tecnologia/series-temporales-que-son-c>.
- [6] *Series temporales*. (Última visita el: 02/09/2024). URL: <https://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/EDescrip/tema7.pdf>.
- [7] Ricardo Moya. *Selección del número óptimo de clusters*. (Última visita el: 24/07/2024). 2016. URL: <https://jarroba.com/seleccion-del-numero-optimo-clusters/>.
- [8] J Zúñiga. “El algoritmo k-means aplicado a clasificación y procesamiento de imágenes”. En: *Recuperado el 15* (2021).
- [9] Arturo de Pablo y Elena Romera. *CÁLCULO III. Apuntes. Grado en Ingeniería en Tecnologías Industriales. Tema 5*. (Última visita el: 04/09/2024).
- [10] William T Cochran et al. “What is the fast Fourier transform?” En: *Proceedings of the IEEE* 55.10 (1967), págs. 1664-1674.
- [11] Charles Van Loan. *Computational frameworks for the fast Fourier transform*. SIAM, 1992.
- [12] Mariano Maisonnave. (Última visita el: 20/08/2024). URL: <http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-Mariano%20Maisonnave.pdf>.
- [13] Sabando María Virginia. (Última visita el: 29/07/2024). URL: <http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-SabandoMariaVirginia.pdf>.

- [14] *Transformadas de Fourier: Definición y propiedades*. (Última visita el: 29/07/2024). 2015. URL: https://caminos.udc.es/info/asignaturas/master_iccp/miccp511/images/Imagenes_complementarios/resumen_transformada_Fourier4.pdf.
- [15] David Reig Albiñana. “Implementación de algoritmos para la extracción de patrones característicos en Sistemas de Reconocimiento De Voz en Matlab”. Tesis doct. Universitat Politècnica de València, 2015. URL: <https://riunet.upv.es/handle/10251/59092>.
- [16] Alfonso Fernández Sarriá. “Estudio de técnicas basadas en la transformada wavelet y optimización de sus parámetros para la clasificación por texturas de imágenes digitales”. Tesis doct. Universitat Politècnica de València, 2008.
- [17] Elena Rupérez Cerezo. *Wavelets. Proyecto Sistemas Informáticos*. 2012.
- [18] Samir Kouro y Rodrigo Musalem. “Tutorial introductorio a la Teoría de Wavelet”. En: *Chile: Universidad Técnica Federico Santa María* (2002).
- [19] Grupo de Investigación en Sistemas Inteligentes. Facultad de Estudios Superiores Cuautitlán. Universidad Nacional Autónoma de México. *Transformada wavelet*. (Última visita el: 14/08/2024). 2018. URL: https://virtual.cuautitlan.unam.mx/intar/?page_id=1108.
- [20] Natalia Nieto y Diana Marcela Orozco. “El uso de la transformada wavelet discreta en la reconstrucción de señales senosoidales”. En: *Scientia et technica* 14.38 (2008), págs. 381-386.
- [21] Elene Rupérez Cerezo. *Wavelets*. (Última visita el: 21/08/2024). 2012. URL: <https://docta.ucm.es/rest/api/core/bitstreams/45678c73-8b81-4c35-8ee2-354aa980ba3d/content>.
- [22] AEMET. *Interpretación del Índice de Precipitación Estandarizado (SPI)*. (Última visita el: 09/08/2024). URL: https://www.aemet.es/es/serviciosclimaticos/vigilancia_clima/vigilancia_sequia/ayuda#:~:text=El%20%C3%ADndice%20SPI%20se%20define%20a%20una%20distribuci%C3%B3n%20normal..
- [23] Guillermo Podestá et al. “Descripción de índices para el monitoreo de sequía meteorológica implementados por el Centro Regional del Clima para el Sur de América del Sur”. En: *Reporte Técnico CRC-SAS-2015-001. Centro Regional del Clima para el Sur de América del Sur. Buenos Aires, Argentina* (2015).
- [24] M Svoboda, M Hayes y D Wood. “Índice normalizado de precipitación”. En: *Guía del Usuario. OMM 1090* (2012).
- [25] *Visual Studio Code*. (Última visita el: 09/07/2024). URL: https://code.visualstudio.com/?wt.mc_id=DX_841432.
- [26] Jupyter. *Project Jupyter*. (Última visita el: 09/07/2024). URL: <https://jupyter.org/>.

- [27] Inc. o sus filiales. Amazon Web Services. *¿Qué es Python?* (Última visita el: 09/07/2024). URL: <https://aws.amazon.com/es/what-is/python/#:~:text=Python%20es%20un%20lenguaje%20de%20ejecutar%20en%20muchas%20plataformas%20diferentes..>
- [28] Soluciones Inába. “Bibliotecas, Librerías y Módulos en Programación: Diferencias e Importancia”. En: *Categoría general. Desarrollo Web*. (2023). URL: <https://www.inabaweb.com/bibliotecas-librerias-y-modulos-en-programacion-diferencias-e-importancia/#:~:text=En%20resumen%2C%20mientras%20que%20las,unidades%20de%20organizaci%C3%B3n%20y%20encapsulamiento..>
- [29] Numpy. *Numpy*. (Última visita el: 09/07/2024). URL: <https://numpy.org//>.
- [30] Pandas. *Pandas documentation*. (Última visita el: 09/07/2024). URL: <https://pandas.pydata.org/pandas-docs/stable/index.html>.
- [31] Matplotlib development team. *Matplotlib: Visualization with Python*. (Última visita el: 09/07/2024). URL: <https://matplotlib.org/>.
- [32] *La librería Matplotlib*. (Última visita el: 09/07/2024). URL: <https://aprendeconalf.es/docencia/python/manual/matplotlib/>.
- [33] AEMET. *Avance Climático Nacional del año 2022*. (Última visita el: 06/08/2024). 2023. URL: <https://www.aemet.es/documentos/es/noticias/2023/Avanceclimaticonaciona2022.pdf>.
- [34] Germán Portillo. *Sequías en España*. (Última visita el: 27/08/2024). 2024. URL: <https://www.ecologiaverde.com/sequias-en-espana-4431.html#:~:text=Las%20peores%20sequ%C3%ADas%20en%20Espa%C3%B1a%20fueron%20la%20sequ%C3%ADa%20de%201749,pa%C3%ADs%20y%20en%20su%20gente..>
- [35] *Fenómeno de el Niño*. (Última visita el: 23/08/2024). URL: <https://www.eltiempo.es/noticias/meteorologia/fenomeno-de-el-nino>.
- [36] Carlos Fernández Freire y Julio Aristizábal Arteaga. “El niño”. En: (2015).
- [37] Ricardo Garcia Herrera et al. “Influence of the North Atlantic oscillation on the Canary Islands precipitation”. En: *Journal of Climate* 14.19 (2001), págs. 3889-3903.
- [38] John M Wallace y Peter V Hobbs. *Atmospheric science: an introductory survey*. Vol. 92. Elsevier, 2006.
- [39] Dennis L Hartmann. *Global physical climatology*. Vol. 103. Newnes, 2015.
- [40] METEORED. *Los impactos de El Niño y de La Niña*. (Última visita el: 23/08/2024). URL: <https://www.tiempo.com/ram/123032/los-impactos-de-el-nino-y-de-la-nina/>.

- [41] *¿Qué Es El Fenómeno Del Niño Y La Niña Y Como Afecta A España?* (Última visita el: 23/08/2024). URL: <https://agrarfertilizantes.es/que-es-el-fenomeno-del-nino-y-la-nina-y-como-afecta-a-espana/#:~:text=El%20fen%C3%B3meno%20del%20Ni%C3%B1o,-El%20fen%C3%B3meno%20del&text=Estos%20efectos%20incluyen%20sequ%C3%ADas%20en,el%20clima%20y%20la%20meteorolog%C3%ADa..>
- [42] Josué M Polanco Martínez. “Análisis espectral de wavelet: una concisa revisión.” En: *Revista de Climatología* 22 (2022).
- [43] Ana Morata Gasca et al. “Análisis wavelet de la precipitación en la Península Ibérica”. En: (2001).