

eman ta zabal zazu



Universidad Euskal Herriko
del País Vasco Unibertsitatea

Facultad de Ciencia y Tecnología
Departamento de Física Aplicada II

Aplicación de técnicas estadísticas en el estudio de fenómenos ambientales y ecosistémicos

Memoria presentada para optar al Grado de Doctor por:

Josué Moisés Polanco Martínez

Dirigida por: Dr. Jon Saénz Aguirre & Dr. Unai Ganzedo López

Leioa, Septiembre de 2011.

josuemoises.polanco@ehu.es

© Servicio Editorial de la Universidad del País Vasco (UPV/EHU)
- Euskal Herriko Unibertsitateko Argitalpen Zerbitzua (UPV/EHU)
- University of the Basque Country Press (UPV/EHU)

ISBN: 978-84-9860-812-0

Resumen

En esta memoria se presentan cuatro casos de estudio que tratan sobre tres temas relacionados con los fenómenos ambientales y ecosistémicos. El primer tema está relacionado con la contaminación atmosférica y la calidad del aire, el segundo con el clima del Jurásico tardío, los paleoindicadores biogeoquímicos y los ciclos orbitales de la Tierra y el tercero con las fluctuaciones de especies marinas y su relación con el clima. El factor común entre los diferentes casos de estudio, es la física del clima y el uso de sofisticadas técnicas estadísticas de análisis de datos que permitan desentrañar algunas de las problemáticas presentes en los datos ambientales y ecosistémicos (como el ruido, el no equiespaciado temporal, la no estacionariedad, etc.). Esto, con el fin de identificar procesos físicos de interés.

En el primer caso de estudio se evalúa una red de control y vigilancia de la calidad del aire debida a inmisiones de SO_2 en Bilbao metropolitano durante el periodo 1996-2001 mediante el Análisis de Componente Principales. En el segundo, se realiza una búsqueda de señal orbital en isótopos estables (valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$) de belemnites recolectados en la cuenca Vasco-Cantábrica pertenecientes al Jurásico tardío mediante el Periodograma de Lomb-Scargle. En el tercero, se cuantifican los efectos de las variables climáticas en las fluctuaciones de capturas de atún rojo en el sur de España entre 1525-1756 por análisis de devianza mediante Modelos Lineales Generalizados. Por último, se analiza la relación entre las capturas del pulpo común de las Islas Canarias y las variables climáticas (temperatura superficial del mar y el índice de la Oscilación del Atlántico Norte) durante el periodo 1989-2007 utilizando análisis espectral vía la Transformada Continua de Wavelet.

Agradecimientos

Oficiales:

- Al H. A. de Tizimín, Yucatán, México (período 2004-2007) y a la Secretaría de Educación Pública del Estado de Yucatán (período 2001-2007) por la financiación para mi desplazamiento inicial.
- A la Universidad del País Vasco/Euskal Herriko Unibersitatea por la financiación durante el período 2006-2009 para la realización de esta tesis doctoral.
- Al proyecto GIC07/53-IT-334-07 del Gobierno Vasco mediante el grupo *UPV/EHU Econometrics Research Group* del cual he obtenido financiación desde octubre de 2009 hasta el presente.
- Al Instituto de Economía Pública de la UPV/EHU por brindar un espacio para la investigación.

Profesionales:

A los directores de esta memoria Dr. Prof. Jon Sáenz Aguirre y Dr. Unai Ganzedo López por su atinada dirección de este trabajo y de quienes he aprendido muchas cosas, lo cual no se limita sólo al ámbito profesional.

Al Dr. Prof. Sergio Robles Orozco (EHU/UPV), al Dr. Prof. Gabriel Ibarra Berastegui (EHU/UPV), al Dr. Prof. José J. Castro Hernández (ULPGC), a la Dra. Angela M. Caballero Alfonso (ULPGC), al Dr. Eduardo Zorita Calvo (GKSS) y a la Dra. Idoia Rosales Franco (IGME) por su inestimable ayuda en varios de los trabajos de investigación aquí presentados.

A los profesores y compañeros del dpto. de Física Aplicada II, Materia Condensada, Economía Aplicada III y del Instituto de Economía Pública de la EHU/UPV.

Al director del dpto. de Física Aplicada II de la EHU/UPV, Dr. Prof. Carlos Santamaría y al profesor (actualmente felizmente jubilado) Dr. Prof. Eduardo Bocanegra, quienes me dieron la bienvenida hace 6 años.

Al Dr. Prof. Javier Fernández Macho del dpto. de Economía Aplicada III y director del Instituto de Economía Pública de la EHU/UPV.

Se agradece a Harish Bhanderi y al dpto. de Ingeniería de la Universidad de Cambridge por permitir utilizar su código fuente LaTeX para escribir este documento.

Especiales:

A mis padres (y aunque uno de ellos ya no está presente, estoy seguro le hubiera gustado compartir conmigo este logro).

De manera muy especial a María Arrugaeta Lage que ha estado en las buenas, en las malas y en las peores, y por supuesto a su familia (en particular a sus padres y abuelos).

A mis amigos Berta Ibáñez, Guillermo Aravena, Jorge Feuchtwanger, Txema Posse, Mikel Eguiraun, Abdu Faik, Juan Carlos, Gaizka.

A este pedacito de tierra (y a su maravilloso tiempo atmosférico) y, sobre todo, a su gente que me han hecho sentir como en casa.

A todas las personas que de una u otra forma han hecho posible este trabajo, pero que en estos momentos por las prisas, olvido mencionar o simplemente no hago para evitar una lista que podría ser más extensa que la tesis misma.

Índice general

Resumen	III
Agradecimientos	VII
Lista de abreviaciones	VII
1. Introducción	1
1.1. Principales características de los sistemas ambientales	1
1.2. Las aproximaciones para estudiar los sistemas ambientales	3
1.2.1. La modelización numérica determinista	4
1.2.2. La modelización numérica probabilística	6
1.2.3. El análisis estadístico de datos	7
1.3. Problemática en el análisis de datos	8
1.4. Motivación y antecedentes de esta memoria	13
1.5. Técnicas estadísticas utilizadas	25
1.6. Programas computacionales utilizados	27
1.7. Objetivos y estructura de esta memoria	28
2. Principales aspectos metodológicos	30
2.1. Análisis de Componentes Principales	30
2.1.1. Calculando las componentes principales	32
2.1.2. Interpretación de las componentes principales	35
2.1.3. Sobre el número de componentes principales	37
2.1.4. Sobre el uso del PCA para identificación de modos	38
2.2. Análisis espectral vía el Periodograma de Lomb-Scargle .	39
2.2.1. Definición del periodograma de Lomb-Scargle	40

2.2.2.	El periodograma clásico vs. el PLS	41
2.2.3.	Estimación de las frecuencias naturales	43
2.2.4.	El periodograma: crudo vs. suavizado	45
2.2.5.	Contraste de hipótesis del PLS debido a ruido rojo	49
2.3.	Análisis de devianza mediante Modelos Lineales Generalizados	52
2.3.1.	Los modelos lineales generalizados	53
2.3.2.	La componente temporal	54
2.3.3.	La partición de la varianza	55
2.4.	Análisis espectral vía la transformada continua de Wavelet	57
2.4.1.	Definición formal de Wavelet	58
2.4.2.	La transformada continua de Wavelet	60
2.4.3.	La estimación del espectro	65
2.4.4.	La estimación del espectro cruzado	69
2.4.5.	Pruebas de significación	71
3.	Evaluación de una red de control de la calidad del aire (inmisión de SO₂) entre 1996-2001 en Bilbao metropolitano por medio del PCA	73
3.1.	Introducción	73
3.1.1.	Factores meteorológicos que actúan en la dispersión y en el transporte del SO ₂	76
3.1.2.	Principales aproximaciones para estudiar la dispersión y el transporte del SO ₂	79
3.1.3.	Técnicas estadísticas utilizadas para la evaluación de redes de control de la calidad del aire	81
3.2.	Area de estudio y antecedentes	82
3.3.	Datos y métodos	85
3.4.	Resultados	87
3.5.	Discusión	90

4. Búsqueda de señal orbital en isótopos estables de belemnites del Jurásico tardío de la cuenca Vasco-Cantábrica mediante el PLS	92
4.1. Introducción	92
4.2. Paleogeografía y paleoceanografía	96
4.3. Material y métodos	97
4.3.1. Series temporales de valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$	97
4.3.2. Métodos	101
4.4. Resultados	105
4.5. Discusión	106
 5. Principales causas de las fluctuaciones de capturas de atún rojo entre 1525-1756 para el sur de España explicado por análisis de devianza mediante GLM	 113
5.1. Introducción	113
5.2. Datos y métodos	117
5.2.1. Serie temporal de capturas de atún rojo	117
5.2.2. Datos climáticos	118
5.2.3. Métodos	121
5.3. Resultados	124
5.3.1. Análisis descriptivo de la serie de capturas	124
5.3.2. Análisis de correlaciones entre los datos	124
5.3.3. Resultados del análisis de devianza mediante GLM	125
5.4. Discusión	128
 6. La relación entre variables climáticas y las fluctuaciones de capturas de pulpo común entre 1989-2007 en las Islas Canarias mediante la CWT	 132
6.1. Introducción	132
6.2. Datos y preprocesamiento	137
6.2.1. Serie temporal de capturas del <i>Octopus vulgaris</i>	137
6.2.2. Datos climáticos	139
6.3. Metodología	141
6.4. Resultados	145
6.5. Discusión	155

ÍNDICE GENERAL

7. Conclusiones y trabajo futuro	160
A. Lista de publicaciones en las que ha participado el autor de esta memoria	166
Referencias	208

Lista de abreviaciones utilizadas en esta memoria

a.	Años
ANOVA	Analysis of variance
AO	The Arctic Oscillation
AOGCM	Atmosphere-Ocean General Circulation Model
AQMN	Air Quality Monitoring Network
BCB	The Basque-Cantabrian Basin
CA	Cluster analysis
CCF	Cross-Correlation Function
CPUE	Catch Per Unit Effort
CSM	Climate System Model
CWA	Cross Wavelet Analysis
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DOF	Degrees of freedom
EMICS	Earth-System Models of Intermediate Complexity
EOF/PCA	Empirical Orthogonal Functions/Principal Component Analysis
GHG	Greenhouse gases
GLM	Generalized linear model
LSFT	Lomb-Scargle Fourier Transform
LSP/PLS	Lomb-Scargle Periodogram/Periodograma de Lomb-Scargle
LSSA	Least-Squares Spectral Analysis
MLE	Maximum Likelihood Estimation
NAO	North Atlantic Oscillation
NOAA	National Oceanic and Atmospheric Administration
OAE	Oceanic Anoxic Event
PBL	Planetary Boundary Layer
PC/CP	Principal Component/Componente Principal
PDB	Pee Dee Belemnite
PDF	Probability Density Function
SOM	Self Organizing Map (Kohonen map)
SVD	Singular Value Decomposition
TLS	Transcontinental Laurasian Seaway
WCO	Wavelet Coherence
WOSA	Welch's Overlapped Segment Averaging

Capítulo 1

Introducción

1.1. Principales características de los sistemas ambientales

Los sistemas ambientales son sistemas dinámicos, por tanto, los procesos físicos, químicos, biológicos y geológicos que suceden en ellos pueden ocurrir en diferentes escalas espacio-temporales [260, 363, 365]. Estas escalas son unas medidas representativas de las variaciones en el tiempo y el espacio de los procesos ambientales y, ambas son necesarias para hacer una adecuada descripción de estos procesos [260, 365]. Las escalas espaciales y temporales abarcan desde la escala atómica y molecular hasta unos cuantos miles de kilómetros o millones de años [260, 363, 365].

Los sistemas ambientales no están aislados debido a que están bajo la influencia de diversos forzamientos externos a ellos. Estos forzamientos pueden tener un origen astronómico, como las variaciones en la luminosidad solar [260, 328] o la variación de los parámetros orbitales de la Tierra [23, 63], o también pueden tener un origen antropogénico. Investigaciones realizadas en las últimas décadas han mostrado que las actividades humanas relacionadas con la utilización de combustibles fósiles y al uso de la tierra, están aumentando considerablemente las concentraciones atmosféricas de CO_2 , CH_4 y NO , entre otros compuestos. Como

1.1 Principales características de los sistemas ambientales

consecuencia de ello se está alterando el balance de energía del sistema climático y está aumentando la temperatura global promedio del aire y del océano [228, 337].

Los sistemas ambientales también pueden estar sujetos a las influencias de otras componentes de otros sistemas ambientales [228]. Estos sistemas pueden interactuar entre sí, esto es, son sistemas abiertos. Cuando interactúan hay intercambio de flujos de masa, energía y momento. La interacción es a todos los niveles espaciales y temporales, es decir, pueden interactuar desde los pequeños sistemas ambientales hasta las componentes del sistema climático, como la interacción océano-atmósfera. Sin embargo, el sistema climático global, tomado como el conjunto de todos sus subsistemas, se considera un sistema cerrado [76, 256].

El sistema climático es cerrado, por lo que los flujos de masa se mantienen en un bucle cerrado. Sin embargo, asociados a estos flujos de masa (el caso más claro es el flujo de agua entre atmósfera, océano, criosfera y litosfera), aparecen transferencias muy importantes de energía entre los diferentes subsistemas (flujo de calor latente) y hacia y desde el exterior (en forma de albedo y radiación saliente de onda larga, por ejemplo). Por todo ello, aparecen ciclos de realimentación entre los diferentes subsistemas del sistema climático [256, 285]. De este modo, cada subsistema afecta la respuesta de los otros subsistemas y por tanto del sistema climático. Estas interacciones, entre las diferentes componentes del sistema climático, muestran características típicas de un sistema no lineal. Esto es, las interacciones múltiples entre diferentes componentes no satisfacen el principio de superposición [256, 285, 363].

La no linealidad que existe, debido a las interacciones entre sistemas o subsistemas ambientales, aparece reflejada en las ecuaciones diferenciales que representan diferentes procesos fundamentales, ya sean físicos, químicos o biológicos [288]. Quizá, el ejemplo más conocido, sean los términos no lineales de advección en las ecuaciones utilizadas para representar el movimiento del aire o del océano [260, 365] o los términos no lineales que aparecen en los modelos de poblaciones [199].

La no linealidad de los sistemas ambientales trae consigo varias consecuencias, como la creación de estructuras o eventos no esperados en la forma de transiciones abruptas. Un ejemplo típico, son los cambios climáticos abruptos que alternan periodos glaciales e interglaciales, como los ocurridos durante el último periodo

1.2 Las aproximaciones para estudiar los sistemas ambientales

glacial, conocidos como las oscilaciones de Dansgaard-Oeschger (DO) [6, 285]. Otra de las consecuencias que aparece debido a la existencia de forzamientos externos de tipo oscilatorio, es que la respuesta espectral de un sistema no lineal, puede exhibir frecuencias no presentes en el forzamiento [285]. Los sistemas no lineales son propensos a presentar un comportamiento caótico. Esto es, pequeños cambios en las condiciones iniciales pueden causar grandes diferencias cuantitativas en el resultado (*el efecto mariposa*) [195, 285].

Otra de las características de los sistemas ambientales que los diferencia de otras áreas de la ciencia como la química, la biología o la medicina, es la imposibilidad de realizar experimentos en condiciones controladas de laboratorio, donde se tengan en cuenta todas las variables y procesos de un sistema ambiental [228]. Es cierto que, bajo ciertas condiciones muy restrictivas, es posible hacer experimentos de laboratorio, sin embargo, no se puede tener en cuenta todo el sistema [363]. La opción de experimentar directamente y a gran escala queda descartada, porque independientemente de las implicaciones éticas o la inviabilidad técnica o económica, como son sistemas altamente complejos, no es posible saber cómo serían sus reacciones ni el riesgo medioambiental que la experimentación conllevaría [43, 314, 359].

Otra propiedad de los sistemas ambientales es que son únicos. Esto es, no existen dos sistemas ambientales que sean lo suficientemente parecidos como para inferir las propiedades de uno del otro o viceversa [228]. Un sistema ambiental nunca repetirá un estado posterior idéntico a uno anterior, debido a las múltiples transferencias difusivas de energía, masa o momento que dotan a los sistemas de su carácter caótico [228, 363].

1.2. Las aproximaciones para estudiar los sistemas ambientales

Las características de los sistemas ambientales, mencionadas en los párrafos anteriores, proporcionan una idea de lo complejo que es estudiar y entender los fenómenos y los procesos ecosistémicos y ambientales. Tradicionalmente, el estudio de estos fenómenos ha sido de tipo descriptivo. Sin embargo, esta aproxi-

1.2 Las aproximaciones para estudiar los sistemas ambientales

mación no explica los mecanismos causales que dan lugar a estos fenómenos. Otra de las herramientas más utilizadas en el estudio de fenómenos ecosistémicos, en particular, con el fin de averiguar el origen de las fluctuaciones de los stocks de peces y otras especies marinas (como los moluscos), son los modelos de dinámica de poblaciones. Sin embargo, muchos de estos modelos o no tienen en cuenta los forzamientos ambientales (como los meteorológicos, oceanográficos o climáticos) o no siempre se dispone de todos los parámetros necesarios de entrada de estos modelos (como la tasa de nacimiento y mortalidad natural, madurez, etc.) [22, 300].

En principio, los sistemas ambientales están gobernados por leyes o principios físicos, químicos o biológicos fundamentales, y aunque son considerados sistemas deterministas, la presencia de procesos caóticos crea un tipo de variabilidad que no permite distinguirla de las variaciones aleatorias (ruido) [363, 365]. Por ello es necesario un conjunto de herramientas adecuadas que puedan manejar estos inconvenientes. Por un lado, y especialmente con la llegada de los ordenadores en la segunda mitad del siglo XX, la modelización (simulación) numérica en cierto modo ha suplido el vacío de no poder hacer experimentos en condiciones controladas [228, 260, 363]. Por otro lado, han sido necesarias herramientas que utilicen ideas y técnicas probabilísticas y estadísticas para describir y estudiar adecuadamente a los sistemas ambientales. Sin embargo, la combinación de estas aproximaciones permite obtener mejores resultados en la investigación de fenómenos, procesos y sistemas ambientales [76, 345, 363, 365].

1.2.1. La modelización numérica determinista

La modelización numérica determinista es un tipo de experimentación virtual que es considerada como una herramienta indispensable en el estudio de fenómenos, procesos y sistemas ambientales [260, 267, 363]. Entiéndase por modelización numérica determinista al procedimiento descrito por [240], esto es: *fenómeno/proceso/sistema ambiental (propuesto o existente) para ser modelado, modelo conceptual del sistema ambiental, modelo matemático del modelo conceptual, discretización y selección algorítmica del modelo matemático, programación (codificación) del modelo discretizado, solución numérica del programa computacional,*

1.2 Las aproximaciones para estudiar los sistemas ambientales

representación e interpretación de la solución numérica.

Existe una amplia variedad de modelos numéricos deterministas que pueden utilizarse para estudiar los fenómenos, procesos y sistemas ambientales. Sin embargo, al estudiar el sistema climático, estos modelos suelen agruparse principalmente en dos categorías. Los modelos conceptuales (también conocidos como cognitivos) y los cuasi-realísticos [228, 363]. Los primeros, son tipos de modelos altamente simplificados e idealizados y sólo describen los procesos físicos, químicos o biológicos de primer orden. Están representados por medio de ecuaciones matemáticas, las cuales requieren el uso de ordenadores para ser resueltas [227, 363]. Por ejemplo, el modelo del sistema caótico de Lorenz o el modelo de Lotka-Volterra [195, 363].

Los segundos, intentan reproducir la realidad lo más cercana posible, sujeta a las limitaciones conceptuales y a los medios computacionales disponibles [216, 363]. Estos modelos han evolucionado de lo más simple a lo más complejo y su evolución está estrechamente relacionada con el desarrollo de las tecnologías computacionales (en especial con la capacidad de procesamiento y de almacenamiento de los ordenadores) [227]. Están clasificados de acuerdo a su nivel de complejidad, esto es, de acuerdo a su resolución espacial y al número de componentes del sistema climático que representan [363].

De modo general, se pueden considerar tres tipos de estos modelos. Los modelos de circulación general de la atmósfera o del océano (conocidos como A/O GCMs) (*v. gr.* el ECHAM5 o el MOM4p1) [112, 216, 295], los modelos de complejidad intermedia (conocidos como EMICs) que representan de modo detallado alguna componente del sistema climático, por ejemplo, el océano, mientras que parametrizan la dinámica de las otras componentes (*v. gr.*, el CLIMBER-2 o el Planet Simulator) [56, 94, 268], y los modelos acoplados multicomponente {Océano + Atmósfera + Criosfera + procesos terrestres (ciclos biogeoquímicos) + aerosoles} de circulación general, conocidos como Climate System Models (CSM) (*v. gr.*, el GISS GCM ModelE) que representan el estado del arte de la modelización del sistema climático [216, 313].

La modelización numérica determinista, sin lugar a dudas ha aportado mucho al entendimiento de los sistemas ambientales, en especial del sistema climático [216]. Sin embargo, los modelos numéricos deterministas no son perfectos porque

1.2 Las aproximaciones para estudiar los sistemas ambientales

presentan ciertas críticas e inconvenientes. Una de las principales críticas viene de la filosofía de la ciencia. Esto es, los modelos numéricos de sistemas ambientales no pueden ser verificados ni validados en el sentido matemático del término. Aunque en ciertos casos es posible corroborar los resultados (predicciones) de las simulaciones con observaciones, estas confirmaciones son parciales. Los modelos numéricos sólo pueden ser evaluados en términos relativos y su valor predictivo siempre está abierto a discusión [248, 341, 363].

En este sentido se han sugerido nuevos enfoques en la filosofía de la ciencia con lo que respecta a la modelización (simulación) numérica determinista de sistemas ambientales, de modo específico a los modelos cuasi-realísticos del sistema climático [26, 267]. Estos enfoques están relacionados en cómo manejar la complejidad y la jerarquía de los modelos, la falsabilidad (prueba de hipótesis), el límite de la predictibilidad, la incertidumbre y la evaluación de las simulaciones [26, 267, 326, 333, 341].

Cada uno de estos enfoques presenta diferentes tipos de problemas. Por ejemplo, no existe ningún modelo ambiental determinista que pueda resolver de forma explícita, en base a ecuaciones que representan primeros principios teóricos, todos los procesos que suceden en escalas más pequeñas que la malla del modelo y por tanto, la parametrización siempre es necesaria [267, 326]. Es imposible tener en cuenta todas las variables del sistema y existen procesos que no están bien entendidos, lo cual lleva a hacer simplificaciones [288]. Podrían existir procesos relevantes en escalas temporales mayores que no están considerados en los modelos [267, 288]. Problemas debido a incertidumbres en las condiciones iniciales, en los datos que alimentan los modelos, en la estructura del modelo relacionada con los procesos ambientales resueltos y no resueltos, y en los aspectos técnicos del modelo (*v. gr.*, en la discretización o en la codificación de los algoritmos) [267, 341].

1.2.2. La modelización numérica probabilística

Desde la aparición de la teoría del caos, se conoce bien que los sistemas afectados de sensibilidad a las condiciones iniciales no pueden predecirse de forma

1.2 Las aproximaciones para estudiar los sistemas ambientales

perfecta desde un punto de vista estrictamente determinista. Por tanto, se ha desarrollado el concepto de predicción probabilística [85, 186], en la cual se produce una predicción del estado de la atmósfera, o una predicción hidrológica o el tipo de predicción ambiental que se desee y una estimación de la incertidumbre, de tal manera que el usuario final pueda evaluar con más información el valor de la predicción [72, 169, 170, 171].

La generación de predicciones probabilísticas tiene muchas variantes, como pueden ser los métodos frecuentistas basados en estimaciones de funciones de densidad de probabilidad paramétricas o experimentales [117, 241], métodos de Monte Carlo aplicados a las salidas finales del proceso predictivo [210] o métodos dinámicos basados en la identificación de las direcciones de máxima inestabilidad del operador de predicción, por medio de vectores singulares o vectores *bred*, que se utilizan ampliamente en los servicios meteorológicos mundiales en la actualidad. Una alternativa reside en el uso de métodos bayesianos [57, 170, 171, 343] de calibración y asimilación de predicciones probabilísticas, ya que el teorema de Bayes permite formular los conceptos probabilísticos de una forma muy natural.

1.2.3. El análisis estadístico de datos

Las técnicas estadísticas han sido tradicionalmente utilizadas en el análisis de datos ambientales y ecosistémicos para identificar las relaciones fundamentales entre componentes o para asignar intervalos de confianza a afirmaciones científicas ante la certeza de que las medidas tienen un nivel de ruido no despreciable. Incluso en algunos casos, se han desarrollado nuevas técnicas estadísticas para estudiar ciertos fenómenos ambientales. Así fue con el análisis armónico que formuló Lord Kelvin en 1867 para estudiar las mareas, o el análisis espectral introducido por Schuster en 1890 y más tarde formalizado por Tukey en 1965 para el estudio de series temporales de eventos sísmicos y de actividad solar [76]. Otro de los ejemplos más significativos entre la relación del uso y el desarrollo de técnicas estadísticas y el estudio de fenómenos ambientales, fue el estudio pionero de Sir Gilbert Walker sobre El Niño y la Oscilación del Sur y la búsqueda de un modelo estadístico que exhibiera un comportamiento cuasi-oscilatorio, lo cual condujo, en colaboración con George U. Yule, a las ecuaciones de Yule-Walker [160].

1.3 Problemática en el análisis de datos

Las técnicas estadísticas permiten hacer una descripción de cualquier conjunto de datos a estudio, hacer inferencias o identificar relaciones entre variables de un sistema ambiental que podrían no ser evidentes debido a la complejidad de estos sistemas [345, 363]. Sin embargo, hay que tener en cuenta que, cualquier conjunto de datos ambientales y ecosistémicos, contiene un grado de incertidumbre (error) porque representa una discretización muestral espacio-temporal de un sistema continuo. Por consiguiente, siempre existirán escalas (espaciales o temporales) de variabilidad que no estén resueltas, las cuales contaminan los datos [225, 345, 392]. Por estas razones, el uso de técnicas estadísticas es adecuado porque ayuda a cuantificar los efectos de la incertidumbre, ya sea en términos de observación y medidas o en términos de nuestro entendimiento de los procesos ambientales [345, 376, 392], ya que permite centrar el análisis en aquellas relaciones que superen el umbral del ruido.

La aplicación de técnicas estadísticas es de carácter universal porque se pueden aplicar a cualquier tipo de datos ambientales o ecosistémicos. Tales como los datos observacionales, los obtenidos de medidas instrumentales, de sensores remotos, de reconstrucciones paleoambientales, de fuentes documentales, de salidas de simulaciones numéricas de modelos climáticos, etc. [189, 363, 376, 392]. Y es esta aproximación la que se ha utilizado en los cuatro casos de estudio presentados en esta memoria.

1.3. Problemática en el análisis de datos

El análisis estadístico de datos ambientales y ecosistémicos, como cualquier otra herramienta científica, no está libre de problemas. Existen varios inconvenientes propios de los datos ambientales y ecosistémicos que pueden complicar la aplicación de las técnicas estadísticas y la interpretación de resultados. A continuación se indican algunos de los problemas o inconvenientes que han estado presentes en los casos de estudio mostrados en esta memoria.

Ruido. Desde un punto de vista general, una serie temporal ambiental o ecosistémica puede considerarse como la suma de una componente dinámica de tipo determinista (señal) más otra componente de origen estocástico (ruido) [365, 385]. Uno de los objetivos principales en el análisis estadístico de series temporales es

detectar y describir las características de la componente dinámica cuando una componente estocástica está presente, con el fin de discernir entre señal y ruido [345, 365]. Sin embargo, para poder determinar el tipo de ruido subyacente en un sistema ambiental, hay que tener en cuenta las características físicas del comportamiento del sistema [365, 392].

En lo que respecta al sistema climático, Hasselmann [124] demostró que el tipo de ruido de fondo de este sistema es rojo [124, 364]. El ruido rojo es común encontrarlo en una amplia variedad de registros climáticos, así como en los resultados generados por medio de simulaciones numéricas del clima [107, 203, 317, 367]. Sin embargo, es importante recalcar que el ruido rojo es uno de los posibles tipos de espectro que puede presentar el sistema climático. El tipo de ruido de fondo puede variar dependiendo de la escala temporal o de la velocidad de respuesta de las componentes del sistema climático. Por ejemplo, para el océano profundo (por debajo de los 1000 m) o para eventos climáticos que contienen oscilaciones altamente no lineales, como los eventos Dansgaard-Oeschger, sus espectros de fondo no parecen ser rojos [40, 364, 366].

Series temporales cortas y pocos elementos. Otra de las características de las series temporales ambientales y ecosistémicas es que suelen ser cortas y tienen pocos elementos en comparación con la escala temporal del fenómeno a estudio [371, 385]. Como consecuencia de la corta longitud de las series, los análisis estadísticos no siempre tienen un amplio margen de confianza [371, 385]. Por otro lado, la disponibilidad de muy pocos elementos de una serie temporal implica no solamente una pobre resolución temporal, sino también espectral [349, 371]. Un problema adicional asociado con la pobre resolución temporal o espacial de los datos es que se puede producir el fenómeno del *aliasing* [273, 386].

Aliasing. Una serie temporal ambiental o ecosistémica es una representación discreta de una variable continua, donde cada elemento de la serie es obtenido en cada intervalo temporal Δt constante. Como una consecuencia de la discretización del tiempo con un periodo de muestreo Δt demasiado grande, surge el efecto del aliasing [365, 371]. Este efecto también puede aparecer en el muestreo de campos espaciales, pero en esta memoria nos vamos a referir siempre al caso temporal.

El efecto del aliasing puede ser ilustrado de una mejor manera con ayuda de la Figura 1.1. La serie temporal $x(t)$ (curva sinusoidal continua en la Figura 1.1)

1.3 Problemática en el análisis de datos

ha sido muestreada con un intervalo temporal Δt . Los puntos discretos obtenidos son compatibles con la curva sinusoidal discontinua y aliaseda en la Figura 1.1, de tal forma que se muestrean puntos en común (x_1, \dots, x_6) a las dos curvas. Como se observa, no es posible determinar sin ambigüedad a qué función sinusoidal pertenecen estos puntos. El problema con esto es que, al trabajar con fenómenos ambientales o ecosistémicos, no siempre es posible controlar el Δt de las muestras (datos) y, debido a la existencia de eventos periódicos de muy alta frecuencia (como el ciclo diario), cabe la posibilidad de estar trabajando con series temporales que contengan alguna señal aliaseda de alta frecuencia. Las señales de alta frecuencia muestreadas con valores demasiado largos de Δt aparecen en el espectro como correspondientes a frecuencias más bajas, aunque esa varianza corresponde en origen a señales de alta frecuencia.

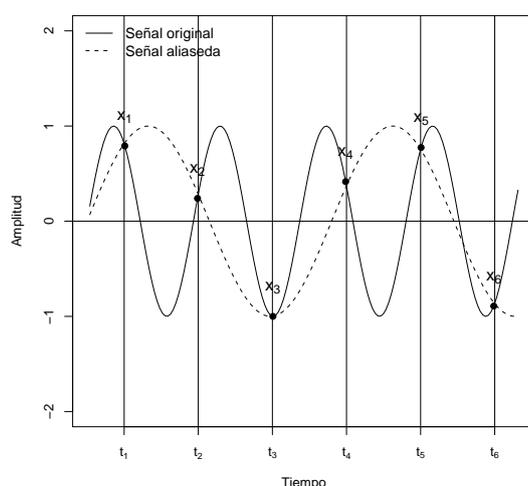


Figura 1.1: Representación esquemática del aliasing utilizando dos funciones sinusoidales con diferentes periodos e iguales amplitudes. Basado en Bendat y Pierson [20].

No equiespaciadas temporalmente. Otro de los inconvenientes que se encuentra frecuentemente al analizar series temporales ambientales y ecosistémicas, es que los tiempos de muestreo no siempre son equidistantes, *i. e.*, las series suelen ser no equiespaciadas temporalmente [225, 349, 371]. Existen dos tipos de series no equiespaciadas temporalmente. Las series irregulares que aparecen debido a la

1.3 Problemática en el análisis de datos

pérdida de datos, *v. gr.*, a fallos en los instrumentos u otros impedimentos en la obtención de los datos y las series no equiespaciadas temporalmente debido a la presencia de una componente estocástica en la localización espacio-temporal de los datos (muestras), *v. gr.*, las series temporales paleoclimáticas que provienen de microfósiles o de experimentos no controlados respecto al tiempo [232, 315, 371].

La forma más común utilizada para resolver este inconveniente es la interpolación en el tiempo de la serie temporal bajo análisis. Sin embargo, esto es una *mala praxis*, porque la interpolación es una herramienta muy dependiente de la estructura de la serie temporal y del método de interpolación empleado [387] y sobre todo, porque estos métodos pueden introducir información inexistente en la serie original en la nueva serie temporal equiespaciada. Como consecuencia pueden sesgar los resultados del análisis estadístico de los datos interpolados ya que no son independientes [225, 317, 371, 378]. Por ello, es recomendable utilizar técnicas estadísticas que puedan trabajar directamente con las series temporales no equiespaciadas sin recurrir a la interpolación.

Autocorrelación, persistencia o memoria. Las variables ambientales frecuentemente exhiben una dependencia estadística (están autocorrelacionadas) con sus valores pasados o con sus valores futuros. Esta dependencia es conocida como persistencia o memoria y se clasifica, desde un punto de vista estadístico, en dos tipos: de corta y de larga duración [365, 376]. La primera tiene una función de autocorrelación que decae exponencialmente, mientras que la segunda tiene una función que decae lentamente de acuerdo a una ley de potencias [304]. Ambos tipos de memoria han sido encontradas en los análisis estadísticos de registros ambientales, de reconstrucciones paleoambientales y de simulaciones numéricas [30, 304, 365, 376]. El problema con las series temporales que presentan autocorrelación es que no cumplen la condición de independencia en las observaciones, un requisito necesario de las técnicas de análisis estadístico clásico (un conjunto de herramientas frecuentemente utilizadas para el análisis de datos ambientales y ecosistémicos) [182, 365].

Tendencia. Las series temporales ambientales y ecosistémicas, por lo general, exhiben un tipo de tendencia, *i. e.*, sus propiedades estadísticas, como la media, crecen o disminuyen a lo largo del tiempo [365, 371]. Debido a que el origen de la tendencia puede tener varias causas, *v. gr.*, efectos antropogénicos (como

los efectos de la urbanización), cambios en los forzamientos externos del sistema climático (como la luminosidad solar o concentraciones de GHG) o la variabilidad natural de baja frecuencia del sistema climático, la presencia de tendencia en una serie temporal no implica necesariamente un cambio en las características físicas del fenómeno que representan [365, 385].

La determinación de la tendencia en una serie temporal es una de las tareas más complicadas en el análisis estadístico de datos, sobre todo cuando las series son cortas y ruidosas [365, 383, 385]. Sin embargo, es recomendable remover la tendencia antes de realizar un análisis estadístico y trabajar con los residuos [365, 371]. También es recomendable no utilizar la inspección visual para determinar si existe o no tendencia en una serie y para averiguar el tipo de tendencia, porque se puede llegar a conclusiones erróneas [365, 385]. Aunque existen diferentes tipos de tendencias, es muy común remover una tendencia de tipo lineal al trabajar con datos ambientales y ecosistémicos [371, 383].

La presencia de *outliers*. Las series temporales ambientales y ecosistémicas pueden presentar valores atípicos o extremos, debido a las características propias de los datos (*v. gr.*, errores inherentes en los datos, como los de medición) o por causas naturales [349, 365]. Por lo general, los valores extremos suelen ser removidos de las series a estudio porque se consideran como errores en los datos o porque las técnicas estadísticas clásicas para el análisis de datos no contemplan la existencia de valores extremos [225, 365]. Sin embargo, estos valores pueden contener información sobre eventos extremos del fenómeno a estudio [225]. Para el tratamiento de eventos extremos existen técnicas estadísticas adecuadas para ello. Una descripción de estas técnicas puede encontrarse en [167, 225].

Incertidumbres en la escala temporal. La escala temporal de las series paleoclimáticas contiene incertidumbres, ya sea debido a los errores de datación (*v. gr.*, de las muestras) o al construir los modelos de edad-profundidad (sedimentológica) [226, 348]. La incertidumbre en los modelos edad-profundidad se convierte en la principal fuente de error al utilizar técnicas estadísticas en el análisis de series temporales paleoclimáticas, de modo especial cuando se utilizan técnicas de análisis espectral [226, 378].

1.4. Motivación y antecedentes de esta memoria

En esta memoria se presentan cuatro casos de estudio que tratan sobre tres temas relacionados con los fenómenos ambientales y ecosistémicos. El primer tema está relacionado con la contaminación atmosférica y la calidad del aire, el segundo con el clima del Jurásico tardío, los paleoindicadores biogeoquímicos y los ciclos orbitales de la Tierra y el tercero con las fluctuaciones de especies marinas y su relación con el clima. A continuación se presentan las principales motivaciones y antecedentes de los tres temas tratados en los casos de estudio de esta memoria.

La contaminación atmosférica y la calidad del aire.

El término contaminación atmosférica se refiere a la presencia en el aire de materias o formas de energía que impliquen riesgo, daño o molestias graves para los humanos y para los bienes materiales [10]. Es importante tener en cuenta que la contaminación atmosférica de origen natural siempre ha existido debido a procesos biológicos, geológicos, químicos y físicos que generan partículas o gases contaminantes, como las erupciones volcánicas, incendios forestales, tormentas de arena, fermentaciones biológicas, etc. Con el descubrimiento del fuego por el hombre se origina la contaminación atmosférica antropogénica. Este tipo de contaminación ha ido adquiriendo importancia desde la revolución industrial y por el uso masivo de combustibles fósiles como fuentes de energía [10, 369].

El campo de estudio de la contaminación atmosférica es muy amplio, incluye desde los estudios de los gases de efecto invernadero y su relación con el sistema climático, la destrucción de la capa de ozono debido a los clorofluorocarbonos, el impacto de liberaciones accidentales de contaminantes químicos, biológicos o radionucleidos a la atmósfera, hasta estudios de la calidad del aire [323, 340]. Sin embargo, en esta sección se hace énfasis en lo que respecta a la calidad del aire porque tiene relación con uno de los casos de estudio presentados en esta memoria (capítulo 3).

Uno de los primeros estudios descriptivos relacionados con la calidad del aire data de mediados del siglo XVII, el *Fumifugium*, publicado por Johan Evelyn en 1648. El *Fumifugium* trata sobre el impacto del uso del carbón como un combustible en el medio ambiente de Londres y algunas medidas para combatir este

1.4 Motivación y antecedentes de esta memoria

tipo de contaminación. Posteriormente, en 1692, Robert Boyle realizó estudios pioneros sobre la composición química atmosférica. Con la llegada de la revolución industrial, el número de estudios relacionados con la contaminación del aire fue en aumento. En este periodo destacan los trabajos de Robert A. Smith en la segunda mitad del siglo XIX sobre la lluvia ácida y porque organizó una red de seguimiento de contaminantes atmosféricos, considerada como la precursora de las actuales redes de control y calidad del aire [323, 340].

Con el incremento masivo del tráfico rodado y el rápido crecimiento industrial y poblacional en diferentes lugares alrededor del mundo entre principios y mediados del siglo XX, las emisiones antropogénicas de contaminantes a la atmósfera se dispararon considerablemente y con ello sus consecuencias [209, 323]. De modo especial hay que considerar el impacto en la salud humana, como es el caso de los incidentes mortales causados por episodios agudos de contaminación atmosférica, tales como la Gran niebla de Londres en 1952 [323, 340].

Durante la década de los 60 y 70, los problemas de contaminación atmosférica fueron ocasionados principalmente por emisiones de SO_2 y por partículas suspendidas en el aire, emitidas principalmente por fuentes industriales y, en menor medida, por núcleos urbanos [209]. Estos problemas de contaminación de origen antropogénico llamaron la atención tanto de la comunidad científica como de la sociedad en general, lo cual condujo a la regulación mediante políticas medioambientales [128, 209].

Como consecuencia de la regulación legal para el control de contaminantes atmosféricos, los niveles de emisiones de SO_2 durante los últimos 25 años han ido disminuyendo en la mayoría de los países occidentales industrializados [237]. Por ejemplo, para la región geográfica a la que pertenece el caso de estudio del capítulo 3, para el periodo de 1997 a 2007 y en comparación al nivel de 1990, hay una clara tendencia a la disminución [239]. Sin embargo, a pesar de esta tendencia a la baja, aún hay problemas relacionados con emisiones locales en combinación con factores meteorológicos adversos y topográficos. En contraste, las emisiones de SO_2 están incrementando en países con economías emergentes [237].

Otra de las consecuencias de las políticas medioambientales es que motivaron el desarrollo de métodos de captación (muestreo) y de análisis para medir tanto las emisiones como las inmisiones [128, 209]. Con lo que respecta a los métodos

1.4 Motivación y antecedentes de esta memoria

para captar y analizar la evolución de las inmisiones de diversos tipos de contaminantes y poder llevar a cabo un control y vigilancia de la calidad del aire, hoy en día se utilizan principalmente los analizadores automáticos [209, 381]. Estos analizadores tienen varias ventajas con respecto a otros métodos de seguimiento de contaminantes. Por ejemplo, tienen una muy alta resolución temporal (se pueden obtener datos cada hora o a menor resolución), pueden ser instalados en cualquier lugar adecuado para ello, etc. Sin embargo, su coste no es bajo, tienden a ser más susceptibles a problemas técnicos cuando no se cuenta con el mantenimiento adecuado y con personal técnico cualificado y requieren de una constante evaluación [209].

Al conjunto de analizadores automáticos (nodos) que miden inmisiones de contaminantes y que forman una red de muestreo, se les conoce como una red de control y vigilancia de la calidad del aire (AQMN, *Air Quality Monitoring Network*) [209]. Estas redes permiten hacer un estudio y seguimiento adecuado de la calidad del aire. Sin embargo, requieren de una constante evaluación para averiguar y garantizar que cada uno de sus nodos proporcionen una caracterización adecuada de la calidad del aire en la zona donde muestrea cada sensor [272]. Las constantes evaluaciones ayudan a determinar el número adecuado de nodos de la red de tal modo que no se obtenga información redundante, a detectar fallos en alguno de los nodos o a detectar una inadecuada localización espacial de los nodos [176, 271].

Una de las herramientas útiles para evaluar las AQMN son las técnicas estadísticas multivariantes, como el análisis por componentes principales (PCA, *Principal Component Analysis*) o técnicas matemáticas de clasificación como el análisis de cluster (CA, *Cluster Analysis*). Estas técnicas han sido utilizadas previamente para la evaluación y el manejo adecuado de redes de seguimiento de la calidad del agua [329, 332, 384]. El uso combinado de ambas técnicas en la evaluación de una AQMN ha sido recientemente llevado a cabo en el año 2006 [111] para determinar la tendencia estacional y la distribución espacial de PM_{10} y O_3 . Posteriormente, en 2008, [272] aplicaron el PCA y el CA para valores de SO_2 y PM_{10} para una AQMN que cubre el área metropolitana de Oporto, encontrando que de los 10 nodos de la red 6 son suficientes para caracterizar las concentraciones de SO_2 .

1.4 Motivación y antecedentes de esta memoria

En lo que respecta a uno de los casos de estudio presentado en esta memoria (capítulo 3), la primera aplicación de técnicas estadísticas multivariantes para evaluar una AQMN para el área metropolitana de Bilbao, fue realizada por Ibarra-Berastegi *et al.* [149] en 2009. Ellos también aplicaron mapas auto-organizativos de Kohonen (SOM) y análisis de cluster y, encontraron que todos los nodos de la red son necesarios para una correcta caracterización regional de las concentraciones mensuales de SO₂.

El clima del Jurásico tardío, los paleoindicadores biogeoquímicos y los ciclos orbitales de la Tierra

La configuración geográfica (paleogeografía) del Jurásico tardío (hace unos 200 millones de años) (Figura 1.2) fue muy diferente en comparación a la actual geografía de la Tierra. Se sabe que existió una gran masa de tierra conocida como Pangea (subdividida en Laurasia, al Norte y en Gondwana, al Sur), la cual se extendía desde los 80°N hasta los 80°S y centrada sobre el Ecuador. Pangea estaba rodeada por un océano enorme, conocido como Pantalasa cuyas aguas envolvían prácticamente a todo el mundo. También existió un océano interior tropical de nombre Tetis y que estuvo localizado en la parte oriental de Pangea [11, 108, 320, 321, 390]. Este océano epicontinental y poco profundo en su parte occidental contenía numerosos islotes [11, 108, 390].

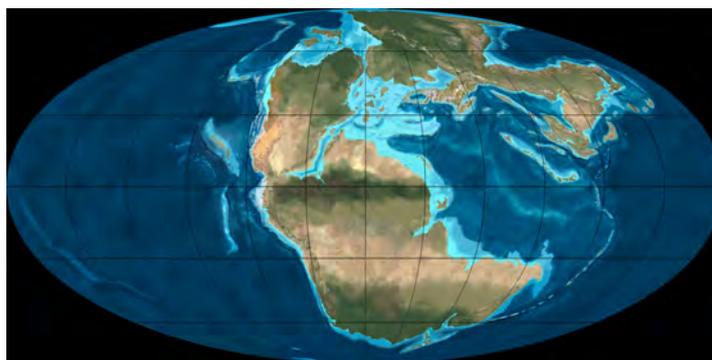


Figura 1.2: Paleogeografía global del Jurásico tardío de acuerdo con Ron Blakey. Fuente: <http://jan.ucc.nau.edu/~rcb7/>, último acceso julio de 2011.

1.4 Motivación y antecedentes de esta memoria

De igual manera que la paleogeografía del Jurásico tardío fue muy diferente a la actual geografía de la Tierra, el clima del Jurásico tardío también fue muy diferente al de hoy en día. En base a diversos estudios paleoclimáticos, se sabe que la temperatura promedio del aire en superficie estuvo entre 5 y 10°C por encima de la temperatura promedio actual de la Tierra (*i. e.*, existió un calentamiento global en comparación al clima actual de la Tierra) y existieron grandes áreas áridas en bajas y medias latitudes en el interior de la parte occidental de Pangea [12, 52]. Prevalció un clima de tipo monzónico que dominaba a lo largo de las medias latitudes, por lo cual, existieron grandes contrastes de temperaturas entre diferentes regiones geográficas y estaciones del año (primavera y verano) y probablemente altas tasas de precipitación (lluvias) que afectaron principalmente a las latitudes medias de Pangea [11, 12, 52].

El Jurásico tardío ha llamado mucho la atención a la comunidad científica debido, en parte, a que durante ese intervalo temporal tuvo lugar la separación de Pangea [12]. De igual modo han llamado la atención dos periodos temporales del Jurásico tardío, el Plienbachiense y el Toarciense (datados entre 182 a 190 millones de años aproximadamente) [12, 346]. Este interés se debe a que tuvieron lugar grandes crisis y cambios ambientales, *v. gr.*, un aumento excepcional en la deposición de material con alto contenido de carbono orgánico a nivel global, eventos oceánicos anóxicos, un breve intervalo de calentamiento global, alta actividad volcánica, liberación de grandes cantidades de metano, un aumento del nivel eustático del mar y extinciones masivas de la biodiversidad [12, 19, 118, 250, 346].

Para el estudio del clima del Jurásico tardío se han utilizado principalmente dos aproximaciones. Las simulaciones numéricas, principalmente modelos acoplados de circulación general de Océano-Atmósfera (AOGCM) [11, 29, 52, 302] y el análisis estadístico de paleoindicadores biogeoquímicos [11, 12, 155]. En esta sección se hace énfasis en la segunda aproximación porque se ha utilizado en uno de los casos de estudio aquí reportados (capítulo 4).

Uno de los obstáculos al utilizar paleoindicadores en el estudio del clima de épocas geológicas del pasado lejano, es la disponibilidad de material geológico. Por ejemplo, para el Jurásico tardío, hoy en día existen pocas rocas de esa época y las existentes no siempre están libres de efectos diagénicos [12, 297]. En parte, por estas razones, se han utilizado los esqueletos fosilizados (macrofósiles) y bien

1.4 Motivación y antecedentes de esta memoria

preservados (sin efectos diagénicos), principalmente de moluscos, braquiópodos y belemnites. De todos ellos se extraen ciertos valores isotópicos (como el $\delta^{18}\text{O}$, $\delta^{13}\text{C}$, Sr/Ca, etc.) para, a partir de esa información, inferir variables paleoambientales, como la temperatura o la salinidad del océano del área geográfica a la que pertenecieron [11, 12, 213, 297].

En lo que respecta a los belemnites, han sido considerados como uno de los grupos de fósiles más adecuados para estimaciones de paleotemperaturas y paleosalinidades del océano de esa época [12, 155, 212, 274, 297]. También han sido utilizados en menor medida y con cierta incertidumbre, para reconstrucciones de la circulación oceánica, paleoproductividad y nivel del mar [109, 274, 298]. Sin embargo, los paleoindicadores, como el $\delta^{18}\text{O}$, $\delta^{13}\text{C}$ o Sr/Ca, extraídos de las rostra fosilizadas de belemnites sufren de un particular inconveniente, contienen ruido [213, 274]. Este ruido puede representar cambios en los procesos ambientales ocurridos en el pasado (frecuentemente de corto periodo) en especies de belemnites adaptables a estos cambios. Sin embargo, también pueden reflejar el rango vertical y lateral del hábitat de los belemnites, incluso pueden reflejar los efectos de las especies, del género o la edad biológica del ejemplar del belemnite del cual procede el macrofósil [213, 274].

Otro de los inconvenientes en el uso de los belemnites en estudios paleoclimáticos es que la localización espacio-temporal de las muestras presentan un cierto grado de aleatoriedad (esto no es exclusivo de los belemnites, es una particularidad de los macrofósiles) [213, 371]. Por esta razón, las series temporales de los paleoindicadores, como los valores medios de los isótopos estables $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$, no suelen estar equiespaciadas temporalmente [371]. Debido a que el número de muestras en una sección estratigráfica no siempre es grande, las series temporales de los paleoindicadores son cortas, y también debido a las incertidumbres en las dataciones de las muestras, el tipo de estudios paleoambientales ha estado limitado más al análisis cualitativo que al cuantitativo.

Quizá, por estas razones, hasta la fecha, el uso de las series temporales de paleoindicadores obtenidos de muestras fosilizadas de belemnites del Jurásico tardío, no han sido utilizados para averiguar si estos isótopos estables pudieron haber registrado algún tipo de evento periódico relacionado con los ciclos orbitales de la Tierra [23, 175]. De manera especial, en la búsqueda del ciclo de excentricidad de

1.4 Motivación y antecedentes de esta memoria

largo periodo de los 400 Ka [23, 63], porque se sabe que ha permanecido invariable desde el Mesozoico hasta el presente [136, 175] y porque ha sido identificado en registros estratigráficos del Mesozoico recolectados en diferentes áreas geográficas [90, 245, 247, 372].

No obstante, a pesar de las limitaciones anteriormente expuestas, en el capítulo 4 de esta memoria se presenta un estudio donde se aplica una técnica robusta de análisis espectral que puede manejar de modo directo (sin previa interpolación en el tiempo) series temporales no equiespaciadas en el tiempo, contaminadas por ruido y teniendo en cuenta los errores de datación de las muestras. Esta aplicación tiene como objetivo identificar la presencia de ciclos orbitales que se hayan podido registrar en series temporales de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ provenientes de muestras de belemnites recolectadas en la cuenca Vasco-Cantábrica [296, 297, 299]. La identificación de ciclos orbitales podría ser útil porque implicaría la presencia de forzamientos externos en el clima de esa época, aclararía ciertos aspectos paleoecológicos de esta especie, incluso podría ayudar a verificar los modelos astronómicos de los parámetros orbitales de la Tierra correspondientes al Jurásico Tardío, tales como el modelo de Laskar *et al.* [175].

La relación entre las fluctuaciones de especies marinas y el clima.

Los stocks de especies marinas (como los peces o los moluscos) son altamente variables a diferentes escalas espacio-temporales [138, 300]. Desde hace más de un siglo, la comunidad científica ha intentado explicar y entender los mecanismos subyacentes que causan estas variaciones [300]. A día de hoy existen dos puntos de vista para explicarlas. El primero considera que las causas principales de las fluctuaciones son debidas a cambios en los factores endógenos de su propia dinámica poblacional, *v. gr.*, cambios en las tasas de crecimiento o reproducción, la mortalidad de origen natural, etc. El segundo considera como las causas principales a los factores exógenos, como la mortalidad debida a la pesca realizada por el hombre, la interacción entre las pesquerías, las interacciones biológicas o los forzamientos de tipo ambiental, (*v. gr.*, meteorológicos, oceanográficos o climáticos) [64, 73, 143].

Uno de los factores exógenos que más ha llamado la atención en las últimas décadas, es el impacto de la variabilidad climática sobre las fluctuaciones de es-

1.4 Motivación y antecedentes de esta memoria

pecies marinas, tales como el atún rojo o el pulpo común (véanse los capítulos 5 y 6, respectivamente). Se sabe que la variabilidad climática determina la distribución, migración y abundancia de las especies marinas y que también afecta a las capturas y a las pesquerías [79, 138, 184, 249]. Por ejemplo, a escalas temporales cortas (de días a semanas, incluso a escala estacional), se conoce relativamente bien el impacto de las variables meteorológicas sobre las pesquerías y sobre las especies marinas. Por un lado, el tiempo atmosférico condiciona las actividades pesqueras, lo cual puede verse reflejado en las capturas [184]. Por otro lado, los procesos ambientales (reflejados en las variables ambientales como la temperatura, la densidad del mar, el estrés del viento, etc.) afectan a las poblaciones principalmente mediante su habilidad para encontrar alimento [138].

A escalas temporales mayores, como la estacional, un incremento en el estrés del viento puede modificar la surgencia costera alterando la inyección de nutrientes a la zona fótica. Este incremento de nutrientes puede cambiar rápidamente la biomasa fitoplanctónica y, debido a la transferencia trófica, afectar al zooplankton, y a su vez a los peces que se alimentan de él, y así hasta alcanzar niveles superiores de la cadena trófica marina [79, 138].

A escalas temporales superiores a la estacional, los efectos climáticos sobre las especies marinas son cada vez más complejos debido a que aumentan los procesos ambientales involucrados [138, 184, 249]. En la escala interanual, los principales procesos ambientales están relacionados con el océano y con la interacción entre el océano y la atmósfera. En lo que respecta al Hemisferio Norte, a medias y altas latitudes, una de las principales y recurrentes características de la variabilidad climática es la Oscilación del Atlántico Norte (NAO, *North Atlantic Oscillation*) (Figura 1.3). La NAO produce grandes cambios en la velocidad y dirección media del viento sobre el Atlántico, en el transporte de humedad y calor entre el Atlántico y los continentes cercanos, y en la intensidad y el número de tormentas y sus trayectorias [145, 342].

Los patrones de circulación atmosférica asociados con la NAO pueden afectar directamente e indirectamente a las pesquerías y a las capturas de diversas especies marinas a escala interanual [184]. Es decir, por un lado, las condiciones meteorológicas, cuando la pesca en pequeña escala podría llevarse a cabo de manera más eficiente son coincidentes con un rango de valores en el índice NAO que

1.4 Motivación y antecedentes de esta memoria

describen los rangos característicos del estrés del viento y de las alturas de las olas de superficie [275]. Por otro lado, debido a que las especies marinas son dependientes de la temperatura del océano [79, 200], y ésta depende de varios factores aparte de la NAO pero relacionados con ésta (estrés del viento, profundidad de la capa de mezcla, cobertura nubosa, etc.) [33, 236, 292, 293, 355, 360].

El estudio del impacto de la variabilidad climática en las fluctuaciones de los stocks de especies marinas se realiza principalmente por medio de dos aproximaciones. La primera utiliza modelos de dinámica poblacional y relación stock-reclutamiento, como los modelos propuestos por Beverton y Holt [27] y Ricker [287] y sus respectivas mejoras, hasta modelos más sofisticados como los modelos propuestos por Navrotsky [233] o el modelo no lineal propuesto por Solari *et al.* [335, 336]. Sin embargo, una de las principales críticas de los modelos clásicos de poblaciones, y como bien apunta Solari [335], “es la capacidad limitada para incluir factores fundamentales de situaciones específicas y unir las dinámicas internas con los factores externos (como los ambientales o la pesca)”. El inconveniente de los modelos más sofisticados, es que no siempre se tiene toda la información de los parámetros intrínsecos (tasa de nacimiento o mortalidad natural, etc.) de las poblaciones de las especies marinas a estudio [22, 300].

La segunda aproximación está relacionada con el análisis estadístico de series temporales (la aproximación utilizada en los casos de estudio presentados en los capítulos 5 y 6). Esto es, el estudio de la relación entre variables ambien-

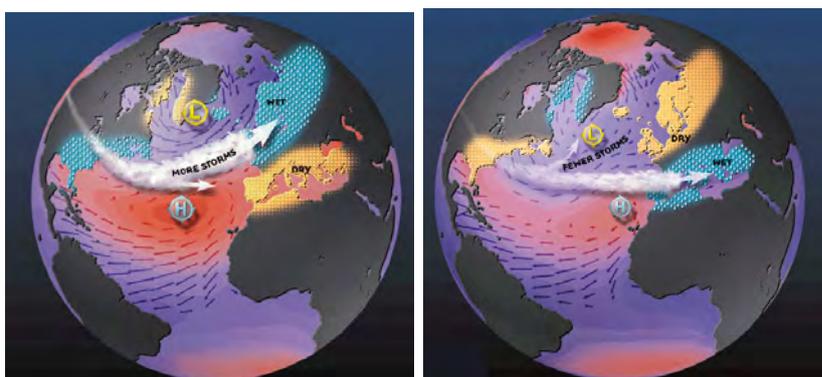


Figura 1.3: La Oscilación del Atlántico Norte (NAO) en su fase positiva (altas) y negativa (bajas). Tomado de <http://www.ldeo.columbia.edu/res/pi/NAO/>

1.4 Motivación y antecedentes de esta memoria

tales y las poblaciones de especies marinas, medidas como capturas o capturas por unidad de esfuerzo (CPUE, *Catch Per Unit Effort*) [50, 115, 144, 300]. Sin embargo, esta aproximación no está libre de inconvenientes. Tales como que la existencia de co-ocurrencia no implica necesariamente causalidad, la presencia de correlaciones físico-biológicas entre series temporales relativamente cortas y temporalmente autocorrelacionadas y la incertidumbre presente en los datos, debido a que representan una discretización muestral espacio-temporal de un sistema continuo. Todos ellos pueden llevar a errores de interpretación de los resultados [225, 249, 345, 392].

El caso de estudio presentado en el capítulo 5, está enfocado en averiguar la relación entre las fluctuaciones de capturas de atún rojo en el Sur de España, obtenidas por medio de almadrabas entre los siglos XVI y XVIII, con algunas variables climáticas. Uno de los antecedentes a este trabajo fue realizado por Ganzedo en 2005 [103]. Ahí se presentó una revisión minuciosa de todos los posibles factores ambientales, políticos y sociales que pudieron haber influenciado las capturas en las almadrabas del atún rojo. Se concluyó que las condiciones climáticas de esa época, y de modo especial las bajas temperaturas durante el mínimo de Maunder (evento climático conocido como la Pequeña Edad de Hielo, entre 1640-1715), pudieron haber reducido de modo significativo el reclutamiento y la abundancia del atún rojo en el Atlántico Norte y en el Mediterráneo [103].

Un poco más tarde, Solari [335] analizó, entre otros casos de estudio, las series de capturas de atún rojo (las mismas analizadas por Ganzedo [103]) utilizando un nuevo método desarrollado por ellos [336] de dinámica poblacional y stock-reclutamiento de tipo no lineal, de tal modo que las capturas de atún y la temperatura anual del Hemisferio Norte fueron consideradas como testigos de abundancia y de cambio climático, respectivamente [335]. Solari [335] sugirió que esta pesquería pudo estar controlada tanto por la tendencia de temperaturas máximas como mínimas, de tal modo que durante el mínimo de Maunder dio origen a un colapso de esta pesquería [335].

Sin embargo, a pesar de la importancia de estos estudios [103, 335], no es posible saber qué porcentaje de variabilidad de las capturas es debido a los factores climáticos, poblacionales y al efecto combinado de todos estos factores. Una forma

1.4 Motivación y antecedentes de esta memoria

de cuantificar los efectos de estos factores es por medio de un análisis de devianza, utilizando modelos lineales generalizados (GLM, *Generalized Linear Models*) ajustando las capturas contra las variables ambientales [53, 115]. Utilizando esta metodología y los mismos datos de las capturas de atún rojo que Ganzedo [103], el caso de estudio presentado en el capítulo 5 [102] de esta memoria determina los porcentajes de varianza debido a los diversos factores involucrados en las capturas.

El otro caso de estudio que relaciona fluctuaciones de especies marinas y clima se presenta en el capítulo 6 de esta memoria. En ese estudio se averigua la relación entre las fluctuaciones de CPUE de pulpo común de las Islas Canarias y su relación con las variables climáticas (temperatura superficial del mar y el índice de la Oscilación del Atlántico Norte) para el periodo 1989-2007. Estudios previos han demostrado que el pulpo común en el inicio de su ciclo de vida (en estado plantónico) es altamente dependiente de la temperatura ambiente [132, 200] y que, durante su posterior estadio (béntico), las oscilaciones estacionales de la temperatura del mar lo afectan considerablemente [159]. Sin embargo, aunque es un hecho aceptado que la temperatura del mar es uno de los factores que afectan la abundancia del pulpo común, no siempre está claro qué tipo de correlación existe entre esta variable y las CPUE.

En un estudio reciente realizado por Solari [335] en la zona de afloramiento de la región Noroccidental de África, encontró que las fluctuaciones en las capturas del pulpo común (datos de FAO, área 34) para los últimos 50 años están relacionadas con la NAO. En esta misma dirección, Sims *et al.* [331] en un estudio realizado en el canal de la Mancha, aunque para otra especie de cefalópodo, encontraron que la variación temporal de los picos de máxima abundancia de este cefalópodo es dependiente de la temperatura del mar, la cual a su vez está relacionada con la NAO.

Uno de los primeros trabajos realizados para estudiar la relación de las fluctuaciones de las capturas del pulpo común con variables climáticas, en las pesquerías en las Islas Canarias fue realizado por Hernández-García *et al.* [130]. Un poco más recientemente, Hernández-García *et al.* [131] llevaron a cabo un estudio sobre el pulpo común de las costas de las Islas Canarias, aunque enfocado a la reproducción de esta especie. Ellos encontraron algunas relaciones entre la temperatura

1.4 Motivación y antecedentes de esta memoria

superficial del mar y la CPUE y reportaron la existencia de dos cohortes anuales. La primera cohorte nace en abril y la segunda entre octubre y noviembre, aunque el instante preciso en que ocurren estos picos fluctúa de año en año.

Como una ampliación del estudio de Hernández-García *et al.* [130] y de Hernández-García *et al.* [131], Caballero-Alfonso *et al.* [45] estudiaron el papel de las variables climáticas (SST y NAO) en las fluctuaciones estacionales observadas en las capturas de pulpo común durante 1989-2007 en las Islas Canarias (parte de esos resultados se presentan en el capítulo 6 de esta memoria). Sin lugar a dudas, el trabajo de Caballero-Alfonso *et al.* [45] ayudó a entender un poco más la relación entre la variabilidad climática y las fluctuaciones de las capturas del pulpo común. Sin embargo, el análisis de ese estudio se limita a la escala estacional y asume un comportamiento estacionario en los datos analizados.

Una aproximación simple para estudiar los fenómenos ambientales y ecosistémicos es considerarlos lineales y estacionarios, aunque la no linealidad y la no estacionariedad son más bien la regla que la excepción [142, 300]. Además, el principal requisito de muchas técnicas estadísticas para el análisis de series temporales asumen que las series bajo estudio son estacionarias [279, 365, 371]. Sin embargo, las series temporales ambientales y ecosistémicas no siempre satisfacen este requerimiento, y la creciente evidencia suporta la necesidad de considerar la importancia de la dinámica de los sistemas que contienen componentes de corta duración en el estudio de procesos/fenómenos ambientales y ecosistémicos [50, 142, 300]. Una técnica estadística que enfrenta el problema de la no estacionariedad es el análisis espectral mediante la transformada continua de Wavelet [113, 206, 207, 350].

Como una extensión del trabajo de Caballero-Alfonso *et al.* [45], en el capítulo 6 de esta memoria también se estudia la relación entre las variables climáticas (SST y NAO) y las fluctuaciones de las capturas (CPUE) del pulpo común en las Islas Canarias para el periodo 1989-2007, considerando diversas escalas temporales y considerando la posible no estacionariedad de esta variable [275]. Para tal fin, se analizan las relaciones (en términos de coherencia espectral de wavelet) entre las variables climáticas (SST y NAO) y las CPUE.

1.5. Técnicas estadísticas utilizadas

En esta memoria se hace uso de varias técnicas estadísticas para el análisis de diferentes series temporales ambientales y ecosistémicas. Las técnicas utilizadas van desde las más simples, como el análisis de componentes principales, hasta técnicas más avanzadas, como el análisis de Wavelet. Las primeras, consideran *de facto* características estadísticas como la linealidad, la normalidad o la estacionariedad, se utilizan desde hace varias décadas y, a pesar de su simplicidad, son de gran ayuda en el estudio de diversos fenómenos ambientales y ecosistémicos [365, 376]. Con respecto a las segundas, tienen en cuenta propiedades estadísticas como la no-estacionariedad o la no-linealidad, y aunque han demostrado su utilidad en diversos estudios ambientales, no son de uso frecuente, aunque empiezan a ganar popularidad [76, 392]. A continuación se hace una descripción a grandes rasgos de las técnicas utilizadas en los casos de estudio presentados en esta memoria (para detalles técnicos consultar el capítulo 2).

1. Análisis de Componentes Principales (PCA). Una de las técnicas estadísticas de más amplio uso en estudios ambientales es el análisis de componentes principales (conocida también como Funciones Empíricas Ortogonales). Esta técnica consiste en la descomposición lineal de un conjunto de variables correlacionadas en términos de funciones de base ortogonal, de tal modo que reducen el número de variables y eliminan la correlación entre ellas [365, 376]. El PCA es utilizado en una amplia gama de aplicaciones en el estudio de fenómenos ambientales, desde el análisis de campos meteorológicos y climáticos (tales como temperaturas, alturas geopotenciales o precipitación) [365, 376], la identificación de potenciales modos físicos asociados con la variabilidad climática¹ [392], hasta la evaluación de redes de control y vigilancia de la calidad del aire [111, 272].

¹Es importante tener siempre presente al utilizar el PCA para la identificación de modos físicos de campos meteorológicos o climáticos, que las estructuras espaciales ortogonales obtenidas del PCA no están construidas *a priori* para extraer información acerca de la dinámica del sistema físico a estudio, sino que simplemente ofrecen una forma eficiente de representación de la variabilidad [392].

2. Análisis espectral vía el periodograma de Lomb-Scargle (PLS).

El análisis espectral es una de las herramientas estadísticas más utilizadas en el análisis de series temporales ambientales y ecosistémicas, ya sea para averiguar la posible existencia de eventos periódicos o para estudiar el comportamiento de la serie en el dominio de las frecuencias [105, 387]. Existe una gran variedad de métodos de análisis espectral, pero la mayoría requieren que la serie esté equiespaciada temporalmente [105, 371, 387], y este es un requisito que no siempre se cumple con las series temporales ambientales o ecosistémicas. La forma clásica para enfrentar el hecho de que la serie no es equiespaciada, es interpolar la serie original y convertirla a una serie temporalmente equiespaciada. Sin embargo, la interpolación altera la estimación del espectro, de tal modo que tiende a sobrestimar en las bajas frecuencias y a subestimar en las altas frecuencias [316, 317].

Para evitar la interpolación se han desarrollado diversos métodos que pueden trabajar directamente con las series no equiespaciadas temporalmente, como el periodograma de Lomb-Scargle (también conocido como *Lomb-Scargle Fourier Transform*). Aunque derivado originalmente para trabajar con series temporales astronómicas [15, 193, 311, 312], a finales de los noventa fue adaptado por Schulz y Stattegger [317] en combinación con la técnica WOSA (*Welch-Overlapped-Segment-Averaging* [374]) para trabajar con series temporales (paleo)climáticas. Un poco más recientemente Schulz y Mudelsee [316] hicieron mejoras al trabajo de Schulz y Stattegger para tener en cuenta el tipo de ruido de fondo que se presenta en las series ambientales (de modo particular, en series paleo/climáticas) [107, 124].

3. Análisis de devianza mediante Modelos Lineales Generalizados (GLM).

Borcard *et al.* [36], Borcard y Legendre [35] y Anderson y Gribble [8] realizaron una partición de la variación total en datos ecológicos en diferentes porciones relacionadas a descriptores (componentes) espaciales, temporales y ambientales, demostrando la importancia de considerar el posible solapamiento entre las diferentes componentes. A diferencia de estos autores que utilizaron el análisis de correspondencia canónica para particionar la variación presente en los datos. Ganzedo *et al.* [102] propusieron utilizar modelos lineales generalizados para relacionar variables climáticas con las capturas de atún rojo y realizar una

1.6 Programas computacionales utilizados

partición de la variación, considerando sólo las componentes ambiental, temporal y la combinación de estas componentes.

4. Análisis espectral cruzado vía la transformada Continua de Wavelet (CWT). El análisis espectral, por medio de la CWT (*Continuous Wavelet Transform*), es una herramienta poderosa utilizada en diferentes áreas de la ciencia e ingeniería. El atractivo y la versatilidad de esta herramienta recae en sus propiedades de descomposición tiempo-escala y en su capacidad para trabajar con series temporales no estacionarias [69, 114, 263]. Aunque el análisis de Wavelet es mucho más que la estimación del espectro (ya sea uni o bi-variante), muchas de las aplicaciones en estudios ambientales y ecosistémicos han sido utilizadas para este fin [172, 219, 350].

La utilización de la CWT en el estudio de fenómenos ambientales obtuvo popularidad debido al trabajo de Torrence y Compo [350]. Sin embargo, las primeras aplicaciones se remontan a estudios de Farge [88], Meyers *et al.* [219], Liu [192], Lau y Weng [177], entre otros. Algunas mejoras significativas (en los test de significación y en la estimación del espectro cruzado) han sido contribuidas por Maraun *et al.* [206, 207] y por Grinsted *et al.* [113]. Entre las aplicaciones más recientes de la CWT se encuentran los estudios sobre las variaciones de abundancia en las capturas de diferentes especies de peces (atún, anchoa, etc.) y sus relaciones con diferentes variables ambientales (temperatura, precipitación, etc.) e índices climáticos (NAO, ENSO, etc.) [144, 218, 300].

1.6. Programas computacionales utilizados

Para poder aplicar las técnicas estadísticas utilizadas en esta memoria se requiere de software potente y flexible. Existen varias opciones que cubren estas características, como Matlab o IDL, sin embargo, no siempre es posible utilizarlas debido al precio que cobran por el derecho de uso y por las restricciones de sus licencias. Una alternativa que supera estas restricciones la proporciona el software libre [28, 161]. En el caso específico para el análisis de datos, una de esas herramientas es R [280]. R está formada por un módulo básico con un conjunto de aplicaciones estadísticas de uso común, tiene la opción de añadir paquetes para aplicaciones específicas y es también un poderoso lenguaje de programación [280],

1.7 Objetivos y estructura de esta memoria

permite la posibilidad de interactuar con otros lenguajes de programación (como **C**, **Fortran**, **Python**, **Matlab**, etc.), tiene una base creciente de usuarios y se está convirtiendo en la *lingua franca* en el análisis de datos [148, 310].

Siguiendo con la tradición acumulada por el grupo de Clima, Meteorología y Medio Ambiente de la Universidad del País Vasco en el desarrollo de paquetes computacionales de libre distribución para el análisis de datos ambientales, *v. gr.*, **Pyclimate** [306], **BISTAPOF_DEMO** [147] o **rairmon** [149], se han adaptado y utilizado algunos paquetes computacionales, principalmente en lenguaje **R**, para llevar a cabo la aplicación de las técnicas estadísticas utilizadas en esta memoria. A continuación se presenta una breve descripción de cada uno de los programas computacionales utilizados en esta memoria.

Para realizar el PCA se ha utilizado el paquete **FactoMineR** [178] en combinación con el **Scatterplot3d** [188] (ambos en **R**). El programa computacional utilizado para estimar el espectro suavizado vía el PLS es **REDFIT ver. 3.5** [316]. Originalmente, **REDFIT** fue escrito (en Fortran) para trabajar sobre windows, pero el autor de esta memoria ha realizado una migración para que trabaje en sistemas GNU/Linux. Por otro lado, también se han escrito un conjunto de scripts en lenguaje **R** y en combinación con **REDFIT**, para considerar los errores de datación de las muestras de los belemnites. Para ajustar los GLM se ha utilizado la función **glm** del paquete **stats** de **R** [280, 358]. Para estimar el espectro y el espectro cruzado vía la CWT se ha utilizado el paquete informático **SOWAS** (en lenguaje **R**) ver. 0.93 para GNU/Linux [204, 205].

1.7. Objetivos y estructura de esta memoria

El objetivo general de esta memoria es estudiar algunos fenómenos ambientales y ecosistémicos por medio de la aplicación de sofisticadas técnicas estadísticas que permitan desentrañar algunas de las problemáticas presentes en los datos ambientales y ecosistémicos (como el ruido, el no equiespaciado temporal, la no estacionariedad, etc.). Esto, con el fin de identificar procesos físicos de interés. Los objetivos particulares son los siguientes:

1.7 Objetivos y estructura de esta memoria

1. Determinar si los sensores de una red de control y vigilancia de la calidad del aire miden (inmisión de SO_2) información redundante. El área de estudio se sitúa sobre Bilbao metropolitano y corresponde al periodo 1996-2001. La técnica estadística aplicada es el PCA, cuyos resultados se comparan con los correspondientes a mapas autoorganizativos de Kohonen y análisis de cluster [149].
2. La búsqueda de ciclos de largo periodo en valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ de belemnites y una posible relación con ciclos orbitales astronómicos. El área de estudio se sitúa en la cuenca Vasco-Cantábrica y corresponde al Jurásico tardío (hace unos 200 millones años). La técnica estadística aplicada es el análisis espectral mediante el PLS.
3. Cuantificar los efectos de las variables climáticas en las fluctuaciones de capturas de atún rojo. El área de estudio se sitúa en el sur de España y corresponde al periodo 1525-1756. La técnica estadística aplicada es el análisis de devianza mediante GLM [102].
4. Identificar la relación entre las capturas del pulpo común de las Islas Canarias y las variables climáticas (temperatura superficial del mar y el índice de la Oscilación del Atlántico Norte) para el periodo 1989-2007. La técnica estadística aplicada es el análisis espectral cruzado (coherencia espectral) mediante la CWT [275].

La estructura de esta memoria es la siguiente. En este capítulo se presenta una introducción general, los objetivos y la estructura de esta memoria. En el capítulo 2 se presentan los aspectos metodológicos de las técnicas estadísticas utilizadas. En los capítulos 3, 4, 5 y 6 se presentan los casos de estudio que llevan a cabo los objetivos particulares de esta memoria, respectivamente. En el capítulo 7 se presentan las conclusiones para los capítulos 3-6, una conclusión general de la memoria y oportunidades de trabajo futuro. En el anexo A se presenta una lista de las publicaciones y de los trabajos presentados en congresos por el autor de esta memoria.

Capítulo 2

Principales aspectos metodológicos

Este capítulo se presentan las principales características técnicas de las cuatro herramientas estadísticas utilizadas en esta memoria. Esto, con el fin de proporcionar al lector desde la parte conceptual hasta la parte aplicada, para tener presente cómo se utilizan de modo adecuado y cómo se evalúan los resultados en el análisis de datos ambientales y ecosistémicos.

2.1. Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA) es una de las técnicas estadísticas multivariantes más populares y antiguas en el análisis de datos. Esta técnica fue desarrollada por Karl Pearson en 1901¹ [255], pero no fue hasta 1939 cuando Hotelling hizo una presentación mucho más formal y acuñó el término de *componente principal* (PC) [2, 141]. El PCA también recibe otros nombres, dependiendo de su campo de aplicación, *v. gr.*, en la teoría de procesos estocásticos se conoce como la expansión o transformada de Karhunen-Loève [222], en turbulencia como descomposición ortogonal propia [24], en ciencias sociales y económicas como vectores principales [163], en ciencias atmosféricas como Funciones Empíricas Ortogonales [365, 376].

¹Aunque algunas investigaciones establecen que sus orígenes se remontan hasta Cauchy (en 1829) y Jordan (en 1874) [2].

2.1 Análisis de Componentes Principales

Antes de presentar una definición formal y describir sus principales características matemáticas, podemos mencionar que el PCA es un tipo de transformación lineal aplicada a un conjunto de datos multivariantes habitualmente correlacionados entre sí, para convertirlos en un menor número de variables no correlacionadas y ortogonales¹ entre sí, que representen de forma decreciente la varianza de los datos originales [156, 376]. Los principales objetivos del análisis de componentes principales de acuerdo con Abdi y Williams [2] y Jolliffe [156], es extraer la información más importante de un conjunto de datos multivariantes, comprimir un conjunto de datos multivariantes manteniendo sólo la información que se considere importante (reducir la dimensionalidad de los datos), simplificar la descripción de un conjunto de datos y analizar la estructura de las observaciones y de las variables.

Notación matemática y conceptos preliminares

Antes de exponer la parte metodológica del PCA es importante mantener una nomenclatura adecuada y consistente. La presentación metodológica está basada y es análoga a la de Abdi y Williams [2] y Hannachi *et al.* [120]. Para denotar matrices, vectores y elementos, se usarán mayúsculas en negritas, minúsculas en negritas y minúsculas en itálicas, respectivamente. Matrices, vectores y elementos de la misma matriz usarán la misma letra, *v. gr.*, \mathbf{X} , \mathbf{x} , x . La transpuesta de una matriz se representará con el superíndice T . La matriz identidad estará denotada por \mathbf{I} . El vector columna de unos y de longitud I viene dado por $\mathbf{1}_{(I \times 1)}$.

Los datos ambientales que van a ser analizados mediante PCA contienen I observaciones (muestras)² y J variables. El número de observaciones es mayor que el número de variables en el caso de estudio donde se usa el PCA (capítulo 3) en esta memoria, aunque esto no es necesario. Cada observación (muestra) es obtenida en los tiempos $t_i, i = 1, 2, \dots, I$ y están representadas por la matriz $\mathbf{D}_{(I \times J)}$ (filas por columnas), donde un ij -ésimo elemento viene dado por $d_{i,j}$. Antes de aplicar el PCA, los datos ambientales requieren ser preprocesados. El

¹En el seno del PCA, el concepto de ortogonalidad de series temporales (las PCs lo son) corresponden al concepto de series incorrelacionadas.

²Se utiliza el término observación sólo por cuestiones prácticas, los datos bajo análisis vía PCA no se limitan a datos observacionales.

2.1 Análisis de Componentes Principales

tipo de procesamiento frecuentemente utilizado consiste en sustraer la media de las observaciones de cada variable y trabajar con las anomalías $\mathbf{X}_{(I \times J)}$ [120, 376], esto es:

$$\mathbf{X}_{(I \times J)} = \mathbf{D}_{(I \times J)} - \mathbf{1}_{(I \times 1)} \bar{\mathbf{D}}_{(1 \times J)} \quad (2.1)$$

donde $\bar{\mathbf{D}}$ (vector de medias muestrales) viene dado por

$$\bar{\mathbf{D}}_{(1 \times J)} = (\bar{\mathbf{d}}_1, \bar{\mathbf{d}}_2, \dots, \bar{\mathbf{d}}_J) = \frac{1}{I} \mathbf{1}_{(1 \times I)}^T \mathbf{D}_{(I \times J)} \quad (2.2)$$

Si sustituimos el tercer miembro de la relación 2.2 en la relación 2.1 y factorizando, las anomalías también pueden representarse, de acuerdo con [120], del siguiente modo:

$$\mathbf{X}_{(I \times J)} = \left(\mathbf{I}_{(I \times I)} - \frac{1}{I} \mathbf{1}_{(I \times 1)} \mathbf{1}_{(1 \times I)}^T \right) \mathbf{D}_{(I \times J)} = \mathbf{M}_{(I \times I)} \mathbf{D}_{(I \times J)} \quad (2.3)$$

donde $\mathbf{M}_{(I \times I)}$ es la matriz de centrado de orden I [120]. De aquí en adelante no utilizaremos (a menos que sea necesario) los subíndices dimensionales I, J para simplificar la notación matemática.

2.1.1. Calculando las componentes principales

La forma más usual como se presenta el cálculo de las componentes principales (PC) en los textos de análisis de datos ambientales (específicamente datos climáticos o meteorológicos), es mediante la solución de un problema de autovalores a través de la matriz de covarianzas de las anomalías \mathbf{X} de los datos \mathbf{D} bajo estudio [120, 365, 376]. Este procedimiento puede expresarse matemáticamente y de acuerdo con Hannachi *et al.* [120] de la siguiente forma. La matriz de covarianzas muestral de la matriz de anomalías \mathbf{X} (relación 2.3) está definida [120] por la relación

$$\mathbf{S} = \frac{1}{I} \mathbf{X}^T \mathbf{X} \quad (2.4)$$

donde cada elemento de \mathbf{S} está formado por las covarianzas entre cada par de variables de \mathbf{X} de dimensiones $I \times J$.

2.1 Análisis de Componentes Principales

Ahora bien, el objetivo del análisis por componentes principales es encontrar un nuevo conjunto de variables (combinaciones lineales) no correlacionadas entre sí que expliquen la máxima varianza. Esto equivale a encontrar un vector unitario $\mathbf{q} = (q_1, \dots, q_J)^T$ tal que $\mathbf{X}\mathbf{q}$ tenga la máxima variabilidad [120, 365]. Esto es

$$\max\{\mathbf{q}^T \mathbf{S}\mathbf{q}\} \quad (2.5)$$

sujeta a la condición $\mathbf{q}^T \mathbf{q} = 1$.

Los autovectores o funciones empíricas ortogonales (EOFs) se obtienen, de acuerdo con Hannachi *et al.* [120], como una solución al problema de autovalores

$$\mathbf{S}\mathbf{q} = \lambda\mathbf{q} \quad (2.6)$$

donde los autovalores $\lambda_l, l = 1, 2, \dots, N$ con $N = \min(I, J)$, vienen dados por

$$\lambda_l = \mathbf{q}_l^T \mathbf{S}\mathbf{q}_l = \frac{1}{I} \|\mathbf{X}\mathbf{q}_l\|^2 \quad (2.7)$$

Los autovalores λ_l proporcionan una medida de la varianza de \mathbf{X} en la dirección de \mathbf{q}_l . Una vez que se resuelve el problema de autovalores (relación 2.6), por regla general, éstos son ordenados de modo decreciente [120], esto es, $\lambda_1 \geq \lambda_2 \dots \geq \lambda_N$. Una forma común de expresar el porcentaje de varianza correspondiente a cada autovalor es por medio de la relación

$$\frac{100\lambda_l}{\sum_{l=1}^N \lambda_l} \quad (2.8)$$

Las l -ésimas componentes principales están dadas por la proyección de \mathbf{X} sobre el l -ésimo autovector $\mathbf{q}_l = (q_{1l}, q_{2l}, \dots, q_{Jl})^T$ [120] y se expresan mediante la siguiente relación

$$\mathbf{p}_l = \mathbf{X}\mathbf{q}_l \quad (2.9)$$

Los elementos $(p_{tl}, t = 1, \dots, I)$ de la relación 2.9 pueden expresarse como

2.1 Análisis de Componentes Principales

$$p_{tl} = \sum_{j=1}^J x_{tj} q_{jl} \quad (2.10)$$

De tal modo que el l -ésimo autovalor λ_l representa la varianza de la l -ésima componente principal $\mathbf{p}_l = (p_{l1}, p_{l2}, \dots, p_{lI})^T$ [120].

Sin embargo, para calcular las componentes principales, en la práctica, no se calcula la matriz de covarianza muestral (relación 2.4) ni se resuelve el problema de autovalores (relación 2.7), sino que se calculan por medio de la descomposición de \mathbf{X} por valores singulares (SVD) [2, 120]. Este procedimiento se describe a continuación [2, 120]. La matriz $\mathbf{X}_{(I \times J)}$ se descompone por valores singulares de la siguiente manera:

$$\mathbf{X}_{(I \times J)} = \mathbf{P}_{(I \times r)} \mathbf{\Sigma}_{(r \times r)} \mathbf{Q}_{(r \times J)}^T \quad (2.11)$$

donde las matrices \mathbf{P} (conocida como la matriz de las componentes principales) y \mathbf{Q} (conocida también como matriz de cargas o de proyección) son tales que sus columnas $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_r$ y $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_r$, son ortogonales y se llaman vectores singulares izquierdo y derecho, respectivamente. El rango de \mathbf{X} es r y $\mathbf{\Sigma} = \text{Diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ es una matriz diagonal cuyos elementos $\sigma_1 \geq \sigma_2 \dots \geq \sigma_r \geq 0$ son los valores singulares de \mathbf{X} [2, 120]. Ahora bien, debido a la propiedad de ortogonalidad de los autovectores (constituyen una base), la relación 2.11 se puede expandir (descomponer) como una combinación lineal [120, 376], esto es

$$\mathbf{X} = \sum_{l=1}^r \sigma_l \mathbf{p}_l \mathbf{q}_l^T \quad (2.12)$$

También es posible expresar la matriz de covarianzas \mathbf{S} en términos de la SVD [2, 120]. Si se sustituye la relación 2.11 en 2.4 se obtiene:

$$\mathbf{S} = \frac{1}{I} \mathbf{Q} \mathbf{\Sigma}^2 \mathbf{Q}^T \quad (2.13)$$

donde $\mathbf{\Sigma}^2 = \text{Diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2)$ y los valores singulares están ordenados de acuerdo a $\sigma_1^2 \geq \sigma_2^2 \dots \geq \sigma_r^2$. Los autovalores se relacionan con los valores singulares mediante $\lambda_l = \frac{\sigma_l^2}{I}$, $l = 1, \dots, r$.

Para terminar esta subsección es importante considerar un par de cuestiones. Desde un punto de vista computacional, hay que tener en cuenta que tanto \mathbf{q}_j (columna j -ésima de \mathbf{Q}) como $-\mathbf{q}_j$ son soluciones adecuadas de la SVD de la matriz \mathbf{X} . Si el paquete computacional proporciona como solución \mathbf{q}_j o $-\mathbf{q}_j$, el signo de la columna j -ésima de \mathbf{P} aparecerá cambiado. En ocasiones se realiza un paso ulterior, que consiste en la rotación de la matriz de proyección \mathbf{Q} mediante una matriz \mathbf{R} de rotación, esto es, $\mathbf{Q}' = \mathbf{R}^T \mathbf{Q}$ [2, 120]. La matriz de rotación \mathbf{R} puede ser ortogonal (rotación ortogonal) o no (rotación oblicua). En esta memoria la rotación no ha ayudado en la interpretación de resultados (capítulo 3), por lo que no se aplica.

2.1.2. Interpretación de las componentes principales

Contribución de una observación a una componente principal

La importancia de una observación para una componente principal, puede ser obtenida por la razón del cuadrado de la componente principal correspondiente a esta observación entre el autovalor asociado con esta componente. Esta razón es conocida como la *contribución* de la i -ésima observación a la l -ésima componente [2], es denotada por $ctr_{i,l}$, y se expresa como

$$ctr_{i,l} = \frac{p_{i,l}^2}{\sum_i p_{i,l}^2} = \frac{p_{i,l}^2}{\lambda_l} \quad (2.14)$$

donde $p_{i,l}$ es la l -ésima componente principal y λ_l es su autovalor asociado. $ctr_{i,l}$ puede tomar valores entre 0 y 1 y para una determinada l -ésima componente la suma de las contribuciones de todas las observaciones es igual a 1. Una sugerencia útil, es basar la interpretación de una componente en las observaciones cuya contribución es mucho mayor que el promedio de la contribución, *i. e.*, observaciones cuya contribución sea mayor que $1/I$ (donde I es el número de observaciones). Las observaciones con altas contribuciones y diferentes signos pueden no ser de mucha ayuda en la interpretación de la componente, porque estas interpretaciones representan los extremos de la componente [2]. Por lo tanto, al analizar el significado del análisis PCA, hay que tener en cuenta los signos de las contribuciones por separado.

Contribución de una componente principal a una observación

La importancia de una componente para una observación dada puede ser estimada por medio del *coseno cuadrado*, e indica la contribución de una componente a la distancia al cuadrado de la observación al origen. Esto corresponde al cuadrado del coseno del ángulo del triángulo rectángulo formado con el origen, la observación y su proyección en la componente principal, y es calculado de este modo

$$\gamma_{i,l}^2 = \frac{p_{i,l}^2}{\sum_l p_{i,l}^2} = \frac{p_{i,l}^2}{d_{i,g}^2} \quad (2.15)$$

donde $d_{i,g}^2$ es el cuadrado de la distancia de una observación dada al origen. En otras palabras, $d_{i,g}^2$ es calculado como la suma de los cuadrados de todas las componentes principales para esta observación. Componentes con grandes valores de $\gamma_{i,l}^2$ contribuyen a una buena parte de la distancia total, por tanto, esas componentes tienen importancia para esa observación [2, 156].

Correlación de una componente y una variable. Factores de carga.

La correlación entre una componente principal y una variable es conocida en la jerga de las PCA como factores de carga o coeficientes espaciales (*factor loadings*). Nótese que la suma de los cuadrados de los coeficientes de correlación entre una variable y todas las componentes es igual a 1. Por esta razón, y porque los factores de carga al cuadrado indican la proporción de la varianza de las variables explicadas por las componentes, los factores de carga al cuadrado son más fáciles de interpretar que los factores de carga [2].

Especial cuidado hay que tener cuando se usa el término *loading* porque tiene varias interpretaciones y puede ser una fuente potencial de confusión. Por esta razón es importante corroborar el significado de *loading* ya sea en la lectura de textos sobre la técnica PCA o en las salidas del programa computacional empleado para calcular PCA [2, 376]. En este sentido es bastante útil la Tabla 9.3 del Wilks [376] donde se proporciona una variedad de terminología y sus sinónimos para el PCA.

2.1.3. Sobre el número de componentes principales

Una de las principales aplicaciones del análisis de componentes principales es reducir la dimensionalidad de las variables de un conjunto de datos, por tanto, un punto problemático es determinar cuántas componentes principales han de ser retenidas (téngase en cuenta que el número máximo de componentes principales que pueden ser retenidas es igual al número de variables) [78, 376]. La decisión del número de componentes puede llevar a una pérdida de información (sub-estimación) o introducir ruido aleatorio (sobre-ajuste) [78]. A pesar de que la técnica PCA tiene ya más de un siglo de existencia, este problema permanece abierto hasta hoy en día. Sin embargo, existen algunas reglas conocidas como *stopping rules* que pueden servir de ayuda para determinar el número de componentes principales a retener [156, 376].

Recientemente Peres-Neto *et al.* [265] hacen una recopilación de 20 reglas objetivas que dividen en dos grandes familias. Las reglas basadas en intervalos de confianza (*v. gr.*, el análisis paralelo, métodos de *bootstrap* basados en autovalores o el test de Bartlett) y las reglas basadas en valores promedios de *tests* estadísticos (*v. gr.*, la regla de Kaiser-Guttman, el modelo *broken-stick* o de la mínima correlación parcial promedio). Al mismo tiempo Peres-Neto *et al.* [265] hacen un análisis comparativo entre estas reglas. Ellos encontraron al utilizar datos simulados que, más que la *stopping rule* utilizada o el número de elementos y el grado de Gaussianidad de los datos analizados, estas reglas son mucho más dependientes de las correlaciones existentes entre las observaciones o entre las variables [265]. Estos resultados (dependencia de la correlación) son parecidos a los obtenidos en estudios anteriores pero con datos ambientales (datos climáticos) [276, 376].

Con lo que respecta al uso de las *stopping rules* en estudios similares al caso de estudio presentado en esta memoria (capítulo 3), los dos criterios frecuentemente utilizados son el criterio de Kaiser (Guttman) y el criterio del porcentaje de varianza acumulada (PVA) [176, 272]. El criterio de Kaiser sólo considera retener las PCs con autovalores mayores de uno [156, 376]. El problema con este criterio es que puede ser muy restrictivo (aun teniendo en cuenta la sugerencia de Jolliffe [156], esto es, retener las PCs cuyos autovalores sean iguales o mayores que 0.7) [176, 272]. El criterio PVA considera retener las PCs cuyos porcentajes de varianza

acumulados superen un determinado valor. Por ejemplo, algunos autores [271, 272] recomiendan tener en cuenta sólo las PCs que superen al menos el 90 % de la varianza acumulada (PVA_{90}). Este criterio (PVA_{90}) se utilizó en el caso de estudio del capítulo 3 de esta memoria.

2.1.4. Sobre el uso del PCA para identificación de modos

Recientemente, el PCA (en su versión clásica o rotada) ha sido utilizado para extraer modos individuales de variabilidad que puedan ser físicamente relevantes, como la Oscilación Ártica (AO), también conocidas como teleconexiones [120, 392]. Sin embargo, existen algunos problemas en la interpretación individual de los modos estadísticos (los modos o patrones espaciales que se obtienen vía PCA) identificados por el PCA como potenciales modos físicos [75, 222].

Uno de estos problemas está asociado con las limitaciones técnicas del PCA. Se sabe que, por construcción, el conjunto de autovectores o EOFs (lo que en la literatura sobre análisis de datos climáticos se conoce como ‘patrones espaciales’ [365, 376]) son mutuamente ortogonales en el espacio y el conjunto de las componentes principales están incorrelacionadas [120]. Estas características son de mucha utilidad práctica, *v. gr.*, la matriz de covarianza de cualquier subconjunto de componentes principales retenidas siempre es diagonal [121]. Sin embargo, éstas características pueden también actuar en contra del análisis que se pretende realizar.

Por ejemplo, si un conjunto de datos se puede representar como la superposición de dos patrones que no sean ortogonales, el PCA no será capaz de hallar esos patrones [7]. El problema con esto, es que los modos físicos no tienen porqué ser ortogonales [7, 75, 120]. Al mismo tiempo, los EOFs muestran una fuerte tendencia a tener la más simple estructura espacial posible dentro del dominio. Esta tendencia lleva a una fuerte dependencia de los EOFs en la forma del dominio espacial (problema de dominio-dependencia) [7, 120]. Por tanto, los modos individuales obtenidos vía PCA no tienen porqué representar por lo general modos normales del sistema dinámico a estudio [222].

Otro problema relacionado con la interpretación de los modos estadísticos y debido a la propiedad de ortogonalidad, es que el análisis vía PCA es no local,

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

esto es, dos puntos espaciales diferentes en un autovector reflejan los aspectos globales del conjunto de datos más que el comportamiento local [7]. El problema con esto es que contrasta con la técnica estadística multivariante de correlación utilizada para definir teleconexiones, donde los patrones pueden ser interpretadas localmente [7, 222].

Estos problemas mencionados en los párrafos anteriores pueden causar dificultades o limitaciones para cualquier posible interpretación física de los modos estadísticos o patrones obtenidos vía PCA [7, 75]. Por todo esto, es importante mantener presentes estas cuestiones al interpretar los modos estadísticos obtenidos mediante PCA [7, 75, 222].

2.2. Análisis espectral vía el Periodograma de Lomb-Scargle

El periodograma de Lomb-Scargle (PLS), también conocido como *Lomb-Scargle Fourier Transform* (LSFT), es una técnica de análisis espectral que puede ser aplicada a series temporales no equiespaciadas en el tiempo, y forma parte de los métodos de análisis espectral por mínimos cuadrados (*Least-Squares Spectral Analysis* (LSSA) [193, 311, 312, 356]. El origen del PLS, por lo general, está asociado a Lomb [193] y Scargle [311]. Sin embargo, retornando un poco hacia el pasado, en 1969, Barning [15] desarrolló una técnica similar, y poco más tarde, en 1971, Vanicek [356] propuso la idea de mínimos cuadrados para determinar el espectro.

Lomb [193] realizó una ampliación del trabajo de Barning [15], de tal modo que estudió las propiedades estadísticas del análisis de frecuencias por mínimos cuadrados para series no equiespaciadas en el tiempo. Entre los resultados estadísticos más importantes del trabajo de Lomb, está la demostración de que las alturas de los picos del espectro de un ruido Gaussiano tienen una distribución del tipo χ_2^2 y que existe una correlación entre la altura de este tipo de espectro para dos frecuencias cualesquiera f_1 y f_2 [193]. Un poco más recientemente, Scargle [311] demostró que el espectro obtenido del ajuste de una función armónica por mínimos cuadrados es equivalente al espectro obtenido vía el periodograma.

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

Scargle también realizó un estudio muy detallado para estimar la significación estadística de la detección de una señal.

El Periodograma de Lomb-Scargle se hizo ampliamente conocido debido al trabajo de Press *et al.* [277]. La adaptación del PLS (originalmente utilizado para la búsqueda de periodicidades en series temporales astronómicas) para el análisis de datos (paleo)climáticos fue realizada por Schulz y Stattegger [317], un poco más tarde, Schulz y Mudelsee [316] implementaron un algoritmo (REDFIT) para tener en cuenta el tipo de ruido de fondo de las series (paleo)climáticas. Recientemente, Mudelsee *et al.* [226] presentaron algunas mejoras a REDFIT y Mudelsee [225] proporcionó, entre otras temáticas, una revisión de los principales conceptos teóricos del análisis espectral y, en particular, de este algoritmo.

En esta sección se presentan los principales aspectos teóricos del Periodograma de Lomb-Scargle y de la metodología utilizada por REDFIT de acuerdo con Schulz y Stattegger [317], Schulz y Mudelsee [316] y Mudelsee *et al.* [226]. Esto con el propósito de proporcionar los elementos necesarios para un análisis adecuado de series temporales (paleo)climáticas no equiespaciadas en el tiempo. Estas consideraciones también se aplican a diferentes tipos de series temporales medioambientales no equiespaciadas en el tiempo, siempre y cuando el tipo de ruido de fondo sea rojo.

2.2.1. Definición del periodograma de Lomb-Scargle

Para una serie temporal $X(t_n)$, donde $n = 1, 2, \dots, N$ (número de elementos de la serie temporal) con media cero y varianza σ^2 (*i. e.*, estacionaria de segundo orden) y sin la presencia de tendencia y de valores extremos, el periodograma de Lomb-Scargle (normalización de Lomb) se define [140, 193, 277] como:

$$P(\omega_i = 2\pi f_i) = \frac{\hat{S}_\tau(\omega_i)}{2\sigma^2} = \frac{1}{2\sigma^2} \left[\frac{XC_\tau^2(\omega_i)}{CC_\tau(\omega_i)} + \frac{XS_\tau^2(\omega_i)}{SS_\tau(\omega_i)} \right] \quad (2.16)$$

donde ω_i es la frecuencia angular y $f_i > 0$, $i = 1, 2, \dots, k$ forman el conjunto de frecuencias en las que se desea calcular el periodograma. Cada término de la relación 2.16 está definido por

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

$$XC_{\tau}(\omega_i) = \sum_{n=1}^N X(t_n) \cos[\omega_i(t_n - \tau(\omega_i))] \quad (2.17)$$

$$CC_{\tau}(\omega_i) = \sum_{n=1}^N \cos^2[\omega_i(t_n - \tau(\omega_i))] \quad (2.18)$$

$$XS_{\tau}(\omega_i) = \sum_{n=1}^N X(t_n) \sen[\omega_i(t_n - \tau(\omega_i))] \quad (2.19)$$

$$SS_{\tau}(\omega_i) = \sum_{n=1}^N \sen^2[\omega_i(t_n - \tau(\omega_i))] \quad (2.20)$$

donde τ , la constante que asegura la invariancia en el tiempo [311], está definida por

$$\tau(\omega_i) = \frac{1}{2\omega_i} \arctan \left[\frac{\sum_{n=1}^N \sen(2\omega_i t_n)}{\sum_{n=1}^N \cos(2\omega_i t_n)} \right] \quad (2.21)$$

2.2.2. El periodograma clásico vs. el PLS

En esta sección se presentan algunas diferencias entre el periodograma clásico y el periodograma de Lomb-Scargle. La definición de periodograma clásico [311, 318] para el caso de muestreos equiespaciados en el tiempo es

$$p(f_i) = \frac{1}{N} \left(\left[\sum_{n=1}^N X(t_n) \cos(\omega_i t_n) \right]^2 + \left[\sum_{n=1}^N X(t_n) \sin(\omega_i t_n) \right]^2 \right) \quad (2.22)$$

donde $\omega_i = 2\pi f_i$, para $i=1, 2, \dots, k$, los tiempos de muestreo están dados por $t_n = n \cdot \Delta t$, para $n=1, 2, \dots, N$ y Δt es la longitud constante del intervalo de muestreo.

1. Frecuencias ortogonales. Tradicionalmente el periodograma clásico para series equiespaciadas en el tiempo, está definido [279, 311] para un conjunto discreto

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

de frecuencias ortogonales f_r , para las cuales las estimaciones del periodograma son independientes entre sí, esto es:

$$f_r = \frac{r}{T_{Tot}} \quad (2.23)$$

para $r = 0, 1, \dots, \lfloor N/2 \rfloor$. Donde $\lfloor N/2 \rfloor$ es el número máximo de frecuencias y está definido por la función máximo entero (*i. e.*, se redondea al entero más próximo para el caso cuando N es impar) y, donde $T_{Tot} = t_N - t_1$, *i. e.*, la duración de la serie temporal y para el caso equiespaciado $T_{Tot} = (N - 1) \cdot \Delta t$. Las frecuencias f_r al ser utilizadas en la relación 2.22, aseguran que el poder espectral correspondiente a diferentes frecuencias sea independiente para cualquier par de frecuencias del conjunto de frecuencias ortogonales y que sea retenida la máxima cantidad de información posible de los datos originales [354]. Sin embargo, para el caso no equiespaciado en el tiempo, no es posible encontrar un conjunto de frecuencias ortogonales para las cuales el poder espectral sea independiente [311, 354].

2. La frecuencia (razón) de muestreo y la frecuencia de Nyquist. Para el caso equiespaciado en el tiempo la frecuencia de muestreo f_s puede definirse como [277]

$$f_s = \frac{1}{\Delta t}, \quad (2.24)$$

por lo cual la frecuencia de Nyquist f_{Nyq} viene dada por

$$f_{Nyq} = \frac{f_s}{2} = \frac{1}{2\Delta t} \quad (2.25)$$

Para el caso no equiespaciado en el tiempo, no hay definida una frecuencia de muestreo y por tanto, una frecuencia de Nyquist [311, 354]. No obstante, hay varias propuestas para manejar este problema. Por ejemplo, Van-Dongen *et al.* [354] proponen utilizar la frecuencia real de Nyquist calculada por medio de la ventana espectral (*periodogram window*) (veáse la sección 2.2.3 para más detalles). Bretthorst [41] propone la frecuencia crítica de Nyquist, definida como

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

$$f_{cNyq} = \frac{1}{2\Delta t'} \quad (2.26)$$

donde $\Delta t'$ es siempre menor o igual que el más pequeño intervalo de tiempo de muestreo. Por último, Schulz y Statteger [317] proponen una frecuencia de Nyquist promedio, esto es

$$\langle f_{Nyq} \rangle = \frac{1}{2 \langle \Delta t \rangle} \quad (2.27)$$

donde $\langle \Delta t \rangle$ es el promedio de los intervalos temporales de muestreo.

3. El teorema de Parseval. Una propiedad importante del periodograma clásico, para series temporales equiespaciadas en el tiempo, es el teorema de Parseval (la energía total de una señal real es la misma ya sea si es calculada en el dominio del tiempo o en el de la frecuencia [277]). Sin embargo, cuando las series no están equiespaciadas en el tiempo este teorema no se cumple [354].

2.2.3. Estimación de las frecuencias naturales

Como se mencionó anteriormente, las frecuencias definidas por la relación 2.23 para el caso no equiespaciado no son ortogonales. Sin embargo, de acuerdo con Fullerton [100], estas frecuencias pueden considerarse aproximadamente ortogonales para el caso no equiespaciado [354]. A este conjunto de frecuencias, utilizadas para calcular el periodograma, se conocen como *frecuencias naturales* y están definidas [311, 354] por

$$f_r' = \frac{r'}{T_{Tot}} \quad (2.28)$$

para $r' = 0, 1, 2, \dots, M_{max}$. Donde M_{max} es el número máximo de frecuencias naturales.

Para el caso donde los muestreos no están equiespaciados en el tiempo, M_{max} depende de los tiempos de muestreo y la frecuencia $f_{M_{max}}$ no necesariamente corresponde con la frecuencia de Nyquist que aparecería si las muestras estuvieran equiespaciadas temporalmente. Por tanto, es necesario utilizar una alternativa

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

válida a la frecuencia de Nyquist [311, 354]. Por ejemplo, si se opta por utilizar la *frecuencia real de Nyquist* sugerida por Van-Dongen *et al.* [354], el procedimiento es el siguiente.

La frecuencia real de Nyquist y la ventana espectral. Para poder determinar la frecuencia real de Nyquist f_{aNyq} se hace uso de una función matemática conocida como *ventana espectral* o *función de respuesta espectral* [354]. Esta función depende sólo de los tiempos de muestreo, $\{t_n\}_{n=1}^N$ y de la diferencia de las frecuencias $\Delta f = f_a - f_b$, donde f_a y f_b son dos frecuencias consecutivas cualesquiera. Para el caso cuando el muestreo está equiespaciado en el tiempo, existe una fórmula exacta para representar la ventana espectral, pero para el caso correspondiente a un muestreo no equiespaciado en el tiempo, no existe una fórmula exacta para su representación [311, 354]. Sin embargo, es posible utilizar la siguiente aproximación [193, 311, 354]

$$W(\Delta f) = \frac{1}{N^2} \left[\left(\sum_{n=1}^N \cos(2\pi\Delta f(t_n - \theta)) \right)^2 + \left(\sum_{n=1}^N \sen(2\pi\Delta f(t_n - \theta)) \right)^2 \right] \quad (2.29)$$

donde $\theta = (t_1 + t_N)/2$. Los valores de $W(\Delta f)$ están definidos entre 0 y 1 y es una función simétrica con respecto a Δf , *i. e.*, $W(-\Delta f) = W(\Delta f)$ [354].

Para determinar la f_{aNyq} se hace uso del fenómeno conocido como *aliasing*. Esto es, para dos frecuencias cuya suma o diferencia es un múltiplo entero de la frecuencia de muestreo, es posible construir una pareja de sinusoides que, después del muestreo, no son distinguibles una de la otra debido a que contienen las mismas muestras (Figura 1.1).

Las copias debido al fenómeno del *aliasing* aparecerán cada $2mf_{aNyq}$, donde $m = \dots, -2, -1, 1, 2, \dots$. Por ejemplo, para el caso donde la serie temporal es cercanamente equiespaciada en el tiempo (las muestras están uniformemente distribuidas), se puede utilizar una f_{aNyq} equivalente al caso de que la serie estuviera equiespaciada. A continuación se realiza una inspección visual, para ello se dibuja la ventana espectral (relación 2.29) utilizando como límite de las frecuencias a la f_{aNyq} y un valor de m mayor o igual a 2 (‘réplicas’). Cuando la ventana espectral se repita

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

en un determinado valor de m , entonces el valor de f_{aNyq} es el límite para las frecuencias donde se puede calcular el periodograma de Lomb-Scargle [140, 354].

En el caso que las muestras no equiespaciadas temporalmente formen grupos que estén espaciados más o menos regularmente, entonces la máxima frecuencia natural estará determinada por el número de grupos n' [140, 354], esto es:

$$f_{aNyq} = \frac{(n' - 1)/2}{T_{Tot}} \quad (2.30)$$

Nótese que M_{max} no está determinado por el número de muestras. Las muestras individuales dentro de un grupo, pueden ser consideradas como medidas redundantes obtenidas aproximadamente al mismo tiempo, por ello, más muestras dentro de cada grupo no incrementan la f_{aNyq} . Sin embargo, incrementan la significación estadística del periodograma [354].

2.2.4. El periodograma: crudo vs. suavizado

El periodograma crudo, *i. e.*, el que se obtiene utilizando la relación 2.16, es un estimador inconsistente del espectro [20]. Esto es, el radio de la desviación estándar o coeficiente de variación del periodograma no tiende a cero a medida que el número de elementos de la serie aumenta [20, 225]. Un método para mejorar la estimación del espectro en términos de tendencia, varianza y supresión de falsos picos espectrales, de tal modo que se obtenga un estimador consistente fue propuesto por Welch [374]. Esta técnica es conocida como *Welch-Overlapped-Segment-Averaging* (WOSA) y fue originalmente diseñada para el caso correspondiente a series temporales equiespaciadas en el tiempo [374]. Para el caso de series no equiespaciadas temporalmente, Schulz y Statteger [317] hicieron una adaptación de la técnica de Welch. El método propuesto por Schulz y Statteger [317], consiste en partir la serie temporal a estudio (de longitud N) en varios segmentos (n_{50}) de longitud $N_{Seg} = 2N/(n_{50} + 1)$, de tal modo que cada uno de estos segmentos se solapen durante la mitad de su longitud (véanse Figuras 2.1 y 2.2).

Para evitar los efectos en el espectro de las irregularidades de los extremos de los N_{Seg} segmentos al usar WOSA, se ha de aplicar (multiplicar) algún tipo de función-ventana (conocida como *taper*) a cada segmento n_{50} en el dominio

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

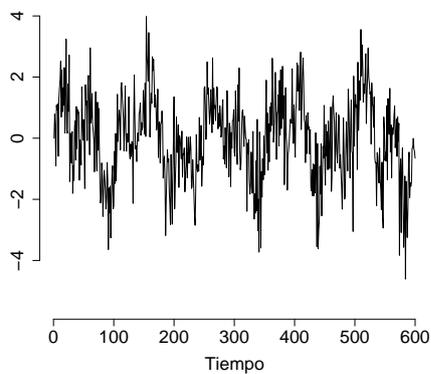


Figura 2.1: Serie temporal ejemplo (suma de dos funciones armónicas con periodos $2\pi/50$ y $2\pi/120$ más ruido Gaussiano $\mathcal{N}(0, 0.95)$) con N elementos a la cual se le aplica la técnica WOSA.

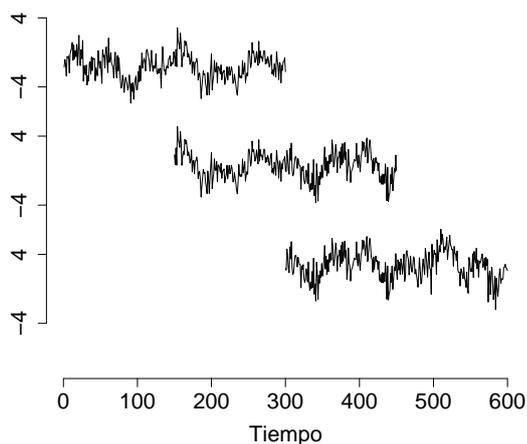


Figura 2.2: Diagrama donde se muestra el procedimiento de la técnica WOSA utilizando la serie de la Figura 2.1, con $n_{50} = 3$ y $N_{Seg} = N/2$.

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

del tiempo [317, 374]. Existen diferentes tipos de ventanas (una amplia revisión puede encontrarse en [123, 238]), pero la que se ha utilizado en este trabajo es la ventana de Hann o de Hanning y está definida [123] como

$$w(n) = 0.5 \left[1 - \cos \left(\frac{2\pi n}{N_{Seg} - 1} \right) \right] \quad (2.31)$$

para $n=0, 1, \dots, N_{Seg} - 1$.

La función-ventana es escalada de esta forma $\sum w_n^2 = N_{Seg}$. A continuación, para los n_{50} segmentos de la serie temporal a los cuales se les ha aplicado la función ventana escalada, se calcula el periodograma de Lomb-Scargle mediante 2.16. Por último, se calcula el promedio de los n_{50} espectros crudos. El espectro promedio resultante es un estimador consistente del (auto)espectro [20, 317], se le conoce como espectro suavizado y se define de este modo

$$\hat{G}_{xx}(f_i) = \frac{2}{n_{50}\Delta f N_{seg}} \sum_{n=1}^{n_{50}} |X_n(f_i)|^2, \quad i = 1, 2, \dots, k \quad (2.32)$$

donde $\Delta f = 1/(N_{seg} \langle \Delta t \rangle)$ y $|X_n(f_i)|^2$ es el PLS definido en la relación 2.16.

Dado el escalado del espectro suavizado (relación 2.32), se cumple la siguiente relación con la varianza de la serie temporal:

$$\Delta f \sum \hat{G}_{xx} = \sigma_x^2 \quad (2.33)$$

donde Δf es la frecuencia fundamental [317].

Debido a que las componentes de los espectros crudos siguen una distribución del tipo χ^2 [262, 279] con 2 grados de libertad, 2.32 también sigue ese tipo de distribución [317]. Cada uno de los n_{50} espectros de la ecuación 2.32 incrementa los grados de libertad, por ello, se reduce el error estándar de la estimación del espectro. Sin embargo, el solapamiento de los n_{50} segmentos introduce una correlación entre los segmentos y un número efectivo de segmentos n_{eff} . Donde el n_{eff} es mucho menor que el n_{50} y la relación entre ellos viene dada por

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

$$n_{eff} = n_{50} \left(1 + 2c_{50}^2 - \frac{2c_{50}^2}{n_{50}} \right)^{-1} \quad (2.34)$$

donde c_{50} es una constante que depende del tipo de la función ventana, por ejemplo, para la ventana de Hanning $c_{50} = 0.167$ [123]. Por tanto, los grados de libertad ν para 2.32 dependerán del n_{eff} (relación 2.34) y por tanto de la constante c_{50} , esto es:

$$\nu = 2n_{eff} \quad (2.35)$$

La distribución del (auto)espectro 2.32 viene dada por la relación [262, 279]

$$\frac{\hat{G}_{xx}(f_i)}{G_{xx}(f_i)} = \frac{\chi^2}{\nu} \quad (2.36)$$

de aquí se sigue que un $(1 - \alpha)$ intervalo de confianza para $G_{xx}(f_i)$, basado en la estimación de $\hat{G}_{xx}(f_i)$, puede ser calculada por [20]

$$\left[\frac{\nu \hat{G}_{xx}(f_i)}{\chi_{\nu; \alpha/2}^2} \leq G_{xx}(f_i) \leq \frac{\nu \hat{G}_{xx}(f_i)}{\chi_{\nu; 1-\alpha/2}^2} \right] \quad (2.37)$$

Algunos autores [20, 262, 279] aplican una transformación logarítmica del tipo $10 \log_{10}$ a la relación 2.37, esto con el objeto de que los intervalos de confianza sean independientes de la frecuencia y al graficar el \hat{G}_{xx} se obtenga una mejor visualización.

La mínima resolución de las frecuencias en \hat{G}_{xx} , *i. e.*, el ancho de banda espectral, B_w , se calcula como

$$B_w = \beta_w \Delta f \quad (2.38)$$

donde β_w es el ancho de banda normalizado y depende del tipo de ventana (*taper*) utilizada, en este trabajo tiene un valor de 2 (ventana de Hanning) [123, 238, 317].

2.2.5. Contraste de hipótesis del PLS debido a ruido rojo

Es bien sabido que muchos fenómenos ambientales (climáticos, meteorológicos, etc) o geofísicos, presentan un espectro de fondo rojo [107, 124, 316]. Hasselmann [124] propuso que un proceso estocástico autoregresivo de primer orden (AR1), es un modelo adecuado para representar ese espectro de fondo rojo, y puede ser utilizado como una hipótesis nula en un test de significación. Si $R(t_r)$ es un proceso AR1, donde los tiempos de muestreo t_r ($r=1,2,\dots,N$) no están equiespaciados, la distribución del periodograma de Lomb-Scargle para este proceso no puede ser calculada analíticamente [316, 226]. Por tanto, es necesario utilizar métodos de simulación de Monte-Carlo para inferir las propiedades estadísticas de la distribución del periodograma [226]. El sesgo en el espectro es debido a que las componentes individuales del PLS no son independientes unas de las otras [193, 311].

Schulz y Mudelsee [316] demostraron por medio de simulaciones de Monte Carlo, que el periodograma de Lomb-Scargle de un proceso AR1 no equiespaciado temporalmente, sobrestima las amplitudes espectrales localizadas en las altas frecuencias (*i. e.*, las frecuencias al final del espectro). Ellos también desarrollaron un algoritmo computacional conocido como REDFIT (utilizado en esta memoria) para corregir ese sesgo. A continuación se presentan los conceptos estadísticos necesarios para luego proporcionar la descripción del REDFIT. Más detalles pueden consultarse en [224, 226, 316].

Un proceso discreto autoregresivo de orden 1 (AR1) y no equiespaciado temporalmente, está definido [226, 291, 316] por la siguiente relación

$$\begin{aligned}
 R(t_1) &= \mathcal{E}_{N(0,1)}, & \text{para } r=1 \\
 R(t_r) &= \rho_r R(t_{r-1}) + \mathcal{E}_{N(0,\sigma_{\mathcal{E}})}(t_r), & \text{para } r=2, 3, \dots, N \\
 \rho_r &= \exp(-(t_r - t_{r-1})/\tau_m)
 \end{aligned}
 \tag{2.39}$$

donde τ_m es la escala temporal correspondiente al proceso AR1 (una medida de su memoria) y \mathcal{E} es un ruido Gaussiano de media cero y varianza

$$\sigma_{\mathcal{E}}^2 = 1 - \exp(-2(t_r - t_{r-1})/\tau_m)
 \tag{2.40}$$

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

El valor de σ_{ϵ}^2 asegura que el proceso AR1 sea estacionario y tenga varianza unidad [316].

El espectro teórico¹ $G_{rr}(f_i)$ de un proceso AR1 (ecuación 2.39) está definido [262, 316] por

$$G_{rr}(f_i) = G_o \frac{1 - \langle \rho \rangle^2}{1 - 2 \langle \rho \rangle \cos(\pi f_i / \langle f_{Nyq} \rangle) + \langle \rho \rangle^2} \quad (2.41)$$

para las frecuencias $f_i, j = 1, 2, \dots, K$, desde la frecuencia fundamental Δf (véase la relación 2.32) hasta la frecuencia promedio de Nyquist [316]. Donde G_o es la amplitud espectral promedio, $\langle \rho \rangle$ es el coeficiente de autocorrelación promedio, *i. e.*, $\langle \rho \rangle = \exp(-\langle \Delta t \rangle / \tau_m)$, donde $\langle \Delta t \rangle$ es la media de los intervalos temporales del proceso AR1 [316] y el valor de τ_m se estima mediante el algoritmo propuesto por Mudelsee [224] con las frecuencias utilizadas para calcular el espectro.

La desviación sistemática entre un espectro teórico de ruido rojo (relación 2.41) y otro espectro estimado, para una serie temporal de la forma 2.39, por medio del periodograma de Lomb-Scargle, depende de los tiempos de muestreo en el intervalo $[t_1, t_N]$ [193, 311]. Para algunas distribuciones arbitrarias de tiempos de muestreo, la carencia de una solución analítica para la desviación (sesgo), impide aplicar una corrección directa al periodograma de Lomb-Scargle [316]. Para lograr superar este obstáculo, se utilizan métodos de Monte-Carlo en combinación con técnicas de *bootstrap*. Basadas en los tiempos de muestreo, se genera un ensemble de N series temporales sintéticas AR1 (ecuación 2.39). La desviación del espectro promedio del ensemble con respecto al espectro teórico (ecuación 2.41) es utilizada para hacer la corrección [316, 226].

A continuación se describe el algoritmo REDFIT tal como es presentado por Schulz y Mudelsee [316].

¹Para el caso equiespaciado, el término coseno del espectro teórico de un AR1 está definido [152] por $\cos(2\pi f_i \Delta t)$, para $-\frac{1}{2\Delta t} \leq f_i \leq \frac{1}{2\Delta t}$. Por esta razón, el término coseno en la relación 2.41 al sustituir $2 \langle \Delta t \rangle$ por $\frac{1}{\langle f_{Nyq} \rangle}$ se expresa como $\cos(\pi f_i / \langle f_{Nyq} \rangle)$.

2.2 Análisis espectral vía el Periodograma de Lomb-Scargle

1. Se estima τ_m para la serie temporal a estudio $X(t_r)$ ($r = 1, 2, \dots, N$) utilizando el algoritmo de Mudelsee [224]. En caso de utilizar varios n_{50} segmentos (WOSA), se calcula el promedio de los τ_m 's de todos los n_{50} segmentos.
2. Se estima mediante el PLS el espectro $\widehat{G}_{xx}(f_i)$ de $X(t_r)$ en el intervalo $[0, < f_{Nyq} >]$ y se determina el área bajo $\widehat{G}_{xx}(f_i)$ para estimar la varianza de $X(t_r)$.
3. Se realizan $p=1$ a N_{sim} simulaciones de Monte-Carlo
 - Se crea la p -ésima serie temporal AR1 por medio de 2.39 usando los tiempos t_r 's de la serie $X(t_r)$, el τ_m estimado en el paso 1 y un conjunto independiente de $\mathcal{E}(t_r)$'s.
 - Se estima el espectro $\widehat{G}_{rr}(f_i)$ para la serie AR1
 - Se escala $\widehat{G}_{rr}(f_i)$ de tal modo que su área bajo el espectro sea igual al área bajo el espectro de $\widehat{G}_{xx}(f_i)$.

Se determina la media aritmética de los N_{sim} espectros correspondientes a las series AR1, $< \widehat{G}_{rr}(f_i) >$

4. Se calcula el espectro teórico $G_{rr}(f_i)$ de un proceso AR1 utilizando la relación 2.41 y el valor estimado de τ_m .
5. Se calcula G_o (relación 2.41) de tal modo que el área bajo $G_{rr}(f_i)$ sea igual al área bajo $\widehat{G}_{xx}(f_i)$.
6. Se calcula un factor de corrección $c(f_i)$ para ajustar la desviación del espectro Lomb-Scargle, *i. e.*

$$c(f_j) = \frac{\langle \widehat{G}_{rr}(f_i) \rangle}{G_{rr}(f_i)} \quad (2.42)$$

7. Se calcula el espectro de Lomb-Scargle corregido

$$\widehat{G}'_{xx}(f_i) = \frac{\widehat{G}_{xx}(f_i)}{c(f_i)} \quad (2.43)$$

8. Para calcular la significación estadística de un pico espectral, el intervalo superior de confianza del ruido AR1 es calculado para varios niveles de significación (*v. gr.*, 0.95 o 0.99) siguiendo una distribución del tipo χ^2 con ν grados de libertad (calculados con 2.35). También es posible calcular niveles de significación a partir de los percentiles del ensemble de Monte-Carlo.

2.3 Análisis de devianza mediante Modelos Lineales Generalizados

9. Se verifica si el modelo AR1 es adecuado para describir $X(t_r)$ probando la igualdad de $G_{rr}(f_i)$ y $\widehat{G}'_{xx}(f_i)$ usando el test no paramétrico de Wald-Wolfowitz (más comúnmente conocido como *runs test*) [20].

2.3. Análisis de devianza mediante Modelos Lineales Generalizados

La variación espacio-temporal en la distribución y abundancia de diferentes organismos, entre ellas las especies marinas, es una propiedad inherente de los sistemas ecológicos [8, 36]. Por ello, y como bien apuntan Anderson y Gribble [8], una valoración de la relativa importancia de los factores bióticos y abióticos que tienen influencia sobre las especies debería de considerar, siempre que sea posible, ambas variaciones. Sin embargo, cuantificar el papel de los factores bióticos y abióticos en la distribución y abundancia de las especies no es una fácil tarea. Para ello han sido desarrolladas o adaptadas diversas metodologías estadísticas. Entre ellas destacan, principalmente, los trabajos de Borcard *et al.* [36] y Borcard y Legendre [35].

Borcard *et al.* [36] y Borcard y Legendre [35] propusieron el uso del análisis de correspondencia canónica para particionar la variación¹ presente en datos ecológicos en diferentes porciones relacionadas a descriptores (componentes) espaciales y ambientales. También demostraron la importancia de considerar el posible solapamiento entre las componentes espaciales y ambientales. Para evitar la redundancia de información en los análisis, realizaron una partición de la variación total en datos ecológicos utilizando cuatro diferentes componentes [35, 36]. 1) La variación debida sólo a factores ambientales, *i. e.*, la fracción de la variación que puede ser explicada por los descriptores ambientales independientemente de cualquier estructura espacial, 2) la variación debida sólo a las estructuras espaciales independientes de cualquier variable ambiental, 3) la variación espacial debida al ambiente, *i. e.*, la componente cuya variación se solapa entre la variación

¹Nótese que se utiliza partición de la variación en lugar de partición de la varianza, esto es debido a que el método utilizado por Borcard *et al.* [36], y de acuerdo con Anderson y Gribble [8], no es estrictamente una partición de varianza (análisis de varianza, ANOVA).

2.3 Análisis de devianza mediante Modelos Lineales Generalizados

debida a factores ambientales y espaciales y 4) la variación no explicada por los factores ambientales y las estructuras espaciales.

Posteriormente Anderson y Gribble [8] extendieron los trabajos de Borcard *et al.* [36] y Borcard y Legendre [35], incorporando una componente temporal y las combinaciones con las otras componentes (en total 8 combinaciones). Esta componente temporal, representa los patrones temporales presentes en los datos que son independientes de cualquier variable ambiental o espacial incluida en el análisis [8]. Adaptación posteriores de estos trabajos previos [8, 35, 36] utilizó modelos lineales generalizados para relacionar variables climáticas con las capturas de atún rojo y realizar una partición de la varianza, considerando solamente las componentes ambiental, temporal y la combinación de ambas [102]. Esta metodología se utilizó en uno de los casos de estudio (capítulo 5) presentados en esta memoria. Por tanto, se presentan a continuación los principales aspectos de esta metodología, siguiendo principalmente a [8, 35, 36, 115, 116, 181, 379].

2.3.1. Los modelos lineales generalizados

Los Modelos Lineales Generalizados (GLM, *Generalized Linear Models*) tienen sus orígenes a principios de la década de los setenta, cuando fueron implementados por Nelder y Wedderburn [234]. Estos modelos de regresión extienden y generalizan los modelos lineales clásicos de dos maneras [74, 81]. La primera, consiste en que estos modelos permiten una mayor variedad de distribuciones (no sólo la distribución Normal o Gaussiana) para la variable respuesta (dependiente). Por ello, el ajuste de los modelos se puede intentar con diferentes tipos de distribuciones (*v. gr.*, la gamma, la exponencial, la binomial, etc.). Así se evita la transformación de los datos a estudio. La segunda, es que permiten establecer una relación entre la variable dependiente y la combinación lineal de predictores por medio de una función apropiada del valor esperado de esta función.

A continuación se presentan las componentes de los GLM de acuerdo con McCullagh y Nelder [215]. Para una introducción elemental a los GLM puede consultarse Dobson [74] y para conceptos más avanzados puede consultarse Dumett [81], McCullagh y Nelder [215] o más recientemente Wood [379].

Las tres componentes principales de un GLM [215] son:

2.3 Análisis de devianza mediante Modelos Lineales Generalizados

1. **La componente aleatoria \mathbf{Y}** , cuyas componentes están independientemente distribuidas con $\mathbb{E}(\mathbf{Y}) = \boldsymbol{\mu}$ y varianza constante σ^2 . \mathbf{Y} es una variable aleatoria que pertenece a la familia exponencial de distribuciones, tales como la Gaussiana, binomial, Poisson, Gamma o Gaussiana inversa [234]. Es importante tener en cuenta que los datos a estudio representados por el vector \mathbf{y} de n observaciones, se asume que son una realización de la variable aleatoria \mathbf{Y} .
2. **La componente sistemática** está definida por

$$\boldsymbol{\eta} = \sum_{j=1}^p \mathbf{x}_j \beta_j \quad (2.44)$$

donde \mathbf{x}_j es la j -ésima variable explicatoria (o independiente) que en conjunto con otros predictores y unos coeficientes β_j produce el predictor lineal $\boldsymbol{\eta}$. β_j es una componente de un vector de parámetros desconocidos cuyos valores son estimados de forma óptima (máxima verosimilitud) utilizando los datos a estudio.

3. **La función de enlace $g(\cdot)$** , que relaciona la componente aleatoria con la sistemática, y puede ser expresada como

$$\eta_i = g(\mu_i) = g(\mathbb{E}(Y_i)) \quad (2.45)$$

donde $g(\cdot)$ es una función monótona diferenciable. Ejemplos de funciones de enlace, pueden ser la función *identidad*, *log*, *recíproca*, *logit* o *probit*. Sin embargo, en la aplicación de los GLM en el capítulo 5 de esta memoria se utiliza la función identidad ($\eta = \mu$) porque se asume que la distribución de los errores es Gaussiana.

2.3.2. La componente temporal

Los efectos temporales representan una medida de la autocorrelación temporal de la serie a estudio (capturas de atún rojo en caso de estudio del capítulo 5) y en cierto sentido representan de un modo muy simplificado la dinámica poblacional de la especie a estudio, donde las capturas pueden considerarse como un testigo de abundancia [102, 335]. La forma de estimar la variación temporal está basada

2.3 Análisis de devianza mediante Modelos Lineales Generalizados

en cómo se estima la variación espacial en Legendre [180], Borcard *et al.* [36] y Anderson y Gribble [8], pero desde un punto de vista unidimensional (en términos del tiempo). Para ello, se ajusta mediante modelos lineales generalizados un polinomio de grado N de la forma

$$f(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots \beta_N t^N \quad (2.46)$$

donde t son los años donde se realizan las capturas. El grado del polinomio viene dado por el mejor ajuste polinómico entre tiempos y capturas. Por ejemplo, para el caso de las capturas de atún rojo (capítulo 5) $N = 3$.

2.3.3. La partición de la varianza

La partición de la varianza de la serie a estudio, *i. e.*, de la serie de capturas de atún rojo (capítulo 5), se realizó teniendo en cuenta cuatro componentes [102].

1. La componente que refleja la variación debida sólo a los efectos de las variables ambientales (A),
2. la componente que refleja la variación debida sólo a los efectos de las variables temporales (T),
3. la componente que refleja la variación debida al efecto combinado entre las variables ambientales y las temporales (C), y
4. la componente que refleja la variación no explicada por las variables ambientales, temporales o por la combinación de éstas (NE).

La varianza explicada de las capturas, se obtiene de cada una de las devianzas explicadas D^2 , obtenidas al ajustar los modelos lineales generalizados a las variables que conforman la componente ambiental $D^2(\text{GLM}_A)$, temporal $D^2(\text{GLM}_T)$ y la combinación de estas $D^2(\text{GLM}_C)$. La devianza explicada D^2 , es la cantidad de varianza explicada por un modelo lineal generalizado, proporciona una medida de la variabilidad de los datos a estudio y es equivalente al coeficiente de determinación R^2 utilizado en los modelos lineales ajustados por mínimos cuadrados

2.3 Análisis de devianza mediante Modelos Lineales Generalizados

[116, 234]. La devianza explicada D^2 y de acuerdo a [116] se obtiene mediante la relación

$$D^2 = 1 - (\text{devianza residual})/(\text{devianza nula}) \quad (2.47)$$

En primer lugar se obtiene la varianza en común debido a los efectos combinados de la componente ambiental y la temporal. Pero debido a que existe un solapamiento entre la componente ambiental y temporal, que se puede visualizar mejor con un diagrama de Venn (Figura 2.3), la varianza explicada en común viene dada por

$$\text{Var.Común} = D^2(\text{GLM}_A) + D^2(\text{GLM}_T) - D^2(\text{GLM}_C) \quad (2.48)$$

La varianza explicada por la componente ambiental solamente viene dada por

$$\text{Var.Amb} = D^2(\text{GLM}_A) - \text{Var.Común} \quad (2.49)$$

La varianza explicada por la componente temporal solamente viene dada por

$$\text{Var.Temp} = D^2(\text{GLM}_T) - \text{Var.Común} \quad (2.50)$$

Por último, la varianza no explicada viene dada por

$$\text{Var.NoExp} = 1 - \text{sum}(\text{Var.Amb}, \text{Var.Temp}, -\text{Var.común}) \quad (2.51)$$

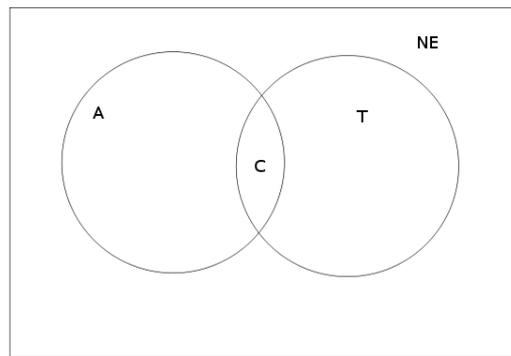


Figura 2.3: Diagrama de Venn que muestra la partición de variación de acuerdo a dos conjuntos de variables independientes, la componente ambiental (A) y la temporal (T). Basado en Anderson y Gribble [8].

2.4. Análisis espectral vía la transformada continua de Wavelet

El análisis espectral vía la transformada continua de Wavelet (CWT, *Continuous Wavelet Transform*) es una herramienta poderosa para enfrentar el problema de la no estacionariedad o la presencia de eventos transitorios en una serie temporal. Esta técnica realiza una descomposición espectral localizada de la serie a estudio determinando los modos dominantes de variabilidad y cómo varían éstos en tiempo y escala (frecuencia) [114, 350]. El análisis espectral vía la CWT incluye la estimación del poder espectral (análisis univariante), análisis espectral cruzado, la coherencia y la fase de la coherencia (análisis bivariante) [113, 206, 207, 350].

Las primeras aplicaciones de esta metodología a series temporales ambientales (climáticas, hidro-meteorológicas, etc.) se remontan a principios de los noventa [88, 172, 177, 219]. Sin embargo, no fue hasta la publicación del trabajo de Torrence y Compo [350] que esta herramienta llegó a ser popular. Quizá, en parte por la forma didáctica de esta publicación, pero también porque ellos hicieron pública una herramienta computacional para realizar el análisis espectral vía la CWT utilizando diferentes funciones wavelet (entre ellas la de Morlet).

Con lo que respecta a las aplicaciones del análisis espectral vía la CWT a series ecológicas (en particular, series que representan poblaciones de especies marinas) y su relación con variables ambientales, son menos frecuentes y más recientes. Uno de los primeros trabajos fue realizado por Machu *et al.* [198], quienes estudiaron la distribución del fitoplancton. Posteriormente, Ménard *et al.* [218] estudiaron la relación entre la variabilidad climática y la población del atún en el Océano Índico. Rouyer *et al.* [301] estudiaron ampliamente diferentes modelos de ruido de fondo y los aplicaron a series de capturas de atún rojo. Más recientemente, Hsieh *et al.* [144] estudiaron la relación de abundancia de la anchoa (en estado larval) y las variables ambientales en la costa Sudoeste de Taiwan. Una revisión sobre la aplicación de esta metodología a series ecológicas fue realizada por Cazelles *et al.* [50].

A continuación se presentan los principales aspectos técnicos del análisis espectral vía la transformada continua de Wavelet, tal y como se presentan en las referencias básicas sobre este tema [50, 113, 173, 177, 206, 207, 219, 350].

2.4.1. Definición formal de Wavelet

Los Wavelets son un tipo de funciones reales o complejas $\psi(t)$ que, al igual que las funciones seno o coseno, pueden utilizarse como una base en un espacio de funciones. Tienen forma de pequeñas ondas que oscilan alrededor de cero, están definidas sobre el eje real $(-\infty, \infty)$ y satisfacen las siguientes condiciones [263].

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.52)$$

$$\int_{-\infty}^{\infty} \psi^2(t) dt = 1 \quad (2.53)$$

La primera condición, conocida como de admisibilidad, implica que la función Wavelet tenga media cero. La segunda condición asegura una localización suficiente en el dominio del tiempo y de la escala [88, 350]. Con respecto a la condición de admisibilidad, adicionalmente, la transformada de Fourier $\Psi(f)$ de una función wavelet $\psi(t)$ debe satisfacer [88, 263] la siguiente condición:

$$0 < \int_0^{\infty} \frac{|\Psi(f)|^2}{f} df < \infty \quad (2.54)$$

Existen diferentes tipos de funciones Wavelet, *v. gr.*, ortogonales, no ortogonales, continuas, discretas, reales, complejas, etc. Las pautas de qué tipo de función Wavelet se utilizará, dependerá del tipo de proceso (fenómeno) a estudio. Por ejemplo, en ciencias ambientales y geociencias es muy común el uso de funciones Wavelet continuas de tipo complejo, en especial el Wavelet de Morlet [50, 173, 350]. La razón de esto es que en el análisis de este tipo de señales (o series temporales, dependiendo del contexto) es necesario extraer información de la amplitud y de la fase del proceso a estudio [50, 173].

El Wavelet de Morlet

El Wavelet de Morlet (Figura 2.4) es esencialmente una función exponencial compleja, cuya amplitud está modulada por una función proporcional a la función

2.4 Análisis espectral vía la transformada continua de Wavelet

de densidad de probabilidad de una distribución Gaussiana (estándar) [172, 263], esto es

$$\psi(t) = C e^{-i\omega_0 t} (e^{-t^2/2} - \sqrt{2} e^{-\omega_0^2/4} e^{-t^2}) \quad (2.55)$$

donde el valor de la constante C es seleccionado de tal modo que 2.55 pueda satisfacer la condición de admisibilidad 2.53 para un determinado valor de ω_0 [263]. Esto sucede cuando $\omega_0 \geq 5$ y, si se cumple la igualdad, entonces $C = 0.7528 \sim \pi^{-1/4}$. Si $\omega_0 \geq 5$, el segundo término de la expresión entre paréntesis de 2.55 se aproxima a cero. Por ello, es posible expresar la relación 2.55 [172, 263] como se hace habitualmente:

$$\psi(t) = \pi^{-1/4} e^{-i\omega_0 t} e^{-t^2/2}, \quad \omega_0 \geq 5 \quad (2.56)$$

A la expresión 2.56 se le conoce como el Wavelet abreviado de Morlet, aunque es común utilizar simplemente el término Wavelet de Morlet [172]. La transformada de Fourier del Wavelet de Morlet viene definida [172] por

$$\hat{\Psi}(\omega) = \pi^{-1/4} e^{-(\omega - \omega_0)^2/2}, \quad \omega_0 \geq 5 \quad (2.57)$$

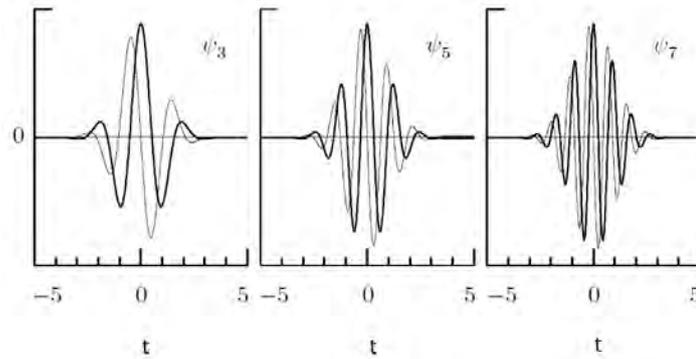


Figura 2.4: El Wavelet de Morlet con tres diferentes valores de ω_0 , 3, 5 y 7, respectivamente. Nótese la parte Real (en línea gruesa) y la parte compleja (en línea delgada) del Wavelet. El parámetro ω_0 controla (resolución tiempo/escala) la frecuencia del exponencial complejo (2.56) que está modulado por una PDF Gaussiana. Si ω_0 aumenta, entonces el número de oscilaciones del Wavelet también aumenta. Tomada de Percival y Walden [263].

2.4 Análisis espectral vía la transformada continua de Wavelet

Funciones Wavelet escaladas

En la literatura sobre análisis de wavelet (veáse, por ejemplo, [173, 177, 207]), es común que las funciones Wavelet estén escaladas de la forma siguiente

$$\psi_{b,a}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \quad a > 0 \quad (2.58)$$

donde ψ es la función Wavelet, a es un parámetro de escala conocido como *voice* y b es un parámetro de localización temporal (traslación en el tiempo) [173, 177]. Nótese que, si $a > 1$ tiene un efecto de dilación en la relación 2.58 y si $a < 1$ tiene un efecto de contracción.

Para el caso cuando la función Wavelet es la de Morlet, también se expresa la relación 2.56 de la forma 2.58 [204, 350]. Este caso es equivalente a ajustar la resolución temporal, esto es, el ancho de la función Gaussiana con respecto a la frecuencia analizada ($\omega' = \omega_0/\omega$), de tal modo que $a = 2\pi/\omega$ (se consideran escalas en lugar de frecuencias) y $\omega_0 = 2\pi$ [350], esto es

$$\psi_{b,a}(t) = \pi^{-1/4}e^{-i\omega_0(\frac{t-b}{a})}e^{-(t-b)^2/2a^2} \quad (2.59)$$

2.4.2. La transformada continua de Wavelet

La transformada continua de Wavelet para una serie temporal o señal $f(t)$, en una localización temporal b y escala a , con respecto a una función Wavelet $\psi_{b,a}(t)$, puede ser definida [172] como una transformada integral, esto es

$$W_f(b, a) = \int_{-\infty}^{\infty} f(t)\psi_{b,a}^*(t)dt, \quad a > 0 \quad (2.60)$$

donde $\psi_{b,a}^*(t)$ es la compleja conjugada de $\psi_{b,a}(t)$ [172].

El nombre de continua deriva de que se asume que los parámetros a y b toman valores continuos. Es por esta razón que las funciones wavelet utilizadas en la CWT no forman una base ortogonal. Por tanto, los coeficientes de la transformada Wavelet en las escalas a y las localizaciones temporales b , contienen información redundante y están correlacionados [206].

Desde un punto de vista práctico, los coeficientes de $W_f(b, a)$ representan la contribución de las escalas (los valores a) a la señal en diferentes instantes

2.4 Análisis espectral vía la transformada continua de Wavelet

temporales (valores de b). Es posible considerar la transformada de Wavelet como una correlación cruzada de una señal $f(t)$ con un conjunto de funciones Wavelet de varias escalas a centradas en diferentes instantes b [50]. Adicionalmente, la CWT puede considerarse como la convolución de $f(t)$ con el conjunto de funciones wavelet $\psi_{b,a}$, *i. e.*, $W_f(b, a) = \langle f, \psi_{b,a} \rangle$ [350].

Una función $f(t)$ puede ser reconstruida, teóricamente hablando, desde su transformada de Wavelet por medio de una fórmula de inversión. La transformada inversa de Wavelet $M_{\psi_{b,a}}$ para la función $W_f(b, a)$ (en este caso la transformada de Wavelet) se puede expresar [50, 172] de la siguiente manera

$$f(t) = M_{\psi_{b,a}} W_f(b, a)(t) \quad (2.61)$$

$$= \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{a^2} W_f(b, a) \psi_{b,a}(t) da db \quad (2.62)$$

donde $C_\psi = \int_0^{\infty} \frac{|\Psi(f)|^2}{f} df$, de tal modo que satisface la condición 2.54, y $\Psi(f)$ es la transformada de Fourier de $\psi(t)$. La transformada inversa de Wavelet (relación 2.61) puede ser vista de dos formas diferentes. Como una forma de reconstrucción de $f(t)$, una vez que se ha calculado su transformada de Wavelet, o como una forma de expresar $f(t)$ como una superposición de funciones (base) Wavelet $\psi_{b,a}$, donde los coeficientes de esta superposición vienen dados por la transformada Wavelet de $f(t)$ [69].

La transformada de Wavelet cumple el teorema de Parseval [172, 263], preserva la energía o varianza de la señal, esto es

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} \frac{1}{a^2} |W_f(b, a)|^2 da db \quad (2.63)$$

Nótese que, el primer lado de la ecuación 2.63, es utilizado para definir la energía (varianza) en la señal (serie temporal) si la señal está centrada (media cero). En el segundo lado, el término $\frac{W_f^2(b, a)}{a^2}$ define una función de densidad espectral que descompone la energía (varianza) en diferentes escalas a centradas en diferentes posiciones temporales b [263].

Consideraciones prácticas para calcular la CWT.

En la práctica la CWT no se calcula directamente de la relación 2.60 (aunque es

2.4 Análisis espectral vía la transformada continua de Wavelet

posible por integración numérica) porque el tiempo de cómputo puede ser muy alto (el tiempo de integración en un intervalo $0 < a \leq I$, $0 < b \leq J$ es IJ^2 [219]). Por ello, una opción es sacar ventaja del teorema de la convolución, calculando la CWT en el espacio espectral (o de Fourier) [219, 350]. La versión discretizada de 2.60 para una serie temporal $f(t_n)$ de acuerdo con Torrence y Compo [350], es

$$\mathbb{W}_f^{\mathcal{F}}(a) = \sum_{n'=0}^{N-1} f(t_{n'}) \psi^* \left[\frac{(n' - n)\delta t}{a} \right] \quad (2.64)$$

donde ψ^* indica la compleja conjugada de ψ .

La transformada discreta de Fourier $\hat{f}(\omega_k)$ de la serie temporal $f(t_n)$ viene dada por

$$\hat{f}(\omega_k) = \frac{1}{N} \sum_{n=0}^{N-1} f(t_n) e^{-2\pi i k n / N} \quad (2.65)$$

donde el índice k correspondiente a las frecuencias angulares ω_k viene dado por $k = 0, \dots, N - 1$.

Teniendo en cuenta que la transformada de Fourier de una función del tipo $\psi_{b,a}(t/a)$ viene dada por $\hat{\psi}_{b,a}(a\omega)$ y, por el teorema de la convolución [350], la transformada de Wavelet es la transformada inversa de Fourier del producto de las transformadas de Fourier:

$$\mathbb{W}_f^{\mathcal{F}}(a) = \sum_{k=0}^{N-1} \hat{f}(\omega_k) \hat{\psi}^*(a\omega_k) e^{i\omega_k n \delta t} \quad (2.66)$$

donde N es el número de elementos de la serie temporal $f(t_n)$, δt es el intervalo temporal (constante para toda la serie temporal) y las frecuencias angulares ω_k vienen dadas [350] por

$$\omega_k = \begin{cases} \frac{2\pi k}{N\delta t} & : k \leq \frac{N}{2} \\ -\frac{2\pi k}{N\delta t} & : k > \frac{N}{2} \end{cases} \quad (2.67)$$

Utilizando la relación 2.66 y la transformada rápida de Fourier es posible calcular la transformada continua de Wavelet para un conjunto de escalas a y tiempos b de un modo eficiente (el tiempo de cómputo se reduce a $IJ \log_2 J$ [219]) [350].

Normalización de la función Wavelet

Torrence y Compo [350], sugieren normalizar la función Wavelet $\hat{\psi}_{b,a}(a\omega_k)$ para asegurar que la transformada de Wavelet (relación 2.66) sea directamente

2.4 Análisis espectral vía la transformada continua de Wavelet

comparable entre diferentes escalas a , y que también puedan ser comparables las transformadas Wavelet de diferentes series temporales. Entonces, la función Wavelet en cada escala a es normalizada de la siguiente manera para que tenga varianza o ‘energía’ unidad

$$\hat{\psi}^N(a\omega_k) = \left(\frac{2\pi a}{\delta t}\right)^{1/2} \hat{\psi}(a\omega_k) \quad (2.68)$$

Si se usa la relación 2.64, entonces la normalización de la función Wavelet es

$$\psi\left[\frac{(n' - n)\delta t}{a}\right] = \sqrt{\frac{\delta t}{a}} \psi_0\left[\frac{(n' - n)\delta t}{a}\right] \quad (2.69)$$

donde ψ_0 está normalizada [350].

Como una consecuencia de la normalización de la función Wavelet, se tienen las siguientes propiedades matemáticas.

$$\int_{-\infty}^{\infty} |\hat{\psi}(\omega')|^2 d\omega' = 1 \quad (2.70)$$

$$\sum_{k=0}^{N-1} |\hat{\psi}^N(a\omega_k)|^2 = N \quad (2.71)$$

donde N es el número de elementos de la serie temporal [350]

Otro resultado de interés como consecuencia de la normalización, es que la transformada de Wavelet sólo dependerá de los ‘pesos’ de las amplitudes de los coeficientes de la transformada de Fourier de $f(t)$ [350].

Sobre escalas, *octavas* y *voices*

Existen dos conceptos que están relacionados con el número de escalas a al calcular la CWT, *i. e.*, *octavas* y *voices*. La octava es el número de divisiones de la escala a y corresponde a potencias enteras de 2, mientras que el número de *voices*, es el número de divisiones de una octava (también se puede considerar como el número de potencias no enteras de dos octavas consecutivas) [263]. En este apartado se presenta la forma como se escogen las escalas a . Como bien apunta Farge [88] cuando se trabaja con Wavelets ortogonales, el conjunto discreto de

2.4 Análisis espectral vía la transformada continua de Wavelet

escalas a está limitado. Sin embargo para los Wavelets continuos no hay un límite para el conjunto de escalas y pueden ser escogidas de modo arbitrario [350]. Según Torrence y Compo [350] el conjunto de escalas viene dado por la siguiente relación

$$a_j = a_0 2^{j\delta_j}, \quad j = 0, 1, \dots, J \quad (2.72)$$

donde J determina el valor de la máxima escala¹, y viene dada por

$$J = \frac{\log_2\left(\frac{N\delta t}{a_0}\right)}{\delta_j} \quad (2.73)$$

donde δt es el espaciado temporal de los elementos de la serie temporal $f(t)$, a_0 es la escala de menor resolución y debería de ser escogida de tal modo que el equivalente al periodo de Fourier, sea *ca.* $2\delta t$. δ_j proporciona el espaciado entre las escalas a y su resolución dependerá de la función wavelet. Por ejemplo, para el Wavelet de Morlet es alrededor de 0.5 [350].

Relación entre las escalas de Wavelet y las frecuencias de Fourier

La relación entre la escala de Wavelet y el ‘periodo equivalente de Fourier’ (originalmente Meyers *et al.* [219] utilizó longitud de onda en lugar de frecuencias) puede, para algunas funciones Wavelet, derivarse analíticamente sustituyendo la función $e^{i2\pi f_0 t}$ de frecuencia conocida f_0 ya sea en la relación 2.60 o 2.64 y calcular la escala a en la cual el poder espectral de Wavelet tiene su máximo [219, 350]. Para el caso de nuestro interés, *i. e.*, para el Wavelet de Morlet y de acuerdo a Torrence y Compo [350] la relación entre la frecuencia f y la escala a viene dada por

$$\frac{1}{f} = \frac{4\pi a}{\omega_o + \sqrt{2 + \omega_o^2}} \quad (2.74)$$

A efectos prácticos se asume que $\omega_o = 2\pi$ [207, 350], por tanto, 2.74 puede expresarse como

$$\frac{1}{f} \approx a \quad (2.75)$$

¹Nótese que la máxima escala podría ‘compararse’, por analogía, con la frecuencia de Nyquist, en el sentido de que calcular la CWT más allá de J no aporta más información.

2.4 Análisis espectral vía la transformada continua de Wavelet

o de modo equivalente,

$$f \approx \frac{1}{a} \tag{2.76}$$

Este resultado simplifica enormemente la interpretación del análisis de Wavelet al utilizar la función de Morlet, porque es posible reemplazar escala por frecuencia [50].

El cono de influencia (COI)

La estimación del poder espectral por medio de la transformada continua de Wavelet de una serie temporal no periódica, puede generar errores (falsas estructuras espectrales), al principio y al final del espectro Wavelet [219]. La razón de estos artefactos se debe a que la relación 2.66 hace uso de la transformada de Fourier, la cual asume periodicidad en los datos bajo análisis [350]. Hay varias formas de manejar este problema. Por ejemplo, Meyers *et al.* [219] proponen aplicar una función de atenuación de tipo coseno a los datos. Otra opción (utilizada en esta memoria), es hacer un *padding* (rellenar con ceros) a cada lado de la serie temporal, después estimar el espectro Wavelet y por último, eliminar la parte del espectro afectada por los ceros de los segmentos añadidos [219, 350].

Uno de los inconvenientes al aplicar el cero *padding* es que produce discontinuidades en los bordes del espectro Wavelet. La región del espectro Wavelet en la cual los efectos de los bordes llegan a ser importantes es conocida como el cono de influencia (COI) [350]. Torrence y Compo [350] definen al COI como el *e-folding time* para la autocorrelación del espectro Wavelet en cada escala. El *e-folding time* (tiene un valor de $\sqrt{2}a$ para el Wavelet de Morlet) es seleccionado de modo que el espectro Wavelet para una discontinuidad en los bordes, decrezca por un factor de e^{-2} . Las porciones del espectro Wavelet entre el eje del tiempo y el COI deberían considerarse poco fiables [350].

2.4.3. La estimación del espectro

Por analogía al análisis espectral de Fourier (vía el periodograma) se puede definir [50] el espectro local de Wavelet como el módulo al cuadrado de la CWT (relación 2.64) de una serie temporal $f(t)$, esto es

2.4 Análisis espectral vía la transformada continua de Wavelet

$$S_f(b, a) = \|\mathbb{W}_f^{\mathcal{F}}(b, a)\|^2 \quad (2.77)$$

La relación 2.77, conocida como el escalograma [173], estima la distribución de la varianza de una serie temporal $f(t)$ entre las escalas a y diferentes localizaciones temporales b [50]. Sin embargo, y al igual que el periodograma, no es un estimador consistente del espectro [263]. Como una forma de obtener un estimador más robusto del espectro, Torrence y Compo [350] proponen suavizar el escalograma promediando sobre todos los espectros Wavelet locales, variando el parámetro de localización temporal b . Entonces, lo que se obtiene es el espectro Wavelet global, utilizando la nomenclatura de Cazelles *et al.* [50], esto es

$$\bar{S}_f(a) = \frac{\sigma_f^2}{T} \int_0^T \|S_f(b, a)\|^2 db \quad (2.78)$$

Este espectro Wavelet global puede ser comparado al clásico espectro de Fourier (una comparación entre el espectro de Fourier y el espectro de Wavelet puede encontrarse en [266]) y, Percival [264] y Torrence y Compo [350], sugieren que es un estimador consistente del espectro. Del mismo modo, es posible obtener una varianza promedio en cada localización temporal, promediando sobre un rango de escalas (banda, en la jerga del análisis de señales), esto es

$$\bar{S}_f(b) = \frac{\sigma_f^2 \pi^{1/4} \sqrt{b}}{C_\psi} \int_0^\infty \sqrt{a} \|S_f(b, a)\|^2 da \quad (2.79)$$

donde σ_f^2 es la varianza de la serie temporal $f(t)$ y C_ψ ha sido definido en la relación 2.62. Nótese como la relación 2.79 puede ser utilizada como un filtro paso banda (esta propiedad puede ser de ayuda para explorar posibles modulaciones de una frecuencia por otra en una serie temporal, o modulaciones de una serie temporal por otra serie [350]), teniendo en cuenta la relación 2.76, para una banda de frecuencia $f_1 - f_2$ [50, 350].

Como consecuencia de promediar, en escala (frecuencia) o en tiempo, se produce sesgo en el espectro Wavelet [207]. Para minimizar este sesgo, Maraun *et al.* [207] proponen promediar la misma cantidad de información independiente en

2.4 Análisis espectral vía la transformada continua de Wavelet

cada escala a , para ello escogen la longitud promedio del *reproducing kernel* de Morlet como una ventana de suavizado \mathcal{W} (Figura 2.5). Nótese la analogía del suavizado del espectro de Fourier por medio de ventanas. El proceso se resume a continuación y para más detalle [204, 207].

- El promediado en la dirección de las escalas debe realizarse con una ventana de suavizado de longitud constante \mathcal{W}_a para escalas logarítmicas (Figura 2.5). \mathcal{W}_a denota la anchura de la ventana mitad en la misma unidad como el número de voces por octavas (N_{voice}).
- El promediado en la dirección del tiempo se realiza utilizando una ventana que tenga longitud proporcional a la escala (Figura 2.5). Esto es, $a\mathcal{W}_b$ incrementa linealmente con la escala. \mathcal{W}_b denota la ventana de suavizado en la dirección del tiempo y $a\mathcal{W}_b$ denota la anchura de la ventana mitad en unidades de tiempo.

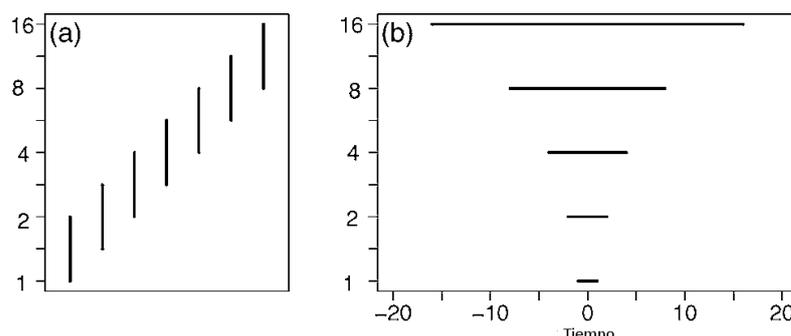


Figura 2.5: Esquema que muestra el proceso de suavizado utilizando el *reproducing kernel*. a) En la dirección de escala. b) En la dirección del tiempo. Tomada de Maraun *et al.* [207].

En resumen, las relaciones 2.78 y 2.79, utilizando el método de suavizado de Maraun *et al.* [207], pueden ser expresadas como

$$\bar{S}_f(b, a) = \mathcal{W}(\|S_f(b, a)\|^2) \quad (2.80)$$

donde \mathcal{W} es el operador de suavizado.

Sobre la normalización del espectro Wavelet

Existen diferentes tipos de normalización para el espectro Wavelet, pero quizá, la más conocida, por su utilidad en ciencias ambientales es la de Torrence y Compo [350], y en menor medida la de Kaiser [157]. Por un lado, la normalización ($\sqrt{\frac{\delta t}{a}}$) sugerida por los primeros, tiene la ventaja de que al calcular el espectro Wavelet para ruido blanco Gaussiano, se obtiene un espectro plano (como debería de esperarse). Sin embargo, si se calcula el espectro Wavelet para varias funciones sinusoidales de igual amplitud, el poder espectral difiere para cada periodo (escalas de oscilación) de los sinusoides (Figura 2.6). Por otro lado, la normalización de Kaiser ($\sqrt{\delta t}$) no permite que el espectro Wavelet de un ruido blanco Gaussiano tenga un comportamiento plano (de hecho tiene un comportamiento del tipo $\sim \frac{1}{a}$). Sin embargo, se conserva el poder espectral en los periodos (escalas de oscilación) correspondientes a los sinusoides (Figura 2.6) [206].

Esta situación, parece un problema serio, sobre todo cuando se sigue la normalización de Torrence y Compo, porque el poder espectral en las altas escalas está sobrestimado y podría llevar a errores de interpretación. Sin embargo, al aplicar un test de significación se consideran los efectos de la normalización [206].

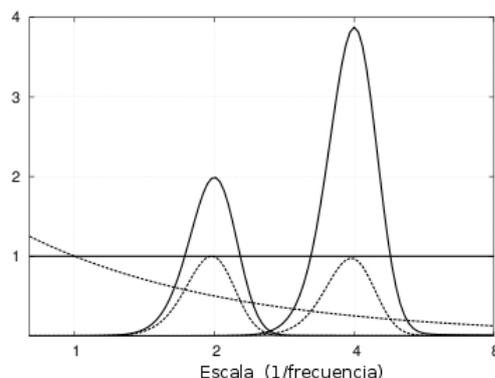


Figura 2.6: Espectro Wavelet global para ruido blanco Gaussiano y para dos sinusoides de igual amplitud y de frecuencias $1/2$ y $1/4$. En el espectro con la línea sólida se ha utilizado la normalización de Torrence y Compo [350]. En el espectro con la línea punteada utiliza la normalización de Kaiser [157]. Tomada de Maraun y Kurths [206].

2.4.4. La estimación del espectro cruzado

De igual manera que en el análisis espectral de Fourier, en el análisis espectral basado en Wavelets, es posible realizar análisis de Wavelet cruzado (CWA, *Cross Wavelet Analysis*). El espectro cruzado de Wavelet para dos series temporales $f(t)$ y $g(t)$ se define como el valor esperado (\mathbb{E}) del producto de los espectros de Wavelet $W_f(b, a)$ y $W_g(b, a)$ de las series [206], esto es

$$WCS_{fg}(b, a) = \mathbb{E}[W_f(b, a)W_g^*(b, a)] \quad (2.81)$$

donde $W_g^*(b, a)$ es la compleja conjugada de $W_g(b, a)$. Para obtener el espectro cruzado en el sentido de obtener las amplitudes espectrales cruzadas de $f(t)$ y $g(t)$ hay que calcular el módulo de $WCS_{fg}(b, a)$. Otra forma de expresar 2.81, descomponiendo en amplitud y fase [206], es

$$WCS_{fg}(b, a) = \|WCS_{fg}(b, a)\|e^{i\phi\{WCS_{fg}(b, a)\}} \quad (2.82)$$

donde la fase $\phi\{WCS_{fg}(b, a)\}$ describe el retraso entre las dos series temporales en la escala a y en torno al instante de tiempo b [206]. Es importante tener cuidado con el espectro cruzado de Wavelet, porque no existe un test de significación. Por tanto, es necesaria una normalización [206, 207].

La normalización del espectro cruzado es conocida como la *Wavelet coherence* (WCO), y viene dada por

$$WCO_{fg}(b, a) = \frac{\|\langle WCS_{fg}(b, a) \rangle\|}{\|\langle W_f(b, a) \rangle\|^{1/2}\|\langle W_g(b, a) \rangle\|^{1/2}} \quad (2.83)$$

donde $\langle \rangle$ denota obligatoriamente un operador de suavizado, tanto para escala como para tiempo [206]. Entonces, $WCO_{fg}(b, a)$ está definido entre 0 y 1. La WCO proporciona información sobre la variabilidad común para cada instante de tiempo y escala entre dos series temporales $f(t)$ y $g(t)$. Un valor de 1 indica una variabilidad igual entre $f(t)$ y $g(t)$ para una determinada escala a y tiempo

2.4 Análisis espectral vía la transformada continua de Wavelet

b . Un valor menor que uno, para una cierta escala a y tiempo b , puede indicar que la variabilidad de ambas series no es análoga. Un valor de cero indica que no hay correlación entre las series para una cierta escala a y en torno al instante de tiempo b [50, 206].

Para obtener información de un posible retraso o adelanto entre dos series temporales para una determinada escala a y tiempo b , se calcula el desfase entre f y g , esto es

$$\phi_{f,g}(b, a) = \tan^{-1} \frac{\Im(\langle WCS_{fg}(b, a) \rangle)}{\Re(\langle WCS_{fg}(b, a) \rangle)} \quad (2.84)$$

Los valores $\phi_{f,g}(b, a)$ estiman la diferencia de fase de dos series temporales para una escala a y tiempo b . Un valor positivo ($0 < \phi_{f,g}(b, a) < \pi/2$) indica un adelanto de $f(t)$ respecto de $g(t)$, mientras que un valor negativo ($-\pi/2 < \phi_{f,g}(b, a) < 0$) indica un retraso de $f(t)$ con respecto a $g(t)$. Un valor de $\phi_{f,g}(b, a) = 0$ implica que, para una determinada escala a y tiempo b , hay una función Wavelet que ajusta de la misma manera para ambas series temporales. Por otro lado, si $\phi_{f,g}(b, a) = \pm\pi/2$ implica que, para una determinada escala a y tiempo b , aparece el máximo retraso (-) o adelanto (+) de $f(t)$ con respecto a $g(t)$ [174]. Para calcular el tiempo de retraso o adelanto $t_{lag}(a)$ se puede utilizar [205] la siguiente relación

$$t_{lag}(a) = \frac{a\phi_{f,g}(b, a)}{2\pi} \quad (2.85)$$

Adicionalmente, se puede calcular la distribución de las fases para un rango de escalas a (o periodos). Una distribución uni-modal indica que hay un valor preferido de $\phi_{f,g}(b, a)$, por ello hay una tendencia estadística de que las dos series temporales $f(t)$ y $g(t)$ estén en fase. Por lo contrario, si la distribución es uniforme no es evidente una asociación entre la fase de $f(t)$ y $g(t)$ [50].

2.4.5. Pruebas de significación

Torrence y Compo [350] propusieron un test de significación para las estimaciones espectrales vía la CWT. Ellos utilizaron un espectro de fondo rojo (relación 2.41) para establecer una hipótesis nula H_o , la cual se define de la siguiente manera [350]. Se asume que la serie temporal a estudio tiene un espectro promedio dado por la relación 2.41. Si un punto escala-tiempo (a, b) del espectro Wavelet está por encima del espectro de fondo, entonces este punto puede ser considerado significativo con un cierto porcentaje de confianza (a este tipo de test se le conoce como *pointwise test*). Como modelo de ruido rojo (véase sección 2.2.5) Torrence y Compo [350] utilizan un proceso estocástico AR1 con tiempos equidistantes (relación 2.39 en su versión equiespaciada temporalmente). El *pointwise test* [207, 350] se resume en los siguientes pasos.

1. Se escoge un nivel de significación α (p. ej. del 0.95)
2. Se ajusta¹ un modelo AR1 (relación 2.39) a la serie temporal bajo estudio $f(t_i)$ equiespaciada temporalmente, con $i = 1, \dots, N$ (el número de elementos).
3. Se estima el percentil 0.95, $W_{q95}(b, a)$, del ensemble de k espectros Wavelet AR1 por medio de simulaciones de Monte-Carlo.
4. Se verifica para cada punto de $W_g(b, a)$ si superan el $W_{q95}(b, a)$. El conjunto de todos esos puntos escala-tiempo (a, b) del espectro Wavelet vienen dados [207] por

$$P_{pw} = \{(b, a) | W_g(b, a) > W_{q95}(b, a)\} \quad (2.86)$$

Los test de significación (tipo *pointwise*) traen asociado el problema de la multiplicidad [207]. Por ejemplo, para un nivel de significación de 0.95, un test aplicado para N espectros Wavelet $W_f(b, a)$ en una determinada escala a y tiempo b , implica un 5%N de falsos positivos. Una solución clásica al problema de la multiplicidad es la aplicación de las correcciones de Bonferroni o de Šidak. Sin

¹Torrence y Compo [350] en su metodología utilizan directamente la relación 2.41. También es posible utilizar técnicas de *bootstrap*, utilizando como modelo de réplicas un proceo AR1 (con las mismas propiedades estadísticas que la serie a estudio). El valor esperado del espectro Wavelet de N realizaciones AR1 converge a la relación 2.41.

2.4 Análisis espectral vía la transformada continua de Wavelet

embargo, para el caso de la estimación espectral vía Wavelet existe un problema adicional. Los coeficientes de la CWT, para diferentes escalas a y tiempos b , están correlacionados. Como consecuencia, los falsos positivos siempre ocurren como ‘parches’ contiguos (estos parches o áreas espectrales espurias reflejan oscilaciones aleatoriamente estables para un periodo corto de tiempo) [204, 207]. Para resolver este inconveniente, Maraun *et al.* [207] desarrollaron un test (*areawise test*) que utiliza información sobre el tamaño y la geometría de las áreas espectrales para discernir si es o no un área significativa. Sin embargo, no siempre es posible aplicar este test porque es necesario una cierta longitud de la serie temporal a estudio. Por ejemplo, en el caso de estudio de esta memoria (véase capítulo 6) donde se aplica la CWT, la serie temporal a estudio contiene pocos elementos (228 datos), por lo cual no es factible aplicar este test.

El test de significación (*pointwise*) para la WCO se describe a continuación [206, 207]. Considerando las longitudes de las series temporales a estudio, se ha utilizado el espectro correspondiente a un proceso AR1 como uno contra el cual la hipótesis nula (‘H0: los procesos no son coherentes’) es contrastada. El test de significación es calculado como sigue: se escoge un nivel de confianza $1 - \alpha$. Se calculan las coherencias (WCO) correspondientes a varios procesos AR1 con la misma autocorrelación de las series temporales a estudio y, por medio de simulaciones de Monte-Carlo, se estiman los valores críticos correspondientes al n -ésimo nivel de confianza ($1 - \alpha$). Cada punto del dominio tiempo-escala es verificado, comprobando si el espectro estimado de las series temporales excede los correspondientes valores críticos aleatorios.

Capítulo 3

Evaluación de una red de control de la calidad del aire (inmisión de SO_2) entre 1996-2001 en Bilbao metropolitano por medio del PCA

3.1. Introducción

La contaminación del aire es un problema presente en la mayoría de las grandes ciudades y es causada principalmente por la industria, los sistemas de calefacción y el tráfico rodado [209, 272]. Un punto en común que comparten estas fuentes de contaminación es que utilizan en muchos casos combustibles fósiles como fuentes de energía, de tal modo que los residuos de los procesos de combustión son emitidos a la atmósfera. Existen además otros procesos industriales, por ejemplo, la siderurgia, las industrias químicas o las refinerías de petróleo, que producen cantidades notables de contaminantes atmosféricos por medios distintos de la combustión de combustibles fósiles [209, 323, 381].

Los residuos de los procesos de combustión están formados principalmente por óxidos de azufre, nitrógeno y de carbono (SO_2 , NO , NO_2 , CO , CO_2) y partículas sólidas, los cuales son conocidos como contaminantes primarios. Estos contami-

nantas, una vez en la atmósfera, transcurrido un cierto tiempo, pueden precipitar (ya sea por deposición seca o húmeda) o reaccionar con otros compuestos presentes en la atmósfera con o sin ayuda de la radiación solar, de tal modo que se convierten en contaminantes secundarios, como el ácido sulfúrico (H_2SO_2), el ácido nítrico (HNO_3) o el ozono (O_3) [209, 369].

En lo que respecta a los compuestos de azufre emitidos a la atmósfera durante la combustión de combustibles fósiles, al menos un 90 % es SO_2 , el resto es en forma de sulfatos, y, en algunos casos, dependiendo de las condiciones de combustión, se puede formar SO_3 [3, 209]. Es importante mencionar que también existen fuentes naturales de SO_2 , como las erupciones volcánicas, los ‘sprays’ marinos, la descomposición biológica o los incendios forestales [381].

El SO_2 es un gas incoloro, no inflamable, denso (1400 kg/m^3), de fuerte olor asfijante y tiene varios efectos en el medio ambiente ya sea como contaminante primario o secundario [209]. Los primeros registros de contaminación atmosférica por SO_2 y otros contaminantes de origen antropogénico se remontan hasta la época de la Revolución Industrial. Sin embargo, no fue hasta la década de los 40 y 50 que se llegó a considerar a este contaminante y a otros como potenciales peligros para la salud humana, debido a su relación con varios incidentes mortales en diferentes lugares del mundo [209, 369].

Existe evidencia real de que la contaminación atmosférica por SO_2 y otros contaminantes, ya sean como primarios o secundarios, aparte de alterar las propiedades atmosféricas, afectan la salud humana, las plantas, los animales, los materiales, los procesos de producción, etc. [323, 369]. Por ejemplo, en lo que respecta a los efectos en las propiedades atmosféricas, los aerosoles resultantes de las reacciones fotoquímicas entre el SO_2 , el NO_x , otras partículas y los hidrocarburos presentes en la atmósfera, pueden formar un tipo de neblina que reduce la visibilidad al dispersar y absorber la radiación solar. Como una consecuencia directa de lo anterior, resulta alterada la cantidad de radiación solar que llega al suelo [344].

Los efectos del SO_2 en la salud humana han sido ampliamente estudiados desde hace ya más de 50 años. Uno de los últimos informes (2005) de la Organización Mundial de la Salud (OMS) [382], establece que una alta exposición al SO_2 (concentraciones de unos $500 \mu\text{g/m}^3$) a corto plazo (menor a un tiempo promedio de 10 minutos [382]) y en estudios controlados, puede ocasionar broncoconstricción y

otros problemas respiratorios [369, 382]. Los efectos del SO₂ a largo plazo (mayor a un día) en la salud humana son de difícil cuantificación porque los estudios epidemiológicos consideran también otros contaminantes presentes en el ambiente. Sin embargo, la OMS recomienda una exposición no mayor a 20 $\mu\text{g}/\text{m}^3$ en 24 horas [382].

El SO₂ no sólo puede afectar a la salud humana, sino también a los animales y a las plantas. Es capaz de producir en varias especies de animales broncoconstricción, y en las plantas producir lesiones, reducción de la fotosíntesis, cambios en la razón de respiración, etc. [272]. Otros efectos del SO₂ se pueden percibir en materiales metálicos, puesto que aceleran el proceso de corrosión al formar ácido sulfúrico ya sea en la atmósfera o en la superficie del metal (un ejemplo típico es el deterioro que sufren las esculturas metálicas expuestas al aire libre). También puede atacar una gran cantidad de materiales relacionados con la construcción, como la caliza, el mármol, etc. (es bien conocido este efecto en monumentos históricos) [272, 369].

Desafortunadamente, el ácido sulfúrico no se limita a dañar estructuras, porque en combinación con el ácido nítrico cuando llegan a la superficie mezclados con el agua de la lluvia, producen una disminución en el pH de la lluvia por debajo del nivel geoquímico normal de 5.0, causando una deposición ácida¹ [344, 389]. Los efectos de la deposición ácida en el ambiente son preocupantes, por un lado, porque produce acidificación de las fuentes naturales de agua, como lagos o ríos [217]. Se sabe que varios tipos de peces (como la trucha o el salmón) son especialmente sensibles a valores bajos de pH, incluso algunos tipos de peces detienen su ciclo reproductivo con valores de pH menores que 5.5. También se ha observado una disminución en el plancton, con lo que se reduce el suministro alimenticio de los peces. Por otro lado, la deposición ácida puede causar una lixiviación de los nutrientes del suelo, lo que produce una pérdida en la producción agrícola o de los bosques, incluso un cambio en la vegetación natural [369].

Durante la década de los 60 y 70, los problemas de contaminación atmosférica ocasionados por emisiones de SO₂ y por partículas suspendidas en el aire

¹Conocido como lluvia ácida, aunque no es un nombre adecuado del todo porque la deposición de estos ácidos no sólo se hace a través de la lluvia, sino también en forma de partículas o polvo (deposición seca) [389].

fueron ampliamente estudiados. Estos problemas de contaminación de origen antropogénico, principalmente por fuentes industriales y en menor medida por núcleos urbanos [209], llamaron la atención tanto de la comunidad científica como de la sociedad en general, lo cual condujo a la regulación de políticas medioambientales. Estas políticas motivaron el desarrollo de sensores para medir tanto las emisiones (*i. e.*, la concentración de contaminantes que salen de un foco emisor) como las inmisiones (*i. e.*, la concentración de contaminantes que están en el ambiente procedentes de una fuente emisora). Todo ello dió lugar a abundantes estudios sobre contaminación atmosférica en diferentes ciudades industrializadas alrededor del mundo [128].

Como consecuencia de la regulación legal para el control de contaminantes atmosféricos, los niveles de emisiones de SO₂ durante los últimos 25 años han ido disminuyendo en la mayoría de los países occidentales industrializados [237]. Por ejemplo, para la región geográfica a la que pertenece el caso de estudio de este capítulo, para el periodo de 1997 a 2007 y en comparación al nivel de 1990, hay una clara tendencia a la disminución [239]. Sin embargo, a pesar de esta tendencia a la baja, aún hay problemas relacionados con emisiones locales en combinación con factores meteorológicos y topográficos. En contraste, las emisiones de SO₂ están incrementando en países con economías emergentes [237].

3.1.1. Factores meteorológicos que actúan en la dispersión y en el transporte del SO₂

El ciclo de vida de un contaminante atmosférico urbano que es emitido por fuentes puntuales y distribuidas (Figura 3.1), como el SO₂, es un proceso complejo porque involucra varios factores y procesos físicos y químicos (incluso algunos no del todo bien entendidos, como la turbulencia). Aunque el ciclo de vida depende en gran medida del tipo de contaminante (gas o partícula), se puede resumir en las siguientes fases (no necesariamente en orden secuencial y varias de estas fases pueden ocurrir al mismo tiempo): *i) emisión, ii) dispersión y transporte, iii) transformaciones químicas y físicas y iv) deposición* [369].

Los factores más importantes que intervienen en cada fase, por ejemplo, para las emisiones, están relacionados con las características físicas de las fuentes

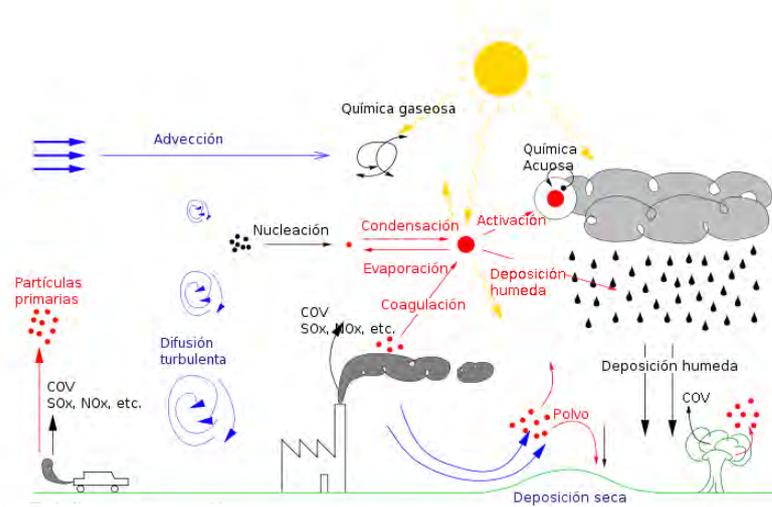


Figura 3.1: Esquema donde se muestra el ciclo de vida de los contaminantes atmosféricos (gases y partículas) y sus principales procesos físicos y químicos. Tomada y adaptada de Sportisse [339].

emisoras (chimeneas, tubos de escape, etc.) y las características aerodinámicas del contaminante, como la forma, peso o tamaño [209, 369]. En lo que respecta a la dispersión y transporte, los dos principales factores involucrados son las condiciones meteorológicas y topográficas locales [209, 369]. Los principales procesos químicos y físicos están relacionados con procesos de difusión y procesos termodinámicos (como la condensación, la evaporación, la solubilidad, etc.). La deposición (húmeda y seca) está condicionada por varios factores, donde, nuevamente, las condiciones meteorológicas juegan un importante papel [369].

Una buena parte de la contaminación del aire ocurre en la parte baja de la atmósfera, conocida como capa límite planetaria (PBL, *Planetary Boundary Layer*) [389]. La PBL es la parte de la atmósfera que experimenta efectos de la superficie en escalas de tiempo rápidas (horas o menos) mediante flujos de momento lineal, calor y humedad. El grosor de la PBL varía con un valor de referencia de 1 kilómetro, dependiendo de la geometría de la superficie (topografía), del tipo de vegetación, de la intensidad y la cizalladura del viento, etc. [340, 389]. Sobre

la tierra, la profundidad de la PBL cambia a lo largo del día en respuesta a los ciclos diurnos de calentamiento y enfriamiento. Por ejemplo, en un día soleado, el calor de la superficie de la tierra y la consecuente mezcla vertical, lleva a un incremento del grosor de la capa límite, alcanzando los máximos valores por la tarde. Pero por la noche el enfriamiento radiativo de la superficie inhibe la mezcla térmica, y por ello se reduce sustancialmente la profundidad de la PBL [256].

La intensidad de un evento contaminante, por ejemplo, las emisiones de SO_2 , está estrechamente relacionada con las condiciones meteorológicas en la parte baja de la PBL. En cambio, la dispersión de los contaminantes emitidos está controlada principalmente por dos procesos distintos: el transporte horizontal debido a los campos de viento y la mezcla vertical debida a la turbulencia atmosférica (la cual está directamente inducida por los efectos de la PBL) [340, 389]. Cuando se trabaja con procesos físicos relacionados con el transporte de contaminantes, es necesario considerar las escalas espacio-temporales. El transporte horizontal de contaminantes puede ocurrir desde unas horas a nivel local, de unos pocos días a nivel continental y de hasta un año a nivel global. El transporte vertical (que depende mucho de la estabilidad atmosférica) en la PBL puede ocurrir desde unas cuantas horas hasta un día [340]. El tiempo de residencia del SO_2 (incluido el ácido sulfúrico) es de unos 2 días [340].

En lo que respecta a las condiciones meteorológicas prevalecientes en el área de estudio de este trabajo, *i. e.*, la Comunidad Autónoma del País Vasco (CAPV) que afectan la dispersión y el transporte del SO_2 y otros contaminantes. La CAPV está bajo la influencia dominante del anticiclón de las Azores y de la variabilidad en el final del *storm-track* atlántico. Por un lado, en invierno, el anticiclón de las Azores se localiza sobre la costa de Portugal de tal modo que el Golfo de Vizcaya queda expuesto al paso de sucesivos frentes asociados a las borrascas atlánticas. Por estas razones, en invierno las condiciones meteorológicas favorecen la ventilación del País Vasco. Por otro lado, durante el verano, el anticiclón migra hacia latitudes septentrionales teniendo bajo su influencia el norte y la mitad occidental de la Península Ibérica. Durante el verano, el menor gradiente de presión sobre el área facilita el desarrollo de circulaciones de aire de origen térmico (brisas marinas). Cuando esta circulación es en el sentido océano-tierra, puede

arrastrar consigo contaminantes emitidos en latitudes más al norte cercanas a la costa hacia el interior de la península [5, 151].

En ocasiones, durante la presencia de un centro de altas presiones al norte (anticiclón británico o escandinavo) se puede registrar un transporte neto de las masas de aire contaminadas procedentes del Norte, Centro y Este de Europa, hacia la Península Ibérica. Los vientos del Sur suelen aparecer asociados a profundas borrascas atlánticas situadas a bajas latitudes, y pueden transportar masas de aire originarias del Norte de África hacia el Norte de la Península Ibérica [151].

3.1.2. Principales aproximaciones para estudiar la dispersión y el transporte del SO₂

La dispersión y el transporte de los contaminantes atmosféricos, tales como el SO₂, dependen de muchos factores como son las características físicas y químicas de los contaminantes, la meteorología local, la ubicación de las fuentes emisoras en relación con las obstrucciones al flujo del aire y la naturaleza del terreno (topografía) y de las estructuras (edificios, torres, etc.) construidas por el hombre que se encuentran en la dirección del viento que viene de las fuentes emisoras [369]. Para estudiar la dispersión y el transporte de los contaminantes atmosféricos existen dos tipos de modelos: los modelos físicos y los modelos matemáticos [209, 389].

Por un lado, los modelos físicos son utilizados para simular procesos atmosféricos por medio de una representación a pequeña escala del sistema a estudio, *v. gr.*, réplicas de un área urbana o una porción de ésta en un túnel de viento. Estos modelos han sido utilizados para entender ciertos procesos puntuales y en la validación de datos para los desarrolladores de modelos matemáticos [389], y es cierto que es una forma muy aproximada de reproducir la realidad, pero tienen problemas derivados de la imposibilidad de replicar adecuadamente todas las escalas de los movimientos atmosféricos. Por ello y por su precio, son de uso limitado [324]. Por otro lado, los modelos matemáticos están basados en un conjunto de algoritmos (numéricos o analíticos) que describen los principales aspectos físicos/químicos del problema a estudio. Sin embargo, requieren del uso de parametrizaciones para los procesos físicos/químicos que no están bien comprendidos o que no se pueden

representar con las escalas explícitamente resueltas por el modelo y el coste computacional es alto [209, 389]. Estos modelos se clasifican en dos grandes grupos: determinísticos y estadísticos [209, 389].

Los modelos determinísticos están basados en descripciones matemáticas de los procesos físicos y químicos de la atmósfera (como el transporte, la difusión, etc.) y su interrelación entre ellos, y constituyen el estado del arte de la modelización de la contaminación atmosférica [209]. Existe una amplia variedad de modelos matemáticos determinísticos que han sido desarrollados y utilizados en estudios de contaminación atmosférica, los cuales pueden ser agrupados principalmente en dos aproximaciones: la Euleriana y la Lagrangiana [324, 339, 389].

La aproximación Euleriana considera un sistema de coordenadas estacionario, *i. e.*, que se mantiene fijo en el espacio, donde las especies de contaminantes entran y salen de las celdas a través de sus paredes y el modelo simula la concentración de las especies en todas las localizaciones como una función del tiempo. Por el contrario, la aproximación Lagrangiana considera un sistema de coordenadas no-estacionario, *i. e.*, que se mueve con el viento local, por tanto no hay un intercambio de masa entre la parcela de aire y los alrededores a excepción de las emisiones que se permiten entrar a la parcela a través de su base. La parcela de aire se mueve continuamente, así que el modelo simula las concentraciones en diferentes localizaciones y tiempos [139, 324].

El modelo determinístico más sencillo, dimensionalmente hablando (modelo 0-dimensional o modelo de caja), es el modelo donde el dominio atmosférico está representado por una única caja en la que la concentración de contaminante a estudio es constante en cualquier parte de la caja. Por ello, la concentración del contaminante i -ésimo es únicamente una función del tiempo, $C_i(t)$ [58, 324, 389]. Debido a que un modelo de transporte de una caja ha sido el modelo conceptual utilizado para interpretar los resultados de este capítulo (sección 3.4), se describen a continuación algunas de sus principales características.

El principio en el que se basa un modelo de transporte de una caja, consiste en identificar un área (generalmente rectangular) de la superficie terrestre como la cara inferior de un cuboide que se extiende hasta la altura de la capa límite atmosférica $H(t)$. Estos modelos asumen que la mezcla de los contaminantes dentro del cuboide es uniforme y se realiza de forma inmediata con el aire disponible

para la dilución. Estos modelos están basados en el principio de la conservación de la masa para una especie de contaminante i que está dentro de una caja de volumen $\Delta x \Delta y H$, cuya concentración media (kg/m^3) es C_i [324, 338].

La definición matemática de este modelo, siguiendo la nomenclatura de Seinfeld y Pandis [324] es:

$$\frac{d}{dt}(C_i \Delta x \Delta y H) = Q_i + R_i \Delta x \Delta y H - S_i + u H \Delta y (C_i^0 - C_i) \quad (3.1)$$

donde Q_i es la razón de emisión del i -ésimo contaminante ($kg\ h^{-1}$), S_i es la razón de eliminación del i -ésimo contaminante ($kg\ h^{-1}$), R_i es la razón neta de las transformaciones químicas del i -ésimo contaminante ($kg\ m^{-3}\ h^{-1}$), C_i^0 es la concentración (de fondo) a sotavento y u es la velocidad del viento (se asume que la dirección del viento es constante y perpendicular al lado YH de la caja sin pérdida de generalidad).

Los modelos estadísticos están basados en relaciones (semi)empíricas entre valores observados de ciertas variables meteorológicas y concentraciones de contaminantes registrados en algún receptor durante algún tiempo, son relativamente simples, tienen bajo coste computacional, y aunque han mostrado buenos resultados en muchos casos, es importante tener en cuenta que no están basados en leyes o principios físicos fundamentales [209, 388].

3.1.3. Técnicas estadísticas utilizadas para la evaluación de redes de control de la calidad del aire

El estudio y seguimiento del impacto de la calidad del aire debido a inmisiones de SO_2 y otros contaminantes puede ser realizado por medio de las redes de control y vigilancia de la calidad del aire (AQMN, *Air Quality Monitoring Network*). Estas redes requieren una constante evaluación para averiguar y garantizar que cada uno de los nodos que integran la red proporcione una caracterización adecuada de la calidad del aire en la zona donde muestrea cada sensor [272]. Los objetivos principales de esta evaluación son los siguientes: determinar el número adecuado de sitios de muestreo (nodos) de la red de tal modo que no se obtenga información redundante, detectar fallos (*fault detection*) en alguno de los nodos

y detectar una inadecuada localización espacial (*fault localisation*) de los nodos [176, 271].

Una de las herramientas que pueden ser de ayuda para evaluar las AQMN son las técnicas estadísticas multivariantes, como el análisis mediante componentes principales (PCA) u otras técnicas matemáticas de clasificación como el análisis de cluster (CA). Estas técnicas han sido utilizadas previamente para la evaluación y el manejo adecuado de redes de seguimiento de la calidad del agua [329, 332, 384]. El uso combinado de ambas técnicas (PCA y CA) en la evaluación de una AQMN ha sido recientemente llevado a cabo en el año 2006 [111] para determinar la tendencia estacional y la distribución espacial de PM_{10} y O_3 . Posteriormente, en 2008, [272] aplicaron el PCA y el CA para valores de SO_2 y PM_{10} para una AQMN que cubre el área metropolitana de Oporto, encontrando que de los 10 nodos de la red 6 son suficientes para caracterizar las concentraciones de SO_2 .

Con lo que respecta al área de estudio de este trabajo y a las inmisiones de SO_2 , en 2009, Ibarra-Berastegi *et al.* [149] aplicaron el PCA, el CA y también mapas auto-organizativos de Kohonen (SOM) para evaluar una AQMN de 4 nodos para el área metropolitana de Bilbao, y encontraron que todos los nodos de la red son necesarios para una correcta caracterización regional de las concentraciones mensuales de SO_2 . Los resultados que se presentan en este capítulo proceden de ese trabajo.

3.2. Area de estudio y antecedentes

La Comunidad Autónoma del País Vasco (CAPV) está localizada en la parte centro norte de la Península Ibérica, frente al mar Cantábrico y entre dos cadenas montañosas, los Pirineos al Este y la cordillera Cantábrica al Oeste. Un complejo de pequeñas montañas y valles con diferentes orientaciones conforman la transición entre la zona costera al norte y los dos principales accidentes topográficos de la Península, la meseta Ibérica y el valle del Ebro [101].

El área de estudio de este trabajo es Bilbao metropolitano, asentada a lo largo del río Nervión, a una distancia aproximada de 16 km desde el centro de la ciudad hasta la costa Cantábrica y en dirección Sureste-Noroeste (Figura 3.2) y tiene una población aproximada de un millón de personas. Bilbao metropolitano pertenece a

3.2 Area de estudio y antecedentes

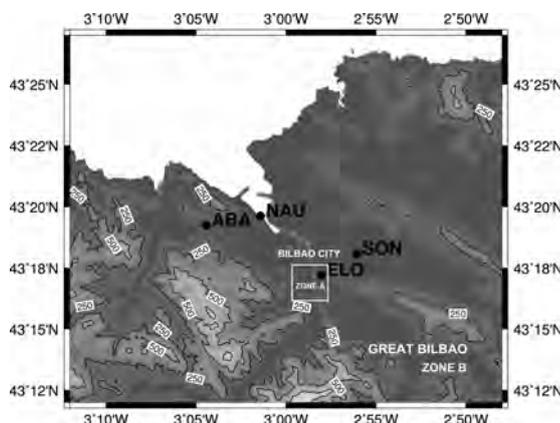


Figura 3.2: Localización de los sensores: Abanto (ABA), Náutica (NAU), Sondika (SON) y Elorrieta (ELO) del área metropolitana de Bilbao. Tomada de Ibarra-Berastegi *et al.* [149]).

una de las principales regiones con fuentes de contaminantes atmosféricos (debido principalmente a la industria y al tráfico rodado) no sólo de la CAPV sino a nivel nacional [101, 147].

Al igual que otras ciudades industrializadas de finales de la década de los setenta y principios de los ochenta, Bilbao metropolitano presentó problemas de contaminación atmosférica de origen industrial. Por esta razón se instaló una red de control y vigilancia de la calidad del aire, la primera a nivel nacional y una de las primeras en Europa [1, 46]. Uno de los primeros objetivos fue dar seguimiento a emisiones de origen industrial de SO_2 . No obstante, en los últimos años los niveles de SO_2 han ido disminuyendo debido a los cambios en el tejido industrial y a los combustibles actualmente utilizados. Sin embargo, la actual legislación ambiental obliga a seguir realizando mediciones de inmisiones de SO_2 [149]. Hoy en día, las estaciones remotas de la red de vigilancia de la calidad del aire que cubren el área urbana y la zona metropolitana de Bilbao (Figura 3.2) forman parte de la red del bajo Nervión con un total de 21 estaciones [1].

Desde la instalación de la AQMN de Bilbao metropolitano hasta la fecha, la estructura de la red ha sufrido modificaciones importantes (por ejemplo, en la localización de las estaciones o en el aumento de éstas) con el objeto de poder captar las campos representativos de las inmisiones de los diversos contaminantes,

3.2 Area de estudio y antecedentes

debido a los cambios de localización o al aumento o disminución de las fuentes de emisiones tanto de SO_2 como de otros contaminantes [5]. Este tipo de cambios no se limitan a la AQMN de Bilbao metropolitano, sino que prácticamente están presentes en la mayoría de las redes de control y vigilancia de la calidad del aire de otras áreas geográficas. Es por estas razones que, tanto el proceso de diseño como el control y la evaluación de una AQMN es un proceso dinámico e interactivo [149]. Por tanto, es importante mantener una constante evaluación de las AQMN para poder captar las trayectorias representativas de las especies de contaminantes, tales como el SO_2 [176, 271, 272].

Un estudio previo [147] mostró que los seis sensores instalados en el área (aproximadamente $3 \times 3 \text{ km}^2$) urbana de Bilbao, *i. e.*, la zona A (Figura 3.2, localización de los nodos no mostrada), miden el mismo patrón de la concentración de contaminantes (entre éstos el SO_2) de ese área, por tanto, se está midiendo información redundante. La explicación física es, por un lado, porque las fuentes emisoras (sistemas domésticos de calefacción y tráfico rodado), al menos para el SO_2 , son casi homogéneas en la zona A. Por otro lado, porque los mecanismos de emisión, dispersión e inmisión de contaminantes dentro de la zona A pueden ser descritos en primera aproximación utilizando un modelo conceptual de una caja, debido a que es posible considerar que el SO_2 se mezcla de modo casi uniforme en la capa límite urbana [149, 324].

En base a los resultados obtenidos por Ibarra-Berastegi *et al.* [147] para el área A, se puede establecer la hipótesis de que en el área B (Figura 3.2) las 4 estaciones podrían también estar midiendo información redundante. Si esto fuera así, incluso quizá podría describirse la concentración de contaminantes en el área B por medio de un modelo de una caja. Sin embargo, es necesario una previa evaluación de esta hipótesis a partir de los datos registrados por la AQMN. Para lograr este objetivo, se recurre al análisis de componentes principales, una técnica estadística multivariante utilizada recientemente con éxito para estudios con objetivos similares a los de este trabajo [176, 272]. Adicionalmente, los resultados obtenidos mediante PCA se comparan con los de otras técnicas estadísticas, como el CA y los SOM [149]. No obstante, los resultados del CA y los SOM no han sido realizados por el autor de esta memoria, por lo que la presentación de resultados se centra especialmente en el caso del PCA.

3.3. Datos y métodos

En este estudio se han usado 68 promedios mensuales de niveles de inmisión de SO_2 ($\mu\text{g m}^{-3}$) (Figura 3.3) para el periodo 1996-2001. Los datos provienen de 4 estaciones automáticas (Abanto, Náutica, Elorrieta y Sondika) de la red de control y vigilancia de la calidad del aire de la CAPV, las cuales están bajo supervisión del Departamento de Medio Ambiente, Planificación Territorial, Agricultura y Pesca, perteneciente al Gobierno Vasco. Se han utilizado promedios mensuales, a pesar de tener datos diarios, porque para los objetivos de este estudio, la escala mensual es suficiente y para evitar el impacto de ciclos de alta frecuencia que podrían crear un efecto de enmascaramiento [149]. Las series temporales prácticamente no tienen datos perdidos, a excepción del año 1996, en el cual faltan 4 datos mensuales. Los datos perdidos se sustituyen por la media de los valores SO_2 correspondientes a la columna (estación) que pertenecen. Las series han sido estandarizadas antes de aplicar el análisis por componentes principales [149].

El PCA es una técnica estadística multivariante que transforma linealmente un conjunto de variables correlacionadas en términos de funciones de base ortogonal, de tal modo que reducen el número de variables originales y eliminan la correlación entre ellas [156, 365, 376]. El PCA es ampliamente utilizado en estudios ambientales, porque permite reducir las dimensiones de los conjuntos de datos a estudio, y por tanto, identificar las características principales del fenómeno de interés [149]. En lo que respecta a la contaminación del aire por emisiones de SO_2 , ha demostrado ser una herramienta adecuada para detectar la variación espacial de campos de SO_2 [196, 286, 307], y, más recientemente, para la evaluación de redes de control de calidad del aire. De tal forma que ha permitido detectar si alguno de los nodos de estas redes miden información redundante o detectar fallos de algunos de los nodos [122, 149, 176, 271, 272].

Una descripción a detalle de los (aspectos metodológicos) pasos implicados en el análisis por componentes principales se encuentra en la sección 2.1 del capítulo 2. Para el análisis PCA se utilizó el paquete computacional `FactoMineR` [178] en R [280]. En nuestro caso, la matriz de datos de anomalías $X_{I \times J}$ del capítulo 2, sección 2.1, consiste en $I = 68$ observaciones mensuales en $J = 4$ estaciones. El número de componentes principales (CPs) retenidas es 4 (lo cual corresponde al

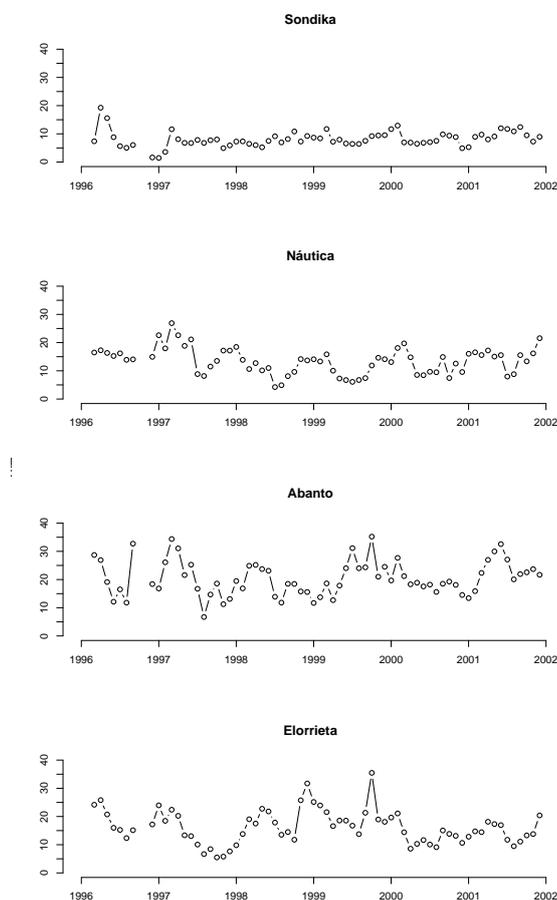


Figura 3.3: Series temporales de valores mensuales de inmisiones de SO_2 para las cuatro estaciones. Sondika, Náutica, Abanto y Elorrieta. Las unidades de las emisiones de SO_2 están en $\mu\text{g m}^{-3}$.

número máximo de CPs para la matriz de anomalías $X_{I \times J}$). Para ello, se utilizó el criterio PVA_{90} propuesto por Pires *et al.* [271, 272] descrito en la subsección 2.1.3 de esta memoria.

La estrategia utilizada para averiguar si las 4 estaciones de la red de control y vigilancia del área metropolitana de Bilbao (área B) son redundantes es la siguiente [149].

1. Identificar qué factores explican la mayor parte de la variabilidad de los

datos mensuales de SO₂ para el área metropolitana de Bilbao.

2. Analizar los factores de carga en cada sensor para detectar cuáles tienen un comportamiento similar. Para hacer un análisis visual, se representan gráficamente los 4 sensores en el espacio de factores de los CPs, donde los factores de carga representan las coordenadas correspondientes a los 4 sensores. Los sensores con tendencia a agruparse en el espacio de factores, se espera que tengan similar comportamiento, mientras que los de tendencia a alejarse podrían estar bajo diferentes regímenes de emisión o dispersión. Por tanto, podrían ser determinadas redundancias en la red de control y vigilancia de la calidad del aire de la zona de estudio.
3. Por medio del análisis de los resultados de las PCA y de los conocimientos *a priori* de las fuentes de emisiones de SO₂ y de las condiciones meteorológicas del área de estudio, se intenta identificar los mecanismos de emisión o dispersión que actúan sobre las inmisiones de SO₂ a escala mensual para el área metropolitana de Bilbao.

3.4. Resultados

Los resultados al aplicar el PCA se presentan en la Tabla 3.1. La sección (a) de esta Tabla muestra los valores propios, los porcentajes de varianza explicados por cada componente principal y el porcentaje de varianza acumulado para las 4 componentes principales. La sección (b) contiene los factores de carga, para la primera CP, tienen el mismo signo y sus valores no difieren mucho entre sí (nótese que ninguna de estas dos características se cumplen para las otras componentes principales). Esto indica que la primera CP representa una variabilidad común de los valores de inmisiones de SO₂ para la zona B (Figura 3.2). De la sección (c) se puede observar que la influencia de la primera CP varía de un sensor a otro, con un rango de 28.19-59.78 % para Náutica y Abanto, respectivamente. Por último, de la sección (d), se establece que la contribución de cada sensor para la primera CP está en un rango de 16.36-34.68 % para Náutica y Abanto, respectivamente [149].

La primera CP se puede interpretar físicamente como el resultado de la mezcla de emisiones que atraviesan la zona B (Figura 3.2). Es decir, representa la contribución de emisiones de SO₂ de diferentes fuentes que se mezclan y se distribuyen uniformemente en la zona B, medidas por los cuatro sensores instalados en esta zona. La segunda CP se puede interpretar físicamente a partir de patrones locales de emisiones o dispersiones de SO₂, los cuales afectan sobre todo a Sondika y Náutica (Tabla 3.1 b) pero de modo opuesto, pues sus factores de carga tienen signos opuestos. Para la tercera CP, la interpretación física no es evidente. Finalmente, la cuarta CP puede asociarse a un efecto local, pues se conoce *a priori* que existen fuentes de emisiones de SO₂ en el área de Abanto (muy cerca hay un polígono petroquímico de grandes dimensiones) [149].

Los factores de carga (Tabla 3.1 b) indican el comportamiento de los sensores, y a excepción de la primera CP, no tienen un comportamiento similar. En la Figura 3.4 se muestra una representación gráfica de los cuatro sensores para las tres primeras CP (84.86 % de la variabilidad total), se puede notar que estos sensores no tienen un comportamiento físico similar más allá del 43.1 % de variabilidad asociada a la primera CP [149]. Estos resultados concuerdan con los obtenidos por Ibarra-Berastegi *et al.* [149] utilizando el análisis de cluster y los mapas auto-organizativos de Kohonen.

3.4 Resultados

Tabla 3.1: Resumen del análisis de componentes principales. (a) Valores propios y porcentaje de la varianza total explicada por cada CP, (b) factores de carga, (c) contribución (%) de cada CP a la variabilidad total de cada sensor, (d) contribución (%) de cada sensor a la variabilidad total dentro de cada CP. Tomado de Ibarra-Berastegi *et al.* [149].

a)

Componente principal	Val. propios	% de var.	% de var. acum
CP-1	1.72	43.09	43.09
CP-2	0.93	23.20	66.29
CP-3	0.74	18.57	84.86
CP-4	0.61	15.14	100.00

b)

Sensor	CP-1	CP-2	CP-3	CP-4
Sondika	0.56	-0.65	0.51	0.11
Náutica	0.53	0.71	0.44	0.14
Abanto	0.77	0.02	-0.19	-0.60
Elorrieta	0.73	-0.04	-0.51	0.46

c)

Sensor	CP-1	CP-2	CP-3	CP-4	Σ
Sondika	30.95	41.96	25.92	1.18	100
Náutica	28.19	50.65	19.25	1.91	100
Abanto	59.78	0.02	3.61	36.59	100
Elorrieta	53.46	0.16	25.50	20.87	100
$\Sigma/100$	1.72	0.93	0.74	0.61	

d)

Sensor	CP-1	CP-2	CP-3	CP-4
Sondika	17.95	45.22	34.89	1.94
Náutica	16.36	54.58	25.92	3.15
Abanto	34.68	0.03	4.86	60.43
Elorrieta	31.01	0.18	34.33	34.48
Σ	100	100	100	100

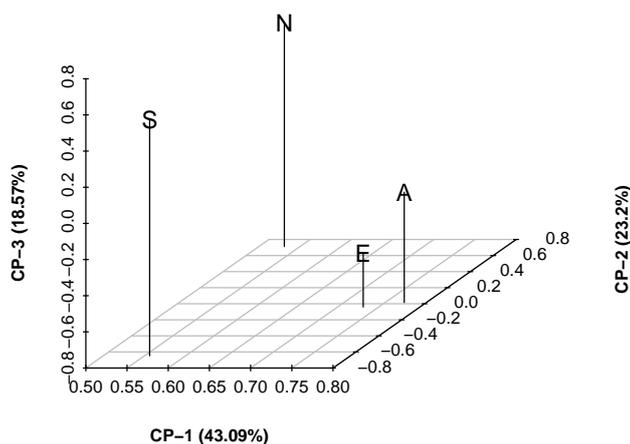


Figura 3.4: Representación gráfica para los cuatro sensores en el espacio de las CP. Abanto (A), Náutica (N), Sondika (S) y Elorrieta (E). Adaptada de Ibarra-Berastegi *et al.* [149].

3.5. Discusión

Los tres sensores que no pertenecen a la zona A están bajo la influencia de diferentes patrones de emisión/dispersión, aunque los factores de carga de las primeras componentes principales (Tabla 3.1 b) sugieren que sólo para una cierta fracción (43.1% de la varianza total) existe un mecanismo del tipo modelo de una única caja actuando para toda la zona B (Figura 3.2). Esto concuerda con el hecho de que el sensor instalado dentro de la zona A, *i. e.*, en Elorrieta (dentro de la ciudad de Bilbao) está influenciado por emisiones de pequeñas fuentes de SO_2 (de tipo urbano), mientras que los sensores instalados en Abanto y Sondika están bajo influencia de otros tipos de fuentes de SO_2 (influencia industrial y en menor medida por el tráfico rodado). La orografía también tiene un papel importante sobre los sensores, *v. gr.*, el sensor instalado en Náutica, el más cercano al mar

(Figura 3.2), tiene mecanismos diferentes de dispersión de los que están en la zona A. Esto significa que emisiones de SO_2 de todas las fuentes en la zona B, no se mezclan perfectamente y no se dispersan con los mismos mecanismos físicos [149].

Es importante tener en cuenta que los resultados aquí obtenidos vía PCA concuerdan con los resultados obtenidos por medio del análisis de cluster [149], en el sentido de que los 4 sensores de la AQMN de la zona B presentan diferentes comportamientos. Del mismo modo, los resultados de las PCA coinciden con los resultados de los SOM, en el sentido de que existe una importante variabilidad en los valores de SO_2 de la zona B, y que toda la información proporcionada por los 4 sensores es necesaria para representar la distribución espacial de este contaminante [149]. Sin embargo, a pesar de que los resultados del CA y los SOM son de fácil interpretación, ambas técnicas no proporcionan una ayuda para la identificación de los posibles mecanismos físicos de emisión o dispersión que actúan sobre las inmisiones de SO_2 .

Capítulo 4

Búsqueda de señal orbital en isótopos estables de belemnites del Jurásico tardío de la cuenca Vasco-Cantábrica mediante el PLS

4.1. Introducción

El Jurásico tardío (Plienbachiense y Toarciense inferior), datado entre los 182-190 Ma (millones de años hacia atrás en el tiempo), fue una época de grandes cambios y crisis ambientales, durante la cual se depositó material orgánico como resultado de varios eventos oceánicos anóxicos (*Oceanic Anoxic Event*, OAE) [12, 153]. Estos cambios ambientales han sido registrados en diversos materiales geológicos provenientes de varios lugares alrededor del mundo [80, 135, 154, 162, 297]. Muchos estudios han sido realizados para saber la causa de estos cambios, y si los mismos se produjeron a escala global. Sin embargo, a pesar de estos esfuerzos, no existe una sola respuesta para explicarlos [12, 346, 353].

Una forma de estudiar los cambios ambientales del pasado, consiste en analizar las propiedades geoquímicas de isótopos estables de material fosilizado, como las rostra de belemnites [109, 213, 297, 305]. Esta especie extinta que vivió en el

Jurásico y Cretácico, ha sido considerada como uno de los grupos de fósiles más adecuados para estimaciones de paleotemperaturas y paleosalinidades del océano de esa época [109, 155, 212, 213, 274].

Rosales *et al.* [296, 297, 299] propusieron que valores del $\delta^{18}\text{O}$ provenientes de las rostra de belemnites (sin efectos diagénicos) de la cuenca Vasco-Cantábrica (*Basque-Cantabrian Basin*, BCB), podrían utilizarse como estimaciones de paleotemperaturas y paleosalinidades para el Jurásico tardío para esta área geográfica. Sin embargo, estos mismos autores probaron que, al menos para este caso, los carbonatos hemipelágicos contenidos en las rocas de alto contenido orgánico de la sucesión estratigráfica donde se recolectaron las muestras de belemnites, aunque ampliamente utilizados para estudios paleoceanográficos, no retienen la composición geoquímica marina después de la diagénesis burial. Por tanto, no es un material adecuado para estudios paleoceanográficos [296].

En un estudio previo realizado en contenido de CaCO_3 de las facies hemipelágicas del Toarciense inferior (biozona *renuicostatum* con una duración temporal de 300 Ka) de la cuenca Vasco-Cantábrica, Berreteaga *et al.* [25] encontraron señal orbital de corto periodo (ciclo de precesión de los equinoccios, *ca.* 20 Ka). Sin embargo, no encontraron estas variaciones cíclicas en el contenido total del Carbono orgánico y en $\delta^{13}\text{C}$ de la roca, aunque sugieren ampliar el intervalo de estudio para confirmar estos resultados. Por otro lado, resultados de otros autores donde se utilizan diferentes paleoindicadores (principalmente contenido de CaCO_3) del Plienbachense y Toarciense, para diferentes zonas geográficas relativamente cercanas a la cuenca Vasco-Cantábrica, confirman la existencia de ciclos de precesión (*ca.* 20 Ka), oblicuidad (*ca.* 40 Ka) y excentricidad (*ca.* 100 Ka) [137, 346, 372, 373].

Una herramienta adecuada para la búsqueda de ciclos orbitales en series temporales paleoclimáticas es el análisis espectral [317, 319, 371]. Esta búsqueda es relativamente trivial cuando las series provienen de muestras paleoclimáticas que están equiespaciadas temporalmente y con alta resolución espacio-temporal de muestreo (como las series de CaCO_2) [371]. Sin embargo, esta búsqueda puede ser complicada cuando las series provienen de material marino calcáreo fosilizado (como las rostra de belemnites), porque existe un cierto grado de aleatoriedad en la localización espacio-temporal de las muestras [232, 371]. Por esta razón,

las series temporales creadas con estos materiales no suelen estar equiespaciadas temporalmente, incluso, pueden presentar grandes diferencias espacio-temporales entre las muestras [371].

Una posible solución para enfrentar el carácter no equiespaciado de las series temporales paleoclimáticas, es interpolar en el tiempo, y así obtener una serie temporal equiespaciada. Sin embargo, la interpolación implica un conocimiento previo del comportamiento de la variable a estudio, y suaviza los datos de tal modo que al calcular el espectro puede suprimir información espectral en las altas frecuencias [232, 316, 317, 371]. Otra solución consiste en utilizar métodos de análisis espectral que puedan aplicarse directamente a las series temporales paleoclimáticas, tales como el periodograma de Lomb Scargle [193, 311, 312, 316, 317], el método CLEAN [134, 289], o más recientemente, los métodos de Mathias *et al.* [211] o Nava [232].

No obstante, en algunos casos, como en el de las series temporales de los valores isotópicos utilizadas en este estudio, las diferencias entre los tiempos de una serie temporal son mucho más grandes que la escala de las mayores oscilaciones de la señal que se espera encontrar. Entonces, ni la interpolación ni los métodos especiales de análisis espectral pueden ser de mucha ayuda [371]. Una posible solución, consiste en partir la serie original en trozos más cortos, utilizando como punto de corte los tiempos que superan las mayores oscilaciones esperadas [319, 371].

Otro factor que puede complicar el análisis espectral de series temporales cortas y no equiespaciadas temporalmente (este problema no es exclusivo de las series temporales paleoclimáticas), como las series de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ de los belemnites, es la presencia de ruido [140, 354]. Se sabe que los paleoindicadores ambientales provenientes de belemnites (*v. gr.*, $\delta^{18}\text{O}$, $\delta^{13}\text{C}$) contienen ruido [274]. Este ruido puede representar cambios en los procesos ambientales ocurridos en el pasado (frecuentemente de corto periodo) en especies de belemnites adaptables a estos cambios (*i. e.*, especies que no hayan migrado como una respuesta a estos cambios). Sin embargo, también pueden reflejar el rango espacial, vertical (profundidad oceánica) y lateral (geográfico) del hábitat de los belemnites, incluso pueden reflejar los efectos de las especies, del género o la edad biológica del belemnite [213, 274].

Un modelo de ruido de fondo comúnmente utilizado en estudios cicloestratigráficos, es un modelo autoregresivo de orden 1 (AR1), también conocido como ‘ruido rojo’ [319, 371]. Una de las justificaciones para el uso del modelo AR1, es que el espectro de muchas series paleoclimáticas es rojo. *i. e.*, el espectro muestra un continuo decrecimiento de la amplitud espectral con el incremento de la frecuencia. Como demostró Hasselmann [124], un modelo AR1 es suficiente para explicar esta característica espectral. Sin embargo, siempre es recomendable verificar si el modelo de ruido es adecuado. Para tal fin, REDFIT utiliza el *runs test* [20, 316].

Un par de razones justifican, en este estudio, el uso del análisis espectral aplicado a las series temporales del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ de los belemnites de la cuenca Vasco-Cantábrica [297, 298, 299]. En primer lugar, porque se pretende averiguar si estos isótopos estables pudieron haber registrado algún tipo de evento periódico relacionado con los ciclos orbitales [23, 175], a pesar de la existencia de una componente aleatoria en la localización espacio-temporal de las muestras, de los errores de datación de las muestras, de la limitación del número de elementos de las series temporales y del ruido de fondo. Especial énfasis se hace en la búsqueda del ciclo de excentricidad de largo periodo de los 400 Ka [23, 63], porque se sabe que ha permanecido invariable desde el Mesozoico hasta el presente [136, 175] y porque ha sido identificado en registros estratigráficos del Mesozoico recolectados en diferentes áreas geográficas [90, 245, 246, 247, 370]. La otra razón es llevar a cabo una comparación de las características espectrales de las series temporales de los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$. Sin embargo, es importante tener en cuenta que los valores de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ tienen una débil correlación cruzada ($r = -0.16$) [297, 298], por lo cual es de esperar que los espectros difieran.

Este capítulo tiene tres objetivos. El primero es explorar las limitaciones impuestas por la componente estocástica en el muestreo de los belemnites en la cuenca Vasco-Cantábrica [296] para detectar algún tipo de señal orbital por medio del análisis espectral (periodograma suavizado de Lomb-Scargle). El segundo es averiguar si las fluctuaciones de los valores del $\delta^{18}\text{O}$ [297, 299] y del $\delta^{13}\text{C}$ [298] podrían reflejar algún tipo de ciclo orbital (astronómico) de largo período (mayor o igual a los 100 Ka). Por último, como tercer objetivo, se pretende comparar entre sí las características espectrales de las series temporales del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$.

4.2. Paleogeografía y paleoceanografía

Hace unos 183 millones de años la distribución física (continentes y océanos) de la Tierra era muy diferente a la de hoy en día. De acuerdo a reconstrucciones paleocontinentales realizadas por Ziegler [390] y Baudin *et al.* [18] para el Jurásico tardío, la cuenca Vasco-Cantábrica (Figura 4.1) estuvo localizada en un mar epicontinental (mar de gran extensión y poco profundo que se extiende sobre un continente), limitada al norte por el macizo Armoricano y al sur por el macizo Ibérico, y conectaba el mar Boreal del norte con el océano Tethys del sur [71, 106, 297].

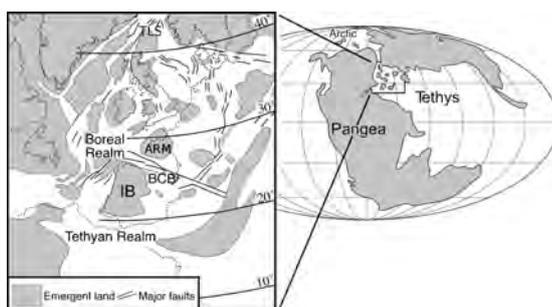


Figura 4.1: Paleogeografía del Jurásico tardío de la región noroeste de Europa. Modificado a partir de Ziegler [390] y Baudin *et al.* [18]. La región de estudio, la cuenca Vasco-Cantábrica (BCB), estuvo localizada entre 20 y 30° latitud Norte. ARM y IB son los macizos Armoricano e Ibérico, respectivamente, y TLS es el Transcontinental Laurasian Seaway.

Las características paleoceanográficas para el Jurásico tardío están basadas en reconstrucciones paleogeográficas, pero con información adicional de análisis paleobiogeográfico y de simulaciones numéricas por medio de modelos acoplados océano-atmósfera de circulación general (AOGCM) [11, 29, 52, 302]. La paleobatimetría para la cuenca Vasco-Cantábrica establece una profundidad máxima de 200 m [29, 390]. La circulación oceánica alrededor de la cuenca Vasco-Cantábrica, muy probablemente fue de Norte-Sur, debida a la diferencia de densidades entre las masas de agua, del Ártico y Tethys [29]. Recientemente, Arias [11], y en base a los resultados de otros autores [52, 254], propuso una representación esquemática

de la circulación superficial del océano de Tethys (Figura 4.2), la cual estuvo dominada por un sistema monzónico, cuyas corrientes superficiales ecuatoriales están dirigidas hacia el Oeste.

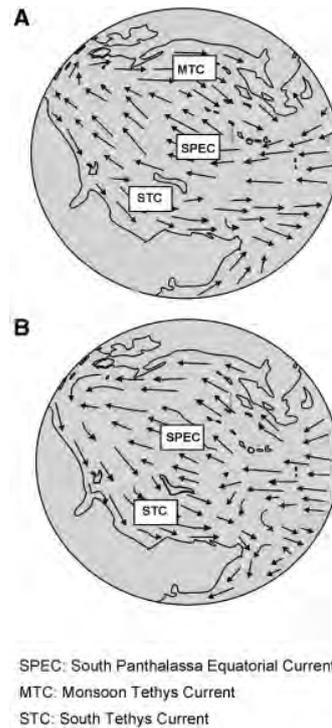


Figura 4.2: Representación esquemática de la circulación de Tethys durante el monzón de verano (A) y el de invierno (B). Tomada de Arias [11].

4.3. Material y métodos

4.3.1. Series temporales de valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$

Los valores medios (se usan valores medios porque en una misma *bed*¹ pueden haber varias partes fosilizadas de un rostrum de belemnites o varias muestras de diferentes especies de belemnites [297]) de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ usados en este estudio

¹La menor resolución espacial de una capa estratigráfica.

4.3 Material y métodos

fueron obtenidos del análisis de muestras fosilizadas de *rostra* de belemnites (Figura 4.3) recolectadas en la cuenca Vasco-Cantábrica, norte de España (Figura 4.4) [296, 297, 298, 299]. Estos valores forman una sección compuesta que abarca 9.6 Ma (de 191.60 a 182.00 Ma) y pertenecen principalmente al Plienbachiense (90 % de los datos) y al Toarciense tardío. Sin embargo, debido a las pocas muestras que se tienen en las biozonas *jamesoni* y *serpentinus* (Figura 4.4), y para satisfacer la propiedad de estacionariedad, ambas series temporales han sido recortadas para formar una sección que va de 188.68 a 183.36 Ma (la sección abarca unos 5.3 Ma) (Figuras 4.5 y 4.6, Tabla 4.1). Donde las áreas grises de estas figuras representan eventos oceánicos anóxicos (OAE) y las líneas verticales discontinuas indican el intervalo temporal donde se aplica el análisis espectral.



Figura 4.3: a) Fotografía de un *rostrum* de belemnite, *i. e.*, la parte fosilizada de donde se obtienen los valores isotópicos del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$. b) Pintura de un belemnite basado en reconstrucciones paleontológicas. Tomadas de *The Octopus News Magazine Online*, <http://www.tonmo.com/science/public/belemnites.php>, último acceso 20/08/2011.

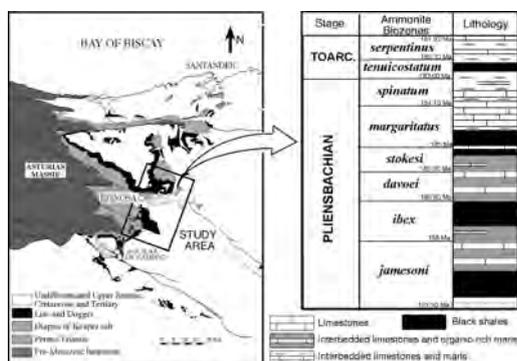


Figura 4.4: Representación geográfica-geológica y sección estratigráfica (simplificada) del área donde las *rostra* de belemnites fueron recolectadas (Norte de España). Tomada y modificada de Rosales *et al.* [297].

El muestreo fue realizado en intervalos estratigráficos de 0.2-5 m y fue calibrado contra bioestratigrafía disponible basada en biozonas de amonites [39, 59]. Las muestras utilizadas en este estudio han sido 176 (originalmente fueron recolectadas 200) y son las que no presentaron efectos diagenéticos [296]. De acuerdo al

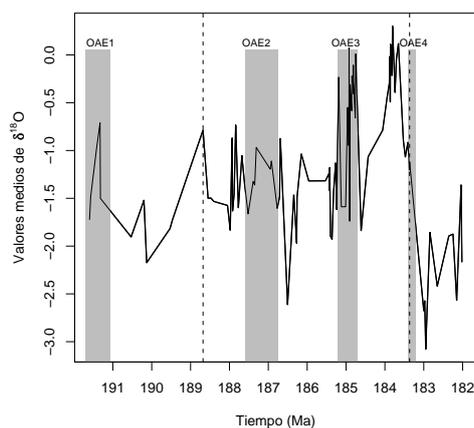


Figura 4.5: Serie temporal (no equiespaciada en el tiempo) de valores medios de $\delta^{18}\text{O}\text{‰}$ PDB de las *rostra* de belemnites del Jurásico tardío usadas en este estudio. Modificada de Rosales *et al.* [297].

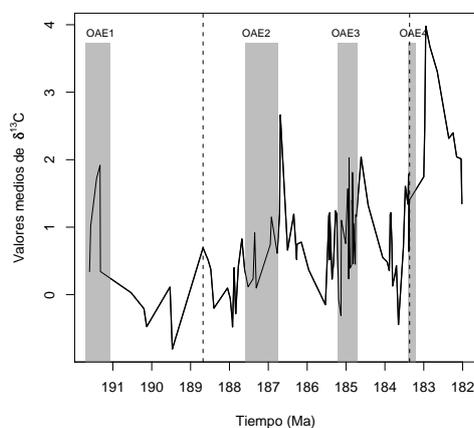


Figura 4.6: Serie temporal (no equiespaciada en el tiempo) de valores medios de $\delta^{13}\text{C}\text{‰}$ PDB de las rostra de belemnites del Jurásico tardío usadas en este estudio. Modificada de Rosales *et al.* [297].

Variable	N	Valor medio	Varianza	Val. extremos
$\delta^{18}\text{O}$	80	-0.9622	0.3896	-2.6120, 0.3015
$\delta^{13}\text{C}$	80	0.7607	0.3785	-0.478, 2.664

Tabla 4.1: Resumen de los principales estadísticos de las series temporales del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ (intervalo temporal 188.68-183.36 Ma), los valores están en ‰ PDB.

análisis bioestratigráfico, la edad de las muestras fue estimada por Rosales *et al.* [297] y está basada en la duración de cada zona y sub-zona de amonites (considerando la razón de sedimentación constante para cada una de ellas) calculadas desde la razón de cambio del $^{87}\text{Sr}/^{86}\text{Sr}$ con edades basadas en McArthur *et al.* [212, 214] y como *tie points* edades radiométricas de Palfy *et al.* [251]. Alrededor de 1-10 mg de calcita de las rostra de belemnites fueron usadas para el análisis isotópico, el cual fue realizado en un espectrómetro de masas VS SIRA-9 [297].

4.3.2. Métodos

Las series temporales de los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ no están equiespaciadas en el tiempo. Para evitar la interpolación se utiliza una versión adaptada por Schulz y Statteger [317] y Schulz y Mudelsee [316] del periodograma de Lomb-Scargle (*Lomb-Scargle Fourier Transform*, LSFT) [193, 311, 312] para analizar series paleoclimáticas. Esta adaptación tiene en cuenta el tipo de ruido de fondo presente en las series temporales paleoclimáticas, *i. e.*, ruido rojo [107, 124, 371] y lleva de nombre REDFIT [316]. Para discernir si un pico espectral de las series temporales a estudio es significativo contra la hipótesis de que ha sido producido por un ruido de fondo del tipo AR1 (ruido rojo), su espectro es comparado con el espectro teórico correspondiente al ruido rojo. Una descripción detallada de este método de análisis espectral puede encontrarse en el capítulo de metodología, sección 2.2 de esta memoria. Los parámetros de entrada de REDFIT utilizados en este caso de estudio están resumidos en la tabla 4.2.

Tabla 4.2: Parámetros de entrada para REDFIT.

Parámetro	Valor
Frec. mínima	$f_o = 1.17e^{-07}$
Frec. máxima	$\langle f_{Nyq} \rangle = 7.52e^{-06}$
$\langle \Delta t \rangle^a$	66,450 años
No. de segmentos WOSA	$N50 = 4$
Oversampling factor	$ofac = 4$
Tipo de ventana	$iwin = \text{Hanning}$
No. de simulaciones de Monte-Carlo	$Nsim = 5000$
dof ^b	8
$\langle \tau_\tau \rangle^c$	35,000 años

^a Intervalo promedio de muestreo

^b Grados de libertad (*degrees of freedom*)

^c Coeficiente promedio de persistencia

La versión de REDFIT (3.5) que se ha utilizado en este trabajo no incorpora errores de datación, aunque la más reciente versión [226] sí tiene en cuenta esta característica, pero no es de libre acceso, uso o distribución. Por lo tanto, se ha

tenido que implementar un programa computacional en lenguaje R [280] en combinación con REDFIT (3.5) para tener en cuenta los errores de datación de las muestras de belemnites para lo cual se utilizan simulaciones de Monte-Carlo. Los errores de datación pueden ser considerados como un proceso Gaussiano, centrado en la mejor estimación de la datación de una muestra y con una desviación estándar δ_d que depende de la abundancia relativa de los fósiles de belemnites dentro de cada *bed*. Este procedimiento puede ser considerado como una adaptación del modelo propuesto por Fleitmann *et al.* [91] y por Mudelsee *et al.* [226], y se resume en los siguientes pasos.

- Sean $\{t(i)\}_{i=1}^N$ los tiempos de las muestras de belemnites y $\{y(i)\}_{i=1}^N$ los valores isotópicos.
- For k=1, Nsims
 1. $t^*(i) = t(i) + \mathcal{E}_{N(0, \delta_d^2)}$, los tiempos perturbados, y donde $\delta_d = 50$ Ka
 2. Se verifica que los nuevos tiempos t^* para cada realización k, tengan un incremento estrictamente monótono.
 3. Se calcula el espectro de $\{t^*(i), y(i)\}_{i=1}^N$ vía REDFIT
- Se calculan los percentiles (95 %) de los Nsims espectros.

donde Nsims es el número de simulaciones de Monte-Carlo (10,000 en este estudio), y los valores de entrada de REDFIT están proporcionados en la tabla 4.2. Nótese que cada vez que REDFIT estima un espectro para cada realización $\{t^*(i), y(i)\}_{i=1}^N$, también estima su correspondiente nivel de ruido rojo con un nivel de confianza del 95 % que se utiliza en el análisis de resultados.

Extendiendo la frecuencia máxima en la estimación del espectro

Un punto que puede ser complicado al utilizar métodos de análisis espectral (no se limita al periodograma de Lomb-Scargle) para series temporales no equiespaciadas en el tiempo, es determinar la frecuencia máxima F_{max} hasta donde se estima el espectro. Para el caso equiespaciado, con $N' = \lfloor N/2 \rfloor$ (donde N es el número de elementos de la serie temporal y $\lfloor \cdot \rfloor$ denota la función máximo entero) la F_{max} es la frecuencia de Nyquist [20, 311], *i. e.*,

$$F_{Nyq} = \frac{1}{2\Delta t} \quad (4.1)$$

donde Δt es el intervalo de tiempo de muestreo.

Sin embargo, para el caso no equispaciado temporalmente, no hay una clara definición de frecuencia de Nyquist [86, 311, 317, 354]. Una opción, un tanto conservadora, para superar este obstáculo ha sido utilizar una frecuencia de Nyquist promedio [277, 317], *i. e.*,

$$\langle F_{Nyq} \rangle = \frac{1}{2 \langle \Delta t \rangle} \quad (4.2)$$

donde $\langle \Delta t \rangle$ es el valor medio de todos los Δt 's. Pero, como bien apunta Scargle [311], no siempre es claro cuál valor medio (la media aritmética, la mediana u otra medida estadística de tendencia central) es el apropiado. Por otro lado, es bien sabido que la media no es una estimación robusta de valor central, y, sin embargo, la mediana sí lo es [20]. En este punto una herramienta que puede ser de ayuda es el análisis exploratorio (histograma) de las diferencias de los tiempos. Como se puede observar en la Figura 4.7, al utilizar la media en lugar de la mediana, en la estimación de la F_{max} , se estaría haciendo una sub-estimación.

Al subestimar la máxima frecuencia, F_{max} , parte de la información espectral de la serie temporal podría no ser considerada [86]. Press *et al.* [277, 289] sugieren que es posible calcular una frecuencia máxima mayor que la $\langle F_{Nyq} \rangle$. Eyer y Bartholdi [86], Koen [165] y Pelt [258] mostraron que la máxima frecuencia puede alcanzar muy altas frecuencias, de tal modo que en muchos casos se puede estimar como $F_{max} = 1/(2 * p)$, donde p es el máximo común divisor de los Δt 's. Sin embargo, es importante tener en cuenta hasta dónde calcular la F_{max} , porque se puede producir *aliasing* o *pseudo-aliasing*, y por ello, introducir falsa información en el espectro ('artefactos' en la jerga del análisis espectral) [221, 226, 354].

Una metodología útil para averiguar la F_{max} sin introducir *aliasing* o *pseudo-aliasing*, es por medio del *periodogram window* (PW) o ventana espectral [70, 86, 193, 311, 354] (véase la sección 2.2 del capítulo de metodología), que se define [86, 354] como

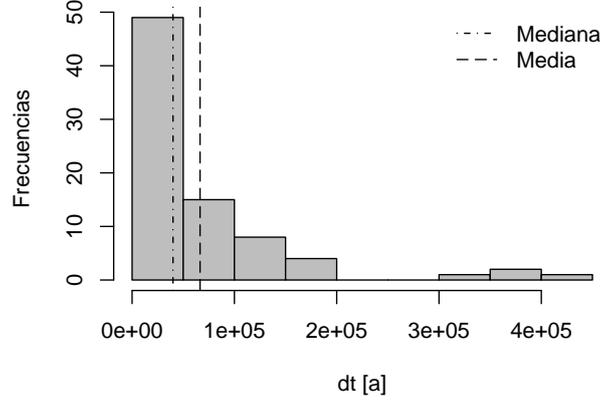


Figura 4.7: Histograma de los Δt 's de las series temporales de los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ de las rostra de belemnites de la cuenca Vasco-Cantábrica.

$$W(\Delta f) = \frac{|\sum_{k=1}^N \exp(i2\pi\Delta f(t_k - \theta))|^2}{N^2} \quad (4.3)$$

donde $\theta = (t_N + t_1)/2$ [354].

En este trabajo, y en base a los argumentos anteriormente presentados, se extiende la F_{max} utilizando la mediana de los valores Δt 's de las series temporales del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$, *i. e.*, $F_{max} = \frac{1}{2 * mediana(\Delta t's)}$. Para evitar *aliasing* o *pseudo-aliasing*, se calcula el PW por medio de la relación 4.3 [86, 165, 354]. Debido a que las muestras utilizadas en este estudio contienen una componente estocástica en su localización espacio-temporal, se incorporan los errores de las dataciones de las muestras en el PW y se determina un intervalo de confianza al 95 % utilizando simulaciones de Monte-Carlo. El intervalo de confianza corresponde al percentil 95 del valor máximo que alcanza el $w(\Delta f)$ calculado desde la frecuencia f_0 y teniendo como límite la frecuencia correspondiente a la frecuencia de Nyquist promedio $< F_{Nyq} >$ (Ecuación 4.2). Como se puede observar en la Figura 4.8 en el intervalo f_0 a F_{max} no se rebasa el intervalo de confianza. El primer valor que rebasa el intervalo de confianza está localizado en la frecuencia 0.0001 a^{-1} ($1/f = 10,000 \text{ a}$), y corresponde a un pseudo-alias.

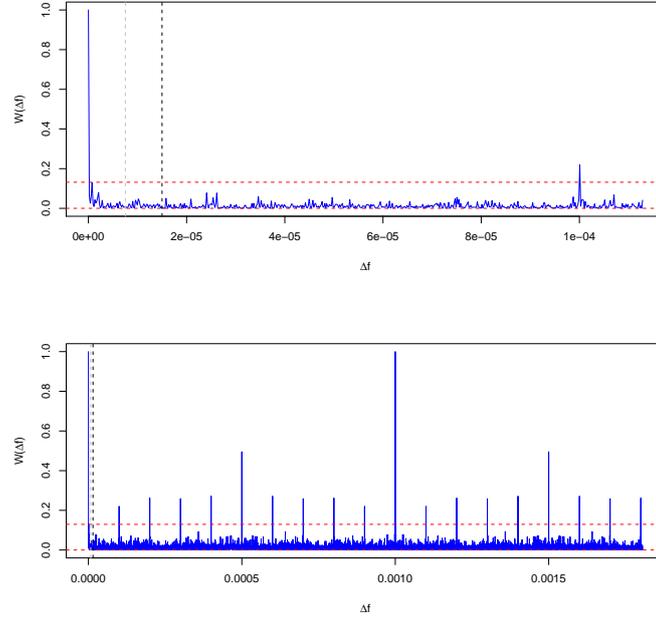


Figura 4.8: *Periodogram window* de las series temporales de los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ para el intervalo 188.68-183.36 Ma, para el caso cuando $F_{max} = 0.00012$ (figura superior), y para el caso cuando $F_{max} = 0.0017$ (figura inferior). La primera línea vertical discontinua (de izquierda a derecha) representa la $\langle F_{Nyq} \rangle$ utilizando un valor medio de los Δt 's. La segunda línea vertical es la ' $\langle F_{Nyq} \rangle$ ' utilizando la mediana de los Δt 's. La línea punteada horizontal superior de color rojo, representa el nivel de confianza al 95 % de los valores $W(\Delta f)$ teniendo en cuenta los errores de datación de las muestras de belemnites.

4.4. Resultados

Antes de hacer una discusión de los resultados, es importante tener en cuenta un par de factores. La series temporales de valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ no han sido interpoladas en el tiempo y no se ha aplicado ningún tipo de *orbital tuning*. Esto es, no se ha aplicado ningún método que ajuste los eventos estratigráficos en el dominio de la profundidad a algún forzamiento de tipo astronómico (*v. gr.*, los ciclos orbitales o de Milankovitch) en el dominio del tiempo [252, 371]. La razón para ello, es que el *orbital tuning* puede introducir razonamiento circular

en el análisis de resultados y porque los periodos (ciclos) utilizados como objetivos para hacer el *tuning* muchas veces están basados en propiedades espectrales de una serie de tiempo más que en los procesos físicos envueltos en su generación [146, 191].

Los resultados obtenidos después de aplicar análisis espectral por medio del periodograma suavizado de Lomb-Scargle (con 8 grados de libertad) a los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ y teniendo en cuenta los errores de datación de las muestras de los belemnites, se muestran en la Figura 4.9. Los picos espectrales que superan el espectro asociado a un ruido de tipo AR1 con un nivel de confianza del 95 %, están localizados *ca.* 1.3 Ma para el $\delta^{18}\text{O}$ y *ca.* 2.1 Ma y 109 Ka para el $\delta^{13}\text{C}$.

4.5. Discusión

Los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ provenientes de carbonatos marinos, como las rostra de belemnites, pueden estar influenciados principalmente por los siguientes factores: procesos diagénicos, procesos fraccionales biogénicos¹, variaciones isotópicas intra-especimen e inter-especimen y factores ambientales [213, 296, 297]. El primer factor puede ser descartado como la principal causa, porque las muestras de belemnites utilizadas para obtener los valores isotópicos medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$, fueron previamente verificados por Rosales *et al.* [296] ante posibles efectos de alteración diagénica.

El segundo factor, *i. e.*, los efectos vitales, son más difíciles de cuantificar porque los belemnites son organismos extintos [213, 296]. Estudios previos han demostrado que las muestras de belemnites provenientes de diferentes lugares y edades no se agrupan isotópicamente de acuerdo a las especies, por ello, es poco probable que exhiban efectos vitales [296, 305, 347]. Por otro lado, se sabe que los belemnites tenían cierto parecido a los actuales cefalópodos [283, 296, 375], y que, en estos últimos, los efectos vitales no influyen en la composición isotópica del oxígeno presente en sus conchas. En base a estos argumentos, Rexfort y Mutterlose [283] y otros autores [375] sugieren que no se espera encontrar efectos de

¹También conocidos como efectos vitales y son producidos cuando los carbonatos orgánicos, de la parte que fosiliza de los belemnites, precipita en desequilibrio isotópico con el agua oceánica [278].

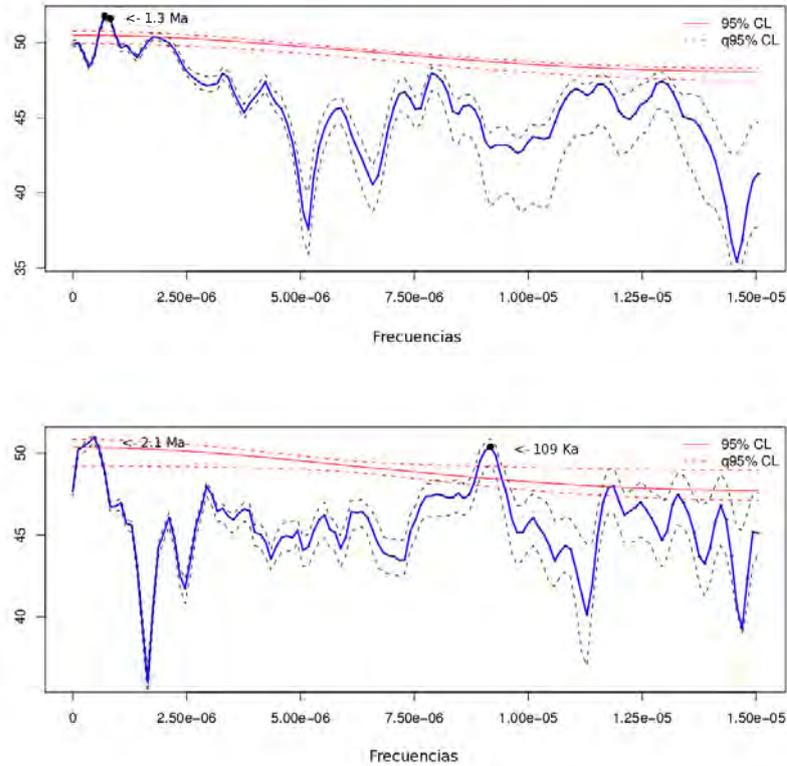


Figura 4.9: Espectros corregidos y suavizados (con 8 grados de libertad) de las series temporales de los valores medios de $\delta^{18}\text{O}$ (figura superior) y $\delta^{13}\text{C}$ (figura inferior) (curvas continuas de color azul) para el intervalo 188.68-183.36 Ma. El espectro fue calculado por medio del programa REDFIT, los parámetros de entrada pueden consultarse en la tabla 4.2, excepto la F_{max} que ha sido calculada utilizando la mediana de los Δt 's. Las curvas discontinuas (en negro) que rodean el espectro, son los intervalos de confianza (5 y 95 %) del espectro al tener en cuenta los errores de datación mediante el método de Monte-Carlo. La curva sólida de color rojo representa el 95 % del nivel de confianza correspondiente a ruido rojo. Las curvas discontinuas de color rojo, son los intervalos de confianza (5 y 95 %) del ruido al tener en cuenta los errores de datación. Se han utilizado 10,000 simulaciones de Monte-Carlo.

biofraccionamiento isotópico en belemnites. Sin embargo, es posible que los efectos paleoecológicos de los belemnites puedan influir en la composición isotópica [213, 283, 296]. Es ampliamente aceptado que los belemnites fueron un tipo de

organismo neotónico, esto quiere decir que muy probablemente hayan podido migrar lateral y verticalmente, a través de masas de agua de diferente temperatura o salinidad [77, 213, 274, 283]. Si este fuera el caso, la composición isotópica de las rostra podría haber cambiado durante su crecimiento en respuesta a los cambios en el ambiente (temperatura y salinidad) [213, 283].

El tercer factor, *i. e.*, las variaciones isotópicas intra-especimen e inter-especimen, son discutidas a continuación. Las series temporales en este estudio corresponden a valores medios de los isótopos estables $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$, esto significa que para una misma *bed* puedan encontrarse diferentes secciones fosilizadas de un mismo belemnite o varias muestras de especies diferentes de belemnites. Rosales *et al.* [296] demostraron que las variaciones de los valores isotópicos ($\delta^{18}\text{O}$ y $\delta^{13}\text{C}$) intra-especimen no superan un 0.3‰, para una misma *bed*, siempre y cuando las muestras no presenten efectos diagénicos. Por ello, no parecen jugar un papel importante en las variaciones de los valores isotópicos. Sin embargo, las variaciones de los valores isotópicos (191.60-182 Ma) inter-especimen, para una misma *bed*, son mucho mayores, hasta un 1.2 ‰ para el $\delta^{18}\text{O}$ y hasta un 2.1 ‰ para el $\delta^{13}\text{C}$ [296].

Rosales *et al.* [296] sugieren que las variaciones de los valores isotópicos inter-especimen para una determinada *bed*, podrían reflejar el comportamiento paleoecológico de diferentes especies de belemnites o podrían estar causadas por variaciones seculares de alta frecuencia como los ciclos de Milankovitch ($10^4 - 10^5$ años). Respecto a la primera hipótesis, Rosales *et al.* [296] sugieren que diferentes especies de belemnites que hayan coexistido en diferentes hábitats, pueden presentar diferentes composiciones isotópicas de $\delta^{18}\text{O}$ si durante sus ciclos de vida hayan habitado de modo preferente en diferentes masas de agua (aguas profundas y frías, aguas superficiales y cálidas) [296]. Aunque los belemnites son considerados organismos neotónicos, hay estudios que sugieren que existieron especies del tipo neotobentónico [9, 278]. Estos argumentos sugieren la importancia de identificar las especies de las muestras de belemnites en el intento de explicar las fluctuaciones de los valores isotópicos, sin embargo, para este estudio, no se tiene esa información.

La segunda hipótesis no se puede descartar, aunque es difícil de sostener porque hay pocas muestras por cada biozona y debido a las restricciones crono y

bio-estratigráficas no se puede resolver temporalmente más allá de una biozona de amonites¹ (Figura 4.4) [296]. Sin embargo, quizá es posible detectar algún tipo de ciclo orbital de más largo periodo que los de Milankovitch (periodo mayor que 10^5 años) pero teniendo en cuenta varias biozonas, como en la serie de los valores isotópicos ($\delta^{18}\text{O}$ y $\delta^{13}\text{C}$) analizada en este estudio.

El cuarto factor, la influencia de los factores ambientales, se discute en las siguientes líneas. En un par de trabajos previos Rosales *et al.* [297, 299] utilizaron exitosamente valores medios de $\delta^{18}\text{O}$ para las reconstrucciones de paleotemperaturas y paleosalinidades. Esto quiere decir que las variaciones de los valores medios de $\delta^{18}\text{O}$ están relacionadas, hasta cierto punto, con cambios en la temperatura/salinidad del océano de esa época [297, 299]. En un trabajo más reciente, Rosales *et al.* [298] muestran que hay un registro en las rostra de belemnites de las principales perturbaciones geoquímicas del agua del mar. Sin embargo, estos resultados no contemplan la posibilidad de que las variaciones isotópicas ($\delta^{18}\text{O}$ y $\delta^{13}\text{C}$) pudieran haber sido influenciadas por forzamientos externos, tales como los ciclos orbitales de largo periodo. Para analizar esta posibilidad, se emplean los resultados de las estimaciones de los espectros suavizados de los valores medios del $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$.

Los únicos picos espectrales significativos, utilizando una frecuencia de Nyquist promedio ($\langle F_{Nyq} \rangle$) en la estimación del espectro, están localizados *ca.* 1.3 Ma (Figura 4.9 superior) para el $\delta^{18}\text{O}$ y *ca.* 2.1 Ma (Figura 4.9 inferior) para el $\delta^{13}\text{C}$. Sin embargo, a pesar de ser significativos y de tener en cuenta en el espectro los errores de datación, estos picos espectrales podrían estar influenciados por la tendencia de las series temporales. Es cierto que la tendencia (lineal) ha sido removida antes de calcular el espectro, pero también hay que tener en cuenta que el cálculo de la tendencia en una serie temporal es algo complicado, y es prácticamente imposible (en especial cuando no se conocen bien los procesos subyacentes del sistema que representan) separar de forma suficientemente precisa, tendencia, periodicidades de muy baja frecuencia y ruido [354, 383]. Más aun, es

¹La sección estratigráfica en este estudio contiene 6 subzonas de amonites, cada subzona tienen una longitud espacio-temporal variable, desde 7 cm hasta 36 cm y de 0.30 Ma hasta 1.4 Ma (los intervalos espaciales no guardan una relación proporcional con los intervalos temporales).

complicado ubicar de forma precisa el ciclo orbital en una u otra frecuencia con una serie temporal tan corta. No obstante, suponiendo que 2.1 o 1.3 Ma representan eventos periódicos, entonces en la sección 188.68-183.36 Ma puede haber a lo sumo 3 y 4 ciclos, respectivamente. Por tanto, es poco fiable afirmar que los picos espectrales *ca.* 2.1 y 1.3 Ma representan eventos periódicos reales. Para poderlo afirmar con más confianza sería necesario disponer de más datos (series temporales más largas).

La ausencia del ciclo orbital de 400 Ka en los resultados del análisis espectral (Figura 4.9) para ambas series temporales (valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$), podría en teoría deberse a la irregularidad en el muestreo y a los errores de datación. Sin embargo, estas causas pueden ser descartadas porque en el análisis espectral se tienen en cuenta estos factores. La razón más plausible de la ausencia del ciclo de 400 Ka son las pocas muestras de belemnites (80) y la longitud de la sección estratigráfica (118 m). Por ejemplo, en el trabajo de Olsen y Kent [246, 247], que cubre una buena parte del Jurásico tardío, hay clara evidencia geológica de la existencia del ciclo orbital de 400 Ka, pero la resolución (espacio-temporal) estratigráfica de sus datos es de 1 m - 2 Ka y con una longitud de la sección estratigráfica de casi 5 km (unos 35 Ma).

Utilizando los resultados del análisis espectral, pero con una extensión de la frecuencia de Nyquist promedio (véase la sección 4.3.2 y Figura 4.8), se puede observar (Figura 4.9 inferior) la existencia de un pico espectral significativo (robusto al ruido de fondo y a los errores de datación) en el espectro del $\delta^{13}\text{C}$ y con una periodicidad *ca.* 109 Ka. Este pico espectral podría estar relacionado con la excentricidad de la Tierra, cuya periodicidad promedio es ~ 100 Ka¹ [23, 63, 175]. Desde el punto de vista de la evidencia geológica, existen varios trabajos para el Jurásico tardío en los cuales se utilizan diversos registros geológicos procedentes de diferentes regiones geográficas, donde han localizado el ciclo promedio de la excentricidad de 100 Ka [90, 245, 246, 346, 370, 372, 373]. Sin embargo, este ciclo no está presente en el espectro del $\delta^{18}\text{O}$ (Figura 4.9), donde en principio, se esperaría encontrarlo.

¹Es importante tener en cuenta que el periodo de 100 Ka, es en realidad un promedio de los 4 periodos dominantes (95, 123, 99 y 130 Ka) de la excentricidad de la Tierra [23, 63, 175].

Si observamos con más detalle la Figura 4.9 superior, utilizando el eje x (frecuencias) como eje de simetría, es posible observar una región, casi a la mitad del eje x, con altos valores de las amplitudes espectrales. El máximo pico de esta región corresponde a una periodicidad de 125 Ka. Esto podría llevar a pensar que representan un mismo evento periódico, sin embargo, si tenemos en cuenta el error de la localización de las frecuencias (tanto debido a los errores de datación como a la técnica utilizada para el análisis espectral), no es posible conciliar estas diferencias. La posibilidad de que el ciclo de 109 Ka sea un artefacto introducido en el espectro, debido a la extensión de la frecuencia promedio de Nyquist, puede ser descartada, porque esta posibilidad ha sido previamente verificada (véase la sección 4.3.2).

Una explicación más sencilla sobre el origen del ciclo de 109 Ka, es debido a un efecto de *aliasing*. Se sabe por otros estudios de series temporales paleoclimáticas [273, 386] que ciclos de más corto periodo, pueden producir picos espectrales significativos de más largo periodo, si la resolución espacial de las muestras es demasiado ancha/grande para resolver las señales de corto periodo. Por ejemplo, Pisias y Mix [273] demostraron que los ciclos de precesión (*ca.* 19 Ka y 23 Ka, valores en el presente) pueden producir picos espectrales significativos de más largo periodo cercanos a 100 Ka cuando se utiliza un intervalo de muestreo de 25 Ka. Lo cuál podría relacionarse erróneamente con el ciclo de 100 Ka de la excentricidad. Si bien es cierto que el origen un tanto aleatorio de las localizaciones espacio-temporales de las muestras de belemnites puede reducir el fenómeno del aliasing, también es cierto que no puede evitarlo por completo [86, 221, 273, 354].

Esta hipótesis se ve reforzada, si tenemos en cuenta el ciclo de 20 Ka, relacionado con la precesión, y encontrado por Berreteaga *et al.* [25] al utilizar registros de CaCO₃ provenientes de la BCB para el Toarciense inferior. Esto se ve agravado por el estilo cosmopolita de vida de los belemnites, es decir, su probable comportamiento migratorio (vertical y lateralmente) como una respuesta a cambios ambientales (temperatura y salinidad del océano) ocurridos durante su tiempo de vida [213, 274]. Es cierto que no se sabe con exactitud el tiempo de vida de los belemnites, pero se estima entre 2 a 8 años [274, 283]. Esto hace factible que los valores isotópicos puedan contener alguna señal ‘aliesada’ de corto periodo, incluso relacionada con algún tipo de evento climático periódico, como el ciclo

anual o estacional. Se sabe por simulaciones numéricas (modelos AOGCM) que el clima del Jurásico tardío fue de tipo monzónico (con gran dominio a lo largo de las medias latitudes), con grandes contrastes de temperaturas entre estaciones y altos valores de temperaturas del aire (entre 5 y 10°C por encima que la temperatura promedio de la Tierra de hoy en día) [11, 29, 52, 302]. Esta variabilidad también podría aparecer aliaseda en las series disponibles.

En base a estos argumentos y debido a la resolución espacio-temporal de las muestras de belemnites utilizadas en este estudio, no es posible discernir con seguridad desde el punto de vista estadístico si el pico de 109 Ka es en realidad un alias de la precesión o de algún otro evento de tipo periódico o casi periódico de frecuencia más alta.

Capítulo 5

Principales causas de las fluctuaciones de capturas de atún rojo entre 1525-1756 para el sur de España explicado por análisis de devianza mediante GLM

5.1. Introducción

El atún rojo (*Thunnus thynnus*) (Figura 5.1), es, entre los diferentes tipos de atunes, el que tiene la mayor distribución geográfica (Figura 5.2) y el único pez pelágico de gran tamaño que puede vivir permanentemente en aguas templadas del Atlántico, con temperaturas entre 3 °C y 30 °C [14, 31, 97]. Se mueve principalmente en zonas de surgencia, como la costa oeste de Marruecos y Portugal, y en estructuras oceanográficas de meso-escala asociadas con la circulación del Atlántico Norte y mares adyacentes, como la costa Este de Norteamérica, el Golfo de Vizcaya y el noroeste del Mediterráneo [14, 32, 37, 377]. Pueden localizarse cerca de la costa o en mar abierto, a profundidades que van desde la superficie del mar o hasta los 1000 m [42, 197].

El atún rojo puede alcanzar una longitud máxima de 4 m, puede pesar hasta 726 kg, aunque el promedio es de 650 kg, y puede alcanzar una edad máxima de

20 años [96, 103]. Los atunes en la fase inicial de su vida (larvas) se alimentan de zooplancton [352], en la etapa juvenil y adulta son depredadores y oportunistas, que consumen principalmente peces pequeños como sardinas, caballas, jureles, etc. e invertebrados como medusas, salpas, cangrejos, etc. [96].



Figura 5.1: Fotografía de un atún rojo (*Thunnus thynnus*). Extraído del Sistema de información global de pesquerías de la página web de la FAO, último acceso mayo de 2011.

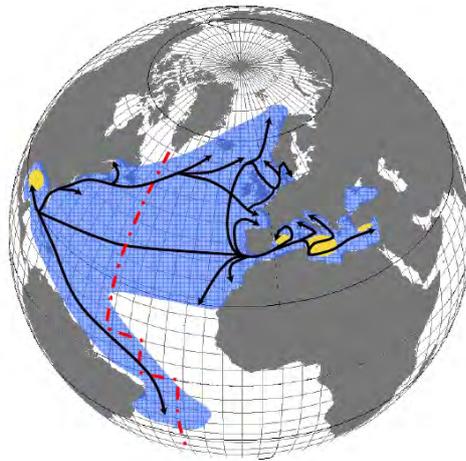


Figura 5.2: Mapa de la distribución espacial del atún rojo (en azul), principales rutas de migración (en negro) y principales áreas de desovamiento (en amarillo). La línea roja discontinua limita los dos principales stocks. Tomado de Fromentin [96].

Una de las características fisiológicas de los túnidos (como el atún rojo), es su capacidad termorreguladora. Estos peces han desarrollado un complejo sistema

5.1 Introducción

para conservar el calor corporal, denominado *rete mirabile*, el cual consiste en una red de vasos capilares a contracorriente. Sin embargo, cuando este sistema termorregulador no es suficiente para refrigerar su cuerpo, entonces estas especies realizan migraciones verticales buscando temperaturas menores del océano [68, 99]. Debido a estas características, este pez puede ser usado como un testigo para estudiar cambios en la temperatura del mar (SST).

El área de estudio de este trabajo (Figura 5.3) está localizada en la costa Atlántica de las provincias de Cádiz y Huelva (sur y suroeste de España) [102]. Las capturas fueron realizadas durante el periodo 1525-1756 por medio de un sistema de almadrabas (Figura 5.4) que pertenecieron al duque de Medina Sidonia [194]. Esta técnica para capturar atún rojo ha sido usada en el sur de España desde la época de los fenicios, aprovechando su comportamiento migratorio, y no ha sufrido cambios significativos [4, 194, 294]. Las mayores capturas anuales para el periodo de estudio se realizaron en 1559, con más de 125,000 atunes (téngase en cuenta que el valor promedio de capturas del periodo de estudio es de 24,000 atunes) entre las almadrabas de Conil y Zahara [102]

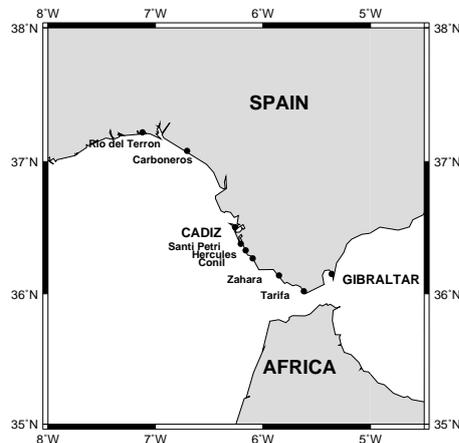


Figura 5.3: Área geográfica de estudio y localización donde estuvieron instaladas las almadrabas del duque de Medina Sidonia. El 95.2% de la captura total de atún se realizó en Conil y Zahara. Tomado de Ganzedo *et al.* [102].



Figura 5.4: Ilustración de una almadraba (área delimitada en negro) utilizada desde el s. XVI hasta principios del s. XX. Fuente: <http://www.historiacocina.com/historia/garum/almadrabas.html>, último acceso mayo de 2011.

La captura de atunes rojos por medio de almadrabas ha tenido grandes fluctuaciones en los siglos recientes [281] y este caso de estudio no es excepción. Diversas hipótesis han sido propuestas para explicar las fluctuaciones de la captura del atún rojo en el sur de España para el periodo 1525-1756. El monje Martín Sarmiento escribió en 1757 un informe por encargo del duque de Medina Sidonia para intentar explicar el porqué de la reducción en la captura de atún rojo de sus almadrabas, sobre todo para los años 1680 a 1756. Sarmiento propuso dos hipótesis sobre el origen de la tendencia decreciente: una que podría deberse a la sobreexplotación y la otra debido a las variaciones climáticas [102, 194].

Hoy en día se sabe que cambios en las condiciones ambientales (*v. gr.*, la temperatura del mar o del aire) en las rutas de migración del atún rojo, pueden afectar su abundancia y comportamiento [127, 166, 380]. Sin embargo, a pesar de estos estudios, no es fácil saber qué porcentaje de variabilidad de las capturas es debido a estos factores. Una forma de cuantificación de estos efectos es por medio de un análisis de devianza de la serie de capturas utilizando modelos lineales generalizados (GLM) [102, 115, 181].

Esta metodología no sólo puede ayudar a conocer qué porcentajes de varianza son debidos a factores climáticos, sino también a otros factores, como la dinámica poblacional de la especie a estudio o una combinación de ambos factores. Para aplicar la partición de varianza en este trabajo se han usado datos de reconstrucciones climáticas y salidas de simulaciones de modelos globales de

océano-atmósfera, como variables explicativas [102]. El objetivo de este trabajo es cuantificar los efectos de las variables climáticas en las fluctuaciones de las capturas de atún rojo.

5.2. Datos y métodos

5.2.1. Serie temporal de capturas de atún rojo

La serie temporal de capturas anuales de atún rojo para el periodo 1525-1756 (Figura 5.5) proviene de López-Capont [194]. Son capturas anuales en número de individuos, obtenidas en cada una de las almadrabas del duque de Medina Sidonia, principalmente (por encima del 95 %) de Conil y Zahara, las cuales estuvieron operativas durante todo el periodo de estudio (1525-1756). La serie cubre un periodo temporal de 232 años, con sólo 12 datos perdidos [102]. Estas pérdidas se deben a que en ciertos años no se hicieron los registros de las capturas o no se armaron las almadrabas [194].

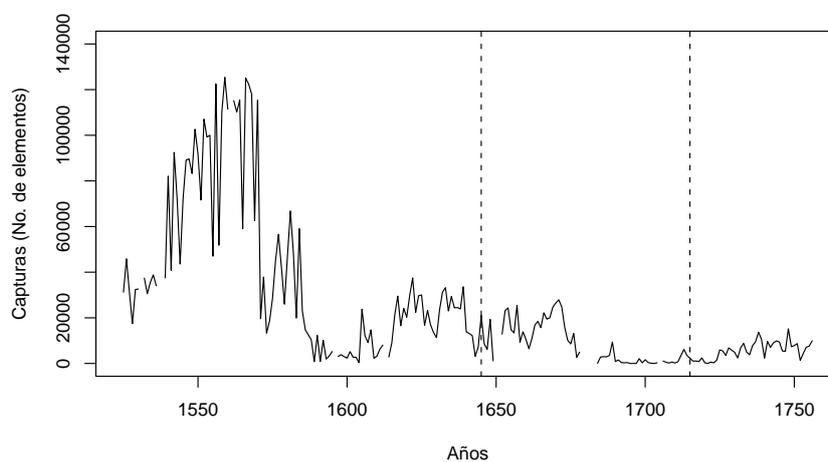


Figura 5.5: Serie temporal de captura de atún rojo para el periodo 1525-1756. El periodo entre las líneas verticales punteadas corresponde al evento climático de la ‘Pequeña Edad de Hielo’.

5.2.2. Datos climáticos

Los datos climáticos utilizados en este estudio representan temperaturas del aire (TA) y del océano (SST) al nivel del mar y se han utilizado porque ambas muestran una relación directa con las pesquerías de atún rojo [61, 261]. Debido a la carencia de datos climáticos observacionales del área geográfica y del periodo de estudio se han utilizado datos provenientes de simulaciones y de reconstrucciones climáticas.

Por un lado, los datos de las simulaciones, SST_s y TA_s (en °C) (Figuras 5.6 superior e inferior, respectivamente) provienen de los resultados generados con el modelo acoplado de Océano-Atmósfera ECHO-G por Zorita *et al.* [391]. El ECHO-G utiliza como modelo atmosférico el ECHAM4 y como modelo oceánico el HOPE-G [183]. Ambas temperaturas fueron promediadas mensualmente dentro de la malla 10°E-15°W, 32°N-42°N, la cual contiene el área de estudio. Un aspecto a considerar es que las temperaturas obtenidas con el ECHO-G no fueron capaces de reproducir la variabilidad climática de alta frecuencia (como el ciclo anual de la temperatura), aunque sí fueron capaces de reproducir la variabilidad climática de baja frecuencia (como la variabilidad a escala decadal) relacionada con forzamientos externos del clima (irradiación solar, vulcanismo y gases de efecto invernadero) [183, 391].

Otro de los aspectos a tener en cuenta es que una simulación climática no depende de las condiciones iniciales *i. e.*, no es un problema de valor inicial, sino más bien depende de otros forzamientos externos a la atmósfera que pertenecen a otros sub-sistemas climáticos con mayor memoria (*v. gr.*, el océano) [228]. La razón del porqué la predictibilidad del clima tiene otros orígenes diferentes al atmosférico, es debida a que ésta tiene un tiempo de predictibilidad muy corto (de 2 a 3 semanas) [253]. A grandes rasgos, la solución de las ecuaciones de evolución del sistema climático contiene una componente estocástica y otra (una función de Green representativa del sistema climático) que acumula la integración temporal del forzamiento sobre la dinámica interna del sistema. La suma de estas dos componentes hace que la variable de estado sea de tipo estocástica, con un término estocástico dependiente de la condición inicial y otro condicionado por los forzamientos externos. Por ello, el resultado de una simulación climática debe

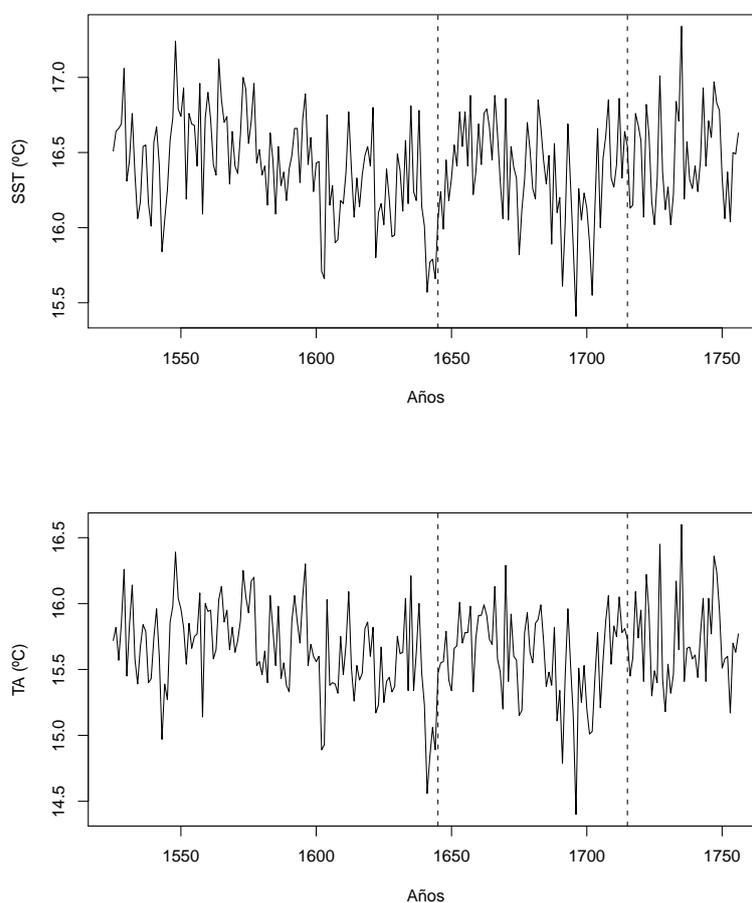


Figura 5.6: Series temporales de SST_s (figura superior) y TA_s (figura inferior) obtenidas de Zorita *et al.* [391] para el periodo de estudio 1525-1756. El periodo entre las líneas verticales punteadas corresponde al evento climático de la ‘Pequeña Edad de Hielo’.

interpretarse en términos de una distribución de probabilidad de la variable de estado (*i. e.*, de las variables climáticas, como la SST, la TA, etc.) [228].

Por otro lado, se han utilizado reconstrucciones paleoclimáticas de anomalías de temperaturas anuales globales del aire (TA_r) (en °C) al nivel del mar (Figura 5.7 superior) provenientes de Mann *et al.* [202]. Estas reconstrucciones están basadas en una calibración multivariante de paleoindicadores climáticos de alta

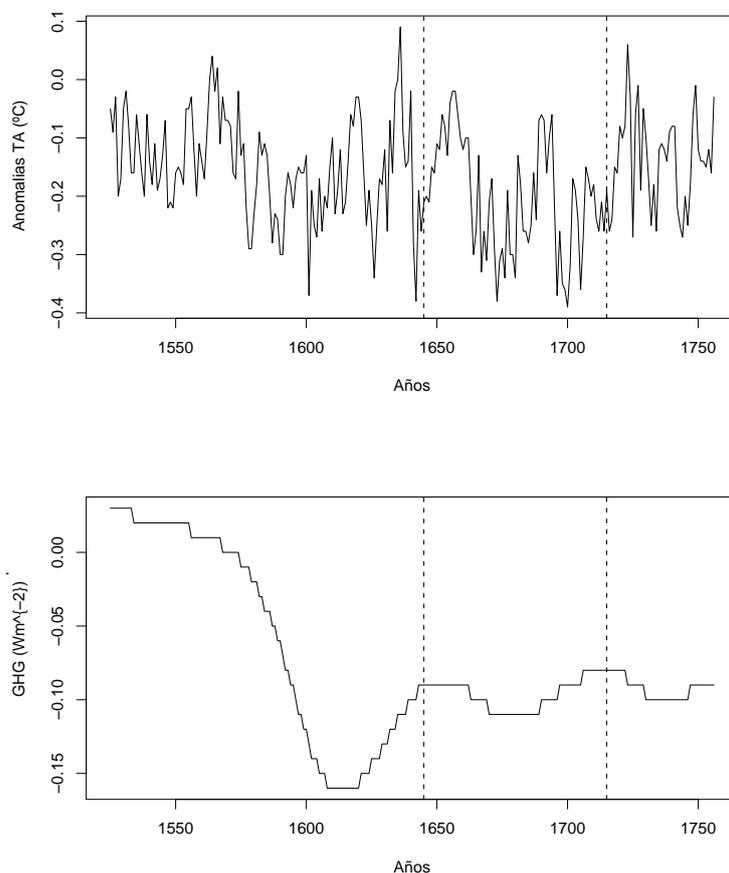


Figura 5.7: Anomalías de temperaturas globales anuales del aire (figura superior) ($^{\circ}\text{C}$) de reconstrucciones (TA_r) provenientes de Mann *et al.* [202] y forzamientos radiativos debido a las concentraciones de gases de efecto invernadero (Wm^{-2}) (figura inferior) provenientes de Crowley [62] para el periodo de estudio 1526-1756. El periodo entre las líneas verticales punteadas corresponde al evento climático de la ‘Pequeña Edad de Hielo’.

resolución provenientes de diferentes áreas geográficas. Estos datos proporcionan elementos para comprender la variabilidad climática natural a diferentes escalas espacio-temporales durante los últimos seis siglos. Del mismo modo como los datos provenientes de simulaciones climáticas, las TA_r de estas reconstrucciones incorporan información sobre los forzamientos externos del clima (irradiancia solar,

vulcanismo y gases de efecto invernadero) [202].

En este estudio también se utilizaron estimaciones de forzamientos radiativos (Wm^{-2}) (Figura 5.7 inferior) debidas a las concentraciones de CO_2 como gas de efecto invernadero (GHG), obtenidas de Crowley [62]. Estas estimaciones fueron calculadas mediante modelos de transferencia radiativa [62, 229]. Este último conjunto de datos ha sido utilizado porque los GHG podrían tener cierta relación con el ciclo de producción primaria, a través de la ganancia de energía y nutrientes durante el estado paralarvae de los atunes. Esta hipótesis fue comprobada por Nakagawa *et al.* [230], aunque utilizando un ambiente artificial controlado.

5.2.3. Métodos

Como primera aproximación se realizó un análisis de correlación simple entre las capturas de atún rojo (no interpoladas temporalmente) y cada una de las series climáticas (TA_s , TA_r , SST_s y GHG) para el periodo de estudio, esto, con el objeto de investigar si existe al menos una relación estadísticamente significativa entre las capturas y las variables climáticas. Para ello, se calculó el coeficiente de correlación por rangos de Spearman [376], esto es

$$r_s = 1 - \frac{6 \sum_{i=1}^N D_i^2}{N(N^2 - 1)} \quad (5.1)$$

donde D_i^2 es la diferencia entre los rangos de las dos variables elevadas al cuadrado y N es el número de muestras expresada en parejas de rangos de las variables a estudio. En el caso cuando existen valores repetidos en cualesquiera de las variables a estudio, se utiliza el rango promedio para cada uno de esos valores [376].

El r_s es un estadístico mucho más robusto (es menos sensible a valores extremos) que otros estadísticos que miden correlación entre variables y es en realidad el coeficiente de correlación de Pearson pero utilizando los rangos de los datos a estudio [365, 376]. Los valores del r_s oscilan entre -1 a 1, de tal modo que los valores negativos indican asociaciones lineales entre variables de modo decreciente, mientras que los valores positivos indican que esta asociación es de modo creciente (nótese que un valor de -1 o 1 implica una perfecta relación lineal

decreciente o creciente, respectivamente). En el caso donde r_s sea igual a cero indica que no existe asociación entre las variables. Para calcular el r_s se utilizó la función `cor` del paquete `stat` de R [280].

Adicionalmente se calculó la correlación cruzada entre las capturas (interpoladas linealmente en el tiempo, para lo cual se utilizó la función `na.interp` del paquete `forecast` en R [280]) y las variables climáticas para el periodo de estudio, esto con el fin de averiguar el grado de correspondencia estadística entre las variables. Para ello se utiliza la función de correlación cruzada (CCF) [330] para dos series temporales conjuntamente estacionarias $x(t_{i+k})$ e $y(t_i)$, con $i = 1, 2, \dots, N$.

La CCF ($\rho_{xy}(k)$) está acotada en el intervalo $[-1, 1]$ y mide la fortaleza y la dirección (retardo o adelanto) de la relación entre $x(t_{i+k})$ e $y(t_i)$. Esta última propiedad es útil para identificar variables causales. Valores positivos de $\rho_{xy}(k)$ indican un relación directa entre $x(t_{i+k})$ e $y(t_i)$, mientras que valores negativos indican una relación inversa. Un valor de 1 indica una perfecta relación lineal directa entre $x(t_{i+k})$ e $y(t_i)$, mientras que un valor de -1 indica una perfecta relación lineal inversa. Para calcular el estimador de la CCF se utilizó la función `ccf` del paquete `Stat` en R [280, 358].

Como segunda aproximación se construyeron cuatro modelos lineales generalizados (GLM l , $l=I,II,III$ y IV) para estudiar la relación entre las capturas de atún rojo (Y) (interpoladas linealmente en el tiempo, para lo cual se utilizó la función `na.interp` del paquete `forecast` en R [280]) y los descriptores climáticos X_i o variables explicativas, esto es, SST_s , TA_s [391], TA_r [202] y GHG [62]. Para ajustar los GLM se ha utilizado la técnica de máxima verosimilitud. Por tanto, estos modelos están optimizados a través de una reducción de la devianza más que de la varianza. Una forma de calcular la devianza explicada D^2 , es mediante la relación

$$D^2 = 1 - (\text{devianza residual})/(\text{devianza nula}) \quad (5.2)$$

la cual es equivalente al coeficiente de determinación (porcentaje de varianza explicada) utilizado en los modelos lineales ajustados por mínimos cuadrados [116].

El primer modelo, GLMI, es utilizado para explicar capturas de atún rojo teniendo en cuenta las variables $GHG + SST_s + TA_s$ [391]. El segundo modelo,

GLMII, es utilizado para explicar las capturas teniendo en cuenta las variables $\text{GHG} + \text{SST}_s$ [391] + TA_r [202]. El tercer modelo, GLMIII, es utilizado para explicar las capturas considerando las variables SST_s [391] + TA_r [202]. El cuarto modelo, GLMIV, es utilizado para explicar capturas considerando las variables $\text{SST}_s + \text{TA}_s$ [391].

Las estructuras temporales en la serie de las capturas podrían contener información de otros efectos históricos, bióticos o ambientales que no están considerados por los forzamientos ambientales externos [102, 182]. Por tanto, se utilizó un polinomio de tercer grado (t, t^2, t^3 , donde t es el tiempo en años de la serie de capturas) [36, 181] para considerar los efectos temporales como un sustituto para la influencia de la dinámica poblacional del atún rojo [102]. El grado del polinomio está determinado por el nivel de significación estadística ($\alpha \leq 0.05$) al ir ajustando término a término los elementos del polinomio en un modelo lineal generalizado, GLMt, que representa los efectos temporales (sin tener en cuenta las variables ambientales). Se calcula la devianza explicada D^2 para el GLMt mediante la relación 5.2.

A continuación se ajustan los cuatro modelos lineales generalizados, GLMs_l ($l=\text{I, II, III y IV}$) utilizando las respectivas variables climáticas de los $\text{GLM}l$ (I,II,III y IV) más todos los términos del polinomio de tercer grado y se calculan las devianzas explicadas D^2 para cada GLMs_l ($l=\text{I, II, III y IV}$) mediante la relación 5.2. La construcción y el ajuste de todos los modelos lineales generalizados utilizados en este trabajo se realizó mediante la función `glm` del paquete `stats` de R [280, 358] y se utilizaron los parámetros por defecto de esta función.

Finalmente, y en base al procedimiento propuesto por Bocard *et al.* [36] y por Legendre [181] (y ampliamente explicado en la sección 2.3 de esta memoria), se realizó una partición de la varianza de la serie de capturas utilizando la devianza obtenida de los modelos lineales generalizados anteriormente mencionados. Esto, con el objeto de determinar la contribución relativa de las variables climáticas comparada con la contribución de la dinámica poblacional [102]. La variación total de las capturas de aún rojo fue descompuesta en cuatro componentes:

- I efecto puro de las variables climáticas,
- II efecto puro de las variables temporales (representan la dinámica poblacional),

III efecto combinado de las variables climáticas con las temporales,

IV variación no explicada.

Utilizando las devianzas explicadas D^2 (en %) de los modelos lineales generalizados GLM l ($l=I,II, III$ o IV), GLMt y GLMs l ($l=I,II, III$ o IV) se calcularon los porcentajes de devianza, (los cuales se pueden interpretar como los porcentajes de varianza de las capturas) para cada componente de la siguiente forma:

$$D^2(\text{común}) = D^2(\text{GLM}l) + D^2(\text{GLMt}) - D^2(\text{GLMs}_l) \quad (5.3)$$

$$\text{Perc.clima} = [D^2(\text{GLM}l) - D^2(\text{común})] * 100 \% \quad (5.4)$$

$$\text{Perc.temp} = [D^2(\text{GLMt}) - D^2(\text{común})] * 100 \% \quad (5.5)$$

$$\text{Perc.común} = [D^2(\text{común})] * 100 \% \quad (5.6)$$

$$\text{Perc.noexp} = [100 - \text{sum}(\text{Perc.clima}, \text{Perc.temp}, -\text{Perc.común})] \% \quad (5.7)$$

donde Perc.clima, Perc.temp, Perc.común y Perc.noexp son los porcentajes de varianza debido solamente al efecto de las variables climáticas, únicamente al efecto de las variables temporales, al efecto combinado de las variables climáticas con las temporales y a la varianza no explicada, respectivamente.

5.3. Resultados

5.3.1. Análisis descriptivo de la serie de capturas

La serie temporal de capturas de atún rojo (Figura 5.5) cubre el periodo de 1525-1756, con máximas variabilidades en las capturas en los periodos 1571-1610 y 1672-1725. La serie tiene una tendencia decreciente, de 72819 (1525-1570) a 1410 (1686-1724) capturas. El registro más alto fue de 125,388 capturas en 1559, y el más bajo, de cero capturas, ocurrió en 1695-1697, 1705 y 1721 [102].

5.3.2. Análisis de correlaciones entre los datos

El resultado de calcular los coeficientes de correlación por rangos de Spearman entre las capturas de atún rojo y cada una de las series climáticas para el periodo de estudio se presenta en la Tabla 5.1. La variable climática que presenta mayor

correlación con las capturas de atún rojo son los GHG. Los resultados de las CCF se presentan en la Figura 5.8. Como se puede observar, todas las variables climáticas presentan correlación significativa, la cual se extiende a varios lags. Nótese otra vez la alta correlación entre las capturas y los GHG (Figura 5.8 inferior derecha).

Tabla 5.1: Coeficientes de correlación por rangos de Spearman y *p-values*.

Variabes climáticas	r_s	<i>p-value</i>
TA _r	0.291	< 0.001
TA _s	0.186	< 0.001
SST _s	0.217	< 0.001
GHG	0.378	< 0.001

5.3.3. Resultados del análisis de devianza mediante GLM

Los resultados de aplicar el análisis de devianza a las capturas de atún rojo utilizando los modelos lineales generalizados se presentan en las Tablas 5.2-5.5 y en la Figura 5.9. Para el GLMI (Tabla 5.2), el porcentaje de varianza atribuido a los efectos climáticos (GHG + SST_s + TA_s) es del 11.6 %, a los efectos temporales es del 13 %, por ambos efectos el 35 % y el porcentaje de varianza no explicado es del 40.4 %. Para el GLMII (Tabla 5.3), el porcentaje de varianza atribuido a los efectos climáticos (GHG + SST_s + TA_r) es del 12.3 %, debido a los efectos temporales es del 12.3 %, por ambos efectos del 35.7 % y el porcentaje de varianza no explicado es del 39.7 %. Para el GLMIII (Tabla 5.4), el porcentaje de varianza atribuido a las variables climáticas (SST_s + TA_r) es del 2.1 %, debido a los efectos temporales es del 38.8 %, por ambos efectos del 9.1 % y un 50 % de varianza no está explicada. Para el modelo GLMIV (Tabla 5.5), el porcentaje de varianza atribuida a variables climáticas (SST_s + TA_s) es del 1.4 %, debido a efectos temporales es del 43.2 %, por ambos efectos del 4.8 % y un 50.6 % de varianza no está explicada [102].

Las variables significativas en los modelos GLMI y GLMII (Tablas 5.2 y 5.3) indican que para las variables climáticas sólo los gases de efecto invernadero

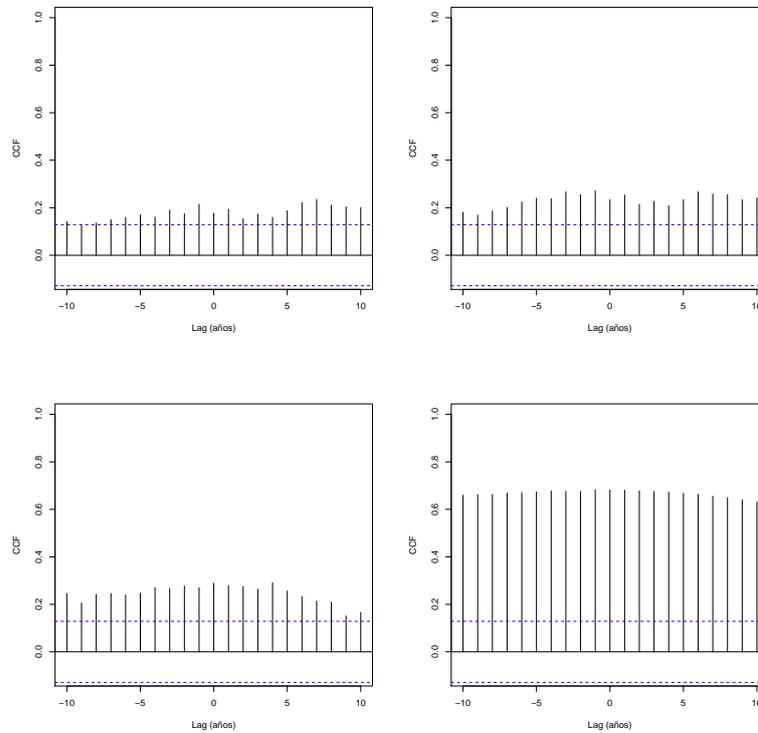


Figura 5.8: Correlación cruzada entre las variables climáticas y las capturas de atún rojo para el periodo 1525-1756. La figura superior izquierda es la CCF utilizando la variable TA_s , la superior derecha la variable SST_s , la inferior izquierda la variable TA_r y la inferior derecha la variable GHG . Las líneas horizontales discontinuas indican el intervalo de confianza al 95 %.

(GHG) son la única variable significativa ($p\text{-value} < 0.0001$). Para el modelo GLMIII las dos variables climáticas (SST_s y TA_r) son significativas ($p\text{-value} < 0.0001$) y para el GLMIV también las dos variables climáticas (SST_s y TA_s) son significativas ($p\text{-value} < 0.0001$). Mientras que, para todos los casos, las tres variables (t, t^2, t^3) de los efectos temporales son significativas ($p\text{-value}$ entre 0.0284 y 0.039). Por último, nótese que debido a la presencia de los GHG (Figura 5.9 superior izquierda y derecha) en los modelos lineales generalizados se obtiene un ajuste mucho más adecuado que para el caso cuando los GHG no se tienen en cuenta (Figura 5.9 inferior izquierda y derecha).

Tabla 5.2: Porcentajes de varianza explicada de las capturas utilizando las variables climáticas del modelo GLMI.

Variable	Términos	Term. selec.	p-value	Var. exp.(%)
Clima	GHG SST _s TA _s	GHG	< .0001	11.6
Temporal	t t^2 t^3	t t^2 t^3	.0389 .0331 .0284	13.0
Ambas				35.0
No explicada				40.4
				$\sum = 100$ %

Tabla 5.3: Porcentajes de varianza explicada de las capturas utilizando las variables climáticas del modelo GLMII. Tomada de Ganzedo *et al.* [102].

Variable	Términos	Term. selec.	p-value	Var. exp.(%)
Clima	GHG SST _s TA _r	GHG	< .0001	12.3
Temporal	t t^2 t^3	t t^2 t^3	.0389 .0331 .0284	12.3
Ambas				35.7
No explicada				39.7
				$\sum = 100$ %

Tabla 5.4: Porcentajes de varianza explicada de las capturas utilizando las variables climáticas del modelo GLMIII. Tomada de Ganzedo *et al.* [102].

Variable	Términos	Term. selec.	p-value	Var. exp.(%)
Clima	SST _s	SST _s	< .0001	2.1
	TA _r	TA _r	< .0001	
Temporal	t	t	.0389	38.8
	t^2	t^2	.0331	
	t^3	t^3	.0284	
Ambas				9.1
No explicada				50.0
				$\Sigma = 100$ %

Tabla 5.5: Porcentajes de varianza explicada de las capturas utilizando las variables climáticas del modelo GLMIV. Tomada de Ganzedo *et al.* [102].

Variable	Términos	Térm. selec.	p-value	Var. exp.(%)
Clima	SST _s	SST _s	< .0001	1.4
	TA _s	TA _s	< .0001	
Temporal	t	t	.0389	43.2
	t^2	t^2	.0331	
	t^3	t^3	.0284	
Ambas				4.8
No explicada				50.6
				$\Sigma = 100$ %

5.4. Discusión

La serie temporal de capturas (Figura 5.5) de atún rojo presenta fuertes fluctuaciones y tiene una clara tendencia a decaer. Por un lado, las fluctuaciones, podrían deberse, en parte, a eventos de tipo socio-económico como guerras, epidemias, crisis económicas, destrucción o mal funcionamiento de las almadrabas y por ello, es posible que hayan tenido una menor capacidad de capturas debido a causas socio-económicas. Ganzedo *et al.* [102] hacen una amplia revisión histórica (para el periodo 1525-1756) sobre este tipo de eventos y encontraron que sólo

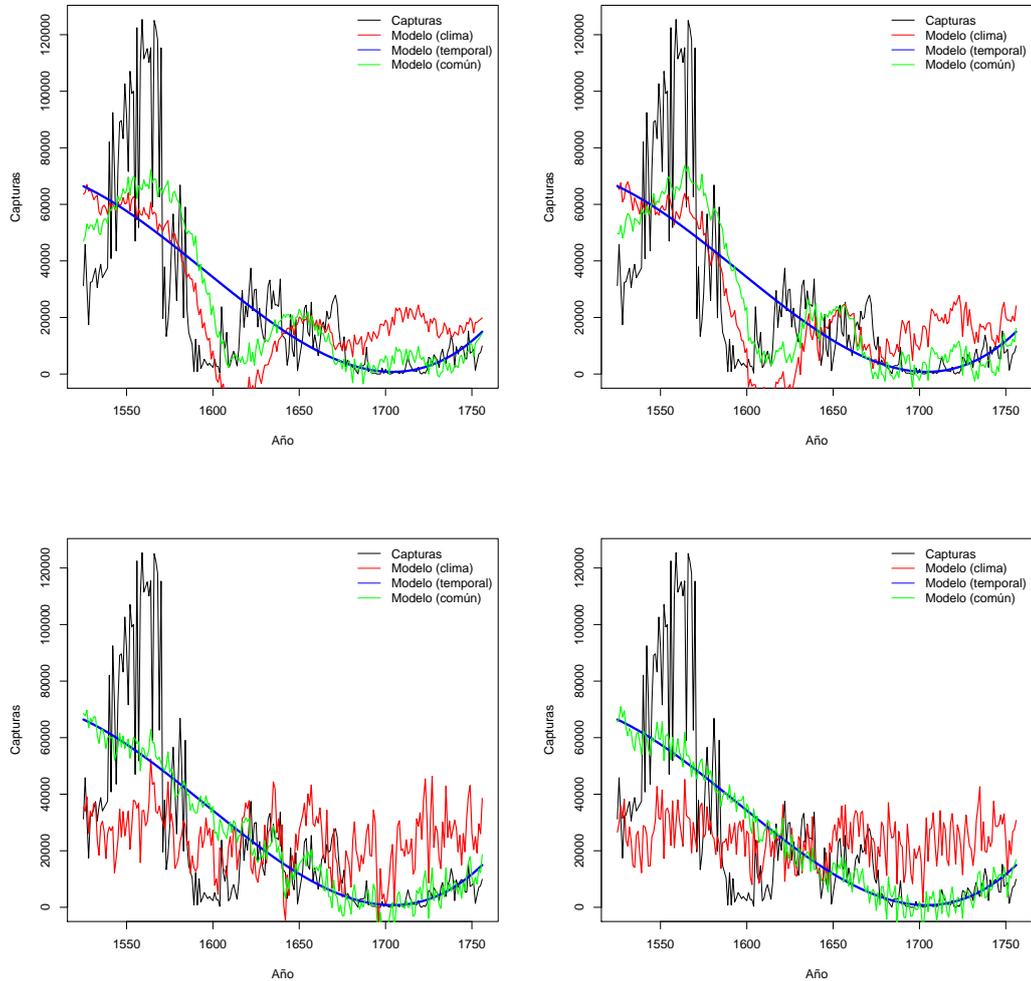


Figura 5.9: Serie temporal de las capturas de atún rojo (curvas en negro) y los valores predichos por los GLM debido a los efectos de las variables climáticas (curvas en rojo), de las variables temporales (curvas en azul) y el efecto combinado de las variables climáticas y las temporales (curvas en verde). La figura superior izquierda corresponde a las variables climáticas ajustadas en el GLMI, la superior derecha al GLMII, la inferior izquierda al GLMIII y la inferior derecha al GLMIV.

unos cuantos años se vieron afectados por estos factores. Por tanto, estos factores no parecen explicar todas las fluctuaciones de la serie de capturas.

Por otro lado, el decremento de las capturas podría deberse a una sobreexplotación, cambios de rutas de migración o a cambios medioambientales. La no

extinción del atún rojo, al menos durante el periodo de estudio, no señala a la sobreexplotación¹ como su principal causa. No hay evidencia de un cambio significativo de las rutas de migración del atún rojo [66, 99, 281], por tanto, tampoco parece ser ésta la causa de esta tendencia. Sin embargo, datos históricos, medidas instrumentales, reconstrucciones de variables medioambientales para áreas geográficas cercanas y Hemisferio Norte, durante 1525-1756, muestran fluctuaciones y un decaimiento en las temperaturas de la superficie del mar (SST) y del aire (TA), y de los gases de efecto invernadero (GHG) (Figuras 5.6 y 5.7) [62, 179, 201, 202, 391].

Durante el periodo 1640-1715, en Europa y Atlántico Norte, ocurrió un evento climático frío². Existen claras evidencias de ello, como el registro de medidas de temperaturas [38], el crecimiento de glaciares de las montañas [269] o el congelamiento ‘anormal’ de algunos ríos y lagos (*v. gr.*, el Támesis) [201]. Un decremento de las SST y de los GHG pudo haber afectado el metabolismo del atún rojo, la disponibilidad de alimento y la producción primaria, y por tanto la supervivencia de las larvas de esta especie [51, 230], resultando en un impacto negativo en las capturas de atún rojo. Estudios en otras áreas del Mediterráneo (sur de Italia y norte de África) y para la misma época, también registraron cambios negativos en sus pesquerías [95, 98, 281, 282].

En base a la anterior línea de argumentación y con los resultados obtenidos del análisis de correlaciones y del análisis de devianza es posible sugerir que las variables climáticas, principalmente los GHG y, en menor medida las temperaturas del océano y del aire al nivel de mar, jugaron un papel en la disminución y en las fluctuaciones de las capturas del atún rojo. Cabe destacar que una de las razones de la menor influencia de las temperaturas (del océano y del aire) en las capturas podría deberse a las limitaciones de las temperaturas, pues provienen de simulaciones y reconstrucciones climáticas y no reproducen adecuadamente la variabilidad climática de alta frecuencia, en especial en las temperaturas TA_s y SST_s , que provienen de las simulaciones climáticas.

Los resultados del análisis de la devianza de las capturas de atún rojo mediante los modelos lineales generalizados (Tablas 5.2-5.5) muestran que a lo sumo

¹Hasta principios del siglo XX las técnicas de pesca no permitían una sobreexplotación [96].

²A éste período se le conoce como la Pequeña Edad de Hielo [201].

las variables climáticas sólo explican hasta un 12.3 %, mientras que la dinámica poblacional (representada por los efectos temporales) explica hasta un 43.2 % de la varianza total de la capturas. Por otro lado, la combinación de los efectos climáticos con los temporales explica hasta un 35.7 %. Sin embargo, esta proporción de varianza contiene una combinación de los efectos climáticos y de los efectos debidos a dinámica poblacional de la especie y no es posible determinar qué porcentaje se debe a cada uno de estos factores [102].

Nótese que en el mejor de los casos sólo se explica un 60.3 % de la varianza de las capturas del atún rojo. Esto podría deberse a varias razones. Por un lado, en este estudio sólo se utilizaron forzamientos radiativos por gases de efecto invernadero (CO_2) y temperaturas del aire y del océano y muy probablemente existieron otras variables climáticas que afectaron de modo negativo (disminuyendo) las capturas en términos de reclutamiento y supervivencia de esta especie. Además, en este estudio no se han tenido en cuenta otros factores de difícil cuantificación que pudieron haber afectado en menor medida a las capturas, tales como las epidemias (*v. gr.*, la epidemia de malaria que ocurrió en 1590), las erupciones volcánicas (*v. gr.*, como las del Huaynaputina en 1600 o el Mount Parker en 1641), los conflictos bélicos (*v. gr.*, la campaña militar anglo-holandesa entre 1703 y 1704 o la guerra de sucesión española entre 1701 y 1713), pérdidas por ataques de piratas o crisis económicas [102].

Capítulo 6

La relación entre variables climáticas y las fluctuaciones de capturas de pulpo común entre 1989-2007 en las Islas Canarias mediante la CWT

6.1. Introducción

El pulpo común (*Octopus vulgaris*) (Figura 6.1) es una especie de molusco cefalópodo y es uno de los principales objetivos de muchas industrias pesqueras en las zonas de afloramiento de la región noroccidental de África [13, 89]. En lo que respecta a las Islas Canarias (Figura 6.2), el área de estudio de este trabajo, es también una de las especies más capturadas por la pesca artesanal, de tal modo que representan entre el 2 y 32 % del total de la biomasa capturada [130, 131].

El *Octopus vulgaris* de las Islas Canarias puede alcanzar una longitud dorsal del manto de unos 400 mm y una longitud total de unos 1600 mm, puede pesar entre 2 y 3 kg (en algunos casos puede alcanzar hasta los 10 kg) tiene un crecimiento muy rápido y puede vivir alrededor de un año [92, 132]. Se alimenta

6.1 Introducción

principalmente de peces, crustáceos y moluscos, aunque el canibalismo tiene un papel importante en su dieta. Se pueden reproducir todo el año, pero tienen bien diferenciados dos periodos de máxima reproducción, enero-julio (pico en abril) y octubre-noviembre, los cuales coinciden con los picos máximos mensuales de la captura por unidad de esfuerzo (CPUE) [130, 131, 132].

Esta relación entre capturas y reproducción es debida, en parte, a que durante los periodos máximos de reproducción, los *Octopus vulgaris* se concentran para

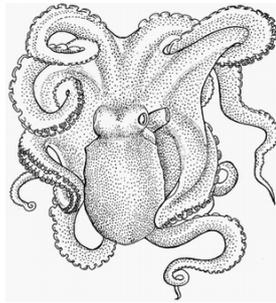


Figura 6.1: *Octopus vulgaris* (pulpo común). Tomado del *Species fact sheets* del Departamento de Acuicultura y Pesquerías de la FAO.

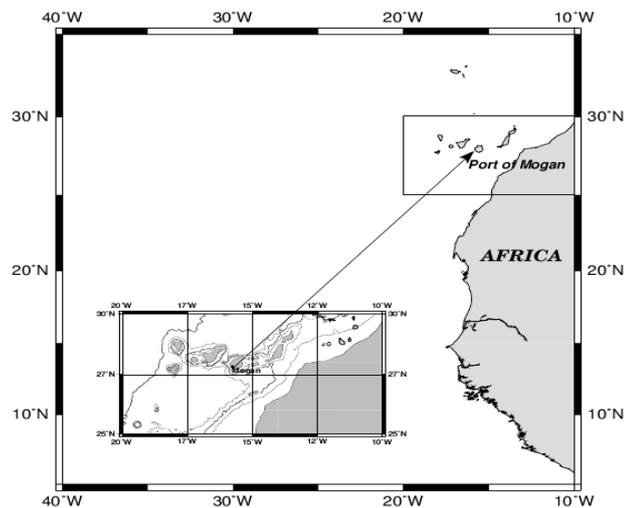


Figura 6.2: Mapa del área de estudio (Islas Canarias). El recuadro indica el área de capturas del *Octopus vulgaris*. Tomado de Polanco *et al.* [275].

la puesta de huevos en aguas someras y, por ello, son objetivos fáciles para la pesca [130]. Sin embargo, el esfuerzo pesquero no es suficiente para justificar estas relaciones. Por estudios realizados en áreas cercanas a las Islas Canarias y en otras áreas geográficas donde se captura *Octopus vulgaris* se sabe que esta especie es sensible a cambios de las variables ambientales, especialmente a la temperatura superficial del mar (SST) [13, 132, 159, 200, 334, 360].

Estudios previos han demostrado que el pulpo común en el inicio de su ciclo de vida (en estado plantónico) es altamente dependiente de la temperatura ambiente [132, 200]. Por otro lado, durante su posterior estadio (béntico), las oscilaciones estacionales de la temperatura del mar lo afectan considerablemente [159]. Sin embargo, aunque es un hecho aceptado que la temperatura del mar es uno de los factores que afectan la abundancia del *Octopus vulgaris*, no siempre es claro qué tipo de correlación existe entre esta variable y las CPUE. Por ejemplo, en un estudio realizado en el Golfo de Cádiz se encontró que los picos máximos de abundancia coinciden con los valores mínimos de la temperatura de la superficie del mar del área de estudio [334]. En el otro extremo, otros estudios han encontrado que las máximas capturas coinciden con los más altos valores de las SST para la costa noroccidental de África (en la zona de pesca conocida como *Sahara bank*) y la costa de Portugal, respectivamente [13, 223, 334].

Para incrementar los escasos y contradictorios resultados de las relaciones entre variabilidad climática y las fluctuaciones de CPUE del *Octopus vulgaris*, en un estudio reciente realizado por Solari [335] en la zona de afloramiento de la región noroccidental de África, se encontró que las fluctuaciones en las capturas del pulpo común para los últimos 50 años están relacionadas con la Oscilación del Atlántico Norte (NAO). En esta misma dirección, Sims y colaboradores [331], en un estudio realizado en el canal de la Mancha, aunque para otra especie de cefalópodo (*Loligo forbesi*), encontraron que la variación temporal de los picos de máxima abundancia de este cefalópodo es dependiente de la temperatura del mar, la cual a su vez está relacionada con la NAO.

En lo que respecta al área de estudio del presente capítulo, los primeros intentos para relacionar las fluctuaciones de las capturas del *Octopus vulgaris* con variables ambientales (climáticas), fueron realizados por Hernández-García *et al.* [130]. Ellos sugirieron que los picos máximos de las capturas (años 1982 y 1989)

podrían estar relacionadas con el fenómeno de ‘El Niño’. Sin embargo, debido a que las series de CPUE eran muy cortas (sólo 8 años de datos), ellos también sugirieron realizar más estudios a medida que aumentara la base de datos.

Un poco más recientemente Hernández-García *et al.* [131] llevaron a cabo un estudio sobre *Octopus vulgaris* de las costas de las Islas Canarias, aunque enfocado a la reproducción de esta especie, encontraron algunas relaciones entre la temperatura superficial del mar y la CPUE. También reportaron la existencia de dos cohortes anuales del pulpo común en las Islas Canarias. La primera cohorte nace en abril y la segunda entre octubre y noviembre. Sin embargo, el instante preciso en que ocurren estos picos fluctúa de año en año.

Como una ampliación del estudio de Hernández-García *et al.* [130] y de Hernández-García *et al.* [131], Caballero-Alfonso *et al.* [45] estudiaron el papel de la variabilidad climática de alta frecuencia (escala estacional) en las fluctuaciones de las capturas (obtenidas por medio de la pesca artesanal) de *Octopus vulgaris* durante 1989-2007. Ellos encontraron una correlación inversa entre la abundancia (medidas como CPUE) y la SST y una relación directa entre la abundancia y la NAO. Al utilizar un modelo lineal, encontraron que durante la primavera la NAO es la principal variable climática que afecta la abundancia de las capturas (explica hasta un 28.64 % de la variabilidad de las CPUE). Por otro lado, durante el otoño, la SST es la variable más importante y significativa, aunque la SST y la NAO tienen casi la misma influencia sobre la abundancia de las capturas (34.2 % y 31.13 %, respectivamente).

Sin lugar a dudas, el trabajo de Caballero-Alfonso *et al.* [45] ayudó a entender un poco más la relación entre la variabilidad climática y las fluctuaciones de las capturas del pulpo común. Sin embargo, ese estudio se limita a la escala estacional. En parte, estas limitaciones son debidas a las técnicas estadísticas utilizadas. Por un lado, los modelos de regresión lineal no son capaces de proporcionar los intervalos en el dominio del tiempo y de la frecuencia donde existen relaciones significativas entre las variables climáticas (como la NAO) y las CPUE. Por otro lado, estas técnicas estadísticas asumen que tanto las series temporales como los fenómenos analizados, son estacionarios y lineales.

Una aproximación simple para estudiar los fenómenos ambientales y ecosistémicos es considerarlos lineales y estacionarios, aunque la no linealidad y la

no estacionariedad son más bien la regla que la excepción [142, 300]. El principal requisito de muchas herramientas estadísticas (*v. gr.*, el análisis espectral vía la transformada de Fourier) es asumir que las series temporales bajo estudio satisfacen la propiedad estadística de la estacionariedad [279, 365, 371]. Sin embargo, las series temporales ambientales y ecosistémicas no siempre satisfacen este requerimiento, y la evidencia creciente indica la necesidad de considerar la importancia de la dinámica de los sistemas que contienen componentes fugaces (de corta duración) en el estudio de los procesos/fenómenos ambientales y ecosistémicos [50, 142, 300].

La dinámica de sistemas que contiene elementos fugaces puede jugar un papel importante en la estructura de los sistemas naturales (ambientales) [22, 125]. Estudios recientes han mostrado que la dinámica de poblaciones puede variar de un estado a otro en escalas multidecadales [50]. Sin embargo, las perturbaciones ambientales no son los únicos mecanismos capaces de activar componentes fugaces. La mezcla de la dinámica no lineal con procesos dinámicos inestables puede aumentar la complejidad de estos sistemas y crear fenómenos de corta duración [50, 67].

El análisis espectral mediante la transformada continua de Wavelet es una herramienta matemática adecuada para enfrentar el problema de la no estacionariedad. Esto es así porque realiza una descomposición espectral localizada de la señal, determinando los modos de variabilidad dominantes y cómo éstos varían en el tiempo y la escala [114, 350]. Esta técnica estadística incluye la estimación del poder espectral (análisis univariante), análisis espectral cruzado, la coherencia y la fase de la coherencia (análisis bivariante) [113, 206, 207, 350]. El análisis espectral de Wavelet ha sido utilizado en diferentes campos de la ciencia, como en la dinámica de fluidos, la climatología, la meteorología, la geofísica o la ecología [50, 88, 113, 177, 350].

Como una extensión del trabajo de Caballero-Alfonso *et al.* [45], Polanco *et al.* [275] estudian el papel de la variabilidad climática de baja frecuencia en las fluctuaciones de las capturas (CPUE) del *Octopus vulgaris*. Para tal fin, se analizan las relaciones (en términos de coherencia espectral de Wavelet) entre las variables climáticas (SST y NAO) y las CPUE. Se asume que los cefalópodos debido a su corto periodo de vida (alrededor de un año) [92, 132, 200] son buenos indicadores

de la variabilidad climática [45, 270, 331, 335]. En especies de vida corta, el efecto de la variabilidad climática podría reflejarse en su biomasa del siguiente año/mes [104]. De esta forma, el reclutamiento anual [290], y en consecuencia la disponibilidad en la pesquería [54] están condicionados por los efectos que la variabilidad climática tiene sobre su largo estadio pelágico paralarvae en comparación con su longevidad [368]. Por ello, es muy probable que aquellas variables climáticas que pueden ser medidas en la superficie del mar (como las SST) puedan tener efectos directos sobre la supervivencia durante su estado pelágico. Esta conexión podría ser explicada a través del crecimiento, la disponibilidad de alimento o la mortalidad natural debida a otras causas [48, 93, 187, 270, 290, 368]. Por estas razones, podría ser más factible diferenciar los efectos climáticos de la influencia de la pesca en esas especies, porque su respuesta a la fluctuaciones ambientales podría ser más rápida [45, 54, 131].

Uno de los objetivos de este capítulo es determinar el papel de la variabilidad climática de alta y baja frecuencia en las fluctuaciones de las capturas (CPUE) del *Octopus vulgaris* de las Islas Canarias. El otro objetivo es averiguar el grado con que las variables ambientales (SST y NAO) y CPUE están relacionadas en diferentes escalas, tanto en el dominio del tiempo como de la frecuencia, haciendo énfasis en las variaciones de baja frecuencia.

6.2. Datos y preprocesamiento

6.2.1. Serie temporal de capturas del *Octopus vulgaris*

La serie temporal de las capturas (medidas como CPUE) mensuales para el periodo 1989-2007 (Figura 6.3 superior) utilizada en este trabajo proviene de Caballero-Alfonso *et al.* [45]. A su vez, esas capturas (obtenidas por medio de técnicas de pesquería artesanal) provienen de capturas diarias registradas por una pesquería en el Suroeste de las Islas Canarias, principalmente del puerto de Mogán [45, 110, 130]. El puerto de Mogán es representativo de todas las Islas, porque es uno de los puertos donde existe mayor actividad pesquera en el archipiélago, tanto en pesca potencial como en el número de capturas. Esta actividad pesquera representa más de un 25 % del total de las capturas anuales de especies bénticas

y peces demersales, y hasta un 10% si son consideradas las capturas de atún y caballa (Gobierno de Canarias, datos no publicados) [45].

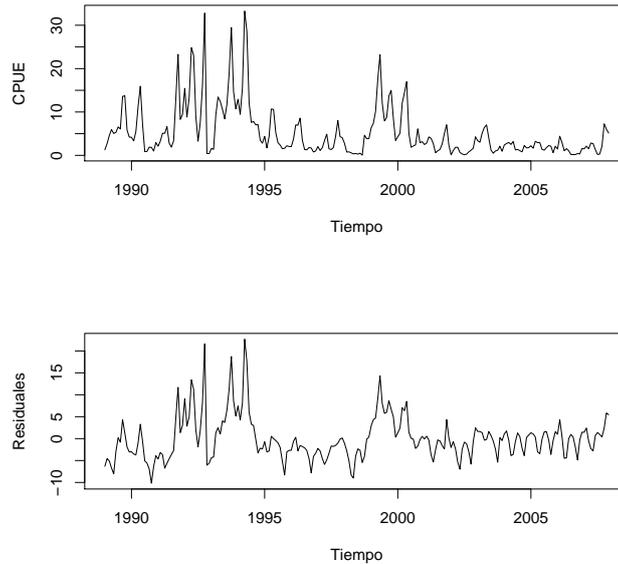


Figura 6.3: La figura superior de este panel corresponde a la serie temporal de las capturas mensuales (228 datos) medidas en CPUE. La inferior son los residuos de la CPUE después de remover la tendencia lineal.

Es importante tener en cuenta que las pesquerías no excluyen ninguna especie en sus registros, por tanto, las capturas son un buen indicador de abundancia [45]. De acuerdo con Caballero-Alfonso *et al.* [45], el esfuerzo pesquero del *Octopus vulgaris* no ha cambiado significativamente en el tiempo durante el periodo 1989-2007. Sin embargo, las capturas reflejan fluctuaciones diferentes al ciclo de vida estacional. Por esta razón, otros factores, como los climáticos, podrían jugar un papel clave en las fluctuaciones del stock.

Para calcular las capturas por unidad de esfuerzo del *Octopus vulgaris* se siguió el procedimiento descrito por Hernández-García *et al.* [130]. Debido a los objetivos de este trabajo (*i. e.*, estudiar la variación climática de alta y baja frecuencia), antes de aplicar las técnicas estadísticas, se realizó un preprocesamiento de la serie mensual de CPUE. Para el estudio de alta frecuencia se

utilizaron anomalías estacionales debido a la existencia de estacionalidad en las CPUE (Figura 6.4) [45]. Para estudiar la variabilidad de baja frecuencia se utilizaron anomalías mensuales de CPUE tras eliminar la tendencia lineal (Figura 6.3 inferior) [275].

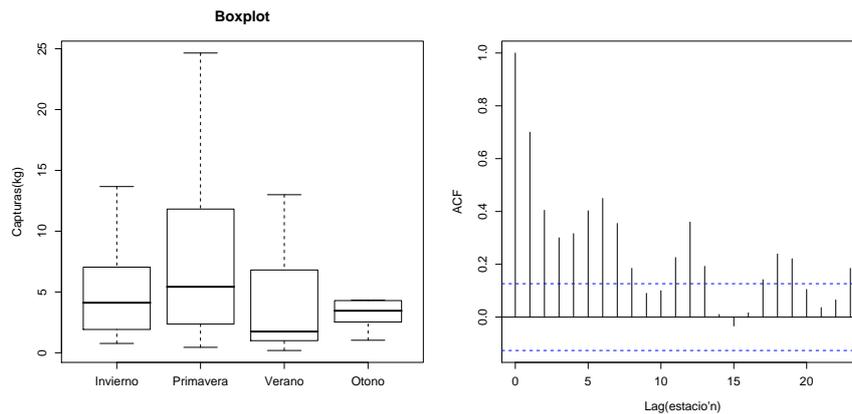


Figura 6.4: Diagrama de caja (izquierda) y función de autocorrelación (ACF) de las CPUE estacionales para el periodo 1989-2007. Tomado de Caballero-Alfonso *et al.* [45].

6.2.2. Datos climáticos

La Oscilación del Atlántico Norte es una de las principales y recurrentes características de la variabilidad climática de medias y altas latitudes en el Hemisferio Norte, de modo especial durante los meses de las estaciones frías. Esta oscilación produce grandes cambios en la velocidad y dirección media del viento sobre el Atlántico, en el transporte de humedad y calor entre el Atlántico y los continentes cercanos, y en la intensidad y el número de tormentas y sus trayectorias [145]. La NAO es un dipolo Norte-Sur de anomalías del campo de presión atmosférica a nivel del mar. Este dipolo lleva a la alternancia de masa atmosférica entre el Ártico (Norte) y el Atlántico subtropical (Sur). En el patrón NAO, uno de los polos que forman parte del dipolo tiene su centro sobre Groenlandia y el otro (de signo opuesto) cubre latitudes centrales del Atlántico Norte, entre las Azores y Portugal (véase Figura 1.3) [145, 342].

6.2 Datos y preprocesamiento

En lo que respecta al área de estudio de este trabajo, las Islas Canarias, están ubicadas muy al límite (al sur) de la influencia de la NAO, por lo cual el impacto sobre el ambiente en las Islas no siempre es fácilmente detectable [103]. Sin embargo, en algunos casos y para ciertos periodos, la influencia de la NAO puede ser lo suficientemente fuerte [45, 103, 275].

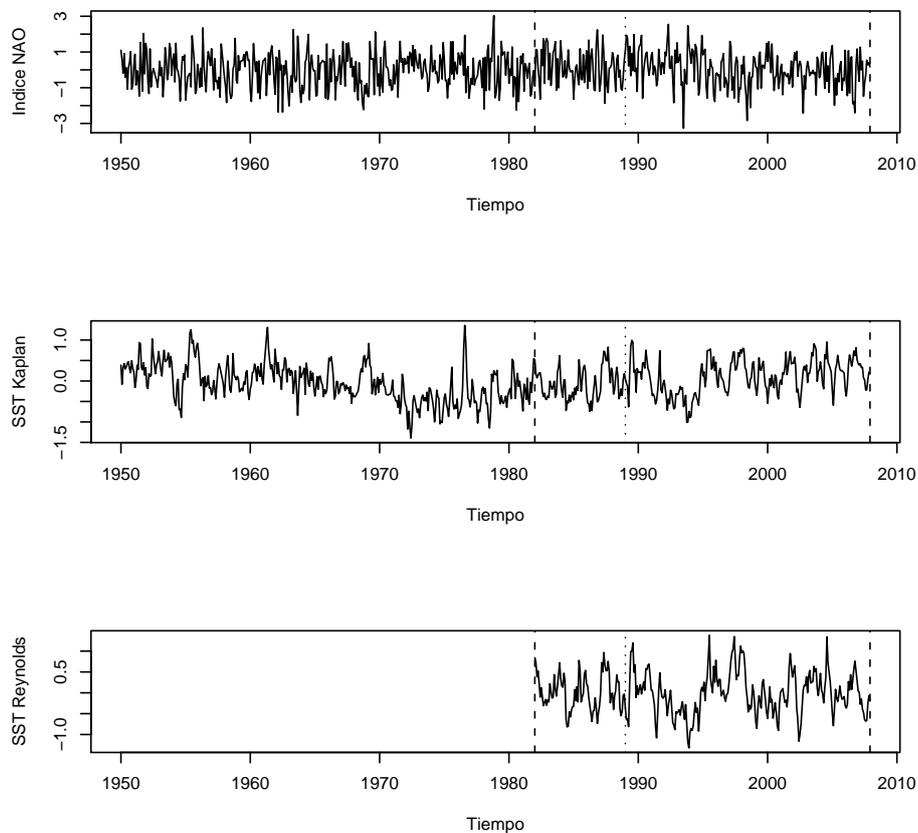


Figura 6.5: Series temporales climáticas. Índice NAO para el periodo 1950-2007, anomalías mensuales de SST (datos de Kaplan *et al.* [158]) para el periodo 1950-2007 y anomalías mensuales de SST (datos de Reynolds *et al.* [284] para el periodo 1982-2007. Ambas series de SST pertenecen al sur de las Islas Canarias (28.5°N/16.5°O). Las líneas verticales discontinuas indican el periodo 1982-2007. La línea vertical punteada y la última línea discontinua indican el periodo 1989-2007.

Para estudiar la influencia de las variables climáticas (NAO y SST) en las altas frecuencias (escala estacional), se utilizaron los siguientes conjuntos de datos. La serie temporal del índice de la Oscilación del Atlántico Norte (NAO) para el periodo 1950-2007 (Figura 6.5 superior) fue obtenida del servidor web de la NOAA (*National Oceanic and Atmospheric Administration*) [243]. Se utilizaron dos conjuntos de datos de anomalías mensuales de temperatura de la superficie del mar (SST) correspondientes al sur de Gran Canaria (28.5°N/ 16.5°O). La base de datos de Kaplan *et al.* v2 extendida [158] para el periodo 1950-2007 (Figura 6.5 central) fue obtenida del servidor web de la NOAA [242]). La base de datos de Reynolds *et al.* v.2 [284] para el periodo 1982-2007 (este conjunto de datos sólo está disponible a partir del 12/1981) (Figura 6.5 inferior) fue obtenida del servidor web de la NOAA [244]. Del mismo modo que con las CPUE, en el análisis estadístico se utilizaron anomalías estacionales de NAO y SST y se eliminó la tendencia lineal de todos estos conjuntos de datos [45].

Para estudiar la variabilidad en las bajas frecuencias, se utilizaron los mismos conjuntos de datos de CPUE, NAO y SST (datos de Kaplan *et al.* v2 extendida [158]) periodo 1989-2007, aunque para estudiar la relación entre la NAO y las SST se utilizó un periodo más amplio (1950-2007) que los utilizados en escala estacional. Se tiene preferencia por los datos de SST de Kaplan *et al.* v2 extendida [158] porque cubren un periodo mayor que los datos de Reynolds *et al.* v.2 [284]. Para hacer los análisis estadísticos se utilizaron anomalías mensuales de CPUE, de NAO y de las SST y se eliminó la tendencia lineal de todos estos conjuntos de datos [275].

6.3. Metodología

La relación entre la CPUE y las variables climáticas NAO y SST (datos de Kaplan *et al.* [158]) se estudió a través de los valores tomados por el coeficiente de correlación de Pearson [376]. Los coeficientes de correlación fueron verificados contra la hipótesis nula (*i. e.*, coeficiente de correlación igual a 0) con un 95 % de confianza. Sin embargo, debido a la existencia de autocorrelación en las series estacionales de CPUE, el test de significación estadística para el coeficiente de correlación de Pearson no tiene una precisión adecuada [84, 185].

Para resolver este inconveniente se consideró la reducción de los grados de libertad debido a la autocorrelación de las series estacionales de CPUE, SST y NAO. Para ello se utilizaron simulaciones de Monte Carlo, de tal modo que se construyeron 500,000 realizaciones de procesos AR1 con el correspondiente valor de la autocorrelación y la longitud (número de elementos) para cada una de las series estacionales. Se calcularon los coeficientes de autocorrelación de las series sintéticas AR1 y se obtuvo un histograma numérico (distribución experimental) (Figura 6.6) de estos coeficientes. Los valores de los coeficientes de correlación por debajo (o por arriba) del percentil 2.5 % (97.5 %) en la distribución experimental fueron considerados significativos [45].

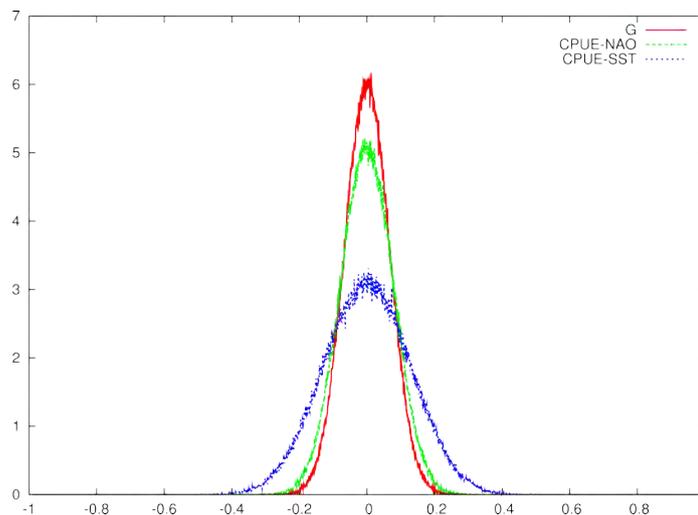


Figura 6.6: Histograma numérico de los coeficientes de autocorrelación obtenidos por medio de simulaciones de Monte Carlo. En rojo la distribución (Monte-Carlo) de los coeficientes de correlación para series gaussianas. En verde y azul son las distribuciones para NAO y SST, respectivamente, considerando sus coeficientes de autocorrelación. El eje de las abscisas determina los nuevos valores de los intervalos de confianza.

Para estudiar la distribución estacional del efecto de la NAO con respecto al SST (datos de Reynolds *et al.* [284]), se utilizó un mapa de correlación de Pearson [21, 365] con isóneas a un 95 % (0.05) y 99 % (0.01) de confianza (Figura 6.7). El área geográfica cubierta por estos mapas está delimitada por el cuadrante 20-50°N y 45°O a 20°E con una resolución espacial de un grado.

Se utilizó un modelo de regresión lineal múltiple (lm) [358] para estimar la relación entre las capturas CPUE (y) y las variables climáticas x_1 y x_2 , para SST (datos de Kaplan *et al.* [158]) y la NAO, respectivamente (donde x_1 y x_2 son las variables independientes e y la dependiente). Los modelos de regresión lineal contienen un término de error (variable aleatoria) añadido (ϵ), por lo que el modelo lineal utilizado tiene la forma $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ (donde y es la variable a explicar y x_1, x_2 las variables explicativas) [379]. Los parámetros β_0, β_1 y β_2 , son los coeficientes de regresión. El coeficiente de determinación (R^2) está definido [379] como

$$R^2 = 1 - \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \bar{y})^2} \quad (6.1)$$

donde \hat{y} es el término predicho por el modelo de regresión (las capturas CPUE) e \bar{y} es la media de las CPUE.

Para averiguar si la varianza explicada aumenta si ambas variables explicativas son consideradas al mismo tiempo o de modo independiente, se utilizaron 3 modelos lineales: 1) SST y NAO, 2) sólo SST y 3) sólo NAO. Se verificaron los residuos de los modelos ajustados para saber si están normalmente distribuidos ($p\text{-value} > \alpha = 0.05$) y si presentan autocorrelación. Para esto, se aplicó el test de Shapiro-Wilk [327] y el test de Durbin-Watson [82, 83], respectivamente.

Para averiguar la relación entre las variables climáticas NAO y SST (datos de Kaplan *et al.* [158]) con las CPUE para el periodo 1989-2007, y la relación entre SST y la NAO para el periodo 1950-2007, se utilizaron dos técnicas estadísticas. La primera es la función de correlación cruzada (CCF) [330], calculada por medio del estimador de la función de correlación cruzada muestral mediante la función `ccf` del paquete `Stat` en R [280, 358].

La segunda técnica estadística utilizada es la coherencia de Wavelet normalizada (WCO), que es una herramienta adecuada para analizar la relación con respecto al tiempo y la escala (frecuencia) entre dos series temporales $x(t_i)$ e $y(t_i)$ [50, 113, 206]. La WCO está definida como la amplitud del espectro cruzado de Wavelet normalizado por el espectro de Wavelet de cada serie temporal $x(t_i)$ e $y(t_i)$ [50, 206], tal y como se describe en la sección 2.4.4 de esta memoria. La WCO puede tomar valores entre 0 y 1. Por un lado, un valor de 0 indica que

no hay relación entre las series temporales en las escalas-tiempo consideradas. En el otro extremo, un valor de 1 indica que existe una perfecta relación lineal [113, 206]. La fase de la coherencia [50] proporciona información acerca del retraso en la relación de dos series temporales, *v. gr.*, si las series están en fase o fuera de fase (ver los detalles técnicos de esta metodología en [50, 206] y en la sección 2.4.4 de esta memoria).

Para calcular la WCO se utilizó el paquete **SOWAS** en R. Este paquete se basa en la metodología de Maraun y Kurths [206] y de Maraun *et al.* [207]. Como función base de Wavelet **SOWAS** utiliza el Wavelet de Morlet. Es importante tener en cuenta que la selección de la función base puede influir el tiempo, la escala y la resolución de la frecuencia en la descomposición de una serie temporal [50, 350]. Sin embargo, el Wavelet de Morlet proporciona un buen balance entre la localización escala (frecuencia)-tiempo [113, 220] y es una de las mejores funciones base en términos de descomposición de la frecuencia de una señal [164]. Ha sido utilizado con éxito en series temporales ambientales, tales como SST, NAO o ENSO [206, 350, 351] y en series ecológicas, tales como capturas de atún [218, 301], de pez espada [60] o de anchoa [144]. Sin embargo, hasta donde sabemos no ha sido utilizado para estudiar la relación de variables ambientales con variaciones de capturas de cefalópodos.

Se aplicó un test de significación estadística de los resultados de la WCO [206, 207]. El primer requisito para aplicar un test de significación es determinar (escoger) un espectro de fondo adecuado. Para muchas de las series temporales ambientales (ecológicas, climáticas, hidrológicas, etc.) un proceso de ruido rojo (AR1) es un modelo razonable [124, 350, 357], aunque pueden existir series temporales ecológicas que tienen una estructura en su autocorrelación que un proceso AR1 no puede describir de modo consistente [65, 150]. Por estas razones, han sido desarrollados varios modelos alternativos de ruido de fondo, como los *surrogates* del tipo I y II o los *beta surrogates* [301]. Sin embargo, el uso de diferentes modelos de ruido de fondo cuando las series son cortas y ruidosas lleva a conclusiones similares [50]. El test de significación utilizado para las WCO en este trabajo se describe en la sección 2.4.5 de esta memoria.

6.4. Resultados

El resultado de aplicar los mapas de correlación de Pearson (Figura 6.7) entre la SST y el índice NAO para el área de estudio, indica una alta y significativa correlación en primavera y otoño. Correlaciones significativas fueron encontradas en la coordenada 28.5°N/16.5°W, en primavera ($r=-0.47$, $p < 0.05$) y en otoño ($r=-0.48$, $p < 0.05$). No se encontraron correlaciones significativas para el invierno ($r=0.16$, $p > 0.05$) y el verano ($r=0.36$, $p > 0.05$).

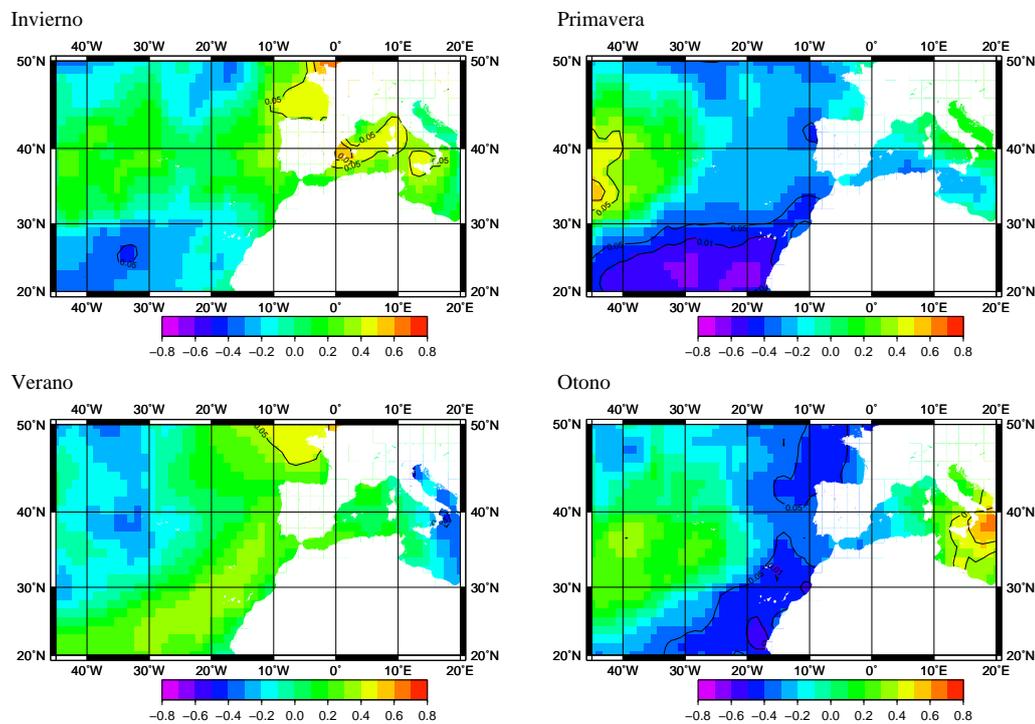


Figura 6.7: Coeficiente de correlación entre el índice NAO y las SST para cada estación del año. Las líneas de contorno representan el p -value a 0.01 (99%) y a 0.05 (95%). Tomado de Caballero-Alfonso *et al.* [45].

Las correlaciones entre las series estacionales de CPUE y la SST se muestran en la tabla 6.1. En esta misma tabla 6.1 se puede apreciar que la relación entre las CPUE y la NAO depende de la estación del año. Para invierno y verano existe una correlación inversa, mientras que para primavera y otoño existe una correlación directa.

6.4 Resultados

	SST(Inv)	SST(Prim)	SST(Ver)	SST(Oto)
CPUE(Inv)	-0.64**			
CPUE(Prim)		-0.4		
CPUE(Ver)			-0.34	
CPUE(Oto)				-0.61**
	NAO(Inv)	NAO(Prim)	NAO(Ver)	NAO(Oto)
CPUE(Inv)	-0.11			
CPUE(Prim)		0.57*		
CPUE(Ver)			-0.37	
CPUE(Oto)				0.59**

Tabla 6.1: Coeficientes de correlación entre CPUE y las variables climáticas (SST y NAO) (** y * indican nivel de significación al 0.01 y 0.05, respectivamente). Tomado de Caballero-Alfonso *et al.* [45].

La tabla 6.2 muestra el resultado de los 3 modelos lineales utilizados: **1m1**: SST-NAO, **1m2**: sólo SST y **1m3**: sólo NAO. En primavera la NAO es la variable que más influencia tiene en la variabilidad de las CPUE, ya sea considerada en conjunto con la SST, modelo **1m1** (24.5 %) o sola, modelo **1m3** (28.65 %). En otoño, la combinación de la NAO y SST, modelo **1m1**, es la que más varianza explica (42.71 %) de las CPUE, lo cual indica colinealidad entre ellas. La SST por si sola, modelo **1m2**, explica un 34.21 %, mientras que la NAO por si sola, modelo **1m3**, explica un 31.13 %. Nótese cómo el porcentaje de varianza explicada por solo la NAO (31.13 %), modelo **1m3**, es relativamente cercano al explicado por la SST solamente (34.21 %), modelo **1m2**, aunque la SST parece tener una influencia un poco mayor.

En la tabla 6.3 se muestran los resultados del análisis de los residuos de los 3 modelos lineales utilizados para las estaciones primavera y otoño. Como se puede apreciar, tanto para primavera como para otoño, los residuales tienen una distribución de tipo normal (test de Shapiro-Wilk) y no presentan autocorrelación ($p > 0.05$) (test de Durbin-Watson).

Término del modelo seleccionado	p-value	R^2 (%)
Término climático (Primavera)		
1m1 SST NAO	0.04(.)	24.5
1m2 SST	0.09(.)	11.12
1m3 NAO	0.01(*)	28.64
Término climático (Otoño)		
1m1 SST NAO	0.0045(.)	42.71
1m2 SST	0.0050(**)	34.21
1m3 NAO	0.0077(**)	31.13

Tabla 6.2: Porcentaje de varianza de las series estacionales de CPUE para primavera y otoño utilizando un modelo lineal (**, * y . indican niveles de significación al 0.01, 0.05 y 0.1, respectivamente). Tomado de Caballero-Alfonso *et al.* [45].

	S-W test		D-W test		
	w	p-value	Autocor.	D-W stat.	p-value
Térm. clim. (Prim.)					
1m1 SST-NAO	0.96	0.51	0.21	1.40	0.19
1m2 SST	0.95	0.39	-0.15	2.22	0.60
1m3 NAO	0.96	0.99	0.25	1.31	0.12
Térm. clim. (Oto.)					
1m1 SST-NAO	0.99	0.99	-0.42	2.83	0.07
1m2 SST	0.97	0.76	-0.20	2.38	0.44
1m3 NAO	0.99	0.99	-0.35	2.67	0.11

Tabla 6.3: Verificación de los residuos de los modelos lineales. S-W corresponde al test de Shapiro-Wilk y D-W al test de Durbin-Watson. Para el test de S-W la hipótesis nula H_0 es: los residuos se distribuyen normalmente. Para el test de D-W la hipótesis nula H_0 es: no existe correlación. Tomado de Caballero-Alfonso *et al.* [45].

La correlación cruzada entre los valores mensuales de las CPUE y la NAO es significativamente diferente de cero en varios *lags* (Figura 6.8 superior). Especial interés se tiene en el *lag* -1 (un mes) y *lag* -2 (dos meses) del índice NAO, con valores positivos de 0.137 ($p < 0.05$) y 0.162 ($p < 0.05$), respectivamente. Nótese que la correlación cruzada no es significativa en *lag* 0. Aunque otros *lags* positivos están presentes debido a que su interpretación no resulta posible en términos de causalidad (las capturas no afectan al índice NAO), son ignorados como falsas alarmas del test. La correlación cruzada entre los valores mensuales de CPUE y la SST (en *lags* menores a un año), es negativa y presenta una correlación cruzada fuertemente significativa hasta los doce meses (Figura 6.8 inferior), lo cual es en parte una manifestación de la persistencia de la SST y la CPUE.

La correlación cruzada entre el índice NAO y la SST para el periodo enero de 1950 a diciembre de 2007 (Figura 6.9 superior), es significativamente diferente de cero en el *lag* 0 (-0.139, $p < 0.01$). La correlación cruzada entre el índice NAO y la SST para el periodo enero de 1989 a diciembre de 2007 es también significativo en el *lag* -1 (-0.156 $p < 0.05$) (Figura 6.9 inferior).

La WCO y la fase entre CPUE y el índice NAO para el periodo enero de 1989 a diciembre de 2007 se muestran en la Figura 6.10. La WCO y la fase entre CPUE y SST (datos de Kaplan *et al.* [158]) para el mismo periodo temporal se muestran en la Figura 6.11. La WCO y la fase entre NAO y SST (datos de Kaplan *et al.* [158]) para el mismo periodo temporal se muestran en la Figura 6.12. La WCO y la fase entre NAO y SST (datos de Kaplan *et al.* [158]) para el periodo enero de 1950 a diciembre de 2007 se muestran en la Figura 6.13. En todos los casos para la WCO se utilizaron 4 configuraciones diferentes de los parámetros de suavizado (escala y tiempo) para averiguar la posible existencia de puntos espectrales de las WCO que pudieran tener un origen espurio. Los valores α indican los niveles de significación para cada una de las configuraciones de los parámetros de suavizado.

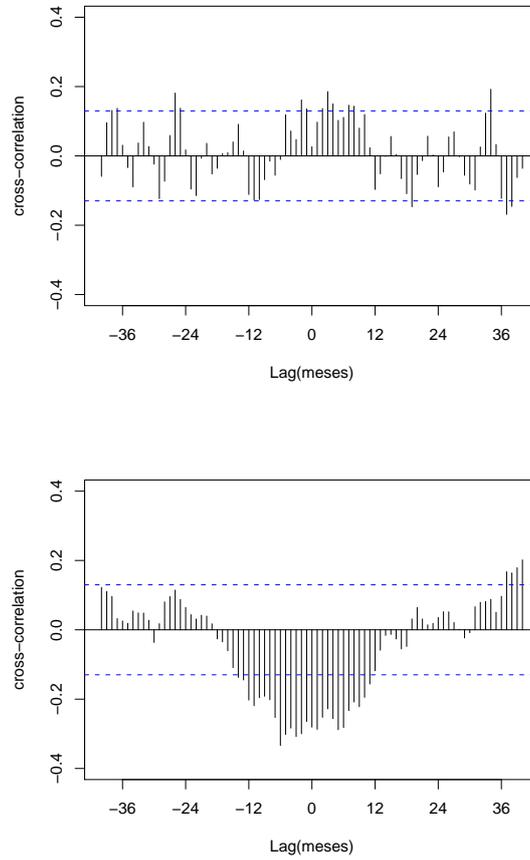


Figura 6.8: Correlación cruzada entre CPUE y NAO (panel superior) y entre CPUE y SST (datos de Kaplan *et al.* [158]) (panel inferior). Las series cubren el periodo de enero de 1989 a diciembre de 2007. Las líneas horizontales discontinuas indican el intervalo de confianza al 95%. Tomado de Polanco *et al.* [275].

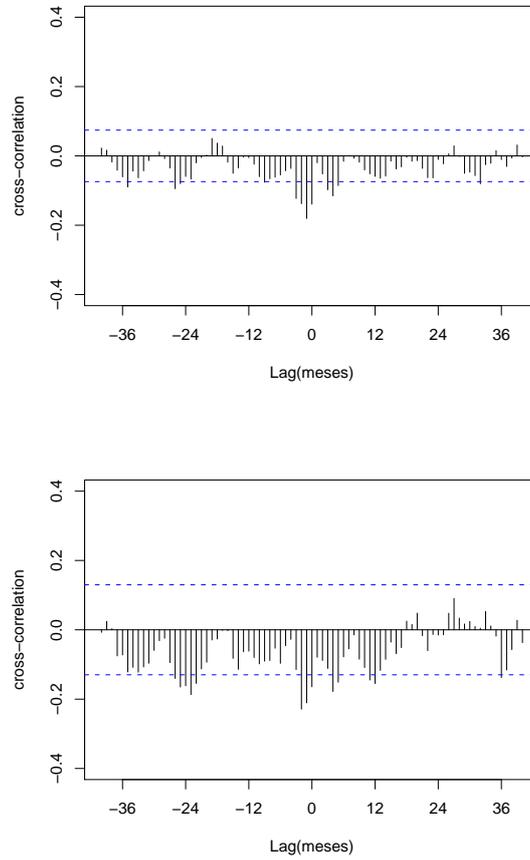


Figura 6.9: Correlación cruzada entre SST y NAO (panel superior) para el periodo enero de 1950 a diciembre de 2007 y entre SST y NAO (panel inferior) para el periodo enero de 1989 a diciembre de 2007. Las SST son datos de Kaplan *et al.* [158]. Las líneas horizontales discontinuas indican el intervalo de confianza al 95 %. Tomado de Polanco *et al.* [275].

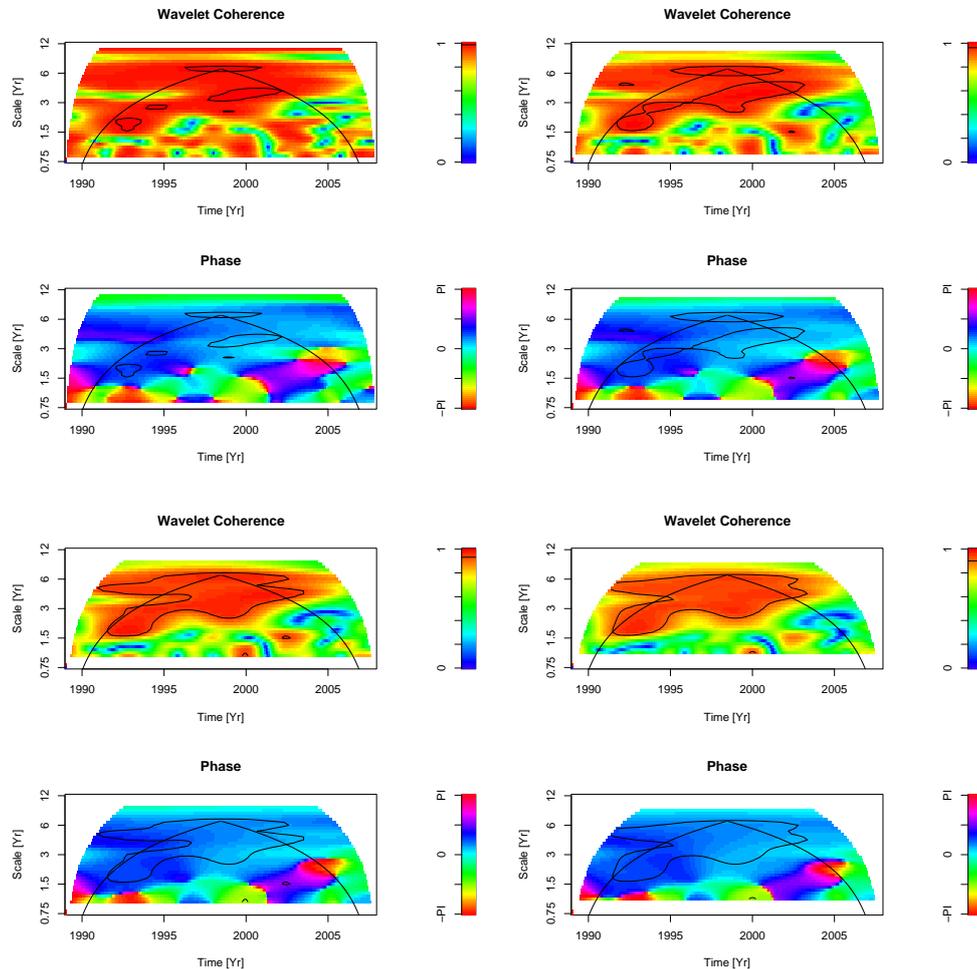


Figura 6.10: WCO entre CPUE e índice NAO. Los parámetros de suavizado para escala y tiempo utilizados para el par de figuras (coherencia y fase) superior izquierda es de 2.4 meses ($\alpha=0.98$), superior derecha es de 3.6 meses ($\alpha=0.96$), inferior izquierda es de 4.8 meses ($\alpha=0.93$) e inferior derecha es de 6 meses ($\alpha=0.90$). Adaptado de Polanco *et al.* [275].

El análisis de la coherencia de Wavelet normalizada (WCO) y de la fase indican una relación no-estacionaria (existencia de señales fugaces) entre la CPUE y las variables climáticas. La coherencia significativa entre CPUE y NAO (coherencia y fase de la Figura 6.10 superior derecha e inferior izquierda) fue encontrada de junio de 1991 a marzo de 2001, con una periodicidad *ca.* 1.9-3.5 años y de fase 0.35-0.92 radianes (de 0.67-1.76 meses), indicando una correlación positiva entre

esas variables, particularmente en esos periodos.

En lo que respecta a la coherencia de Wavelet normalizada (WCO) y de la fase entre la CPUE y las SST (datos de Kaplan *et al.* [158]) (Figura 6.11), la variabilidad de alta frecuencia en las CPUE no parece estar particularmente afectada por la SST, a pesar de que la CCF muestran una cercana relación entre CPUE y SST

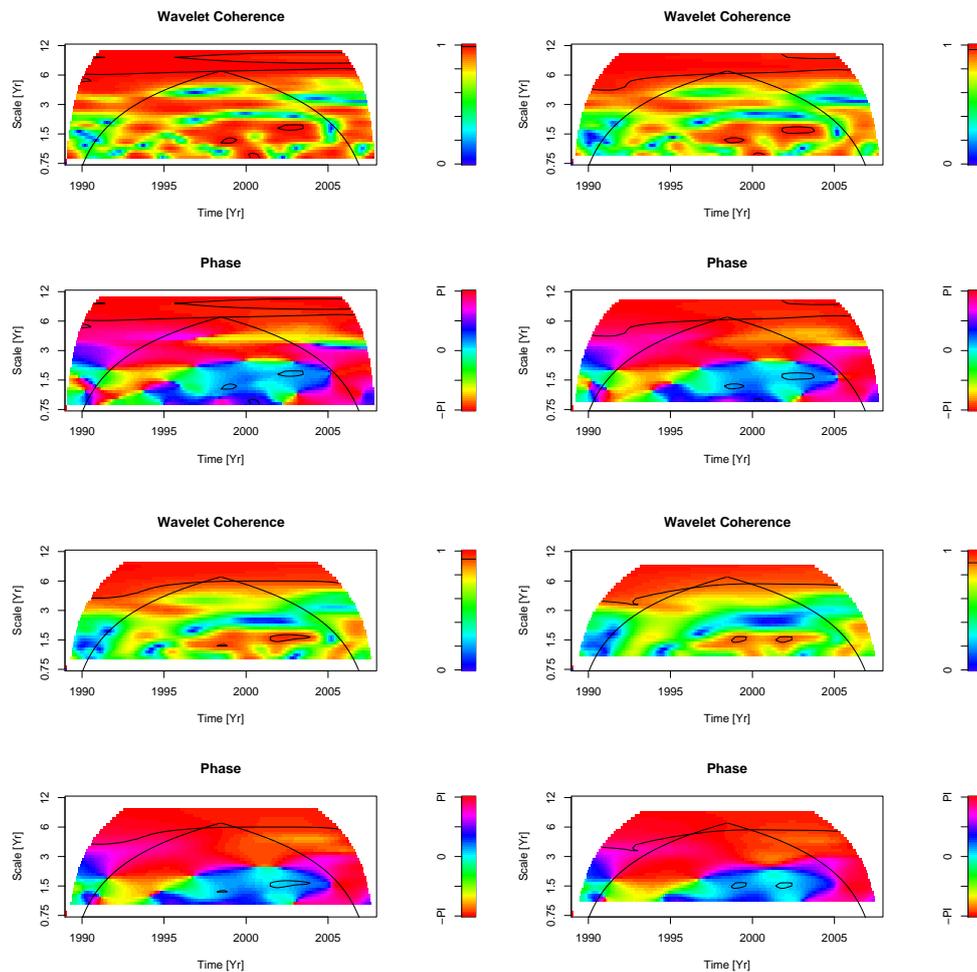


Figura 6.11: WCO entre CPUE y SST (datos de Kaplan *et al.* [158]). Los parámetros de suavizado para escala y tiempo utilizados para el par de figuras (coherencia y fase) superior izquierda es de 2.4 meses ($\alpha = 0.98$), superior derecha es de 3.6 meses ($\alpha = 0.96$), inferior izquierda es de 4.8 meses ($\alpha = 0.93$) e inferior derecha es de 6 meses ($\alpha = 0.90$). Adaptado de Polanco *et al.* [275].

en una escala de tiempo mensual. Los únicos lugares donde la WCO está sobre los niveles que podrían ser esperados por azar, están dentro del cono de influencia y no pueden ser interpretados como relaciones estadísticamente significativas.

Los resultados de calcular la WCO y la fase entre el índice NAO y las SST (datos de Kaplan *et al.* [158]) para el periodo enero de 1989 a diciembre de 2007 (Figura 6.12) no parecen revelar una relación significativa más allá a lo esperado

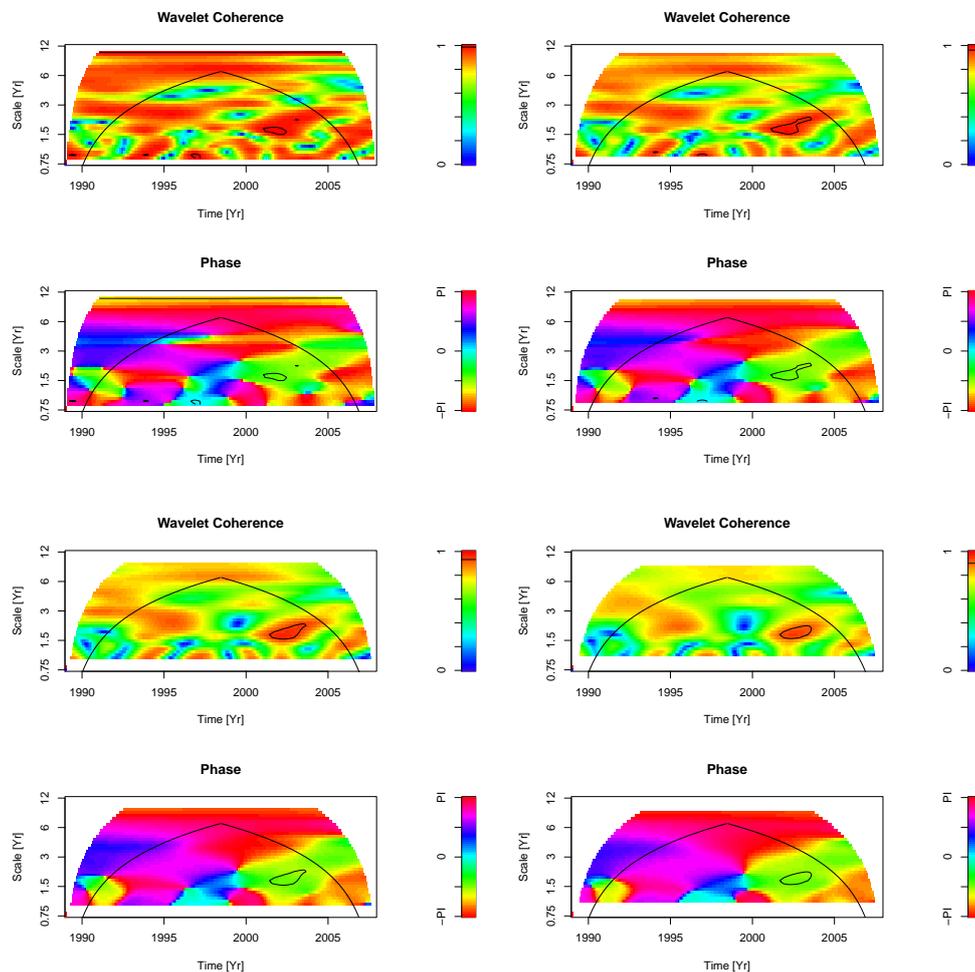


Figura 6.12: WCO entre NAO y SST (datos de Kaplan *et al.* [158]). Los parámetros de suavizado para escala y tiempo para el par de figuras (coherencia y fase) superior izquierda es de 2.4 meses ($\alpha=0.98$), superior derecha es de 3.6 meses ($\alpha=0.96$), inferior izquierda es de 4.8 meses ($\alpha=0.93$) e inferior derecha es de 6 meses ($\alpha=0.90$).

por azar, incluso para el periodo enero de 1950 a diciembre de 2007 (datos de Kaplan *et al.* [158]) (Figura 6.13). Nótese que la presencia de pequeñas concentraciones de puntos espectrales significativos en la WCO no implican necesariamente la existencia de una relación causal entre dos variables (NAO y SST, para este caso) debido al problema de la multiplicidad [206]. Puntos espectrales espurios en WCO podrían aparecer, y aunque existe un test de significación (*areawise test*)

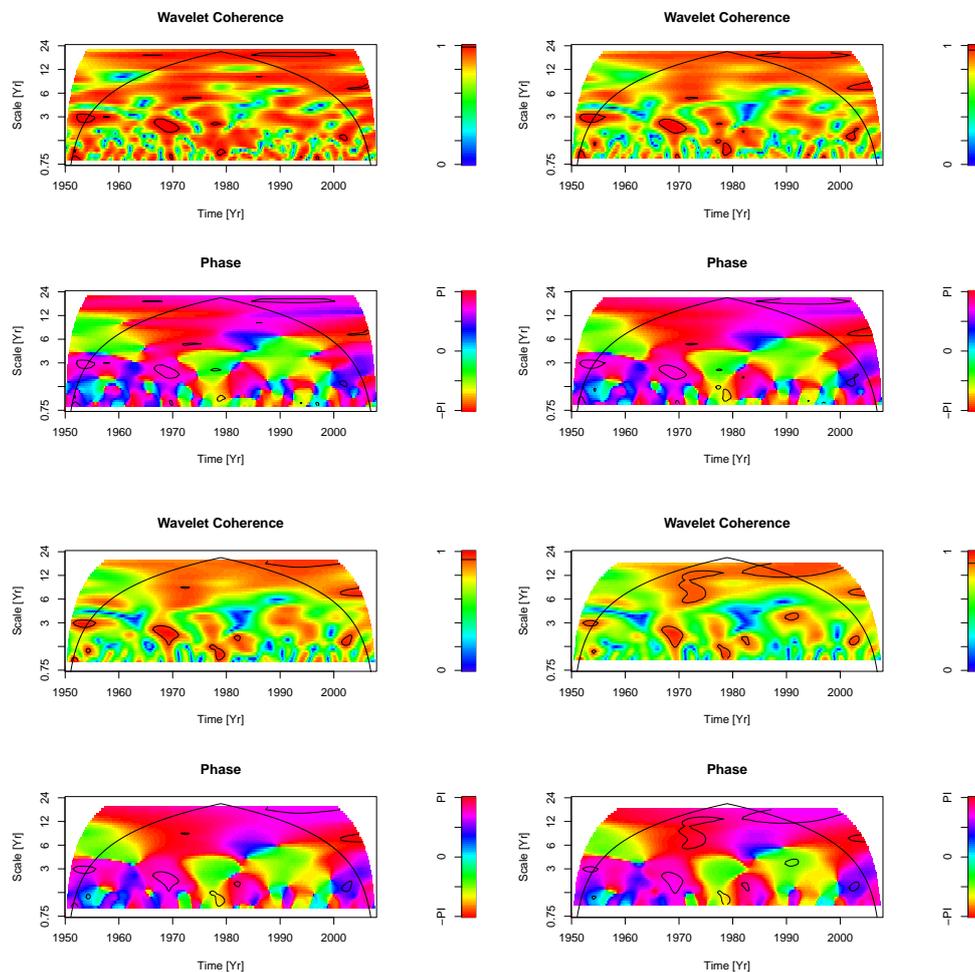


Figura 6.13: WCO entre NAO y SST (datos de Kaplan *et al.* [158]). Los parámetros de suavizado para escala y tiempo para el par de figuras (coherencia y fase) superior izquierda es de 2.4 meses ($\alpha=0.98$), superior derecha es de 3.6 meses ($\alpha=0.96$), inferior izquierda es de 4.8 meses ($\alpha=0.93$) e inferior derecha es de 6 meses ($\alpha=0.90$).

[206, 207] para enfrentar este problema, no se puede aplicar debido a que las series temporales aquí analizadas son cortas y tienen pocos elementos.

6.5. Discusión

Los resultados obtenidos en este estudio muestran que existe un efecto de la variabilidad climática estacional en la abundancia de *Octopus vulgaris* en las Islas Canarias [45]. Las correlaciones entre las CPUE del *Octopus vulgaris* y las SST en diferentes estaciones del año deben ser la consecuencia del efecto de las SST en la supervivencia de los pulpos en su estado paralarvae, la razón de crecimiento, la edad de la población juvenil y la distribución temporal de los picos de reproducción [45]. Relacionado con esto, Hernández-García *et al.* [131] reportaron la existencia de 2 cohortes anuales del pulpo común en las Islas Canarias. La primera durante el mes de abril y la segunda entre octubre y noviembre. Resultados semejantes han sido obtenidos, con lo que respecta al ciclo reproductivo, en otros lugares del Atlántico Noreste [126, 235, 308]. Estos picos reproductivos parecen ocurrir cuando la NAO cambia de bajos a altos valores y viceversa.

La relación entre la SST y el índice NAO ha sido ampliamente establecida en otros estudios desde hace varios años [49, 208, 270, 322, 361, 362]. En esta dirección y en base a los resultados presentados en este capítulo, en primavera y otoño existe una relación inversa entre SST y NAO en las Islas Canarias [45]. Esto puede ser explicado cuando la NAO es considerada desde un punto de vista estacional, donde el invierno es el periodo dominante, pero no en el área cercana a las Islas Canarias, donde la influencia de la NAO es más importante en primavera y otoño. La relación entre factores climáticos (SST y NAO) con la abundancia de *Octopus vulgaris* está además influenciada por las características reproductivas de esta especie de cefalópodo. La relación es más intensa en otoño, dado que en esa época existe una coincidencia en todos los factores que está ausente en el resto de estaciones del año [45].

Sin embargo, la influencia de la NAO y SST va más allá de su efecto sobre las agregaciones reproductivas y la disponibilidad de los *Octopus vulgaris* adultos a las pesquerías. La supervivencia de los pulpos comunes en estado paralarvae, reclutamiento y abundancia están bajo la influencia de muchos factores,

aparte de la temperatura y la disponibilidad de alimento [355, 360]. Estos factores adquieren más o menos relevancia dependiendo de la escala geográfica y de las características locales [89, 129, 200, 325]. Es claro que la biomasa del pulpo común y su accesibilidad a las pesquerías, son la consecuencia de un conjunto de variables ambientales y al mismo tiempo de componentes locales que modulan el éxito de reclutamiento, la distribución, abundancia y comportamiento del *Octopus vulgaris* [87, 119].

En lo que respecta al área geográfica de interés de este trabajo, el apareamiento y desovamiento del pulpo común se presenta en fases de alrededor de un año aunque con picos en primavera y otoño. Sin embargo, y de acuerdo con Hernández-García y colaboradores [131], estos picos fluctúan (con respecto al tiempo) de año en año. Estas fluctuaciones podrían deberse a una combinación entre las estrategias de los métodos de pesca y el ciclo de vida del pulpo común [45]. Esta explicación se ve reforzada por las observaciones realizadas por Hernández-García y colaboradores [130]. Esto es, durante la primavera los pulpos adultos se concentran en aguas someras para desovar, por lo cual son más accesibles para ser capturados. Mientras que durante el verano (la época donde menos capturas se realizan), y tal como sugiere Mangold [200], la población de pulpos podría moverse hacia mayores profundidades, por lo cual son menos accesibles para ser capturados.

A pesar de que, para las fluctuaciones estacionales, el ciclo de vida del *Octopus vulgaris* es el factor de influencia predominante al diseñar las estrategias de capturas de las pesquerías, también es necesario tener en cuenta que los patrones de teleconexión climáticos (NAO) a gran escala pueden influir en los patrones biológicos [325] y en las estrategias de las capturas de las pesquerías en diferentes formas (*v. gr.*, las condiciones meteorológicas u oceanográficas locales adversas pueden impedir la salida de la flota artesanal). Sin embargo, tal y como apunta Hernández-García y colaboradores [131] las influencias directas e indirectas de la NAO en las pesquerías del *Octopus vulgaris* podrían tener cierta importancia. Cabe señalar que la temperatura registrada es una consecuencia del acoplamiento entre la atmósfera y el océano [45].

Como es de esperarse, la NAO conduce los ciclos climáticos y biológicos en el centro-este del océano Atlántico. Es claro que todos sus estados atmosféricos

tienen diferentes grados de influencia sobre el océano y de modo indirecto sobre las poblaciones marinas [361]. Estudios previos han demostrado [61, 168, 309] que la NAO está reflejada en las fluctuaciones de las pesquerías, pero también es cierto que la región sur de las Islas Canarias es menos sensible a sus efectos biológicos [133]. En resumen, cabe destacar la importancia de la NAO, tanto de modo directo como a través de las SST, como un factor que controla la abundancia del *Octopus vulgaris* en el archipiélago Canario. Basado en los resultados de los modelos lineales se sabe que en primavera la NAO es el principal factor climático, la cual explica de modo significativo una varianza del 28.64%. Durante el otoño el efecto de la SST y la NAO no difieren mucho (un 31.13% por la NAO y un 34.21% por la SST) [45].

Hernández-García *et al.* [131] y Caballero *et al.* [45] han mostrado que existe un efecto de la variabilidad climática (SST y la NAO) en altas frecuencias (a escala estacional) sobre la abundancia del *Octopus vulgaris* en las Islas Canarias. La relación (retraso) entre estas variables puede ser explicada como un efecto de la temperatura del océano en el reclutamiento del pulpo común. Debido a que la SST afecta en las fases cruciales de su ciclo de vida, *i. e.*, la supervivencia en estado larval, razón de crecimiento, edad de la población béntica juvenil y la distribución temporal de los picos de máxima reproducción. En bajas frecuencias, la relación SST-CPUE no es tan fuerte y directa como entre la NAO-CPUE, tal como es mostrado por los análisis estadísticos vía la WCO. Por todo ello, la relación entre NAO-CPUE en bajas frecuencias es probablemente debida a otros mecanismos diferentes a la temperatura.

La relación entre la SST y el índice NAO ha sido ampliamente establecida por otros autores [49, 208, 322, 362]. Sin embargo, la CPUE no está particularmente afectada por la SST en bajas frecuencias. Por tanto, la NAO está probablemente relacionada con otras variables ambientales que afectan el ciclo de vida del *Octopus vulgaris* de modo más directo que la temperatura, como podrían ser por ejemplo, las alturas de las olas [47, 190, 303] o la circulación oceánica forzada por el viento [55, 259]. Y aunque el efecto del sistema de afloramiento del Noroeste Africano podría variar de acuerdo a los cambios de la NAO [16, 17, 257], la relación entre la NAO con la zona de afloramiento del Noroeste Africano es un

tanto ambigua debido a una correlación negativa entre el índice NAO y el estrés del viento meridional y de una carencia de correlación con las SST [231]. No obstante, dilucidar este punto queda fuera del ámbito de este estudio.

Los patrones de circulación atmosférica asociados con la NAO pueden afectar directamente a las pesquerías y las capturas de las especies. Es decir, las condiciones meteorológicas, cuando la pesca en pequeña escala podría llevarse a cabo de manera más eficiente son coincidentes con un rango de valores en el índice NAO que describen los rangos característicos del estrés del viento y de las alturas de las olas de superficie. Sin embargo, el índice NAO registrado durante ambos picos reproductivos es diferente (ANOVA de una vía; $F=9.06$; $p < 0.01$), siendo positivo en abril y negativo en octubre, con vientos alisios del norte durante el primer pico y vientos del sur durante el segundo. No obstante, es también cierto que la variación de año en año de las CPUE del pulpo, son también consecuencia del éxito del reclutamiento [335], y por tanto, la razón de supervivencia del pulpo en su estado paralarvae.

El *Octopus vulgaris* en su estado paralarvae es parte del zooplancton durante las primeras (5-12) semanas de vida [200, 360], cambiando gradualmente a un estilo béntico de vida [34]. Este proceso depende de la temperatura [200], y la temperatura del océano depende de varios factores aparte de la NAO, pero relacionados con ésta (estrés del viento, profundidad de la capa de mezcla, cobertura nubosa, etc.). Estos procesos también afectan la disponibilidad de alimento, dispersión, intensidad de los mecanismos de retención de los pulpos en su estado paralarvae alrededor de la isla, etc. [33, 236, 292, 293, 355, 360]. Estos factores son más o menos relevantes dependiendo de la escala geográfica y características locales [130, 200]. De cualquier modo, los efectos de la NAO no parecen ser similares durante todo el periodo de estudio. Estos efectos fueron más prominentes entre 1989 y 1994, coincidiendo con los picos máximos de capturas. Durante este periodo, la coherencia de Wavelet muestra un retraso que oscila entre 1.9 y 3.5 años, el cual podría estar relacionado con el ‘solapamiento’ de varios cohortes en la pesquería tal como es sugerido por Hernández-García y colaboradores [131].

Todas estas interacciones son complejas por la forma en la que las variables climáticas afectan a las especies biológicas (*i. e.*, el desarrollo embrionario, correspondencia y no correspondencia entre paralarvae/plancton, razón de crecimen-

to/desarrollo, edad de la población, edad de la madurez, tiempo de desovamiento, etc.). A pesar de eso, las influencias directas e indirectas de la NAO sobre la pesquería del pulpo podrían tomar diferentes formas, y la SST es una de las más importantes debido a sus efectos en el sistema biológico a diferentes niveles de organización (especies, comunidades y ecosistemas). Esto puede también condicionar la interacción entre el pulpo y los pescadores, *i. e.*, facilitando o complicando el manejo y recuperación de las trampas. Las variaciones en las SST son una consecuencia del acoplamiento entre la atmósfera (y en consecuencia, de la NAO como un fenómeno atmosférico) y el océano, y sus efectos sobre el pulpo varían dependiendo de la fase del ciclo de vida de las especies. No obstante, la NAO no explica el 100 % de la varianza de todos los procesos oceánicos de la zona y, en definitiva, hay otros agentes que aún no se conocen.

Capítulo 7

Conclusiones y trabajo futuro

En este capítulo se presentan las conclusiones de esta memoria, ordenadas conforme a como se han presentado los casos de estudio, seguido de una conclusión general y de algunas líneas de investigación de trabajo futuro.

Conclusiones por capítulos

- El caso de estudio presentado en el capítulo 3, evalúa una red de control de vigilancia y calidad del aire del área metropolitana de Bilbao utilizando promedios mensuales de inmisiones de SO_2 para el periodo 1996-2001 por medio de PCA. En base a los resultados obtenidos se concluye que los cuatros sensores de la AQMN del área metropolitana de Bilbao no miden información redundante. Por tanto, todos los sensores son necesarios para una correcta evaluación de las inmisiones de SO_2 para esta área de estudio [149]. Es importante hacer notar que la misma conclusión fue obtenida con otras técnicas estadísticas diferentes (CA y SOM) aplicadas al mismo conjunto de datos, intervalo temporal y área de estudio [149].

Como conclusión adicional y relacionada con la cuestión metodológica, podemos remarcar que aunque los resultados obtenidos por medio del CA y los SOM son de fácil interpretación, ambas técnicas no proporcionan una ayuda para la identificación de los posibles mecanismos físicos de emisión o dispersión que actúan sobre las inmisiones de SO_2 . Sin embargo, el PCA aporta información sobre cómo es el comportamiento de los sensores, pro-

porcionando información valiosa para la identificación de los posibles mecanismos físicos involucrados [149].

- El caso de estudio presentado en el capítulo 4, realiza una búsqueda de señal orbital de largo periodo en series temporales de los valores medios de $\delta^{18}\text{O}$ y $\delta^{13}\text{C}$ de las rostra de belemnites de la cuenca Vasco-Cantábrica del Jurásico tardío por medio del PLS. En base a los resultados obtenidos se presentan a continuación las principales conclusiones. En primer lugar, se concluye que no se detectaron ciclos orbitales de largo periodo, en particular, el ciclo de los 400 Ka que se esperaba encontrar porque se ha mantenido estable en los últimos 250 Ma. En segundo lugar, se concluye que el ciclo de 109 Ka del espectro del $\delta^{13}\text{C}$ está creado por un efecto de *aliasing* como resultado de ciclos de más corto periodo, muy probablemente causado por el ciclo de la precesión o al ciclo anual. Sin embargo, es difícil determinar qué evento de tipo periódico o cuasi periódico de corto periodo es la causa del ciclo de 109 ka, debido a la resolución temporal de las muestras de los belemnites utilizadas en este estudio.

En tercer lugar, se concluye que la incertidumbre de la estimación espectral al considerar los errores de datación de las muestras de los belemnites aumenta en las bajas frecuencias. En cuarto lugar, se concluye que la extensión de la frecuencia de Nyquist proporciona un margen más amplio para explorar el comportamiento espectral de las series a estudio. Por último, se concluye que es necesario la adquisición de más muestras de belemnites del área de estudio para llegar a conclusiones definitivas.

- El caso de estudio presentado en el capítulo 5, cuantifica los efectos de las variables climáticas en las fluctuaciones de las capturas de atún rojo entre 1525-1756 para el sur de España por medio del análisis de devianza mediante GLM. En base a los resultados obtenidos de este estudio multidisciplinario se puede concluir que, tanto las variaciones climáticas, la dinámica poblacional del atún rojo así como la combinación entre ambas, son los factores que determinan principalmente las fluctuaciones de las capturas del atún rojo. Sin embargo, la dinámica poblacional de esta especie es el factor que

explica el mayor porcentaje de varianza. La variable ambiental que explica más porcentaje de varianza en los GLM son los GHG [102].

- El caso de estudio presentado en el capítulo 6, estudia la variabilidad climática y su relación con las fluctuaciones de capturas por unidad de esfuerzo del pulpo común para el periodo 1989-2007 en las Islas Canarias por medio de la CWT. En base a los resultados obtenidos se presentan a continuación las principales conclusiones. Se concluye que en escalas de alta frecuencia (escala estacional) existe una correlación inversa entre la abundancia (medida en CPUE) y la SST y una relación directa entre la abundancia y la NAO, particularmente en otoño. También se concluye que durante la primavera la NAO es la principal variable climática que afecta a la abundancia de las capturas, mientras que durante el otoño, la SST es la variable más importante y significativa [45].

Se concluye que en escalas de baja frecuencia existe una fuerte relación estadísticamente significativa entre la CPUE y la NAO, aunque la CPUE no está particularmente afectada por la SST. Por tanto, se concluye que las fluctuaciones del pulpo común podrían ser el resultado de las fluctuaciones de las SST pero en sinergia con otras variables ambientales desconocidas, las cuales también están afectadas por la NAO [275].

Conclusiones generales

En esta memoria se han presentado cuatro casos de estudio de tipo multidisciplinario que están relacionados con fenómenos ambientales y ecosistémicos que sucedieron en diferentes escalas espacio-temporales y, aunque tratan sobre temas variados, el factor común es la presencia de la componente climática y la aplicación de técnicas estadísticas. A modo general en esta memoria se concluye que, al utilizar el análisis estadístico de datos ambientales o ecosistémicos es importante tener siempre presente un buen conocimiento de la física del clima, de la componente geológica, en la escala espacio-temporal a estudio y, por supuesto de la componente biológica (de la especie a estudio). Esto, porque en el análisis estadístico de datos, una relación estadísticamente significativa no implica necesariamente una relación causa-efecto y tampoco explica los mecanismos climáticos, geológicos o biológicos involucrados. Por otro lado, con lo que respecta

al uso correcto de las técnicas estadísticas, es importante tener un conocimiento profundo de éstas (fortalezas y limitaciones).

Trabajo futuro

- En lo que respecta al trabajo futuro del caso de estudio presentado en el capítulo 3, cabe destacar lo siguiente. Debido a que la evaluación de las AQMN es un proceso dinámico, es necesaria una constante evaluación de estas redes, y el PCA puede ayudar para llevar a cabo esta tarea. Como trabajo a futuro sería interesante comparar los resultados presentados en esta memoria con un intervalo temporal mucho más reciente, por ejemplo, para los últimos 5 años. Incluso, podría ser más interesante utilizar otros contaminantes, como el O_3 , NO_x o PM_{10} aparte del SO_2 y comparar los resultados obtenidos al evaluar la AQMN del área metropolitana de Bilbao. En principio, debido a las diferentes localizaciones de las fuentes emisoras y al ciclo de vida de los diferentes tipos de contaminantes registrados en la red, cabe la posibilidad de obtener resultados diferentes, en términos de evaluación de la red, para cada contaminante analizado. Esta hipótesis ha sido comprobada por otros autores para otras áreas geográficas [271], pero hasta donde se sabe, no ha sido verificado para el área metropolitana de Bilbao.
- En referencia al trabajo futuro del caso de estudio presentado en el capítulo 4, es necesario obtener más muestras de belemnites para obtener una mejor resolución espacio-temporal y así ganar más confianza en los resultados del análisis espectral de las series temporales de los valores medios de $\delta^{18}O$ y $\delta^{13}C$. Es importante recordar que uno de los principales obstáculos al estudiar el clima de épocas geológicas muy remotas (y la del Jurásico tardío no es la excepción) es la poca disponibilidad de material geológico. Para el caso del Jurásico tardío los belemnites son uno de los pocos registros paleoambientales que nos pueden ayudar en esta tarea y las series aquí analizadas forman parte de un reducido número de series temporales de isótopos estables de belemnites bien datadas que existen a día de hoy.

La serie de $\delta^{13}\text{C}$ plantea un interesante problema futuro, que se podrá abordar en el caso de que se disponga de más muestras de belemnites con mayor resolución temporal. Mediante métodos analíticos o de Monte-Carlo se podría dilucidar de una forma definitiva si los ciclos encontrados son o no reflejo del aliasing de ciclos orbitales de corto periodo (Milankovitch). En cualquiera de los dos casos, se obtendría un resultado interesante, ya que sería el primero en encontrar señal orbital utilizando belemnites e implicaría la presencia de forzamientos externos en el clima de esa época. Por un lado, también ayudarían a esclarecer algunos de los aspectos polémicos con respecto a la localización en frecuencia de algunos de los ciclos de Milankovitch para el Jurásico tardío. Esto es, *v. gr.*, no se conoce a ciencia cierta el valor (o si existieron dos componentes) del ciclo de precesión en el Jurásico tardío, aunque se sitúa en promedio alrededor de los 20 Ka [372]. Por otro lado, también podrían aclarar ciertos aspectos paleoecológicos de esta especie (características migratorias, como los movimientos verticales u horizontales) que no están bien entendidos o que están abiertos a discusión.

- Respecto al trabajo futuro correspondiente al caso de estudio presentado en el capítulo 5, algunas de las cuestiones pendientes a resolver son las siguientes. Existe un alto porcentaje de varianza (entre un 40-50 %) que los modelos GLM no son capaces de explicar, ya sea por efectos climáticos, por la dinámica poblacional del atún rojo o por una combinación de ambas. Recientemente Caballero-Alfonso [44] presentó un estudio sobre las fluctuaciones del atún rojo en el estrecho de Gibraltar y el Mediterráneo occidental para el periodo 1525-1936 por medio de modelos lineales y, de manera similar como en nuestro caso de estudio, hay un alto porcentaje de varianza (entre un 30-60 %) no explicado por los modelos lineales utilizados. Esto podría ser el resultado de otros factores socio-económicos, como conflictos bélicos, crisis económicas, epidemias, etc. o también podría ser debido a fenómenos ambientales a escala local. Por tanto, sería interesante averiguar cuál fue el papel de estos factores. No obstante, ningún modelo estadístico limitado puede aspirar a explicar el 100 % de la varianza de una población de atún, como es bien sabido.

Debido a que la dinámica poblacional juega un importante papel en explicar las fluctuaciones del atún rojo, quizá como trabajo futuro sería interesante aplicar modelos de dinámica poblacional lo más realistas posibles y comparar los resultados con los de los modelos lineales. Por esta línea apunta uno de los casos de estudio presentado por Solari [335], aunque sólo tiene en cuenta los factores climáticos y no hace una comparación de resultados utilizando otras aproximaciones metodológicas, como el análisis estadístico de datos.

- En lo que respecta al trabajo futuro del caso de estudio presentado en el capítulo 6, uno de los trabajos pendientes a realizar consistiría en averiguar qué otras variables ambientales, afectadas también por la NAO, pueden tener influencia en las fluctuaciones del pulpo común en las Islas Canarias. Para esto, sería conveniente obtener datos oceanográficos *in situ*, para poder mejorar la comprensión de cómo afecta la dinámica oceánica y atmosférica a nivel de mesoescala y local al pulpo común.

Se han obtenido buenos resultados por medio del análisis espectral vía la CWT. Tanto en uno de los casos de estudio presentados en esta memoria como en otros trabajos anteriormente publicados [60, 144, 218, 301]. El análisis espectral vía la CWT trae asociado el problema de que las escalas contiguas están correlacionadas. Esto es, proporcionan información redundante, la cual puede en cierto sentido dificultar la interpretación de los resultados. Una forma de enfrentar este problema es utilizar el análisis espectral por medio de la transformada discreta de Wavelet (DWT), en especial la Maximal Overlap Discrete Wavelet Transform (MODWT) [263]. Por ello, como trabajo futuro podría ser interesante aplicar esta herramienta al caso de estudio del capítulo 6 y comparar los resultados de ambas técnicas.

Anexo A

Lista de publicaciones en las que ha participado el autor de esta memoria

1. Ibarra-Berastegi, G., Sáenz, J., Ezcurra, A., Ganzedo, U., Díaz de Argandoña, J., Errasti, I., Fernández-Ferrero, A. y **Polanco-Martínez, J.** (2009). Assessing spatial variability of SO₂ field as detected by an air quality network using Self-Organizing Maps, cluster, and Principal Component Analysis. *Atmospheric Environment*, 102(3):258265.
2. **Polanco-Martínez, J.**, Sáenz, J. y Robles, S (2008). Hunting False Spectral Peaks detected in Unevenly Paleo(Climate) Time Series with Large Gaps using Lomb-Scargle Periodogram. Poster presentado en el *Summer School on Stochastic and Probabilistic methods for atmosphere, ocean, and climate dynamics* organizado por la Universidad de Victoria y por el Pacific Institute for the Mathematical Sciences (PIMS), celebrada en Victoria, Canada. El resumen del trabajo puede consultarse en la memoria del evento: http://www.pims.math.ca/files/Abstracts_0.pdf

-
3. **Polanco-Martínez, J.**, Sáenz, J., Robles, S y Rosales, I. (2008). Long-period orbital cycles of Early Jurassic detected $\delta^{18}\text{O}$ trends of belemnites from the Basque-Cantabrian basin, northern Spain. Poster presentado en la reunión anual de la European Meteorology Society / European Conference on Applied Climatology celebrada en Amsterdam.
 4. Ganzedo, U. and Zorita, E. and Solari, A. P. and Chust, G. and Del Pino, A. S. and **Polanco, J.** and Castro, J. J.. (2009). What drove tuna catches between 1525 and 1756 in southern Europe?. *ICES Journal of Marine Science*, 66(7):1595-1604.
 5. Caballero-Alfonso, A. M. and Ganzedo, U. and Trujillo-Santana, A. and **Polanco, J.** and Santana del Pino, A. and Ibarra-Berastegi, G. and Castro-Hernández, J. J. (2010). The role of climatic variability on the short-term fluctuations of octopus captures at the Canary Islands. *Fisheries Research*, 102(3):258-265.
 6. **J. Polanco** and U. Ganzedo and J. Sáenz and A. M. Caballero-Alfonso and J. J. Castro-Hernández. Wavelet analysis of correlation among Canary Islands octopus captures per unit effort, sea-surface temperatures and the North Atlantic Oscillation. (2011). *Fisheries Research*, 107(1-3):177-183.

Referencias

- [1] Estaciones remotas de la red de vigilancia de la calidad del aire. Departamento de Medio Ambiente, Planificación Territorial, Agricultura y Pesca, Gobierno Vasco. <http://www.ingurumena.ejgv.euskadi.net>, último acceso, agosto 2011. 83
- [2] H. Abdi and L. J. Williams. Principal Component Analysis. *Wiley Interdisc. Rev.: Comp. Stat.*, 2(4):433–459, 2010. 30, 31, 34, 35, 36
- [3] D. M. B. Adams, Carpenter, L. B. A. M, Clarke, and *et al.* SO_x-emissions and control. <http://www.coalonline.org/site/coalonline/content/browser/81432/SOx-emissions-and-control>. 74
- [4] J. Agudo. *Artes y técnicas de pesca tradicionales. El Bajo Guadalquivir*. Junta de Andalucía, Sevilla, 1991. 115
- [5] A. Albizuri. In *Caracterización de patrones meteorológicos a escala regional y local y su relación con los niveles de calidad del aire registrados en la C.A.P.V. Análisis de episodios*. Memorias de la 3a. Jornada técnica sobre contaminación atmosférica, Dept. de Medio Ambiente, Planificación Territorial, Agricultura y Pesca, Gobierno Vasco, 2008. 79, 84
- [6] R. B. Alley, J. Marotzke, W. D. Nordhaus, J. T. Overpeck, D. M. Peteet, R. A. Pielke Jr, R. T. Pierrehumbert, P. B. Rhines, T. F. Stocker, L. D. Talley, and Wallace J. M. Abrupt climate change. *Science*, 299(5615):2005–2010, 2003. 3

- [7] M. H. P. Ambaum, B. J. Hoskins, and D. B. Stephenson. Arctic Oscillation or North Atlantic Oscillation? *J. of Clim.*, 14(16):3495–3507, 2001. [38](#), [39](#)
- [8] M. J. Anderson and N. A. Gribble. Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Australian J. of Eco.*, 23(2):158–167, 1998. [26](#), [52](#), [53](#), [55](#), [56](#)
- [9] T. F. Anderson, B. N. Popp, A. C. Williams, L. Z. Ho, and J. D. Hudson. The stable isotopic records of fossils from the Peterborough Member, Oxford Clay Formation (Jurassic), UK: palaeoenvironmental implications. *J. of the Geol. Soc.*, 151(1):125–138, 1994. [108](#)
- [10] E. Aránguez, J. M. Ordóñez, J. Serrano, N. Aragonés, R. Fernández-Patier, A. Gandarillas, and I. Galán. Contaminantes atmosféricos y su vigilancia. *Revista Española de Salud Pública*, 73(2):123–132, 1999. [13](#)
- [11] C. Arias. Palaeoceanography and biogeography in the Early Jurassic Panthalassa and Tethys Oceans. *Gondwana Research*, 14(3):306–315, 2008. [16](#), [17](#), [18](#), [96](#), [97](#), [112](#)
- [12] T. R. Bailey, Y. Rosenthal, J. M. McArthur, B. van de Schootbrugge, and M. F. Thirlwall. Paleooceanographic changes of the late Pliensbachian-Early Toarcian interval: a possible link to the genesis of an Oceanic Anoxic Event. *Earth & Plan. Sc. Lett.*, 212(3-4):307–320, 2003. [17](#), [18](#), [92](#)
- [13] E. Balguerías, M. E. Quintero, and C. L. Hernández-González. The origin of the Saharan Bank cephalopod fishery. *ICES J. of Mar. Sc.*, 57(1):15–23, 2000. [132](#), [134](#)
- [14] F. X. Bard, P. Bach, and E. Josse. Habitat, Ecophysiologie Des Thons: Quoi de neuf depuis 15 ans? *ICCAT Sci. Pap.*, 50(1):319–342, 1998. [113](#)
- [15] F. J. M. Barning. The numerical analysis of the light-curve of 12 Lacertae. *Bull. of the Astron. Inst. of the Neth.*, 17(1):22–28, 1963. [26](#), [39](#)

- [16] E. D. Barton, J. Arístegui, P. Tett, M. Cantón, J. García-Braun, S. Hernández-León, L. Nykjaer, C. Almeida, J. Almunia, S. Ballesteros, and *et al.* The transition zone of the Canary Current upwelling region. *Prog. in Ocean.*, 41(4):455–504, 1998. [157](#)
- [17] E. D. Barton, J. Arístegui, P. Tett, and E. Navarro-Pérez. Variability in the Canary Islands area of filament-eddy exchanges. *Prog. in Ocean.*, 62:71–94, 2004. [157](#)
- [18] F. Baudin, J. P. Herbin, J. P. Bassoullet, J. Dercourt, G. Lachkar, H. Manivit, and M. Renard. *Distribution of organic matter during the Toarcian in the Mediterranean Tethys and Middle East*. Deposition of Organic Facies, Stud. Geol. Assoc. Pet. Geol., Tulsa, Okla., 30:73-91, 1990. [96](#)
- [19] D. J. Beerling, M. R. Lomas, and D. R. Grocke. On the nature of methane gas-hydrate dissociation during the Toarcian and Aptian oceanic anoxic events. *Ame. J. Sc.*, 302(1):28–49, 2002. [17](#)
- [20] J. S. Bendat and A. G. Piersol. *Random Data. Anal. and Meas. Proc.* Wiley-Interscience, New York, 1986. [10](#), [45](#), [47](#), [48](#), [52](#), [95](#), [102](#), [103](#)
- [21] R. E. Benestad, I. Hanssen-Bauer, and D. Chen. *Empirical-statistical down-scaling*. World Scientific Pub Co Inc, London., 2008. [142](#)
- [22] T. G. Benton, S. J. Plaistow, and T. N. Coulson. Complex population dynamics and complex causation: devils, details and demography. *Proc. of the Roy. Soc. B*, 273(1591):1173–1181, 2006. [4](#), [21](#), [136](#)
- [23] A. Berger and M. F. Loutre. Astronomical theory of climate change. *J. Phys. IV France*, 121:1–35, 2004. [1](#), [18](#), [19](#), [95](#), [110](#)
- [24] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *An. Rev. of Fluid Mech.*, 25(1):539–575, 1993. [30](#)
- [25] A. Berreteaga, S. Robles, I. Rosales, and S. Quesada. Ciclos de precesión en el Toarciense de la Cuenca Vasco-Cantábrica. *Geogaceta*, 32:259–262, 2002. [93](#), [111](#)

- [26] K. Beven. Towards a coherent philosophy for modelling the environment. *Proc. of the Roy. Soc. A*, 458(2026):2465–2484, 2002. [6](#)
- [27] R. J. H. Beverton and S. J. Holt. *On the dynamics of exploited fish populations*. Chapman and Hall, London, 1993. [21](#)
- [28] N. Bezroukov. Open source software as a special type of academic research. *First Monday*, 4(10), 1999. [27](#)
- [29] C. J. Bjerrum, F. Surlyk, J. H. Callomon, and R. L. Slingerland. Numerical paleoceanographic study of the Early Jurassic Transcontinental Laurasian Seaway. *Paleocean.*, 16(4):390–404, 2001. [17](#), [96](#), [112](#)
- [30] R. Blender and K. Fraedrich. Long time memory in global warming simulations. *Geophys. Res. Lett.*, 30(14):CLM7–1, 2003. [11](#)
- [31] B. A. Block, H. Dewar, S. B. Blackwell, T. D. Williams, E. D. Prince, C. J. Farwell, A. Boustany, S. L. H. Teo, A. Seitz, A. Walli, et al. Migratory movements, depth preferences, and thermal biology of Atlantic bluefin tuna. *Science*, 293(5533):1310–1314, 2001. [113](#)
- [32] B. A. Block, S. L. H. Teo, A. Walli, A. Boustany, M. J. W. Stokesbury, C. J. Farwell, K. C. Weng, H. Dewar, and T. D. Williams. Electronic tagging and population structure of Atlantic bluefin tuna. *Nature*, 434(7037):1121–1127, 2005. [113](#)
- [33] G. W. Boehlert, W. Watson, and L. C. Sun. Horizontal and vertical distributions of larval fishes around an isolated oceanic island in the tropical Pacific. *D. Sea Res. Part I. Ocean. Res. P.*, 39(3-4):439–466, 1992. [21](#), [158](#)
- [34] Sv Boletzky. Post-hatching behaviour and mode of life in cephalopods. In Nixon M, Messenger J. B. (eds) *The Biology of Cephalopods Symp. Zool. Soc. Lond.*, 38:557-567, 1977. [158](#)
- [35] D. Borcard and P. Legendre. Environmental control and spatial structure in ecological communities: an example using oribatid mites (Acari, Oribatei). *Env. and Eco. Stat.*, 1(1):37–61, 1994. [26](#), [52](#), [53](#)

- [36] D. Borcard, P. Legendre, and P. Drapeau. Partialling out the spatial component of ecological variation. *Ecology*, 73(3):1045–1055, 1992. [26](#), [52](#), [53](#), [55](#), [123](#)
- [37] A. M. Boustany, D. J. Marcinek, J. Keen, H. Dewar, and B. A. Block. Movements and temperature preferences of Atlantic bluefin tuna (*Thunnus thynnus*) off North Carolina: a comparison of acoustic, archival and pop-up satellite tags. In *Electronic tagging and tracking in marine fisheries: proceedings of the Symposium on Tagging and Tracking Marine Fish with Electronic Devices*. Kluwer Academic Pub, Dordrecht., 2001. [113](#)
- [38] R. S. Bradley and P. D. Jones. ‘Little Ice Age’ summer temperature variations: their nature and relevance to recent global warming trends. *The Holocene*, 3(4):367–376, 1993. [130](#)
- [39] J. C. Braga, M. J. Comas-Rengifo, A. Goy, P. Rivas, and A. Yebenes. El lias inferior y medio en zona central de la cuenca vasco-cantabrica (camino, santander). III Coloquio de Estratigrafía y Paleontología del Jurásico de España, Logroño: Instituto de Estudios Riojanos, Ciencias de la Tierra, Geología 11:17-43, 1988. [99](#)
- [40] H. Braun, P. Ditlevsen, J. Kurths, and M. Mudelsee. Limitations of red noise in analysing Dansgaard-Oeschger events. *Clim. of the Past Disc.*, 5:1803–1818, 2009. [9](#)
- [41] G. Bretthorst. *Frequency Estimation and Generalized Lomb-Scargle Periodograms in Statistical Challenges in Astronomy*. Edited by E. D. Feigelson and G. Jogesh Babu. Springer, New York., 2003. [42](#)
- [42] R. Brill, M. Lutcavage, G. Metzger, P. Bushnell, M. Arendt, J. Lucy, C. Watson, and D. Foley. Horizontal and vertical movements of juvenile bluefin tuna (*Thunnus thynnus*) in relation to oceanographic conditions of the western North Atlantic, determined with ultrasonic telemetry. *Fish. Bull.*, 100(2):155–167, 2002. [113](#)

- [43] M. Bunzl. Researching geoengineering: should not or could not? *Env. Res. Lett.*, 4(4):045104(3pp), 2009. [3](#)
- [44] A. M. Caballero-Alfonso. *Recent and historical climate variability effects on the population dynamics of several marine species*. PhD thesis, Univ. de Las Palmas de Gran Canaria, España, 2011. [164](#)
- [45] A. M. Caballero-Alfonso, U. Ganzedo, A. Trujillo-Santana, **Polanco, J.**, A. Santana del Pino, G. Ibarra-Berastegi, and J. J. Castro-Hernández. The role of climatic variability on the short-term fluctuations of octopus captures at the Canary Islands. *Fish. Res.*, 102(3):258–265, 2010. [24](#), [135](#), [136](#), [137](#), [138](#), [139](#), [140](#), [141](#), [142](#), [145](#), [146](#), [147](#), [155](#), [156](#), [157](#), [162](#)
- [46] E. Cambra, E. Alonso, Cirarda F., and T. Martínez-Rueda. Health impact assessment of air pollution. ENHIS-1 project WP5 health impact assessment, Local City Report Bilbao, 2005. [83](#)
- [47] B. Cañellas, A. Orfila, F. Méndez, A. Álvarez, and J. Tintoré. Influence of the NAO on the northwestern Mediterranean wave climate. *Sci. Mar.*, 74(1):55–64, 2010. [157](#)
- [48] A. Caverivière, F. Domain, and A. Diallo. Observations on the influence of temperature on the length of embryonic development in *Octopus vulgaris* (Senegal). *Aqua. Liv. Res.*, 12(2):151–154, 1999. [137](#)
- [49] D. R. Cayan. Latent and sensible heat flux anomalies over the northern oceans: Driving the sea surface temperature. *J. of Phys. Ocean.*, 22(8):859–881, 1992. [155](#), [157](#)
- [50] B. Cazelles, M. Chavez, D. Berteaux, F. Ménard, J. O. Vik, S. Jenouvrier, and N.C. Stenseth. Wavelet analysis of ecological time series. *Oecologia*, 156(2):287–304, 2008. [22](#), [24](#), [57](#), [58](#), [61](#), [65](#), [66](#), [70](#), [136](#), [143](#), [144](#)
- [51] R. C. Chambers, D. A. Witting, and S. J. Lewis. Detecting critical periods in larval flatfish populations. *J. of Sea Res.*, 45(3-4):231–242, 2001. [130](#)

- [52] M. A. Chandler, D. Rind, and R. Ruedy. Pangaeen climate during the Early Jurassic: GCM simulations and the sedimentary record of paleoclimate. *Geol. Soc. of Ame. Bull.*, 104(5):543–559, 1992. [17](#), [96](#), [112](#)
- [53] R. E. Chandler. On the use of generalized linear models for interpreting climate variability. *Environmetrics*, 16(7):699–715, 2005. [23](#)
- [54] C. S. Chen, G. J. Pierce, J. Wang, J. P. Robin, J. C. Poulard, J. Pereira, A. F. Zuur, P. R. Boyle, N. Bailey, D. J. Beare, et al. The apparent disappearance of *Loligo forbesi* from the south of its range in the 1990s: trends in *Loligo* spp. abundance in the northeast Atlantic and possible environmental influences. *Fish. Res.*, 78(1):44–54, 2006. [137](#)
- [55] K. C. Chhak and A. M. Moore. The North Atlantic Oscillation as a source of stochastic forcing of the wind-driven ocean circulation. *Dyn. of Atm. and Oceans*, 43(3-4):151–170, 2007. [157](#)
- [56] M. Claussen, L. Mysak, A. Weaver, M. Crucifix, T. Fichefet, M. F. Loutre, S. Weber, J. Alcamo, V. Alexeev, A. Berger, et al. Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models. *Clim. Dyn.*, 18(7):579–586, 2002. [5](#)
- [57] C. A. S. Coelho, S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes, and D. B. Stephenson. Forecast calibration and combination: A simple bayesian approach for ENSO. *J. Clim.*, 17(7):1504–1516, 2004. [7](#)
- [58] R. S. Collett and K. Oduyemi. Air quality modelling: a technical review of mathematical approaches. *Meteorol. App.*, 4(3):235–246, 1997. [80](#)
- [59] M. J. Comas-Rengifo, A. Goy, P. Rivas, and A. Yebenes. El Toarciense en Castillo Pedroso (Santander). *III Coloquio de Estratigrafía y Paleontología del Jurásico de España, Logroño: Instituto de Estudios Riojanos, Ciencias de la Tierra, Geología*, 11:63–71, 1988. [99](#)
- [60] A. Corbineau, T. Rouyer, B. Cazelles, J. M. Fromentin, A. Fonteneau, and F. Menard. Time series analysis of tuna and swordfish catches and climate

- variability in the Indian Ocean (1968-2003). *Aqua. Liv. Res.*, 21(3):277–285, 2008. [144](#), [165](#)
- [61] A. A. H. M. Corten. *Herring and climate: changes in the distribution of North Sea herring due to climate fluctuations*. PhD thesis, University of Groningen, The Netherlands, 2001. [118](#), [157](#)
- [62] T. J. Crowley. Causes of climate change over the past 1000 years. *Science*, 289(5477):270–277, 2000. [120](#), [121](#), [122](#), [130](#)
- [63] M. Crucifix, M. F. Loutre, and A. Berger. The climate response to the astronomical forcing. *Spa. Sc. Rev.*, 1(4):213–226, 2006. [1](#), [19](#), [95](#), [110](#)
- [64] J. Csirke. *Introducción a la dinámica de poblaciones de peces*. FAO, Doc. Téc. Pesca 192, 1989. [19](#)
- [65] K. M. Cuddington and P. Yodzis. Black noise and population persistence. *Proc. of the Roy. Soc. B*, 266(1422):969–973, 1999. [144](#)
- [66] P. Cury, O. Anneville, F. X. Bard, A. Fonteneau, and C. Roy. Obstinate North Atlantic Bluefin Tuna (*Thunnus Thynnus Thynnus*) an Evolutionary Perspective to Consider Spawning Migration. *ICCAT Sci. Pap.*, 50:239–247, 1998. [130](#)
- [67] J. M. Cushing, B. Dennis, R. A. Desharnais, and R. F. Costantino. Moving toward an unstable equilibrium: saddle nodes in population systems. *J. of Ani. Eco.*, 67(2):298–306, 1998. [136](#)
- [68] L. Dagorn, P. Bach, and E. Josse. Movement patterns of large bigeye tuna (*Thunnus obesus*) in the open ocean, determined using ultrasonic telemetry. *Mar. Bio.*, 136(2):361–371, 2000. [115](#)
- [69] I. Daubechies. *Ten lectures on wavelets*. Society for Industrial Mathematics, Philadelphia, USA, 1992. [27](#), [61](#)
- [70] T. J. Deeming. Fourier analysis with unequally-spaced data. *Astrop. and Spa. Sc.*, 36(1):137–158, 1975. [103](#)

- [71] J. Dercourt, L. E. Ricou, and B. Vrielynck. *Atlas Tethys palaeoenvironmental maps*. Gauthier-Villars, Paris, France, 1993. [96](#)
- [72] T. Diomede, F. Nerozzi, T. Paccagnella, and E. Todini. The use of meteorological analogues to account for LAM QPF uncertainty. *Hydrol. Earth Syst. Sc.*, 12(5):141–157, 2008. [7](#)
- [73] P. A. Dixon, M. J. Milicich, and G. Sugihara. *Noise and nonlinearity in an ecological system* In *Nonlinear Dynamics and Statistics*. A. I. Mees (ed). Birkhauser, Boston, Mass., 2001. [19](#)
- [74] A. J. Dobson. *An introduction to generalized linear models*. Chapman & Hall, 1990. [53](#)
- [75] D. Dommenget and M. Latif. A cautionary note on the interpretation of EOFs. *J. Clim.*, 15(2):216–225, 2002. [38](#), [39](#)
- [76] R. Donner, S. Barbosa, J. Kurths, and N. Marwan. Understanding the Earth as a Complex System—recent advances in data analysis and modelling in Earth sciences. *T. Eur. Phys. J. Spe. Top.*, 174(1):1–9, 2009. [2](#), [4](#), [7](#), [25](#)
- [77] P. Doyle and D. I. M. MacDonald. Belemnite battlefields. *Lethaia*, 26(1):65–80, 1993. [108](#)
- [78] S. Dray. On the number of principal components: A test of dimensionality based on measurements of similarity between matrices. *Comp. Stat. and Data Analysis*, 52(4):2228–2237, 2008. [37](#)
- [79] K. F. Drinkwater, G. Beaugrand, M. Kaeriyama, S. Kim, G. Ottersen, R.I. Perry, H. O. Pörtner, J. J. Polovina, and A. Takasuka. On the processes linking climate to ecosystem changes. *J. of Mar. Syst.*, 79(3-4):374–388, 2010. [20](#), [21](#)
- [80] L. V. Duarte. Clay minerals and geochemical evolution in the Toarcian-lower Aalenian of the Lusitanian Basin (Portugal). *Cuadernos de Geología Ibérica*, (24):69–98, 1998. [92](#)

- [81] M. Dumett. *Modelos lineales generalizados*. Universidad Nacional de Colombia. Dept. de Matemáticas y estadística, Bogota, 1996. [53](#)
- [82] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. I. *Biometrika*, 37(1-2):409–428, 1950. [143](#)
- [83] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. II. *Biometrika*, 38(1-2):159–178, 1951. [143](#)
- [84] W. Ebisuzaki. A method to estimate the statistical significance of a correlation when the data are serially correlated. *J. clim.*, 10(9):2147–2153, 1997. [141](#)
- [85] E. S. Epstein. Stochastic dynamic prediction. *Tellus*, 21:739–759, 1969. [7](#)
- [86] L. Eyer and P. Bartholdi. Variable stars: which Nyquist frequency? *Astron. and Astrop. Supp. Ser.*, 135(1):1–3, 1999. [103](#), [104](#), [111](#)
- [87] A. Faraj and N. Bez. Spatial considerations for the Dakhla stock of *Octopus vulgaris*: indicators, patterns, and fisheries interactions. *ICES J of Mar. Sc.*, 64(9):1820, 2007. [156](#)
- [88] M. Farge. Wavelet transforms and their applications to turbulence. *Ann. Rev. of Fluid Mech.*, 24(1):395–458, 1992. [27](#), [57](#), [58](#), [63](#), [136](#)
- [89] V. Faure, C. A. Inejih, H. Demarcq, and P. Cury. The importance of retention processes in upwelling areas for recruitment of *Octopus vulgaris*: the example of the Arguin Bank (Mauritania). *Fish. Ocean.*, 9(4):343–355, 2000. [132](#), [156](#)
- [90] A. G. Fischer. *Orbital cyclicity in Mesozoic strata in Cycles and events in stratigraphy*. Springer-Verlag, Berlin, 1991. [19](#), [95](#), [110](#)
- [91] D. Fleitmann, S. J. Burns, M. Mudelsee, U. Neff, J. Kramers, A. Mangini, and A. Matter. Holocene forcing of the Indian monsoon recorded in a stalagmite from Southern Oman. *Science*, 300(5626):1737–1739, 2003. [102](#)

- [92] Food and Agriculture Organization of the United Nations. Species fact sheets: Octopus vulgaris (Lamarck, 1798). <http://www.fao.org/fishery/species/3571/en>, último acceso, agosto 2011. 132, 136
- [93] J. W. Forsythe. *A working hypothesis of how seasonal temperature change may impact the field growth of young cephalopods in Recent advances in cephalopod fisheries biology*. Tokai University Press, Tokyo, 1993. 137
- [94] K. Fraedrich, H. Jansen, E. Kirk, U. Luksch, and F. Lunkeit. The Planet Simulator: Towards a user friendly model. *Meteorologische Zeitschrift*, 14(3):299–304, 2005. 5
- [95] J. M. Fromentin. The East Atlantic and Mediterranean bluefin tuna stock management: uncertainties and alternatives. *Sc. Mar.*, 67(S1):51–62, 2003. 130
- [96] J. M. Fromentin. Description of Atlantic bluefin tuna (BFT). *ICCAT Field Manual*, 2006. 114, 130
- [97] J. M. Fromentin and A. Fonteneau. Fishing effects and life history traits: a case study comparing tropical versus temperate tunas. *Fish. Res.*, 53(2):133–150, 2001. 113
- [98] J. M. Fromentin, A. Fonteneau, and H. Farrugio. Biological reference points and natural long-term fluctuations: The case of the eastern Atlantic bluefin tuna. *ICCAT Sci. Pap.*, 51(1):2072–2084, 2000. 130
- [99] J. M. Fromentin and J. E. Powers. Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. *Fish and Fisheries*, 6(4):281–306, 2005. 115, 130
- [100] A. W. Fullerton. Searching for periodicity in astronomical data. In *The Study of Variable Stars Using Small Telescopes*. Cambridge University Press, Cambridge., 1986. 43

- [101] G. Gangoiti, L. Alonso, M. Navazo, A. Albizuri, G. Pérez-Landa, M. Matabuena, V. Valdenebro, M. Maruri, J. Antonio García, and M. M. Millán. Regional transport of pollutants over the Bay of Biscay: analysis of an ozone episode under a blocking anticyclone in west-central Europe. *Atm. Env.*, 36(8):1349–1361, 2002. [82](#), [83](#)
- [102] U. Ganzedo, E. Zorita, A. P. Solari, G. Chust, A. S. Del Pino, **Polanco, J.**, and J. J. Castro. What drove tuna catches between 1525 and 1756 in southern Europe?. *ICES J. of Mar. Sc.*, 2009. [23](#), [26](#), [29](#), [53](#), [54](#), [55](#), [115](#), [116](#), [117](#), [123](#), [124](#), [125](#), [127](#), [128](#), [131](#), [162](#)
- [103] U. Ganzedo-López. *Efecto de las variaciones climáticas en la distribución espacio-temporal de Thunnus thynnus thynnus (Linnaeus 1758) y Thunnus alalunga (Bonnaterre 1788) en el Océano Atlántico*. PhD thesis, Univ. de Las Palmas de Gran Canaria, España, 2005. [22](#), [23](#), [114](#), [140](#)
- [104] S. Georgakarakos, D. Koutsoubas, and V. Valavanis. Time series analysis and forecasting techniques applied on loliginid and ommastrephid landings in Greek waters. *Fish. Res.*, 78(1):55–71, 2006. [137](#)
- [105] M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. Advanced spectral methods for climatic time series. *Rev. Geophys.*, 40(1):1003, doi:10.1029/2000RG000092, 2002. [26](#)
- [106] W. Gibbons and T. Moreno. *The geology of Spain*. The Geological Society, London., 2002. [96](#)
- [107] D. L. Gilman, F. J. Fuglister, and J. M. Mitchell Jr. On the power spectrum of red noise. *J. Atm. Sc.*, 20(2):182–184, 1963. [9](#), [26](#), [49](#), [101](#)
- [108] J. Golonka. Late Triassic and Early Jurassic palaeogeography of the world. *Palaeogeog., Palaeoclim., Palaeoeco.*, 244(1-4):297–307, 2007. [16](#)
- [109] J. J. Gómez, A. Goy, and M. L. Canales. Seawater temperature and carbon isotope variations in belemnites linked to mass extinction during the Toarcian (Early Jurassic) in central and Northern Spain. Comparison with

- other European sections. *Palaeogeog., Palaeoclim., Palaeoeco.*, 258(1-2):28 – 58, 2008. [18](#), [92](#), [93](#)
- [110] J. A. González, Santana J. I., and J. Carrillo. La pesca en el puerto de mogán (Islas Canarias): flota, artes y análisis de las capturas entre 1980 y 1990. Informe Técnico del Centro de Tecnología Pesquera. Ed. Cabildo de Gran Canaria. Telde, España, 1991. [137](#)
- [111] E. Gramsch, F. Cereceda-Balic, P. Oyola, and D. Von Baer. Examination of pollution trends in Santiago de Chile with cluster analysis of PM10 and Ozone data. *Atm. Env.*, 40(28):5464–5475, 2006. [15](#), [25](#), [82](#)
- [112] S. M. Griffies. Elements of MOM4p1, NOAA. Geophysical Fluid Dynamics Laboratory, Ocean Group Technical Report No. 6, Princeton, 2009. [5](#)
- [113] A. Grinsted, J. C. Moore, and S. Jevrejeva. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonl. Proc. in Geophys.*, 11:561–566, 2004. [24](#), [27](#), [57](#), [136](#), [143](#), [144](#)
- [114] A. Grossman and J. Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. on Math. Anal.*, 15:723–736, 1984. [27](#), [57](#), [136](#)
- [115] A. Guisan, T. C. Edwards, and T. Hastie. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Eco. Mod.*, 157(2-3):89–100, 2002. [22](#), [23](#), [53](#), [116](#)
- [116] A. Guisan and N. E. Zimmermann. Predictive habitat distribution models in ecology. *Eco. Mod.*, 135(2-3):147–186, 2000. [53](#), [56](#), [122](#)
- [117] J. M. Gutiérrez, A. S. Cofiño, R. Cano, and M. A. Rodríguez. Clustering methods in statistical downscaling for short-range weather forecasts. *Mon. Wea. Rev.*, 132(9):2169–2183, 2004. [7](#)
- [118] A. Hallam. Estimates of the amount and rate of sea-level change across the Rhaetian Hettangian and Pliensbachian Toarcian boundaries (latest Triassic to early Jurassic). *J. of Geol. Soc.*, 154(5):773, 1997. [17](#)

- [119] R. T. Hanlon and J. B. Messenger. *Cephalopod behaviour*. Cambridge Univ. Press, Cambridge, 1998. [156](#)
- [120] A. Hannachi, I. T. Jolliffe, and D. B. Stephenson. Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. of Clim.*, 27(9):1119–1152, 2007. [31](#), [32](#), [33](#), [34](#), [35](#), [38](#)
- [121] A. Hannachi, I. T. Jolliffe, D. B. Stephenson, and N. Trendafilov. In search of simple structures in climate: Simplifying EOFs. *Int. J. of clim.*, 26(1):7–28, 2006. [38](#)
- [122] M. F. Harkat, G. Mourot, and J. Ragot. An improved PCA scheme for sensor FDI: Application to an air quality monitoring network. *J. of Proc. Cont.*, 16(6):625–634, 2006. [85](#)
- [123] F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. of the IEEE*, 66(1):51–83, 1978. [47](#), [48](#)
- [124] K. Hasselmann. Stochastic climate models Part I. Theory. *Tellus*, 28(6):473–485, 1976. [9](#), [26](#), [49](#), [95](#), [101](#), [144](#)
- [125] A. Hastings. Transient dynamics and persistence of ecological systems. *Eco. Lett.*, 4(3):215–220, 2001. [136](#)
- [126] H. Hatanaka. Studies on the fisheries biology of common octopus off the northwest coast of Africa. *Bull. F. S. Fish. Res. Lab.*, 17:13–124, 1979. [155](#)
- [127] J. R. Hazel. *Thermal biology In The physiology of Fishies*. CRC press, Boca Raton, Florida, 1993. [116](#)
- [128] R. C. Henry. History and fundamentals of multivariate air quality receptor models. *Chem. and Intel. Lab. Syst.*, 37(1):37–42, 1997. [14](#), [76](#)
- [129] V. Hernández-García and J. J. Castro. Morphological variability in *Illex coindetii* (Cephalopoda: Ommastrephidae) along the north-west coast of Africa. *J. of the Mar. Biol. Assoc. of the UK*, 78(4):1259–1268, 1998. [156](#)

- [130] V. Hernández-García, J. L. Hernández-Lopez, and J. J. Castro. The octopus (*Octopus vulgaris*) in the small-scale trap fishery off the Canary Islands (Central-East Atlantic). *Fish. Res.*, 35(3):183–189, 1998. [23](#), [24](#), [132](#), [133](#), [134](#), [135](#), [137](#), [138](#), [156](#), [158](#)
- [131] V. Hernández-García, J. L. Hernández-López, and J. J. Castro-Hdez. On the reproduction of *Octopus vulgaris* off the coast of the Canary Islands. *Fish. Res.*, 57(2):197–203, 2002. [23](#), [24](#), [132](#), [133](#), [135](#), [137](#), [155](#), [156](#), [157](#), [158](#)
- [132] J. L. Hernández-López, J. J. Castro-Hernández, and V. Hernández-García. Age determined from the daily deposition of concentric rings on common octopus (*Octopus vulgaris*) beaks. *Fish. Bull.*, 99(4):679–684, 2001. [23](#), [132](#), [133](#), [134](#), [136](#)
- [133] R. G. Herrera, D. G. Puyol, E. H. Martín, L. G. Presa, and P. R. Rodríguez. Influence of the North Atlantic oscillation on the Canary Islands precipitation. *J. Clim.*, 14(19):3889–3903, 2001. [157](#)
- [134] D. Heslop and M. J. Dekkers. Spectral analysis of unevenly spaced climatic time series using CLEAN: signal recovery and derivation of significance levels using a Monte Carlo simulation. *Phys. Earth Planet. Int.*, 130(1):103–116, 2002. [94](#)
- [135] S. P. Hesselbo, D. R. Gröcke, H. C. Jenkyns, C. J. Bjerrum, P. Farrimond, H. S. M. Bell, and O. R. Green. Massive dissociation of gas hydrate during a Jurassic oceanic anoxic event. *Nature*, 406(6794):392–395, 2000. [92](#)
- [136] L. A. Hinnov and J. G. Ogg. Cyclostratigraphy and the Astronomical Time Scale. *Stratigraphy*, 4(2-3):239–251, 2007. [19](#), [95](#)
- [137] L. A. Hinnov and J. F. Park. Strategies for assessing Early-Middle (Pliensbachian-Aalenian)-Jurassic cyclochronologies. *Phil. Trans. R. Soc. A*, 357:1831–1859, 1999. [93](#)
- [138] E. E. Hofmann and T. M. Powell. Environmental variability effects on marine fisheries: four case histories. *Eco. App.*, 8(Sup):23–32, 1998. [19](#), [20](#)

- [139] J. R. Holton. *An introduction to dynamic meteorology*. Elsevier Academic Press, Burlington, MA, 2004. [80](#)
- [140] J. H. Horne and S. L. Baliunas. A prescription for period analysis of unevenly sampled time series. *Astrophys. J.*, 302:757–763, 1986. [40](#), [45](#), [94](#)
- [141] H. Hotelling. Analysis of a complex of statistical variables into principal components. *J. of Educational Psychology*, 24(6):417–441, 1933. [30](#)
- [142] C. Hsieh, S. M. Glaser, A. J. Lucas, and G. Sugihara. Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. *Nature*, 435(7040):336–340, 2005. [24](#), [136](#)
- [143] C. Hsieh and M. D. Ohman. Biological responses to environmental forcing: the linear tracking window hypothesis. *Ecology*, 87(8):1932–1938, 2006. [19](#)
- [144] C. H. A. O. Hsieh, C. S. Chen, T. A. I. S. Chiu, K. U. O. T. Lee, F. J. E. N. Shieh, J. I. A. Y. I. Pan, and M. A. N. Lee. Time series analyses reveal transient relationships between abundance of larval anchovy and environmental variables in the coastal waters southwest of Taiwan. *Fish. Ocean.*, 18(2):102–117, 2009. [22](#), [27](#), [57](#), [144](#), [165](#)
- [145] J. W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck. An overview of the North Atlantic oscillation In: Hurrell J. W., Kushnir Y., Ottersen G., Visbeck M. (eds). *The North Atlantic Oscillation: climatic significance and environmental impact*. *Geophysical Monograph*, 134, 2003. [20](#), [139](#)
- [146] P. J. Huybers and C. Wunsch. A Depth-Derived Pleistocene Age-Model: Uncertainty Estimates, Sedimentation Variability, and Nonlinear Climate Change. *Paleocean.*, 19:1–24, 2004. [106](#)
- [147] G. Ibarra-Berastegi, A. Elías, A. Barona, J. Sáenz, A. Ezcurra, and J. Díaz de Argandoña. From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in Bilbao. *Env. Mod. and Soft.*, 23(5):622–637, 2007. [28](#), [83](#), [84](#)

- [148] G. Ibarra-Berastegi, U. Ganzedo, J. Saenz, A. Ezcurra, I. Errasti, A. Elias, A. Barona, and L. Insausti. R: A valuable tool for linking high education and research. *WSEAS Trans. on Adv. in Ing. Edu.*, 6(4):121–130, 2009. [28](#)
- [149] G. Ibarra-Berastegi, J. Sáenz, A. Ezcurra, U. Ganzedo, J. Díaz de Argandoña, I. Errasti, A. Fernández-Ferrero, and **Polanco-Martínez, J.** Assessing spatial variability of SO₂ field as detected by an air quality network using Self-Organizing Maps, cluster, and Principal Component Analysis. *Atm. Env.*, 43(25):3829–3836, 2009. [16](#), [28](#), [29](#), [82](#), [83](#), [84](#), [85](#), [86](#), [87](#), [88](#), [89](#), [90](#), [91](#), [160](#), [161](#)
- [150] P. Inchausti and J. Halley. The long-term temporal variability and spectral colour of animal populations. *Evol. Eco. Res.*, 4(7):1033–1048, 2002. [144](#)
- [151] A. Inza, M. E. Sánchez, M. Menéndez, and L. A. Ortega. In *Análisis de contribución de fuentes en PM₁₀ y PM_{2.5} en un área de fondo urbano con influencia de emisiones industriales (Abanto, Vizcaya)*. Gobierno Vasco. Departamento de Medio Ambiente y Ordenación del Territorio. Dirección de Planificación, Evaluación y Control Ambiental, 2005. [79](#)
- [152] G. M. Jenkins and D. G. Watts. *Spectral analysis and its applications*. Holden-Day, Oakland, CA, 1969. [50](#)
- [153] H. C. Jenkyns. The early Toarcian (Jurassic) anoxic event; stratigraphic, sedimentary and geochemical evidence. *Ame. J. Sc.*, 288:101–151, 1988. [92](#)
- [154] H. C. Jenkyns, D. R. Gröcke, and S. P. Hesselbo. Nitrogen isotope evidence for water mass denitrification during the early Toarcian (Jurassic) oceanic anoxic event. *Paleocean.*, 16(6):593–603, 2001. [92](#)
- [155] H. C. Jenkyns, C. E. Jones, D. R. Grocke, S. P. Hesselbo, and D. N. Parkinson. Chemostratigraphy of the Jurassic System: applications, limitations and implications for palaeoceanography. *J. of the Geol. Soc.*, 159(4):351, 2002. [17](#), [18](#), [93](#)
- [156] I. T. Jolliffe. *Principal component analysis*. Springer-Verlag, New York, 2002. [31](#), [36](#), [37](#), [85](#)

- [157] G. Kaiser. *A friendly guide to wavelets*. Birkh., Boston, 1994. [68](#)
- [158] A. Kaplan, M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan. Analyses of global sea surface temperature 1856-1991. *J. of Geophys. Res.*, 103(18):18567–18589, 1998. [140](#), [141](#), [143](#), [148](#), [149](#), [150](#), [152](#), [153](#), [154](#)
- [159] S. Katsanevakis and G. Verriopoulos. Modelling the effect of temperature on hatching and settlement patterns of meroplanktonic organisms: the case of octopus. *Sc. Mar.*, 70(4):699–708, 2006. [23](#), [134](#)
- [160] R. W. Katz. Sir Gilbert Walker and a connection between El Nino and statistics. *Stat. Sc.*, 17(1):97–112, 2002. [7](#)
- [161] C.M. Kelty. Free software/free science. *First Monday*, 6(12), 2001. [27](#)
- [162] D. B. Kemp, A. L. Coe, A. S. Cohen, and L. Schwark. Astronomical pacing of methane release in the Early Jurassic period. *Nature*, 437(7057):396–399, 2005. [92](#)
- [163] S. M. Kendall. *Multivariate analysis*. Charles Griffin, London, 1980. [30](#)
- [164] J. F. Kirby. Which wavelet best reproduces the Fourier power spectrum? *Computers & Geosciences*, 31(7):846–864, 2005. [144](#)
- [165] C. Koen. The Nyquist frequency for irregularly spaced time-series: a calculation formula. *Mon. N. Roy. Astron. Soc.*, 371(3):1390–1394, 2006. [103](#), [104](#)
- [166] K. E. Korsmeyer and H. Dewar. *Tuna metabolism and energetics*. Academic Press, San Diego, 2001. [116](#)
- [167] J. Kropp and H. J. Schellnhuber. *In Extremis: Trends, Correlations and Extremes in Hydrology and Climate*. Springer-Verlag, Berlin, 2009. [12](#)
- [168] A. S. Krovnin. *A comparative study of climatic changes in the North Pacific and North Atlantic and their relation to the abundance of fish stocks In Climate Change and Northern Fish Populations*. R. J. Beamish Ed. Victoria, B.C., 1995. [157](#)

- [169] S. Kruizinga and A. H. Murphy. Use of an analogue procedure to formulate objective probabilistic temperature forecasts in The Netherlands. *Mon. Wea. Rev.*, 111(11):2244–2254, 1983. [7](#)
- [170] R. Krzysztofowicz. Probabilistic hydrometeorological forecasts: Toward a new era in operational forecasting. *Bull. Amer. Meteor. Soc.*, 79(2):243–251, 1998. [7](#)
- [171] R. Krzysztofowicz. The case for probabilistic forecasting in hydrology. *J. Hydrol.*, 249(1-4):2–9, 2001. [7](#)
- [172] P. Kumar and E. Foufoula-Georgiou. *Wavelet Analysis in Geophysics: An Introduction*. Wavelets in Geophysics, Academic Press, London, 1994. [27](#), [57](#), [59](#), [60](#), [61](#)
- [173] P. Kumar and E. Foutoula-Georgiou. Wavelet analysis for geophysical applications. *Rev. of Geophys.*, 35(4):385–412, 1997. [57](#), [58](#), [60](#), [66](#)
- [174] S. N. Lane. Assessment of rainfall-runoff models based upon wavelet analysis. *Hydrol. Proc.*, 21(5):586–607, 2007. [70](#)
- [175] J. Laskar, P. Robutel, F. Joutel, M. Gastineau, A. C. M. Correia, and B. Levrard. A long-term numerical solution for the insolation quantities of the Earth. *Astron. & Astrop.*, 428:261–285, 2004. [18](#), [19](#), [95](#), [110](#)
- [176] J. Lau, W. T. Hung, and C. S. Cheung. Interpretation of air quality in relation to monitoring station’s surroundings. *Atm. Env.*, 43(4):769–777, 2009. [15](#), [37](#), [82](#), [84](#), [85](#)
- [177] K. M. Lau and H. Weng. Climate signal detection using wavelet transform: How to make a time series sing. *Bull. Amer. Meteor. Soc.*, 76(12):2391–2402, 1995. [27](#), [57](#), [60](#), [136](#)
- [178] S. Lê, J. Josse, and F. Husson. FactoMineR: an R package for multivariate analysis. *J. of Stat. Soft.*, 25(1):1–18, 2008. [28](#), [85](#)

- [179] J. Lean, J. Beer, and R. Bradley. Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.*, 22(23):3195–3198, 1995. [130](#)
- [180] P. Legendre. *Quantitative methods and biogeographic analysis*. In D. J. Garbary and R. G. South, Editors, Evolutionary biogeography of the marine algae of the North Atlantic. NATO Ser. G22, Springer, Berlin, 1990. [55](#)
- [181] P. Legendre. Spatial autocorrelation: trouble or new paradigm? *Ecology*, 74(6):1659–1673, 1993. [53](#), [116](#), [123](#)
- [182] P. Legendre and L. Legendre. *Numerical ecology*. Elsevier Science BV, Amsterdam, 1998. [11](#), [123](#)
- [183] S. Legutke and R. Voss. The Hamburg Atmosphere-Ocean coupled circulation model ECHO-G. DKRZ Technical Report No. 18. 1999. [118](#)
- [184] P. Lehodey, J. Alheit, M. Barange, T. Baumgartner, G. Beaugrand, K. Drinkwater, J. M. Fromentin, S. R. Hare, G. Ottersen, R. I. Perry, et al. Climate variability, fish, and fisheries. *J. Clim.*, 19(20):5009–5030, 2006. [20](#)
- [185] C. E. Leith. The Standard Error of Time-Average Estimates of Climatic Means. *J. of App. Meteor.*, 12:1066–1068, 1973. [141](#)
- [186] C. E. Leith. Theoretical skill of monte carlo forecasts(stochastic atmospheric processes). *Mon. Wea. Rev.*, 102:409–418, 1974. [7](#)
- [187] S. C. Loporati, G. T. Pecl, and J. M. Semmens. Cephalopod hatchling growth: the effects of initial size and seasonal temperatures. *Mar. Bio.*, 151(4):1375–1383, 2007. [137](#)
- [188] U. Ligges and M. Mächler. Scatterplot3d—an R package for Visualizing Multivariate Data. *J. of Stat. Soft.*, 8(11):1–20, 2003. [28](#)
- [189] Z. Lin and M. B. Beck. Understanding complex environmental systems: a dual approach. *Environmetrics*, 18(1):11–26, 2007. [8](#)

- [190] P. Lionello and A. Sanna. Mediterranean wave climate variability and its links with NAO and Indian Monsoon. *Clim. Dyn.*, 25(6):611–623, 2005. [157](#)
- [191] L. Lisiecki. *Paleoclimate time series: New alignment and compositing techniques, a 5.3-Myr benthic $\delta^{18}O$ stack, and analysis of Pliocene-Pleistocene climate transitions*. PhD thesis, Brown University, USA, 2005. [106](#)
- [192] P. C. Liu. Wavelet spectrum analysis and ocean wind waves. In *Wavelets in Geophysics*. Academic Press, New York, 1994. [27](#)
- [193] N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrop. Spa. Sc.*, 39(2):447–462, 1976. [26](#), [39](#), [40](#), [44](#), [49](#), [50](#), [94](#), [101](#), [103](#)
- [194] F. López-Capont. *La etapa pesquera del padre Sarmiento y su época: de los atúnes, y sus transmigraciones y conjeturas, sobre la decadencia de las almadrabas, y sobre los medios para restituirlas*. Caixa de Pontevedra, Pontevedra, 1997. [115](#), [116](#), [117](#)
- [195] E. N. Lorenz. Deterministic Nonperiodic Flow. *J. Atm. Sc.*, 20(2):130–148, 1963. [3](#), [5](#)
- [196] L. A. Lowell, Leonard O. M., and Flocchini R. G. A principal component analysis of sulphur concentrations in the western United States. *Atm. Env.*, 18(4):783–791, 1984. [85](#)
- [197] M. E. Lutcavage, R. W. Brill, G. B. Skomal, B. C. Chase, J. L. Goldstein, and J. Tutein. Tracking adult North Atlantic bluefin tuna (*Thunnus thynnus*) in the northwestern Atlantic using ultrasonic telemetry. *Mar. Bio.*, 137(2):347–358, 2000. [113](#)
- [198] E. Machu, B. Ferret, and V. Garçon. Phytoplankton pigment distribution from SeaWiFS data in the subtropical convergence zone south of Africa: a wavelet analysis. *Geophys. Res. Lett.*, 26(10):1469–1472, 1999. [57](#)
- [199] M. Mangel. *The theoretical biologist’s toolbox*. Cambridge University Press, Cambridge, U.K., 2006. [2](#)

- [200] K. Mangold. *Octopus vulgaris*. In P. R. Boyle (ed) *Cephalopods Life Cycles* Vol. 1. Academic Press, London, 1983. [21](#), [23](#), [134](#), [136](#), [156](#), [158](#)
- [201] M. E. Mann. Little ice age. In *Encyclopedia of Global Environmental Change Vol. 1*. MacCracken & Perry (ed), New York, 2001. [130](#)
- [202] M. E. Mann, R. S. Bradley, and M. K. Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392(6678):779–787, 1998. [119](#), [120](#), [121](#), [122](#), [123](#), [130](#)
- [203] M. E. Mann and J. M. Lees. Robust estimation of background noise and signal detection in climatic time series. *Clim. Ch.*, 33(3):409–445, 1996. [9](#)
- [204] D. Maraun. *What Can We Learn from Climate Data? Methods for Fluctuation, Time/Scale and Phase Analysis*. PhD thesis, University of Potsdam, Potsdam, Germany, 2006. [28](#), [60](#), [67](#), [72](#)
- [205] D. Maraun. F.A.Q. del software SOWAS. <http://tocsy.agnld.uni-potsdam.de/wavelets/>, último acceso, abril 2010. [28](#), [70](#)
- [206] D. Maraun and J. Kurths. Cross wavelet analysis: significance testing and pitfalls. *Nonl. Proc. in Geophys.*, 11(4):505–514, 2004. [24](#), [27](#), [57](#), [60](#), [68](#), [69](#), [70](#), [72](#), [136](#), [143](#), [144](#), [154](#), [155](#)
- [207] D. Maraun, J. Kurths, and M. Holschneider. Nonstationary Gaussian processes in wavelet domain: Synthesis, estimation, and significance testing. *Phys. Rev. E*, 75(1):16707, 2007. [24](#), [27](#), [57](#), [60](#), [64](#), [66](#), [67](#), [69](#), [71](#), [72](#), [136](#), [144](#), [155](#)
- [208] J. Marshall, H. Johnson, and J. Goodman. A Study of the Interaction of the North Atlantic Oscillation with Ocean Circulation. *J. Clim.*, 14(7):1399–1421, 2001. [155](#), [157](#)
- [209] E. M. Martínez-Ataz and Y. D. de Mera-Morales. *Contaminación atmosférica*. Ed. Universidad de Castilla-La Mancha., 2004. [14](#), [15](#), [73](#), [74](#), [76](#), [77](#), [79](#), [80](#), [81](#)

- [210] R. Marty, I. Zin, and Ch. Obled. On adapting PQPFs to fit hydrological needs: the case of flash flood forecasting. *Atm. Sc. Lett.*, 9:73–79, 2008. [7](#)
- [211] A. Mathias, F. Grond, R. Guardans, D. Seese, M. Canela, and H. Diebner. Algorithms for Spectral Analysis of Irregularly Sampled Time Series. *J. of Stat. Soft.*, 11(2):1–27, 2004. [94](#)
- [212] J. M. McArthur, D. T. Donovan, M. F. Thirlwall, B. W. Fouke, and D. Matthey. Strontium isotope profile of the early Toarcian (Jurassic) oceanic anoxic event, the duration of ammonite biozones, and belemnite palaeotemperatures. *Earth & Plan. Sc. Lett.*, 179(2):269–285, 2000. [18](#), [93](#), [100](#)
- [213] J. M. McArthur, P. Doyle, M. J. Leng, K. Reeves, C. T. Williams, R. Garcia-Sánchez, and R. J. Howarth. Testing palaeo-environmental proxies in Jurassic belemnites: Mg/Ca, Sr/Ca, Na/Ca, $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$. *Palaeogeog., Palaeoclim., Palaeoeco.*, 252(3-4):464–480, 2007. [18](#), [92](#), [93](#), [94](#), [106](#), [107](#), [108](#), [111](#)
- [214] J. M. McArthur, R. J. Howarth, and T. R. Bailey. Strontium Isotope Stratigraphy: LOWESS version 3: Best Fit to the Marine Sr-Isotope Curve for 0509 Ma and Accompanying Look-up Table for Deriving Numerical Age. *The J. of Geol.*, 109:155–170, 2001. [100](#)
- [215] P. McCullagh and J. A. Nelder. *Generalized linear models*. Chapman Hall, London, 1989. [53](#)
- [216] K. McGuffie and A. Henderson-Sellers. Forty years of numerical climate modelling. *Int. J. of Clim.*, 21(9):1067–1109, 2001. [5](#)
- [217] D. Melas and U. Fra. *Regional/transboundary transport of air pollution: An introductory science review*. Ed. Universidad de Extremadura, Caceres, 2008. [75](#)
- [218] F. Ménard, F. Marsac, E. Bellier, and B. Cazelles. Climatic oscillations and tuna catch rates in the Indian Ocean: a wavelet approach to time series analysis. *Fish. Ocean.*, 16(1):95–104, 2007. [27](#), [57](#), [144](#), [165](#)

- [219] S. D. Meyers, B. G. Kelly, and J. J. O'Brien. An introduction to wavelet analysis in oceanography and meteorology: With application to the dispersion of Yanai waves. *Mon. Wea. Rev.*, 121(10):2858–2866, 1993. [27](#), [57](#), [62](#), [64](#), [65](#)
- [220] X. Mi, H. Ren, Z. Ouyang, W. Wei, and K. Ma. The use of the Mexican Hat and the Morlet wavelets for detection of ecological patterns. *Plant Ecology*, 179(1):1–19, 2005. [144](#)
- [221] F. Mignard. About the Nyquist Frequency. *Tech. Rep.*, 2005. [103](#), [111](#)
- [222] A. H. Monahan, J. C. Fyfe, M. H. P. Ambaum, D. B. Stephenson, and G. R. North. Empirical orthogonal functions: The medium is the message. *J. Clim.*, 22(24):6501–6514, 2009. [30](#), [38](#), [39](#)
- [223] A. Moreno, J. Pereira, C. Arvanitidis, J. P. Robin, D. Koutsoubas, C. Perales-Raya, M. M. Cunha, E. Balguerias, and V. Denis. Biological variation of *Loligo vulgaris* (Cephalopoda: Loliginidae) in the eastern Atlantic and Mediterranean. *Bull. of Mar. Sc.*, 71(1):515–534, 2002. [134](#)
- [224] M. Mudelsee. TAUEST: A computer program for estimating persistence in unevenly spaced weather/climate time series. *Computers & Geosciences*, 28(1):69–72, 2002. [49](#), [50](#), [51](#)
- [225] M. Mudelsee. *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*. Springer, 2010. [8](#), [10](#), [11](#), [12](#), [22](#), [40](#), [45](#)
- [226] M. Mudelsee, D. Scholz, R. Röthlisberger, D. Fleitmann, A. Mangini, and E. W. Wolff. Climate spectrum estimation in the presence of timescale errors. *Nonl. Proc. in Geophys.*, 16:43–56, 2009. [12](#), [40](#), [49](#), [50](#), [101](#), [102](#), [103](#)
- [227] P. Müller. Constructing climate knowledge with computer models. *Wiley Interdisc. Rev.: Climate Change*, 1(4):565–580, 2010. [5](#)
- [228] P. Müller and H. Von Storch. *Computer modelling in atmospheric and oceanic sciences: building knowledge*. Springer, 2004. [2](#), [3](#), [4](#), [5](#), [118](#), [119](#)

- [229] G. Myhre, E.J. Highwood, K.P. Shine, and F. Stordal. New estimates of radiative forcing due to well mixed greenhouse gases. *Geophys. Res. Lett.*, 25(14):2715–2718, 1998. [121](#)
- [230] Y. Nakagawa, M. Eguchi, and S. Miyashita. Pacific bluefin tuna, *Thunnus orientalis*, larvae utilize energy and nutrients of microbial loop. *Aquaculture*, 267(1-4):83–93, 2007. [121](#), [130](#)
- [231] N. Narayan, A. Paul, S. Mulitza, and M. Schulz. Trends in coastal upwelling intensity during the late 20th century. *Ocean Sc.*, 6:815–823, 2010. [158](#)
- [232] A. Nava. Fourier spectral analysis for unevenly spaced, average value, data. *Computers & Geosciences*, 36(7):853–860, 2010. [11](#), [93](#), [94](#)
- [233] V. Navrotsky. Synergetic Approach and modeling of fish population dynamics. *Annals of Oper. Res.*, 94(1):357–373, 2000. [21](#)
- [234] J. A. Nelder and R. W. M. Wedderburn. Generalized linear models. *J. Roy. Stat. Soc. A*, 135(3):370–384, 1972. [53](#), [54](#), [56](#)
- [235] C. M. Nigmatullin and A. E. Barkovsky. Seasonal grouping of *Octopus vulgaris* off Cap Blanc (Mauritania). In *5th All-Union conference on the Comm. Invertebrates, Minsk (Naroch)*, 1990. [155](#)
- [236] B. L. Norcross and R. F. Shaw. Oceanic and estuarine transport of fish eggs and larvae: a review. *Trans. A. Fish. Soc.*, 113:153–165, 1984. [21](#), [158](#)
- [237] G. Nunnari, S. Dorling, U. Schlink, G. Cawley, R. Foxall, and T. Chatterton. Modelling SO₂ concentration at a point with statistical approaches. *Env. Mod. and Soft.*, 19(10):887–905, 2004. [14](#), [76](#)
- [238] A. Nuttall. Some windows with very good sidelobe behavior. *IEEE Trans. on Acoust., Speech and Sig. Proces.*, 29(1):84–91, 1981. [47](#), [48](#)
- [239] A. Nyíri, M. Gauss, and H. Klein. Transboundary air pollution by main pollutants (S, N, O₃) and PM. Spain. European Monitoring and Evaluation Programme, Meteorological Synthesizing Centre-W, 2009. [14](#), [76](#)

- [240] W. L. Oberkampf, S. M. DeLand, B. M. Rutherford, K. V. Diegert, and K. F. Alvin. Error and uncertainty in modeling and simulation. *Rel. Eng. and Syst. Saf.*, 75(3):333–357, 2002. 4
- [241] C. Obled, G. Bontron, and R. Garçon. Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach. *Atmos. Res.*, 63:303–324, 2002. 7
- [242] National Oceanic and Atmospheric Administration. Kaplan extended sst v2. http://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html, último acceso, Noviembre 2010. 141
- [243] National Oceanic and Atmospheric Administration. The NAO index database. http://www.cpc.noaa.gov/products/precip/CWlink/pna/nao_index.html, último acceso, Noviembre 2010. 141
- [244] National Oceanic and Atmospheric Administration. NOAA optimum interpolation (OI) sea surface temperature (SST) v2. <http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html>, último acceso, Diciembre 2010. 141
- [245] P. E. Olsen. A 40-million-year lake record of Early Mesozoic orbital climatic forcing. *Science*, 234(4778):842–842, 1986. 19, 95, 110
- [246] P. E. Olsen and D. V. Kent. Milankovitch climate forcing in the tropics of Pangea during the Late Triassic. *Palaeogeog., Palaeoclim., Palaeoeco.*, 122(1-4):1–26, 1996. 95, 110
- [247] P. E. Olsen and D. V. Kent. Long-period Milankovitch cycles from the Late Triassic and Early Jurassic of eastern North America and their implications for the calibration of the Early Mesozoic time-scale and the long-term behaviour of the planets. *Phil. Trans. R. Soc. A*, 357:1761–1786, 1999. 19, 95, 110
- [248] N. Oreskes, K. Shrader-Frechette, and K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147):641, 1994. 6

- [249] J. E. Overland, J. Alheit, A. Bakun, J. W. Hurrell, D. L. Mackas, and A. J. Miller. Climate controls on marine ecosystems and fish populations. *J. of Mar. Syst.*, 79(3-4):305–315, 2010. [20](#), [22](#)
- [250] J. Pálffy and P. L. Smith. Synchrony between Early Jurassic extinction, oceanic anoxic event, and the Karoo-Ferrar flood basalt volcanism. *Geology*, 28(8):747, 2000. [17](#)
- [251] J. Pálffy, P. L. Smith, and J. K. Mortensen. A U-Pb and $^{40}\text{Ar}/^{39}\text{Ar}$ time scale for the Jurassic. *Can. J. of Earth Sc.*, 37(6):923–944, 2000. [100](#)
- [252] H. Palike. *Extending the astronomical calibration of the geological time scale*. PhD thesis, University of Cambridge, Cambridge, U.K., 2001. [105](#)
- [253] T. Palmer and R. Hagedorn. *Predictability of weather and climate*. Cambridge Univ Press, Cambridge, U.K., 2006. [118](#)
- [254] T. Parrish and R. L. Curtis. Atmospheric circulation, upwelling, and organic-rich rocks in the Mesozoic and Cenozoic eras. *Palaeogeog., Palaeoclim., Palaeoeco.*, 40(1-3):31–66, 1982. [96](#)
- [255] K. Pearson. On lines and planes of closest fit to systems of points in space. *Phil. Mag.*, 2(11):559–572, 1901. [30](#)
- [256] J. P. Peixoto and A. H. Oort. *Physics of Climate*. AIP Press, New York, 1992. [2](#), [78](#)
- [257] J. L. Pelegrí, J. Arístegui, L. Cana, M. González-Dávila, A. Hernández-Guerra, S. Hernández-León, A. Marrero-Díaz, M. F. Montero, P. Sangra, and M. Santana-Casiano. Coupling between the open ocean and the coastal upwelling region off northwest Africa: water recirculation and off-shore pumping of organic matter. *J. of Mar. Syst.*, 54(1-4):3–37, 2005. [157](#)
- [258] J. Pelt. High Frequency Limits in Periodicity Search from Irregularly Spaced Data. *Baltic Astron.*, 18:83–92, 2009. [103](#)

- [259] T. Penduff, B. Barnier, W. K Dewar, and J. J. O'Brien. Dynamical Response of the Oceanic Eddy Field to the North Atlantic Oscillation: A Model Data Comparison. *J. of Phys. Ocean.*, 34:2615, 2004. [157](#)
- [260] G. Peng, L. M. Leslie, and Y. Shao. *Environmental modelling and prediction*. Springer, 2002. [1](#), [2](#), [4](#)
- [261] P. Pepin. Effect of temperature and size on development, mortality, and survival rates of the pelagic early life history stages of marine fish. *Can. J. of Fish. and Aqua. Sc.*, 48(3):503–518, 1991. [118](#)
- [262] D. B. Percival and A. T. Walden. *Spectral analysis for physical applications*. Cambridge Univ Press, London, 1993. [47](#), [48](#), [50](#)
- [263] D. B. Percival and A. T. Walden. *Wavelet methods for time series analysis*. Cambridge Univ Press, 2006. [27](#), [58](#), [59](#), [61](#), [63](#), [66](#), [165](#)
- [264] D. P. Percival. On estimation of the wavelet variance. *Biometrika*, 82(3):619, 1995. [66](#)
- [265] P. R. Peres-Neto, D. A. Jackson, and K. M. Somers. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Comp. Stat. and Data Analysis*, 49(4):974–997, 2005. [37](#)
- [266] V. Perrier, T. Philipovitch, and C. Basdevant. Wavelet spectra compared to Fourier spectra. *J. of Math. Phys.*, 36(3):1506–1519, 1995. [66](#)
- [267] A. C. Petersen. Philosophy of climate science. *Bull. Amer. Meteor. Soc.*, 81(2):265–271, 2000. [4](#), [6](#)
- [268] V. Petoukhov, A. Ganopolski, V. Brovkin, M. Claussen, A. Eliseev, C. Kubatzki, and S. Rahmstorf. CLIMBER-2: a climate system model of intermediate complexity. Part I: model description and performance for present climate. *Clim. Dyn.*, 16(1):1–17, 2000. [5](#)
- [269] C. Pfister, J. Luterbacher, G. Schwarz-Zanetti, and M. Wegmann. Winter air temperature variations in western Europe during the Early and High Middle Ages (AD 750-1300). *The Holocene*, 8(5):535–552, 1998. [130](#)

- [270] G. J. Pierce, V. D. Valavanis, A. Guerra, P. Jereb, L. Orsi-Relini, J. M. Bellido, I. Katara, U. Piatkowski, J. Pereira, E. Balguerias, et al. A review of cephalopod-environment interactions in European Seas. *Hydrobiologia*, 612:49–70, 2008. [137](#), [155](#)
- [271] J. C. M. Pires, M. C. Pereira, M. C. M. Alvim-Ferraz, and F. G. Martins. Identification of redundant air quality measurements through the use of principal component analysis. *Atm. Env.*, 43(25):3837–3842, 2009. [15](#), [38](#), [82](#), [84](#), [85](#), [86](#), [163](#)
- [272] J. C. M. Pires, S. I. V. Sousa, M. C. Pereira, M. C. M. Alvim-Ferraz, and F. G. Martins. Management of air quality monitoring using principal component and cluster analysis Part I: SO₂ and PM10. *Atm. Env.*, 42(6):1249–1260, 2008. [15](#), [25](#), [37](#), [38](#), [73](#), [75](#), [81](#), [82](#), [84](#), [85](#), [86](#)
- [273] N. G. Pisias and A. C. Mix. Aliasing of the geologic record and the search for long-period Milankovitch cycles. *Paleocean.*, 3(5):613–619, 1988. [9](#), [111](#)
- [274] O. G. Podlaha, J. Mutterlose, and J. Veizer. Preservation of $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ in belemnite rostra from the Jurassic/Early Cretaceous successions. *Ame. J. Sc.*, 298(4):324, 1998. [18](#), [93](#), [94](#), [108](#), [111](#)
- [275] J. Polanco, U. Ganzedo, J. Sáenz, A. M. Caballero-Alfonso, and J. J. Castro-Hernández. Wavelet analysis of correlation among Canary Islands octopus captures per unit effort, sea-surface temperatures and the North Atlantic Oscillation. *Fish. Res.*, 107:177–183, 2011. [21](#), [24](#), [29](#), [133](#), [136](#), [139](#), [140](#), [141](#), [149](#), [150](#), [151](#), [152](#), [162](#)
- [276] R. W. Preisendorfer. *Principal components analysis in Meteorology and Oceanography*. Elsevier, Amsterdam, 1988. [37](#)
- [277] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C*. Cambridge University Press, 1992. [40](#), [42](#), [43](#), [103](#)
- [278] G. D. Price and B. W. Sellwood. Warm palaeotemperatures from high Late Jurassic palaeolatitudes (Falkland Plateau): Ecological, environmental or

- diagenetic controls? *Palaeogeog. Palaeoclim., Palaeoeco.*, 129(3):315–328, 1997. [106](#), [108](#)
- [279] M. B. Priestley. *Spectral analysis and time series*. Academic Press Limited, London, 1981. [24](#), [41](#), [47](#), [48](#), [136](#)
- [280] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. [27](#), [28](#), [85](#), [102](#), [122](#), [123](#), [143](#)
- [281] C. Ravier and J. M. Fromentin. Long-term fluctuations in the eastern Atlantic and Mediterranean bluefin tuna population. *ICES J. of Mar. Sc.*, 58(6):1299–1317, 2001. [116](#), [130](#)
- [282] C. Ravier and J. M. Fromentin. Are the long-term fluctuations in Atlantic bluefin tuna (*Thunnus thynnus*) population related to environmental changes? *Fish. Ocean.*, 13(3):145–160, 2004. [130](#)
- [283] A. Rexfort and J. Mutterlose. Stable isotope records from *Sepia officinalis* key to understanding the ecology of belemnites? *Earth & Plan. Sc. Lett.*, 30:212–221, 2006. [106](#), [107](#), [108](#), [111](#)
- [284] R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang. An improved in situ and satellite SST analysis for climate. *J. Clim.*, 15(13):1609–1625, 2002. [140](#), [141](#), [142](#)
- [285] J. A. Rial, R. A. Pielke, M. Beniston, M. Claussen, J. Canadell, P. Cox, H. Held, N. De Noblet-Ducoudré, R. Prinn, J.F. Reynolds, et al. Nonlinearities, feedbacks and critical thresholds within the Earth’s climate system. *Clim. Ch.*, 65(1):11–38, 2004. [2](#), [3](#)
- [286] M. B Richman. A principal component analysis of sulfur concentrations in the western United States. *Atm. Env.*, 20(3):606–607, 1986. [85](#)
- [287] W. E. Ricker. *Computation and interpretation of biological statistics of fish populations*. Bull. Fish. Res. Bd. Can. 191, 1975. [21](#)
- [288] D. Rind. Complexity and climate. *Science*, 284(5411):105–107, 1999. [2](#), [6](#)

- [289] D. H. Roberts, J. Lehar, and J. W. Dreher. Times series analysis with clean. I. derivation of a spectrum. *The Astron. J.*, 93:968–989, 1987. [94](#), [103](#)
- [290] J. P. Robin and V. Denis. Squid stock fluctuations and water temperature: temporal analysis of English Channel Loliginidae. *J. of App. Eco.*, 36(1):101–110, 1999. [137](#)
- [291] P. M. Robinson. Estimation of a time series model from unequally spaced data. *Stochastic Processes and their applications*, 6(1):9–24, 1977. [49](#)
- [292] J. M. Rodríguez, E. D. Barton, L. Eve, and S. Hernández-León. Mesozooplankton and ichthyoplankton distribution around Gran Canaria, an oceanic island in the NE Atlantic. *D. Sea Res. Part I. Ocean. Res. P.*, 48(10):2161–2183, 2001. [21](#), [158](#)
- [293] J. M. Rodríguez, E. D. Barton, S. Hernández-León, and J. Arístegui. The influence of mesoscale physical processes on the larval fish community in the Canaries CTZ, in summer. *Prog. in Ocean.*, 62:171–188, 2004. [21](#), [158](#)
- [294] J. Rodríguez-Roda. Evolución de la pesquería del atún, *Thunnus thynnus* (L.) del Golfo de Cádiz durante los años 1980 a 1982. *Inv. Pesq*, 47(2):253–262, 1983. [115](#)
- [295] E. Roeckner, G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, et al. The atmospheric general circulation model ECHAM 5. PART I: Model description. Rep. 349, Max-Planck-Institut für Meteorologie, Hamburg, 2003. [5](#)
- [296] I. Rosales, S. Quesada, and S. Robles. Primary and diagenetic isotopic signals in fossils and hemipelagic carbonates: the Lower Jurassic of northern Spain. *Sediment.*, 48(5):1149–1169, 2001. [19](#), [93](#), [95](#), [98](#), [99](#), [106](#), [107](#), [108](#), [109](#)
- [297] I. Rosales, S. Quesada, and S. Robles. Paleotemperature variations of Early Jurassic seawater recorded in geochemical trends of belemnites from the Basque-Cantabrian basin, northern Spain. *Palaeogeog., Palaeoclim.*,

- Palaeoeco.*, 203:253–275, 2004. [17](#), [18](#), [19](#), [92](#), [93](#), [95](#), [96](#), [97](#), [98](#), [99](#), [100](#), [106](#), [109](#)
- [298] I. Rosales, S. Quesada, and S. Robles. Geochemical arguments for identifying second-order sea-level changes in hemipelagic carbonate ramp deposits. *Terra Nova*, 18(4):233–240, 2006. [18](#), [95](#), [98](#), [109](#)
- [299] I. Rosales, S. Robles, and S. Quesada. Elemental and oxygen isotope composition of Early Jurassic belemnites: salinity vs. temperature signals. *J. of Sediment. Res.*, 74(3):342–354, 2004. [19](#), [93](#), [95](#), [98](#), [109](#)
- [300] T. Rouyer, J. M. Fromentin, F. Menard, B. Cazelles, K. Briand, R. Pianet, B. Planque, and N. C. Stenseth. Complex interplays among population dynamics, environmental forcing, and exploitation in fisheries. *Proc. National Academy Sciences, USA*, 105(14):5420, 2008. [4](#), [19](#), [21](#), [22](#), [24](#), [27](#), [136](#)
- [301] T. Rouyer, J. M. Fromentin, N. C. Stenseth, and B. Cazelles. Analysing multiple time series and extending significance testing in wavelet analysis. *Mar. Eco. Prog. Ser.*, 359:11–23, 2008. [57](#), [144](#), [165](#)
- [302] C. M. Rowe, D. B. Loope, R. J. Oglesby, R. Van der Voo, and C. E. Broadwater. Inconsistencies between Pangean reconstructions and basic climate controls. *Science*, 318(5854):1284–1286, 2007. [17](#), [96](#), [112](#)
- [303] G. Rózyński. Long-term evolution of Baltic Sea wave climate near a coastal segment in Poland; its drivers and impacts. *Ocean Eng.*, 37(2-3):186–199, 2010. [157](#)
- [304] D. Rybski, A. Bunde, S. Havlin, and H. Von Storch. Long-term persistence in climate and the detection problem. *Geophys. Res. Lett.*, 33(6):L06718, 2006. [11](#)
- [305] G. Sælen, P. Doyle, and M. R. Talbot. Stable-isotope analyses of belemnite rostra from the Whitby Mudstone, England: surface water conditions during deposition of a marine black shale. *Palaios*, 11(2):97–117, 1996. [92](#), [106](#)

- [306] J. Sáenz, J. Zubillaga, and J. Fernández. Geophysical data analysis using Python. *Computers & Geosciences*, 28(4):457–465, 2002. [28](#)
- [307] M. L. Sánchez, J. L. Casanova, M. C. Ramos, and J. L. Sánchez. Studying the spatial and temporal distribution of SO₂ in an urban area by principal component factor analysis. *Atm. Res.*, 20(1):53 – 65, 1986. [85](#)
- [308] P. Sánchez. Mollusca, Cephalopoda. Fauna Ibérica. *Museo Nacional de Ciencias Naturales, CSIC, Madrid*, 1992. [155](#)
- [309] J. Santiago. The North Atlantic Oscillation and recruitment of temperate tunas. *ICCAT Sci. Pap.*, 97(40):240–249, 1998. [157](#)
- [310] A. Sargin. Statistics and data with R: An applied approach through examples. *J. of Stat. Soft., Book Reviews*, 30(6):1–2, 2009. [28](#)
- [311] J. D. Scargle. Studies in astronomical time series analysis. II- Statistical aspects of spectral analysis of unevenly spaced data. *The Astrop. J.*, 263:835–853, 1982. [26](#), [39](#), [41](#), [42](#), [43](#), [44](#), [49](#), [50](#), [94](#), [101](#), [102](#), [103](#)
- [312] J. D. Scargle. Studies in astronomical time series analysis. III- Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data. *The Astrop. J.*, 343:874–887, 1989. [26](#), [39](#), [94](#), [101](#)
- [313] G. A. Schmidt, I. Aleinov, N. Bell, M. Bauer, S. Bauer, B. Cairns, G. Faluvegi, Y. Hu, N.Y. Kiang, D. Koch, et al. Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *J. Clim.*, 19(2):153–192, 2006. [5](#)
- [314] S. H. Schneider. Geoengineering: could we or should we make it work. *Phil. Trans. R. Soc. A*, 366(1882):3843–3862, 2008. [3](#)
- [315] T. Schneider. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Clim.*, 14:853–871, 2001. [11](#)

- [316] M. Schulz and M. Mudelsee. REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Computers & Geosciences*, 28(3):421–426, 2002. [26](#), [28](#), [40](#), [49](#), [50](#), [94](#), [95](#), [101](#)
- [317] M. Schulz and K. Stattegger. SPECTRUM: Spectral analysis of unevenly spaced paleoclimatic time series. *Computers & Geosciences*, 23(9):929–945, 1997. [9](#), [11](#), [26](#), [40](#), [43](#), [45](#), [47](#), [48](#), [93](#), [94](#), [101](#), [103](#)
- [318] A. Schuster. On the investigation of hidden periodicities with applications to a supposed 26 day period of meteorological phenomena. *Terrestr. Magn.*, 3:13–41, 1898. [41](#)
- [319] W. Schwarzacher. *Cyclostratigraphy and the Milankovitch theory*. Elsevier, Amsterdam, 1993. [93](#), [94](#), [95](#)
- [320] C. R. Scotese. Plate tectonic animation, jurassic to quaternary. PALEOMAP. <http://www.scotese.com>, último acceso, febrero 2011. [16](#)
- [321] C. R. Scotese, R. K. Bambach, C. Barton, R. Van der Voo, and A. M. Ziegler. Paleozoic base maps. *The J. of Geol.*, 87(3):217–277, 1979. [16](#)
- [322] R. Seager, Y. Kushnir, M. Visbeck, N. Naik, J. Miller, G. Krahnmann, and H. Cullen. Causes of Atlantic Ocean Climate Variability between 1958 and 1998. *J. Clim.*, 13(16):2845–2862, 2000. [155](#), [157](#)
- [323] J. H. Seinfeld. *Contaminación atmosférica. Fundamentos físicos y químicos*. Inst. de Estudios de Adm. Local, Madrid, 1978. [13](#), [14](#), [73](#), [74](#)
- [324] J. H. Seinfeld and S. N. Pandis. *Atmospheric chemistry and physics: From air pollution to climate change*. Wiley, New York, 2006. [79](#), [80](#), [81](#), [84](#)
- [325] J. M. Semmens, G. T. Pecl, B. M. Gillanders, C. M. Waluda, E. K. Shea, D. Jouffre, T. Ichii, K. Zumholz, O. N. Katugin, S. C. Leporati, et al. Approaches to resolving cephalopod movement and migration patterns. *Rev. in Fish Bio. and Fish.*, 17(2):401–423, 2007. [156](#)

- [326] S. Shackley, P. Young, S. Parkinson, and B. Wynne. Uncertainty, complexity and concepts of good science in climate change modelling: are GCMs the best tools? *Clim. Ch.*, 38(2):159–205, 1998. [6](#)
- [327] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591, 1965. [143](#)
- [328] M. Sharma. Variations in solar magnetic activity during the last 200 000 years: is there a Sun-climate connection? *Earth & Plan. Sc. Lett.*, 199(3-4):459–472, 2002. [1](#)
- [329] S. Shrestha and F. Kazama. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Env. Mod. and Soft.*, 22(4):464–475, 2007. [15](#), [82](#)
- [330] R. H. Shumway and D. S. Stoffer. *Time series analysis and its applications. With R examples*. Springer, New York, 2000. [122](#), [143](#)
- [331] D. W. Sims, M. J. Genner, A. J. Southward, and S. J. Hawkins. Timing of squid migration reflects North Atlantic climate variability. *Proc. Roy. Soc. B*, 268(1485):2607, 2001. [23](#), [134](#), [137](#)
- [332] K. P. Singh, A. Malik, D. Mohan, and S. Sinha. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) a case study. *Water Res.*, 38(18):3980–3992, 2004. [15](#), [82](#)
- [333] L. A. Smith. What might we learn from climate forecasts? *Proc. National Academy Sciences, USA*, 99(Supp):2487–2492, 2002. [6](#)
- [334] I. Sobrino, L. Silva, J. M. Bellido, and F. Ramos. Rainfall, river discharges and sea temperature as factors affecting abundance of two coastal benthic cephalopod species in the Gulf of Cadiz (SW Spain). *Bull. of Mar. Sc.*, 71(2):851–865, 2002. [134](#)
- [335] A. P. Solari. *New non-linear model for the study and the exploitation of fishery resources*. PhD thesis, Univ. de Las Palmas de Gran Canaria, España, 2008. [21](#), [22](#), [23](#), [54](#), [134](#), [137](#), [158](#), [165](#)

- [336] A. P. Solari, J. M. Martin-Gonzalez, and C. Bas. Stock and recruitment in Baltic cod (*Gadus morhua*): a new, non-linear approach. *ICES J. of Mar. Sc.*, 54(3):427, 1997. [21](#), [22](#)
- [337] S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller. IPCC 2007. Climate Change 2007: the Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007. [2](#)
- [338] B. Sportisse. Box models versus Eulerian models in air pollution modeling. *Atm. Env.*, 35(1):173–178, 2001. [81](#)
- [339] B. Sportisse. A review of current issues in air pollution modeling and simulation. *Comp. Geosc.*, 11(2):159–181, 2007. [77](#), [80](#)
- [340] B. Sportisse. *Fundamentals in air pollution: from processes to modelling*. Springer, Heidelberg, 2010. [13](#), [14](#), [77](#), [78](#)
- [341] D. A. Stainforth, M. R. Allen, E. R. Tredger, and L. A. Smith. Confidence, uncertainty and decision-support relevance in climate predictions. *Phil. Trans. R. Soc. A*, 365(1857):2145–2161, 2007. [6](#)
- [342] N. C. Stenseth, G. Ottersen, J. W. hurrell, A. Mysterud, M. Lima, K. S. Chan, N. G. Yoccoz, and B. Adlandsvik. Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *Proc. Roy. Soc. B*, 270(1529):2087–2096, 2003. [20](#), [139](#)
- [343] D. B. Stephenson, C. A. S. Coelho, F. J. Doblas-Reyes, and M. Balmaseda. Forecast assimilation: a unified framework for the combination of multi-model weather and climate predictions. *Tellus*, 57:253–264, 2005. [7](#)
- [344] A. C. Stern. *Air pollution. Vol. 5, Air quality management*. Academic Press, New York. [74](#), [75](#)
- [345] H. Storch and A. Navarra. *Analysis of Climate Variability: Applications and Statistical Techniques*. Springer, Heilderberg, 1999. [4](#), [8](#), [9](#), [22](#)

- [346] G. Suan, B. Pittet, I. Bour, E. Mattioli, L. V. Duarte, and S. Mailliot. Duration of the Early Toarcian carbon isotope excursion deduced from spectral analysis: Consequence for its possible causes. *Earth & Plan. Sc. Lett.*, 267(3-4):666–679, 2008. [17](#), [92](#), [93](#), [110](#)
- [347] F. C. Tan, J. D. Hudson, and M. L. Keith. Jurassic (callovian) paleotemperatures from Scotland. *Earth & Plan. Sc. Lett.*, 9(5):421–426, 1970. [106](#)
- [348] R. J. Telford, E. Heegaard, and H. J. B. Birks. All age-depth models are wrong: but how badly? *Quat. Sc. Rev.*, 23(1-2):1–5, 2004. [12](#)
- [349] D. J. Thomson. Time series analysis of Holocene climate data. *Phil. Trans. R. Soc. A*, 330(1615):601–616, 1990. [9](#), [10](#), [12](#)
- [350] C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, 79(1):61–78, 1998. [24](#), [27](#), [57](#), [58](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [68](#), [71](#), [136](#), [144](#)
- [351] C. Torrence and P. J. Webster. Interdecadal changes in the ENSO-monsoon system. *J. Clim.*, 12(8):2679–2690, 1999. [144](#)
- [352] I. Uotani, T. Saito, K. Hiranuma, and Y. Nishikawa. Feeding habit of bluefin tuna *Thunnus thynnus* larvae in the western North Pacific Ocean. *Bull. Jap. Soc. Sci. Fish.*, 56(5):713–717, 1990. [114](#)
- [353] B. van de Schootbrugge, J. M. McArthur, T. R. Bailey, Y. Rosenthal, J. D. Wright, and K. G. Miller. Toarcian oceanic anoxic event: An assessment of global causes using belemnite C isotope records. *Paleocean.*, 20, 2005. [92](#)
- [354] H. P. A. Van Dongen, E. Olofsen, J. H. Van Hartevelt, and E. W. Kruyt. A procedure of multiple period searching in unequally spaced time-series with the Lomb-Scargle method. *Bio. Rhy. Res.*, 30(2):149–177, 1999. [42](#), [43](#), [44](#), [45](#), [94](#), [103](#), [104](#), [109](#), [111](#)
- [355] W. F. Van Heukelem. Environmental control of reproduction and life span in Octopus: an hypothesis. In Stancyk S.E. (ed) Reproductive ecology of marine invertebrates, Univ. South Carolina, Columbia, 1979. [21](#), [156](#), [158](#)

- [356] P. Vanicek. Further development and properties of the spectral analysis by least-squares. *Astrop. Sp. Sc.*, 12(1):10–33, 1971. [39](#)
- [357] D. A. Vasseur and P. Yodzis. The color of environmental noise. *Ecology*, 85(4):1146–1152, 2004. [144](#)
- [358] W. N. Venables and B. D. Ripley. *Modern applied statistics with S*. Springer, New York, 4 edition, 2002. [28](#), [122](#), [123](#), [143](#)
- [359] D. G. Victor. On the regulation of geoengineering. *Oxford Rev. of Eco. Pol.*, 24(2):322, 2008. [3](#)
- [360] R. Villanueva. Experimental rearing and growth of planktonic *Octopus vulgaris* from hatching to settlement. *Can. J. of Fish. and Aqua. Sc.*, 52(12):2639–2650, 1995. [21](#), [134](#), [156](#), [158](#)
- [361] M. Visbeck, E. Chassignet, R. Curry, T. Delworth, B. Dickson, and G. Krahnmann. *The North Atlantic Oscillation*, chapter 6. The Ocean’s Response to North Atlantic Oscillation Variability. edited by J. Hurrell, Y. Kushnir, G. Ottersen and M. Visbeck. AGU, 2003. [155](#), [157](#)
- [362] M. Visbeck, H. Cullen, G. Krahnmann, and N. Naik. Ocean model’s response to North Atlantic Oscillation-like wind forcing. *Geophys. Res. Lett.*, 25(24):4521–4524, 1998. [155](#), [157](#)
- [363] H. Von Storch and G. Flöser. *Models in environmental research. Proc. Second GKSS School on Env. Res.* Springer Verlag, 2001. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [8](#)
- [364] H. Von Storch, J. S. Von Storch, and P. Müller. Noise in the climate system ubiquitous, constitutive and concealing. In Björn, E. and Schmid, (Eds.) *Mathematics Unlimited-2001 and Beyond. Part II*. Springer, 2001. [9](#)
- [365] H. Von Storch and F. W. Zwiers. *Statistical analysis in climate research*. Cambridge University Press, Cambridge, U.K., 1999. [1](#), [2](#), [4](#), [8](#), [9](#), [11](#), [12](#), [24](#), [25](#), [30](#), [32](#), [33](#), [38](#), [85](#), [121](#), [136](#), [142](#)

- [366] J. S. Von Storch, P. Müller, and E. Bauer. Climate variability in millennium integrations with coupled atmosphere-ocean GCMs: A spectral view. *Clim. Dyn.*, 17(5):375–389, 2001. [9](#)
- [367] J. S. Von Storch, P. Müller, R. J. Stouffer, R. Voss, and S. F. B. Tett. Variability of deep-ocean mass transport: spectral shapes and spatial scales. *J. Clim.*, 13(11):1916–1935, 2000. [9](#)
- [368] C. M. Waluda, P. N. Trathan, and P. G. Rodhouse. Influence of oceanographic variability on recruitment in the *Illex argentinus* (Cephalopoda: Ommastrephidae) fishery in the South Atlantic. *Mar. Eco. Prog. Ser.*, 183:159–167, 1999. [137](#)
- [369] K. Wark and C. F. Warmer. *Contaminación del aire: origen y control*. Limusa, México, 1994. [13](#), [74](#), [75](#), [76](#), [77](#), [79](#)
- [370] G. P. Weedon. The detection and illustration of regular sedimentary cycles using Walsh power spectra and filtering, with examples from the Lias of Switzerland. *J. of the Geo. Soc.*, 146(1):133, 1989. [95](#), [110](#)
- [371] G. P. Weedon. *Time-series analysis and cyclostratigraphy: examining stratigraphic records of environmental cycles*. Cambridge Univ Press, Cambridge, 2003. [9](#), [10](#), [11](#), [12](#), [18](#), [24](#), [26](#), [93](#), [94](#), [95](#), [101](#), [105](#), [136](#)
- [372] G. P. Weedon and H. C. Jenkyns. Cyclostratigraphy and the Early Jurassic timescale: Data from the Belemnite Marls, Dorset, southern England. *Geo. Soc. Ame. Bull.*, 111:1823–1840, 1999. [19](#), [93](#), [110](#), [164](#)
- [373] G. P. Weedon, H. C. Jenkyns, A. L. Coe, and S. P. Hesselbo. Astronomical calibration of the Jurassic time-scale from cyclostratigraphy in British mudrock formations. *Phil. Trans. R. Soc. A*, 357:1787–1813, 1999. [93](#), [110](#)
- [374] P. Welch. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio & Electroac.*, 15(2):70–73, 1967. [26](#), [45](#), [47](#)

- [375] H. Wierzbowski. Carbon and oxygen isotope composition of Oxfordian-Early Kimmeridgian belemnite rostra: palaeoenvironmental implications for Late Jurassic seas. *Palaeogeog., Palaeoclim., Palaeoeco.*, 203(1-2):153–168, 2004. [106](#)
- [376] D. S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Academic Press, London, 1995. [8](#), [11](#), [25](#), [30](#), [31](#), [32](#), [34](#), [36](#), [37](#), [38](#), [85](#), [121](#), [141](#)
- [377] S. G. Wilson, M. E. Lutcavage, R. W. Brill, M. P. Genovese, A. B. Cooper, and A. W. Everly. Movements of bluefin tuna (*Thunnus thynnus*) in the northwestern Atlantic Ocean recorded by pop-up satellite archival tags. *Mar. Bio.*, 146(2):409–423, 2005. [113](#)
- [378] A. Witt. Interactive comment on Orbital forcings of the Earths climate in wavelet domain by A. V. Glushkov et al. *Clim. of the Past Disc.*, 1:S127–S137, 2005. [11](#), [12](#)
- [379] S. N. Wood. *Generalized additive models: an introduction with R*. Chapman & Hall/CRC Press, Boca Raton, FL., 2006. [53](#), [143](#)
- [380] R. J. Wootton. *Fish ecology*. Blakie & Son Lim., Glasgow, 1992. [116](#)
- [381] W. H. O World-Health-Organization. *Air Quality Guidelines for Europe*. Number 91. WHO Reg. Pub. European series; No. 91, 2000. [15](#), [73](#), [74](#)
- [382] W. H. O World-Health-Organization. *WHO Air Quality Guidelines for particular matter, ozone, dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment*. WHO Publications, Germany, 2005. [74](#), [75](#)
- [383] Z. Wu, N. E. Huang, S. R. Long, and C. K. Peng. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc. National Academy Sciences, USA*, 104(38):14889, 2007. [12](#), [109](#)
- [384] D. A. Wunderlin, M. P. Diaz, M. V. Ame, S. F. Pesce, A. C. Hued, M. A. Bistoni, et al. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). *Water Res.*, 35(12):2881–2894, 2001. [15](#), [82](#)

- [385] C. Wunsch. The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull. Amer. Meteor. Soc.*, 80(2):245–255, 1999. [8](#), [9](#), [12](#)
- [386] C. Wunsch and D. E. Gunn. A densely sampled core and climate variable aliasing. *Geo-marine Lett.*, 23(1):64–71, 2003. [9](#), [111](#)
- [387] P. Yiou, E. Baert, and M. F. Loutre. Spectral analysis of climate data. *Surv. in Geophys.*, 17(6):619–663, 1996. [11](#), [26](#)
- [388] P. Young, S. Parkinson, and M. Lees. Simplicity out of complexity in environmental modelling: Occam’s razor revisited. *J. of App. Stat.*, 23(2-3):165–210, 1996. [81](#)
- [389] P. Zannetti. *Air Pollution Modeling: Theories, computational methods and available software*. WIT Press, New York, 1990. [75](#), [77](#), [78](#), [79](#), [80](#)
- [390] P. A. Ziegler. *Geological atlas of western and central Europe*. Ed. Shell Int. Pet. Geol. Soc., 1990. [16](#), [96](#)
- [391] E. Zorita, J. F. González-Rouco, H. Von Storch, J. P. Montávez, and F. Valero. Natural and anthropogenic modes of surface temperature variations in the last thousand years. *Geophys. Res. Lett.*, 32(8):755–762, 2005. [118](#), [119](#), [122](#), [123](#), [130](#)
- [392] F. W. Zwiers and H. Von Storch. On the role of statistics in climate research. *Int. J. of Clim.*, 24(6):665–680, 2004. [8](#), [9](#), [22](#), [25](#), [38](#)