



**Universitat de les
Illes Balears**

Long-term extreme wave height events in the Balearic Sea: characterization, variability and prediction

Memòria presentada per
Bartomeu Cañellas Moragues
per obtenir el grau de Doctor en Ciències Marines per la
Universitat de les Illes Balears

Director:
Dr. Alejandro Orfila Förster
Científic titular del CSIC
IMEDEA (CSIC-UIB)

Ponent:
Dr. Misericòrdia Ramón Juanpere
Catedràtica de la Universitat de les
Illes Balears

A mis padres

Agradecimientos

En primer lugar me gustaría agradecer a Joaquín Tintoré y Alejandro Orfila por la confianza y dedicación al desarrollo de esta Tesis y por su apoyo en todos los aspectos necesarios para que este proyecto haya terminado siendo una realidad.

Quisiera también agradecer a Fernando Méndez, Salvador Balle, Melisa Menéndez, Tomeu Garau, Biel Jordà, Alberto Álvarez y Lluís Gómez por compartir conmigo una parte de su conocimiento y haber aportado importantes ideas tanto técnicas como teóricas a los trabajos presentados aquí.

Finalmente, me gustaría agradecer a los incondicionales, aquellos que por alguna razón siempre han estado allí durante estos últimos años compartiendo los buenos momentos y porque no decirlo, los malos también (que han sido muy pocos): especialmente a Mar por todas las cosas buenas que me ha dado, a Amaya por haber compartido destino, alegrías y penas en el mundo de la investigación, a Jano por haber sido mucho más que sólo un director de Tesis y a Enrique por redescubrirme la pasión por la fotografía. No podría terminar sin agradecer a mi familia y amigos por todo lo que hemos pasado juntos: Toni, Tomeu, Gudi, Toñin, Pep, Eze, Joan, Emilio, Enric, Pepo, Gonçal, Mamen, Pedro, Anita y Meri, gràcies.

Summary

This Thesis is structured to guide the reader through an extensive work about wave climate in the Balearic Sea, its coastal areas and more generally the Western Mediterranean Sea. Specifically this Thesis studies and characterizes the oceanographic conditions obtained from different sources: numerical models or synthetic data and oceanographic instruments.

This Thesis is structured as follows: Chapter 1 provides an introduction to the ocean surface waves, their characteristics as well as the available data gathered in the studied area. A comparison between the different sources of data has been also carried out to check the accuracy of the numerical models. A brief introduction on time series analysis applied to wind waves is presented in Chapter 2, with special emphasis on the specific methodologies used to analyze wave data. Chapter 3 introduces the study area through a preliminary characterization of the meteorological aspects and wave climate conditions in the Western Mediterranean. Moreover, this chapter presents a description of the wave energy distribution around the Balearic Islands and the joint distributions between wave heights and peak periods that defines the most probable energy conditions in an area.

The first published study is presented in Chapter 4 where a characterization of the long-term extreme Significant Wave Height distributions in the Balearic Sea is made. In order to obtain the long-term extreme statistics the Peak Over Threshold (POT) method is used due to its efficiency with rare extremal populations. A 50-year return period quantile around the Balearic Islands has been determined through the probabilistic distribution of the most severe storms during the period ranging from 1958 to 2001.

The second study, presented in Chapter 5, examines teleconnections between the North Atlantic Oscillation and the wave climate of the Northwestern Mediterranean Sea, defined by its monthly mean significant wave height and the 95th percentile significant wave height. This study makes use of the EOF encoding ability to split the temporal and spatial variability prior to search for teleconnections. The North Atlantic Oscillation has been found to drive some of the variability of the Northwestern Mediterranean wave climate with an instantaneous response at monthly time scales.

Finally, the last study presented in Chapter 6 explores the suitability of a soft computing technique to predict Significant Wave Heights in the Mediterranean Sea. Using the HIPOCAS database that contains meteorological and oceanographic records from year 1958 to year 2001 a Genetic Algorithm was trained to reconstruct the 44-year time series of Significant Wave Height at different coastal points of the Balearic Islands. The equations obtained were used to predict Significant Wave Heights measured at several buoys. We obtained that SWH at each buoy can be predicted with a

1-hour ahead forecast time using only wind data from the previous 12 hours. A general conclusions and future work are presented at the end of this Thesis in Chapters 7 and 8.

Resumen

Esta tesis está estructurada para guiar al lector a través de un amplio trabajo sobre el clima marítimo en el Mar Balear, sus zonas litorales y en general el Mediterráneo occidental. En concreto, esta Tesis estudia y caracteriza las condiciones oceanográficas obtenidas a través de diferentes fuentes: modelos numéricos o datos sintéticos e instrumentos oceanográficos.

Esta tesis está estructurada de la siguiente manera: el Capítulo 1 proporciona una introducción a las ondas oceánicas, sus características principales así como los datos disponibles para su estudio que se han logrado reunir en la zona del Mediterráneo occidental. Se ha realizado una comparación de las diferentes fuentes de datos para comprobar y validar la exactitud de los datos proporcionados por los modelos numéricos. El Capítulo 2 hace hincapié en el análisis de series temporales, con especial interés en el análisis del oleaje producido por el viento. El Capítulo 3 presenta el área de estudio a través de una caracterización preliminar de los aspectos meteorológicos y las principales características del clima marítimo en el Mediterráneo occidental. Por otra parte, este capítulo presenta una breve descripción del contenido energético del oleaje en las Islas Baleares a través las funciones de distribución conjuntas entre la altura de ola y el periodo de pico, que permiten definir el contenido energético del oleaje estadísticamente más probable en el área de estudio.

El Capítulo 4 recoge el primer estudio publicado durante esta Tesis. Este estudio consiste en la caracterización del régimen extremal del oleaje a través de distintas funciones de distribución en el Mar Balear. Para la caracterización estadística del régimen extremal se ha hecho uso de la función de distribución de Picos Sobre Umbral o "Peaks Over Threshold" debido a su gran eficiencia en la caracterización de eventos extremos. A través de este estudio se han obtenido los cuantiles correspondientes a un período de retorno de 50 años alrededor del Mar Balear, mediante el estudio de las tormentas marítimas más severas de los últimos 45 años.

El segundo estudio publicado se presenta en el Capítulo 5, este estudio examina las teleconexiones entre la Oscilación del Atlántico Norte y el clima marítimo del Mediterráneo Noroccidental, caracterizado por la altura de ola media mensual y la altura de ola correspondiente al percentil 95 de la función de distribución empírica. Este estudio hace uso de la habilidad de codificación del análisis EOF antes realizar la búsqueda de correlaciones, para de esta manera separar la variabilidad espacial y temporal del proceso. Se ha hallado que la Oscilación del Atlántico Norte afecta a la variabilidad tanto espacial como temporal del clima marítimo del Mediterráneo Noroccidental a una escala mensual.

Finalmente, el último estudio publicado y presentado en esta Tesis en el Capítulo 6 explora la oportunidad de realizar predicciones de oleaje mediante el uso de algoritmos

genéticos, una herramienta de bajo coste computacional. El uso de la base de datos HIPOCAS ha permitido el entrenamiento del algoritmo genético gracias al amplio registro meteorológico y oceanográfico que contiene. De esta manera ha sido posible reconstruir las propias series de oleaje presentes en la base de datos HIPOCAS así como realizar predicciones a corto plazo basadas en las ecuaciones predictivas obtenidas a través de los algoritmos genéticos. Se ha obtenido además que la altura de ola significativa puede ser predicha en las boyas con un horizonte de predicción de una hora utilizando sólo el registro de viento de las 12 horas anteriores. Las conclusiones generales y las recomendaciones de futuro están incluidas al final de esta Tesis, en los Capítulos 7 y 8.

Motivation and Aims

Oceans are one of the most dynamical systems in the Earth. Many processes interact influencing them from the surface to the deepest abyssal areas. The atmosphere is one of the biggest contributor to the ocean variability through the air-sea boundary. The atmosphere gains much of its heat at the interface in tropical latitudes by back radiation from the heated ocean in the form of latent heat from the evaporation of water at the interface. In turn, the atmosphere heats the ocean surface in higher latitudes. This atmospheric motion at the interface produces an energy transference to the ocean surface generating large scale currents among other important processes.

The wind is another important mechanism that influences the ocean behavior. The wind acts over the ocean surface through the free surface boundary layer where a transfer of energy and momentum is continuously made, giving rise to waves. A key factor of this transfer is the difference in speed of these two fluid layers. The frictional stress exerted by a moving fluid is proportional to the square of the fluid celerity, so the wind stress exerted upon the sea surface is proportional to the square of the wind speed.

Wind records analyzed at large temporal scales (e.g. 30 years) defines the meteorological climatology in an area, and thus the wind climatology will determine the "wave climatology". The "wave climatology", in fact, is what we know as wave climate and it is the counterpart of the meteorological climatology applied to the ocean surface waves. In fact, the definition of Wave Climate can be explained as the succession of sea states that characterizes the typical behavior of the waves in an area.

A Detailed assessment of the wave climate is a previous requirement for all human activities in the coastal zone. Beach nourishment, port design and operability, dispersion and diffusion of pollutants are some examples that require a precise knowledge of the wave climate. On the other hand, wave climate analysis requires a large amount of data to ensure its statistical significance. In the last decades these data have been collected using scalar and directional wave buoys moored at specific locations, providing high temporal resolution records. Conversely satellites have been used to overcome the spatial lack of data but the problem of having a large amount of spatial and temporal wave records were still unresolved (Krogstad and Barstow, 1999).

To date, statistical analysis of wave climate has been carried out with relatively short data sets, e.g. 10 years of data for the satellite altimeters, or 20 years of wave data from deep and shallow water oceanographic buoys. Altimeter data has, as principal shortcoming, its temporal inhomogeneity and a coarse spatial resolution in areas like the Western Mediterranean, marked by a complex orography. Areas with a complex orography usually presents small scale wave climate variabilities not detectable with altimeter. Moreover, altimeter data is not always available when big storms de-

velop hazardous wave fields near the coast. Therefore, the need for new databases in wave climate and analysis is one of the important issues that the scientific community had to deal with.

Nowadays, wave generation models are an alternative to avoid the usual lack of data in ocean and atmospheric studies. Numerical models can be now implemented in very fine grids and have become a powerful tool not only for engineering or prediction purposes but also for climate studies involving large temporal periods.

In this Thesis we make use of the largest temporal oceanographic database available for the Mediterranean Sea, the HIPOCAS database a 44-year oceanographic database that covers the area between Corsica and the Strait of Gibraltar. This database opens new possibilities to study the wave climate and its extreme wave heights events.

This Thesis aims, first to improve the knowledge about wave energy conditions in the Balearic Sea and its coastal areas, quantifying statistically the intensity and recurrence of extreme wave heights that would represent a risk for the management and security of coastal areas and offshore structures. Second, to explore the possible relations between wave climate and distant patterns of climatic variability. Finally, to develop a low-cost operational soft computing technique to predict wave heights in the coastal zone of the Balearic Sea.

Contents

Summary	i
Resumen	iii
Motivation and Aims	v
List of Figures	ix
List of Tables	xiii
1 Ocean surface waves	1
1.1 Characteristics of ocean surface waves	1
1.2 Wind generated waves	2
1.2.1 Wave generation models	3
1.3 Oceanographic wave data	4
1.3.1 The HIPOCAS project: a 44-year wave reanalysis	4
1.3.2 Oceanographic Buoy Data	5
1.3.3 Comparison and validation between the HIPOCAS database and Buoy data.	7
2 Time series analysis applied to wind waves	11
2.1 Time domain	11
2.1.1 Frequency domain analysis	13
2.1.2 Probabilistic approach	16
2.1.3 Long-term analysis	16
2.2 Extreme value distributions	18
2.2.1 The Generalized Extreme value Distribution (GEV)	19
2.2.2 The Peaks Over Threshold method (POT)	20
2.2.3 Parameter estimation	21
2.3 Empirical orthogonal function analysis	22

3	Study area description	25
3.1	The NorthWestern Mediterranean wave climate: preliminary characterization	26
3.1.1	Mean Significant Wave Height distribution	27
3.1.2	Wave Energy distribution in the Balearic Islands coast	30
3.1.3	Joint distribution of SWH and periods	32
4	Extreme wave height estimation in the Balearic Sea	35
4.1	Methodology	36
4.1.1	Threshold and time span selection for the GPD-P	36
4.1.2	Model selection	38
4.2	Results	40
4.2.1	Return Levels for Hs50	40
4.3	Discussion	41
4.4	Conclusions	43
5	Influence of the NAO on the Northwestern Mediterranean wave climate	45
5.1	Introduction	45
5.2	Methodology	48
5.3	Results	49
5.3.1	Spatial Variance EOF modes for monthly mean SWH	49
5.3.2	Spatial Variance EOF modes for 95th percentile monthly mean SWH	51
5.4	Discussion and Conclusions	54
6	Wave Height prediction in the Western Mediterranean using Genetic Algorithms	61
6.1	Introduction	61
6.2	Data and Methods	63
6.3	Results	65
6.3.1	Prediction of HIPOCAS data	66
6.3.2	Prediction of Buoy data	70
6.4	Conclusions	72
7	General Conclusions	75
8	Recommendations for future work	77
	Bibliography	79
	Publications	85

List of Figures

1.1	Approximate distribution of ocean surface wave energy, disturbing and restoration forces. <i>Figure adapted from the Coastal Engineering Manual (Thompson, 2002)</i>	2
1.2	Geographic location of the three different domains covered by the atmospheric and oceanographic models REMO, HAMSOM and WAM.	5
1.3	Geographic location of the study area and 0.125° resolution HIPOCAS grid in the Western Mediterranean. Red triangles correspond to the location of the seven oceanographic buoys available in the Western Mediterranean area. The four buoys used in this Thesis marked as: Maó (MH), Dragonera (DR), Capdepera (CP) and Cabrera (CB)	6
1.4	A moored WAVESCAN buoy.	7
1.5	Aanderaa DB 4700 (Cabrera) oceanographic buoy equipped with meteorological and oceanographic sensors.	8
1.6	Point by point comparison between the SWH recorded by the MH buoy and the SWH time series provided by the closest node to the buoy, the HIPOCAS node 1320.	9
1.7	Point by point comparison between the wind module (U) recorded by the MH buoy and the wind module (U) time series provided by the closest node to the buoy, the HIPOCAS node 1320.	9
1.8	Logn-Normal probability plots for the SWH distribution. HIPOCAS database is shown as a red line and the buoy data is shown in blue.	10
2.1	Example of sea surface elevation from mean water level.	12
2.2	Example of the autocorrelation function for the sea surface elevation	14
2.3	Example of the power density spectra of a wave record, the peak frequency is $f_p = 1.1rad/s$	15
2.4	20-year significant wave height time series, from the HIPOCAS database.	17
2.5	Example of number of selected extremal SWH with the annual maxima method (squares) vs. peaks over threshold method (asterisks)	20

3.1	View of the Western Mediterranean Sea included in the HIPOCAS reanalysis database, different subbasins are shown: the Alboran Sea, the Balearic Basin and the Gulf of Lions, among others.	26
3.2	Mean significant wave height obtained from the 44-year HIPOCAS database in the Western Mediterranean area.	28
3.3	Monthly evolution during the 44 years of the spatial SWH patterns in the Western Mediterranean area. Top-right figure correspond to January, Bottom-left figure correspond to December.	29
3.4	Mean wave energy flux histogram obtained from the HIPOCAS database node 1320, near Maó.	30
3.5	Mean wave energy flux histogram obtained from the HIPOCAS database node 1432, near Dragonera island.	31
3.6	Mean wave energy flux histogram obtained from the HIPOCAS database node 1611, near Cabrera island.	31
3.7	$SWH - Tp$ empirical joint distribution obtained from the HIPOCAS database node 1320, near Maó.	32
3.8	$SWH - Tp$ empirical joint distribution obtained from the HIPOCAS database node 1432, near Dragonera Island.	33
3.9	$SWH - Tp$ empirical joint distribution obtained from the HIPOCAS database node 1611, near Cabrera Island.	33
4.1	Example of mean excess plot over a threshold level u	36
4.2	Quantil-Quantil plot for the W-statistic, $u=4.2$ meters	38
4.3	Quantil-Quantil plot for the W-statistic, $u=1.7$ meters	38
4.4	Spatial distribution of the shape parameter ξ	39
4.5	Spatial distribution of the likelihood ratio test	40
4.6	Spatial distribution of the 50-year significant wave height levels	41
4.7	Intensity and direction of the highest SWH registered in the HIPOCAS database	42
5.1	Geographic location of the study area.	47
5.2	Variance explained by the first six EOFs, for the monthly SWH. Grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.	50
5.3	(a-top) First spatial variance EOF mode for the monthly mean SWH, (b-bottom) Second spatial variance EOF mode for the monthly mean SWH.	51
5.4	(a-top) Time evolution of the spatial mean computed from 528 months for SWH, (b, c) Time amplitudes for the first two EOFs.	52

5.5	Cross correlation between second spatial variance EOF mode time amplitude and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations.	53
5.6	Variance explained by the first six EOFs, for the monthly 95th percentile SWH. Grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.	54
5.7	(a-top left) First spatial variance EOF mode for the 95th percentile SWH, (b-top right) Second spatial variance EOF mode for the 95th percentile SWH, (c-bottom left) Third spatial variance EOF mode for the 95th percentile SWH, and (d-bottom right) Fourth spatial variance EOF mode for the 95th percentile SWH.	55
5.8	(a-top) Time evolution of the spatial mean computed from 528 months for the 95th percentile SWH, (b, c, d, e) 2nd ,3rd and 4th EOF time amplitudes.	56
5.9	(a-top) Cross-correlation between second, third and fourth spatial variance EOF mode time amplitudes and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations. . . .	57
5.10	(a-top) Snapshot corresponding to the HIPOCAS database (12/01/2001), it shows the SWH field associated to the South-West winds, generated near the Corsica Island and propagated towards the Gulf of Lions. (b-bottom) Snapshot corresponding to the HIPOCAS database (12/28/1980), it shows the fetch associated to the severe North-East winds, generated in the Gulf of Genoa and propagated towards the Catalonia coasts, Valencia coasts and the Balearic Islands.	58
6.1	Geographic location of the study area. Black dots correspond to the HIPOCAS coastal nodes and triangles to the oceanographic buoys, MH (Mahon), DR (Dragonera) and CB (Cabrera).	64
6.2	Linear correlation (r^2) obtained at the 49 HIPOCAS coastal nodes, between the genetic algorithm output and the numerical model data. Dashed line corresponds to the mean correlation calculated over the nodes $r^2 = 0.90$. Filled dots correspond to the HIPOCAS nodes presented in the Results section, which are the closest to the available buoys.	67
6.3	Left panels: Hourly time series comparison, between the genetic algorithm predicted SWH and the HIPOCAS nodes SWH (1320, 1432 and 1611). Black triangles represent the GA predicted SWH, grey asterisks correspond to the HIPOCAS database SWH. Right panels: Scatter diagrams comparing GA SWH and the HIPOCAS SWH data, unit diagonal corresponds to a perfect fit.	68

-
- 6.4 Left panels: Contour plot of the absolute error between HIPOCAS database SWH and predicted SWH at the selected HIPOCAS nodes against the wind intensity U , countour lines indicate the percentage of the population with the indicated error. Right panels: idem but compared against the mean wind direction $\bar{\theta}$ 69
- 6.5 Left Panels: SWH measured at the three buoys (MH, DR and CB)(grey asterisks) compared with the predictive equations SWH (black triangles) computed using the buoy's wind records. Right Panels: scatter diagram of these data. 71
-

List of Tables

2.1	Probability distribution functions for long-term statistics extracted from the Coastal Engineering Manual (Thompson, 2002)	18
2.2	The three types theorem limiting functions	19

Chapter 1

Ocean surface waves

1.1 Characteristics of ocean surface waves

Surface waves are always present in the ocean or any body of water when the action of different forces, like the wind, transfers energy to the ocean surface. For coastal engineering and coastal management purposes wind waves are the most investigated since a thorough understanding of the interaction of waves with offshore structures has now become a capital factor in the safe and economical design of such structures. Ocean surface waves controls for instance a large number of physical processes such as the morphodynamic states of sandy beaches or the design and operability of coastal structures. Therefore, a detailed knowledge on surface waves as well their energy distribution is essential for establishing wave conditions around offshore structures and coastal areas.

Ocean waves are the result of forces acting on a fluid tending to deform it, like the superficial tension, and restoration forces that tend to maintain a level fluid surface, like gravity (Massel, 1996). Four types of ocean waves can be distinguished: capillary waves, infragravity waves, long-period waves and transtidal waves. Wave type classification is based on the characteristic frequency range of each group, from capillary waves with periods around 10^2 cycles/second to tidal oscillations with periods of $10^{-5} - 10^{-6}$ cycles/second. The frequency range associated with the generation forces is very wide and ocean surface response covers a broad range of wave lengths and periods. An schematic representation of the energy contained in ocean surface waves and its related primary disturbing force as well as their restoration force is shown in Figure 1.1.

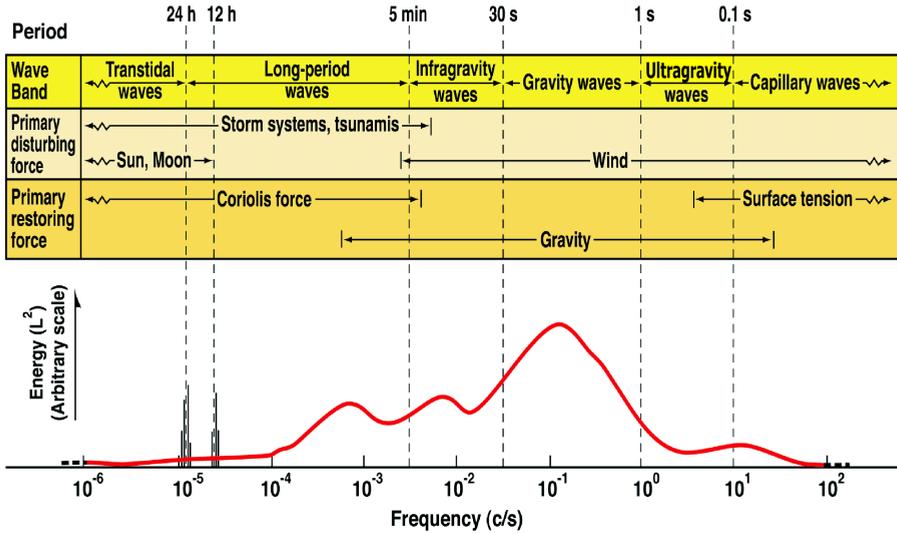


Fig. 1.1: Approximate distribution of ocean surface wave energy, disturbing and restoration forces. Figure adapted from the *Coastal Engineering Manual* (Thompson, 2002)

1.2 Wind generated waves

Wind generated waves, with periods between 1 and 30 seconds, are often the most well-known type of waves, because of their temporal range and energetic surf beat viewable in all coasts across the oceans as well as in any other body of water under the action of gravity and wind stress.

Wind waves belong to the gravity waves group, i.e. waves produced when the surface of the ocean is displaced from the equilibrium state due to the action of shear stresses (e.g. wind or pressure gradients) and the force of gravity acting as the restoration force. The initial stage of wave generation consists on small fluctuations of the atmospheric pressure producing capillary waves with frequencies around 10^2 cycles/second. These waves are usually attenuated due to surface tension when the generating force disappears. As the wind velocity increases, the energy transfer between the atmospheric boundary layer and the sea surface also increases, then waves grow and gravity force is sufficient to support wave motion. The energy increment is not infinite because of the energy dissipation processes such wave breaking and whitecapping. Moreover, when waves reach their limiting steepness they break as whitecaps, spilling or plunging breakers (Massel, 1989; Banner and Peregrine, 1993).

Knowledge on the mechanisms of generation, interaction and dissipation of ocean waves has been accumulated during the last years. Aristotle, 2000 years ago, realized that the wind acting on the sea surface plays a very important role in the development

of waves, but from the time of Aristotle to the early 20th century very little progress was made towards the understanding of the generation and growth of waves in the ocean. It was not until the last 70 years that more fundamental knowledge of what caused waves and how they behaved was developed (Massel, 1996). However, modern understanding of the dynamic air-sea processes involved during the wave generation and dissipation has been developed only within the last 50 years, when a basic understanding of the mechanisms of wave generation was acquired.

Phillips (1957) and Miles (1957) in independent works, established the first solid framework in the understanding of wind-wave generation. Phillips' wind-wave generation model assumes that the atmospheric pressure fluctuations don't interact with generated waves, then the pressure fluctuations and the generated wave will be traveling together at some velocity U generating a resonance phenomenon. However, this model seemed too weak to support the continued growth of wind waves. On the contrary Miles' model growing rate is exponential, because the Miles' mechanism considered the resonant interaction between the wave induced pressure fluctuations and the free surface waves. Experimental results showed that Phillip's theory accounts reasonably well for the initial wave growth but the major portion of the spectral development occurs due to Miles' mechanism, although the theory still predicts energy transfer values that are smaller than measured values. These models made important theoretical simplifications, some of them corrected in the last generation wave models presented in the next section.

1.2.1 Wave generation models

Early forecasts of the sea state were created manually based on empirical relationships between the present state of the sea, the expected wind conditions, the fetch/duration, and the direction of the wave propagation (Wittman and Clancy, 2004). For about 40 years, wind-driven numerical wave prediction models have proven useful for ship routing, offshore structures safety and also more recently, for climate research. Since the pioneering model by Gelci et al. (1957), many wave models have been developed in which the complex nature of wave generation, propagation and decay has been described. For forecasting purposes, it was realized that the random nature of the sea state was best described by a spectral decomposition in which the energy of the waves was attributed to as many wave trains as necessary, each of them with a specific direction and period.

This first numerical model by Gelci et al. (1957) was based on the spectral decomposition of the sea state and it was operated in 1956 by the French Weather Service, and focused on the North Atlantic. This first generation wave models did not consider nonlinear wave interactions. On the contrary, second generation models, available by the early 1980s, parameterized these non-linear interactions leading to the parametric models and the hybrid models (Allender et al., 1985).

Nowadays research in wave generation mechanisms is still very active. In 1988 the *WAMDI* group developed the first third generation wind-wave model including all the improvements proposed by the *SWAMP* group in 1985 (Hasselmann et al., 1988). These third generation wave models explicitly represent all the physics relevant for the development of the sea state in two dimensions. Actually third generation wave models are the state of the art in wave generation models, used in research institutes and meteorological agencies for wave forecasting (Günther et al., 1992, Gómez and Carretero, 1994). Some bias are reported in the outputs of wave models when compared with satellite altimetry, but predictions are good enough for later analysis and scientific studies based on the data from these models. One of the most important features of these models is the capacity to provide long-term datasets without missing data or gaps, and its enhanced spatial coverage that can cover large areas with fine grids. On the other side, the principal shortcoming of these models is the need for highly-intensive computing resources that are not always affordable for all research centers.

1.3 Oceanographic wave data

1.3.1 The HIPOCAS project: a 44-year wave reanalysis

In this Thesis we have used wave data from the HIPOCAS Project (Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe) (Soares, 2008). This database consists on a high resolution, spatial and temporal, long-term hindcasted data set. HIPOCAS reanalysis covers, on an hourly basis, a period ranging from 1958 to 2001 providing 44-years of wave data over an homogeneous grid. HIPOCAS dataset was produced by means of dynamical downscaling from the NCEP/NCAR global re-analysis using the regional atmospheric model REMO (Kalnay et al., 1996, Jacob and Podzun, 1997). Hourly wind fields from the REMO (u_{10}) were used as forcing for the third generation wave model WAM (see Sotillo et al. (2005)). The three different domains covered by the HIPOCAS project are shown in Figure 1.2, where the three models used for the reanalysis, i.e. REMO, HAMSOM and WAM, are displayed with its domain.

As a result, wave data used is the output of the WAM model implemented in a $1/8^\circ$ resolution mesh over the Western Mediterranean Sea. In this Thesis we cover different nodes in an area between 38°N - 42°N and 1°W - 6°E , see Figure 1.3.

In the Western Mediterranean Sea, wave height data from the HIPOCAS database was calibrated using buoy data and achieved through the EOF decomposition of the spatial modes and its temporal amplitudes and then transforming the spatial modes using a non-linear parameterization to fit adequately the satellite and buoy data (Tomás et al., 2008).

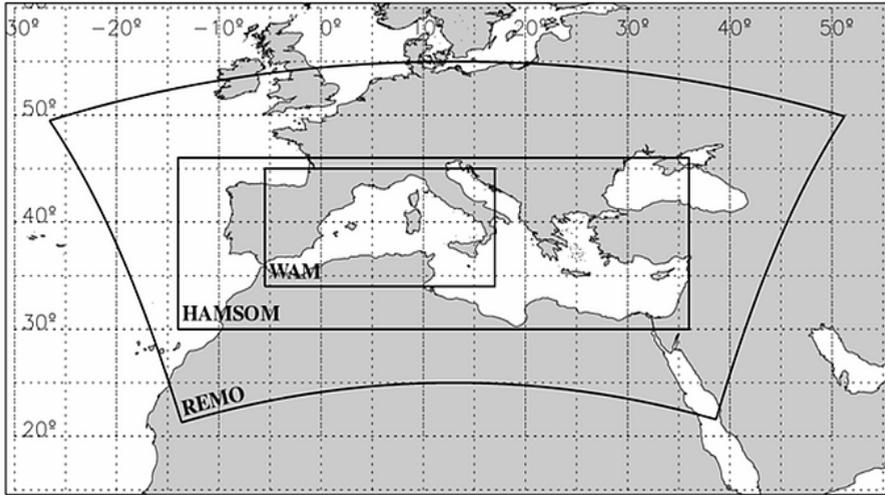


Fig. 1.2: Geographic location of the three different domains covered by the atmospheric and oceanographic models REMO, HAMSOM and WAM.

HIPOCAS allow to perform different analysis due to the long record available in this set. Also, the high resolution (spatial and temporal) of this reanalysis is an important feature to be taken into account. The analysis carried out in this Thesis lean on the high resolution of this database to perform *i*) characterization of return period levels at deep waters as well as their spatial distribution; *ii*) analysis of the temporal and spatial variability associated to those extreme significant wave heights fields with special attention to climatic links and *iii*) develop a real-time prediction tool based on genetic algorithms.

1.3.2 Oceanographic Buoy Data

Deep water buoys providing real time wave and wind parameters have been also used in this Thesis. In order to validate and compare results between HIPOCAS data and real buoy data, we gathered meteorological and oceanographic data from four oceanographic buoys located in the same area as the HIPOCAS nodes in the Balearic Sea (Figure 1.3). This figure shows the 0.125° resolution HIPOCAS grid in the Western Mediterranean area and also the location of the seven oceanographic buoys available in this area.

Two of these buoys, Maó Buoy (MH) (39.73°N - 4.42°E , depth: 300 meters) and Dragonera Buoy (DR) (39.55°N - 2.10°E , depth: 135 meters), are WAVESCAN buoys equipped with a Motion Reference Unit (MRU) for wave direction measurements (Figure 1.4). These kind of sensors offer the advantage of being insensitive to either

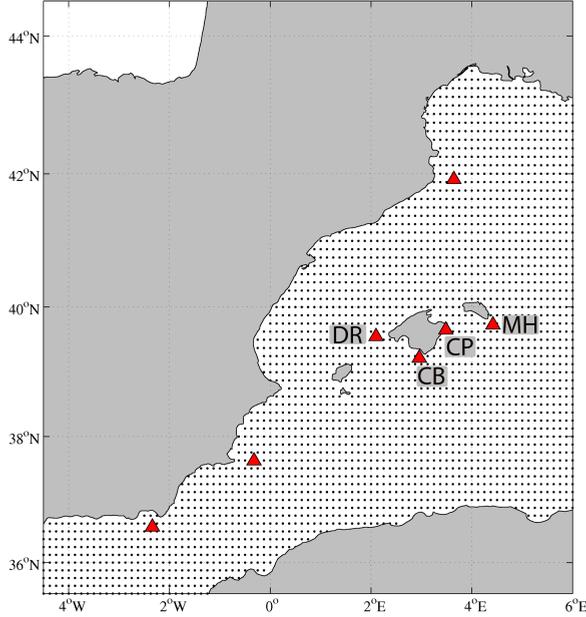


Fig. 1.3: Geographic location of the study area and 0.125° resolution HIPOCAS grid in the Western Mediterranean. Red triangles correspond to the location of the seven oceanographic buoys available in the Western Mediterranean area. The four buoys used in this Thesis marked as: Maó (MH), Dragonera (DR), Capdepera (CP) and Cabrera (CB)

high or low temperatures as well as spinning and rough conditions. The WAVESCAN buoy incorporates a 3-axis compass for buoy orientation measurements and three accelerometers to compute the vertical displacement of the buoy and also the direction of the displacement.

These kind of buoys measure the wave direction (θ), mean wave period (T_m), peak period (T_p) and significant wave height (SWH), among other useful wave parameters. To obtain these parameters the buoy records continuously each hour during 20 minutes at 4 Hz frequency to provide a bidimensional energy spectra.

These two WAVESCAN buoys belong to the REDEXT (Red Exterior) network from Puertos del Estado and to date have produced 3 and 10-year of data, DR and MH buoy respectively. The final data consists on hourly meteorological and oceanographic parameters such as wind module and direction, SWH , wave period and wave direction.

Cabrera Buoy (CB) is a meteorological and oceanographic buoy located between Mallorca and the Cabrera National Park in the Cabrera Island (39.22°N - 2.96°E , depth: 70 meters)(Figure 1.5). The management and maintenance of this buoy is shared on the basis of Collaboration Agreement between CSIC and National Parks



Fig. 1.4: *A moored WAVESCAN buoy.*

Foundation. The buoy installed in 2008 provided a year of meteorological and oceanographic data.

The remaining buoy, Capdepera buoy (CP) ($39^{\circ}39.0\text{ N}$ - $3^{\circ}29.1\text{ E}$, depth: 45 meters) from Puertos del Estado is located near the north-eastern coast of the Mallorca island and was moored in 1989. The major shortcoming of the CP buoy is that it is not equipped with a compass to provide wave directions. Another shortcoming of this buoy is that it provides information about SWH and swell period in shallow waters, where the swell is obviously affected by shoaling, refraction, and diffraction processes.

1.3.3 Comparison and validation between the HIPOCAS database and Buoy data.

As stated in the previous section the HIPOCAS database was calibrated using available buoy data in the Western Mediterranean area by Tomás et al. (2008). The calibration method was applied to monthly long-term distributions of SWH in the Western Mediterranean. After the calibration process, the author has validated the results with satellite data showing how this methodology is useful to improve the results obtained directly from numerical models.

A direct comparison between the SWH calibrated HIPOCAS time series and SWH from buoy data is shown in Figure 1.6. The HIPOCAS node chosen for comparison



Fig. 1.5: *Aanderaa DB 4700 (Cabrera) oceanographic buoy equipped with meteorological and oceanographic sensors.*

is the WAM node 1320, located near the coast of Maó. The instrumental observations were obtained from the Maó Buoy deployed by Puertos del Estado, near this HIPOCAS node. A point by point comparison between the wind intensity obtained from the buoy and the wind intensity in the HIPOCAS database is shown in the next Figure 1.7. Both figures show a good agreement between the compared datasets. The temporal resolution found in the HIPOCAS series is up to 1 hour while in the Maó Buoy the temporal resolution is 3 hours between data. An additional advantage of the HIPOCAS database is the absence of gaps, for instance in the case of the MH buoy we found a 9% of missing data during the whole record. This fact can be observed in the wind intensity comparison shown in Figure 1.7, where the wind intensity record presents a large gap during the last days.

On the other hand a validation and comparison between HIPOCAS SWH and buoy SWH data is also presented here from the probabilistic point of view Figure 1.8. Both distributions show a similar behavior for the whole range of SWH levels in the distribution, although a slight difference is shown for the SWH below the mean value of 1.2 meters. Anyway the HIPOCAS node reproduces the observations obtained at the oceanographic buoy well.

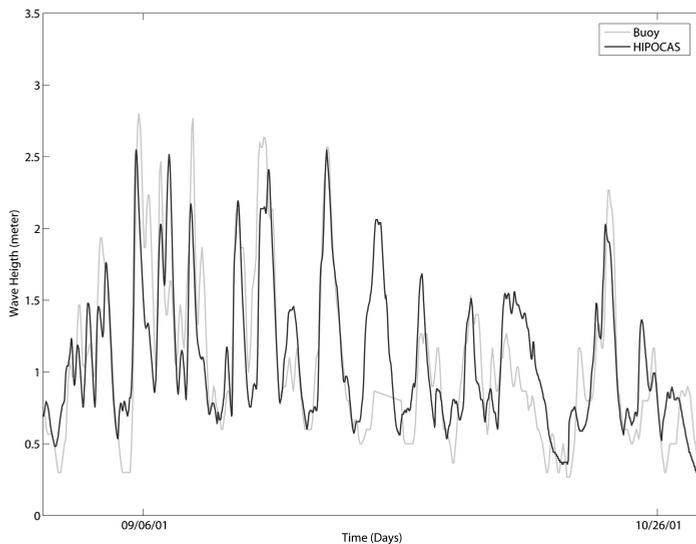


Fig. 1.6: Point by point comparison between the SWH recorded by the MH buoy and the SWH time series provided by the closest node to the buoy, the HIPOCAS node 1320.

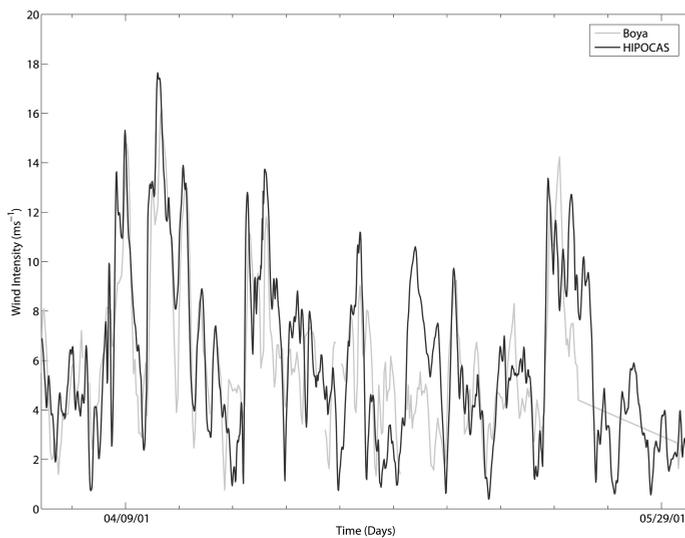


Fig. 1.7: Point by point comparison between the wind module (U) recorded by the MH buoy and the wind module (U) time series provided by the closest node to the buoy, the HIPOCAS node 1320.

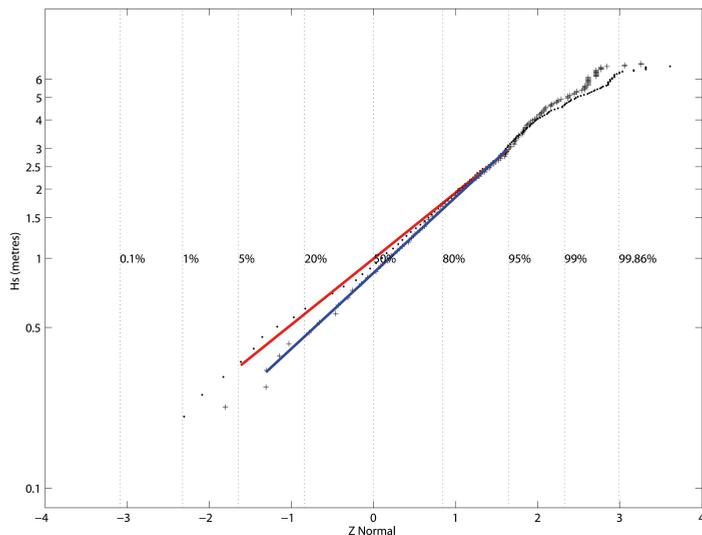


Fig. 1.8: *Logn-Normal probability plots for the SWH distribution. HIPOCAS database is shown as a red line and the buoy data is shown in blue.*

Thus, the use of the HIPOCAS database is an excellent alternative to study extreme wave height probability distributions as well as their spatial variability in the Western Mediterranean area. Also, the large amount of data contained in these series will provide the opportunity to improve our knowledge about the interactions between meteorological and oceanographic parameters, including large scale interactions such as climatic teleconnections with distant geophysical phenomena.

Chapter 2

Time series analysis applied to wind waves

When someone looks at the sea surface at first sight notice their randomness and quasi-stationarity. This is one of the most important properties of wind induced surface waves, their irregularity. This characteristic is shown in a real sea surface register recorded during a storm in the Western Mediterranean (Figure 2.1). The analysis and prediction of wave parameters can be achieved only through stochastic analysis of the sea surface at concrete points. The basic domains of the stochastic analysis span these three domains: time, frequency and probability (Massel, 1996). These domains will be further discussed in the next subsections.

2.1 Time domain

Principal tools for the analysis in the time domain consists on the evaluation of the auto-correlation or cross-correlation functions from the wave records. The autocorrelation function is a measure of the correlation-connection between two values $\eta(t)$ and $\eta(t+\tau)$, being η a random variable. If we consider an ensemble of k records $\zeta_k(t)$ taken under identical conditions we cannot expect that these records will be identical or even similar. The family $\zeta_k(t)$ represents k realizations of the stochastic process $\zeta(t)$. Stochastic processes are grouped in one of these three categories: stationary and ergodic, stationary and non-stationary.

We define a process as stationary in the wide sense if,

$$E[\zeta(t)] = \bar{\zeta} = const, \quad (2.1)$$

and

$$K(t_1, t_2) = K(t_1 - t_2) = E[\zeta(t_1)\zeta(t_2)] = K(\tau), \quad \tau = t_1 - t_2. \quad (2.2)$$

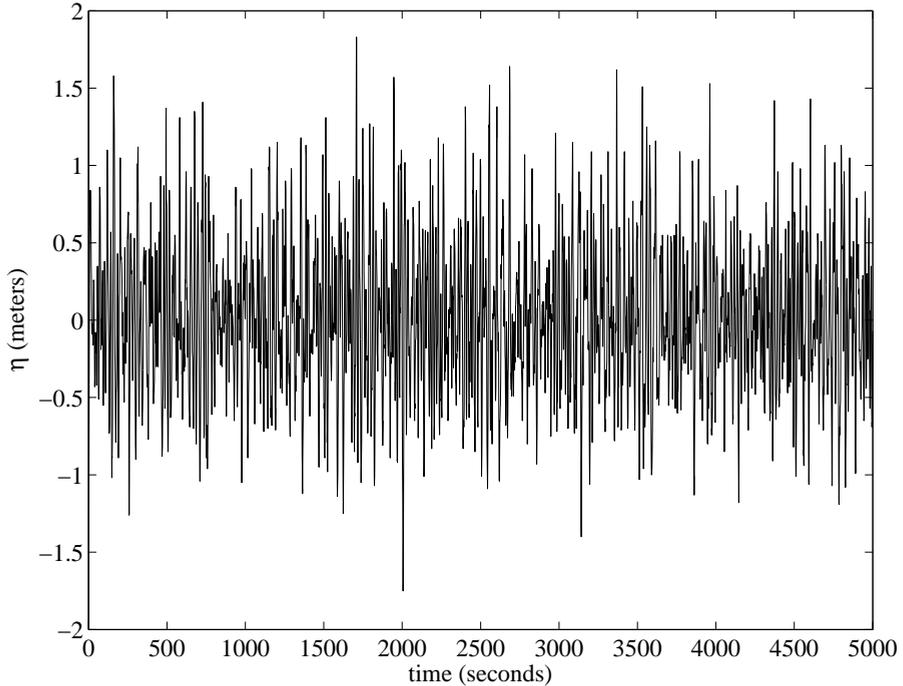


Fig. 2.1: *Example of sea surface elevation from mean water level.*

being E and K the statistical expectation and the autocorrelation function, respectively.

A random process is stationary if all statistical moments are invariant at all times $t = (t_1, t_2, \dots, t_n)$. The use of an ensemble of $\zeta_k(t)$ records, allows the approximation of the values for the ensemble mean and variance over k . If $t = t_1$, we obtain the ensemble mean as,

$$E[\zeta_k(t_1)]_k = \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N \zeta_k(t_1)}{N}, \quad (2.3)$$

and the variance as,

$$E\{[\zeta_k(t_1)]^2\}_k = \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N [\zeta_k(t_1)]^2}{N}. \quad (2.4)$$

One of the common problems in Equations 2.3 and 2.4 is that they are based on an ensemble of records theoretically done under the same macroscopic conditions. This replicability condition is impossible to achieve in field experiments. To overcome these difficulties, Kinsman (1965), proposed the ergodic theorem. This theorem allows the ensemble averages to be replaced with time averages using a sufficiently long

realization of the process. The ergodic stationary process should satisfy the following equality,

$$E[F\{\zeta_k(t = t_*)\}]_k = E[F\{\zeta_{k=k^*}(t)\}]_t. \quad (2.5)$$

This theorem allows the ensemble averages to be replaced with time averages, which is very useful when dealing with real sea surface wave records. In real conditions, a wave record at a single point is never stationary or ergodic, but during short time periods (i.e: 1 hour) and in reduced areas, this process could be seen as a temporally stationary and spatially ergodic process.

A long record from a single point can be divided in many equal subrecords based on a fixed subrecord time length to assume stationarity and ergodicity. Every subrecord is now considered as a “sea state”. We have now defined an important concept in ocean surface wave analysis: the sea state, defined as the temporal period when the statistical moments (e.g. standard deviation, variance, mean, etc) remains constant. The sea state allows to substitute a long wave data record with a discrete and shorter statistical information for each sea state. This statistical concept is a powerful tool due to its encoding capability allowing to define a wave record with few statistical parameters. This process is known as short-term analysis.

The analysis of the evolution in time of the different sea states is known as long-term analysis and the analysis of only the highest sea states in a record is known as extreme long-term analysis. Both kind of analysis will be seen in detail in next chapters.

2.1.1 Frequency domain analysis

The simplest representation of the random behavior of the sea surface would be the linear superposition of infinite number of wave trains traveling in different directions. The wave profile, in its discrete form, can be expressed as,

$$\zeta(x, y, t) = \sum_{l=1}^N a_l \cos[k_l(x_l \cos \alpha_l + y_l \cos \alpha_l) - \omega_l t + \varphi_l]. \quad (2.6)$$

where for each l component a is the wave amplitude, $k = \frac{2\pi}{L}$ the wave number with L the wave length, α the incoming angle, $\omega = 2\pi/T$ the angular frequency with T the wave period and φ the phase.

The assumption of the superposition of infinite harmonics has a direct applicability in wave analysis, through the Fourier integral analysis or Fourier transform. The Fourier transform decomposes a given function into the basic pieces, its harmonics, and it is closely related to the autocorrelation function $K(\tau)$ through the Wiener-Khinchine theorem. This theorem states that a function in time can be represented in the frequency domain, through the Fourier transform of the autocorrelation function. As an example Figure 2.2 shows the autocorrelation function for the previous

Sea Surface Elevation data record, showing a rapid decay of the correlation at lag $\tau = 3$.

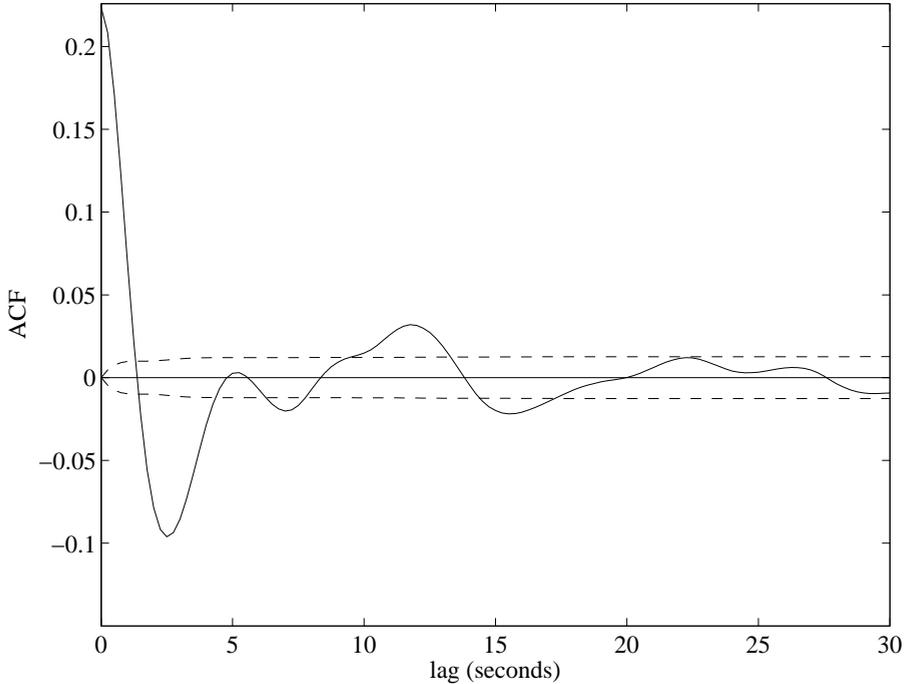


Fig. 2.2: Example of the autocorrelation function for the sea surface elevation

More precisely, through the Fourier transformation we obtain the power density spectrum of the analyzed function, or the energy contained in each of the frequency spectrum bands. Figure 2.3 shows the computed density spectrum for the previous autocorrelation figure, where the peak frequency is located at 1.1rad/s being the peak period $T_p = 5.7$ seconds. The filtered spectrum obtained through the application of a parzen window is shown in red.

Once the power density spectrum is obtained, much more information can be extracted from the signal through the spectral moments. The r^{th} spectral moment is defined as,

$$r^{\text{th}} = \int_0^{\infty} \omega^r S(\omega) d\omega. \quad (2.7)$$

The $zero^{\text{th}}$ spectral moment is the total variance of the process, σ_{ζ}^2 . If the process represented is the sea surface displacement, we can introduce one of the most important parameters in wave climate analysis, the spectral significant wave height, SWH

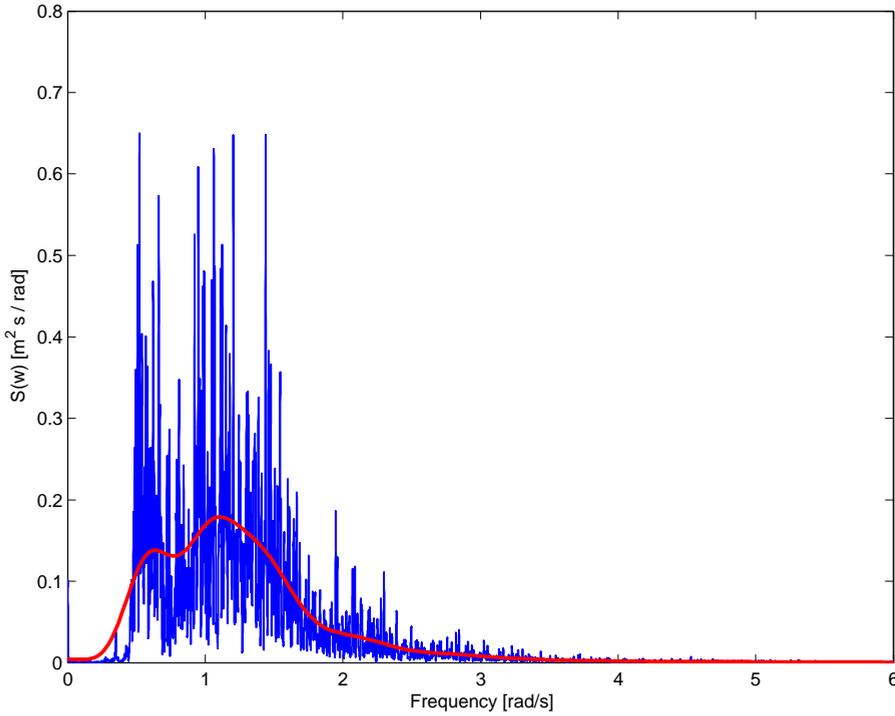


Fig. 2.3: Example of the power density spectra of a wave record, the peak frequency is $f_p = 1.1 \text{ rad/s}$

or H_{m0} . It is defined as four times the root square of the variance (the *zeroth* spectral moment),

$$H_{m0} = 4\sqrt{m_0}. \quad (2.8)$$

Other useful spectral moments are the mean angular frequency $\bar{\omega}$, the mean period $\bar{\tau}$ and the mean zero crossing period $\bar{\tau}_z$,

$$\bar{\omega} = \frac{m_1}{m_0}; \quad (2.9)$$

$$\bar{\tau} = 2\pi \frac{m_0}{m_1}; \quad (2.10)$$

$$\bar{\tau}_z = 2\pi \sqrt{\frac{m_0}{m_2}}. \quad (2.11)$$

The information contained in the power density spectrum can be expressed with these spectral parameters, showing the principal characteristics of every sea state

defined by H_{m0} , $\bar{\omega}$ and $\bar{\tau}$. These parameters can be also obtained without any spectral analysis, only through the statistical approach. This will be done in next section.

2.1.2 Probabilistic approach

As seen in the previous section the spectral wave parameters allow the characterization of a sea state analyzing the energetic content of the energy density spectrum. For the short-term analysis of a wave record we also can introduce some statistical parameters based on a probabilistic approach of the problem. Statistical parameters are almost equivalent to the spectral ones, and due to the stationarity, gaussianity and ergodicity assumption a process can be completely determined by the first and second statistical moments (Massel, 1996).

The most important statistical parameter in the short-term analysis is the significant wave height H_s , introduced by Sverdrup and Munk (1947). This statistical parameter relates the visual wave height with the amplitude obtained with instrumental data records. Statistically, H_s is calculated as the mean of the one-third highest waves in the record,

$$H_s = \frac{3}{N} \sum_{i=1}^{N/3} H_i. \quad (2.12)$$

This parameter is closely related to the spectral parameter H_{m0} , since they are approximately equal except when water depth is small ($kh < \pi/10$, being k the wavenumber and h depth) or waves are very steep. For shallow waters or very steep waves, a deviation from gaussianity is more evident than in deep waters. This is due to the positive values of the higher order statistical parameters, kurtosis and skewness, indicating nonlinearities in the wave field (Longuet-Higgins, 1963). In these cases, the spectral parameter H_{m0} gives a better estimation.

Other statistical parameters are obtained through the mean values of height \bar{H} and period $\bar{\tau}$. These parameters are defined in the field of the short-term analysis.

With these two different approaches (spectral and probabilistic), we can generate another time series based on the time evolution of the spectral or statistical parameters obtained through the preliminary short-term analysis. Figure 2.4 displays a 20-year time series inferred from the energy spectrum obtained from the WAM model.

2.1.3 Long-term analysis

Long-term analysis of wave data is the statistical analysis of a time series containing spectral/statistical wave parameters obtained previously from the raw data. This analysis allows to introduce a new concept: wave climate. Wave climate is, the long-term statistical characterization of the behavior of different sea state parameters in a given area. For H_{m0} , the long-term distribution contains information about the

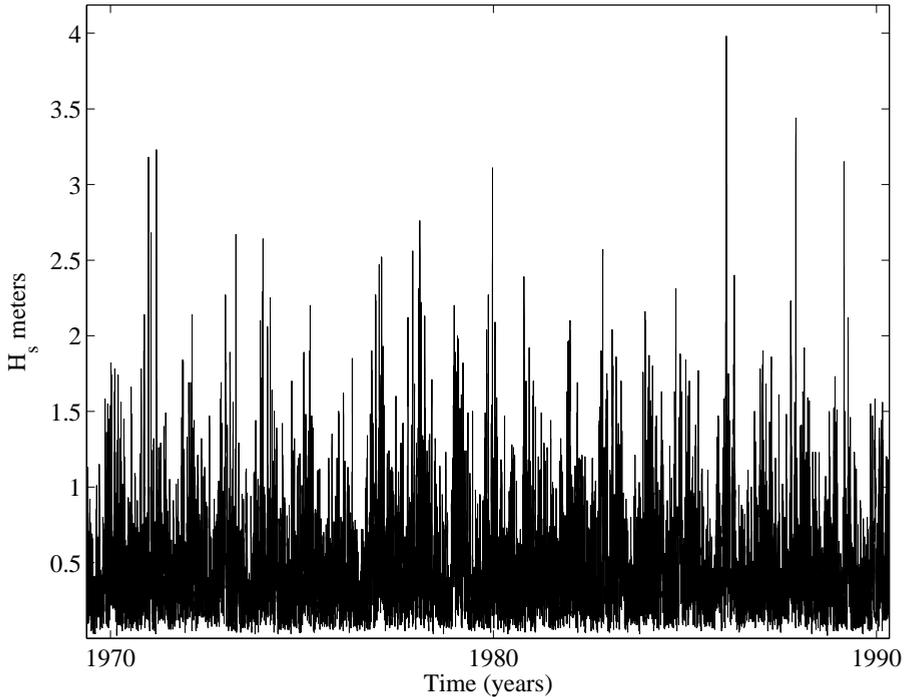


Fig. 2.4: 20-year significant wave height time series, from the HIPOCAS database.

probability of non-exceedance of a given significant wave height level during a mean time period (monthly, seasonal or annual).

Long-term analysis is based on the statistical analysis of a parameter (e.g. H_{m0}), through its probability distribution function,

$$F(x) = \Pr\{X \leq x\}. \quad (2.13)$$

If the distribution function F is differentiable, we also can define the probability density function as,

$$f(x) = \frac{dF}{dx}, \quad (2.14)$$

and,

$$F(x) = \Pr\{X \leq x\} = \int_{-\infty}^x f(x)dx. \quad (2.15)$$

There are some probability distributions actually used in wave climate to describe the long-term distribution of a sea state parameter. The most widely used are the log-normal distribution and the biparametric Weibull distribution. In Spain, the recommendation for maritime works published in the ROM (Maritime Works Recommendations; EPPE, 1991) makes use of the log-normal distribution function for

the significant wave height characterization in harbors and deep water buoys located in the coastal areas. Other authors, also make use of the log-normal distribution to analyze the monthly mean and standard deviation for the significant wave height in low-energetic sandy beaches of Mallorca (Gómez-Pujol et al., 2007). There are, however, other well-known long-term probability distributions as shown in Table 2.1.

Distribution Function	Mathematical Expression
Fisher-Tippett I	$F(x) = e^{-e^{-\left(\frac{x-B}{A}\right)}}$
Weibull	$F(x) = 1 - e^{-\left(\frac{x-B}{A}\right)^k}$
Fisher-Tippett II	$F(x) = e^{-\left(\frac{x}{A}\right)^{-k}}$
Log-Normal	$F(x) = \frac{1}{Ax\sqrt{\pi}} e^{-\left(\frac{\ln x - B}{A}\right)^2}$
Log Pearson	$F(x) = \frac{1}{Ax\Gamma(k)} \left(\frac{\ln x - B}{A}\right)^{(k-1)} e^{-\left(\frac{\ln x - B}{A}\right)}$
Pearson	$F(x) = \frac{1}{A\Gamma(k)} \left(\frac{x-B}{A}\right)^{(k-1)} e^{-\left(\frac{x-B}{A}\right)}$
Binomial	$F(x) = \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x}$
Poisson	$F(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

Table 2.1: Probability distribution functions for long-term statistics extracted from the Coastal Engineering Manual (Thompson, 2002)

2.2 Extreme value distributions

Extreme wave heights are rare by definition. One of the shortcomings of the long-term probability distributions is their poor accuracy in the right hand tail of the distribution, corresponding to the largest SWH levels and the lower probabilities of exceedance. Thus, long-term extreme statistical methods must deal with the problem of using a variable sample of records to estimate extreme wave height values. Another important problem that one has to deal with when analyzing extreme events is the lack of data of these rare extreme events. The estimation of unusual events is marked by the complexity of gather sufficient data in the extremal right hand tail of the distribution. In next sections will be shown the most common methods to analyze time series in the search for extremal populations in SWH records.

The basis of the extreme value prediction methods is the use of the asymptotic behavior of the probability distribution function for the maxima/minima of a random variable. Fisher and Tippett (1928) and Gnedenko (1943), proposed and developed respectively, the three types theorem for the study of extremal population. This theorem states that any non-degenerate distribution (a distribution function which is not either 0 or 1) belongs to one of the three limiting functions or families, Table 2.2,

The three families of distributions are known as the extreme value distributions;

Family	Distribution Function
I	$G(z) = \exp\{-\exp[-(\frac{z-b}{a})]\}$ $-\infty < z < \infty$,
II	$G(z) = \exp\{-(\frac{z-b}{a})^{-\alpha}\}$ $z > b$,
III	$G(z) = \exp\{-[-(\frac{z-b}{a})^\alpha]\}$ $z < b$.

Table 2.2: *The three types theorem limiting functions*

where Family I correspond to the Gumbel family; Family II to the Fréchet family and Family III to the Weibull family. Each distribution presents a scale parameter a and location parameter b . Additionally, the Fréchet and Weibull families present a shape parameter α . The three types theorem is essentially an extreme value analog of the central limit theorem (Coles, 2001).

2.2.1 The Generalized Extreme value Distribution (GEV)

The three families I, II, III, can be combined into a single family of models, known as the Generalized Extreme Value Distribution or GEV. The classical approach to perform an extreme value analysis is to fit the annual maxima series with the Generalized Extreme Value (GEV) cumulative distribution function,

$$G(z) = \exp\{-[1 + \xi(\frac{z - \mu}{\sigma})]^{-1/\xi}\}, \quad (2.16)$$

where μ is the location parameter, $\sigma > 0$ is the so-called scale parameter and ξ is a shape parameter which determines the tail of the distribution. When $\xi \rightarrow 0$ the GEV distribution corresponds to the Gumbel family, conversely for $\xi > 0$ the Fréchet form is adopted and for $\xi < 0$ the Weibull form is adopted. This unification into a single family removes the necessity to make a subjective aprioristic judgement about which of the families has to be adopted, because data by themselves determine the most appropriate type of tail behavior.

The annual maxima method developed by Gumbel (1958) considers only the largest extreme value for each year. The principal shortcoming of this method is that using only the maximum value of each year leads to the loss of information contained in other large sample data for a given period (Castillo and Hadi, 1997). To solve the problem of working only with a data per year the Generalized Pareto Distribution (GPD) was introduced by Pickands (1975). The GPD method, models all values larger than a given threshold u . The differences between these values and the threshold value are called exceedances over the threshold, and it is assumed to follow a GPD (σ, ξ) distribution whose cumulative distribution function is defined by,

$$G(y; \sigma, \xi) = 1 - (1 + \xi \frac{y}{\sigma})^{-1/\xi}, \quad (2.17)$$

where y are the exceedances over the threshold u , ($y = x - u > 0$). This method mitigates the annual maxima method problem because it considers all values larger than a given threshold u , allowing to obtain larger extremal populations from the same data record. As an example of the GPD method, Figure 2.5 shows a 2-year time series for the SWH . The annual maxima method only takes 2 data points from this series marked with black squares. On the contrary, with the Peaks Over Threshold method more than 20 data could be extracted from this short record, allowing better estimations from the distribution function. However the threshold selection method is not trivial and in next sections we will focus on it.

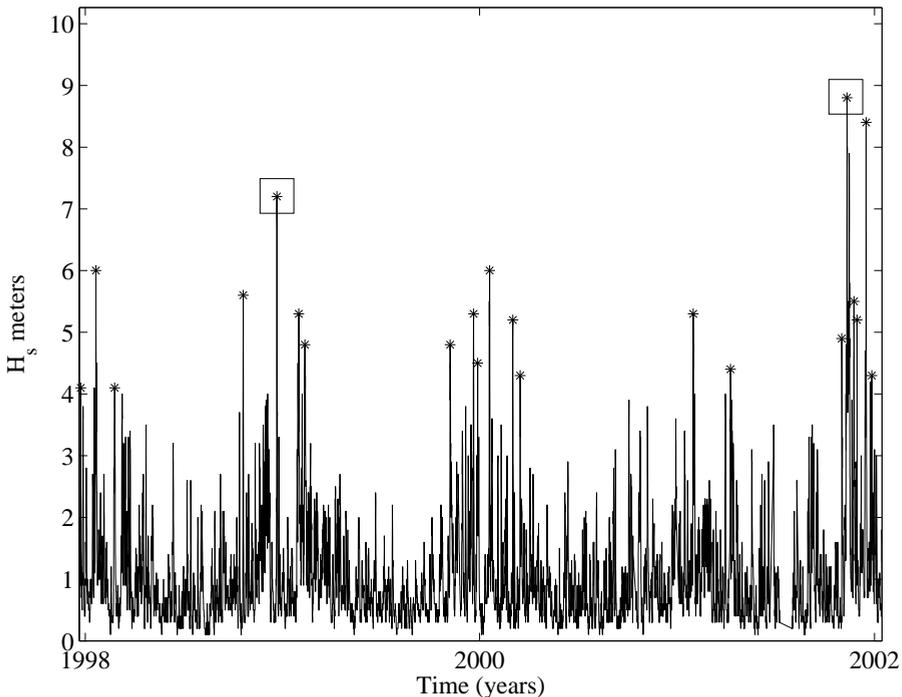


Fig. 2.5: Example of number of selected extremal SWH with the annual maxima method (squares) vs. peaks over threshold method (asterisks)

2.2.2 The Peaks Over Threshold method (POT)

A modification to the model defined in Equation 2.17 is the Poisson-GPD model for exceedances, originally developed by hydrologists which is closely related to the Peaks Over Threshold (POT) method. This model is a joint distribution between the GPD for the exceedances values y , and a Poisson distribution for the number of exceedances over a level threshold u in any given year. With this model, one can estimate not

only the intensity of the exceedances but also the frequency of these exceedances over the selected threshold.

The POT method assumes that the number N of exceedances of the level u in any given year follows a Poisson distribution with mean λ . Moreover, the exceedances $\{y_i\}_{i=1}^N$ are independent and identically distributed from the GPD (Smith, 2003). Under these hypothesis, the probability that the annual maximum of the GPD-P process is lower than a value x , with $x > u$, is given by,

$$F(G(y)) = e^{-\lambda\tau(1-G(y))} = \exp -\lambda\left(1 + \xi\frac{y}{\sigma}\right)^{-1/\xi}, \quad (2.18)$$

where λ is the Poisson parameter, σ the scale parameter and ξ the shape parameter are to be determined.

2.2.3 Parameter estimation

The GPD-P model reduces to the determination of the three unknown parameters, $\lambda > 0$, $\sigma > 0$ and $-\infty < \xi < \infty$. The scale and shape parameters arise from the GPD and λ from the Poisson distribution. These three parameters are estimated using the Maximum Likelihood Method (MLM).

The Maximum Likelihood Estimators are the values of the unknown parameters that maximise the log-likelihood function. In practise these are local maxima found by nonlinear optimization. In this Thesis we used the MLM because of its efficiency in large samples sizes (e.g. $N > 500$). As shown by Hosking and Wallis (1987) the values of the shape parameter is ususally restricted to $-0.5 < \xi < 0.5$, because this is the range of values commonly observed in practical applications. Moreover, these values avoid convergence problems when computing an algorithm for the MLE.

The log-likelihood function for the GPD-P, if N exceedances are observed over a τ -year period is given by,

$$\ell(y; \lambda, \sigma, \xi) = N \log \lambda - \lambda\tau - N \log \sigma - \left(\frac{1 + \xi}{\xi}\right) \sum_{i=1}^N \log \left(1 + \xi\frac{y_i}{\sigma}\right) \quad (2.19)$$

Maximizing the log-likelihood function respect to $\theta = (\lambda, \sigma, \xi)$ in the GPD-P or $\theta = (\mu, \sigma, \xi)$ in the GEV case, leads to the maximum likelihood estimates. To maximize the log-likelihood function an efficient global optimization procedure can be used, namely the SCE-UA algorithm (Duan et al., 1992). This optimisation algorithm have been already used successfully in the calibration of highly non-linear problems.

One advantage of the maximum ML method is that approximate standard errors for estimated parameters and design values can be automatically produced via the observed information matrix (Katz et al., 2002). The elements of the Hessian matrix of ℓ , evaluated at the maximum likelihood estimators is known as the observed

information matrix, defined as,

$$I_o = -\frac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j} \quad (2.20)$$

The covariance matrix of the maximum likelihood estimators can be constructed obtaining the inverse of the observed information matrix $I_o(\theta)$. Then the standard errors $se(\hat{\theta}_i)$ of the estimators are the square root of the diagonal entries.

An approximate $(1 - \alpha)$ confidence interval for θ_i would be,

$$\hat{\theta}_i = \pm z_{\alpha} se(\hat{\theta}_i), \quad (2.21)$$

where for $z_{0.95} = 1.96$ gives a 95% confidence interval for the ML estimators. Then the confidence bounds for the design levels, i.e. the extreme wave heights associated to a probability level, can be obtained through this method.

2.3 Empirical orthogonal function analysis

Traditionally, Empirical Orthogonal Functions (EOFs) also known as Principal Component Analysis (PCA) are used to identify the preferred modes of variability of a dynamical system (Preisendorfer, 1988). EOFs are another important tool for SWH analysis because they allow to search structures that explain the maximum amount of variability in a two dimensional data set. One dimension in the data set represents the dimension in which we are seeking to find a structure, and the other dimension represents the dimension in which realizations of this structure are sampled. When searching characteristic spatial structures that vary with time, for example, we would use space as the structure dimension and time as the sampling dimension, this analysis produces a set of structures in the first dimension, which we call the EOFs, and which we can think of as being the structures in the spatial dimension (Hartmann, 2008).

EOFs are the eigenfunctions of a covariance data matrix which, when ordered by eigenvalue, represent the dominant patterns of the variance. Usually, EOFs are calculated from data in which the time average mean at each location is removed. This leads to an EOF decomposition of the temporal variance of the data, named covariance EOF analysis. Similarly, an EOF decomposition of the spatial variance (gradient EOF analysis) is obtained if the spatial average is removed from each map. EOF decomposition of the temporal variance describes those spatial patterns with the highest contribution to time variability. On the other hand, if the EOF decomposition of the spatial variance is carried out spatial patterns with the strongest spatial gradients will be obtained (Álvarez, 2003).

Given a matrix of data D with dimensions $M \times N$, where M is the number of spatially distributed points and N is the number of points over time, its singular value decompositions are given by,

$$D = \Gamma + U_1 \Sigma_1 V_1^T, \quad (2.22)$$

$$D = \Lambda + U_2 \Sigma_2 V_2^T, \quad (2.23)$$

where U are matrices of left singular vectors, Σ matrices of singular values, V are matrices of right singular vectors, Γ is a matrix with identical columns containing the temporal mean, and Λ is a matrix with identical rows, containing the spatial mean at each sampling time. Superscript T indicates a matrix transpose. In both cases, the column vectors of the U matrices are spatial EOFs and the products ΣV^T are the principal components. Similarly, the row vectors in the V^T matrices are temporal EOFs and the products $U \Sigma$ are matrices of EOF coefficients. Thus, the definition of what is EOF and principal component is somewhat arbitrary. Other EOF decompositions, where temporal and spatial means are removed or no mean is removed, are possible.

In this Thesis we will make use only of temporal and spatial variance EOF decompositions. If an EOF decomposition of the temporal variance is done, the time mean is subtracted from the time series and the resultant mode eigenvalues sum the temporal variance. Thus, top-ranked modes will be associated with features with strong temporal variability. Conversely, if all the SWH maps have the spatial average removed, the modes decompose the spatial variance showing the spatial patterns with the strongest horizontal gradients. In both cases, the spacetime variability contained in the time series of SWH data is reduced to combinations of a small number of EOFs and their respective amplitude functions (Álvarez, 2003).

Chapter 3

Study area description

The Mediterranean Sea, Figure 3.1, a semi-enclosed sea, is located on the western side of the European continental area. The Mediterranean consists of two sub-basins, the Western and the Eastern Mediterranean. The WM is connected to the EM by the Sicily Strait, and to the Atlantic Ocean by the Gibraltar Strait, the thermohaline circulation in this basin is strongly influenced by these connections (Astraldi et al., 1999). Its area, excluding the Black Sea, is about 2.5 million km^2 ; its extent is about 3700 km in longitude, 1600 km in latitude and surrounded by 21 African, Asian and European countries.

The average depth is 1500 meters with a maximum depth of 5150 meters in the Ionian Sea. The Mediterranean Sea is an almost completely closed basin, being connected to the Atlantic Ocean through the narrow Gibraltar strait (14.5 km wide, less than 300m deep at the seal). These morphologic characteristics are rather unique. In fact, most of the other marginal basins have much smaller extent and depth or they are connected through much wider openings to the open ocean. Examples of the first type are the Black and Baltic seas, of the second the Gulf of Mexico and the Arabian Sea, among others.

Moreover, high mountain ridges surrounds the Mediterranean Sea on almost every side (Lionello et al., 2004). These characteristics have important consequences on air masses and atmospheric circulation at the regional scale: the Mediterranean sea is an important heat reservoir representing an important source of energy and moisture for cyclone development and its complex land topography plays a crucial role in steering air flow, so that energetic meso-scale features are present in the atmospheric circulation. The Mediterranean climate region evolves on the north to the Marine West Coast Climate and on the south to the Subtropical Desert Climate. The southern part of the region is mostly under the influence of the descending branch of the Hadley cell, while the Northern part is more linked to the mid-latitude variability, characterized by the NAO and other mid latitude teleconnections patterns (Lionello et al., 2004).



Fig. 3.1: View of the Western Mediterranean Sea included in the HIPOCAS reanalysis database, different subbasins are shown: the Alboran Sea, the Balearic Basin and the Gulf of Lions, among others.

However, the climate variability patterns (teleconnections) present a large amount of synoptic to meso-scale spatial variability, inter-seasonal and multi-decadal to centennial time variability. An important consequence is that the analysis of the Mediterranean Climate can be used to identify changes in the intensity and extension of global scale climate pattern like NAO, ENSO and the monsoons and their region of influence.

3.1 The NorthWestern Mediterranean wave climate: preliminary characterization

The so-called Northwest Mediterranean is the area that stretches from Italy to the Strait of Gibraltar between latitudes 35°N - 44°N and longitudes between 5°W - 6°E approximately (Figure 3.1). The north-western part and central part of the Balearic Sea are forced by northerly winds (Tramontana and Mistral) during the main part of the year, while the eastern part is generally modulated by a seasonal variability.

Gale forced mistrals often develop over the Gulf of Genova and the Gulf of Lions extending the effects over the whole basin. The role of the Pyrenees in the west part and the Alps in the north-east area are decisive boundaries for the wind and pressure distribution over the whole Western Mediterranean basin (Orfila et al., 2005).

The local winds in the Mediterranean region are affected by the mountain borders that can channel the flow toward the sea, increasing their speeds. These high winds are a major cause of storms over the Mediterranean between Italy and the Balearic Islands (Flamant et al., 2003).

3.1.1 Mean Significant Wave Height distribution

In order to present a general idea of the wave climate in the Western Mediterranean area, here a preliminary description of the wave climate in this area is provided. One of the best indicators to study the wave climate in an area is through the mean significant wave height levels, or the spatial distribution of swell energy. The spatial distribution of this mean wave height levels around the Balearic Islands is shown in Figure 3.2, obtained from the 44-year HIPOCAS database.

As seen, the mean SWH levels ranges between 0.5-0.6 meters in the coastal areas of Spain, to the 1.2 meters displayed by the contour line just below the area between the Gulf of Lions and the french city of Marseille. This spatial east-west gradient shows how the SWH levels are not equally distributed along the Western Mediterranean basin due to the variability of the wind forcing during the year period and the shoaling effect due to the change in water depth near the coast.

Spatial wave height gradients are originated in areas where the wind force blows with different intensities and/or direction over large areas. In the Western Mediterranean area this east-west spatial gradient in SWH is produced during the winter storms where winds coming from the north are blowing from the Gulf of Lions and the Gulf of Genoa creating a large east-west direction fetch. This fetch is characterized by exposed areas on the east boundary and a sheltered area in the west boundary and some coastal areas facing the west direction or sufficiently far from the fetch area.

This fact can be better explained observing the evolution in time of these spatial patterns obtained from the 44 year aggregation in twelve months of the SWH monthly mean, Figure 3.3. This figure shows how the intensity of the spatial gradient changes over the year but the spatial distribution of maximum and minimum SWH is constant all over this time period. As seen, SWH values vary throughout the year presenting high values in the autumn-winter months with values above 1 meter except for the coastal areas. During these months the east-west spatial gradient is more visible because the local winds, like the Tramontana wind, blows during these months with

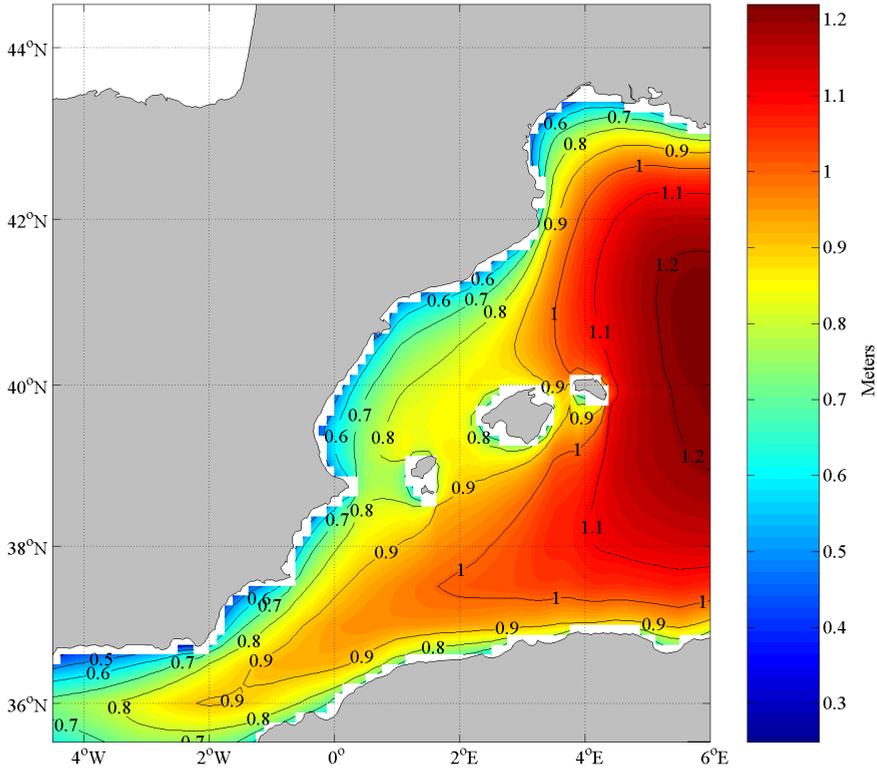


Fig. 3.2: Mean significant wave height obtained from the 44-year HIPOCAS database in the Western Mediterranean area.

more intensity revealing an area below the Gulf of Lions visibly affected by these winds. In this particular area the mean SWH values are above 1.8 meters in December. On the contrary, during the spring-summer months the SWH intensity drops below 1 meter in the whole area especially between the Balearic Islands and the Catalan coast reaching mean SWH values of only 0.3 meters. However, the spatial distribution of SWH remains constant despite of the SWH intensity drop.

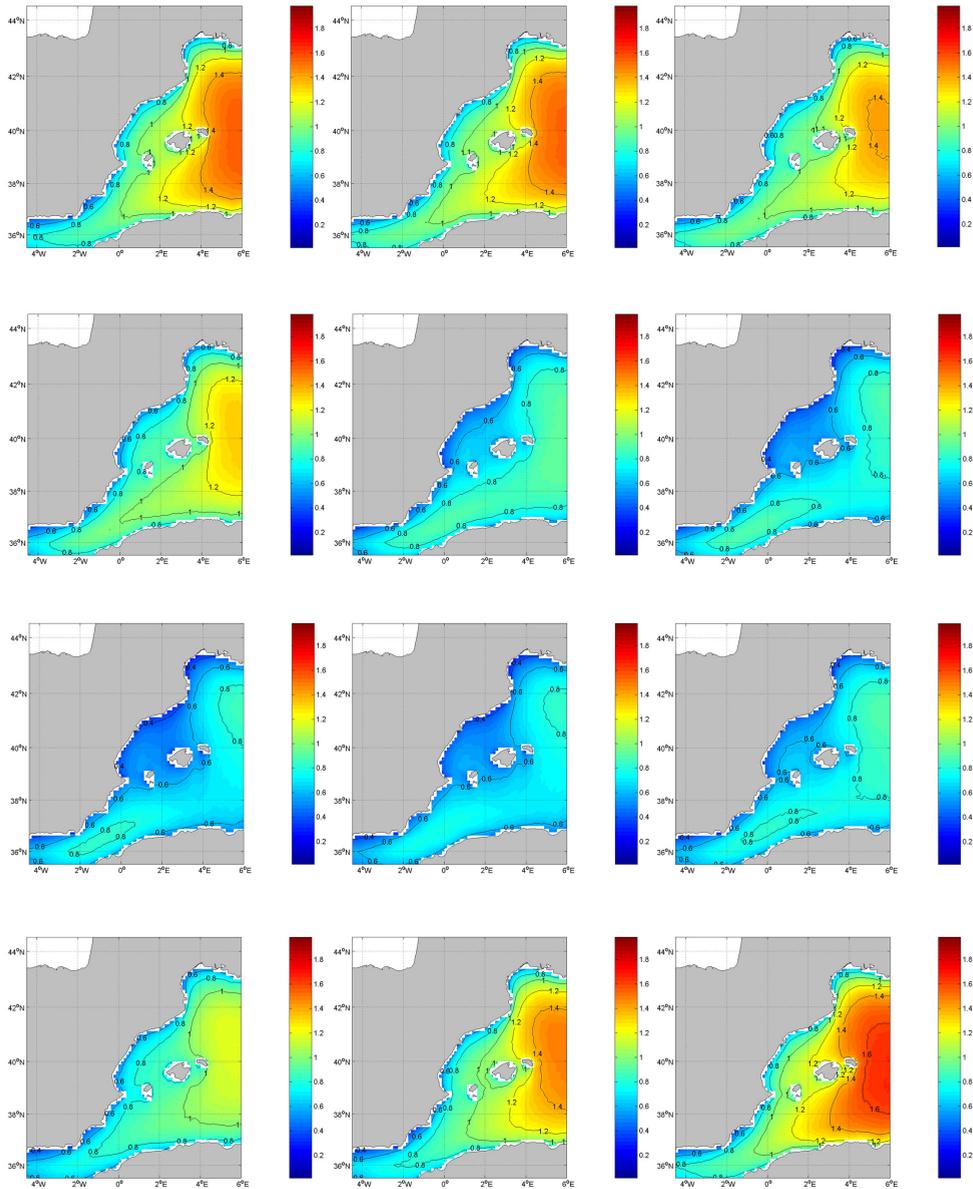


Fig. 3.3: Monthly evolution during the 44 years of the spatial SWH patterns in the Western Mediterranean area. Top-right figure correspond to January, Bottom-left figure correspond to December.

3.1.2 Wave Energy distribution in the Balearic Islands coast

Wave energy is an indicator of the direction and the amount of energy contained in the sea. It is possible to display this distributions in a bidimensional histogram to show how the energy it is not equally distributed, both in direction and energy, in the different coastal areas analyzed. For this purpose three different HIPOCAS nodes were used to characterize the amount of energy contained in each analyzed area of the Balearic Islands coast. The first area corresponds to the HIPOCAS node 1320 located near the coast of Maó, the energy distribution for this point is shown in Figure 3.4. The analysis revealed that most of the energy is concentrated in waves coming from the north with a range of energy between 0.5 and 5 kW/m .

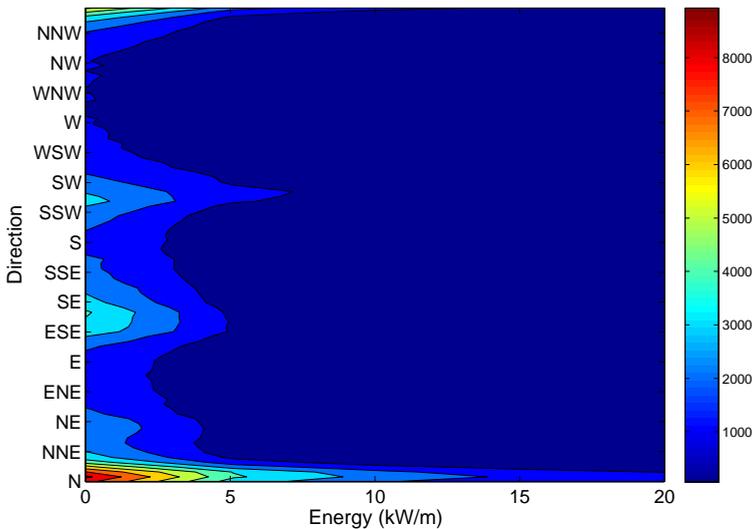


Fig. 3.4: Mean wave energy flux histogram obtained from the HIPOCAS database node 1320, near Maó.

The next histogram corresponds to the HIPOCAS node 1432 located near the Dragonera island in the south-west coast of Mallorca, this Figure 3.5 shows how the energy in this case is divided between two main directions, *NNE* and *SSW*. In the *NNE* direction the range of energy is between 0.5 and 10 kW/m , on the other hand the *SSW* direction shows less energetic SWH with a range between 0.5 and 7 kW/m .

Finally, from the analyzed HIPOCAS nodes we can conclude that the less energetic area corresponds to the south-east area of Mallorca, where the energy range values are around 0.1 and 4 kW/m distributed between two main SWH directions, the *ESE* and the *SW*. The energy distribution in this case is shown in Figure 3.6.

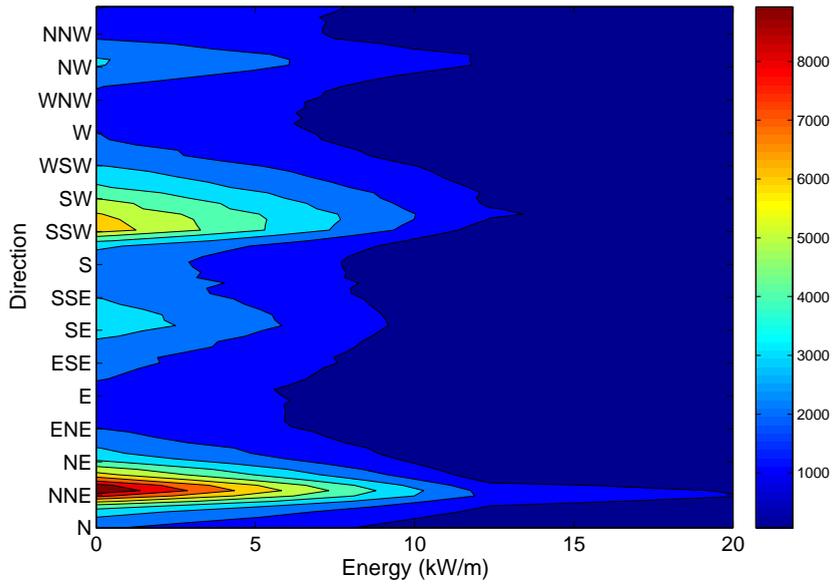


Fig. 3.5: Mean wave energy flux histogram obtained from the HIPOCAS database node 1432, near Dragonera island.

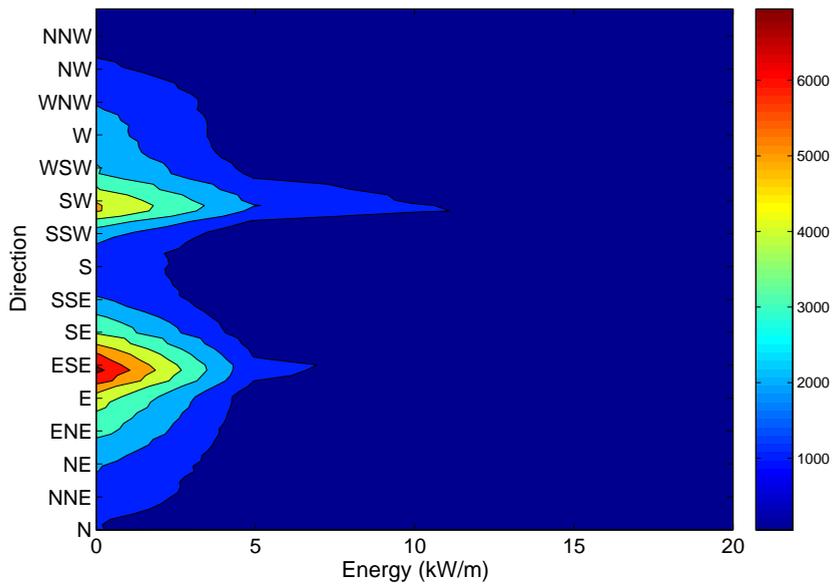


Fig. 3.6: Mean wave energy flux histogram obtained from the HIPOCAS database node 1611, near Cabrera island.

3.1.3 Joint distribution of SWH and periods

For coastal engineering and oceanography it is important not only the distribution of the energy and its direction but the joint distribution of the SWH and peak period (T_p), this is what defines the amount of energy per wavelength in a point. The empirical joint distribution of these parameters in a bidimensional histogram reveals important information about the most probable sea state in an area and its associated to the most probable SWH and peak period. As seen before the analyzed HIPOCAS nodes presents a range of energy between 0.5 and 10 kW/m in the Mallorca coast, this amount of energy is considered as a low energetic range of energy when compared with the surrounding seas.

As an example of high energetic seas, the buoy from Puertos del Estado located in the Cantabric Sea near Bilbao registered a mean SWH energy flux of 30 kW/m . In this area of the spanish coast the most probable sea state in terms of SWH and T_p is $1 < SWH < 1.5$ meters and $9 < T_p < 10$ seconds. On the other hand as we will see in the next Figures 3.7, 3.8 and 3.9 the most probable sea state in the Balearic Sea is in the range of $0.5 < SWH < 1$ meters and $3 < T_p < 5$ seconds, as shown in $SWH - T_p$ histograms.

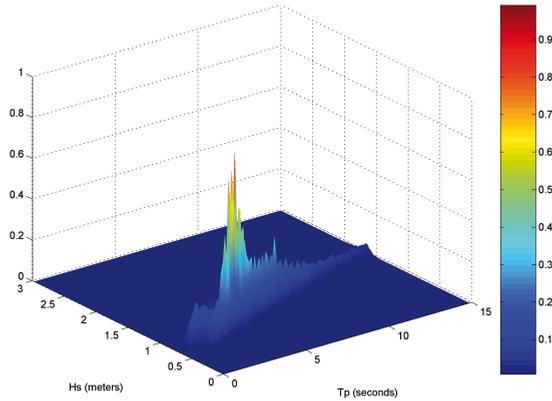


Fig. 3.7: $SWH - T_p$ empirical joint distribution obtained from the HIPOCAS database node 1320, near Maó.

These joint distributions explains why the SWH energy in the analyzed Mediterranean nodes is so low, showing that it is not only due to a low mean SWH but also due to low peak periods. These low values are typical values for areas where the dominant sea state correspond to a fetch-limited sea state due to the short distances between coasts. This short distances between the different Mediterranean coasts produces a limitation in the area where the wind is able to transfer the energy to the sea resulting in "young" seas with short periods and the absence of large swells with

periods between 7 and 14 seconds, normal values in fully-developed seas recorded in areas like the Cantabric Sea.

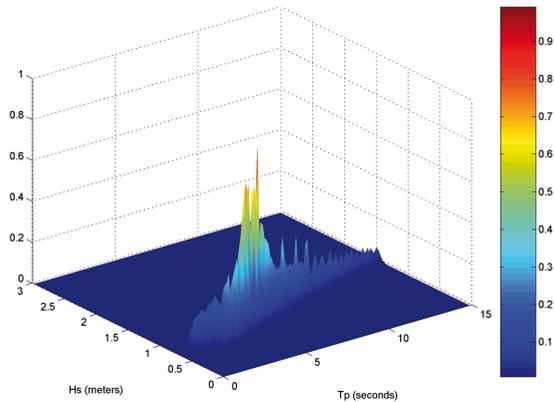


Fig. 3.8: *SWH – Tp empirical joint distribution obtained from the HIPOCAS database node 1432, near Dragonera Island.*

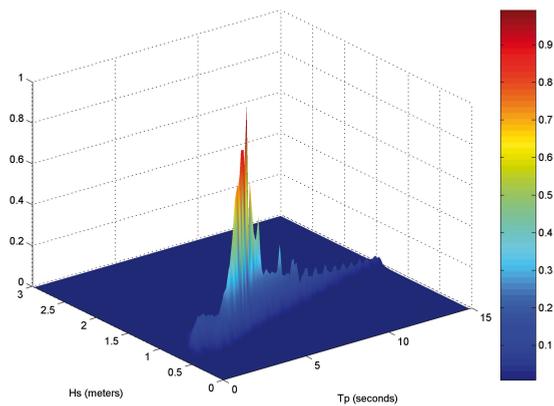


Fig. 3.9: *SWH – Tp empirical joint distribution obtained from the HIPOCAS database node 1611, near Cabrera Island.*

Chapter 4

Extreme wave height estimation in the Balearic Sea

Part of this chapter has been published as an article in the Journal of Coastal Research, Special Issue vol. 50, at the International Coastal Symposium 2007.

Extreme value wave climate analysis at a particular site requires predicting long-term wave height levels from short duration records. In the present work we used the Peak Over Threshold (POT) model, assuming the frequency as a Poisson process and the intensity to be Pareto distributed, to characterize the spatial variability of the long-term extreme value wave climate along the Balearic Sea.

Wave data used is part of the HIPOCAS database, a 44-years high resolution, spatial and temporal, wave hindcast covering an area between 38°N-42°N and 1°W-6°E, of the western Mediterranean Sea. The use of data from a homogeneous grid, instead of a single location wave data record, allows to describe the spatial variability of the long-term extreme wave height levels over the whole Balearic basin. Results show that extreme values for the 50-year return period level are around 11 m in the north sector of the Balearic Islands, while in the southern part lower extreme values are found due to the shadow effect of the islands under severe north (Tramontana) storms. This work is the first statistical characterization of the long-term extreme wave climate in the Balearic Sea.

4.1 Methodology

4.1.1 Threshold and time span selection for the GPD-P

An important issue when modeling threshold excesses with the GPD-P is to choose correctly the threshold value u and the minimum time span δt between successive extreme events.

The extreme events are identified by considering all values larger than a given threshold u and with a minimum time span δt between the storms, to ensure the meteorological independence of the observed excesses. It is not an easy aspect because it requires a balance between bias and variance caused by the selected threshold. If the selected threshold is too low we will violate the asymptotic basis of the model, causing bias. On the contrary a too high threshold will produce few excesses over the selected threshold causing a high variance in the estimated values (Méndez et al., 2006).

Some tools are available to choose the correct threshold. For example an aprioristic test like the mean excess plot can be used (Coles, 2001), Figure (4.1),

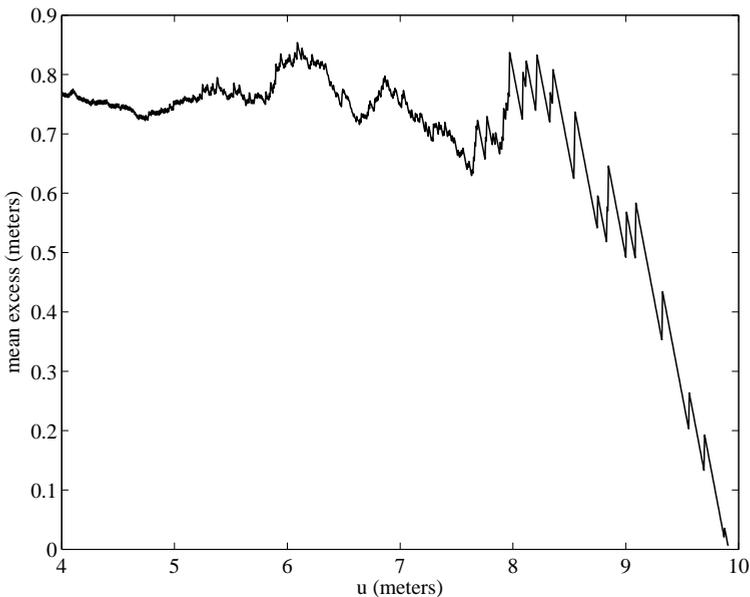


Fig. 4.1: *Example of mean excess plot over a threshold level u*

This test leads to a quick estimation of the shape parameter ξ . Assuming that Y follows a GPD distribution, the mean excess over threshold u , is a linear function of u with slope $\xi(1 - \xi)$. However, sometimes, the mean excess plot can be difficult to interpret, making the decision subjective. As an example, in Figure (4.1), it is shown

a mean excess plot for different thresholds u , from this plot it is difficult to interpret at what level the plot is taking a linear tendency. To avoid the subjectivity in the threshold selection, we use an alternative diagnostic method known as the W-statistic plot (Smith, 2003). The W-statistic is defined as,

$$W_i = \frac{1}{\xi} \log \left\{ 1 + \frac{\xi Y_i}{\sigma + \xi(u - \mu)} \right\} \quad (4.1)$$

This method is based on, if all assumptions are correct, including the selected threshold u and the time span δt , then W_i are also independent and exponentially distributed variables with mean 1.

Figure (4.2) and Figure (4.3) shows the quantile plot (QQ-plot) for the W-statistic, for a selected threshold $u = 4.2$ metres and $u = 1.7$ metres respectively, $\delta t = 72$ hours. As seen in the first figure (4.2), expected values for W are close to the observed ones with a slope near the unit diagonal, indicating the suitability of the selected parameters u and δt . On the contrary the second figure (4.3) shows a clear deviation from the unit diagonal, because the selected threshold is too low and caused bias in the selected probability distribution.

The W-statistic was applied at each grid point for different time spans, between 12 and 144 hours. Finally, a time span $\delta t = 72$ hours was selected for the whole area and a value for the threshold u corresponding to the 99.5% percentile of the empirical distribution was chosen at each grid point.

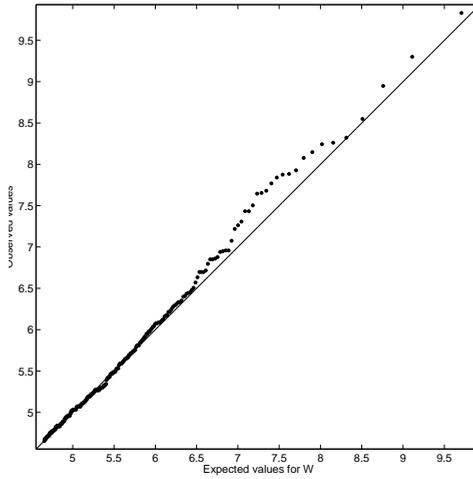


Fig. 4.2: *Quantil-Quantil plot for the W -statistic, $u=4.2$ meters*

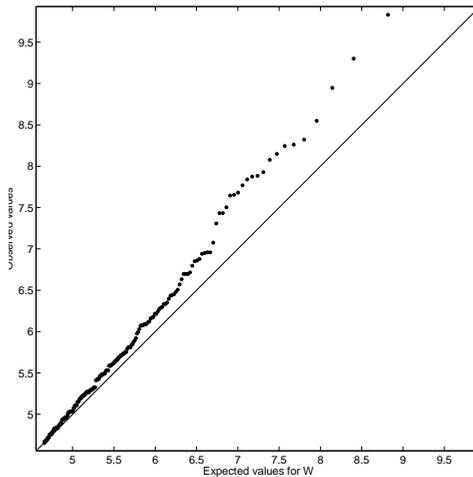


Fig. 4.3: *Quantil-Quantil plot for the W -statistic, $u=1.7$ meters*

4.1.2 Model selection

The selection of the simplest possible model that fits the data sufficiently well is important. Therefore, we check for every point if the contribution of the shape parameter is statistically significant. This is performed using the likelihood ratio test (Coles, 2001).

With nested models $M_A \subset M_B$ (M_B including the shape parameter and M_A with the shape parameter $\xi = 0$), we can assure that model explains substantially (at the α -level of significance) more variability in the data than M_A if,

$$2[\ell_B(M_B)] - \ell_A(M_A) > \chi_{k,1-\alpha}^2, \quad (4.2)$$

where $\ell_B(M_B)$ and $\ell_A(M_A)$ are the maximized log-likelihood functions under models M_B and M_A , respectively, and $\chi_{k,1-\alpha}^2$ is the $1-\alpha$ quantile of the χ^2 distribution with k degrees of freedom, then for $\alpha = 0.05$ and $k = 1$, $\chi_{1,0.95}^2 \approx 3.84$.

Figure (4.4) shows the estimated shape parameter for the whole area and Figure (4.5) shows the statistical significance of the inclusion of the shape parameter.

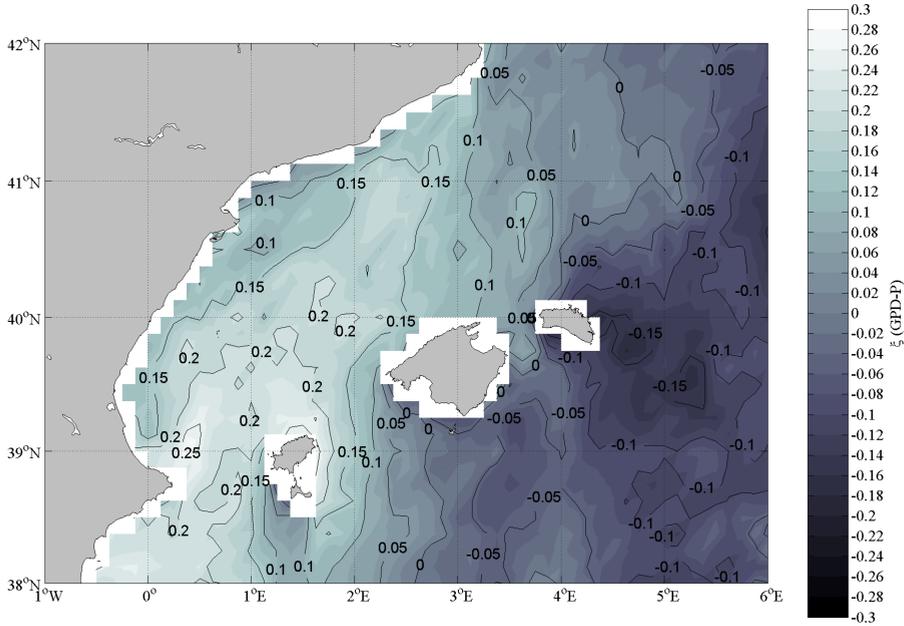


Fig. 4.4: *Spatial distribution of the shape parameter ξ*

As seen in Figure (4.4), we can distinguish three different areas, depending on the value of the shape parameter. For the eastern area of the Balearic Sea the shape parameter correspond to a Weibull tail $\xi < 0$, conversely the western area which correspond to the coasts of Valencia and Catalonia belongs to the Fréchet domain of attraction $\xi > 0$.

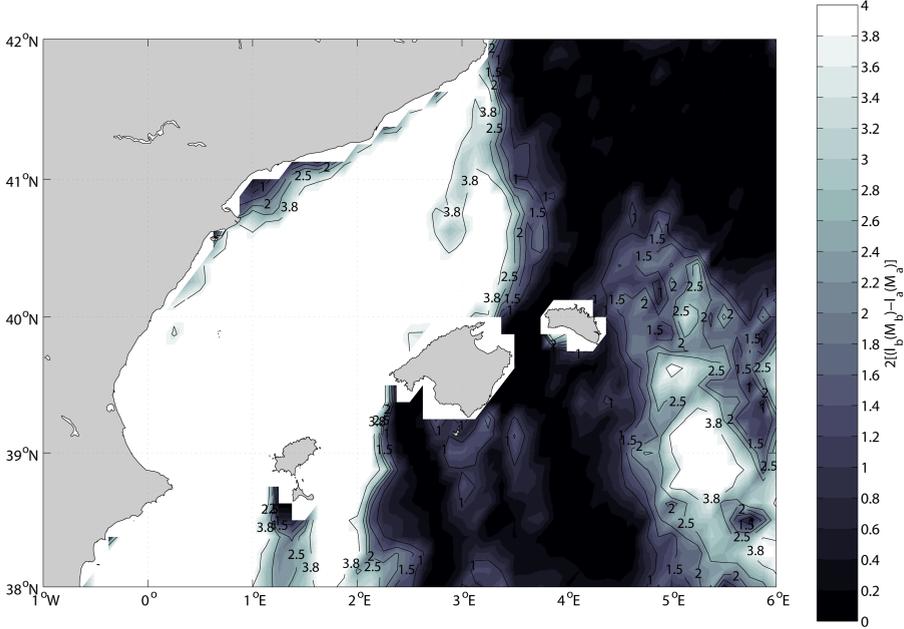


Fig. 4.5: Spatial distribution of the likelihood ratio test

In next figure (Figure (4.5)) the darker area reveals that for this zone the inclusion of the shape parameter has no statistical significance, thus this area corresponds to the Gumbel family for the GEV distribution or the exponential distribution for the GPD, therefore, for the area where the inclusion of the shape parameter ξ is not significant we fix the value $\xi \approx 0$.

4.2 Results

4.2.1 Return Levels for Hs50

We can define the N-year return level as the average time interval in years between successive events of an extreme significant wave height being equalled or exceeded. So, the probability that H_s will be exceeded in any given year is,

$$Pr(H_s \geq x) = 1 - \frac{1}{N} \quad (4.3)$$

being $Pr(H_s \geq x)$ the values of a cumulative probability distribution function and N the desired return period. Therefore a fifty-year return period is equivalent to $Pr(H_s \geq x) = 0.98$. For the 1387 HIPOCAS grid points the GPD-P return period level could be estimated as,

$$H_{sN} = u - \frac{\sigma \lambda^\xi}{\xi \ln(1 - 1/N)^\xi} - \frac{\sigma}{\xi} \quad (4.4)$$

Results for the 50-year return period significant wave height H_{s50} are shown in Figure 3.7. These values are around 11 metres in the northern quadrant of the islands while in the southern part are less than 8 metres. This is the result of the shadow effect of the islands over the intense north fetch produced by the storms. Extreme wave heights over the Catalan coast are significantly lower than those obtained in the north of the Islands due to the angular spreading of the most severe storms.

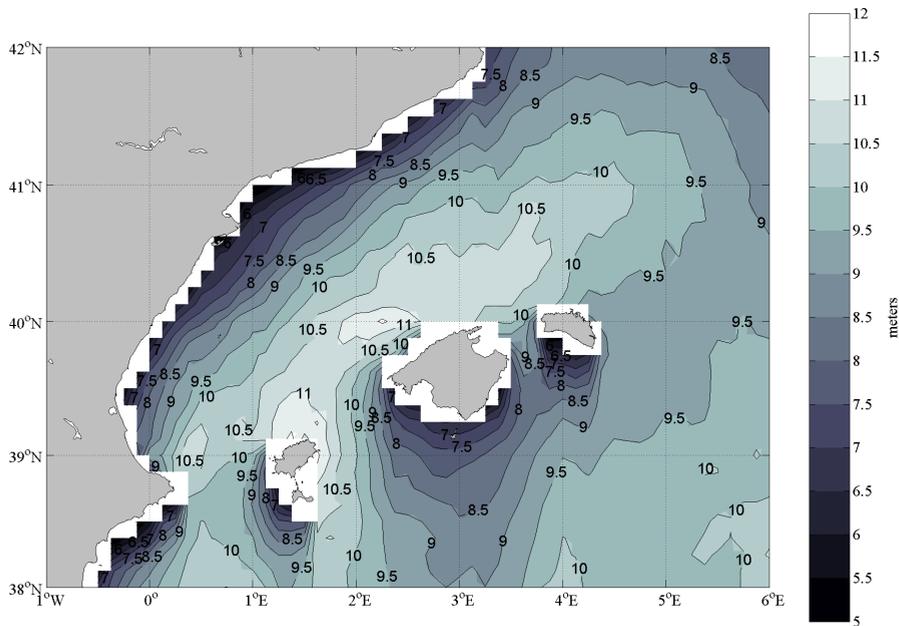


Fig. 4.6: *Spatial distribution of the 50-year significant wave height levels*

4.3 Discussion

Deep water wave climate over the Balearic Sea has in general a complex pattern as a result of the complex orography of the surrounding area. The Mediterranean Sea is well known to be one of the most active cyclogenetic area in the world where the climate is mainly conditioned by severe atmospheric forcing during winters.

The mountains range in the vicinity is a key factor controlling the storm track. The role of the Pyrenees in the west part and the Alps in the north-east area are decisive boundaries for the wind and pressure distribution over the whole basin. The north-western part and central part of the Balearic Sea are forced by northerly winds (Mistral) during the main part of the year, while the eastern part is generally modulated by a seasonal variability. Gale forced mistrals often develop over the Gulf of Genova and the Gulf of Lions extending the effects over the whole basin.

In order to have a rough idea of the behaviour of the storms, the intensity and direction of the maximum significant wave height for the 44 years data are shown in Figure (4.7).

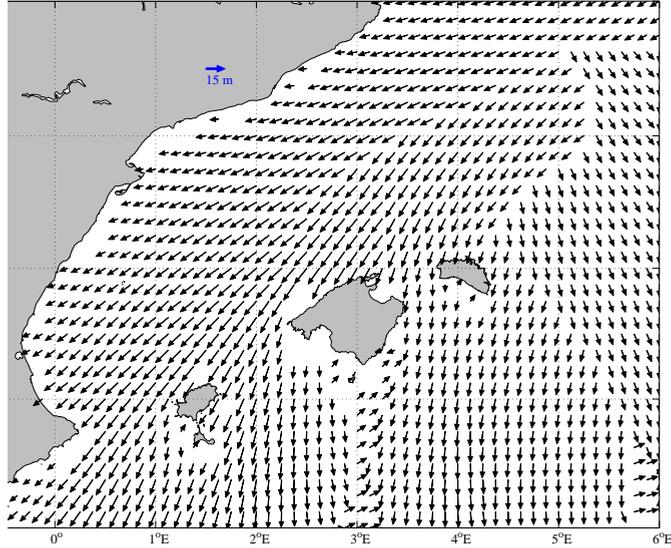


Fig. 4.7: *Intensity and direction of the highest SWH registered in the HIPOCAS database*

As seen, prevailing directions for the maximum *SWH* are from the northeast along the Balearic Sea. This result was previously observed by Sotillo et al. (2005) where a high wind area was identified along the Western Mediterranean from the Gulf of Lions to Northern Algeria and Tunisia.

In the work of Sotillo et al. (2005) wind speeds for a 100-year return period shows a maximum located in the Gulf of Lions with levels of winds up to 30m/s. In the eastern part, differences in wave directions are obtained as a result of the different storm track pattern over this area. As seen, maximum significant wave heights are reached in the Balearic Channel as was obtained from the GPD-P distribution, in our long-term extreme wave climate characterization.

Results also show a spatial variability of the tail of the extreme value distribution: in the western area a bounded tail (Weibull distribution) is detected. This can be associated to a homogeneous distribution in the intensity of the extreme events. On the other hand, along the Spanish coast, the extreme waves tend to be heavy tail distributed (Fréchet distribution). This aspect can be related on the two main storms that affected this area in December, 1981 and November, 2001. This last storm was

driven by heavy Mistral winds (up to 40 m/s) with a large associated fetch, going from the Liguria Sea to the Balearic channel. HIPOCAS nodes located in the area, show maximum significant wave heights over 12 metres at deep waters.

Data used in this work show that the use of a 44-year database can improve the estimation of the extreme SWH values. This improvement is achieved using large populations of extremal data and reducing the uncertainty inherent to the statistical estimation techniques.

4.4 Conclusions

The most important conclusions in this work are enumerated below, including a brief description:

- HIPOCAS data is an adequate database for the extreme wave climate characterization in the Balearic Sea.
- The GPD-P distribution provided more accurate results, in the estimated parameters, when compared to the GEV model.
- The W-statistic method allowed to apply an objective method for the determination of the proper time span δt and threshold value u for each HIPOCAS node.
- The analysis of the shape parameter ξ shows three different areas, corresponding each one to a different family or domain of attraction.
- The inclusion of the ξ parameter is not always statistically significant. Two different areas can be shown, the Western area, between Valencia and Mallorca, and the Eastern area between Mallorca and 6°E.
- For the Balearic basin and the Catalonia coast the spatial distribution of the H_{s50} parameter have been obtained.

The HIPOCAS database is an adequate tool for climate studies in areas like the Mediterranean. The REMO modelling of the NCEP/NCAR global reanalysis data is an important improvement, allowing global reanalysis data being downscaled to a regional scale. This improvement in the spatial resolution of the atmospheric model results in a better estimation of the wind forcing. Therefore, the hourly wave fields of the last 40 years have been hindcasted with a high spatial and temporal resolution.

The GEV distribution has as principal shortcoming the use of only a data per year in the Annual Maxima Method, causing high a variance in the estimated values. From the analysis we can conclude that the GPD-P distribution provides a better estimation than the GEV distribution, for the extreme wave climate analysis.

In the Peaks Over Threshold Method the extreme events are identified by considering all values larger than a given threshold u and with a minimum time span δt between the storms. After analyzing all the HIPOCAS nodes, a good balance between bias and variance is achieved when the threshold level u correspond to the 95th percentile SWH at each point and a time span δt of 72 hours.

Three different areas can be distinguished attending the type of tail behaviour. The heavy tailed area between Catalonia and Mallorca is the response of the GPD-P distribution to a two different large storms occurred in this area. The tail type distribution is an essential information for characterizing the wave climate.

After analyzing the statistical significance of the shape parameter ξ we can demonstrate that in some areas a simpler model can be used setting the shape parameter to $\xi \approx 0$. The shadow effect of the Balearic Islands and the angular spreading of the storms produce a reduction in the magnitude of the higher return levels in the coastal areas of Catalonia and the southern coasts of the Balearic Islands. The analysis performed provided a good estimation of the spatial variability of the 50-year return period significant wave height for the Balearic Sea and Catalonian coasts

This is the first wave climate characterization carried out in the Western Mediterranean using data from a high resolution hindcasted database.

Chapter 5

Influence of the NAO on the Northwestern Mediterranean wave climate

Part of this chapter has been published as an article in Scientia Marina, vol. 74, num. 1.

This study examines teleconnections between the North Atlantic Oscillation (NAO) and the wave climate of the Northwestern Mediterranean Sea (NWM), defined by its monthly mean significant wave height (SWH) and the 95th percentile significant wave height (95th percentile SWH), in the period ranging from 1958 to 2001. Analyzed data is part of the multidecadal hindcast over Europe carried out during the HIPOCAS project. In order to avoid fictitious cross-correlations, data were prewhitened by fitting a p-order autoregressive model. To split the temporal and spatial variability, an EOF encoding technique was applied to residuals prior to search for teleconnections. The Northwestern Mediterranean wave climate has been found to be influenced by the North Atlantic Oscillation (NAO) with an instantaneous response. When the NAO is in its positive phase, positive anomalies in SWH and the 95th percentile SWH appear in the area between the Balearic Islands, the Gulf of Lions and the Catalonia coast.

5.1 Introduction

Teleconnections are interactions between widely separated parts of the ocean and the atmosphere occurring at different time scales. These interactions originate recurring and persistent modes of low-frequency variability in large-scale patterns of atmospheric pressure and ocean circulation anomalies (Orfila et al., 2005, Vignudelli et al., 1999). These anomalies in large-scale circulation are manifested in surface waves, driving some of the variability of wave climate, which has consequences, among others,

in the operation and safety of shipping, offshore activities, and coastal development (Kushnir et al., 1997). Therefore, the search and recognition of teleconnections is an important issue for the prediction and characterization of the wave climate variability over time.

A fraction of the interannual variability in the climate over Europe is associated with a single pattern of pressure anomalies over the north eastern Atlantic area, known as the North Atlantic Oscillation (NAO) (Hurrell, 1995). The NAO has long been identified as a major factor in climate variability and oceanic variability in many areas of the northern hemisphere (e.g. Rogers and Jeffrey (1997); Rodwell et al. (1999); Vignudelli et al. (1999); Hurrell et al. (2001); Lionello and Sanna (2005); Lionello and Galati (2008)). It can be described as an irregular oscillatory transfer of mass between two main centers of action, the Icelandic Low and the Azores High, that modify the paths of the storms crossing the north Atlantic from the east coast of America to Europe (Woolf et al., 2002).

Wave climate analysis requires a large amount of data to ensure the statistical significance (Canellas et al., 2007). Traditionally, these data have been collected using scalar and directional wave buoys moored at specific locations providing high temporal resolution records.

In the last decade, satellites have been used to overcome the spatial lack of data (Krogstad and Barstow, 1999) but the problem of having a large amount of spatial and temporal wave records were still unresolved. Queffeuou (2005) and Queffeuou and Bentamy (2007), used altimeter data to perform an analysis of the wave height variability over the Mediterranean Sea but, altimeter data has as a shortcoming its temporal inhomogeneity and a coarse spatial resolution in areas like the Western Mediterranean, marked by a complex orography.

Alternatively, wave generation models are another option to avoid the usual lack of data in ocean and atmospheric studies. Models are initialized with real conditions and the deviation due to the nonlinearity of the governing equations corrected with the assimilation of data. Numerical models can be now implemented in very fine grids. Although, usually hindcasted model data, underestimate the actual sea wave height (SWH) when compared with buoy observations, satellite data and simulations forced by higher resolution wind fields, they provide a reliable representation of the real space and time variability (Lionello and Sanna, 2005). For this reason, these hindcast models have become a powerful tool not only for engineering or prediction scales but for climate studies involving large temporal periods.

The Mediterranean Sea is well known to be one of the most active cyclogenetic areas in the world. The orography of the surrounding area with the Pyrenees in the Western part and the Alps in the North-Eastern part are natural barriers controlling

the generation and track patterns of cyclones over the Western Mediterranean area.

In this area, climate is mainly conditioned by severe atmospheric forcing during winter season, when the polar front jet is more likely to influence Mediterranean cyclogenesis and weather (Trigo et al., 2002). Most of the strong winds observed in the Mediterranean belong to the category of local winds and are originated by air-flow/mountain interaction, as down slope flows or due to channeling effects. Concretely the Northwestern Mediterranean area is forced by northerly and north-westerly winds during the main part of the year, while a less intense cyclogenetic activity is observed during the rest of the year (Trigo et al., 2002, Canellas et al., 2007).

This study explores the influence of the monthly NAO index as a general descriptor of the pressure field variability in the North Atlantic areas, over the monthly average SWH fields in the northwestern Mediterranean Sea, and the 95th percentile significant wave height, in the period ranging from 1958 to 2001.

To do that, 44 years of hourly wave data with 0.125° spatial resolution obtained from HIPOCAS project (Ratsimandresy et al., 2008) in the northwestern Mediterranean Sea (Figure 5.1) has been analyzed.

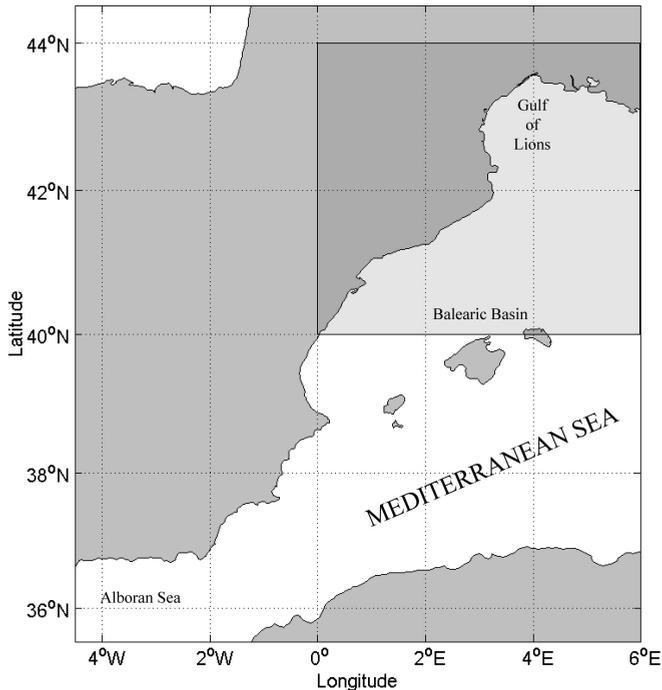


Fig. 5.1: *Geographic location of the study area.*

The underlying assumption is that wave climate for the northwestern Mediter-

anean area is influenced by large-scale atmospheric patterns of variability, like the NAO phenomenon, originated in the adjacent Atlantic Ocean. The article is structured as follows: Section 2 describes the methodology employed in the study. Results obtained from the EOF analysis are shown in Section 3. Section 4 discusses and concludes the work.

5.2 Methodology

Monthly NAO index data were provided by the Climate Analysis Section, NCAR (Hurrell, 1995). The NAO index used in this study has been calculated as the difference in the normalized pressure between Ponta Delgada in Azores and Stykkisholmur in Reykjavik (Iceland). The wave parameters chosen for the correlation with the NAO index are the monthly mean significant wave height (SWH) and the 95th percentile SWH.

The monthly mean and standard deviation for SWH were obtained through fitting hourly time series to a Log-normal probability distribution to the entire HIPOCAS database (Jaspers, 1956). There is no theoretical basis for selecting the initial unknown probability distribution function to represent the observed data. Some other probability distributions are also valid for this purpose, being the most common the bi-parametric Weibull distribution and the Log-Normal distribution. These two distributions are good enough to describe the mean behaviour of the SWH time series.

However, we did not use the whole record to compute the mean and standard deviation to avoid a possible contamination in the SWH statistics, due to the influence of extremal maxima/minima. Then, fitted data correspond only to SWH between the 5-95th percentiles of the cumulative distribution function, in order to exclude SWH that correspond to the extremal range of probability.

As a result, the mean and standard deviation for each month at each node were obtained from this probability distribution. Moreover, the 95th percentile SWH (corresponding to those wave heights only surpassed during 5% of time each year) was obtained using the same theoretical probability function. With this parameter, only the highest SWH are retained, being a good indicator for the most extreme long term SWH episodes.

Once hourly data are aggregated into the monthly statistical parameters, data are encoded with Empirical Orthogonal Functions (EOFs) to identify the preferred modes of variability (Preisendorfer, 1988). EOF decomposition of the temporal variance describes those spatial patterns with the highest contribution to time variability. On the other hand, if the EOF decomposition of the spatial variance is carried out spatial patterns with the strongest spatial gradients will be obtained (Álvarez, 2003).

In this study, we have focused on spatial variance EOF modes since we are interested in the possible influence of NAO in the variability of the spatial patterns of wave climate. Time series from the temporal amplitudes associated to those EOFs are used to cross correlate with the monthly NAO index.

Cross-correlation functions strongly depend on the structure of the original time series through the autocorrelation, implying that nonzero values of the cross-correlation function do not necessarily represent a relationship between two time series if they are auto-correlated (Katz, 1988). In order to avoid possible artificial cross correlations, a prewhitening of each time series was carried out. Prewhitening fits a p-order autoregressive model AR(p) of the form,

$$x(t) = a_1x(t-1) + \dots + a_px(t-p) + \Phi(t) \quad (5.1)$$

to the original time series, sufficient to reduce the residuals $\phi(t)$ to white noise. The order of the autoregressive is selected by the Akaike Information Criterion (Akaike, 1974). The Akaike Information Criterion (AIC) determines the model order p by minimizing a theoretical information function given by,

$$AIC(p) = N \ln(\sigma_x^2(p)) + 2p \quad (5.2)$$

where N is the number of samples and σ_x^2 is the estimated variance of the white driving noise and the term $2p$ in equation 5.2 is a "penalty" for the use of extra AR coefficients that do not substantially reduce the prediction error (Konyaev, 1990). Once obtained the value of the filter order p by minimizing equation 5.2 the residuals obtained after fitting the AR(p) process in equation 5.1 are not autocorrelated time series.

5.3 Results

5.3.1 Spatial Variance EOF modes for monthly mean SWH

The first two EOFs account for the 91.03% of the total variance (Figure 5.2) and are related to the spatial patterns with the highest contribution to the spatial variance of the significant wave height fields.

For the first EOF, which accounts for the 82.79% of the total variance (Figure 5.3a), the resulting EOF decomposition shows a strong spatial gradient in the west-east direction, with a maximum anomaly located on the East boundary.

Conversely second spatial EOF mode, which accounts for the 8.25% of the total variance, (Figure 5.3b), shows no spatial gradients. This pattern displays an isostatic response of the whole area, modulated throughout the year by the temporal ampli-

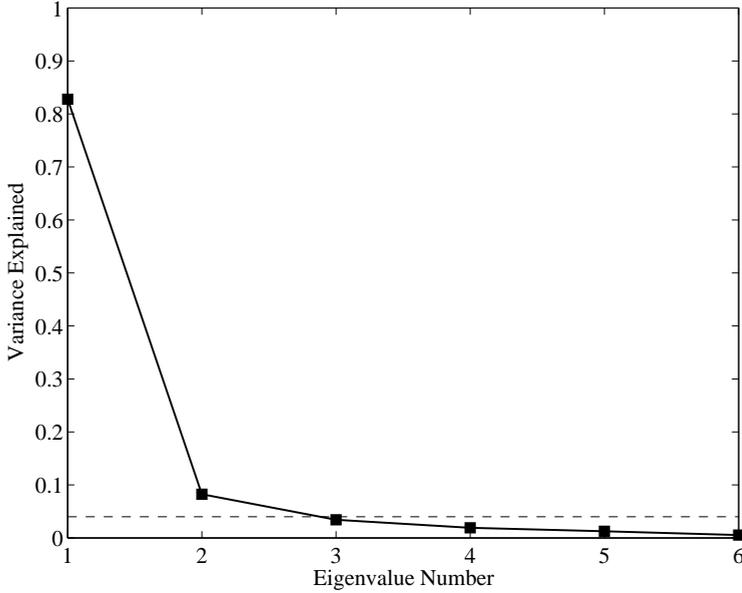


Fig. 5.2: Variance explained by the first six EOFs, for the monthly SWH. Grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.

tude time series.

The order applied for the prewhitening attending the AIC is, $p=2$ for the NAO index and 12 for the temporal amplitudes of both EOFs. The spatial mean time series, computed for each of the 528 months (not prewhitened), is shown in Figure 5.4a.

Maximum values are reached during the Autumn-Winter months. For the monthly mean SWH, the autocorrelation function (not shown) is negative autocorrelated at time lag $\tau = 6$ indicating a strong seasonality.

For the first EOF, the temporal amplitude presents an annual cycle highly influenced by seasonality (Figure 5.4b,5.4c), with Autumn-Winter months taking the highest positive amplitudes (Figure 5.4b), while spring-summer months presents less intense significant wave height fields in the whole area analyzed.

Cross-correlation between first temporal amplitude time series and the NAO index provides, as expected, no significant correlation at the 95% level of confidence. The power density spectrum (not shown) for this amplitude was calculated showing a dominant peak at twelve months indicating that this mode explains the variability introduced by the annual cycle. On the other hand, significant correlation was found

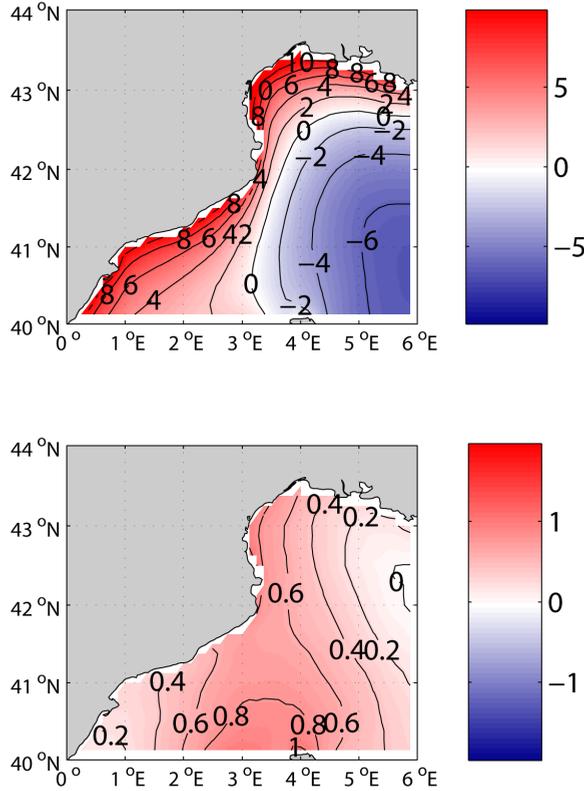


Fig. 5.3: (a-top) First spatial variance EOF mode for the monthly mean SWH, (b-bottom) Second spatial variance EOF mode for the monthly mean SWH.

(at the 95% confidence bounds) between the NAO index and the amplitude of the second EOF. Maximum correlation is at time lag $\tau = 0$ (Figure 5.5), implying an instantaneous response at monthly time scales.

When the time amplitude is in its maximum positive values, the response of this second EOF mode can contribute up to 0.20 m to the mean SWH in the area between the Balearic Islands, the Gulf of Lions and the Catalonia coast.

5.3.2 Spatial Variance EOF modes for 95th percentile monthly mean SWH

For the 95th percentile SWH the first four EOFs modes were retained. These modes explain the 92.38% of the total variance (Figure 5.6).

Like in the monthly mean SWH results, the 95th percentile SWH variability is mainly reproduced by the first mode which accounts for 72.54% of the total vari-

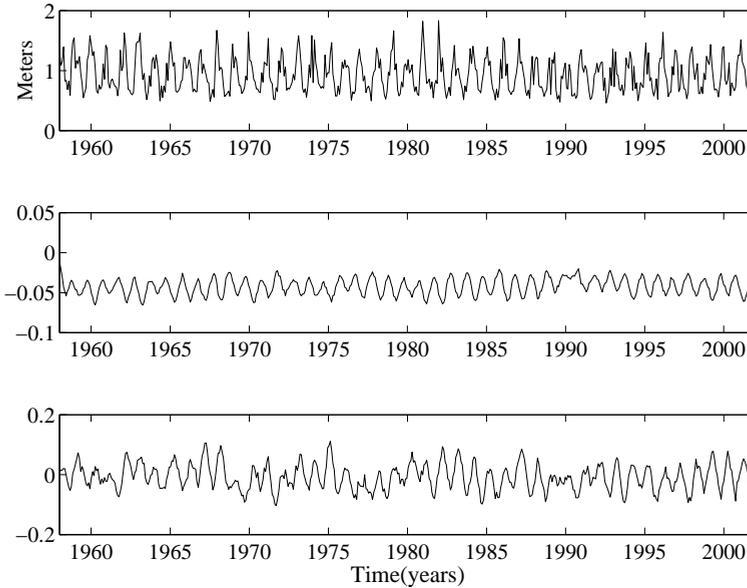


Fig. 5.4: (a-top) Time evolution of the spatial mean computed from 528 months for SWH, (b, c) Time amplitudes for the first two EOFs.

ance. The spatial pattern for this mode (Figure 5.7a) shows a longitudinal gradient.

For negative amplitudes we can observe negative anomalies along the coast and positive anomalies between the Balearic Islands and the Gulf of Lions. The second EOF mode accounts for 10.87% of the total variance and its spatial pattern displays the same isostatic response as seen in the monthly mean SWH results (Figure 5.7b). The third EOF mode, accounts for 5.10% of the total variance and its spatial pattern displays a maximum negative anomaly in the East boundary (Figure 5.7c). The location and orientation of this pattern reflects the fetch associated to SW winds. The fourth EOF mode accounts for 3.87% of the total variance and its spatial pattern displays a maximum anomaly between the continental shelf and the Balearic Islands (Figure 5.7d).

Prewhitening was applied to the 95th percentile time amplitude time series. The filtering order p , is again 12 for the four temporal amplitudes analyzed. The spatial mean time series for the 95th percentile SWH is shown in Figure 5.8a.

Maximum values are reached during the Autumn-Winter months. Similarities were found between the monthly mean SWH and the 95th percentile SWH, not only in the spatial mean time evolution, but also in the two first EOF modes. The auto-correlation function (not shown) anti-correlates the mean at lag 6, indicating negative correlations when comparing values for the Summer-Winter months. Both wave pa-

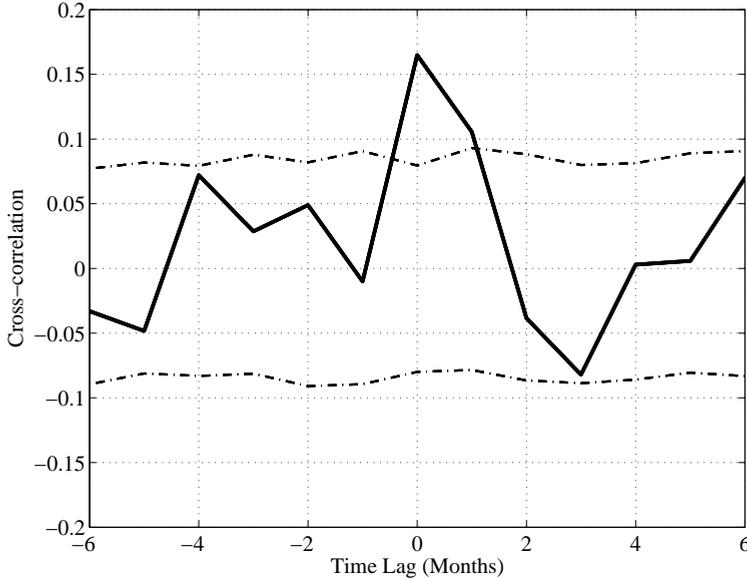


Fig. 5.5: *Cross correlation between second spatial variance EOF mode time amplitude and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations.*

rameters (monthly mean SWH and the 95th percentile SWH) show an identical time evolution, with maximum and minimum values at the same time lag. However, main differences between them are found in their intensity, between their spatial patterns as well as in the variance explained by each mode.

Strong seasonality is also present in the first temporal amplitude, showing maximum values during winters, and minimum values during summer. Cross-correlation between first temporal amplitude for the 95th percentile SWH (Figure 5.8b) and the NAO index provides, as expected, no significant correlation at the 95% level of confidence (not shown). On the contrary significant correlations are found in the 2nd, 3rd and 4th temporal amplitudes. The 2nd EOF amplitude (Figure 5.8c) is correlated with the NAO index at time lag $\tau = 0$, (Figure 5.9a). The maximum anomaly reaches 0.60 meter in the whole area when the amplitude is in its maximum.

Time amplitude for the 3rd mode (Figure 5.8d) is also correlated with the NAO index at time lag $\tau = 0$ (Figure 5.9b). The contribution of this mode indicates that when the NAO index is in its positive phase, a negative spatial anomaly appears in eastern area, indicating less intense extreme SWH fields. When time amplitude is in its maximum values the maximum anomaly reaches 0.60 meter in the area corresponding to a fetch from the Corsica Island to the Gulf of Lions.

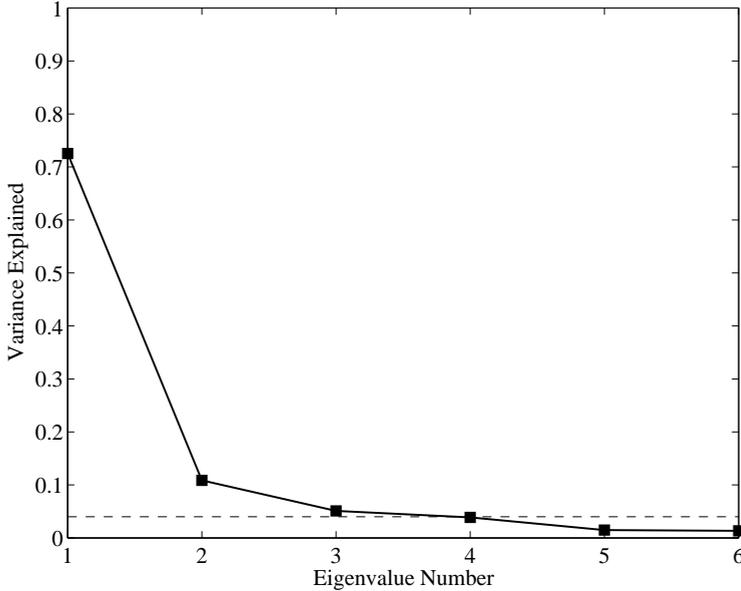


Fig. 5.6: Variance explained by the first six EOFs, for the monthly 95th percentile SWH. Grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.

Finally, the fourth EOF temporal amplitude (Figure 5.8e), as previous amplitudes analyzed, shows a significant correlation with the NAO index at time lag $\tau = 0$ (Figure 5.9c). The basic difference with other amplitudes is its intensity. This amplitude has a value close to zero during the whole record indicating that this mode might not contribute to the reconstructed field, for most of the time. However, the maximum amplitude obtained falls in this mode (December 1980) contributing to the spatial field with 2.50 meter in the 95th percentile SWH. This anomaly is constrained to the area between the Catalonia coasts and the Balearic Islands.

5.4 Discussion and Conclusions

In this study was shown that the NAO phenomenon contributes to the spatial variability found in the Northwestern Mediterranean. When the NAO is in its positive phase a reinforcement of the northerly cold and dry air masses from the arctic regions appear, generating more severe weather conditions over the Northwestern Mediterranean Sea (Hurrell, 1995, Hurrell et al., 2001). On the contrary the negative phase of the NAO exhibits a weakening of the westerlies all across the North Atlantic onto Europe.

The results obtained in this study were computed with one of the several indicators

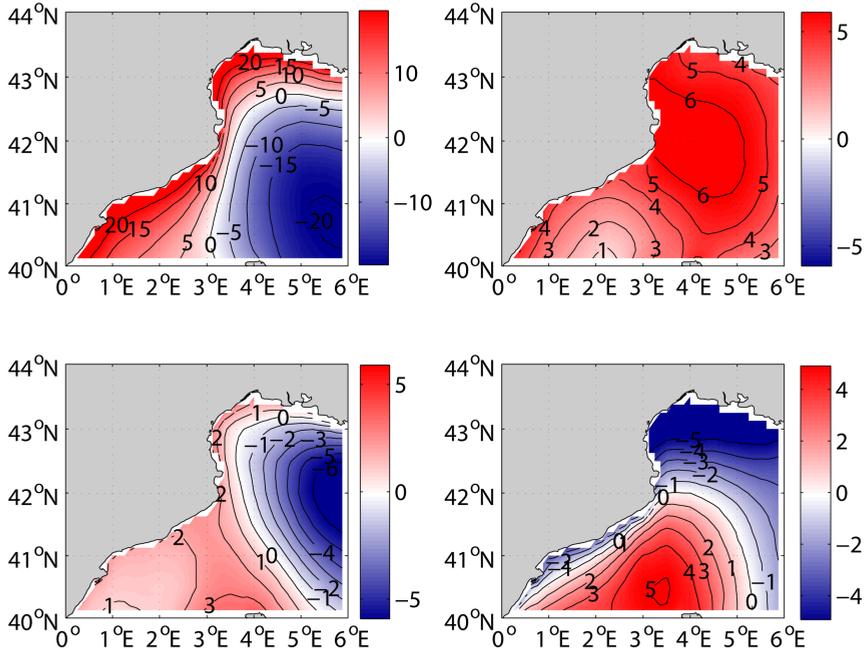


Fig. 5.7: (a-top left) First spatial variance EOF mode for the 95th percentile SWH, (b-top right) Second spatial variance EOF mode for the 95th percentile SWH, (c-bottom left) Third spatial variance EOF mode for the 95th percentile SWH, and (d-bottom right) Fourth spatial variance EOF mode for the 95th percentile SWH.

of the NAO activity. Other useful indexes are also available in the NCAR database. One of them is the DJFM (December to March) NAO index. This DJFM NAO index when was correlated with the December to March SWH aggregated returned higher values because it is in this period of the year when the NAO signature is much more noticeable. On the contrary, in the December to March aggregation the loss of information is evident, masking other important key controlling atmospheric patterns, such the intraannual atmospheric/oceanographic cycle and the monthly evolution of the studied wave climate parameters.

The Northwestern Mediterranean climatology has a high cyclogenetic activity during the Autumn-Winter months, associated with severe weather conditions and driving intense winds from the northeast and northwest (Campins et al., 2000, Picornell et al., 2001, Trigo et al., 2002). On the contrary, in summer the wind pattern changes, and the direction and intensity is more variable. This is in agreement with the results obtained for the first temporal amplitudes in both analyzed fields (monthly mean SWH and 95th percentile SWH), with maximum amplitudes located in the winter months. This pattern was already observed and characterized by Lionello and Sanna

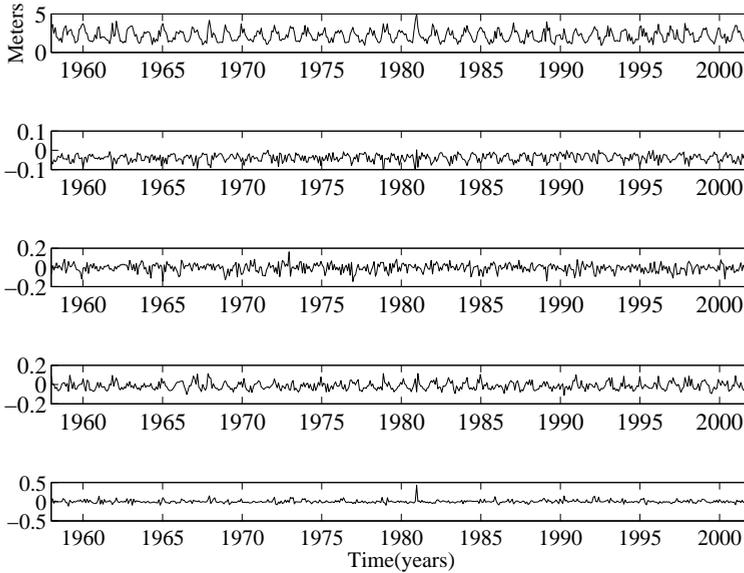


Fig. 5.8: (a-top) Time evolution of the spatial mean computed from 528 months for the 95th percentile SWH, (b, c, d, e) 2nd ,3rd and 4th EOF time amplitudes.

(2005) as the annual cycle for the monthly mean SWH.

First mode explains around the 70-80% of the total variance, for both fields. After analyzing the spatial decomposition of monthly SWH statistical parameters, power spectral densities, autocorrelations and cross-correlations, we suggest that first modes explain only the variability introduced by the Summer-Winter intraannual cycle. We could not detect any influence of the NAO phenomena in the first mode which explains most of the intraannual variability.

Correlation between NAO and the second EOF amplitude has been found for the spatial variance in monthly mean SWH, this significant correlation between the second EOF amplitude for the monthly mean SWH and the NAO index was already observed by Lionello and Sanna (2005). On the other hand the second, third and fourth EOF amplitudes for the 95th percentile SWH are significantly correlated as well. The response of these modes to the NAO was found to be instantaneous at monthly time scale, for both wave parameters. Both second EOFs modes for the monthly mean SWH and the 95th percentile SWH show clear similarities reflecting an isostatic response of the whole area modulated throughout the data records by the time amplitude. The anomalies for both modes, in the monthly mean SWH and 95th percentile monthly mean SWH, contribute in 0.20 meter and 0.60 meter, respectively.

For the 95th percentile SWH two different patterns are shown for the third and

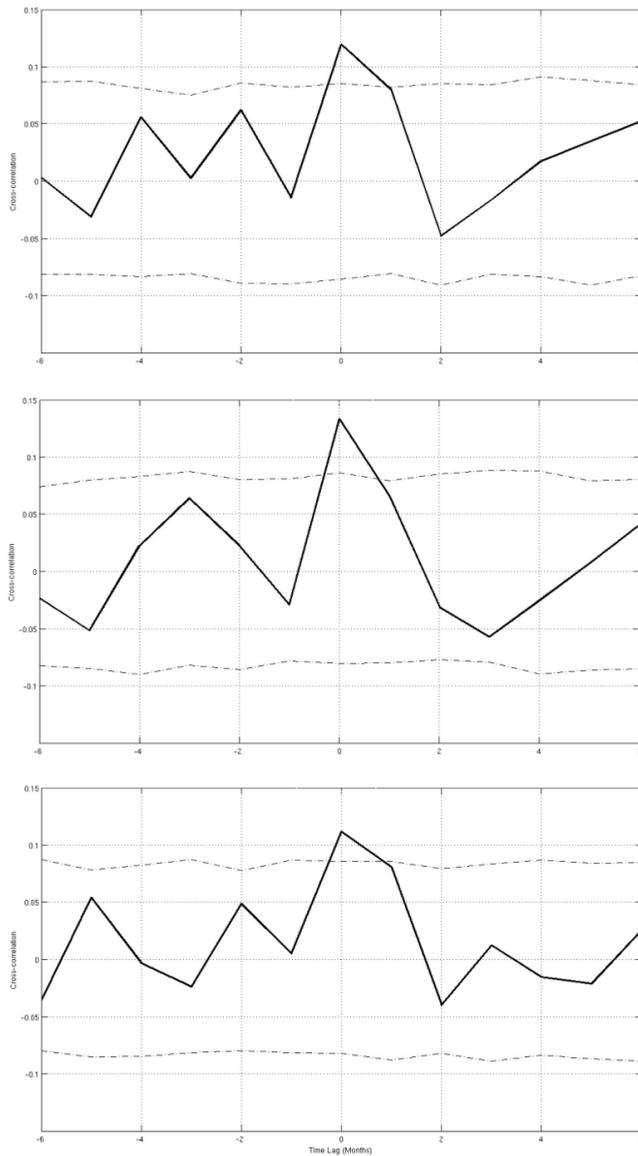


Fig. 5.9: (a-top) Cross-correlation between second, third and fourth spatial variance EOF mode time amplitudes and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations.

fourth modes. Third mode displays a dipole shaped pattern with a maximum anomaly in the east boundary. The location and orientation of this pattern reflects the fetch associated to the intense South-West winds originated in the Corsica Island and prop-

agated to the Gulf of Lions, most common during the wintertime (Figure 5.10a).

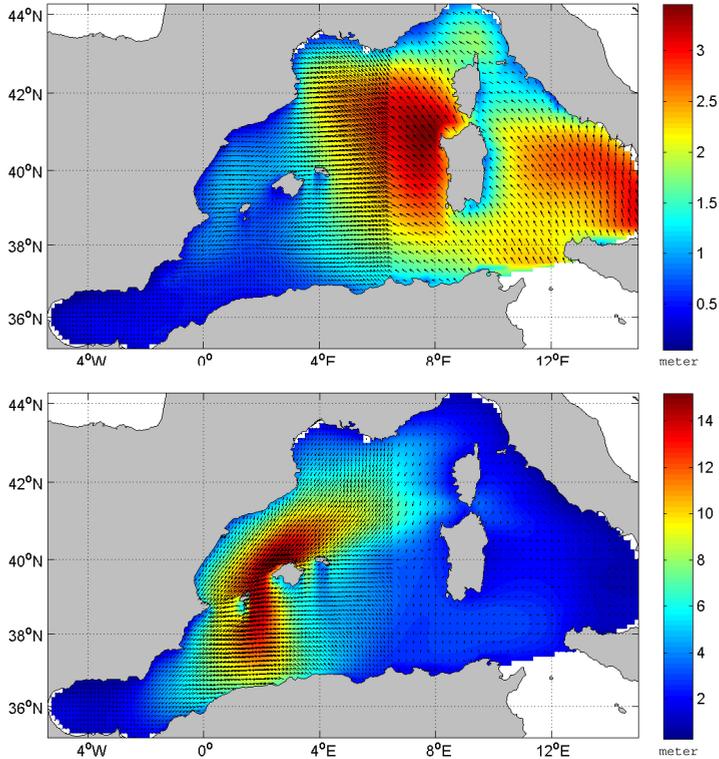


Fig. 5.10: (a-top) Snapshot corresponding to the HIPOCAS database (12/01/2001), it shows the SWH field associated to the South-West winds, generated near the Corsica Island and propagated towards the Gulf of Lions. (b-bottom) Snapshot corresponding to the HIPOCAS database (12/28/1980), it shows the fetch associated to the severe North-East winds, generated in the Gulf of Genoa and propagated towards the Catalonia coasts, Valencia coasts and the Balearic Islands.

This negative anomaly in the eastern area during the positive phase of the NAO reduces the monthly mean SWH values up to 0.60 meter. Fourth EOF for the 95th percentile SWH mode also displays a dipole shaped pattern with maximum values for the anomaly between the Spanish continental shelf and the Balearic Islands. The location and orientation of this pattern reflects the fetch associated to the intense winds from the North-East which are often associated to the most severe sea storms recorded in the Western Mediterranean. These storms usually cause coastal damage and large economic loss in locations like Mallorca and Valencia.

In December 1980 (as well as November 2001, not shown) a severe storm was

recorded in the area between Catalonia and the Balearic Islands, see for instance the amplitude in the 4th EOF for the 95th SWH percentile (Figure 5.8e). In this event, wind waves were originated in the vicinity of the Gulf of Genoa with a direction of propagations to the South-West (Figure 5.10b). The observed fetch has the same spatial pattern than the fourth EOF mode. This mode suggests that the variability corresponding to these extreme events are reflected in this spatial pattern.

The correlation of long term wave climate parameters with the NAO phenomena indicates that some of the variability of wave height is associated with the NAO in the Northwestern Mediterranean Sea. However a strong seasonal signal was found in both parameters, monthly mean SWH and the 95th percentile monthly mean SWH, as expected from previous studies of the cyclogenetic activity of the Mediterranean. This fact has been already noted by Lionello and Sanna (2005). Both analyzed parameters provided good results for the analysis, however, we have seen that the 95th percentile of SWH provides more dynamical information (patterns of variability in higher order EOFs) than the monthly mean SWH.

Chapter 6

Wave Height prediction in the Western Mediterranean using Genetic Algorithms

Part of this chapter has been accepted for publishing as an article in Ocean Engineering.

This study explores the suitability of a soft computing technique to predict (forecast and hindcast) Significant Wave Heights in some areas of the Mediterranean Sea. Using a 44-year database of meteorological and oceanographic records we train a Genetic Algorithm, to reconstruct a 44-year time series of Significant Wave Height at different coastal points of the Balearic Islands (Western Mediterranean). The correlation obtained between the numerical model data and the algorithm output is $0.80 \leq r^2 \leq 0.97$. The equations obtained by the Genetic Algorithm are then used to predict Significant Wave Heights measured at several buoys. We make use of the equation from the node closest to the buoy using as input wind data from the meteorological station at each buoy. We obtained that SWH at each buoy can be predicted with a 1-hour ahead forecast time using only wind data from the previous 12 hours with a correlation of $0.87 \leq r^2 \leq 0.91$.

6.1 Introduction

Prediction of ocean surface waves is an important issue for some of the most relevant human activities in the coastal zone. Beach nourishment, port design and operability, dispersion and diffusion of pollutants are some examples that require a precise knowledge of predicted Significant Wave Height (SWH) fields as well as their extreme values (Sverdrup and Munk, 1947). SWH is an important oceanographic parameter that is

required to define a sea state or estimate the wave climate in an area. Traditionally, prediction of ocean surface waves requires the use of numerical models which integrate the fluid governing equations. These models have become a powerful tool not only for engineering or prediction works but also for climate studies involving large temporal periods. However, one of the major drawbacks of the numerical models is the large computational resources required for integrating the primitive equations over fine grids or large areas.

An alternative to prediction based on first physical principles is the use of past observations to predict the evolution of a dynamical system. Takens (1981), provided the theoretical background to build prediction models from past observations of chaotic time series. Specifically, Taken's theorem states that given a deterministic time series of observations $\{x_i\} = \{x(i \cdot \tau) \mid i = 1 \dots M\}$, where τ is the sampling time, there exists a map $P(\cdot)$ satisfying,

$$x_m = P(x_{m-1}, x_{m-2}, \dots, x_{m-d}), \quad (6.1)$$

being d the embedding dimension.

For instance, Álvarez (2003) used Genetic Algorithms (GA) to forecast the space-time variability of the Sea Surface Temperature in the Alboran Sea (Western Mediterranean); Basu et al. (2005), predicted significant wave heights in an area of the Arabian Sea, using different time series of surface wind speeds and SWH from three different oceanographic buoys as input for a GA; Gaur and Deo (2008), used an alternative soft computing tool known as genetic programming (GP, (Koza, 1992)) to explore the usefulness of this method for real-time wave forecasting in a limited area of the Gulf of Mexico, also this method is actually used by Guven (2009) for time series modelling of daily flow rate at a river in US; Ustoorikar and Deo (2008), compared two different empirical prediction methods (GP and artificial neural networks (ANN)) to fill up gaps in wave data from four wave rider buoys maintained by the US National Data Buoy Center in the Gulf of Mexico. The ANN method was also used by Arena and Puca (2004) to perform a reconstruction of SWH records from two correlated buoys from the NOAA in California.

GA are one of the most extended techniques used for empirical prediction. They are based on the evolutionary genetic laws, e.g. reproduction, crossover and mutation where Darwinian theories of natural selection and survival are used to approximate a map that best describes a physical chaotic system.

In this work we apply a multivariable GA to a set of buoy and numerical model wave data to predict directly from past values of wind module and wind direction the SWH in the coastal areas of the Balearic Sea (Western Mediterranean Sea). Moreover, this study explores the possibility to forecast SWH in coastal areas using a GA trained with wind at 10 meters from a meteorological model and SWH from a numer-

ical wave model.

Our results demonstrate the possibility not only to reproduce a 44-year time-series from a numerical database, but to also predict SWH directly from wind buoy observations, using the equations obtained previously by training an evolutionary algorithm.

6.2 Data and Methods

Wave and wind data used to train the algorithm are part of the HIPOCAS Project (Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe) (Soares, 2008). This database consists on a high resolution, spatial and temporal, long-term hindcasted data set covering on an hourly basis a period ranging from 1958 to 2001 providing 44-years of wave data over an homogeneous grid. HIPOCAS dataset was produced by means of dynamical downscaling from the NCEP/NCAR global re-analysis using the regional atmospheric model REMO (Kalnay et al., 1996, Jacob and Podzun, 1997). Hourly wind fields from the REMO were used as the forcing for a third generation wave model WAM, see Sotillo et al. (2005), for more details. SWH data used in this work is the output of the WAM model implemented in a 0.125° resolution mesh over the Western Mediterranean Sea. A total of 49 coastal nodes were analyzed, covering an area between 38°N - 40.5°N and 1°E - 4.5°E , see Figure (6.1).

For each HIPOCAS node, the equation that best describes the dynamics of the SWH is obtained by means of a GA (Alvarez et al., 2001), which is trained using hourly mean wind module (U), hourly mean wind direction ($\bar{\theta}$) and SWH data from the HIPOCAS database.

The algorithm for predicting the state variable $\{x_i\}$ considers an initial population of randomly generated equations $P_j(\dots)$ derived from random combinations of explanatory variables $\vec{V}_i = \vec{V}(i \cdot \tau)$, $i = 1 \dots M$, randomly chosen constants and the four basic operators ($+$, $-$, \times , \div). For a μ -step ahead prediction with embedding dimension d , the fitness of the j -candidate equation map $P_j(\cdot)$ is computed as,

$$\Delta_j^2 = \sum_{m=\mu+d}^M \left[x_m - P_j(\vec{V}_{m-\mu}, \dots, \vec{V}_{m-\mu-d+1}) \right]^2 \quad (6.2)$$

where x_m and \vec{V}_m denote the value of the explained variable and the vector of explanatory variables, respectively, at time $t_m = m\tau$.

The fitness to the data, Δ_j^2 , establishes the strength of each individual in the selection process. Specifically, the strength index for the j -equation string is expressed as,

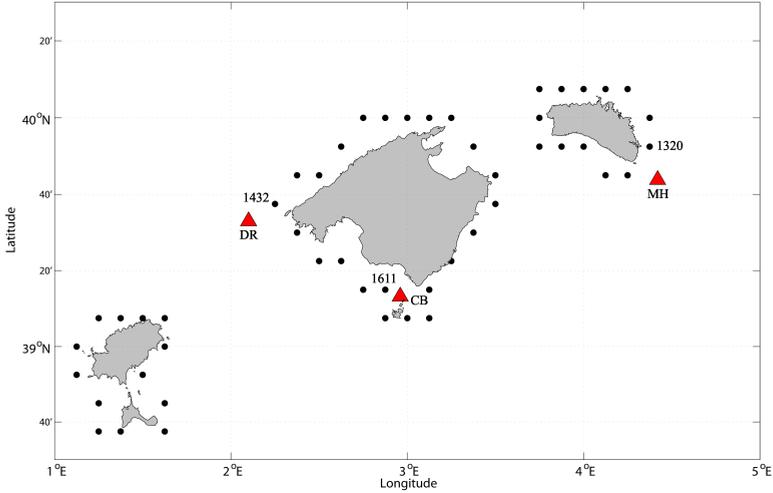


Fig. 6.1: Geographic location of the study area. Black dots correspond to the HIPOCAS coastal nodes and triangles to the oceanographic buoys, MH (Mahon), DR (Dragonera) and CB (Cabrera).

$$F_j^2 = 1 - \frac{\Delta_j^2}{\sum_{m=\mu+d}^M (x_m - \bar{x})^2} \quad (6.3)$$

where \bar{x} is the mean value of the training data. F_j can be interpreted as the percentage of the training set's total variance explained by the j -equation string and so, for a given individual, the higher this quantity, the better the data are represented by the corresponding equation.

This population of potential solutions is next subject to an evolutionary process, by which those individuals that best fit the data are selected from the initial population. This process is carried out as follows. First, part of the time series is selected as the training set, i.e. the set of data that will be used in computing the best approximation to the map $P(\cdot)$ in, Eq. (6.1).

By examining the fitness of $P_j(\cdot)$ the strongest individuals are then selected to exchange parts of the character strings between them (reproduction and crossover), while the individuals less fitted to the data are discarded. Finally, a small percentage of elements in the equation strings are mutated at random, except for the top ranked equation strings in order to avoid losing their information. As a result of this pro-

cess a new set of equations is obtained. The evolutionary steps are repeated with a new generation and finished after a number of generations determined by the user. Afterwards, a validation is carried out using the data not included in the training set to test the goodness of the candidate equation for reproducing the deterministic signal.

A total of 49 coastal nodes from the HIPOCAS database have been used as the input for the training and validation of the GA. For this particular case the use of GA in a vectorial configuration made possible the inclusion of the wind module and the mean direction as input variables, because usually the evolution of a chaotic natural system is not restricted to the influence of only one observed variable, but a nonlinear interaction among several variables is quite common.

To check that SWH is driven by local meteorological conditions we have performed a multivariate regression analysis of the SWH time series of HIPOCAS against wind parameters (intensity U and direction $\bar{\theta}$). In the present case the multivariable GA in its vectorial form can be expressed as,

$$SWH_m = P(\vec{V}_{m-\mu}, \dots, \vec{V}_{m-\mu-d+1}) \quad (6.4)$$

SWH_m being the significant wave height at time t_m and

$$\vec{V}_m = (U_m, \cos \theta_m, \sin \theta_m) \quad (6.5)$$

is the vector of explanatory variables at time t_m .

Moreover we have used data from three deep ocean buoys in order to demonstrate the capability of the algorithm to retrieve a nonlinear equation that can mimic not only the non-linear interactions between meteorological and oceanographic variables from a numerical model database, but also capable of generating a valid equation for real-time forecasting at offshore stations, using the same exact equation as in the nearest HIPOCAS node.

After the training and validation process, the best ranked equations at three coastal nodes (numbered nodes in Figure 6.1) were used at the closest three available deep water buoys to test the ability of these equations to forecast SWH data directly from the buoy wind observations.

6.3 Results

As a first approximation to the problem of finding a predictive equation that would provide the future evolution of the SWH, a multiple regression analysis of SWH at time t_m was performed using as input the information at time t_{m-1} contained in the meteorological series of the HIPOCAS database. The mean correlation coefficient over all the HIPOCAS nodes is 0.88 with a standard deviation of 0.02, which implies that most of the total variance in the HIPOCAS SWH time series can be explained

by the local meteorologic conditions —wind intensity and direction— at the previous time.

Hence we applied the GA in a vectorial configuration to train the algorithm with a multivariate set of time series containing only the module of the wind U and the wind direction $\bar{\theta}$. The GA algorithm requires a large number of parameters to be configured when the underlying dynamical system is unknown. In our case, this configuration required a large trial and error process. Finally the selected parameters were $d = 12$ for the embedding dimension and $\tau = 1$ for the time lag. The number of individuals in each population has been set to 120 and the number of iterations to 10000.

For the training process the selected total number of points of the time series was 3000 including a validation set fixed at 1000 data points. This parameters were the most efficient ones in terms of the quality of the output and the CPU time needed to reach the desired performance, which was the achievement of optimum fitness strength $F^2 \geq 0.9$.

6.3.1 Prediction of HIPOCAS data

The cross-correlation coefficient for the 49 HIPOCAS coastal nodes (r^2) calculated using the 44-year time series from the numerical model and the 44-year predicted SWH from the GA, is shown in Figure (6.2).

As seen, the GA is able to reproduce the dynamical behaviour of the hourly SWH time series, with a cross-correlation coefficient that remains between values of 0.70 and 0.97 with a mean correlation calculated over all the nodes of $r^2 = 0.90$. At this point, we want to remark that less than 1% of the total data were used in the learning process of the GA.

To provide a detailed view of the performance of the GA, SWH at three selected coastal nodes (HIPOCAS 1320, HIPOCAS 1432 and HIPOCAS 1611, see Figure 6.1) are compared with the SWH predicted by the GA. Figure 6.3 (top left, middle left and bottom left) shows the performance of the GA output displaying one month of data (12/08/1968 - 01/27/1969) with a 1-hour ahead prediction for the nodes 1320, 1432 and 1611.

It can be seen the good agreement between the numerical model SWH and the GA predicted SWH. The best ranked equations provided by the GA for these locations are respectively,

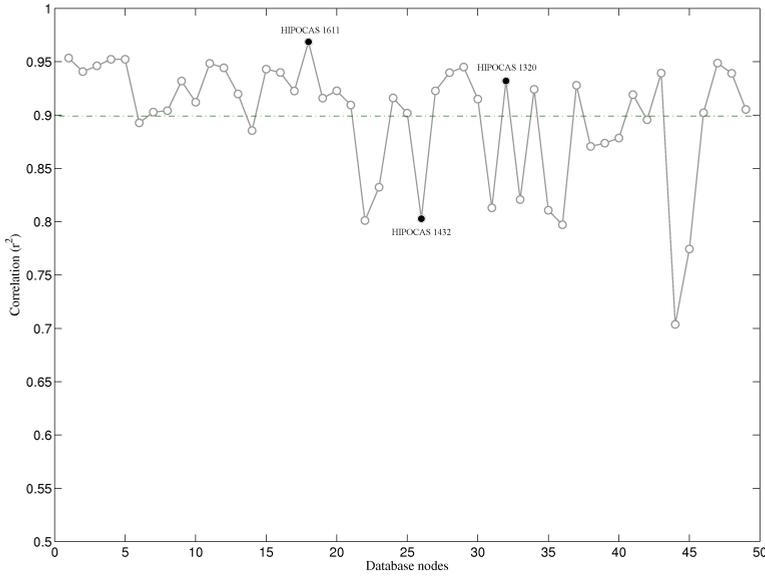


Fig. 6.2: Linear correlation (r^2) obtained at the 49 HIPOCAS coastal nodes, between the genetic algorithm output and the numerical model data. Dashed line corresponds to the mean correlation calculated over the nodes $r^2 = 0.90$. Filled dots correspond to the HIPOCAS nodes presented in the Results section, which are the closest to the available buoys.

$$x_0(t) = 0.12x_1(t-10) \frac{0.22x_1(t-1)^2}{2.11 + x_3(t-12)x_3(t-5) + x_3(t-2)x_3(t-5)} \quad (6.6)$$

$$x_0(t) = 0.14(x_1(t-1) + x_1(t-7) - \frac{7.79x_3(t-10)x_3(t-4)}{2.59 + 0.15x_1(t-2)x_3(t-1)x_3(t-4)}) \quad (6.7)$$

$$x_0(t) = \frac{x_1(t-2)}{8.09 - 0.18x_3(t-1) - x_2(t-6) - 0.18x_1(t-10) - 0.18x_2(t-12)} \quad (6.8)$$

being t the time instant, x_0 the SWH, x_1 the wind module, x_2 the cosinus of the wind direction and x_3 the sinus of the wind direction.

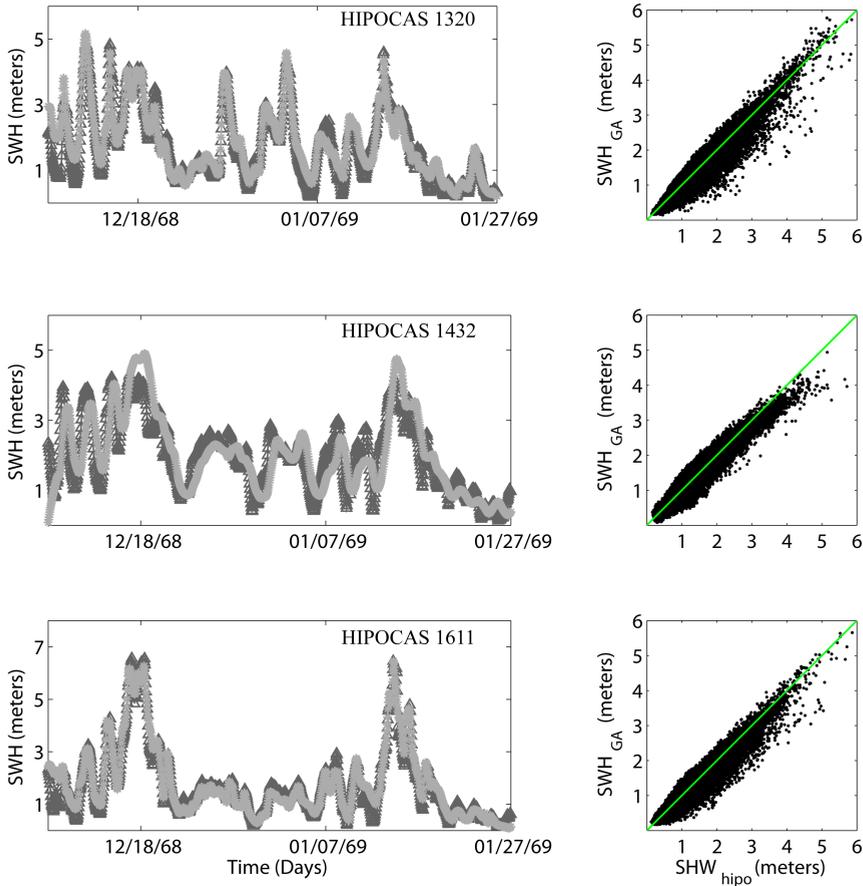


Fig. 6.3: Left panels: Hourly time series comparison, between the genetic algorithm predicted SWH and the HIPOCAS nodes SWH (1320, 1432 and 1611). Black triangles represent the GA predicted SWH, grey asterisks correspond to the HIPOCAS database SWH. Right panels: Scatter diagrams comparing GA SWH and the HIPOCAS SWH data, unit diagonal corresponds to a perfect fit.

In order to evaluate the errors reported by the GA with the 1-hour ahead prediction, scatter diagrams of the SWH from the numerical model database, plotted against the SWH obtained through the nonlinear equation of the GA, are shown in Figure 6.3 (top right, middle right and bottom right) for the three nodes. These scatter diagrams provide information about the behaviour of each equation used at the different coastal nodes. It can be seen how in these particular HIPOCAS nodes a general underestimation of the SWH from the numerical model database exist. The larger underestimation arises in the range of SWH between 2 and 5 meters, but in general we can see how the errors are equally distributed along the ideal fit of slope $k = 1$, except for the largest SWH that present stronger deviations.

The absolute error in SWH estimation related to wind intensity U and direction $\bar{\theta}$, computed as the difference between SWH from the database and predicted SWH from the GA, is shown in Figure 6.4. The error (in meters) plotted against the wind module U for HIPOCAS nodes 1320 is displayed in Figure 6.4 (top left).

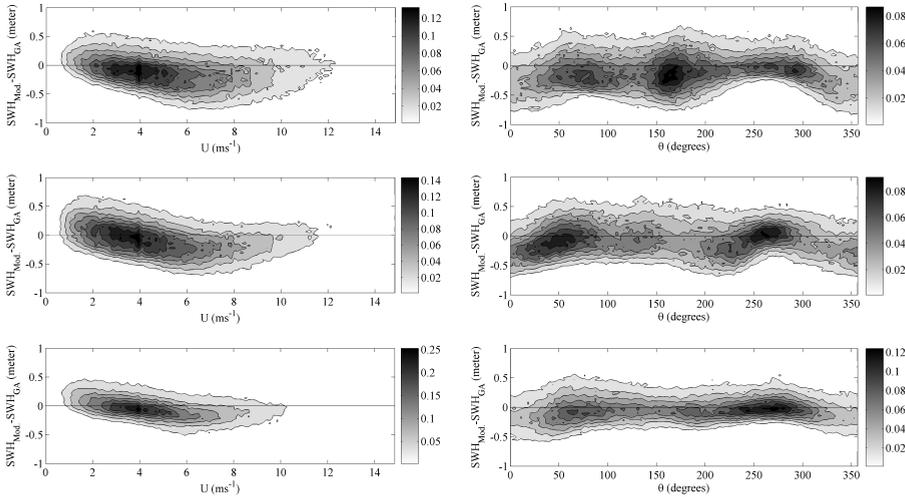


Fig. 6.4: Left panels: Contour plot of the absolute error between HIPOCAS database SWH and predicted SWH at the selected HIPOCAS nodes against the wind intensity U , countour lines indicate the percentage of the population with the indicated error. Right panels: *idem* but compared against the mean wind direction $\bar{\theta}$.

We can observe how the dispersion of the error is larger as U increases. In this case the error shows a rotation around $U = 5ms^{-1}$. An underestimation error is observed for wind module values under $U = 5ms^{-1}$, but for wind module values larger than $U = 5ms^{-1}$ the nonlinear equation overestimates SWH from the numerical model database.

The error plot for HIPOCAS node 1432 is shown in Figure 6.4 (middle left). This plot shows the same behaviour in the error dispersion for wind speeds over $4ms^{-1}$. In this case the equation obtained is also overestimating SWH at high wind speeds over $U = 4ms^{-1}$.

The same error behaviour as in the node 1320 and 1432, is also observed for the HIPOCAS node 1611 (Figure 6.4, bottom left), with SWH observations underestimated for wind module values under $U = 4ms^{-1}$ and overestimated SWH values when $U > 4ms^{-1}$.

As seen, wind module plays an important role in the predicted SWH and its error. The error by underestimation is generally produced by wind speeds $U < 5ms^{-1}$. On the contrary the error due to the overestimation of the numerical model database

values generally arises from wind speeds larger than $U > 5ms^{-1}$. The dispersion in error increases in both directions as the wind module $U \geq 6ms^{-1}$.

The evaluation of the error dispersion related to the wind direction also seems to be important. To show how the error is also dependant on the wind direction $\bar{\theta}$, Figure 6.4 (top right, middle right and bottom right), shows the error in SWH estimation against the wind direction $\bar{\theta}$. We can observe how the error is equally distributed along the whole range of directions at each node. Also it is noticeable a range of directions where the percentage slightly increases. The errors arise from directions ranging between $30-100^\circ$, $150-200^\circ$ and $200-350^\circ$, where the observed error is given by the overestimation of the numerical model database, except for the middle-right plot (node 1432) in the $250-300^\circ$ where it is shown a slight underestimation of the numerical model SWH.

Finally, a forecasting skill test was also performed for the HIPOCAS node 1320 to asses the goodness of the method for long-term prediction. This test considered different threshold values in equation 6.2 by modifying the μ -step parameter, hereby allowing for μ -hours ahead predictions. The values used in equation 6.2 for the μ parameter were 3,6,9 and 12 (hours).

The obtained correlation coefficients (r^2) between the prediction and the original HIPOCAS time series were: 0.88, 0.87, 0.77 and 0.58 respectively. Obviously, the 1-hour ahead prediction fits the data with a higher degree of correlation ($r^2 = 0.94$), but it is worth remarking that high values ($r^2 > 0.80$) are still obtained even for a 6-hour ahead prediction.

6.3.2 Prediction of Buoy data

In order to test the ability of the obtained equations to predict real wave data, three deep water buoys providing real time wave and wind parameters have been used. Two of these buoys are part of the real time observing system from Ente Publico de Puertos del Estado (EPPE), known as the Spanish deep ocean buoy network (REDEXT), while the third one belongs to the Mediterranean Institute for Advanced Studies (IMEDEA). These three oceanographic buoys were the only available real data in the studied area and the selected HIPOCAS nodes in Figure(6.1) are the closest nodes to the buoys.

Two of these buoys, Mahon Buoy (MH) ($39.73^\circ\text{N}-4.42^\circ\text{E}$) and Dragonera Buoy (DR) ($39.55^\circ\text{N}-2.10^\circ\text{E}$), belong to the the REDEXT network and contain 3 and 10-year of data, respectively. The data consists on hourly meteorological and oceanographic parameters such as wind module and direction, SWH, wave period and wave direction. Cabrera Buoy (CB) is another deep water meteorological and oceanographic

graphic buoy located between Mallorca and the Cabrera Island (39.22°N - 2.96°E). This buoy belongs to IMEDEA and contains only a year of oceanographic data covering 2008.

The performance of the GA obtained at the REDEXT buoys is shown in Figure 6.5.

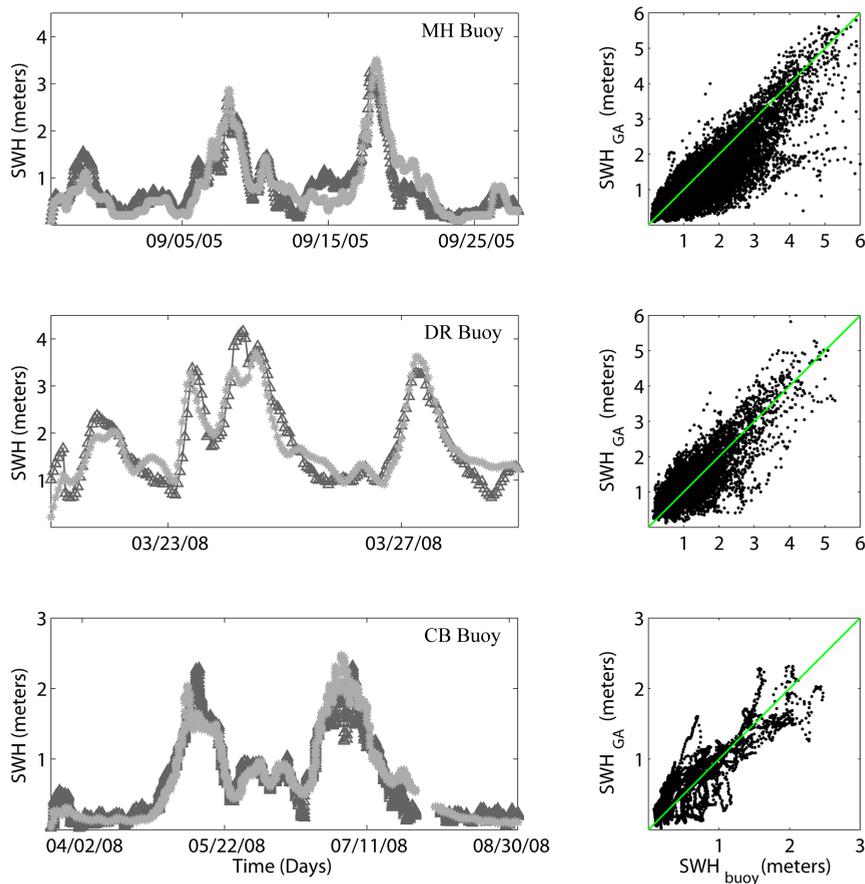


Fig. 6.5: Left Panels: SWH measured at the three buoys (MH, DR and CB) (grey asterisks) compared with the predictive equations SWH (black triangles) computed using the buoy's wind records. Right Panels: scatter diagram of these data.

The time period chosen for the reconstruction of the buoy data is constrained to one month for the MH buoy, one week for the DR buoy and 4 months for the CB buoy, in order to avoid the large gaps found in the time series. To test the strength of the GA under extreme maximum and minimum wind and wave conditions, we also selected a time period in the series where the wave climate conditions presented a high deviation from the mean SWH values to include different wave climate conditions with wave

heights ranging from 0.30 to 4.0 meters, and wind speeds ranging from 0.7 to 15 ms^{-1} .

The prediction time series at the MH buoy was obtained using a month of wind data from the buoy and with the same nonlinear equation found through the analysis of the HIPOCAS node 1320. As shown in Figure 6.5 (top left) the hourly forecast represents an accurate reproduction of the original time series. The correlation obtained after analyzing a month of data is $r^2 = 0.89$.

The results for the DR and CB buoys are shown in Figure 6.5 (middle left and bottom left). DR buoy and CB buoy were predicted using the equations obtained at the HIPOCAS nodes 1432 and 1611, respectively. The correlation obtained at these buoys was $r^2 = 0.91$ for the DR buoy and $r^2 = 0.87$ for the Cabrera buoy. After the reconstruction of the three buoys the mean correlation obtained was $r^2 = 0.89$.

To determine the behaviour of the error from the forecasted SWH, three scatter diagrams are shown in Figure 6.5 (top, middle and bottom right). These scatter plots are generated using the entire record available at each buoy. As in the numerical model database prediction we can observe how the error arises principally from the underestimation of buoy data. The range of data where the algorithm provides less accurate predictions is located between $3.0 \leq \text{SWH} \leq 6.0$ meters for the MH buoy. On the contrary scatter diagrams for the DR buoy (Figure 6.5, middle left) and the CB buoy (bottom left) present a well balanced error with equal deviations at both sides of the ideal fit $k = 1$. Generally we observed how the large errors associated to SWH prediction are due to a slight delays in the prediction causing a large dispersion in the scatter diagrams results.

6.4 Conclusions

In this work we have used a Genetic Algorithm to reconstruct the dynamical model of HIPOCAS SWH in 49 deep water locations. The training process was designed to forecast SWH using the wind intensity and wind direction at each location. GA has been trained using a small amount of the total available data (385000 hourly data). After this process the equation that describes the evolution of the SWH at each point was obtained (e.g. the equation that mimics the dynamical system). The equation is then used to reproduce the rest of the time series.

It was found that the mean 1-hour ahead forecast skill for the 49 locations is above $0.70 \leq r^2 \leq 0.97$ for the whole 44-year period. In addition, the forecasting skill test performed at node HIPOCAS 1320 indicates that the GA in the present configuration can provide satisfactory predictions up to 6-hours ahead. This good agreement between the SWH obtained by the dynamical equation and the corresponding SWH from the database is not surprisingly since both wind data (the forcing) and SWH (the response) were used in the GA learning process. To test the validity of the ob-

tained equations, we have used three locations with wave records from deep water buoys. Applying the corresponding equation obtained with the HIPOCAS database to the measured wind at the buoy and comparing with the measured SWH we get a forecasting skill $0.87 \leq r^2 \leq 0.91$.

The good performance of the presented approach suggests that both, numerical models and GA, can be combined to increase forecasting skills. Although the GA approach—as other predictive approaches based on statistical fitting of time series such as Neural Networks or Genetic Programming—may not provide insight on the physical connection between wind and wave response, they can quickly convey us with reliable information at low computational cost for operational purposes.

Chapter 7

General Conclusions

This Thesis, analyzes the Western Mediterranean wave climate have. It has been shown how the Mediterranean area is affected by a strong seasonality both in the wind and wave regimes, generating a temporal pattern of variability that evolves from the high winter SWH in the Northern areas of the Balearic Islands to extreme minima in SWH fields along the Southern coasts of the Islands during summer months. This pattern clearly corresponds to the "normal" conditions observed in the mid-latitude climatologies, such as the Western Europe climatology, where intense northerly winds affect the area during the winter months while milder conditions are found during the summer months.

The first spatial characterization for the extreme wave height in the Western Mediterranean has been carried out taking advantage of the HIPOCAS database. This database has been shown to be an adequate tool for climate studies in areas like the Mediterranean. The REMO data was an important improvement, allowing global reanalysis data being downscaled to a regional scale. This improvement in the spatial resolution of the atmospheric model resulted in a better estimation of the wind forcing allowing a higher spatial and temporal resolution in the SWH fields analyzed. Hereby, this Thesis presents a high spatial resolution in the estimations obtained for the extreme SWH patterns along the Balearic Sea.

An additional improvement presented is the use of the Pareto-Poisson distribution function. This method allows to gather an extremal population up to 3 times larger than the population obtained from other "classic" methodologies. Thus, the uncertainties associated to any statistical estimation method have been reduced thanks to the larger number of data available for the analysis. We want also to remark that this is the first wave climate characterization that is able to show the spatial distribution of the extreme SWH in the Western Mediterranean area.

Extreme wave heights are directly related to the extreme wind conditions in the

boundary layer between the ocean and the atmosphere. The behavior and variability of this boundary layer will determine the energy transfer to the ocean surface. It has been observed that the distant patterns of variability affecting pressure patterns over the Atlantic sea affects also the wind patterns in the Western Europe regions. This effects are extended to the pluviometry, wind regimes, temperatures, and other important geophysical variables. This Thesis have shown how these distant patterns of variability specifically the North Atlantic Oscillation affects the SWH mean and extreme values instantly at monthly time scales. When the NAO is in its positive phase a reinforcement of the northerly cold and dry air masses from the arctic regions appear, generating more severe wave climate conditions over the Northwestern Mediterranean Sea. On the contrary the negative phase of the NAO exhibits a weakening of the west-erlies all across the North Atlantic bringing moist air into the Mediterranean and cold air to northern Europe. The correlation obtained for the long term wave climate parameters with the NAO phenomena indicated that some of the variability of wave height is associated with the NAO in the Northwestern Mediterranean Sea.

However a strong seasonal signal was found in both parameters, monthly mean SWH and the 95th percentile monthly mean SWH, as expected from previous studies of the cyclogenetic activity of the Mediterranean. Both analyzed parameters provided good results for the analysis. Despite this, we obtained that the 95th percentile of SWH provides more dynamical information (patterns of variability in higher order EOFs) than the monthly mean SWH, indicating that this is an important parameter to study hidden patterns of variability.

Nowadays, one of the relevant issues in oceanography and engineering studies is related to the prediction of future conditions for wave climate parameters. This thesis explores SWH prediction through the use of GA, a low computational prediction technique that "learns" from past values of wind and waves observations. This method has been shown to be a useful alternative to the highly computer-intensive methods implemented in research centers to predict wave climate. The presented method provided an excellent performance of predicted vs. observed SWH even under extreme conditions, like high wind intensities or highly variable wind directions. This fact suggests that both, numerical models and GA, can be combined to increase forecasting skills at low operational costs.

Chapter 8

Recommendations for future work

Prediction of ocean surface waves requires both the knowledge on the present state of the ocean (initial conditions) as well as the evolution of the forcing parameters to adequately integrate the numerical models. Besides, numerical models demand large computational resources since they need to be integrated over large domains, during long time periods or both conditions at a time. Nevertheless, to date these models provide the most complete overview of the ocean surface waves.

Since many operational or engineering activities are based on statistical/spectral wave parameters some initiatives around the world have been developed in order to build large databases of oceanic parameters including water wave parameters. These databases provide a unique opportunity to test the possibility of integrating soft computing techniques in operational wave prediction systems in many different coastal and oceanic areas. Soft computing techniques have the benefit of being a low cost system that can be implemented operationally to make predictions in places where the required technological infrastructure is not available. Moreover, the presented approach is able to provide wave forecast of large areas in a low computational time.

The introduction of nearshore wave models (e.g. SWAN - Simulating Waves Nearshore) in the wave climate analysis field has to be explored for a better understanding of the generation and propagation processes of extreme SWH in the Balearic Islands. These models allow not only to analyze or simulate wave generation processes but also are the best available tools for wave prediction in coastal areas. The use of genetic algorithms to obtain shallow water conditions from deep water wave models or buoys is an approach that opens new possibilities. In this sense, an operational system can be developed to predict wave propagation linking numerical models, buoy data and genetic algorithms. This tool could be very useful for search and rescue

teams in the Balearic Islands and also for management of coastal areas.

The climatic change is noticeable at different spatial and temporal scales and all over different climatic patterns such as the North Atlantic Oscillation and the El Niño-Southern Oscillation (ENSO). The study of the effects of the NAO and ENSO on the variability of the Mean Wave Energy Flux in the Western Mediterranean is a line that deserves the scientific attention. This study can contribute to the understanding of the morphodynamic variability observed in different coastal areas of the Mediterranean Sea. The Mean Wave Energy Flux is a key feature to describe and predict the evolution of the beach planform. With this study a correlation between the the NAO index and a Mean Wave Energy Flux Index could be found, indicating a relationship between the morphodynamic states of a beach and the NAO positive and negative phases. This study would give also the opportunity of predicting future morphodynamic states through the prediction of the NAO index and the ENSO index.

Bibliography

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Institute of Statistical Mathematics. 19, 716–723.
- Allender, J., Barnett, T., Bertotti, L., Bruinsma, J., Cardone, V., L. Cavaleri, J. E., Golding, B., Greenwood, A., Guddal, J., Günther, H., Hasselmann, K., Hasselmann, S., Joseph, P., Kawai, S., Komen, G., Lawson, L., Linne, H., Long, R., M. Lybanon, E. M., Rosenthal, W., Toba, Y., Uji, T., de Voogt, W., 1985. Sea wave modeling project (swamp) , an intercomparison study of wind wave predictions models, part 1 : principal results and conclusions. Tech. rep.
- Álvarez, A., 2003. Performance of satellite-based ocean forecasting (soft) systems: A study in the adriatic sea. *Journal of Atmospheric and Oceanic Technology* 20, 717–729.
- Alvarez, A., Orfila, A., Tintore, J., May 2001. DARWIN: An evolutionary program for nonlinear modeling of chaotic time series. *Computer Physics Communications* 136, 334–349.
- Arena, F., Puca, S., 2004. The reconstruction of significant wave height time series by using a neural network approach. *ASME Journal of Offshore Mechanics and Arctic Engineering* 126 (3), 213–219.
- Astraldi, M., Balopoulos, S., Candela, J., Font, J., Gacic, M., Gasparini, G., Manca, B., Theocaris, A., Tintore, J., 1999. The role of straits and channels in understanding the characteristics of mediterranean circulation. *Progress in Oceanography* 44, 65–108.
- Banner, M., Peregrine, D., 1993. Wave breaking in deep water. *Annual Review of Fluid Mechanics* 22, 373–344.
- Basu, S., Sarkar, A., Satheesan, K., Kishitawal, C. M., Sep. 2005. Predicting wave heights in the north Indian Ocean using genetic algorithm. *Geophysical Research Letters* 32, 17608.
- Campins, J., Genoves, A., Jansa, A., Guijarro, J., Ramis, C., 2000. A catalogue and a classification of surface cyclones for the western mediterranean. *International Journal of Climatology* 20, 969–984.

- Canellas, B., Orfila, A., Méndez, F. J., Menéndez, M., Gómez-Pujol, L., Tintoré, J., 2007. Application of a pot model to estimate the extreme significant wave height levels around the balearic sea (western mediterranean). *Journal of Coastal Research* 50, 329–333.
- Castillo, E., Hadi, A., 1997. Fitting the generalised pareto distribution to data. *Journal of the American Statistical Association* 92, 1609–1620.
- Coles, S., 2001. An introduction to statistical modeling of extreme values. Springer-Verlag, University of Bristol, 227.
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28, 1015–1031.
- EPPE, 1991. Maritime Works Recommendations.ROM 0.3-91. Puertos del Estado, Ministerio de Fomento.
- Fisher, R. A., Tippett, L. H. C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Vol. 24. p. 180.
- Flamant, C., Pelon, J., Hauser, D., Quentin, C., Drennan, W. M., Gohin, F., Chapron, B., Gouarrion, J., 2003. Analysis of surface wind and roughness length evolution with fetch using a combination of airborne lidar and radar measurements. *Journal of Geophysical Research* 108, 8058, 26 PP.
- Gaur, S., Deo, M., 2008. Real-time wave forecasting using genetic programming. *Ocean Engineering* 35 (11-12), 1166 – 1172.
- Gelci, R., Cazale, H., Vassal, J., 1957. Prediction of waves, the method of frequency-directional spectral densities. *Bull. Inform. Comite Central Oceanogr. d'Etude Côtes* 9, 416–435.
- Gnedenko, B., 1943. On the limiting distribution of random data. *Annals of Mathematics*.
- Gómez, M., Carretero, J., 1994. Implementation of a high resolution third generation model on the atlantic coast of spain. *PIANC Bulletin*, Belgique.
- Gómez-Pujol, L., Orfila, A., Canellas, B., Alvarez-Ellacuria, A., Méndez, F., Medina, R., Tintoré, J., 2007. Morphodynamical classification of sandy beaches in a microtidal, low energy marine environment. *Marine Geology* 242, 235–246.
- Gumbel, E., 1958. *Statistics of extremes*. Columbia University Press, New York, 375.
- Günther, H., Lionello, P., Janssen, P., Bertotti, L., Brning, C., Carretero, J., Cavaleri, L., Guillaume, A., Hansen, B., Hasselman, S., Hasselman, K., de las Heras, M., Hollingsworth, A., Holt, M., Lefevre, J., Portz, R., 1992. Implementation of a third generation ocean wave model at the european centre for medium-range weather forecasts. Tech. rep., ECMWF, UK.
-

- Güven, A., 2009. Linear genetic programming for time-series modeling of daily flow rate. *Earth System Science* 118 (2), 137–146.
- Hartmann, D., 2008. Matrix methods: Eof and svd. ATM notes, University of Washington, Department of atmospheric Sciences 552, 57–98.
- Hasselmann, S., Hasselmann, K., Janssen, P., Komen, G., Bertotti, L., Lionello, P., Guillaume, A., Cardone, V., Greenwood, J., Reistad, M., Zambresky, L., Ewing, J., 1988. The wam model - a third generation ocean wave prediction model. *Journal of Physical Oceanography* 18, 1775–1810.
- Hosking, J., Wallis, J., 1987. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* 29, 339–349.
- Hurrell, J., 1995. Decadal trends in the north atlantic oscillation: Regional temperatures and precipitation. *Science* 269, 676–679.
- Hurrell, J., Kushnir, Y., Visbeck, M., 2001. The north atlantic oscillation. *Science* 291, 603–605.
- Jacob, D., Podzun, R., 1997. Sensitivity studies with the regional climate model remo. *Meteorology and Atmospheric Physics* 63 (1-2).
- Jaspers, N., 1956. Statistical distribution patterns of ocean waves and of wave induced stresses and motions with engineering applications. *Transactions-The Society of Naval Architects and Marine Engineers* 64, 375–432.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Joseph, K., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The ncep/ncar 40-years reanalysis project. *Bulletin of the American Meteorological Society* 77, 437–471.
- Katz, R., 1988. Use of cross correlations in the search for teleconnections. *Journal of Climate* 8, 241–253.
- Katz, R., Parlange, M., Naveau, P., 2002. Statistics of extremes in hydrology. *Advances in Water Resources* 25, 1287–1304.
- Kinsman, B., 1965. Wind waves, their generation and propagation on the ocean surface. Englewood Cliffs: Prentice-Hall, 676.
- Konyaev, K., 1990. Spectral analysis of physical oceanographic data. Vol. 81. Russian Translation Series, Rotterdam.
- Koza, J., 1992. Genetic programming: On the programming of computers by means of natural selection. The MIT Press, Cambridge, MA.
-

- Krogstad, H., Barstow, S., 1999. Satellite wave measurements for coastal applications. *Coastal Engineering* 37, 283–307.
- Kushnir, Y., Cardone, V., Greenwood, J., Cane, M., 1997. The recent increase in north atlantic wave heights. *Journal of Climate* 10 (8), 21072113.
- Lionello, P., Galati, M. B., Jun. 2008. Links of the significant wave height distribution in the mediterranean sea with the northern hemisphere teleconnection patterns. *Advances in Geosciences* 17, 13–18.
- Lionello, P., Malanotte-Rizzoli, P., Boscolo, R., Luterbacher, J., 2004. The mediterranean climate: Basic issues and perspectives. *White Paper on Mediterranean Climate Variability and Predictability*, 5–17.
- Lionello, P., Sanna, A., Nov. 2005. Mediterranean wave climate variability and its links with nao and indian monsoon. *Climate Dynamics* 25, 611–623.
- Longuet-Higgins, M., 1963. The effect of nonlinearities on statistical distributions in the theory of sea waves. *Journal of Fluid Mechanics* 17, 459–480.
- Massel, S., 1996. *Ocean surface waves: Their physics and prediction*. World Scientific, Singapore, 508.
- Méndez, F., Menéndez, M., Luceno, A., Losada, I., 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model. *Journal of Geophysical Research* 111 (C7), 13 PP.
- Miles, J., 1957. On the generation of surface waves by shear flows. *Journal of Fluid Mechanics* 3, 185–204.
- Orfila, A., Álvarez, A., Tintoré, J., Jordi, A., Basterretxea, G., Aug. 2005. Climate teleconnections at monthly time scales in the ligurian sea inferred from satellite data. *Progress in Oceanography* 66, 157–170.
- Phillips, O., 1957. On the generation of waves by turbulent wind. *Journal of Fluid Mechanics* 2, 417–445.
- Pickands, J., 1975. Statistical inference using extreme order statistics. *Annals of Statistics* 3, 119–131.
- Picornell, M., Jansá, A., Genovés, A., Campins, J., 2001. Automated database of mesocyclones from the hirlam(inm)-0.5 analyses in the western mediterranean. *International Journal of Climatology* 21, 335–354.
- Preisendorfer, R., 1988. *Principal component analysis in meteorology and oceanography*. Elsevier, Amsterdam.
- Queffelec, P., 2005. Wave height variability over the mediterranean sea, using altimeter data. In: *Proceedings of the 5th International Symposium on Ocean Wave Measurements and Analysis (Madrid, Spain)*. No. 21, CD-ROM.
-

- Queffeuou, P., Bentamy, A., 2007. Analysis of Wave Height Variability Using Altimeter Measurements: Application to the Mediterranean Sea. *Journal of Atmospheric and Oceanic Technology* 24, 2078.
- Ratsimandresy, A., Sotillo, M., Albiach, J. C., lvarez Fanjul, E., Hajji, H., 2008. A 44-year high-resolution ocean and atmospheric hindcast for the mediterranean basin developed within the hipocas project. *Coastal Engineering* 55 (11), 827 – 842.
- Rodwell, M. J., Rowell, D. P., Folland, C. K., Mar. 1999. Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature* 398, 320–323.
- Rogers, A., Jeffrey, C., 1997. North atlantic storm track variability and its association to the north atlantic oscillation and climate variability of northern europe. *Journal of Climate* 10 (7), 1635–1647.
- Smith, R., 2003. *Statistics of extremes, with applications in environment, insurance and finance*. University of North Carolina, Department of Statistics.
- Soares, C., 2008. Hindcast of dynamic processes of the ocean and coastal areas of europe. *Coastal Engineering* 55, 825–826.
- Sotillo, M., Ratsimandresy, A., Carretero, J., Bentamy, A., Valero, F., González-Rouco, F., 2005. A high-resolution 44-year atmospheric hindcast for the mediterranean basin: contribution to the regional improvement of global reanalysis. *Climate Dynamics* 25, 219–236.
- Sverdrup, H., Munk, W., 1947. *Wind, sea and swell: theory of relations for forecasting*. U.S Navy Hydrographic Office.
- Takens, F., 1981. Detecting strange attractors in turbulence, lecture notes in mathematics. *Lecture Notes in Mathematics* 898, 366.
- Thompson, E., 2002. *Coastal engineering manual, part 2*. U.S. Army Engineer Research and Development Center.
- Tomás, A., Méndez, F., Losada, I., 2008. A method for spatial calibration of wave hindcast data bases. *Continental Shelf Research* 28 (3), 391–398.
- Trigo, I. F., Bigg, G. R., Davies, T. D., 2002. Climatology of cyclogenesis mechanisms in the mediterranean. *Monthly Weather Review* 130, 549.
- Trigo, R., Osborn, T., Corte-Real, J., 2002. The north atlantic oscillation influence on europe: climate impacts and associated physical mechanisms. *Climate Research* 20, 9–17.
- Ustoorikar, K., Deo, M., 2008. Filling up gaps in wave data with genetic programming. *Marine Structures* 21 (2-3), 177 – 195.
-

- Vignudelli, S., Gasparini, G. P., Astraldi, M., Schiano, M. E., Mar. 1999. A possible influence of the north atlantic oscillation on the circulation of the western mediterranean sea. *Journal of Geophysical Research* 26, 623–626.
- Wittman, P., Clancy, M., 2004. Thirty years of operational ocean wave forecasting at fleet numerical meteorology and oceanography center. In: *Symposium on the 50th Anniversary of Operational Numerical Weather Prediction*, University of Maryland.
- Woolf, D. K., Challenor, P. G., Cotton, P. D., Oct. 2002. Variability and predictability of the north atlantic wave climate. *Journal of Geophysical Research (Oceans)* 107, 3145.
-

Publications

Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean)

B. Cañellas[†], A. Orfila[†], F. J. Méndez[‡], M. Menéndez[‡], Gómez-Pujol. L. and J. Tintoré[†]

[†]IMEDEA (CSIC – UIB)
07190, Illes Balears, Spain

[‡] Ocean & Coastal Research Group, IH Cantabria
Universidad de Cantabria
39005, Santander, Spain



ABSTRACT

CAÑELLAS, B., ORFILA, A., MÉNDEZ, F.J., MENÉNDEZ, M., GÓMEZ-PUJOL, L. and TINTORÉ, J., 2007. Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). Journal of Coastal Research, SI 50 (Proceedings of the 9th International Coastal Symposium), 329 – 333. Gold Coast, Australia, ISSN 0749.0208

Extreme value wave climate analysis at a particular site requires predicting long-term wave height levels from short duration records. In the present work we used the Peak Over Threshold (POT) model, assuming the frequency as a Poisson process and the intensity to be Pareto distributed, to characterise the spatial variability of the long-term extreme value wave climate along the Balearic Sea. Wave data used is part of the HIPOCAS database, a 44-years high resolution, spatial and temporal, wave hindcast covering an area between 38°N-42°N and 1°W-6°E, of the western Mediterranean Sea. The use of data from a homogeneous grid, instead of a single location wave data record, allows describing the spatial variability of the long-term extreme wave height levels over the whole Balearic basin. Results show that extreme values for a 50-year return period level around 11 m are found in the north sector of the Balearic Islands while in the southern part much lower extreme values are found, due to the shadow effect of the islands under severe north-eastern storms.

ADDITIONAL INDEX WORDS: *extreme wave climate, HIPOCAS, spatial variability.*

INTRODUCTION

Detailed assessment of wave climate is a previous requirement for all human activities in the coastal zone. Beach nourishment, port design and operability, dispersion and diffusion of pollutants are some examples that require a precise knowledge of the long-term distribution of significant wave height, H_s , and mean period, T_m , as well as the long-term extreme value distribution. On the other hand, wave climate analysis requires a large amount of data to ensure the statistical significance. These data have been collected in the last decades using scalar and directional wave buoys moored at specific locations providing high temporal resolution records. In the last decade, satellites have been used to overcome the spatial lack of data (KROGSTAD AND BARSTOW, 1999) but the problem of having a large amount of spatial and temporal wave records were still unresolved.

QUEFFEULOU (2005) used altimeter data to perform an analysis of the wave height variability over the Mediterranean Sea but, as noted for some authors, altimeter data has, as a shortcoming, its temporal inhomogeneity and a coarse spatial resolution in areas like the Western Mediterranean, marked by a complex orography.

Statistical analysis of wave climate has been thus carried out with relatively short data sets, e.g. 10 years of data for the satellite altimeters (TOPEX, ERS-1/2) used in the analysis carried out by QUEFFEULOU (2005).

Alternatively, wave generation models are another option to avoid the usual lack of data in ocean and atmospheric studies. Models are initialised with real conditions and the deviation due to the nonlinearity of the governing equations corrected with the assimilation of data. Numerical models can be now

implemented in very fine grids. These hindcast models have become a powerful tool not only for engineering or prediction scales but for climate studies involving large temporal periods. The 44 years of hourly wave data base with 0.125° spatial resolution obtained from HIPOCAS project (SOARES *et al.*, 2002) in the Western Mediterranean is used in the present work.

In order to characterise the long-term extreme value distribution of significant wave height in the Balearic Sea, we use the Peak Over Threshold (POT) method, which is widely used in the definition of the extreme behaviour of severe storms. The spatial variability of the extreme wave events is obtained determining the 50-year return period quantile in every node.

The paper is structured as follows. In the first section, we present the wave data as well as the POT model. The next section deals with the estimation of the extreme value return levels. The 50-year return period levels over the Balearic Sea is presented and discussed in the third section. Finally we conclude the work.

DATA AND METHODOLOGY

The HIPOCAS Data

Wave data is part of the HIPOCAS Project (Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe). This database consists on a high resolution, spatial and temporal, long-term hindcasted data set (SOARES *et al.*, 2002). This re-analysis covers, on an hourly basis, a period ranging from 1958 to 2001 providing 44-years of wave data over an

homogeneous grid. HIPOCAS dataset was produced by means of dynamical downscaling from the NCEP/NCAR global re-analysis using the regional atmospheric model REMO. Hourly wind fields from the REMO (U_{10}) were used as forcing for the third generation wave model WAM. As a result, wave data used are the output of the WAM model implemented in a $1/8^\circ$ resolution mesh over the Western Mediterranean Sea. In this work we cover 1387 nodes in an area between 38°N - 42°N and 1°W - 6°E (Figure 1). This dataset provides the opportunity to perform a significant analysis of return period levels at deep waters and their spatial distribution, which can help to understand the severity of the storms in this particular area.

The Generalised Extreme Value and Generalised Pareto Distribution

The classical approach to perform an extreme value analysis is to fit the annual maxima values with the Generalised Extreme Value (GEV) cumulative distribution function,

$$G(x) = \exp\left\{-\left(1 + \xi \frac{x - \mu}{\psi}\right)_+^{-1/\xi}\right\} \tag{1}$$

where μ is the location parameter, $\psi > 0$ is the so-called scale parameter and ξ is a shape parameter which determines the tail of the distribution. When $\xi \approx 0$ the GEV distribution corresponds to the Gumbel family, conversely for $\xi > 0$ the Fréchet form is adopted and for $\xi < 0$ the Weibull form is adopted. The annual maxima method developed by GUMBEL (1958) considers only the largest value for each year. There is some criticism in the use of this approach because using only the maximum value per year leads to the loss of information contained in other large-sample values for a given period (CASTILLO AND HADI, 1997).

To solve the problem of working only with a data per year the Generalised Pareto Distribution (GPD) was introduced by PICKANDS (1975). The GPD method models all values larger than a given threshold u . The differences between these values and the threshold u are called exceedances over the threshold and it is assumed to follow a GPD (σ, ξ) distribution whose Cumulative Distribution Function is defined by,

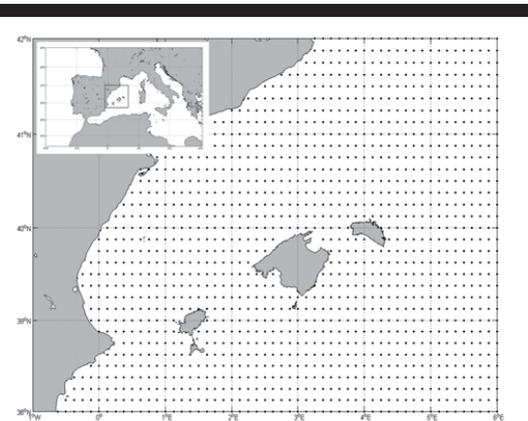


Figure 1. Geographic location of the study area and the $1/8^\circ$ resolution HIPOCAS grid over the Balearic Sea.

$$G(y; \sigma, \xi) = \begin{cases} 1 - (1 + \xi y / \sigma)^{-1/\xi}; \xi \neq 0, \sigma > 0 \\ 1 - \exp(-y / \sigma); \xi = 0, \sigma > 0 \end{cases} \tag{2}$$

where $\sigma > 0$ is the location parameter, $-\infty < \xi < \infty$ is the shape parameter and y are the exceedances over the threshold u ($y = x - u > 0$).

The Poisson-GPD Model

A modification to the model defined in Eq.(2) is the Poisson-GPD model for exceedances, originally developed by hydrologists, which is closely related to the Peaks Over Threshold (POT) method. This model is a joint distribution, the GPD, for the exceedances values y , and a Poisson distribution for the number of exceedances over a level u in any given year. With this model, one can estimate not only the intensity of the exceedances but also the frequency of these events.

Therefore we assume that the number, N , of exceedances of the level u in any one year has a Poisson distribution with mean λ , and the exceedances $\{y_i\}_{i=1}^N$ are independent and identically distributed from the GPD.

Under these hypothesis, the probability that the annual maximum of the GPD-P process is lower than a value x , with $x > u$, is given by,

$$F(x) = \exp\left\{-\lambda \left(1 + \xi \frac{x - u}{\sigma}\right)_+^{-1/\xi}\right\} \tag{3}$$

where the Poisson parameter λ , the scale parameter σ and the shape parameter, ξ are to be determined.

Parameter Estimation

The GPD-P model reduces to the determination of the three unknown parameters, $\lambda > 0$, $\sigma > 0$ and $-\infty < \xi < \infty$. The scale and shape parameters arise from the GPD (PICKANDS, 1975) and λ from the Poisson distribution. These three parameters are estimated using the Maximum Likelihood Method (MLM). The Maximum Likelihood Estimators are the values of the unknown parameters that maximise the log-likelihood function. In practise these are local maxima found by nonlinear optimisation.

The log-likelihood function for the GPD-P, if N exceedances are observed over a τ -year period is given by,

$$l(y; \lambda, \sigma, \xi) = N \log \lambda - \lambda \tau + \log \left(\frac{1}{\sigma^N}\right) + \log \prod_{i=1}^N \left(1 + \xi \frac{y_i}{\sigma}\right)^{-\frac{1-\xi}{\xi}} \tag{4}$$

Maximising $l(y; \theta_k)$ respect to $\theta = (\lambda, \sigma, \xi)$ in the GPD-P leads to the maximum likelihood estimate $\hat{\theta} = (\hat{\lambda}, \hat{\sigma}, \hat{\xi})$. On the other hand a useful relation between the GEV and GPD-P parameters is found in literature (SMITH, 2003),

$$\sigma = \psi + \xi(u - \mu) \quad \lambda = \left(1 + \xi \frac{u - \mu}{\psi}\right)^{-1/\xi} \tag{5}$$

Then through a simple parameter substitution (eq.5 into eq.4) we could express the model in terms of the GEV parameters (μ, ψ, ξ) and, consequently, fit the GPD-P model.

Threshold Selection and Time Span for the GPD-P

An important issue when modelling threshold excesses with the GPD-P is to choose correctly the threshold value u and the minimum time span Δt between successive extreme events. Then the extreme events are identified by considering all values larger than a given threshold u and with a minimum time span Δt between the storms, to ensure the meteorological independence of the observed excesses. It is not an easy aspect because it requires a balance between bias and variance caused by the selected threshold. If the selected threshold is too low we will violate the asymptotic basis of the model, causing bias. On the other hand a too high threshold will produce few excesses over the selected threshold causing a high variance in the estimated values (MENDEZ *et al.*, 2006).

Some tools are available to choose the “correct” threshold. For example an *a priori* test like the mean excess plot (COLES, 2001) can be used (Figure 2). This test leads to a quick estimation of the shape parameter ξ . Assuming that Y follows a GPD distribution, the mean excess over threshold u , is a linear function of u with slope $\xi / (1 - \xi)$.

However, sometimes, the mean excess plot can be difficult to interpret, making the decision subjective (COLES, 2001). As an example, in Figure 2, a linear tendency for a threshold above $u > 5$ metres is observed at the HIPOCAS node 1193, located at Menorca’s island north shore but a more precise value is difficult to obtain.

To avoid the subjectivity in the threshold selection, we use an alternative diagnostic method known as the W-statistic plot (SMITH, 2003). The W-statistic is defined as,

$$W_i = \frac{1}{\xi} \log \left\{ 1 + \frac{\xi Y_i}{\psi + \xi(u - \mu)} \right\} \quad (6)$$

This method is based on, if all assumptions are correct, including the selected threshold u and the time span Δt , then W_i are also independent and exponentially distributed variables with mean 1. Figure 3 shows the quantile plot (QQ-plot) for the W-statistic at the same HIPOCAS node, for the selected

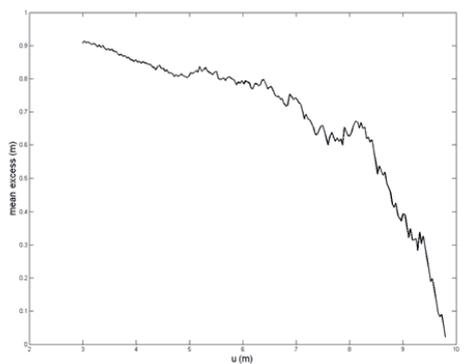


Figure 2. Mean excess plot for different thresholds values.

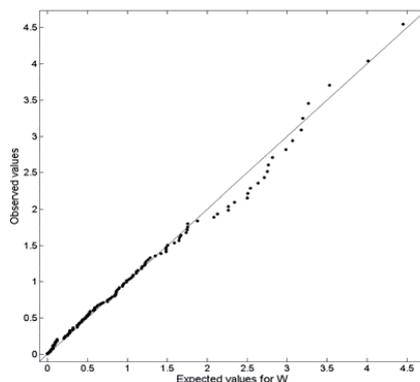


Figure 3. W statistic quantile plot for $u=5.2$ meters and a 72 hours time span at the HIPOCAS node 1193.

threshold $u = 5.2$ metres and $\Delta t = 72$ hours. As seen in this figure, expected values for W are close to the observed ones with a slope near the unit diagonal, indicating the suitability of the selected parameters u and Δt .

The W-statistic was applied at each grid point for different time spans, between 12 and 144 hours. Finally, a time span $\Delta t = 72$ hours was selected for the whole area and a value for the threshold u corresponding to the 99.5% percentile of the empirical distribution was chosen at each grid point.

Model Selection

The selection of the simplest possible model that fits the data sufficiently well is important. Therefore, we check for every point if the contribution of the shape parameter ξ is statistically significant. This is performed using the likelihood ratio test (COLES, 2001). With nested models $M_A \subset M_B$ (M_B including the shape parameter and M_A with the shape parameter $\xi = 0$), we can assure that model M_B explains substantially (at the α -level of significance) more variability in the data than M_A if $2[l_B(M_B) - l_A(M_A)] > \chi^2_{k,1-\alpha}$, where $l_B(M_B)$ and $l_A(M_A)$ are the maximised log-likelihood functions under models M_B and M_A , respectively, and $\chi^2_{k,1-\alpha}$ is the $1-\alpha$ quantile of the χ^2 distribution with k degrees of freedom ($\chi^2_{1,0.05} \approx 3.84$).

Figure 4(a) shows the estimated shape parameter for the whole area and 4(b) shows the statistical significance of the inclusion of the shape parameter. As seen, we can distinguish three different areas, depending on the value of the shape parameter ξ . The dark colour in Figure 4(b) reveals that for this area, it is not significant the inclusion of the shape parameter (which corresponds to the Gumbel family for the GEV distribution or the exponential for the GPD distribution). The western area corresponds with a Weibull tail ($\xi < 0$) and the extreme waves in the area around Spain has a Frechet tail ($\xi > 0$). Therefore, for the area where the inclusion of the shape parameter ξ is not significant we fix the value ($\xi = 0$).

Return Levels for H_{s50}

The N -year return level is the average time interval in years between successive events of an extreme significant wave height being equalled or exceeded. So, the probability that H_s will be exceeded in any given year is,

$$F(H_s \geq x) = 1 - \frac{1}{N} \tag{7}$$

being $F(x)$ the values of a cumulative probability distribution function and N the return period. Therefore a fifty-year return period is equivalent to $F(H_s \geq x) \approx 0.98$.

For the 1387 HIPOCAS grid points the GPD-P case H_{sN} , could be estimated as,

$$H_{sN} = u - \frac{\sigma \lambda^\xi}{\xi \ln(F(H_s \geq x))^\xi} - \frac{\sigma}{\xi} \tag{8}$$

Results for the 50-year return period significant wave height are shown in Figure (5). These values are around 11 metres in the northern quadrant of the islands while in the southern part are less than 8 metres. This is the result of the shadow effect of the islands over the intense north fetch produced by the storms. Extreme wave heights over the Catalan coast are significantly lower than those obtained in the north of the Islands due to the angular spreading of the more severe storms.

DISCUSSION AND CONCLUSION

Deep water wave climate over the Balearic Sea has in general a complex pattern as a result of the complex orography of the surrounding area. The Mediterranean Sea is well known to be one of the most active cyclogenetic area in the world where the climate is mainly conditioned by severe atmospheric forcing during winters. The mountains range in the vicinity is a key factor controlling the storm track. The role of the Pyrenees in the west part and the Alps in the north-east area are decisive boundaries for the wind and pressure distribution over the whole basin. The north western part and central part of the Balearic Sea are forced by northerly winds (mistral) during the main part of the year, while the eastern part is generally modulated by a seasonal variability. Gale forced mistrals often develop over the Gulf of Genova when the passage of the 500

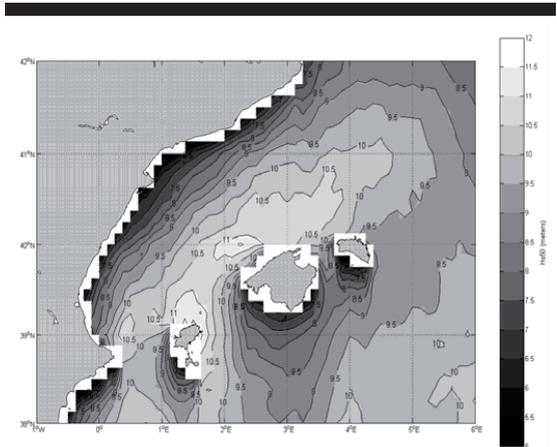


Figure 5. Spatial distribution of H_{s50}

mb thought cross the south eastern part of France extending the effects over the whole basin.

In order to have a rough idea of the behaviour of the storms, the intensity and direction of the maximum significant wave height for the 44 years data are shown in Figure 6. As seen, prevailing directions are from the northeast along the Balearic Sea. This result was previously observed by SOTILLO *et al.* (2006) where a high wind area was identified along the Western Mediterranean from the Gulf of Lions to Northern Algeria and Tunisia. Wind speeds for a 100-year return period shows a maximum located in the Gulf of Lions with levels of winds up to 30m/s. In the eastern part, differences in wave directions are obtained as a result of the different storm track pattern over this area. As seen, maximum significant wave heights are reached in the Balearic Channel as was obtained from the GPD-P distribution.

From the analysis we can conclude that the GPD-P distribution provides a good estimation for wave climate analysis. The analysis performed provided the spatial variability of the 50-year return period significant wave height. The shadow effect of the islands and the angular spreading of

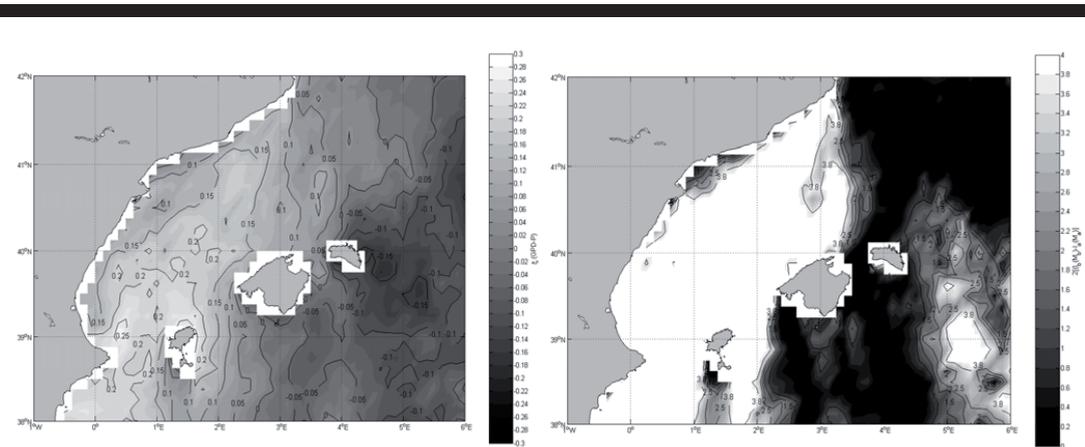


Figure 4. Spatial distribution of (a) shape parameter ξ and (b) statistical significance of the shape parameter ξ

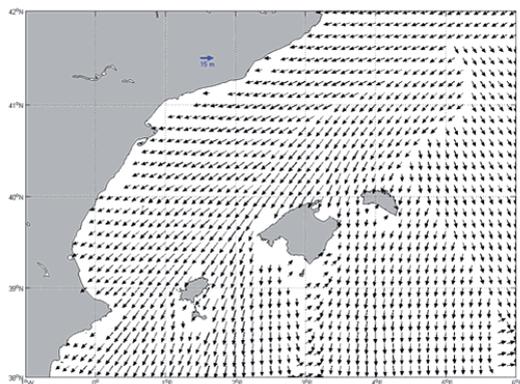


Figure 6. Intensity and direction for the highest value at each point during the 44 years data.

the storms produce a reduction in the magnitude of the higher return levels.

Results also show a spatial variability of the tail of the extreme value distribution: in the western area a bounded tail (Weibull) is detected. This can be associated to a homogeneous distribution in the intensity of the extreme events. On the other hand, along the Spanish littoral, the extreme waves tend to be heavy tail distributed (Fréchet). This aspect can be related on the two main storms that affected this area in December, 1981 and November, 2001. This last storm was driven by heavy Mistral winds (up to 40 m/s) with a large associated fetch, going from the Liguria Sea to the Balearic channel. HIPOCAS nodes located in the area, show maximum significant wave heights over 12 metres at deep waters.

The use of data from the HIPOCAS re-analysis provides a powerful tool for the estimation of the extreme events for a risk analysis in the western Mediterranean Sea.

ACKNOWLEDGEMENTS

The authors thanks to Puertos del Estado (Ministerio de Fomento) for the use of HIPOCAS data base. The authors would also like to thank financial support from Govern Balear Project (UGIZC). AO would like to thank financial aid from CSIC (2006-3-01-068). FJM and MM are indebted to the

Spanish Ministry of Education and Science for the funding provided in the "Ramon y Cajal" program and FPI program, respectively. FJM and MM acknowledge the support of the Spanish Dirección General de Investigación under grants ENE2004-08172 and CGL2005-05365/CLI. BC would like to thank Obdulio Serrano for his kindly support.

LITERATURE CITED

- CASTILLO, E. and HADI, A.S., 1997. Fitting the generalised Pareto distribution to data. *Journal of the American Statistical Association*, 92, 1609-1620.
- COLES, S., 2001. *An introduction to statistical Modelling of Extreme Values*. London: Springer-Verlag, 208 p.
- DAVISON, A.C. and SMITH, R.L., 1990. Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society*, B 52, 393-442.
- GUMBEL, E.J., 1958. *Statistics of Extremes*. New York: Columbia University Press, 375 p.
- KROGSTAD, H.E. and BARSTOW, S., 1999. Satellite wave measurements for coastal applications. *Coastal Engineering*, 37, 283-307.
- MÉNDEZ, F. J.; MENÉNDEZ, M; LUCEÑO, A., and LOSADA, I. J., 2006. Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model. *Journal of Geophysical Research*, 111, C7.
- PICKANDS, J., 1975. Statistical inference using extreme order statistics. *Annals of Statistics*, 3, 119-131.
- QUEFFEULOU, P., 2005. Wave height variability over the Mediterranean Sea, using altimeter data. *Proceedings of the 5th International Symposium on Ocean Wave Measurements and Analysis* (Madrid, Spain), Paper 21, CD-ROM.
- SMITH, R.L., 2003. *Statistics of extremes*, with applications in environment, insurance and finance. Chappel Hill, North Carolina: University of North Carolina, Department of Statistics.
- SOARES, C.G.; CARRETERO, J.C.; WEISSE, R., and ÁLAVAREZ, E., 2002. A 40 years hindcast of wind, sea-level and waves in European waters. *Proceedings of OMAE 2002: 21st International Conference on Offshore Mechanics and Arctic Engineering* (Oslo, Norway), Paper OMAE 2002-28604.
- SOTILLO, M.G.; AZNAR, R., and VALERO, F., 2006. Mediterranean offshore extreme wind analysis from the 44-year HIPOCAS database: different approaches towards the estimation of return periods and levels of extreme values. *Advances in Geosciences*, 7, 275-2

Influence of the NAO on the northwestern Mediterranean wave climate

BARTOMEU CAÑELLAS¹, ALEJANDRO ORFILA¹, FERNANDO MÉNDEZ²,
ALBERTO ÁLVAREZ¹ and JOAQUÍN TINTORÉ¹

¹IMEDEA (CSIC- UIB), 07190 Esporles, Illes Balears, Spain.

²Environmental Hydraulic Institute, IH Cantabria, Universidad de Cantabria, 39005 Santander, Spain.

SUMMARY: This study examines teleconnections between the North Atlantic Oscillation (NAO) and the wave climate of the northwestern Mediterranean Sea (NWM), defined by the monthly mean significant wave height (SWH) and the 95th percentile significant wave height (95th percentile SWH), in the period ranging from 1958 to 2001. The data analyzed comes from the multidecadal hindcast over Europe carried out during the HIPOCAS project. In order to avoid fictitious cross-correlations, data were prewhitened by fitting a p-order autoregressive model. To split the temporal and spatial variability, an EOF encoding technique was applied to residuals before searching for teleconnections. We found the northwestern Mediterranean wave climate to be influenced by the North Atlantic Oscillation (NAO) with an instantaneous response. When the NAO is in its positive phase, positive anomalies in the SWH and the 95th percentile SWH appear in the area between the Balearic Islands, the Gulf of Lions and the Catalanian coast.

Keywords: teleconnections, wave climate, NAO, EOF, Mediterranean Sea, variability.

RESUMEN: INFLUENCIA DE LA NAO EN EL CLIMA MARÍTIMO DEL MEDITERRÁNEO NOROCCIDENTAL. – El presente trabajo estudia las teleconexiones entre la Oscilación del Atlántico Norte (NAO) y el clima marítimo en el Mediterráneo noroccidental, definido por la media mensual de la altura significativa (SWH) y la media mensual de la altura significativa correspondiente al percentil 95 (percentil 95 SWH), en el período comprendido entre 1958 y 2001. Los datos analizados corresponden al retroanálisis de 44 años de datos atmosféricos, dentro del marco del proyecto HIPOCAS. Estos datos han sido codificados en forma de EOFs para obtener la variabilidad espacio-temporal asociada a la NAO. Los datos utilizados han sido previamente blanqueados para evitar correlaciones ficticias entre las series, mediante el ajuste a un modelo autorregresivo de orden p. Los resultados muestran como el clima marítimo del Mediterráneo noroccidental está influenciado a escala mensual por las distintas fases de la NAO. Cuando la NAO está en su fase positiva, podemos observar anomalías positivas en la altura significativa media mensual, así como en la altura de ola correspondiente al percentil 95, estas anomalías aparecen en la zona comprendida entre las Islas Baleares, el Golfo de León y la costa catalana.

Palabras clave: teleconexiones, clima marítimo, NAO, EOF, variabilidad, mar Mediterráneo.

INTRODUCTION

Teleconnections are interactions between widely separated parts of the ocean and the atmosphere occurring at different time scales. These interactions lead to recurring and persistent modes of low-frequency variability in large-scale patterns of atmospheric pressure and ocean circulation anomalies

(Vignudelli *et al.*, 1999; Orfila *et al.*, 2005). These anomalies in the large-scale circulation are manifested in surface waves and drive some of the variability of the wave climate, which has consequences for the operation and safety of shipping, offshore activities, and coastal development, among other things (Kushnir *et al.*, 1997). Therefore, it is important to search for and recognize teleconnections in order to

predict and characterize the wave climate variability over time.

A fraction of the interannual variability in the climate in Europe is associated with a single pattern of pressure anomalies in the north eastern Atlantic area, known as the North Atlantic Oscillation (NAO) (Hurrell, 1995). The NAO has long been identified as a major factor in climate variability and oceanic variability in many areas of the northern hemisphere (e.g. Rogers, 1997; Rodwell *et al.*, 1999; Vignudelli *et al.*, 1999; Hurrell *et al.*, 2001; Lionello and Sanna, 2005; Lionello and Galati, 2008). It can be described as an irregular oscillatory transfer of mass between two main centres of action, the Icelandic Low and the Azores High, that modify the paths of the storms crossing the North Atlantic from the east coast of America to Europe (Woolf *et al.*, 2002).

Wave climate analysis requires a large amount of data to ensure statistical significance (Cañellas *et al.*, 2007). Traditionally, these data have been collected using scalar and directional wave buoys moored at specific locations, which provide high temporal resolution records. In the last decade, satellites have been used to overcome the spatial lack of data (Krogstad and Barstow, 1999) but the problem of having a large amount of spatial and temporal wave records is still unresolved. Queffeuilou (2005, 2007) used altimeter data to analyze the wave height variability over the Mediterranean Sea. However, altimeter data have a shortcoming: they are temporally inhomogeneous and has a coarse spatial resolution in areas like the western Mediterranean, marked by a complex orography. Alternatively, wave generation models are another option to avoid the usual lack of data in ocean and atmospheric studies. Models are initialized with real conditions and the deviation due to the nonlinearity of the governing equations is corrected by data assimilation. Numerical models can now be implemented in very fine grids. Although, hindcasted model data usually underestimate the actual sea wave height (SWH) when compared with buoy observations, satellite data and simulations forced by higher resolution wind fields, they provide a reliable representation of the real space and time variability (Lionello and Sanna, 2005). Therefore, these hindcast models have become a powerful tool not only for engineering and predicting scales but for carrying out climate studies involving large temporal periods.

The Mediterranean Sea is well known to be one of the most active cyclogenetic areas in the world.

The orography of the surrounding area with the Pyrenees in the western part and the Alps in the north-eastern part are natural barriers controlling the generation and track patterns of cyclones over the western Mediterranean area. In this area, the climate is mainly conditioned by severe atmospheric forcing during the winter season, when the polar front jet is more likely to influence Mediterranean cyclogenesis and weather (Trigo *et al.*, 2002a). Most of the strong winds observed in the Mediterranean belong to the category of local winds and are originated as down slope flows by air-flow/mountain interaction, or are due to channelling effects (Buzzi *et al.*, 2005). The northwestern Mediterranean area is forced by northerly and north-westerly winds during most of the year, while less intense cyclogenetic activity is observed during the rest of the year (Trigo *et al.*, 2002b; Cañellas *et al.*, 2007).

This study explores the influence of the monthly NAO index as a general descriptor of the pressure field variability in North Atlantic areas, on the monthly average SWH fields in the northwestern Mediterranean Sea, and the 95th percentile significant wave height, in the period from 1958 to 2001.

44 years of hourly wave data with 0.125° spatial resolution obtained from the HIPOCAS project (Ratsimandresy *et al.*, 2008) in the northwestern Mediterranean Sea (Fig. 1) were analyzed. The underlying assumption is that the wave climate for the northwestern Mediterranean area is influenced by large-scale atmospheric patterns of variability, like the NAO phenomenon, originated in the adjacent Atlantic Ocean. The article is structured as follows: Section 2 describes the data and methodology employed in the study. Results obtained from the EOF analysis are shown in Section 3. Section 4 discusses and concludes the work.

MATERIALS AND METHODS

Wave data is part of the HIPOCAS Project (Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe) (Soares *et al.*, 2008). This database consists in a high resolution (spatial and temporal), long-term hindcasted dataset that covers the period from 1958 to 2001 on an hourly basis. It provides 44 years of wave data over a homogeneous grid. The HIPOCAS dataset was produced by means of dynamical downscaling from the NCEP/NCAR global reanalysis (Kalnay *et al.*, 1996) using the regional

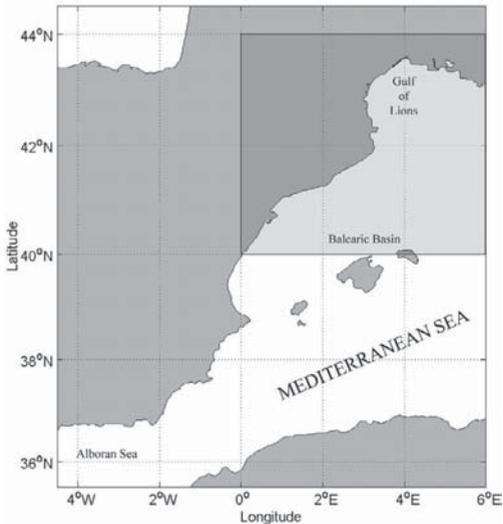


FIG. 1. – Geographic location of the study area.

atmospheric model REMO (Jacob and Podzun, 1997). Hourly wind fields from REMO were used as the forcing for the third generation wave model WAM (see Sotillo *et al.*, 2006, for more details). In the present paper, we use 2021 HIPOCAS nodes in an area between 40°N–44°N and 0°–6°E (Fig. 1), with a spatial resolution of 1/8°.

In the western Mediterranean Sea, SWH data from the HIPOCAS database were extensively calibrated using altimeter and buoy data. The calibration was achieved through the EOF decomposition of the spatial modes and their temporal amplitudes, which were then transformed using a non-linear parameterization to fit the satellite and buoy data adequately (for more details please see, Cavaleri *et al.*, 2005; Tomas *et al.*, 2008).

Monthly NAO index data were provided by the Climate Analysis Section, NCAR (Hurrell, 1995). The NAO index used in this study has been calculated as the difference in the normalized pressure between Ponta Delgada in the Azores and Stykkisholmur in Reykjavik (Iceland). The wave parameters chosen for the correlation with the NAO index are the monthly mean significant wave height (SWH) and the 95th percentile of the SWH.

The monthly mean and standard deviation for the SWH were obtained by fitting hourly time series to a log-normal probability distribution of the entire HIPOCAS database (Jaspers, 1956). There is no theoretical basis for selecting the initial unknown

probability distribution function to represent the observed data. Some other probability distributions are also valid for this purpose: the most common is the bi-parametric Weibull distribution and the Log-Normal distribution. These two distributions are good enough to describe the mean behaviour of the SWH time series. However, we did not use the whole record to compute the mean and standard deviation to avoid a possible “contamination” in the SWH statistics due to the influence of extremal maxima/minima. Then, fitted data correspond only to the SWH between the 5 and 95th percentiles of the cumulative distribution function in order to exclude SWH that correspond to the extremal range of probability. As a result, the mean and standard deviation for each month at each node were obtained from this probability distribution. Moreover, the 95th percentile SWH (which corresponds to those wave heights that are only surpassed during 5% of each year) was obtained using the same theoretical probability function. With this parameter, only the highest SWH are retained, and are a good indicator for the most extreme long term SWH episodes.

Once hourly data are aggregated into the monthly statistical parameters, data are encoded with Empirical Orthogonal Functions (EOFs) to identify the preferred modes of variability (Preisendorfer, 1988). EOF decomposition of the temporal variance describes these spatial patterns with the highest contribution to time variability. However, if the EOF decomposition of the spatial variance is carried out, spatial patterns with the strongest spatial gradients will be obtained (Álvarez, 2003). In this study, we have focused on spatial variance EOF modes since we are interested in the possible influence of the NAO on the variability of the spatial patterns of the wave climate. Time series from the temporal amplitudes associated with these EOFs are used to cross correlate with the monthly NAO index.

Cross-correlation functions strongly depend on the structure of the original time series through auto-correlation, which implies that nonzero values of the cross-correlation function do not necessarily represent a relationship between two time series if they are auto-correlated (Katz, 1988). In order to avoid possible artificial cross correlations, a prewhitening of each time series was carried out. Prewhitening fits a p-order autoregressive model (AR (p)) of the form

$$x(t) = a_1x(t-1) + \dots + a_px(t-p) + \Phi(t) \quad (1)$$

to the original time series, which is sufficient to reduce the residuals $\{\Phi(t)\}$ to white noise. The order of the autoregressive is selected by the Akaike Information Criterion (Akaike, 1974). The Akaike Information Criterion (AIC) determines the model order “p” by minimizing a theoretical information function given by,

$$AIC(p) = N \ln(\sigma_x^2(p)) + 2p \quad (2)$$

where N is the number of samples and σ_x^2 is the estimated variance of the white driving noise and the term 2p in Equation (2) is a “penalty” for the use of extra AR coefficients that do not substantially reduce the prediction error (Konyaev, 1990). Once the value of the filter order “p” is obtained by minimizing Equation (2) the residuals obtained after fitting the AR (p) process in Eq. (1) are not autocorrelated time series.

RESULTS

Spatial Variance EOF modes for monthly mean SWH

The first two EOFs account for 91.03% of the total variance (Fig. 2) and are related to the spatial patterns with the highest contribution to the spatial variance of the significant wave height fields.

For the first EOF, which accounts for 82.79% of the total variance (Fig. 3A), the resulting EOF decomposition shows a strong spatial gradient in the west-east direction, with a maximum anomaly located on the east boundary. Conversely, the second spatial EOF mode, which accounts for 8.25% of the total variance, (Fig. 3B), shows no spatial gradients. This pattern displays an isostatic response of the whole area, modulated throughout the year by the temporal amplitude time series.

The order applied for the prewhitening attending the AIC is $p = 2$ for the NAO index and 12 for the temporal amplitudes of both EOFs. The spatial mean time series, computed for each of the 528 months (not prewhitened), is shown in Figure 4A. Maximum values are reached during the autumn-winter months. For the monthly mean SWH, the autocorrelation function (not shown) is negatively autocorrelated at time lag $\tau = 6$, which indicates strong seasonality.

For the first EOF, the temporal amplitude presents an annual cycle that is highly influenced

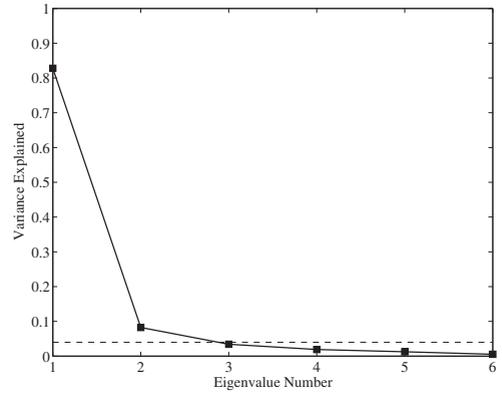


FIG. 2. – Variance explained by the first six EOFs for the monthly SWH. The grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.

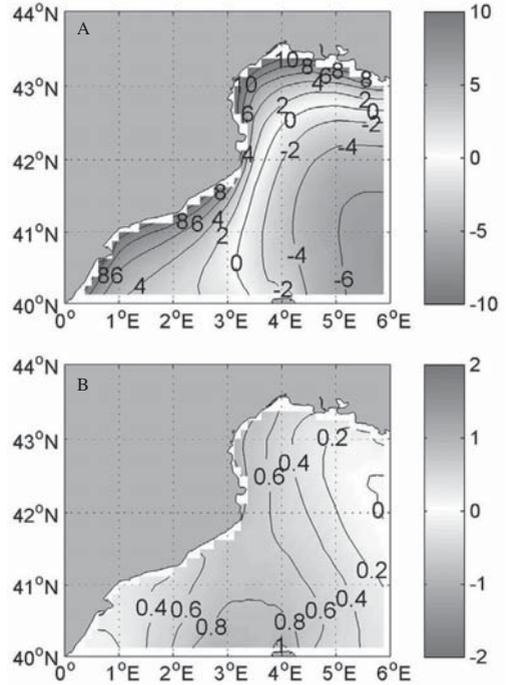


FIG. 3. – A, first spatial variance EOF mode for the monthly mean SWH; B, second spatial variance EOF mode for the monthly mean SWH.

by seasonality (Fig. 4B, C): the autumn-winter months have the highest positive amplitudes (Fig. 4B), while spring-summer months have less intense significant wave height fields in the whole area analyzed.

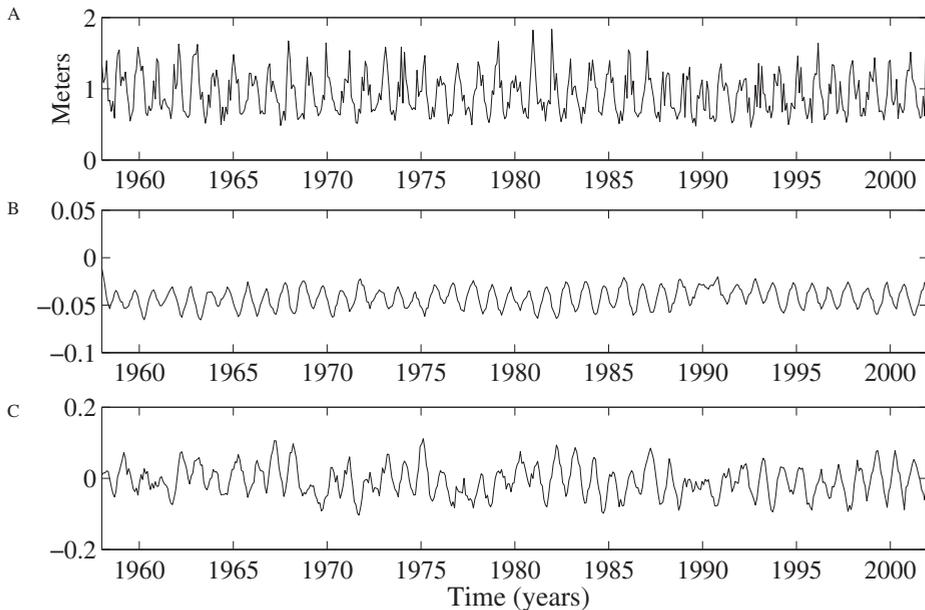


FIG. 4. – (a-top) Time evolution of the spatial mean computed from 528 months for SWH, (b, c) Time amplitudes for the first two EOFs.

Cross-correlation between first temporal amplitude time series and the NAO index provides, as expected, no significant correlation at the 95% level of confidence. The power density spectrum (not shown) for this amplitude was calculated showing a dominant peak at twelve months, which indicates that this mode explains the variability introduced by the annual cycle.

However, a significant correlation was found (at the 95% confidence bounds) between the NAO index and the amplitude of the second EOF. The Maximum correlation is at time lag $\tau = 0$ (Fig. 5), implying an instantaneous response at monthly time scales.

When the time amplitude is at its maximum positive values, the response of this second EOF mode can contribute up to 0.20 m to the mean SWH in the area between the Balearic Islands, the Gulf of Lions and the Catalanian coast.

Spatial Variance EOF modes for the 95th percentile monthly mean SWH

The first four EOF modes were retained for the 95th percentile SWH. These modes explain 92.38% of the total variance (Fig. 6).

Like in the monthly mean SWH results, the 95th percentile SWH variability is mainly re-

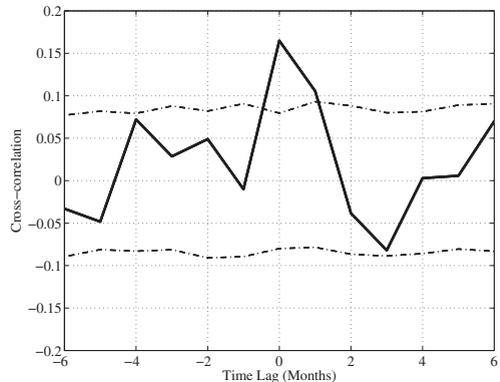


FIG. 5. – Cross correlation between second spatial variance EOF mode time amplitude and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations.

produced by the first mode, which accounts for 72.54% of the total variance. The spatial pattern for this mode (Fig. 7A) shows a longitudinal gradient. For negative amplitudes we can observe negative anomalies along the coast and positive anomalies between the Balearic Islands and the Gulf of Lions. The second EOF mode accounts for 10.87% of the total variance and its spatial pattern displays the same isostatic response as seen

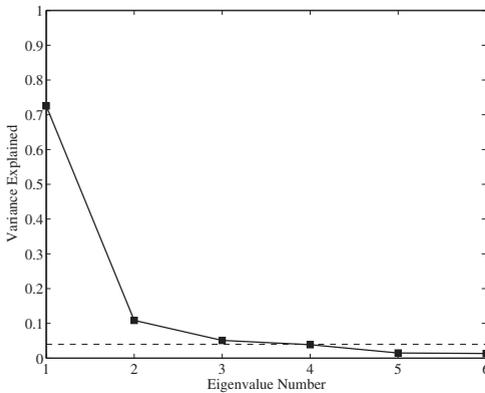


FIG. 6. – Variance explained by the first six EOFs for the monthly 95th percentile SWH. The grey dotted line corresponds to the 0.95 cumulative explained variance. Only EOFs with explained variance above this line were selected for the analysis.

in the monthly mean SWH results (Fig. 7B). The third EOF mode accounts for 5.10% of the total variance and its spatial pattern displays a maximum negative anomaly on the east boundary (Fig. 7C). The location and orientation of this pattern reflects the fetch associated with SW winds. The fourth EOF mode accounts for 3.87% of the total variance and its spatial pattern displays a maxi-

imum anomaly between the continental shelf and the Balearic Islands (Fig. 7D).

Prewhitening was applied to the 95th percentile time amplitude time series. The filtering order “p” is again 12 for the four temporal amplitudes analyzed.

The spatial mean time series for the 95th percentile SWH is shown in Figure 8A. Maximum values are reached during the autumn-winter months. Similarities were found between the monthly mean SWH and the 95th percentile SWH, not only in the spatial mean time evolution, but also in the two first EOF modes. The autocorrelation function (not shown) anti-correlates the mean at lag 6, which indicates negative correlations when values for the summer and winter months are compared. The two wave parameters (monthly mean SWH and the 95th percentile SWH) show identical time evolutions, with maximum and minimum values at the same time lags. However, the main differences between them are found in their intensity, between their spatial patterns as well as in the variance explained by each mode.

There is also strong seasonality in the first temporal amplitude, which shows maximum values during winter, and minimum values during summer. Cross-correlation between the first temporal amplitude for the 95th percentile SWH (Fig. 8B) and the NAO in-

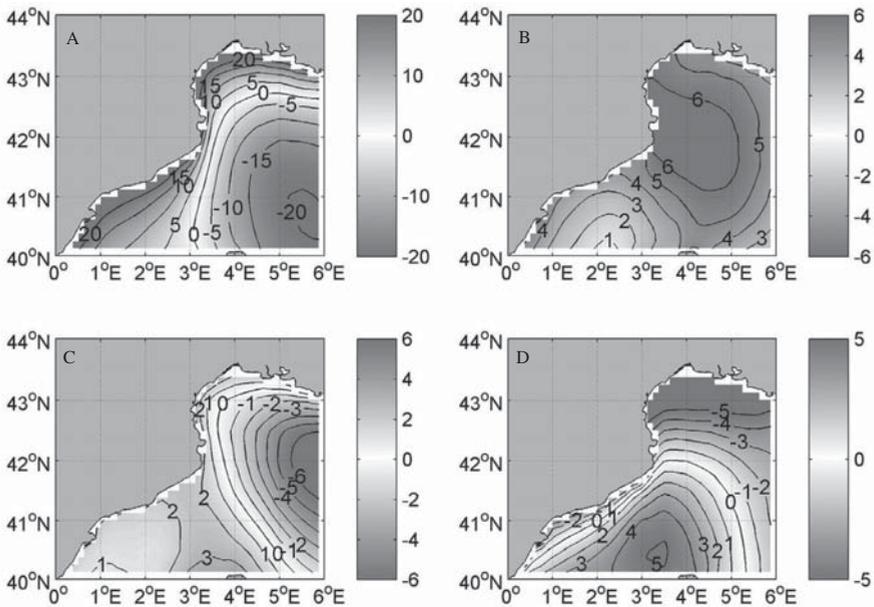


FIG. 7. – A, first spatial variance EOF mode for the 95th percentile SWH; B, second spatial variance EOF mode for the 95th percentile SWH; C, third spatial variance EOF mode for the 95th percentile SWH; and D, fourth spatial variance EOF mode for the 95th percentile SWH.

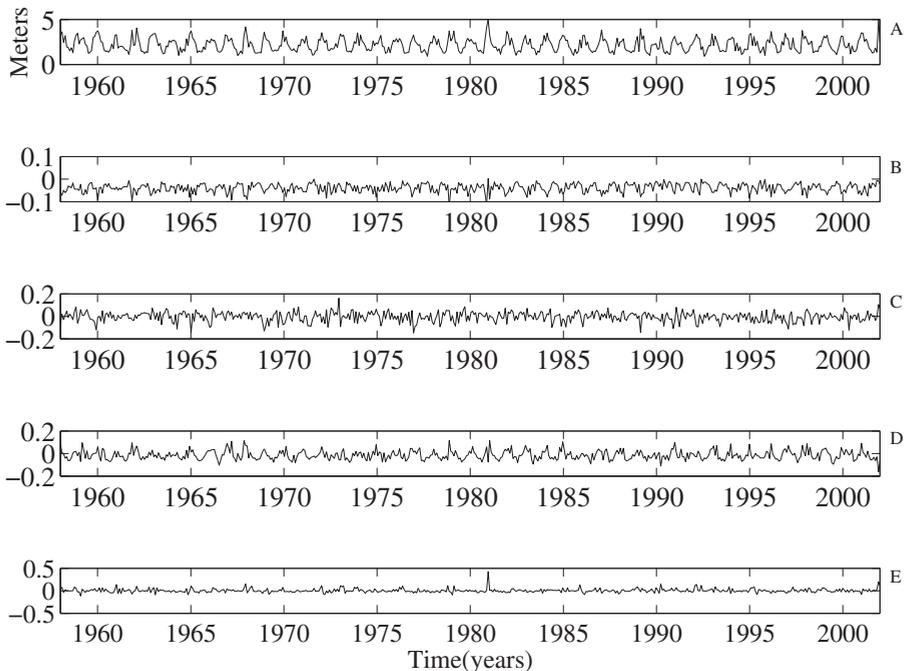


FIG. 8. – Time evolution of the spatial mean computed from 528 months for the 95th percentile SWH (A), 2nd, 3rd and 4th EOF time amplitudes (B, C, D, E).

dex provides, as expected, no significant correlation at the 95% level of confidence (not shown).

However, significant correlations are found in the 2nd, 3rd and 4th temporal amplitudes. The 2nd EOF amplitude (Fig. 8C) is correlated with the NAO index at time lag $\tau = 0$, (Fig. 9A). The maximum anomaly reaches 0.60 m in the whole area when the amplitude is at its maximum.

The time amplitude for the 3rd mode (Fig. 8D) is also correlated with the NAO index at time lag $\tau = 0$ (Fig. 9B). The contribution of this mode indicates that when the NAO index is in its positive phase, a negative spatial anomaly appears in the eastern area, which indicates less intense extreme SWH fields. When the time amplitude is at its maximum values the maximum anomaly reaches 0.60 m in the area corresponding to a fetch from the island of Corsica to the Gulf of Lions.

Finally, the fourth EOF temporal amplitude (Fig. 8E), like previous amplitudes analyzed, shows a significant correlation with the NAO index at time lag $\tau = 0$ (Fig. 9C). The basic difference with other amplitudes is its intensity. This amplitude has a value close to zero during the entire record, which indicates that,

for most of the time, this mode might not contribute to the reconstructed field. However, the maximum amplitude obtained falls into this mode (December 1980) and contributes to the spatial field with 2.50 m in the 95th percentile SWH. This anomaly is constrained to the area between the Catalanian coast and the Balearic Islands.

DISCUSSION

This study shows that the NAO phenomenon contributes to the spatial variability found in the northwestern Mediterranean. When the NAO is in its positive phase there is a reinforcement of the northerly cold and dry air masses from the arctic regions, which generates more severe weather conditions over the northwestern Mediterranean Sea (Hurrell, 1995; Hurrell *et al.*, 2001). However, the negative phase of the NAO exhibits a weakening of the westerlies over the entire North Atlantic into Europe.

The results obtained in this study were computed with one of the several indicators of NAO activity. Other useful indexes are also available in the NCAR

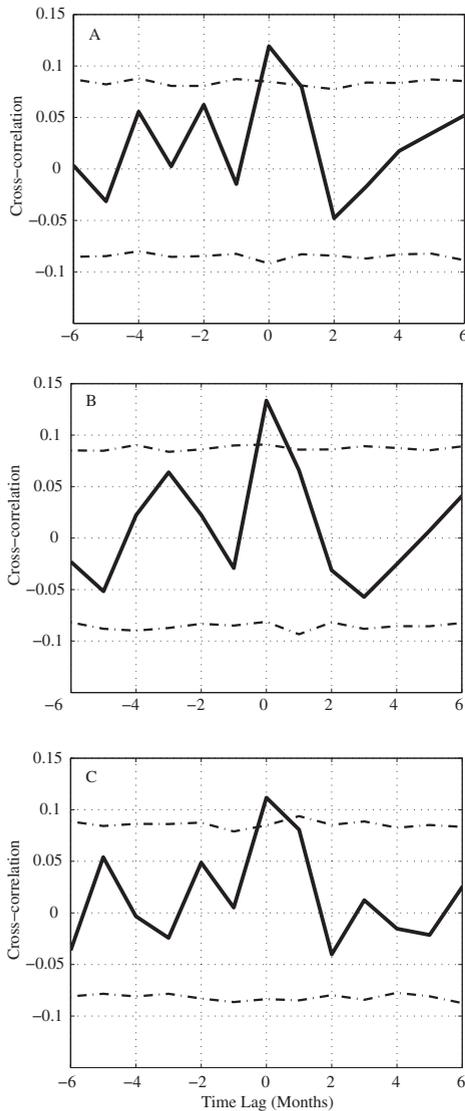


FIG. 9. – (a-top) Cross-correlation between second, third and fourth spatial variance EOF mode time amplitudes and the monthly NAO index time series. Significant positive correlation found at time lag $\tau = 0$. Confidence bounds at 95% obtained through Monte-Carlo simulations.

database. One of them is the DJFM (December to March) NAO index. When this DJFM NAO index was correlated with the December to March aggregated SWH, it returned higher values because it is in this period of the year when the NAO signature is much more noticeable. However, in the December to March aggregation the loss of information is

evident, masking other important key controlling atmospheric patterns, such the intrannual atmospheric/oceanographic cycle and the monthly evolution of the studied wave climate parameters.

The northwestern Mediterranean climatology has high cyclogenetic activity during the autumn-winter months associated with severe weather conditions that drive intense winds from the northeast and northwest (Campins *et al.*, 2000; Picornell *et al.* 2001; Trigo *et al.* 2002b). However, in summer the wind pattern changes, and the direction and intensity is more variable. This is in agreement with the results obtained for the first temporal amplitudes in the two analyzed fields (monthly mean SWH and 95th percentile SWH), with maximum amplitudes located in the winter months. This pattern was already observed and characterized by Lionello and Sanna (2005) as the annual cycle for the monthly mean SWH.

The first mode explains around 70 to 80% of the total variance for both fields. The analysis of the spatial decomposition of monthly SWH statistical parameters, power spectral densities, autocorrelations and cross-correlations implies that the first modes explain only the variability introduced by the summer-winter intrannual cycle. We could not detect any influence of the NAO phenomena in the first mode, which explains most of the intrannual variability.

A correlation between NAO and the second EOF amplitude was found for the spatial variance in monthly mean SWH; Lionello and Sanna (2005) had already observed a significant correlation between the second EOF amplitude for the monthly mean SWH and the NAO index. However, the second, third and fourth EOF amplitudes for the 95th percentile SWH are also significantly correlated. The response of these modes to the NAO was found to be instantaneous at a monthly time scale for both wave parameters.

The two second EOF modes for the monthly mean SWH and the 95th percentile SWH show clear similarities, which reflect an isostatic response of the whole area modulated throughout the data records by the time amplitude. The anomalies for both modes, in the monthly mean SWH and 95th percentile monthly mean SWH, contribute with 0.20 m and 0.60 m respectively.

For the 95th percentile SWH two different patterns are shown for the third and fourth modes. The third mode displays a dipole shaped pattern with a maximum anomaly on the east boundary. The location and orientation of this pattern reflects the fetch associated

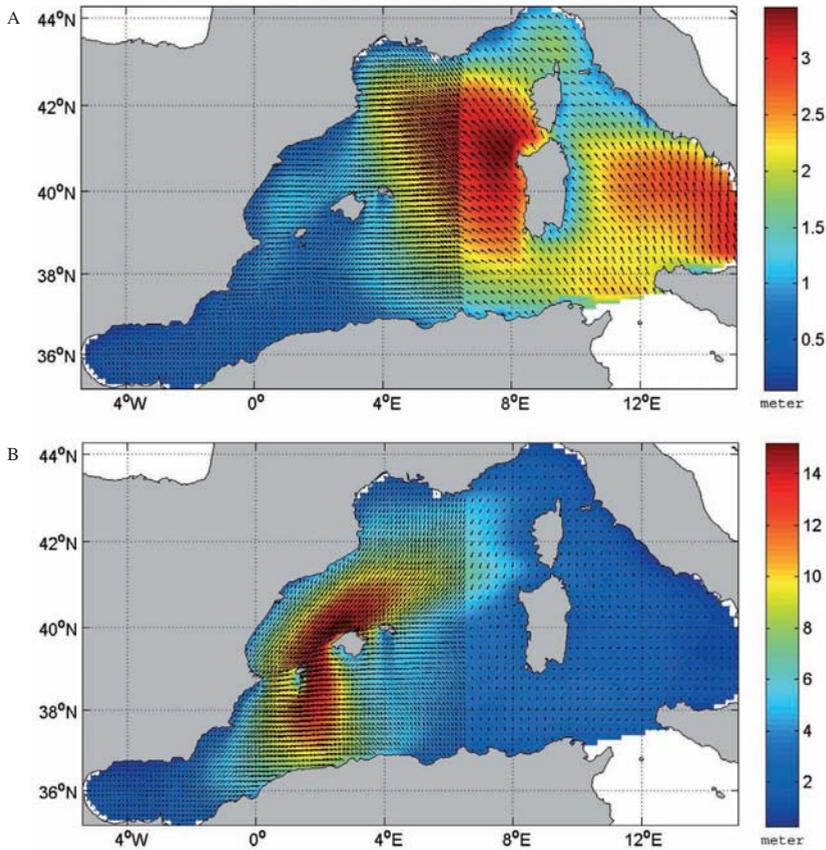


FIG. 10. – A, snapshot from the HIPOCAS database (12/01/2001), which shows the SWH field associated with the south-west winds generated near the island of Corsica and propagated towards the Gulf of Lions. B, snapshot from the HIPOCAS database (12/28/1980), which shows the fetch associated with the severe north-east winds, generated in the Gulf of Genoa and propagated towards the Catalanian coast, Valencia coast and the Balearic Islands.

with the intense south-west winds originating in the island of Corsica and propagated to the Gulf of Lions, which are most common during the wintertime (Fig. 10A). This negative anomaly in the eastern area during the positive phase of the NAO reduces the monthly mean SWH values by up to 0.60 m. The fourth EOF for the 95th percentile SWH mode also displays a dipole shaped pattern with maximum values for the anomaly between the Spanish continental shelf and the Balearic Islands. The location and orientation of this pattern reflects the fetch associated with the intense winds from the north-east, which are often associated with the most severe sea storms recorded in the western Mediterranean. These storms usually cause coastal damage and large economic loss in locations like Mallorca and Valencia.

In December 1980 (as well as November 2001, not shown) a severe storm was recorded in the area between Catalonia and the Balearic Islands, see for instance the amplitude in the 4th EOF for the 95th SWH percentile (see Fig. 8E). In this event, wind waves originated in the vicinity of the Gulf of Genoa with a south-west propagation direction (Fig. 10B). The observed fetch has the same spatial pattern as the fourth EOF mode. This mode suggests that the variability corresponding to these extreme events is reflected in this spatial pattern.

The correlation of long term wave climate parameters with the NAO phenomena indicates that some of the variability of wave height is associated with the NAO in the northwestern Mediterranean Sea. However, a strong seasonal signal was found

in both parameters, monthly mean SWH and the 95th percentile monthly mean SWH, as suggested by previous studies of the cyclogenetic activity of the Mediterranean. This fact has already been noted by Lionello and Sanna (2005).

The two analyzed parameters provided good results for the analysis; however, we found that the 95th percentile SWH provides more dynamic information (patterns of variability in higher order EOFs) than the monthly mean SWH.

ACKNOWLEDGMENTS

The authors would like to thank Puertos del Estado (Ministerio de Fomento) for the use of the HIPOCAS database. The authors appreciate the financial support from the Govern Balear Project (UGIZC). A. Orfila and J. Tintore appreciate the financial support from the MICIN Project CTM2006-12072/MAR. Comments on the manuscript from two anonymous referees are gratefully acknowledged.

REFERENCES

- Álvarez, A. – 2003. Performance of Satellite-Based Ocean Forecasting (SOFT) Systems: A Study in the Adriatic Sea. *J. Atmos. Ocean. Technol.* 20(5): 717-729, doi: 10.1175/1520-0426.
- Akaike, H. – 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.*, 19: 716-723.
- Campins, J., A. Genoves, A. Jansa, J.A. Guijarro and C. Ramis. – 2000. A Catalogue and A Classification of Surface Cyclones for the Western Mediterranean. *Int. J. Climatol.*, 20: 969-984.
- Cañellas, B., A. Orfila, F.J. Mendez, M. Menendez, L. Gomez-Pujol and J. Tintore. – 2007. Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). *J. Coast. Res.*, 50: 329-333.
- Cavaleri, L. and M. Sclavo. – 2006. The calibration of wind and wave model data in the Mediterranean Sea. *Coastal Eng.*, 53(7): 613-627.
- Hurrell, J.W. – 1995. Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation. *Science*, 269(5224): 676-679, doi: 10.1126/science.269.5224.676.
- Hurrell, J.W., Y. Kushnir and M. Visbeck. – 2001. The North Atlantic Oscillation. *Science*, Volume 291 (5504): 603-605.
- Jacob, D. and R. Podzun. – 1997. Sensitivity studies with the regional climate model REMO. *Meteorol. Atmos. Phys.*, 63 (1-2), doi: 119-129, 10.1007/BF01025368.
- Jaspers, N.H. – 1956. Statistical distribution patterns of ocean waves and of wave induced stresses and motions with engineering applications. *Trans. Soc. Nav. Archit. Mar. Eng.*, 64: 375-432.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, R. Jenne and D. Joseph. – 1996. The NCEP/NCAR 40-years Reanalysis Project. *Bull. Am. Meteorol. Soc.*, 77: 437-471.
- Katz, R.W. – 1988. Use of cross correlations in the search for teleconnections. *J. Climate*, 8: 241-253.
- Konyayev, K.V. – 1990. Spectral analysis of physical oceanographic data. In: *Russian Translation Series*. (Vol. 81). A.A. Balkema, Rotterdam.
- Kushnir, Y., V.J. Cardone, J.G. Greenwood and M.A. Cane. – 1997. The recent increase in North Atlantic Wave Heights. *J. Climate*, 10(8): 2107-2113, doi:10.1175/1520-0442.
- Krogstad, H.E. and S. Barstow. – 1999. Satellite wave measurements for coastal applications. *Coastal Eng.*, 37: 283-307, doi:10.1016/S0378-3839(99)00030-7
- Lionello, P. and A. Sanna. – 2005. Mediterranean wave climate variability and its links with NAO and Indian Monsoon. *Clim. Dyn.*, 25(6): 611-623, doi:10.1007/s00382-005-0025-4.
- Lionello, P. and M.B. Galati. – 2008. Links of the significant wave height distribution in the Mediterranean Sea with the Northern Hemisphere teleconnection pattern. *Adv. Geosci.*, 17: 13-18.
- Linonello, P., S. Cogo, M.B. Galati and A. Sanna. – 2008. The Mediterranean surface wave climate inferred from future scenario simulations. *Global Planet. Change*, 63(2-3): 152-162, doi:10.1016/j.gloplacha.2008.03.004.
- Orfila, A., A. Alvarez, J. Tintore, A. Jordi and G. Basterretxea. – 2005. Climate teleconnections at monthly time scales in the Ligurian Sea inferred from satellite data. *Progr. Oceanogr.*, 66(2-4): 157-170, doi:10.1016/j.physletb.2003.10.071.
- Picornell, M.A., A. Jansa, A. Genoves and J. Campins. – 2001. Automated Database from the HIRLAM(INM)-0.5° Analyses in the Western Mediterranean. *Int. J. Climatol.*, 21: 335-354, doi:10.1002/joc.621.
- Preisendorfer, R.W. – 1988. *Principal component analysis in meteorology and oceanography*. Elsevier Science Publishing Company, New York.
- Queffelecoul, P. – 2005. Wave height variability over the Mediterranean Sea, using altimeter data. *Proceedings of the 5th International Symposium on Ocean Wave Measurements and Analysis* (Madrid, Spain), Paper 21. CD-ROM.
- Queffelecoul, P. and A. Bentamy. – 2007. Analysis of wave height variability using altimeter measurements: application to the Mediterranean Sea. *J. Atmos. Ocean. Tech.*, 24(12): 2078-2092, doi:10.1175/2007JTECH0507.1.
- Ratsimandresy, A.W., M.G. Sotillo, J.C. Carretero Albiach, E. Álvarez Fanjul and H. Hajji. – 2008. A 44-year high resolution ocean and atmospheric hindcast for the Mediterranean Basin developed within the HIPOCAS Project. *Coastal Eng.*, 55: 827-842, doi:10.1016/j.coastaleng.2008.02.025.
- Rodwell, M.J., D.P. Rowell and C.K. Folland. – 1999. Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, 398: 320-323.
- Rogers, J.C. – 1997. North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe. *J. Climate*, 10: 1635-1647, doi: 10.1175/1520-0442(1997)010<1635:NASTVA>2.0.CO;2.
- Soares, C.G. – 2008. Hindcast of dynamic processes of the ocean and coastal areas of Europe. *Coastal Eng.*, 55: 825-826, doi:10.1016/j.coastaleng.2008.02.007.
- Sotillo, M.G., R. Aznar and F. Valero. – 2006. Mediterranean offshore extreme wind analysis from the 44-year HIPOCAS database: different approaches towards the estimation of return periods and levels of extreme values. *Adv. Geosci.*, 7: 275-278.
- Tomas, A., F.J. Mendez and I.J. Losada. – 2008. A method for spatial calibration of wave hindcast data bases. *Cont. Shelf Res.*, 28(3): 391-398.
- Trigo, R.M., T.J. Osborn and J.M. Corte-Real. – 2002a. The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Climate Res.*, 20(1): 9-17, doi: 10.3354/cr020009.
- Trigo, I.F., G.R. Bigg and T.D. Davies. – 2002b. Climatology of cyclogenesis mechanisms in the Mediterranean. *Mon. Weather Rev.*, 130(3): 549-569.
- Vignudelli, S., G. P. Gasparini, M. Astraldi and M. Schiano. – 1999. A possible influence of the North Atlantic Oscillation on the circulation of the Western Mediterranean Sea. *Geophys. Res. Lett.*, 26(5): 623-626.
- Woollf, D. K., P.G. Challenor and P.D. Cotton. – 2002. The variability and predictability of North Atlantic wave climate. *J. Geophys. Res.*, 107(C10): 9.1-9.14.

Scient. ed.: J. Font.
Received September 2, 2008. Accepted March 3, 2009.
Published online November 19, 2009.



Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Wave height prediction in the Western Mediterranean using genetic algorithms

B. Cañellas*, S. Balle, J. Tintoré, A. Orfila

IMEDEA (UIB-CSIC), Miquel Marqués, 07190 Esporles, Spain

ARTICLE INFO

Article history:

Received 30 September 2009

Accepted 9 February 2010

Keywords:

Genetic algorithms
Prediction
Significant wave height
Western Mediterranean

ABSTRACT

This study explores the suitability of a soft computing technique to predict (forecast and hindcast) significant wave heights in some areas of the Mediterranean Sea. Using a 44-year database of meteorological and oceanographic records we train a genetic algorithm, to reconstruct a 44-year time series of significant wave height at different coastal points of the Balearic Islands (Western Mediterranean). The correlation obtained between the numerical model data and the algorithm output is $0.80 \leq r^2 \leq 0.97$. The equations obtained by the genetic algorithm are then used to predict significant wave heights measured at several buoys. We make use of the equation from the node closest to the buoy using as input wind data from the meteorological station at each buoy. We obtained that SWH at each buoy can be predicted with a 1-h ahead forecast time using only wind data from the previous 12 h with a correlation of $0.87 \leq r^2 \leq 0.91$.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Prediction of ocean surface waves is an important issue for some of the most relevant human activities in the coastal zone. Beach nourishment, port design and operability, dispersion and diffusion of pollutants are some examples that require a precise knowledge of predicted significant wave height (SWH) fields as well as their extreme values (Sverdrup and Munk, 1947). SWH is an important oceanographic parameter that is required to define a sea state or estimate the wave climate in an area. Traditionally, prediction of ocean surface waves requires the use of numerical models which integrates the fluid governing equations. These models have become a powerful tool not only for engineering or prediction works but also for climate studies involving large temporal periods. However, one of the major drawbacks of the numerical models is the large computational resources required for integrating the primitive equations over fine grids or large areas.

An alternative to prediction based on first physical principles is the use of past observations to predict the evolution of a dynamical system. Takens (1981) provided the theoretical background to build prediction models from past observations of chaotic time series. Specifically, Taken's theorem states that given a deterministic time series of observations $\{x_i\} = \{x(i \cdot \tau) \mid i = 1 \dots M\}$, where τ is the sampling time, there exists a map $P(\cdot)$ satisfying,

$$x_m = P(x_{m-1}, x_{m-2}, \dots, x_{m-d}), \quad (1)$$

being d the embedding dimension.

For instance, Alvarez (2003) used genetic algorithms (GA) to forecast the space–time variability of the Sea Surface Temperature in the Alboran Sea (Western Mediterranean); Basu et al. (2005) predicted significant wave heights in an area of the Arabian Sea, using different time series of surface wind speeds and SWH from three different oceanographic buoys as input for a GA; Gaur and Deo (2008) used an alternative soft computing tool known as genetic programming (GP, Koza, 1992) to explore the usefulness of this method for real-time wave forecasting in a limited area of the Gulf of Mexico, also this method is actually used by Guven (2009) for time series modelling of daily flow rate at a river in US; Ustoorikar and Deo (2008) compared two different empirical prediction methods (GP and artificial neural networks (ANN)) to fill up gaps in wave data from four wave rider buoys maintained by the US National Data Buoy Center in the Gulf of Mexico. The ANN method was also used by Arena and Puca (2004) to perform a reconstruction of SWH records from two correlated buoys from the NOAA in California.

GA are one of the most extended techniques used for empirical prediction. They are based on the evolutionary genetic laws, e.g. reproduction, crossover and mutation where Darwinian theories of natural selection and survival are used to approximate a map that best describes a physical chaotic system.

In this work we apply a multivariable GA to a set of buoy and numerical model wave data to predict directly from past values of wind module and wind direction the SWH in the coastal areas of the Balearic Sea (Western Mediterranean Sea). Moreover, this study explores the possibility to forecast SWH in coastal areas using a GA trained with wind at 10 m from a meteorological model and SWH from a numerical wave model.

Our results demonstrate the possibility not only to reproduce a 44-year time-series from a numerical database, but to

* Corresponding author. Tel.: +34 971 610 508.
E-mail address: tomeu.canellas@gmail.com (B. Cañellas).

also predict SWH directly from wind buoy observations, using the equations obtained previously by training an evolutionary algorithm.

2. Data and methods

Wave and wind data used to train the algorithm are part of the HIPOCAS Project (Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe) (Soares, 2008). This database consists on a high resolution, spatial and temporal, long-term hindcasted data set covering on an hourly basis a period ranging from 1958 to 2001 providing 44-years of wave data over an homogeneous grid. HIPOCAS dataset was produced by means of dynamical downscaling from the NCEP/NCAR global re-analysis using the regional atmospheric model REMO (Kalnay et al., 1996; Jacob and Podzun, 1997). Hourly wind fields from the REMO were used as the forcing for a third generation wave model WAM, see Sotillo et al. (2005), for more details. SWH data used in this work are the output of the WAM model implemented in a 0.125° resolution mesh over the Western Mediterranean Sea. A total of 49 coastal nodes were analyzed, covering an area between 38°N–40.5°N and 1°E–4.5°E, see Fig. 1.

For each HIPOCAS node, the equation that best describes the dynamics of the SWH is obtained by means of a GA (Alvarez et al., 2001), which is trained using hourly mean wind module (U), hourly mean wind direction (θ) and SWH data from the HIPOCAS database.

The algorithm for predicting the state variable $\{x_i\}$ considers an initial population of randomly generated equations $P_j(\dots)$ derived from random combinations of explanatory variables $\vec{V}_i = \vec{V}(i \cdot \tau)$, $i = 1 \dots M$, randomly chosen constants and the four basic operators (+, -, ×, ÷). For a μ -step ahead prediction with embedding dimension d , the fitness of the j -candidate equation map $P_j(\cdot)$ is computed as

$$A_j^2 = \sum_{m=\mu+d}^M \left[x_m - P_j(\vec{V}_{m-\mu}, \dots, \vec{V}_{m-\mu-d+1}) \right]^2 \tag{2}$$

where x_m and \vec{V}_m denote the value of the explained variable and the vector of explanatory variables, respectively, at time $t_m = m\tau$.

The fitness to the data, A_j^2 , establishes the strength of each individual in the selection process. Specifically, the strength index for the j -equation string is expressed as

$$F_j^2 = 1 - \frac{A_j^2}{\sum_{m=\mu+d}^M (x_m - \bar{x})^2} \tag{3}$$

where \bar{x} is the mean value of the training data. F_j can be interpreted as the percentage of the training set's total variance explained by the j -equation string and so, for a given individual, the higher this quantity, the better the data are represented by the corresponding equation.

This population of potential solutions is next subject to an evolutionary process, by which those individuals that best fit the data are selected from the initial population. This process is carried out as follows. First, part of the time series is selected as the training set, i.e. the set of data that will be used in computing the best approximation to the map $P(\cdot)$ in, Eq. (1).

By examining the fitness of $P_j(\cdot)$ the strongest individuals are then selected to exchange parts of the character strings between them (reproduction and crossover), while the individuals less fitted to the data are discarded. Finally, a small percentage of elements in the equation strings are mutated at random, except for the top ranked equation strings in order to avoid losing their information. As a result of this process a new set of equations is obtained. The evolutionary steps are repeated with a new generation and finished after a number of generations determined by the user. Afterwards, a validation is carried out using the data not included in the training set to test the goodness of the candidate equation for reproducing the deterministic signal.

A total of 49 coastal nodes from the HIPOCAS database have been used as the input for the training and validation of the GA. For this particular case the use of GA in a vectorial configuration made possible the inclusion of the wind module and the mean direction as input variables, because usually the evolution of a chaotic natural system is not restricted to the influence of only one observed variable, but a nonlinear interaction among several variables is quite common.

To check that SWH is driven by local meteorological conditions we have performed a multivariate regression analysis of the SWH time series of HIPOCAS against wind parameters (intensity U and

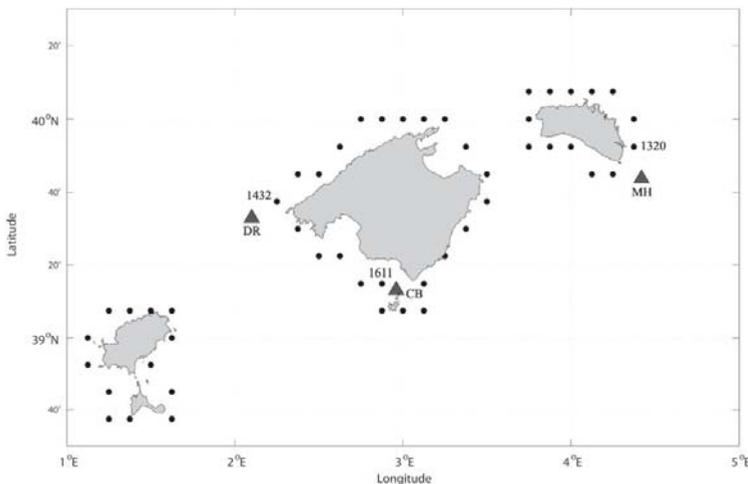


Fig. 1. Geographic location of the study area. Black dots correspond to the HIPOCAS coastal nodes and triangles to the oceanographic buoys, MH (Mahon), DR (Dragonera) and CB (Cabrera).

direction $\bar{\theta}$). In the present case the multivariable GA in its vectorial form can be expressed as

$$SWH_m = P(\vec{V}_{m-\mu}, \dots, \vec{V}_{m-\mu-d+1}) \quad (4)$$

SWH_m being the significant wave height at time t_m and

$$\vec{V}_m = (U_m, \cos\theta_m, \sin\theta_m) \quad (5)$$

is the vector of explanatory variables at time t_m .

Moreover we have used data from three deep ocean buoys in order to demonstrate the capability of the algorithm to retrieve a nonlinear equation that can mimic not only the non-linear interactions between meteorological and oceanographic variables from a numerical model database, but also capable of generating a valid equation for real-time forecasting at offshore stations, using the same exact equation as in the nearest HIPOCAS node.

After the training and validation process, the best ranked equations at three coastal nodes (numbered nodes in Fig. 1) were used at the closest three available deep water buoys to test the ability of these equations to forecast SWH data directly from the buoy wind observations.

3. Results

As a first approximation to the problem of finding a predictive equation that would provide the future evolution of the SWH, a multiple regression analysis of SWH at time t_m was performed using as input the information at time t_{m-1} contained in the meteorological series of the HIPOCAS database. The mean correlation coefficient over all the HIPOCAS nodes is 0.88 with a standard deviation of 0.02, which implies that most of the total variance in the HIPOCAS SWH time series can be explained by the local meteorologic conditions—wind intensity and direction—at the previous time.

Hence we applied the GA in a vectorial configuration to train the algorithm with a multivariate set of time series containing only the module of the wind U and the wind direction $\bar{\theta}$. The GA algorithm requires a large number of parameters to be configured when the underlying dynamical system is unknown. In our case, this configuration required a large trial and error process. Finally the selected parameters were $d=12$ for the embedding dimension and $\tau=1$ for the time lag. The number of individuals in each population has been set to 120 and the number of iterations to 10000.

For the training process the selected total number of points of the time series was 3000 including a validation set fixed at 1000 data points. These parameters were the most efficient ones in terms of the quality of the output and the CPU time needed to reach the desired performance, which was the achievement of optimum fitness strength $F^2 \geq 0.9$.

3.1. Prediction of HIPOCAS data

The cross-correlation coefficient for the 49 HIPOCAS coastal nodes (r^2) calculated using the 44-year time series from the numerical model and the 44-year predicted SWH from the GA is shown in Fig. 2. As seen, the GA is able to reproduce the dynamical behaviour of the hourly SWH time series, with a cross-correlation coefficient that remains between values of 0.70 and 0.97 with a mean correlation calculated over all the nodes of $r^2=0.90$. At this point, we want to remark that less than 1% of the total data were used in the learning process of the GA.

To provide a detailed view of the performance of the GA, SWH at three selected coastal nodes (HIPOCAS 1320, HIPOCAS 1432 and HIPOCAS 1611, see Fig. 1) are compared with the SWH predicted by the GA. Fig. 3 (top left, middle left and bottom left) shows the performance of the GA output displaying one month of data (12/08/1968–01/27/1969) with a 1-h ahead prediction for the nodes 1320, 1432 and 1611. It can be seen the good

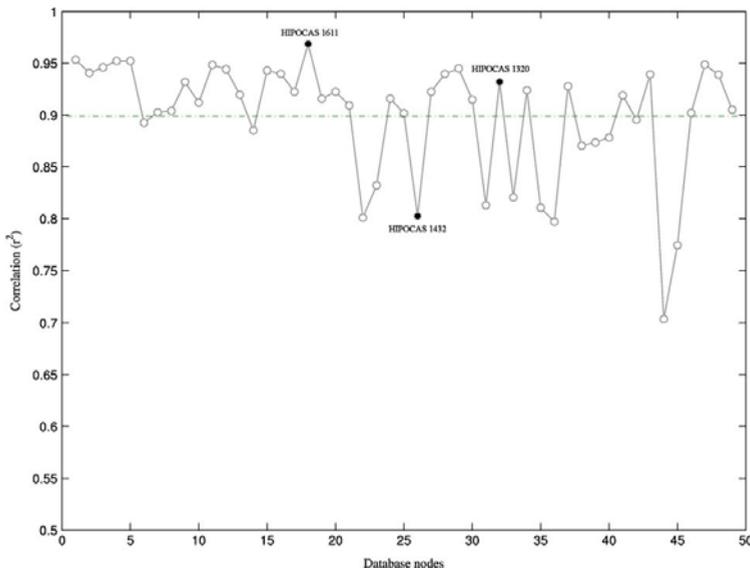


Fig. 2. Linear correlation (r^2) obtained at the 49 HIPOCAS coastal nodes, between the genetic algorithm output and the numerical model data. Dashed line corresponds to the mean correlation calculated over the nodes $r^2=0.90$. Filled dots correspond to the HIPOCAS nodes presented in the Results section, which are the closest to the available buoys.

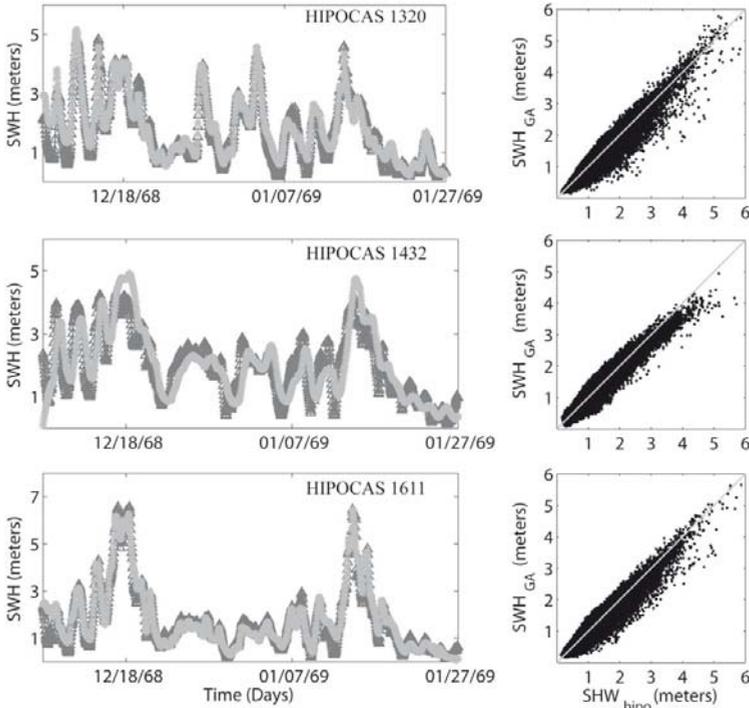


Fig. 3. Left panels: hourly time series comparison, between the genetic algorithm predicted SWH and the HIPOCAS nodes SWH (1320, 1432 and 1611). Black triangles represent the GA predicted SWH, grey asterisks correspond to the HIPOCAS database SWH. Right panels: scatter diagrams comparing GA SWH and the HIPOCAS SWH data, unit diagonal corresponds to a perfect fit.

agreement between the numerical model SWH and the GA predicted SWH. The best ranked equations provided by the GA for these locations are, respectively,

$$x_0(t) = 0.12x_1(t-10) \frac{0.22x_1(t-1)^2}{2.11 + x_3(t-12)x_3(t-5) + x_3(t-2)x_3(t-5)} \quad (6)$$

$$x_0(t) = 0.14 \left(x_1(t-1) + x_1(t-7) - \frac{7.79x_3(t-10)x_3(t-4)}{2.59 + 0.15x_1(t-2)x_3(t-1)x_3(t-4)} \right) \quad (7)$$

$$x_0(t) = \frac{x_1(t-2)}{8.09 - 0.18x_3(t-1) - x_2(t-6) - 0.18x_1(t-10) - 0.18x_2(t-12)} \quad (8)$$

being t the time instant, x_0 the SWH, x_1 the wind module, x_2 the cosinus of the wind direction and x_3 the sinus of the wind direction.

In order to evaluate the errors reported by the GA with the 1-h ahead prediction, scatter diagrams of the SWH from the numerical model database, plotted against the SWH obtained through the nonlinear equation of the GA, are shown in Fig. 3 (top right, middle right and bottom right) for the three nodes. These scatter diagrams provide information about the behaviour of each equation used at the different coastal nodes. It can be seen how in these particular HIPOCAS nodes a general underestimation of the SWH from the numerical model database exist. The larger underestimation arises in the range of SWH between 2 and 5 m, but in general we can see how the errors are equally distributed along the ideal fit of slope $k=1$, except for the largest SWH that present stronger deviations.

The absolute error in SWH estimation related to wind intensity U and direction $\bar{\theta}$, computed as the difference between SWH from the database and predicted SWH from the GA, is shown in Fig. 4. The error (in meters) plotted against the wind module U for HIPOCAS nodes 1320 is displayed in Fig. 4 (top left). We can observe how the dispersion of the error is larger as U increases. In this case the error shows a rotation around $U=5 \text{ m s}^{-1}$. An underestimation error is observed for wind module values under $U=5 \text{ m s}^{-1}$, but for wind module values larger than $U=5 \text{ m s}^{-1}$ the nonlinear equation overestimates SWH from the numerical model database.

The error plot for HIPOCAS node 1432 is shown in Fig. 4 (middle left). This plot shows the same behaviour in the error dispersion for wind speeds over 4 m s^{-1} . In this case the equation obtained is also overestimating SWH at high wind speeds over $U=4 \text{ m s}^{-1}$.

The same error behaviour as in the node 1320 and 1432 is also observed for the HIPOCAS node 1611 (Fig. 4, bottom left), with SWH observations underestimated for wind module values under $U=4 \text{ m s}^{-1}$ and overestimated SWH values when $U > 4 \text{ m s}^{-1}$.

As seen, wind module plays an important role in the predicted SWH and its error. The error by underestimation is generally produced by wind speeds $U < 5 \text{ m s}^{-1}$. On the contrary the error due to the overestimation of the numerical model database values generally arises from wind speeds larger than $U > 5 \text{ m s}^{-1}$. The dispersion in error increases in both directions as the wind module $U \geq 6 \text{ m s}^{-1}$.

The evaluation of the error dispersion related to the wind direction also seems to be important. To show how the error is also dependant on the wind direction $\bar{\theta}$, Fig. 4 (top right, middle

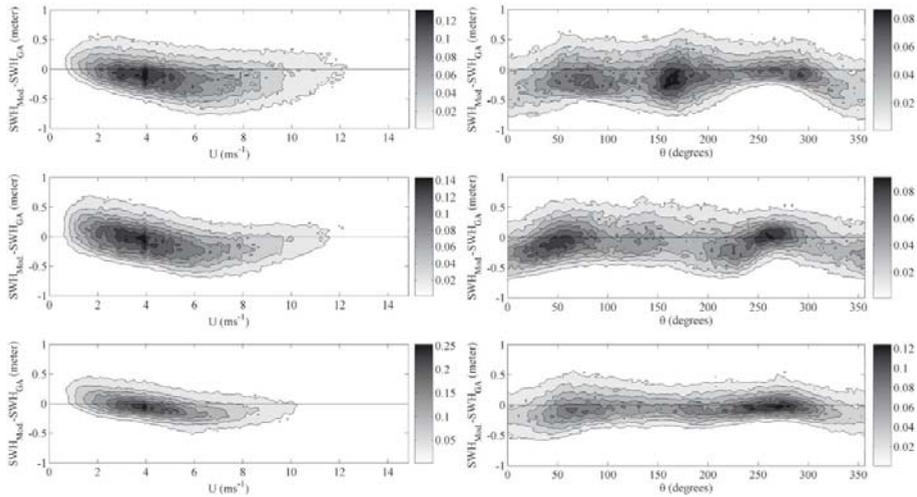


Fig. 4. Left panels: contour plot of the absolute error between HIPOCAS database SWH and predicted SWH at the selected HIPOCAS nodes against the wind intensity U , contour lines indicate the percentage of the population with the indicated error. Right panels: idem but compared against the mean wind direction $\bar{\theta}$.

right and bottom right), shows the error in SWH estimation against the wind direction $\bar{\theta}$. We can observe how the error is equally distributed along the whole range of directions at each node. Also it is noticeable a range of directions where the percentage slightly increases. The errors arise from directions ranging between $30\text{--}100^\circ$, $150\text{--}200^\circ$ and $200\text{--}350^\circ$, where the observed error is given by the overestimation of the numerical model database, except for the middle-right plot (node 1432) in the $250\text{--}300^\circ$ where it is shown a slight underestimation of the numerical model SWH.

Finally, a forecasting skill test was also performed for the HIPOCAS node 1320 to assess the goodness of the method for long-term prediction. This test considered different threshold values in Eq. (2) by modifying the μ -step parameter, hereby allowing for μ -hours ahead predictions. The values used in Eq. (2) for the μ parameter were 3, 6, 9 and 12 (h).

The obtained correlation coefficients (r^2) between the prediction and the original HIPOCAS time series were: 0.88, 0.87, 0.77 and 0.58, respectively. Obviously, the 1-h ahead prediction fits the data with a higher degree of correlation ($r^2 = 0.94$), but it is worth remarking that high values ($r^2 > 0.80$) are still obtained even for a 6-h ahead prediction.

3.2. Prediction of buoy data

In order to test the ability of the obtained equations to predict real wave data, three deep water buoys providing real time wave and wind parameters have been used. Two of these buoys are part of the real time observing system from Ente Publico de Puertos del Estado (EPPE), known as the Spanish deep ocean buoy network (REDEXT), while the third one belongs to the Mediterranean Institute for Advanced Studies (IMEDEA). These three oceanographic buoys were the only available real data in the studied area and the selected HIPOCAS nodes in Fig. 1 are the closest nodes to the buoys.

Two of these buoys, Mahon Buoy (MH) ($39.73^\circ\text{N}\text{--}4.42^\circ\text{E}$) and Dragonera Buoy (DR) ($39.55^\circ\text{N}\text{--}2.10^\circ\text{E}$), belong to the REDEXT network and contain 3 and 10-year of data, respectively. The data consist on hourly meteorological and oceanographic parameters

such as wind module and direction, SWH, wave period and wave direction. Cabrera Buoy (CB) is another deep water meteorological and oceanographic buoy located between Mallorca and the Cabrera Island ($39.22^\circ\text{N}\text{--}2.96^\circ\text{E}$). This buoy belongs to IMEDEA and contains only a year of oceanographic data covering 2008.

The performance of the GA obtained at the REDEXT buoys is shown in Fig. 5. The time period chosen for the reconstruction of the buoy data is constrained to one month for the MH buoy, one week for the DR buoy and 4 months for the CB buoy, in order to avoid the large gaps found in the time series. To test the strength of the GA under extreme maximum and minimum wind and wave conditions, we also selected a time period in the series where the wave climate conditions presented a high deviation from the mean SWH values to include different wave climate conditions with wave heights ranging from 0.30 to 4.0 m, and wind speeds ranging from 0.7 to 15 m s^{-1} .

The prediction time series at the MH buoy was obtained using a month of wind data from the buoy and with the same nonlinear equation found through the analysis of the HIPOCAS node 1320. As shown in Fig. 5 (top left) the hourly forecast represents an accurate reproduction of the original time series. The correlation obtained after analyzing a month of data is $r^2 = 0.89$.

The results for the DR and CB buoys are shown in Fig. 5 (middle left and bottom left). DR buoy and CB buoy were predicted using the equations obtained at the HIPOCAS nodes 1432 and 1611, respectively. The correlation obtained at these buoys was $r^2 = 0.91$ for the DR buoy and $r^2 = 0.87$ for the Cabrera buoy. After reconstruction of the three buoys the mean correlation obtained was $r^2 = 0.89$.

To determine the behaviour of the error from the forecasted SWH, three scatter diagrams are shown in Fig. 5 (top, middle and bottom right). These scatter plots are generated using the entire record available at each buoy. As in the numerical model database prediction we can observe how the error arises principally from the underestimation of buoy data. The range of data where the algorithm provides less accurate predictions is located between $3.0 \leq \text{SWH} \leq 6.0$ meters for the MH buoy. On the contrary scatter diagrams for the DR buoy (Fig. 5, middle left) and the CB buoy (bottom left) present a well balanced error with equal deviations at both sides of the ideal fit $k=1$. Generally we observed how the

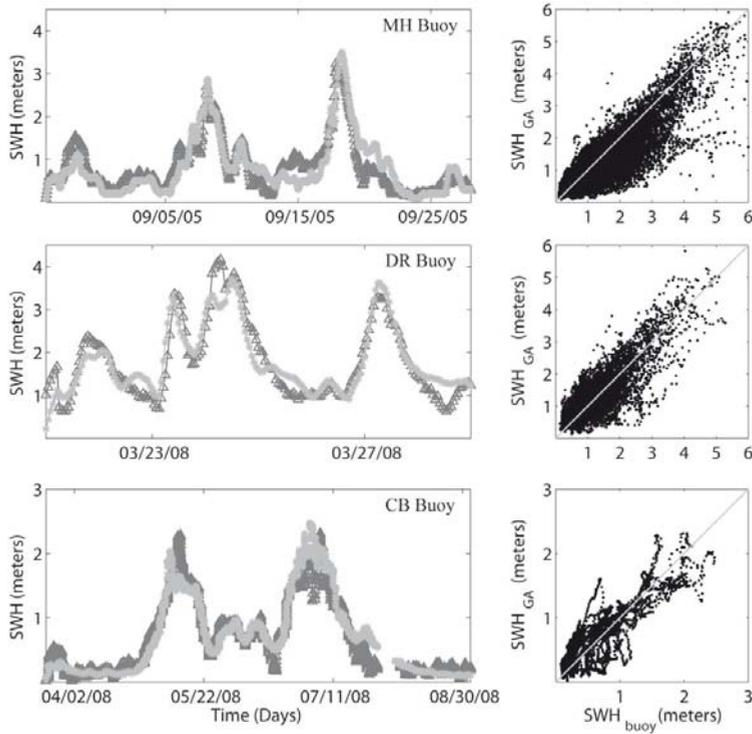


Fig. 5. Left panels: SWH measured at the three buoys (MH, DR and CB) (grey asterisks) compared with the predictive equations SWH (black triangles) computed using the buoy's wind records. Right panels: scatter diagram of these data.

large errors associated to SWH prediction are due to a slight delays in the prediction causing a large dispersion in the scatter diagrams results.

4. Conclusions

In this work we have used a genetic algorithm to reconstruct the dynamical model of HIPOCAS SWH in 49 deep water locations. The training process was designed to forecast SWH using the wind intensity and wind direction at each location. GA has been trained using a small amount of the total available data (385 000 hourly data). After this process the equation that describes the evolution of the SWH at each point was obtained (e.g. the equation that mimics the dynamical system). The equation is then used to reproduce the rest of the time series.

We find that the mean 1-h ahead forecast skill for the 49 locations is above $0.70 \leq r^2 \leq 0.97$ for the whole 44-year period. In addition, the forecasting skill test performed at node HIPOCAS 1320 indicates that the GA in the present configuration can provide satisfactory predictions up to 6-h ahead. This good agreement between the SWH obtained by the dynamical equation and the corresponding SWH from the database is not surprisingly since both wind data (the forcing) and SWH (the response) were used in the GA learning process. To test the validity of the obtained equations, we have used three locations with wave records from deep water buoys. Applying the corresponding equation obtained with the HIPOCAS database to the measured wind at the buoy and comparing with the measured SWH we get a forecasting skill $0.87 \leq r^2 \leq 0.91$.

The good performance of the presented approach suggests that both, numerical models and GA, can be combined to increase forecasting skills. Although the GA approach—as other predictive approaches based on statistical fitting of time series such as neural networks or genetic programming—may not provide insight on the physical connection between wind and wave response, they can quickly convey us with reliable information at low computational cost for operational purposes.

Acknowledgements

The authors acknowledge Puertos del Estado for providing us with the HIPOCAS database. Useful comments from Dr. A. Alvarez, as well as from three anonymous referees are also acknowledged. We also thank Govern de les Illes Balears and MICIN for financial support through Projects UGIZC and CTM2006-12072/MAR.

References

- Alvarez, A., 2003. Performance of satellite-based ocean forecasting (soft) systems: a study in the adriatic sea. *Journal of Atmospheric and Oceanic Technology* 20 (5), 717–729.
- Alvarez, A., Orfila, A., Tintore, J., 2001. DARWIN: an evolutionary program for nonlinear modeling of chaotic time series. *Computer Physics Communications* 136, 334–349.
- Arena, F., Puca, S., 2004. The reconstruction of significant wave height time series by using a neural network approach. *ASME Journal of Offshore Mechanics and Arctic Engineering* 126 (3), 213–219.
- Basu, S., Sarkar, A., Satheesan, K., Kishitawal, C.M., 2005. Predicting wave heights in the north Indian Ocean using genetic algorithm. *Geophysical Research Letters* 32 17608-+.

- Gaur, S., Deo, M., 2008. Real-time wave forecasting using genetic programming. *Ocean Engineering* 35 (11–12), 1166–1172.
- Güven, A., 2009. Linear genetic programming for time-series modeling of daily flow rate. *Earth System Science* 118 (2), 137–146.
- Jacob, D., Podzun, R., 1997. Sensitivity Studies with the Regional Climate Model Remo. *Meteorology and Atmospheric Physics*.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Joseph, K., Joropoulos, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*.
- Koza, J., 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press, Cambridge, MA.
- Soares, C., 2008. Hindcast of dynamic processes of the ocean and coastal areas of Europe. *Coastal Engineering* 55 (11), 825–826.
- Sotillo, M., Ratsimandresy, A., Carretero, J., Bentamy, A., Valero, F., González-Rouco, F., 2005. A high-resolution 44-year atmospheric hindcast for the Mediterranean basin: contribution to the regional improvement of global reanalysis. *Climate Dynamics* 25, 219–236.
- Sverdrup, H., Munk, W., 1947. *Wind, Sea and Swell: Theory of Relations for Forecasting*. U.S. Navy Hydrographic Office.
- Takens, F., 1981. Detecting strange attractors in turbulence. In: *Lecture Notes in Mathematics*, vol. 898, 1981, p. 366.
- Ustoorikar, K., Deo, M., 2008. Filling up gaps in wave data with genetic programming. *Marine Structures* 21 (2–3), 177–195.

Morphodynamic classification of sandy beaches in low energetic marine environment

L. Gómez-Pujol^{a,*}, A. Orfila^a, B. Cañellas^a, A. Alvarez-Ellacuria^a,
F.J. Méndez^b, R. Medina^b, J. Tintoré^a

^a IMEDEA (CSIC-UIB), Miquel Marqués, 21, 07190 Esporles, Spain

^b Ocean & Coastal Research Group, IH Cantabria, Universidad de Cantabria, Avda. Castros s/n 39005 Santander, Spain

Received 3 July 2006; received in revised form 24 February 2007; accepted 3 March 2007

Abstract

Morphodynamic classification of beaches has achieved widespread acceptance in both geological and geomorphological literature. In this sense, the present work classifies twelve Mediterranean low energetic beaches according to the dimensionless fall parameter (Ω) parameter in the Island of Mallorca. Propagation of 44 yr of wave data as well as a detailed sediment study allows to provide probabilities for morphodynamical beach state on an annual and seasonal basis. Consequently, beaches in Mallorca fall between three major categories which are (a) truly reflective, (b) reflective skewed to intermediate and (c) intermediate beaches. The Mallorcan beach position in the morphodynamical scheme is close related to the physiographical and geological framework. Comparison of observed values with those obtained in the analysis leads that for gross beach classification there is agreement between predicted and real state. However on a seasonal classification, mainly during summer, there is no agreement between the predicted state and the real one. As the model does not incorporate the role of summer sea breezes, beach reflective states are highlighted. Real beach configurations correspond to more energetic wave dynamic conditions and to intermediate state scenarios. © 2007 Elsevier B.V. All rights reserved.

Keywords: beach morphology; beach morphodynamics

1. Introduction

The dynamical study of beaches, has adopted the model of a system moving towards a state of dynamic equilibrium under steady forcing conditions (Woodroffe, 2003). According to Wright and Thom (1977), beach morphology relates the mutual adjustment between topography and fluid dynamics. The morphological makeup of beach systems is not accidental

because the arrangement and association of forms occur in an organized contextual space and time (Sonu and van Beek, 1971; Sonu and James, 1973; Lipman and Holman, 2000). Since the classification derived by Wright and Short (1983), obtained from the analysis of the evolution during 6 yr in a number of Australian study sites, beach systems are comprehended in terms of three-dimensional morphodynamic models that include quantitative parameters (wave breaking height, sediment fall velocity, wave period and beach slope) and boundary conditions for definable form-processes association (e.g. presence or absence of bars as well as its type). This has led to the classification of beaches

* Corresponding author. Present address: IMEDEA (CSIC-UIB).
Tel.: +34 971611759; fax: +34 971611761.

E-mail address: lgomez-pujol@uib.es (L. Gómez-Pujol).

into three main categories relating the beach state observations with the physical forcing (Short, 1999): dissipative, intermediate (from the intermediate–dissipative domain to the intermediate–reflective domain) and reflective modes. This classification is quantified by means of a dimensionless fall velocity parameter (abbreviated to DFVP below), which is defined as:

$$\Omega = \frac{H_b}{w_s T} \quad (1)$$

where H_b is the wave breaking height, T is the wave period and w_s is the sediment fall velocity. The DFVP was first proposed by Gouraly (1968) and rewritten by Dean (1973). Values of Ω less than one are associated with reflective states, values between 1 and 6 to intermediate states and values greater than 6 related to dissipative states (Short, 1999).

The DFVP can be seen as a predictive equation that indicates which beach type will occur under certain ranges of waves and grain size parameters assuming that the beach will fully respond to governing parameters which may take days (e.g. associated with storm periods), or to about a year (e.g. modifications of sediment size and type by nourishment projects) (Benedet et al., 2004). Nevertheless, it is not clear whether the DFVP based solely on wave characteristics and sediment size, should really be termed a morphodynamic parameter or just as a dynamic one.

In this way limitations in applying the Wright and Short approach are recognized particularly for intermediate phases prediction. Wright et al. (1987) found only a 36% of agreement between observed and predicted beach states. DFVP fall velocity parameter is useful in discriminating between extreme beach states, but it does not characterize adequately intermediate situations.

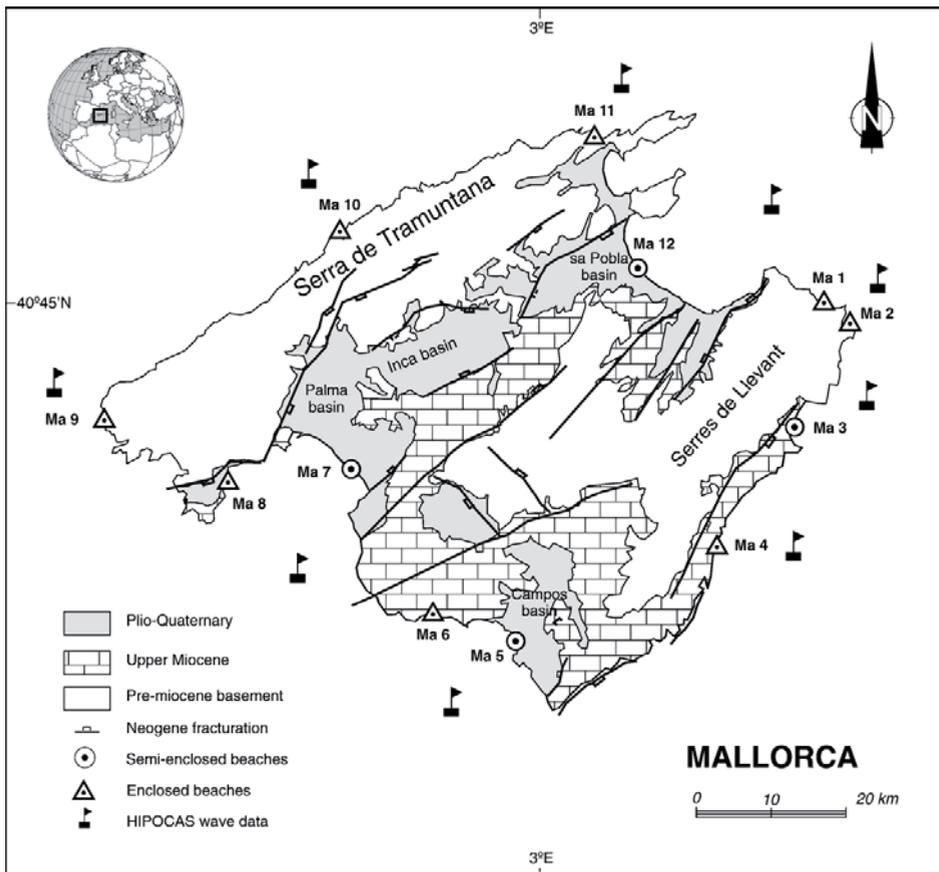


Fig. 1. Geographic location of the study area (a). Position of the HIPOCAS wave data (squares) and beaches studied (circles) (b).

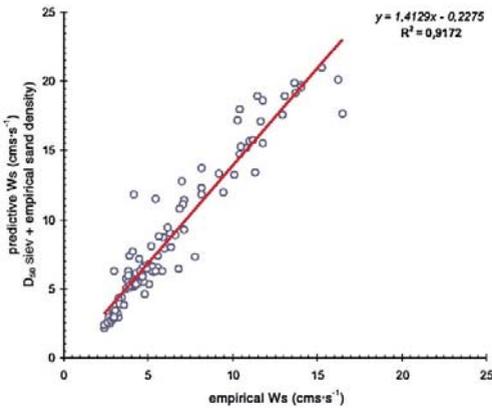


Fig. 2. Agreement between sediment fall velocity computed by empirical methods and predictive Gibb's equation incorporating empirical sand density.

Ranasinghe et al. (2004), justify this fact because there is a lack of accuracy of beach state models and the degree of subjectivity involved in their identification, but most of the temporal variability in DFVP are related to fluctuations in breaking wave height (Anthony, 1998). Further, restrictions on the applicability of this model relates with the considerations of tidal range effects. Despite the additional parameters that incorporate tide-induced migration of hydrodynamic processes across beach profile (Masselink and Short, 1993), the prediction fails in both the higher extreme of energetic parameters at the megatidal beaches (Levoy et al., 2000; Masselink and Hegge, 1995) and in the lower ones, when it concerns to sheltered microtidal beaches (Masselink and Pattiaratchi, 2001; Goodfellow and Stephenson, 2005). In addition, the effect of sea breezes on beach morphology is not considered and appears as a distortion in the DFVP prediction (Masselink and Pattiaratchi, 1998). Sanderson and Eliot (1999) pointed out that, beach state models are not always practical if complications such as the presence of nearshore reefs exist. Geological factors, as underlying geology (bedrock, accommodation space, inheritance,

etc.) and nature and source of beach materials (grain shape, packing, composition, etc.) are factors explaining possible discrepancies between predicted and observed beach states (Jackson et al., 2005; Smith and Cheung, 2002). The Wright and Short model rely largely on dynamic factors, which may be appropriate in wave-dominated linear coasts of Australia but further research is necessary for sheltered beaches because is very difficult to include them in the previous classification (Klein and de Menezes, 2001). Anthony (1998), argues that for a full validation of the DFVP it has to be tested against a wide range of natural environments particularly, within lower energy beach systems with a long time response.

The main goals of this paper are (a) to elucidate a beach morphodynamic sequence and classification for microtidal, low-energy, carbonate sand beaches with headlands and bay geomorphology; and (b) to address the utility of Wright and Short model in this type of environments.

2. Study area

The island of Mallorca, the largest of the group called as Balearics, is located in the western Mediterranean Sea. These islands are the eastern emergent part of the Balearic Promontory; a thickened continental crustal unit forming the NE continuation of the Alpine Betic thrust and fold belt build during Middle Miocene (Gelabert et al., 1992).

Beaches represent 10% of the coasts and are closely related to the basins disposition although they appear sometimes as pocket beaches spread along the cliffs coasts of the island. Mallorca beaches are composed by medium to fine sand with large percentage (more than 70%) of bioclastic sediments most of which derive from organisms associated with the endemic reef-building seagrass *Posidonia oceanica* (Fornós and Ahr, 1997). *P. oceanica* is the dominant seagrass in the Mediterranean Sea, where it covers about 50,000 km² (Bethoux and Cópín-Motegut, 1986) of coastal sandy and occasionally rocky, areas from less than 1 m to about 40 m water depth. The climate is typical from the Mediterranean Sea with hot dry summers and mild wet winters. The annual mean

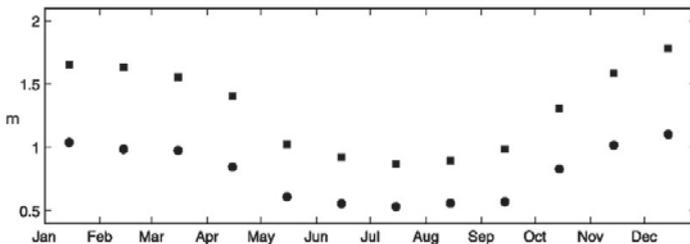


Fig. 3. Standard deviation (circles) and mean (squares) from the log-normal distribution of H_b at the beach (only shown one point).

Table 1
Physiography frame and sediment textural and compositional properties for each study sites

Code	# samples	Mean (μm)	Sorting	Skewness	Kurtosis	Type of sediment	Lithoclasts (%)	Bioclasts (%)	Physiography frame
Ma1	7	574.740	1.568	-0.107	0.976	Coarse sand	10.29	89.71	Enclosed
Ma2	10	355.950	1.584	0.083	1.133	Medium sands	0.88	99.13	Enclosed
Ma3	15	241.484	0.744	-0.116	1.018	Medium/fine sands	1.53	98.47	Semi-enclosed
Ma4	4	454.025	1.557	0.029	0.948	Medium sands	0.76	99.24	Enclosed and sheltered
Ma5	15	215.780	1.619	0.118	1.197	Fine sands	1.64	98.35	Semi-enclosed
Ma6	4	374.375	1.814	0.040	0.966	Medium/fine sands	1.31	98.69	Enclosed and sheltered
Ma7	20	196.984	1.614	0.090	1.029	Fine-very fine sands	4.32	95.68	Semi-enclosed
Ma8	12	242.592	1.648	0.193	0.965	Fine sands	4.47	95.53	Enclosed
Ma9	8	207.225	1.677	0.118	1.228	Fine sands	11.76	88.24	Semi-enclosed
Ma10	8	298.225	1.803	0.196	1.073	Medium-fine sands	28.06	71.94	Enclosed and sheltered
Ma11	4	481.650	1.546	0.112	0.748	Medium sands	3.57	96.43	Enclosed
Ma12	15	247.50	1.66	0.15	0.88	Fine sands	1.70	98.30	Semi-enclosed

temperature is approximately 17 °C with mean winter and summer temperatures around 10° and 25 °C respectively. The mean annual precipitation is about 500 mm and is mostly concentrated in autumn.

The Western Mediterranean presents a temperate, oligotrophic, clear sea environment. Wave heights rarely exceed 8 m with typical wavelength less than 50 m. These values are considerably reduced nearshore where the maximum height is about 4 m and usually recorded when winds between 6–8 Beaufort scale blow. The prevailing winds during the year are mainly from the north-west and are weak or moderate, sometimes rising to gale force during winter. During autumn–winter, the south-western direction is also important, with variable fetches. Tides are almost negligible in the Mediterranean with a spring tidal range of less than 0.25 m, although changes in atmospheric pressure and wind stress can account for a considerable portion of sea level fluctuations. These

physical settings conform coastal areas in the Balearic Islands as low-energy systems where significant morphological changes are restricted to severe weather episodes or long moderate events when wave related processes enhance sediment dynamics controlled by wave-storm events (Basterretxea et al., 2004).

3. Data and methods

3.1. Sediment characterization

A total of 122 sand samples were collected at 12 sandy beaches (Fig. 1). Samples were taken at several cross-shore elevations in transects perpendicular to the shore at locations with different morphological features (e.g. beach face, surf domain, troughs or bars). Samples were rinsed with fresh water, dried 24 h in the oven at 95 °C and divided into sub-samples for sieving and settling analysis.

Table 2
Waves and sediment parameters for the Wright and Short Ω calculations

Code	Study site	Summer (April–September)		Winter (October–March)		D_{50} (μm)	w_s (cm/s)	ρ (g/cm^3)
		H_b (m)	T_p (s)	H_b (m)	T_p (s)			
Ma1	Cala Mesquida	0.93	4.73	0.81	4.36	600.590	20.47	2.87
Ma2	Cala Agulla	0.27	4.48	0.54	4.84	346.000	11.80	2.78
Ma3	Cala Millor	0.21	4.72	0.42	7.03	272.307	8.75	2.79
Ma4	Estany d'en Mas	0.60	4.84	1.00	5.33	450.425	15.60	2.78
Ma5	Es Trenc	0.18	4.00	0.75	4.12	207.840	6.78	2.67
Ma6	Cala Pi	0.24	4.36	0.48	5.93	352.925	12.21	2.87
Ma7	S'Arenal	0.27	3.75	0.90	4.12	199.737	6.18	2.78
Ma8	Magalluf	0.33	3.75	0.39	3.75	231.858	7.74	2.81
Ma9	Sant Elm	1.00	4.12	1.09	4.36	200.813	6.53	2.81
Ma10	Platja d'en Repic	0.15	4.36	0.18	4.48	271.888	9.11	2.68
Ma11	Cala Molins	0.54	4.00	0.57	4.24	465.125	16.14	2.88
Ma12	Es Comú de Muro	0.12	4.48	0.57	5.69	239.475	8.06	2.74

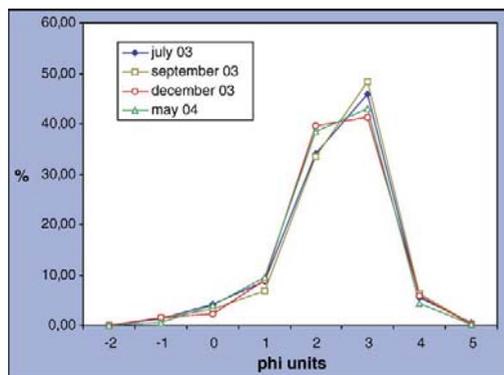


Fig. 4. Seasonal sediment size distribution at Cala Millor (Ma3).

Dry sieve analysis was performed using a series of sieves ranging in mesh size from 0.063 mm to 4.76 mm. Grain size distributions were determined using the GRADISTAT© package (Blott and Pye, 2001).

For each fraction a textural analysis was performed. Using a binocular microscope, 100 grain counts of loose particles distributed on a microscope slide were made for each settling fraction giving a total of 700 counts per sieved sub-sample. Fourteen component classes were identified (foraminifera, gastropods, bivalves, rock fragments, indeterminate grains, etc.). The percentage weight of components of individual fractions were summed and expressed as a percentage of the total sub-sample (Kench, 1997).

The density of sand samples were measured from the volumetric displacement of 75 g of dry sand in a graduated cylinder containing 50 ml of water at 20 °C (Smith and Cheung, 2003). The mean settling velocity (w_s) of some samples were determined using a settling tower, 2 m long and with internal diameter of 0.11 m with a differential pressure sensor (Giró and Maldonado, 1985). The correlation between sieve size and mean settling velocity was calculated and used to estimate the mean settling velocity for those samples that were not assessed in the settling tower. We have found a good agreement with the sediment velocity values predicted by Gibbs equation (Gibbs et al., 1971) although using D_{50} sieve size and empirical sand density rather than quartz (Fig. 2).

3.2. Nearshore wave propagation

To characterize wave climate at the beach, 44 yr of hourly data were analyzed. This database is part of the HIPOCAS project (Soares et al., 2002) where a wave hindcast of the Mediterranean Sea was carried out in a high resolution mesh for the period between 1958 and 2001. To cover all beaches, ten grid points located at deep

waters were chosen according the orientation and proximity of the study areas (see Fig. 1). However, as waves propagate from deep to shallow water they undergo some changes in their spatial energy (i.e. diffraction, refraction, shoaling, etc.) that have to properly be modeled to obtain the wave climate in the near shore region.

Since propagation of the 385,000 HIPOCAS data is unreliable, selected combinations of the long term probability distribution of significant wave height ($0.5 < H_s < 5$ m) and wave period ($3 < T_p < 12$ s) were propagated to the beaches using a mild slope parabolic model (OLUCA). The model solves in a discretized finite difference domain the mild slope equation,

$$\frac{\partial^2 A}{\partial y^2} + \left(2ik_0 + \frac{1}{cc_g} \frac{\partial cc_g}{\partial x} \right) \frac{\partial A}{\partial x} + \frac{1}{cc_g} \frac{\partial cc_g}{\partial y} \frac{\partial A}{\partial y} + \left(\frac{\omega^2}{g} - k_0^2 + \frac{ik_0}{cc_g} \frac{\partial cc_g}{\partial x} \right) A = 0 \quad (2)$$

where A is the wave amplitude (i.e. $\eta = Ae^{ik_0x}$), k_0 is a characteristic wave number, ω the frequency, c_g the group velocity and c the wave celerity. For each beach a mesh of 15×15 m resolution was obtained by interpolation from the IHM nautical charts.

To obtain a criterion for the wave height at the breaking point, from the HIPOCAS data-set, we compute the wave height that did not exceed more than 12 hours per year (H_{s12}) and then propagated to each beach. From this study, we found that the breaking depth at all study sites is around 5 m. Consequently, the wave height at this depth was chosen as the breaking wave height (H_b) in Eq. (1).

3.3. Nearshore wave climate

At the beach, a bivariate empirical histogram was built with the wave breaking heights H_b and their corresponding periods T_p . This joint distribution contains information about the annual rate of occurrence of a concrete sea state for a given value of H_b and T_p .

Table 3
ANOVA test evaluating Cala Millor (Ma3) seasonal sediment size distributions variability

Source	Sum of squares	df	Mean square	F-ratio	Significant
Between season samples	0.077	3	0.025	0.073	0.974
Whiting season samples	42.996	121	0.355		
Total	43.074	124			

The statistics conclude that there is not a significant difference between mean beach size seasonal distributions.

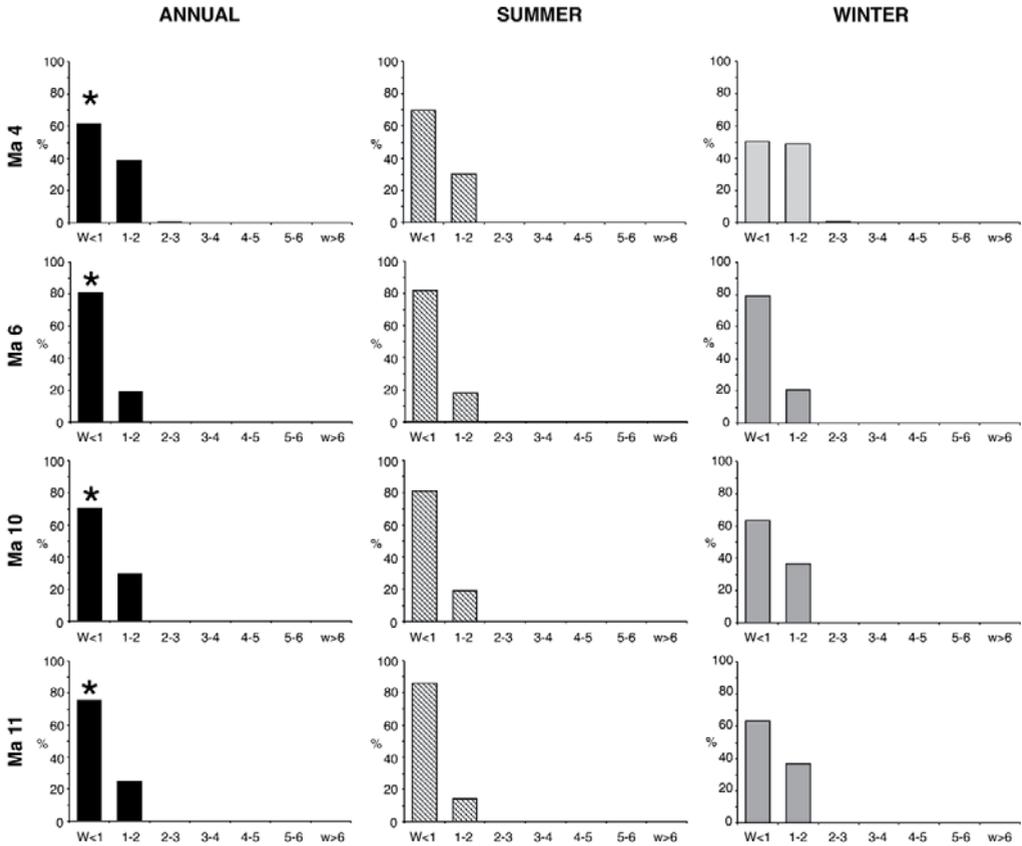


Fig. 5. Probabilities of occurrence of beach states for enclosed and sheltered beaches of Estany d'en Mas (Ma4), Cala Pi (Ma6), Platja d'en Repic (Ma10) and Cala Molins (Ma11). Stars indicate the modal state.

Trends between summer and winter were modeled using a log-normal probability distribution function. This function is characterized by the so-called location parameter μ^* and the scale parameter σ^* as (Castillo et al., 2005),

$$f(H_b|\mu^*, \sigma^*) = \frac{1}{H_b \sigma^* \sqrt{2\pi}} \exp\left(\frac{-(\ln(H_b) - \mu^*)^2}{2\sigma^{*2}}\right) \quad (3)$$

From the estimated log-normal parameters μ^* and σ^* , we calculated the mean and the standard deviation as,

$$\mu = \exp\left(\mu^* + \frac{\sigma^{*2}}{2}\right) \quad (4)$$

$$\sigma = \{\exp(2\mu^* + \sigma^{*2})(\exp(\sigma^{*2}) - 1)\}^{1/2} \quad (5)$$

As shown in Fig. 3, the evolution of the estimated mean and standard deviation leads to the conclusion that to an-

alyze correctly the beach state, seasonality has to be taken in to account. In this sense, two periods were proposed, the summer covering from April to September and the winter wave climate period from October to March. For each beach the combined (H_b , T_p) empirical bivariate histogram for the summer and winter periods was obtained from the relative wave frequencies. The wave breaking criterion chosen has been Hs12. This criterion allows to propagate 99.86% of the entire wave climate. At the breaking point the density spectra was reconstructed using the entire propagated data and therefore it can be seen as a modal state value if the interval higher probability is considered.

4. Results

4.1. Sediment characteristics

The analyzed beaches consist of medium to fine sands moderately to poorly sorted. Binocular microscopic

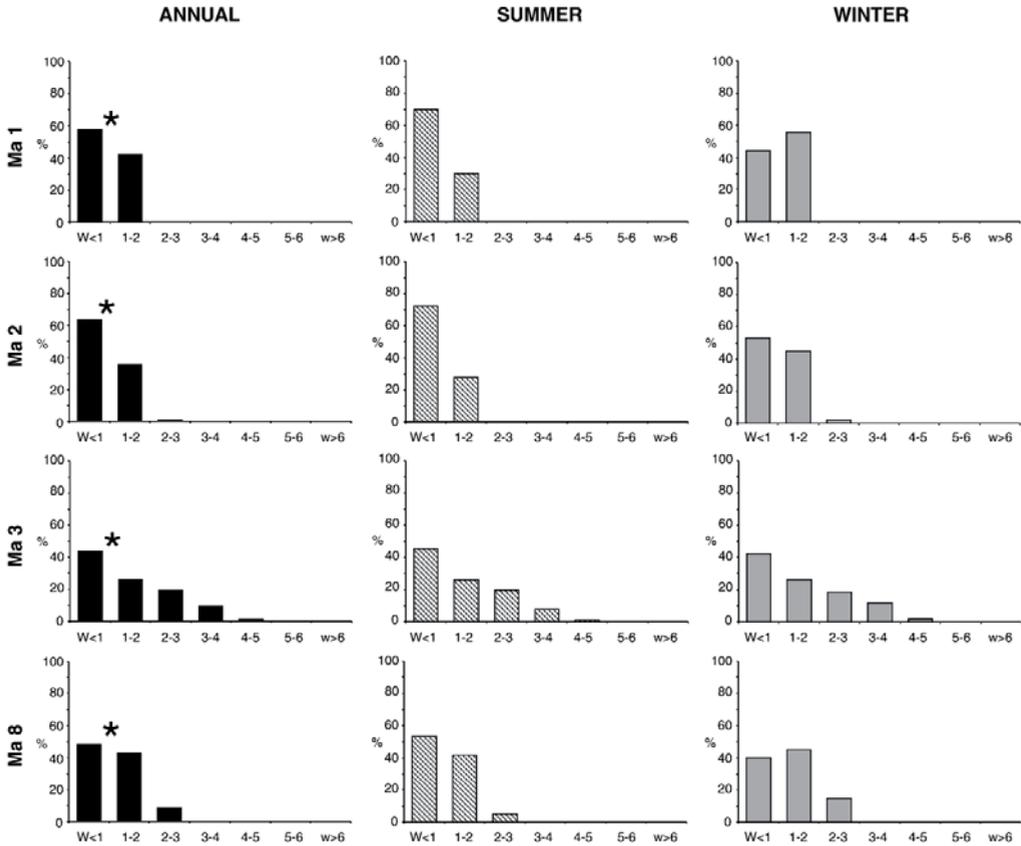


Fig. 6. Probabilities of occurrence of beach states for enclosed beaches of Cala Mesquinda (Ma1), Cala Agulla (Ma2), Cala Millor (Ma3) and Magalluf (Ma8). Stars indicate the modal state.

examination shows that the sediments are composed by a mixture of siliciclastic and biogenic materials mainly composed of foraminifera and shell fragments of gastropods, bivalves which sums around the 94.50% of the bulk sediment. The rest, a 5.5%, relates to quartz grains and cliff face detached fragments (Table 1).

Sand density values range from 2.68 to 2.88 g/cm³ which are in agreement with carbonates and dolomite mineral densities. Beaches such as Ma1 or Ma11 have the biggest density values because they receive a considerable input of carbonate rock fragments from cliffs. The lower density values correspond to Ma5 with a very important foraminifera bulk component, and also to Ma10 which falls in the pre-Miocene basement domain although is quite rich in quartz and non-carbonate grains, its density being lower.

Settling velocities range from 20.47 cm/s to 6.53 cm/s (Table 2). Higher velocities correspond with beach sands

rich in non-biogenic and coarse grains, and the lower velocities with the biogenic components. Despite the sand density being similar between the different samples, there are differences between them according to the settling velocity. This fact can be explained by the shape and the settling pattern of biogenic grains (Paphitis et al., 2002; Smith and Cheung, 2003).

Sediment size variation for DFVP calculations are poorly understood because beach mean sediment size and mean size distribution is known to be a conservative property (Fig. 4). Mean grain size and settling velocities may also vary little over time on beaches with mature sediment suites, e.g. constant water temperature and salinity, lack of fluvial sediment contributions, etc. (Anthony, 1998). Historical surveys carried in Ma 3 (Cala Millor), highlighted that if different samples are mixed there is not a significant variation between seasonal sampling in terms of mean size (Table 3). Nevertheless

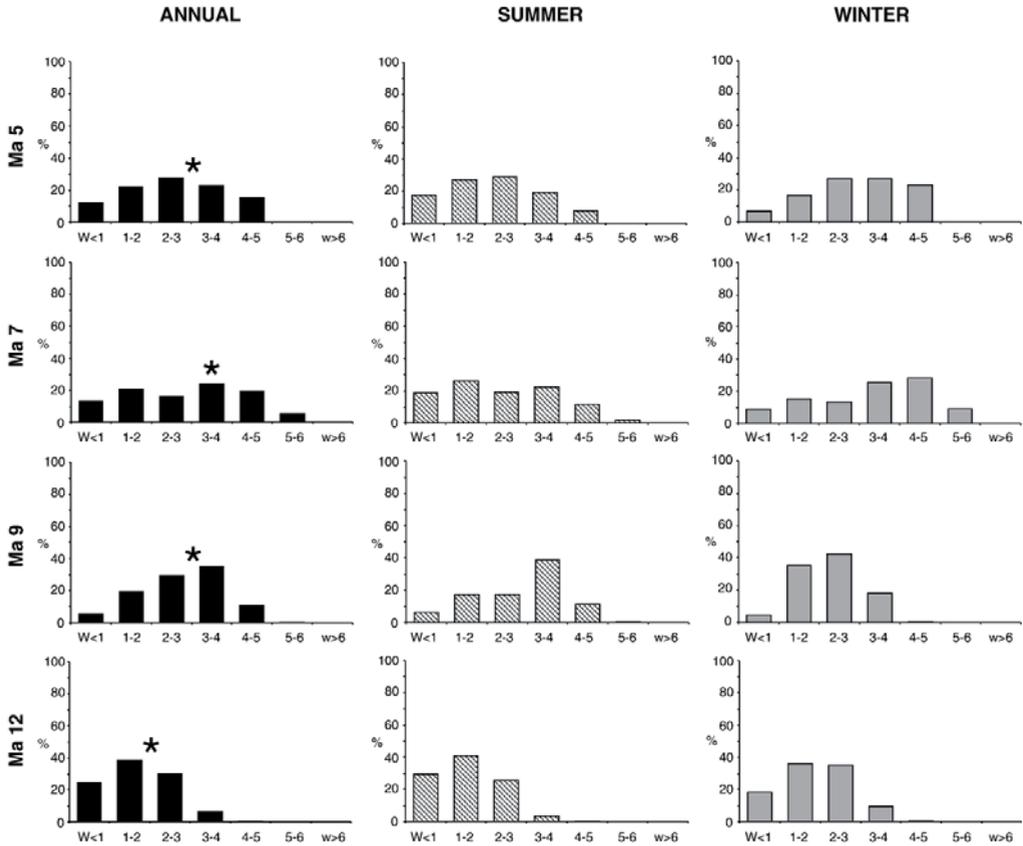


Fig. 7. Probabilities of occurrence of beach states for enclosed beaches of Es Trenc (Ma5), s'Arenal (Ma7), Sant Elm (Ma9) and Es Comú de Muro (Ma12). Stars indicate the modal state.

beach sediment size distribution from summer is slightly coarser and less classified than winter size distribution (Fig. 4).

4.2. Nearshore waves

A seasonal energetic displacement is observed for both, propagated wave heights and peak period, due to the existence of different behaviors throughout a climatic year. The winter energetic displacement is represented by an increment in the rate of occurrence of higher crest periods and higher significant wave heights. This effect is accentuated at north oriented beaches, which are more affected by the energetic north swells. A propitious period for this situation is during fall and winter, when severe storm affect the western balearic basin. Results for the different beaches are shown in Table 2, where wave

heights and wave periods correspond to the intervals with the maximum frequency in the bivariate histogram for each beach.

4.3. Morphodynamic classification

The morphodynamic state for the studied beaches has been predicted according to the model by Wright and Short (1983) where beach grain size and wave statistics were combined to calculate the Ω parameter in annual and seasonal basis. According to wave energy and sediment properties, beaches in Mallorca tend to fall in an intermediate state. Roughly, three groups of beaches can be separated according to their modal conditions. The first one corresponds to truly reflective beaches which are enclosed beaches or sheltered beaches. Ma4, Ma6, Ma10 and Ma11 which have a 75% of probability

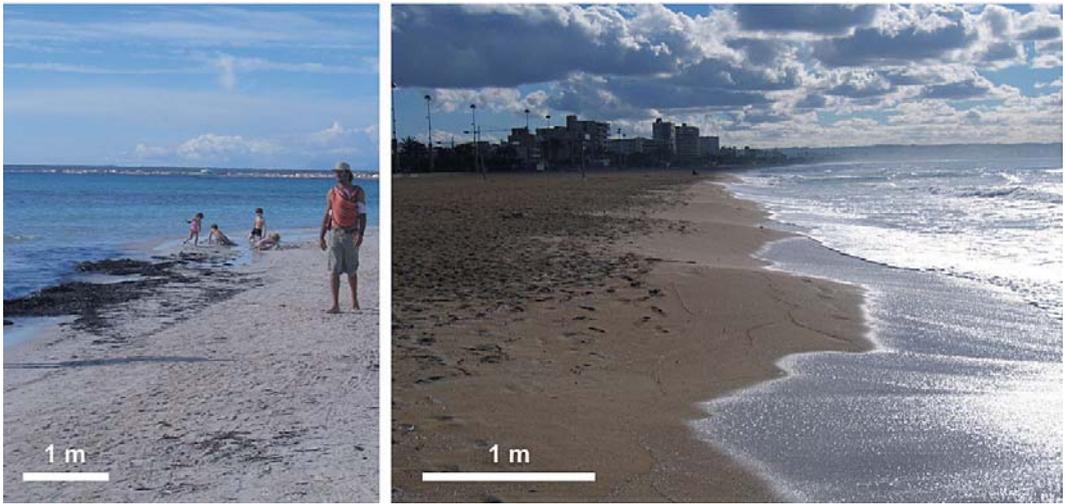


Fig. 8. Fieldwork observations of diagnostic beach features during summer season when Wright and Short's model predict reflective states. At left, a bar joined to the subaerial beach during night at Es Trenc (Ma5). Children are playing just on top. At right cusp at horns at s'Arenal (Ma7).

to fail below an Ω value of 1 (Fig. 5). The second group of beaches present modal conditions in the intermediate states although skewed to reflective positions. Beaches as Ma1, Ma2, Ma3 and Ma8 are good examples of this behavior. Probabilities to be in a Ω value of 2 and lower than 1, are between 60 and 80% of modal states (Fig. 6). This type of configuration relates with ridge–runnel and incipient transverse bars beach configuration. The rest of the beaches belong to intermediate states. Probabilities for each predictive state are sparse although Ω values between 3 and 4 achieve near the 30% of probability, nearest Ω intervals have probabilities larger than 10% (Fig. 7). This kind of beaches corresponds with semi-enclosed beaches, usually backed by a field of coastal dunes (Ma5, Ma7, Ma9 and Ma12). During fieldwork longshore, crescentic and transverse bars have been observed, and some of them were just emerged as a berm in the subaerial beach (Fig. 8). Distribution probabilities allow to assess the dynamic behavior of each beach according wave climate seasonality. Thus, from Figs. 5, 6 and 7 we can check that as in winter the probability for a morphodynamic state are higher than in summer. For instance, in the in Ma5, Ma7 and Ma12 during winter the probability and distribution for morphodynamic state is skewed towards reflective stages. Conversely, the same beaches are towards a dissipative position of the spectrum during summer. Another group of beaches do not present significant differences between winter and summer distributions, although the weights of the more dissipative stages are higher in

winter than in summer. This can be clearly observed for enclosed beaches as Ma1, Ma2, Ma3 and Ma4, and even for very sheltered beaches as Ma6, Ma11 o Ma11 (Figs. 5 and 6).

5. Discussion

The environmental setting of nearshore and foreshore in the study sites differs from many parameters described for low-energy beaches according to Jackson et al. (2002). Significant wave heights are greater than 0.21 m and the mild winds, corresponding to the sea breezes regime during summer, reach values of 5 m/s and occasionally can increase up to 10 m/s (Ramis et al., 1990). From late summer to early winter, storm episodes are frequent with more intensive winds blowing with a SE or NE component, and waves can achieve significant heights of 2.5 m. Nevertheless environmental setting also differs from high energy and open ocean beaches despite the presence of the diagnostic morphodynamic characteristics such as nearshore bar–rip morphology, cusps and steps (Lipman and Holman, 2000; Masselink and Hegge, 1995).

Recent studies maintain the poor application of high-energy models in predicting two-dimensional morphology in low-energy environments (Masselink and Pattiaratchi, 2001; Jackson et al., 2005; Eliot et al., 2006). However, in the case of Ma3, Ma8, Ma11 and Ma12 empirical studies involving beach profile monitoring bathymetries and sedimentological analysis have found

an optimal agreement between the Short and Wright model prediction and beach configuration. Thus, Ma3 is a beach with a set of crescentic bars which during summer evolve to a transverse bar until the bar joins the beachface (Fig. 9). The Short and Wright prediction for Ma3 is a combination of ridge–runnel and reflective states, although with considerable presence of transverse bar to the subaerial beach. Ma8 is a beach characterized by a small variability according to the protection that exerts the *Posidonia oceanica* meadows, but reflective and ridge–runnel stages are quite common and cusps on the emerged beach are a conspicuous form. This configuration ties with the prediction where reflective and ridge–runnel morphodynamical states have similar maximum probabilities of occurrence. All of that, points out that, in the absence of empirical profile and bathymetry-based studies, Wright and Short's model, is a useful parameter for gross classifications of Mediterranean beaches. However, when predictions are assessed from a seasonal point of view according to wave seasonality, Ω predictions tend to fail, mainly, in semi-enclosed beaches.

Predictions give a reflective configuration for summer season, nevertheless field observations in Ma3, Ma5, Ma7 and Ma12, points up that the beach profile slope is smoother (slopes between 0.018 and 0.027). Diagnostic features, as transverse and rhythmic bar corresponding to more dynamic states, are also present (Fig. 8). The reason

for this situation is that the Wright and Short model does not incorporate sea breeze effects on beach morphodynamics which seems to drive beach dynamics during the relative mild summer wave conditions. The sea breezes induce changes in the incidents wave field that may affect beach morphology and the associated processes, inducing longshore transport and overlaying series of daily mini-storm cycles characterized by erosion in the afternoon and beach accretion occurring the rest of the day (Masselink, 1996; Masselink and Pattiaratchi, 1998).

To check this hypothesis in 2004 during the summer sea breeze and winter conditions, a field survey was carried out in Cala Millor (Ma3). In this study, a wave gage measured wave height at 5 m depth simultaneously with the data provided at deep water by the WAM model at the same HIPOCAS point used for the analysis. This data, once propagated to the beach, were in 83% of cases 44% smaller than the wave height measured by the wave gage. Conversely, during winter, differences between wave heights propagated from the numerical model and the measured ones are only up to 20%. This leads to the conclusion that local sea breeze which is not included in the deep water wave model has to be taken into account to properly characterize morphodynamic Mallorcan beaches.

Major differences between beaches seem to be related to geological factors. The geological and physiographic framework controls the spectrum and the angle of incident



Fig. 9. Cala Millor (Ma3) aerial photography. Note the configuration of transverse bars along the coastline as well as rips and cusps. Dark areas correspond to *Posidonia oceanica* meadows.

waves; in fact enclosed beaches present less dynamic states than semi-enclosed beaches.

6. Conclusions

The morphodynamic model of Wright and Short has been used to classify beach morphology on Mediterranean low-energy beaches using 44 yr of wave data propagated to the wave breaking depth as well as detailed sediment analysis for each beach. Results relate to the probabilities for modal morphological state in annual and seasonal basis. The Ω parameter leads to separate Mallorcan beaches in three major groups according to the geological framework configuration. Thus modal reflective beaches relates to enclosed and sheltered beaches (a), reflective beaches lightly skewed to intermediate states are common in enclosed beaches exposed to main wave energetic directions (b), and intermediate beaches are hosted in semi-enclosed localities. Other works pointed out the discrepancies between predicted and observed beach states under various conditions. The analyses presented are in agreement for a gross annual classification as well as for winter predictions, when waves have large peak periods and bigger significant wave height. However, Ω fails in summer predictions because it does not incorporate the effect of summer sea breezes which exert an important influence in beach behavior when the action of sea waves is negligible. For this reason reflective stages are over represented with respect to the more dynamic features observed on the beaches, including transverse bars, cusps and rips. The Wright and Short model rely largely on dynamic factors, but it just introduces averaged wave statistics and sediment size. Our observations show that, for Mediterranean beaches, further research is required in order to associate the summer sea breezes and associated hydrodynamics to elucidate the controls on beach morphology classification (Figs. 5 and 6).

Acknowledgments

Financial support from the Balearic Island Government is greatly acknowledged (UGIZ project). Support in experimental sediment characterization from J. Guillén from CMIMA (CSIC) and J.J. Fornós from UIB is greatly appreciated. Field support from B. Casas and P. Balaguer is also acknowledged. Critical comments from Dr. E. Anthony and two anonymous referees have improved substantially the first manuscript. The authors thank Puertos del Estado (Ministerio de Fomento) and Mr. Enrique Álvarez for the use of HIPOCAS database. FJM is indebted to the Spanish

Ministry of Education and Science for the funding provided in the “Ramon y Cajal” Program.

References

- Anthony, E., 1998. Sediment-wave parametric characterization of beaches. *J. Coast. Res.* 14, 347–352.
- Basterretxea, G., Orfila, A., Jordi, A., Casas, B., Lynett, P., Liu, P.L.-F., Duarte, C.M., Tintore, J., 2004. Seasonal dynamics of a microtidal pocket beach with *Posidonia oceanica* seabeds (Mallorca, Spain). *J. Coast. Res.* 20, 1155–1164.
- Benedet, L., Finkl, C.W., Campbell, T., Klein, A., 2004. Predicting the effect of beach nourishment and cross-shore sediment variation on beach morphodynamic assessment. *Coast. Eng.* 51, 839–861.
- Bethoux, J., Cópín-Motegut, G., 1986. Biological fixation of atmospheric nitrogen in the Mediterranean sea. *Limnol. Oceanogr.* 31 (6), 1353–1358.
- Blott, S.J., Pye, K., 2001. Gradistat: a grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surf. Process. Landf.* 26, 1237–1248.
- Castillo, E., Hadi, A.S., Balakrishnan, N., Sarabia, J.M., 2005. Extreme Value and Related Models with Applications in Engineering and Science. Wiley, New York.
- Dean, R.G., 1973. Heuristic models of sand transport in the surf zone engineering dynamics in the surf zone. *Proc. First Aust. Conf. Coastal Eng., Inst. Eng., Australia*, pp. 208–214.
- Eliot, M.J., Travers, A., Eliot, I., 2006. Morphology of a low-energy beach, Como Beach, Western Australia. *J. Coast. Res.* 22, 63–77.
- Fornós, J.J., Ahr, W.M., 1997. Temperate carbonates on modern, low-energy, isolated ramp: the balearic platform, Spain. *J. Sediment. Res.* 67, 364–373.
- Gelabert, B., Sàbat, F., Rodríguez-Perea, A., 1992. An structural outline of the Serra de Tramuntana of Mallorca (Balearic Islands). *Tectonophysics* 203, 167–183.
- Gibbs, R.J., Matthews, M.D., Link, D.A., 1971. The relationship between sphere size and settling velocity. *J. Sediment. Petrol.* 41, 7–18.
- Giró, S., Maldonado, A., 1985. Análisis granulométrico por métodos automáticos: tubo de sedimentación y sedigraph. *Acta Geol. Hisp.* 20, 95–102.
- Goodfellow, A., Stephenson, W.J., 2005. Beach morphodynamics in a strong-wind bay: a low energy environment? *Mar. Geol.* 214, 101–116.
- Gouraly, M.R., 1968. Beach and Dune Erosion Tests, Laboratory Report M935/M936. Delft Hydraulics.
- Jackson, N.L., Nordstrom, K.F., Eliot, I., Masselink, G., 2002. Low energy” sandy beaches in marine and estuarine environments: a review. *Geomorphology* 48, 147–162.
- Jackson, D.W.T., Cooper, J.A.G., del Rio, L., 2005. Geological control of beach morphodynamic state. *Mar. Geol.* 216, 297–314.
- Kench, P.S., 1997. Contemporary sedimentation in the Cocos (Keeling) Islands, Indian Ocean: interpretation using settling velocity analysis. *Sediment. Geol.* 114, 109–130.
- Klein, A.H.F., de Menezes, J.T., 2001. Beach morphodynamic and profile sequence for a headland bay coast. *J. Coast. Res.* 17, 812–835.
- Levoy, F., Anthony, E.J., Monfort, O., Larssonneur, C., 2000. The morphodynamics of megatidal beaches in Normandy, France. *Mar. Geol.* 171, 39–59.
- Lipman, T.C., Holman, R.A., 2000. The spatial and temporal variability of sand bar morphology. *J. Geophys. Res.* 95 (C7), 11575–11590.
- Masselink, G., 1996. Sea breeze activity and its effect on coastal processes near Perth, Western Australia. *J. R. Soc. West. Aust.* 79, 199–205.

- Masselink, G., Hegge, B., 1995. Morphodynamics of meso- and macrotidal beaches: examples from central Queensland, Australia. *Mar. Geol.* 129, 1–23.
- Masselink, G., Pattiaratchi, C.B., 1998. The effect of sea breeze on beach morphology, surf zone hydrodynamics and sediment resuspension. *Mar. Geol.* 146, 115–135.
- Masselink, G., Pattiaratchi, C.B., 2001. Seasonal changes in beach morphology along the sheltered coastline of Perth, Western Australia. *Mar. Geol.* 172, 243–263.
- Masselink, G., Short, A.D., 1993. The effect of tide range on beach morphodynamics, a conceptual model. *J. Coast. Res.* 9, 785–800.
- Paphitis, D., Collins, M.B., Nash, L.A., Wallbridge, S., 2002. Settling velocities and entrainment thresholds of biogenic sands (shell fragments) under unidirectional flow. *Sedimentology* 49, 211–225.
- Ramis, C., Jansà, A., Alonso, S., 1990. Sea breeze in Mallorca. A numerical study. *Meteorol. Atmos. Phys.* 42, 249–258.
- Ranasinghe, R., Symonds, G., Black, K., Holman, R., 2004. Morphodynamics of intermediate beaches: a video imaging and numerical modelling study. *Coast. Eng.* 51, 629–655.
- Sanderson, P.G., Eliot, I., 1999. Compartmentalisation of beachface sediments along the southwestern coast of Australia. *Mar. Geol.* 162, 145–174.
- Short, A.D., 1999. In: Short, A.D. (Ed.), *Wave-dominated Coasts, Handbook of Beach and Shoreface Morphodynamics*. Wiley, Chichester, pp. 173–203.
- Smith, D.A., Cheung, K.F., 2002. Empirical relationships for grain size parameters of calcareous sand on Oahu, Hawaii. *J. Coast. Res.* 18, 82–93.
- Smith, D.A., Cheung, K.F., 2003. Settling characteristics of calcareous sand. *J. Hydraul. Eng.* 129, 479–483.
- Soares, C.G., Weisse, R., Carretero, J.C., Alvarez, E., 2002. A 40 years hindcast of wind, sea level and waves in European waters. Proc. 21st International Conference on Offshore Mechanics and Arctic Engineering 23–28 June 2002. Norway, Oslo.
- Sonu, C.J., James, W.R., 1973. A Markov model for beach profile changes. *J. Geophys. Res.* 78, 1462–1471.
- Sonu, C.J., van Beek, J.L., 1971. Systematic beach changes on the outer banks, North Carolina. *J. Geol.* 79, 416–425.
- Woodroffe, C.D., 2003. *Coasts: form, process and evolution*. Cambridge University Press, Cambridge, UK. 623 pp.
- Wright, L.D., Short, A.D., 1983. Morphodynamics of beaches and surf zones in Australia, *Handbook of Coastal Processes and Erosion*. In: Komar, P.D. (Ed.), CRC Press, Boca Raton, FL, pp. 35–64.
- Wright, L.D., Thom, B.G., 1977. Coastal depositional landforms, a morphodynamical approach. *Prog. Phys. Geogr.* 1, 412–459.
- Wright, L.D., Short, A.D., Boon, J.D., Hayden, B., Kimball, S., List, J.H., 1987. The morphodynamic effects of incident wave groupiness and tide range on an energetic beach. *Mar. Geol.* 74, 1–20.

Wave energy and the upper depth limit distribution of *Posidonia oceanica*

Eduardo Infantes*, Jorge Terrados, Alejandro Orfila, Bartomeu Cañellas and Amaya Álvarez-Ellacuría

Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC-UIB), Miquel Marqués 21, 07190, Esporles, Spain, e-mail: eduardo.infantes@uib.es

* Corresponding author

Abstract

It is widely accepted that light availability sets the lower limit of seagrass bathymetric distribution, while the upper limit depends on the level of disturbance by currents and waves. The establishment of light requirements for seagrass growth has been a major focus of research in marine ecology, and different quantitative models provide predictions for seagrass lower depth limits. In contrast, the influence of energy levels on the establishment, growth, and maintenance of seagrasses has received less attention, and to date there are no quantitative models predicting the evolution of seagrasses as a function of hydrodynamics at a large scale level. Hence, it is not possible to predict either the upper depth limit of the distribution of seagrasses or the effects that different energy regimes will have on these limits. The aim of this work is to provide a comprehensible methodology for obtaining quantitative knowledge and predictive capacity for estimating the upper depth limit of seagrasses as a response to wave energy dissipated on the seafloor. The methodology has been applied using wave data from 1958 to 2001 in order to obtain the mean wave climate in deep water seaward from an open sandy beach in the Balearic Islands, western Mediterranean Sea where the seagrass *Posidonia oceanica* forms an extensive meadow. Mean wave conditions were propagated to the shore using a two-dimensional parabolic model over the detailed bathymetry. The resulting hydrodynamics were correlated with bottom type and the distribution of *P. oceanica*. Results showed a predicted near-bottom orbital velocity of between 38 and 42 cm s⁻¹ as a determinant of the upper depth limit of *P. oceanica*. This work shows the importance of interdisciplinary effort in ecological modeling and, in particular, the need for hydrodynamical studies to elucidate the distribution of seagrasses in shallow depths. Moreover, the use of predictive models would permit evaluation of the effects of coastal activities (construction of ports, artificial reefs, beach nutrient-input, dredging) on benthic ecosystems.

Keywords: near-bottom orbital velocity; parabolic model; *Posidonia oceanica* upper depth limit; seagrass distribution; wave energy.

Introduction

Seagrasses are marine flowering plants that cover large areas in shallow coastal waters; the meadows they form are among the most biologically diverse and productive components of coastal systems. *Posidonia oceanica* (L.) Delile, an endemic seagrass species of the Mediterranean Sea, grows between the depths of 0.5 and 45 m (Procaccini et al. 2003) and covers an estimated surface area between 2.5 and 5 million ha, approximately 1–2% of the 0–50 m depth zone (Pasqualini et al. 1998).

The study of light requirements for seagrass growth has been a major focus of research in marine ecology, and different quantitative models provide predictions of the seagrass lower depth limits (Dennison and Alberte 1985, Dennison 1987, Duarte 1991, Kenworthy and Fonseca 1996, Koch and Beer 1996, Greve and Krause-Jensen 2005). Seagrasses can thrive up to depths where the irradiance at the top of the leaf canopy is above 11% of surface irradiance (Duarte 1991), or where the number of hours with values of irradiance above photosynthetic saturation is >6–8 h (Dennison and Alberte 1985). The upper depth limit of distribution of seagrasses has been related to their tolerance of desiccation at low tide (Koch and Beer 1996) and to ice scour (Robertson and Mann 1984). In seas with low tidal range and no ice formation, such as the Mediterranean Sea, the upper depth limit of seagrasses is determined mainly by their tolerance of wave energy (Koch et al. 2006). However, quantitative estimates of the wave energy that sets the upper depth limit of seagrasses are still scarce to date.

Studies in flumes have evaluated the influence of seagrass on unidirectional water flow (Fonseca et al. 1982, Gambi et al. 1990, Folkard 2005), sediment stabilization (Fonseca and Fisher 1986), and wave attenuation (Fonseca and Cahalan 1992). Field measurements and experiments have also been performed to study water flow in seagrass meadows (Fonseca 1986, Worcester 1995, Koch and Gust 1999) and evaluate their effect on particle resuspension (Terrados and Duarte 2000, Gacia and Duarte 2001). However, few studies have analyzed the effects of local hydrodynamics on seagrasses.

Storms can resuspend sediments and uproot seagrasses, whereas sediment deposition can bury them and cause mortality (Williams 1988, Preen et al. 1995, Duarte et al. 1997, Bell et al. 1999). Conceptual models have been proposed to explain differences in shape, bottom relief, and cover of seagrass meadows (Fonseca et al. 1983, Fonseca and Kenworthy 1987) or the depth distribution of intertidal seagrass (van Katwijk et al. 2000) as a function of wave energy and/or current velocity. Local hydrodynamics (indirectly estimated from the weight loss of clod cards) have been related to depth and seagrass presence and have been used to identify the habitat requirements of South Australian seagrass species

(Shepherd and Womersley 1981). Keddy (1984) developed a relative wave exposure index (REI) in order to quantify the degree of wave exposure by using wind speed, direction, and fetch measurements. REI values have been correlated to different attributes of seagrass meadows, such as the content of silt-clay and organic matter of the sediment, seagrass cover, shoot density (Fonseca and Bell 1998, Fonseca et al. 2002, Krause-Jensen et al. 2003, Frederiksen et al. 2004), and biomass (Hovel et al. 2002). Shallow seagrass populations tend to be more spatially fragmented in wave-exposed environments than in wave-sheltered environments (Fonseca and Bell 1998, Frederiksen et al. 2004), and temporal changes in seagrass cover are also greater at the most wave-exposed sites (Frederiksen et al. 2004). Plaster clod cards and REI provide only semi-quantitative idiosyncratic estimates of current and/or wave energy and, therefore, comparison between studies is difficult. Additionally, REI estimates do not consider the influence of depth on wave damping (but see van Katwijk and Herms 2000). Direct quantitative measurements of the energy of waves and currents impinging on seagrasses are necessary to elucidate their effects on them, to identify their habitat requirements, to predict their spatial distribution and their response to both natural (i.e., storms, hurricanes, etc.) and anthropogenic (i.e., dredging and beach eutrophication, coastal development, etc.) disturbances.

The main goal of this study was to develop a methodology to estimate the wave energy impinging on seagrass meadows (Figure 1) and to obtain a quantitative and testable relationship between wave energy and the upper depth limit of the Mediterranean seagrass *Posidonia oceanica*. Wave energy estimation was obtained by analyzing 44 years of wave data. We consider long-term historical wave data, rather than present-day wave measurements, are more appropriate to link wave energy to the presence of *P. oceanica*, given the low rates of vegetative growth and space occupancy in this seagrass species (Marbà and Duarte 1998). Aerial color photographs and bathymetric data were also used for an accurate mapping of the meadow and to delimit the upper

depth boundary of *P. oceanica*. We provide a testable estimate of the near-bottom current velocity that could set a threshold for the presence of *P. oceanica* and, therefore, determine the upper depth limit of this seagrass.

Materials and methods

Study areas and regional settings

This work was carried out in Cala Millor, located on the northeast coast of Majorca Island (Figure 2A,B). The beach is in an open bay with an area of ca. 14 km². Near the coast to 8 m depth, there is a regular slope indented with sand bars near the shore (Figure 2C); these bars migrate from offshore to onshore between periods of gentle wave conditions. At depths from 6 m to 35 m, the seabed is covered by a meadow of *Posidonia oceanica*. This area was chosen for this study because of the availability of data from previous studies. The tidal regime is microtidal, with a spring range of less than 0.25 m (Gómez-Pujol et al. 2007). The bay is located on the east coast of the island of Majorca and it is therefore exposed to incoming wind and waves from NE to ESE directions. According to the criteria of Wright and Short (1983), Cala Millor is an intermediate barred sandy beach formed by biogenic sediments with median grain values ranging between 0.28 and 0.38 mm at the beach front (L. Gómez-Pujol and A. Orfila, unpublished data).

HIPOCAS database (1958–2001) and deep water wave characterization

Wave data used are part of the Hindcast of Dynamic Processes of the Ocean and Coastal Areas of Europe (HIPOCAS) project. This database consists of a high resolution, spatial and temporal, long-term hindcast dataset (Soares et al. 2002, Ratsimandresy et al. 2008). The HIPOCAS data were collected hourly for the period 1958 to 2001, providing 44 years of wave data with a 0.125° spatial resolution; this is the most complete wave data base currently available for the Mediterranean Sea. These hindcast models have become powerful tools, not only for engineering or predictive purposes, but also for long-term climate studies (Cañellas et al. 2007). These data were produced by the Spanish Harbor Authority by dynamical downscaling from the National Center for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) global reanalysis using the regional atmospheric model REMO. We used the data from HIPOCAS node 1433 (see Figure 2B) located 10 km offshore at 50 m depth, which is the closest HIPOCAS node to Cala Millor. The long-term distribution of significant wave height and wave direction at this node (Figure 3) shows that the most energetic waves usually come from N–NNE. These wave directions are also the most frequent during the 44-year dataset.

Data contained in the HIPOCAS node consist of a set of sea states (one per hour) defined by their significant wave height, spectral peak period, and direction. An estimation of the long-term distribution of the mean significant wave height (H_m) and its standard deviation was

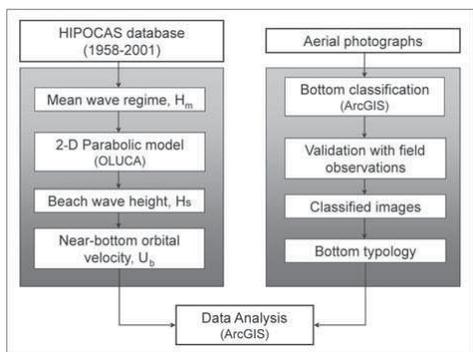


Figure 1 Schematic diagram of the approach presented. HIPOCAS, hindcast of dynamic processes of the ocean and coastal areas of Europe.

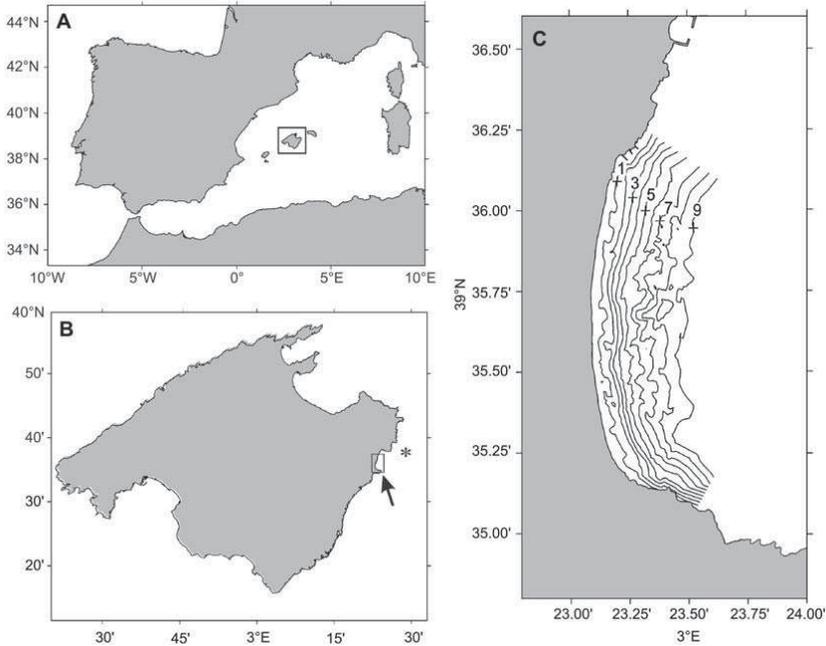


Figure 2 (A) Location of Majorca in the Mediterranean Sea. (B) Location of the study area of Cala Millor in Majorca. The asterisk (*) indicates HIPOCAS node 1433. (C) Bathymetry of Cala Millor with isobaths (in meters).

carried out making use of the lognormal probability distribution (Castillo et al. 2005). The long-term probability distribution identifies the most probable sea state for the 44-year period. Before estimation of this wave regime, data were pre-selected taking into account their incoming directions. Only sea states directed towards the beach were included in the analysis. Mean wave climate for Hipocas node 1433 provides an H_m value of 1.53 ± 0.96 (± 1 SD) m, a peak period of 7.3 s, and a direction of 11.25° . This mean (most probable) wave cli-

mate was propagated to the shore using a parabolic model.

Shallow-water wave conditions

As water waves propagate from the region where they are generated to the coast, both wave amplitude and wavelength are modified. The surf zone is a highly dynamic area where energy from waves is partially dissipated through turbulence in the boundary layer and transformed into short and long waves, mean sea level variations, and currents (Dean and Dalrymple 2002). In the present work, waves were propagated using a gentle slope parabolic model (OLUCA-MC), which includes refraction-diffraction effects as well as energy losses due to wave breaking (Kirby and Dalrymple 1983, GIOC 2003). Detailed bathymetry obtained with echo-sounding was used to generate the numerical mesh. The model solves continuity and momentum equations assuming a smooth bottom (e.g., variations of the bottom negligible within a wave length) and converting the hyperbolic system to a parabolic system (e.g., with wave propagation in one direction).

Two grids were generated: the external grid (122x81 nodes), which covers the deeper area with 75 m resolution between nodes, and the internal grid (140x311 nodes), which covers the shallow area with 15 m resolution. The model output provides the wave field (significant wave height and direction of the mean flux energy) in the whole grid.

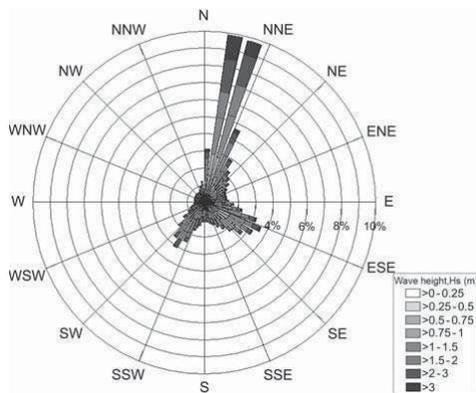


Figure 3 Directional wave histogram for HIPOCAS node 1433 (50 km from Cala Millor).

Maximum near-bottom orbital amplitudes (A_b) were calculated following the linear wave theory:

$$A_b = \frac{Hs}{2\sinh(2\pi D/\lambda)} \quad (1)$$

where D is the water depth and λ is the wave length calculated iteratively as:

$$\lambda = \frac{T_p^2 g}{2\pi} \tanh\left(\frac{2\pi D}{\lambda}\right) \quad (2)$$

where T_p is the peak period and g is the acceleration of gravity. The maximum near-bottom orbital velocity (U_b) is:

$$U_b = 2\pi A_b / T_p \quad (3)$$

Bottom typology and bathymetry

Remote sensing of the seabed from air or space is commonly used for mapping seagrass habitats over a wide range of spatial scales (McKenzie et al. 2001). Satellite spectral images from Ikonos are suitable for the detection and mapping of the upper depth limit of seagrass distribution in shallow clear waters (Fornes et al. 2006). Aerial color photographs have been used in some studies to describe temporal changes in the distribution of seagrasses (Hine et al. 1987), to evaluate the effect of wave exposure (Frederiksen et al. 2004), and the influence of anthropogenic activities (Lerliche et al. 2006). We used aerial photographs for mapping the area covered by *Posidonia oceanica*, dead *P. oceanica* rhizome, rocks, and sand up to 11 m depth. Aerial color photographs were taken in August of 2002 with a resolution of 0.4 m. Polygons were drawn around the different areas with Arc/GIS software (Arc/Info and Arc/Map v9.0, ESRI) and classified by bottom types. In those areas where bottom recognition was not possible from the aerial photographs, field surveys were carried out to identify the typology of the substratum (some areas tend to accumulate seagrass leaves which can lead to false interpretation of aerial images). Image classification was validated with bathymetric filtered echo-soundings and field observations.

During 2005, an acoustic survey was carried out to determine the bathymetry of the inner mesh and to test the classification of the seagrass coverage from the aerial photographs. Acoustic mapping of *Posidonia oceanica* was performed with a ship-mounted Biosonics DE-4000 echo sounder (BioSonics, Inc., Seattle, USA) equipped with a 200 KHz transducer. The draught of the boat allowed sampling up to depths of approximately 0.5 m. Inshore-offshore echo-sounding transects were sampled perpendicularly to the bathymetric gradient, with a separation of 50 m between transects. Acoustic pulse rate was set to 25 s⁻¹ and the sampling speed was set to 3 knots, which allowed for a horizontal resolution of 1 m (Orfila et al. 2005). Bottom typology was estimated as the most probable after echogram examination using the first to second bottom echo ratio technique (Orlowski 1984, Chivers et al. 1990). The resulting echo sounding points were filtered, averaged (1 output equals 20 pings) and

clustered into three groups (*P. oceanica* meadows, sandy, and hard bottoms) taking into consideration calibrations performed for previously classified bottoms. Hard bottoms include rocks, *P. oceanica* rhizome mats, and zones with poor seagrass coverage (Figure 4). Afterwards, this map was verified by direct observation at random points, also distinguishing those bottom types that the algorithm was not able to identify (i.e., dead rhizome and rocks). The map of bottom typology in Figure 4 represents the final classification (aerial photograph verified with echo-sounding and direct SCUBA observations).

Bathymetric data were interpolated using the kriging technique to create 1-m scale depth contours that were overlaid with bottom typology. Similarly, maximum near-bottom orbital velocity data were interpolated to create 5-cm s⁻¹ scale U_b contours that were overlaid with bottom typology. Percent coverage of rocks, sand, dead rhizomes of *Posidonia oceanica*, and *P. oceanica* were calculated for each depth and near-bottom orbital velocity interval. A total of 400 points were randomly selected throughout the study area, and the corresponding bottom type and near-bottom orbital velocity interval were used to calculate the average near-bottom orbital velocity for each bottom type. Additionally, 400 points along the upper limit of the *P. oceanica* meadow were randomly selected and the corresponding near-bottom orbital velocities were used to estimate an average along that edge. Differences in near-bottom orbital velocity between each bottom type and the upper limit of *P. oceanica* were evaluated using Kruskal-Wallis non-parametric analysis of variance (due to heterogeneity of data variance).

Results

The study site has a total area of 121.44 ha wherein sand and *Posidonia oceanica* meadow are the most abundant substrata. Shallow bottoms in Cala Millor (depths between 0 and 6 m) are mostly sandy, with some patches of rock, particularly in the center and south parts of the beach (Table 1). The upper depth limit of *P. oceanica* is located between 5 and 6 m and the meadow continues down to 30–35 m depth (acoustic survey data not shown). The limit between seagrass and sand is irregular and has several areas of hard substratum between 4 and 7 m that correspond to dead rhizomes of *P. oceanica* partly covered by sand and algae, which can be indicative of meadow regression (Figure 4). *P. oceanica* reaches the highest percentage cover at depths greater than 8 m, and no stands are found in depths shallower than 4 m. Most of the rocky bottom is in the shallowest water, between 0 and 2 m. Two large sand fingers cross the seagrass meadow at the center of the study area; these may have resulted from sand transport from the exposed beach after storm events.

Mean wave conditions propagated over the beach resulted in wave heights between 0.2 and 0.4 m, with significant wave breaking in the shallow sandy area (between 0.5 and 1 m) (Figure 5A) and near-bottom orbital velocities up to 110 cm s⁻¹. The highest velocities occurred in the shallow sandy part of the beach (between 1 and 3 m) (Figure 5B).

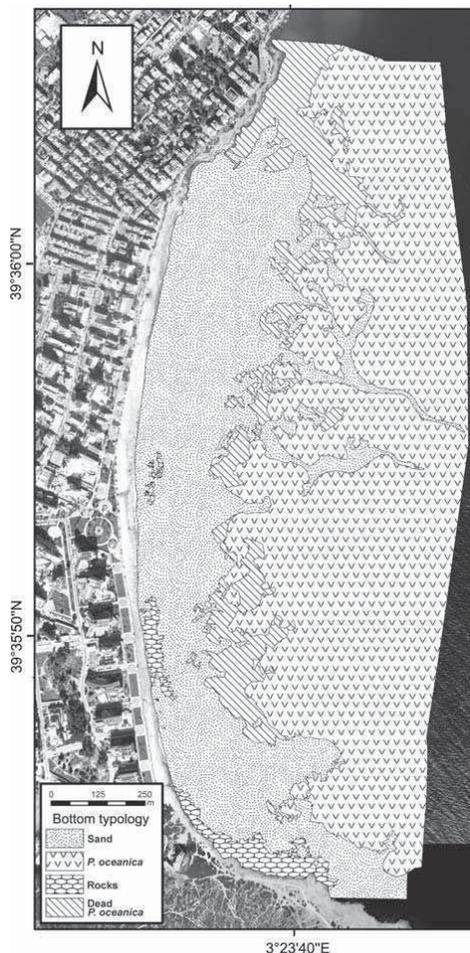


Figure 4 Bottom typology off Cala Millor (Majorca, western Mediterranean Sea). *P.*, *Posidonia*.

Sandy areas are associated with high values of near-bottom orbital velocities, while *Posidonia oceanica* is associated with lower velocities (Figure 6). *P. oceanica* is

not present in areas with velocities higher than 38–42 cm s⁻¹. This velocity interval might be considered a first estimate of the threshold near-bottom orbital velocity that allows the *P. oceanica* to occur in Cala Millor.

Variance of near-bottom orbital velocity was higher in sand, rock, and dead *Posidonia oceanica* than in the *P. oceanica* meadow (Figure 7). Kruskal-Wallis non-parametric analysis of variance detected significant differences in the average near-bottom orbital velocity between bottom types [H (4, n=800)=310.34, p<0.001], and post hoc multiple comparisons of mean ranks showed that velocities were lower in *P. oceanica* stands and at the *P. oceanica* upper limit than in rock, sand, and dead *P. oceanica* (Figure 7).

Discussion

In this study, we present a methodology to estimate the wave energy that determines the upper depth limit of *Posidonia oceanica*. Our data show that an increase in wave energy is related to a decrease of *P. oceanica* cover, and that above a threshold level of wave energy seagrass is not present. It is important to emphasize that the evidence provided in this study is correlative and applies to our study site only. Other locations, with different sediment characteristics and wave climates might provide different threshold values. Additional sources of disturbance (both natural and anthropogenic) will also introduce variability in the threshold estimates. The near-bottom orbital velocities are computed from numerical model predictions and real velocities within the meadow could be lower due to wave attenuation by the seagrass meadow and wave-current interactions. However, the predicted velocities where *P. oceanica* is not present provide an estimate of the threshold velocities that would impede the persistence of this seagrass species.

The usefulness of this methodology is that it provides quantitative estimates of wave energy that sets the upper depth limit of *Posidonia oceanica* and, therefore, these can be compared to those obtained in other locations.

We obtained an estimate of the threshold value of near-bottom orbital velocity that allows the long-term persistence of *Posidonia oceanica* (38–42 cm s⁻¹). It has been suggested that *Zostera marina* L. can tolerate unidirectional

Table 1 Percent cover of bottom type at different depths.

Depth (m)	<i>Posidonia oceanica</i>	Sand	Rock	Dead <i>Posidonia oceanica</i>
1–2	0	96.63	3.24	0.13
2–3	0	97.73	1.75	0.52
3–4	0	97.31	0	2.69
4–5	0.46	83.36	0	16.17
5–6	11.55	51.03	0	37.42
6–7	42.34	25.27	0	32.39
7–8	73.87	21.10	0	5.03
8–9	83.82	14.64	0	1.54
9–10	94.40	5.59	0	0
10–11	97.98	2.02	0	0
Total area (m ²)	565,349.42	528,241.5	25,192.7	95,632.92

Total area of coverage of each bottom type is presented in the last row.

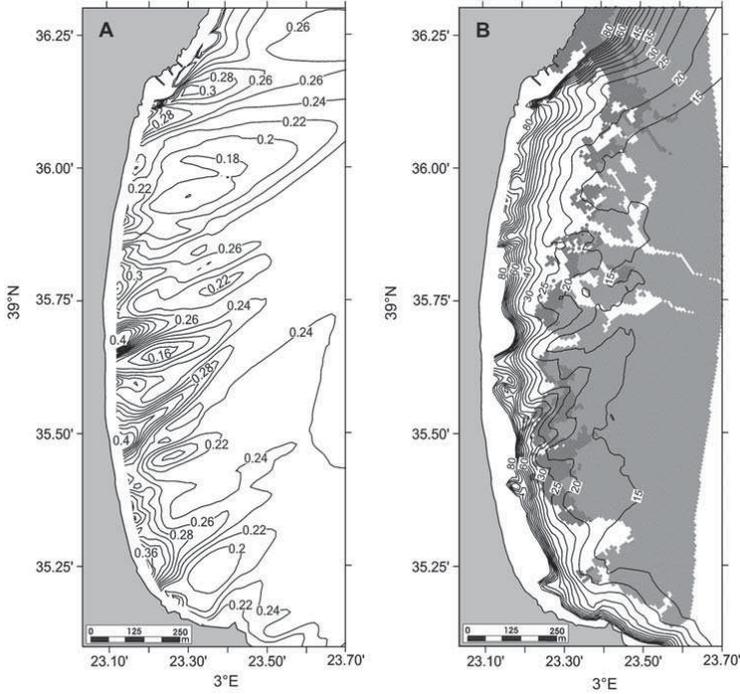


Figure 5 (A) Distribution of wave heights (m) at the beach derived from mean wave conditions (11.25° , $H_m=1.53$ m, $T_p=7.30$ s). (B) Distribution of near-bottom orbital velocities (U_b , $m\ s^{-1}$) at the beach and cover of the *Posidonia oceanica* meadow (gray) and dead rhizomes (dark gray).

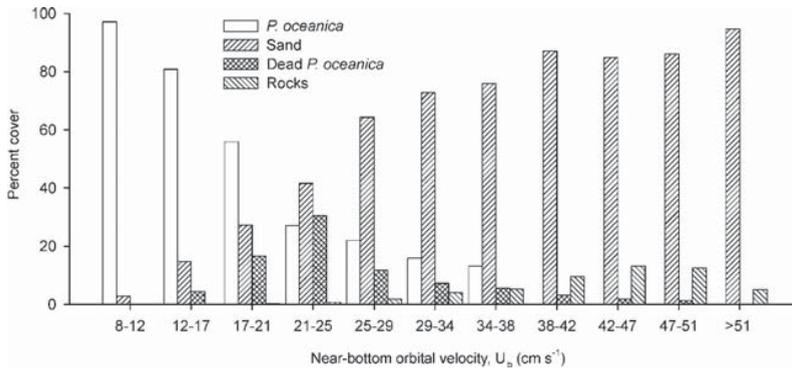


Figure 6 Percent coverage of the different bottom types in each of the near-bottom orbital velocity (U_b) intervals established in Cala Millor (Majorca, Western Mediterranean Sea). *P.*, *Posidonia*.

tional current velocities up to 120–150 $cm\ s^{-1}$ and that the meadows formed by this species become spatially fragmented at tidal current velocities of 53 $cm\ s^{-1}$ (Fonseca et al. 1983). Shallow mixed meadows of *Z. marina* and *Halodule wrightii* Asch. seem to remain spatially fragmented at tidal current speeds above 25 $cm\ s^{-1}$ (Fonseca and Bell 1998). Experimental transplantations of *Z. marina* along depth gradients in intertidal zones indicate that this species cannot persist at sites where the maximum

bottom orbital velocity during the tidal cycle reaches 53–63 $cm\ s^{-1}$. Furthermore, this species can survive when exposed to waves for less than 60% of the time and maximum orbital velocity is less than 40 $cm\ s^{-1}$ (van Katwijk and Harmus 2000). Hence, our estimate of the threshold value of near-bottom orbital velocity that allows the persistence of *P. oceanica* is within the range of values of current velocity proposed for other seagrass species.

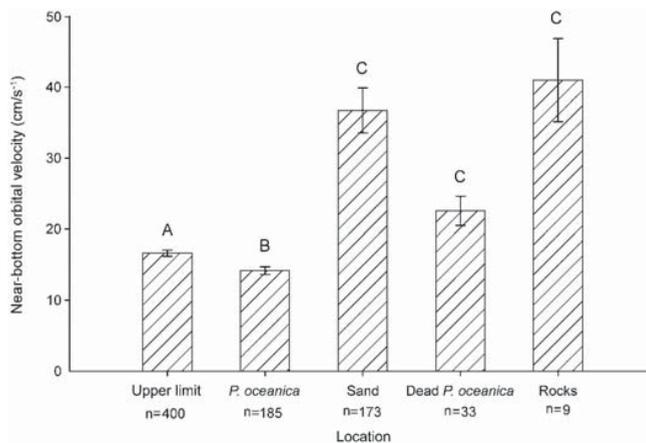


Figure 7 Mean near-bottom orbital velocity above each bottom type in Cala Millor (Majorca, Western Mediterranean Sea). Different capital letters indicate significant differences between bottom types (post hoc multiple pairwise comparison of mean ranks, $p < 0.05$). Error bars show 95% confidence intervals. N values indicate the number of points selected randomly in each bottom type. Differences in N values between bottom types are driven by the differences in percentage covers of each bottom type in the study area. *P.*, *Posidonia*.

Frederiksen et al. (2004) used aerial photographs to follow changes in the distribution of *Zostera marina* from 1954 to 1999 and showed that seagrass landscapes can change extensively over long periods of time, especially in the more wave-exposed areas. Comparison of aerial photographs taken in 1956 and 2004 indicates that the upper depth limit of *Posidonia oceanica* in Cala Millor has regressed in the south part of the beach (IMEDEA 2005), which is a sector of the beach well exposed to the most energetic waves (those from the N–NNE). We do not rule out, however, that other processes may determine the upper depth limit of *P. oceanica*. This seagrass species is able to grow in certain locations, usually sheltered, almost to sea level (Ribera et al. 1997), which suggests that neither photo-inhibition nor temperature fluctuations associated with shallow depths are important in setting the upper depth limit. Except for the desiccation effects associated with exposure to air at low tides (negligible in the Mediterranean Sea), the consensus is that the upper depth limit of seagrasses is set by the energy of currents and waves (Koch et al. 2006). Wave breaking might have a role in establishing the upper limit, but the position at which waves break will change depending on wave height, wave period, and presence of submerged sandbars. In addition, the effect of waves of certain energy may depend on the amount of time the plants are exposed to different levels of wave energy. Average wave conditions have been used to estimate the average field of near-bottom orbital velocities, but other wave statistics could be used, such as extreme events (De Falco et al. 2008), and elaborate a prediction of those velocities that allow the long-term persistence of *P. oceanica*. We used 44 years of data, as this would be an adequate period of time over which to average the effects of waves on the slow-growing *P. oceanica* (Marbà and Duarte 1998).

We note that sediment density and grain size characteristics have not been considered in this study. These

granulometric characteristics may play an important role in determining the location of seagrasses, because wave energy will move and resuspend the sediment, which might bury or erode a seagrass meadow. Differences in sediment characteristics between locations will then contribute to the variability of the estimate of the threshold wave energy that sets the upper depth limit of *Posidonia oceanica*.

We conclude that the methodology presented here is a useful tool to estimate wave energy on the bottom and to identify the level that sets the upper depth limit of *Posidonia oceanica* meadows. Our approach provides testable estimates of the threshold level of wave energy that allows this species to persist and indicates a research program to validate them. This study also highlights the importance of interdisciplinary and multidisciplinary research in ecological modeling and, in particular, the need for hydrodynamical studies to elucidate the distribution of seagrasses at shallow depths.

Acknowledgements

E. Infantes would like to thank the Spanish Ministerio de Educación y Ciencia (MEC) FPI scholarship program (BES-2006-12850) for financial support. Financial support from grants CTM2005-01434, CTM2006-12072/MAR, and the Integrated Coastal Zone Management (UGIZC) project are acknowledged. The authors would like to thank Puertos del Estado for the HIPO-CAS dataset and Marta Fuster for assistance with the data.

References

- Bell, S.S., B.D. Robbins and S.L. Jensen. 1999. Gap dynamics in a seagrass landscape. *Ecosystems* 2: 493–504.
- Cañellas, B., A. Orfila, F.J. Méndez, M. Menéndez and J. Tintore. 2007. Application of POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). *J. Coast. Res.* 50: 329–333.

- Castillo, M.C., A. Baquerizo and M.A. Losada. 2005. Temporal and spatial statistical variability of the wave height in the surf zone. *29th International Conference on Coastal Engineering, Lisbon, Portugal, 2004*: 997–1008.
- Chivers, R., N. Emerson and D. Burns. 1990. New acoustic processing for underway surveying. *Hydrogr. J.* 56: 8–17.
- Dean, R.G. and R.A. Dalrymple. 2002. *Coastal processes with engineering applications*. Cambridge University Press, Cambridge. pp. 475.
- De Falco, G., M. Baroli, A. Cucco and S. Simeone. 2008. Intra-basinal conditions promoting the development of a biogenic carbonate sedimentary facies associated with the seagrass *Posidonia oceanica*. *Continental Shelf Res.* 28: 797–812.
- Dennison, W.C. 1987. Effects of light on seagrass photosynthesis, growth and depth distribution. *Aquat. Bot.* 27: 15–26.
- Dennison, W.C. and R.S. Alberte. 1985. Role of daily light period in the depth distribution of *Zostera marina* (eelgrass). *Mar. Ecol. Prog. Ser.* 25: 51–61.
- Duarte, C.M. 1991. Seagrass depth limits. *Aquat. Bot.* 40: 363–377.
- Duarte, C.M., J. Terrados, N.S.R. Agawin, M.D. Fortes, S.S. Bach and W.J. Kenworthy. 1997. Response of a mixed Philippine seagrass meadow to experimental burial. *Mar. Ecol. Prog. Ser.* 147: 285–294.
- Folkard, A.M. 2005. Hydrodynamics of model *Posidonia oceanica* patches in shallow water. *Limnol. Oceanogr.* 50: 1592–1600.
- Fonseca, M.S. and J.S. Fisher. 1986. A comparison of canopy friction and sediment movement between 4 species of seagrass with reference to their ecology and restoration. *Mar. Ecol. Prog. Ser.* 29: 15–22.
- Fonseca, M.S. and W.J. Kenworthy. 1987. Effects of current on photosynthesis and distribution of seagrasses. *Aquat. Bot.* 27: 59–78.
- Fonseca, M.S. and J.A. Cahalan. 1992. A preliminary evaluation of wave attenuation by 4 species of seagrass. *Est. Coast. Shelf Sci.* 35: 565–576.
- Fonseca, M.S. and S.S. Bell. 1998. Influence of physical settings on seagrass landscapes near Beaufort, North Carolina, USA. *Mar. Ecol. Prog. Ser.* 171: 109–121.
- Fonseca, M.S., J.S. Fisher, J.C. Zieman and G.W. Thayer. 1982. Influence of the seagrass, *Zostera marina* L., on current flow. *Est. Coast. Shelf Sci.* 15: 351–364.
- Fonseca, M.S., J.C. Zieman, G.W. Thayer and J.S. Fisher. 1983. The role of current velocity in structuring eelgrass (*Zostera marina* L.) meadows. *Est. Coast. Shelf Sci.* 17: 367–380.
- Fonseca, M.S., P.E. Whitfield, N.M. Kelly and S.S. Bell. 2002. Modeling seagrass landscape pattern and associated ecological attributes. *Ecol. Appl.* 12: 218–237.
- Fornes, A., G. Basterretxea, A. Orfila, A. Jordi, A. Álvarez and J. Tintore. 2006. Mapping *Posidonia oceanica* from IKONOS. *J. Photogramm. Remote Sens.* 60: 315–322.
- Frederiksen, M., D. Krause-Jensen, M. Holmer and J.S. Laursen. 2004. Spatial and temporal variation in eelgrass (*Zostera marina*) landscapes: influence of physical setting. *Aquat. Bot.* 78: 147–165.
- Gacia, E. and C.M. Duarte. 2001. Sediment retention by a Mediterranean *Posidonia oceanica* meadow: the balance between deposition and resuspension. *Est. Coast. Shelf Sci.* 52: 505–514.
- Gambi, M.C., A.R.M. Nowell and P.A. Jumars. 1990. Flume observations on flow dynamics in *Zostera marina* (eelgrass) beds. *Mar. Ecol. Prog. Ser.* 61: 159–169.
- GIOC. 2003. *Spectral wave propagation model (Oluca-SP)*. State Coastal Office, Spanish Environmental Ministry and University of Cantabria, Santander. pp. 170 (in Spanish).
- Gómez-Pujol, L., A. Orfila, B. Canellas, A. Alvarez-Ellacuría, F.J. Mendez, R. Medina, et al. 2007. Morphodynamic classification of sandy beaches in low energetic marine environment. *Mar. Geol.* 242: 235–246.
- Greve, T. and D. Krause-Jensen. 2005. Predictive modelling of eelgrass (*Zostera marina*) depth limits. *Mar. Biol.* 146: 849–858.
- Hine, A.C., M.W. Evans, R.A. Davis and D.F. Belknap. 1987. Depositional response to seagrass mortality along a low-energy, barrier-island coast: west-central Florida. *J. Sediment. Petrol.* 57: 431–439.
- Hovel, K.A., C.R. Fonseca, D.L. Myer, W.J. Kenworthy and P.E. Whitfield. 2002. Effects of seagrass landscape structure, structural complexity and hydrodynamic regime on macrofaunal densities in North Carolina seagrass beds. *Mar. Ecol. Prog. Ser.* 243: 11–24.
- IMEDEA. 2005. Variabilidad y dinámica sedimentaria de las playas de Cala Millor y Cala Sant Vicent. Volumen 1. Govern de les Illes Balears, Cala Millor. pp. 103.
- Keddy, P.A. 1984. Quantifying a within-lake gradient of wave energy in Gillfillan Lake, Nova-Scotia. *Can. J. Bot.* 62: 301–309.
- Kenworthy, W.J. and M.S. Fonseca. 1996. Light requirements of seagrasses *Halodule wrightii* and *Syringodium filiforme* derived from the relationship between diffuse light attenuation and maximum depth distribution. *Estuaries* 19: 740–750.
- Kirby, J.T. and R.A. Dalrymple. 1983. The propagation of weakly nonlinear waves in the presence of varying depth and currents. In: *Proceedings of the XXth Congress I.A.H.R., Moscow 7, Moscow, September, 1983*: 198–202.
- Koch, E.W. and S. Beer. 1996. Tides, light and the distribution of *Zostera marina* in Long Island Sound, USA. *Aquat. Bot.* 53: 97–107.
- Koch, E.W. and G. Gust. 1999. Water flow in tide- and wave-dominated beds of the seagrass *Thalassia testudinum*. *Mar. Ecol. Prog. Ser.* 184: 63–72.
- Koch, E.W., J.D. Ackerman, J. Verduin and M. van Keulen. 2006. Fluid dynamics in seagrass ecology – from molecules to ecosystems. In: (A.W.D. Larkum, R.J. Orth, C.M. Duarte, eds). *Seagrasses: biology, ecology and conservation*. Springer, Dordrecht. pp. 193–224.
- Krause-Jensen, D., M.F. Pedersen and C. Jensen. 2003. Regulation of eelgrass (*Zostera marina*) cover along depth gradients in Danish coastal waters. *Estuaries* 26: 866–877.
- Leriche, A., V. Pasqualini, C.F. Boudouresque, G. Bernard, P. Bonhomme, P. Clabaut, et al. 2006. Spatial, temporal and structural variations of a *Posidonia oceanica* seagrass meadow facing human activities. *Aquat. Bot.* 84: 287–293.
- Marbà, N. and C.M. Duarte. 1998. Rhizome elongation and seagrass clonal growth. *Mar. Ecol. Prog. Ser.* 174: 269–280.
- McKenzie, L.J., M.A. Finkbeiner and H. Kirkman. 2001. Methods for mapping seagrass distribution. In: (F.T. Short and R.G. Coles, eds). *Global seagrass research methods*. Elsevier, Amsterdam. pp. 101–121.
- Orfila, A., A. Jordi, G. Basterretxea, G. Vizoso, N. Marba, C.M. Duarte, et al. 2005. Residence time and *Posidonia oceanica* in Cabrera Archipelago National Park, Spain. *Continental Shelf Res.* 25: 1339–1352.
- Orlowski, A. 1984. Application of multiple echo energy measurements for evaluation of sea bottom type. *Oceanologia* 19: 61–78.
- Pasqualini, V., C. Pergent-Martini, P. Clabaut and G. Pergent. 1998. Mapping of *Posidonia oceanica* using aerial photographs and side scan sonar: application off the Island of Corsica (France). *Est. Coast. Shelf Sci.* 47: 359–367.
- Preen, A.R., W.J. Lee Long and R.G. Coles. 1995. Flood and cyclone related loss, and partial recovery, of more than 1000 km² of seagrass in Hervey Bay, Queensland, Australia. *Aquat. Bot.* 52: 3–17.
- Procaccini, G., M.C. Buia, M.C. Gambi, M. Perez, G. Pergent, C. Pergent-Martini and J. Romero. 2003. The seagrasses of the Western Mediterranean. In: (E.P. Green, F.T. Short, eds). *World atlas of seagrasses, UNEP World Conservation Monitoring Centre*. University of California Press, Berkeley, CA. pp. 48–58.

- Ratsimandresy, A.W., M.G. Sotillo, J.C. Carretero Albiach, E. Álvarez Fanjul, H. Hajji. 2008. A 44-year high-resolution ocean and atmospheric hindcast for the Mediterranean Basin Developer within the HIPOCAS Project. *Coast. Eng.* 55: 825–826.
- Ribera, G., M. Coloreu, C. Rodríguez-Prieto and E. Ballesteros. 1997. Phytobenthic assemblages of Addaia Bay (Menorca, Western Mediterranean): composition and distribution. *Bot. Mar.* 40: 523–532.
- Robertson, A.I. and K.H. Mann. 1984. Disturbance by ice and life-history adaptations of the seagrass *Zostera marina*. *Mar. Biol.* 80: 131–141.
- Shepherd, S.A. and H.B.S. Womersley. 1981. The algal and seagrass ecology of Waterloo Bay, South-Australia. *Aquat. Bot.* 11: 305–371.
- Soares, C.G., J.C. Carretero, R. Weisse and E. Alvarez. 2002. A 40 years hindcast of wind, sea level and waves in European waters. In: *Proceedings of OMAE 2002: 21st International Conference on Offshore Mechanics and Artic Engineering, Oslo, Norway 2000*: 1–7.
- Terrados, J. and C.M. Duarte. 2000. Experimental evidence of reduced particle resuspension within a seagrass (*Posidonia oceanica*) meadow. *J. Exp. Mar. Biol. Ecol.* 243: 45–53.
- van Katwijk, M.M. and D.C.R. Hermus. 2000. Effects of water dynamics on *Zostera marina*: transplantation experiments in the intertidal Dutch Wadden Sea. *Mar. Ecol. Prog. Ser.* 208: 107–118.
- van Katwijk, M.M., D.C.R. Hermus, D.J. de Jong, R.M. Asmus and V.N. de Jonge. 2000. Habitat suitability of the Wadden Sea for restoration of *Zostera marina* beds. *Helgol. Mar. Res.* 54: 117–128.
- Williams, S.L. 1988. Disturbance and recovery of a deep-water Caribbean seagrass bed. *Mar. Ecol. Prog. Ser.* 42: 63–71.
- Worcester, S.E. 1995. Effects of eelgrass beds on advection and turbulent mixing in low current and low shoot density environments. *Mar. Ecol. Prog. Ser.* 126: 223–232.
- Wright, L.D. and A.D. Short. 1983. Morphodynamics of beaches and surf zones in Australia. In: (P.D. Koma, ed). *Handbook of coastal processes and erosion*. CRC Press, Boca Raton, FL. pp. 35–64.

Received 26 February, 2008; accepted 23 September, 2008; online first 13 February, 2009