UNIVERSITAT POLITÈCNICA DE CATALUNYA

BARCELONA TECH

Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports de Barcelona

Laboratori d'Enginyeria Marítima

Program of Doctorate of Philosophy in Civil Engineering

Multivariate characterization of wave storms in coastal areas

Ph.D. thesis as a compendium of publication

PhD Thesis presented by

Jue LIN-YE

for the degree of DOCTOR OF PHILOSOPHY IN CIVIL ENGINEERING

Barcelona, March, 2018

Laboratori d'Enginyeria Marítima

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Supervisor

Dr. Vicente Gràcia Garcia

To my beloved ones.

"Dilegua, o notte!... Tramontate, stelle!... "–*Turandot*, Puccini

Contents

List of Figures

List of Tables

Symbol Index

Abbreviatures

EA	East Atlantic pattern
GCM	General circulation model
GPD	Generalized Pareto distribution
HAC	Hierarchical Archimedean copula
movMF	Mixture of von Mises-Fisher distribution functions
NAO	North Atlantic oscillation
NW	Northwestern
PACF	Partial AutoCorrelation Function
RCM	Regional circulation model
RCP	Representative concentration pathway
SC	Scandinavian (atmospheric) pattern
SWAN	Simulating WAves Nearshore (spectral wave-model)
VGAM	Vectorial generalized additive model
VGLM	Vectorial generalized linear model

Variables

θ_p^*	Wave-direction at the peak of the wave-storm
E_u	Unitary energy at the peak of the wave-storm
H_p	Significant wave-height at the peak of the wave-storm
T_p	Peak wave-period at the peak of the wave-storm
Tr	Return period
D	Duration of the wave-storm
E	Total energy of the wave-storm

Abstract

Wave-storms are the responsible of the main changes in the Coast. Their detailed characterization results in a better design of any marine structure. The most common approach to describe wave-storms is to simplify the event by taking the significant wave-height (H_p), peak period (T_p) and direction (θ_p^*) at the peak of the storm and treating these variables independently. However, it is well accepted that some relationship should exist between them. What is more, the development of sophisticated numerical models in Coastal Engineering are demanding additional variables such as the duration of the wave-storm (D), the amount of associated energy (E), temporary evolution of the variables and their relationship to atmospheric climate-indices, to better reproduce the simulated processes.

The main objectives and results of this thesis are as follows. First, wave-storms in the present wave climate of the Catalan Coast are characterized, assuming stationarity. The wave-storm variables modelled are: the energy at the peak of the wave-storm (E_u), T_p, E, D, θ_p^* and the proportions of time from the start to the storm peak and from the storm peak to the end (growth-decay rates). E, E_u, T_p and D are fit by generalized Pareto distributions (GPD). Their joint probability structure is characterized by a hierarchical Archimedean copula (HAC). θ_p^* is characterized through a mixture of von Mises-Fisher probability distribution functions and related to E, T_p and D through a multinomial logistic model. The growth-decay rates are related to D through third degree polynomials. A triangle and an irregular-trapezoid are proposed to model the wave-storm shape.

In the present climate of the Catalan Coast, the constructed statistical model can serve to generate synthetic wave-storms. The most predominant θ_p^* are north and east. The most appropriate geometric figure to describe the evolution of the wave-height is a irregular-trapezoid. For D over 100h, the peak of the wave-storm is generally closer to the end of the wave-storm than to the start.

After establishing a stationary model, non-stationarity is incorporated into the characterization of wave-storms in the Catalan Coast. E, H_p, T_p and D are characterized through non-stationary GPDs. The wave-storm threshold, the wave-storminess and the parameters of the GPDs are related to North Atlantic Oscillation (NAO), East Atlantic pattern (EA) and Scandinavian pattern (SC) and their first two time-derivatives, through Vectorial generalized additive models. The joint probability structure is characterized by a pseudo-time-dependent HAC. A severe greenhouse gas emission scenario is considered.

The mean values of all wave-storm variables decrease in the 21^{st} century, except for D in the northern part of the coast. A negative NAO may cause an increase

in wave-storminess; the wave-storm threshold and the GPD parameters are most influenced by the dynamics of the climate-patterns, rather than by the climate-patterns themselves.

The non-stationary methodology is repeated in the northwestern Black Sea, considering both a mild and a severe emission scenarios. Here, wave-storminess is not affected by the proposed climate-patterns, whereas the wave-storm threshold is strongly influenced by SC and EA. The average value of the wave-storm variables seem to have a more positive trend than in the Catalan Coast, and it is observed that an increase in mean values is related to an increase in variance. SC and EA also strongly influence the parameters of the GPDs.

In the two study areas, the dependence between E and D is high, while the general dependence among the wave-storm variables is medium. In the Catalan Coast, it is expected that the dependence between E and D should increase with time. In the northwestern Black Sea, it is the dependence among all the wave-storm variables that increases with time, in both emission scenarios; the severe emission scenario presents less dependence among wave-storm variables.

Resumen

Los temporales de mar provocan gran parte de las cambios a la costa. Sus caracterización detallada tiene como consecuencia un mejor diseño de cualquier estructura marina. El enfoque más común para describir los temporales de mar es simplificar este suceso tomando la altura de ola significante (H_p), el periodo pico (T_p) y la dirección (θ_p^*) en el pico de la tormenta y tratándolas de forma independiente. Sin embargo, está ampliamente aceptado que exista al menos alguna relación entre ellas. Es más, el desarrollo de sofisticados modelos numéricos en la Ingeniería de Costas pide variables adicionales como la duración de tormenta (D), la cantidad de energía asociada (E), la evolución temporal de las variables y su relación con índices climáticos atmosféricos, todo para una mejor reproducción de los procesos simulados.

Los objetivos y resultados principales de esta tesis son lo siguiente. Primero, se caracteriza tormentas de mar en el clima de oleaje presente, de la costa catalana, suponiendo estacionalidad. Las variables modeladas son: la energía unitaria en el pico del temporal (E_u), T_p, E, D, θ_p^* y la proporción de tiempo desde el inicio hasta el pico y desde el pico al final del temporal (ratios de crecimiento-decrecimiento). Se caracteriza E, E_u, T_p y D con distribuciones generalizadas de Pareto (GPD), y se caracteriza la estructura de probabilidades conjunta de estas variables vía una cópula jerárquica arquimedeana (HAC). Se caracteriza θ_p^* con una combinación de distribución de probabilidad de von Mises-Fisher y se le relaciona con E, T_p y D a través de un modelo logístico multinomial. Se propone una forma triangular o trapezoide-irregular para modelar la forma del temporal.

En el clima presente de la coasta catalana, el model estadístico construido puede generar temporales sintéticos. Las θ_p^* principales son el norte y el este. La figura geométrica que mejor describe la evolución de la altura de ola es un trapezoide irregular. Para D mayor que 100h, el pico del temporal está generalmente más cerca del final que del principio. La media de cada variable decrece en el siglo 21, excepto la de D, en el norte de la costa. Una NAO negativa puede causar una subida de la tormentosidad. Además, el umbral de tormenta y los parámetros de GPD están influenciados principalmente por la dinámica de los patrones climáticos, en vez de serlo por los propios patrones climáticos.

Después de establecer un modelo estacionario, se incorpora la no estacionalidad a la caracterización de temporales de mar en la costa catalana. Se caracteriza E, H_p, T_p y D con GPDs no estacionarios. El umbral de temporal, la tormentosidad y los parámetros de los GPDs están relacionados con la Oscilación de Atlántico norte (NAO), el Patrón de Atántico oriental (EA) y el Patrón escandinavo (SC)

y sus primeras dos derivadas temporales, a través de Modelos aditivos generalizados vectoriales. Se caracteriza la estructura de probabilidades conjunta con un HAC pseudo-dependiente del tiempo. Se considera un escenario grave de Cambio climático.

Se repite la metodología no estacionaria en el noroeste del Mar negro, considerando tanto un escenario suave de Cambio climático como otro grave. En el noroeste del Mar negro, la tormentosidad de mar no está afectada por los patrones climáticos propuestos, todo y que el umbral de temporal está fuertemente influenciada por SC y EA. Los valores medios de las variables de temporal parecen tener una tendencia más positiva que en la costa catalana, y se observa que una subida de los valores medios se relaciona con otra subida de las varianzas. SC y EA afectan fuertemente a los parámetros de los GPDs.

En las dos zonas de estudio, la dependencia entre E y D es alta, mientras que la dependencia general entre las variables de temporal es media. En la costa catalana, se espera que la dependencia entre E y D crezca con el tiempo. En el noroeste del Mar negro, es la dependencia entre todas las variables de temporal la que crece con el tiempo, en ambos escenarios de Cambio climático; el escenario grave presenta menos dependencia entre las variables.

Acknowledgments

I would like to sincerely thank my advisor, Dr. Vicenç Gràcia, for his invaluable experience as a scientific researcher and his constant battle to provide a neophyte like me with funding (FP7-ENV-603396 and ECHO/SUB/2013/661009 European Projects, CTM2013-45141-R and CTM2017-88036-R Spanish National Projects). Prof. Agustín Sánchez-Arcilla, for believing in me and including me in the H2020-EO-2016-730030-CEASELESS European project. The other professors and staff of the same centre, Dr. Manuel Espino, Dr. Francesc Xavier Gironella, Prof. José Antonio Jiménez, Dr. Marc Mestres, Dr. César Mösso, Prof. Joan Pau Sierra, Mr. Joaquim Sospedra, Dr. Daniel González, Ms. Genoveva Comas, Ms. Emilia Baños, Ms. Marisol Esplandiu, Ms. Alba Serra and Ms. Carmen García.

Dr. Adrian Stanica and Dr. Irina Dinu, from GeoEcoMar had been most helpful in receiving me in their centre for a visit during summer of 2016. As for expertise on Statistical aspects of the thesis, I would like to kindly mention Dr. Maribel Ortego; Prof. Juan José Egozcue, for remarks on mathematical concepts; Dr. Ostap Okhrin and Dr. Alexander Ristig for early correspondence on hierarchical Archimedean copulas; Dr. Bettina Grün, Dr. Peter Jupp and Dr. Alan Gelfand for early correspondence or talk on directional statistics.

I would also like to thank Prof. Piero Lionello for providing data for the second article of this thesis. Additionally, the support of Puertos del Estado, Institut Cartogràfic i Geològic de Catalunya (ICGC), European Centre for Medium-Range Weather Forecasts (ECMWF) and the Copernicus Marine Environment Monitoring Service has been essential to made this dissertation possible.

I would like to thank both my family of origin (father, mother, Jiji, my grandparents) and my in-laws (Mr. Manuel García Gamarra, Ms. Rosario León and Ms. Roser García). I deeply appreciate the intellectual environment created by personnel from the same university, such as Dr. Anna Mujal, Yanxin Mao, Heng Yang, Xiao Lin, Dr. Dalila Sánchez-Escobedo, among others.

I would like to express special gratefulness to my husband, Manuel García León. Appart from the help and affection in the personal aspects of the thesis, he has also provided technical support. He has the role of the postdoctoral fellow, only that he has done it while still being a Ph.D. candidate. He has shown a professional and mental maturity far beyond what is expected from a Ph.D. candidate. I wish him all the best for his forthcoming graduation.

1. Introduction

Human populations exhibit a predisposition to occupy the coastal fringe. The projections for population at the coastal zone show more demographic growth than inland areas (Merkens et al. (2016); Neumann et al. (2015)). There is a variety of reasons for this phenomenon: compared to continental regions, coastal areas present benefits such as a milder weather, better communications and cheaper means of transportation, by means of maritime transportation, as well as fishery as a resource of food (Small and Nicholls (2003)). In some countries, the sea water is even a resource of fresh water for daily use. The coast is often a vulnerable environment, which natural ecosystem needs constant care and protection against natural hazards such as wave storms. Other elements on the coast, such as harbours and promenades, are also prone to destruction by the same hazard. The action of wave storms require special attention, for their severity (Sánchez-Arcilla et al. (2008b)). However, due to their scarcity in number, they are not well understood, which lead to the fact that they are not easily predictable.

A first question to ask oneself should be "what is an extreme event?" (Embrechts et al. (1997)). They are those episodes that are extraordinary in nature, based on the criteria of one indicator. In engineering, the importance of wave-storms are remarked as they are the design values to be used for infrastructure. However, this argument only holds valid in a world with a single stressor (the hydrodynamic forcing) or within a stationary framework. The actual assessment should be multivariate, which grows complicated when the variables have joint dependence of non-linear nature. Also, Climate Change make it necessary to consider a non-stationary framework for future wave climate (Wang et al. (2015a)). These two approaches are explored in this thesis.

1.1. Motivation

Wave storms strongly perturb the state of coastal environments. Such changes become concomitant, with episodic coastal hazards such as marine flooding and erosion. These extreme phenomena drive complex hydrodynamic processes, the understanding of which is basic for proper infrastructure design (Goda (2010)). Statistical characterization of storms is required to provide high quality boundary conditions for the increasingly need for assessing and forecasting the impacts of storms (Sánchez-Arcilla et al. (2014); Gràcia et al. (2013)). The conventional design is usually based on the probabilistic definition of a single parameter, typically the

wave height. Other wave-storm variables, like duration of the storm, storm total energy and associated wave period can all influence the final response of a beach or the progression of the damage of a structure (Martín-Soldevilla et al. (2015); Melby and Kobayashi (2011)). These variables are atually mutualy dependent (Salvadori et al. (2007); de Waal and van Gelder (2005)), but the classical methodology either assumes one variable to be stochastic and the other ones to be deterministic or assumes all variables to be stochastic but completely independent.

A common modelling approach is to hindcast high energy events or to synthesize storms to a representative extreme sea-state, which is generally predisposed by the degree of knowledge of the area. For the latter case, dependency structures among the hydrodynamic variables pose a hurdle, because they are usually unknown. Exploratory methods such as 2D scatter plots, have been widely used as a rule-of-thumb for the most frequent problem, wave height vs. wave period. However, the interpretation of existing co-dependences among several variables becomes more difficult, disregarding their mutual dependences. Hard to interpret are those cases in which a wide scatter cloud can mislead biased co-dependence structures, due to subjective criteria. Storm modelling requires to consider a multivariate analysis of storm parameters (Corbella and Stretch (2012)), as univariate analyses may oversimplify coastal processes, thus leading to overestimation of the storm induced damages.

Another aspect of concern is that wave-storms may eventually be exacerbated as a consequence of Climate Change. Such changes in the local climates may affect patterns in wave-storminess, which has a significant role in enhancing the destruction of dunes, transportation of sediments beyond the surf-zone to unusual depth or burial of benthic biota (Wang et al. (2015b); Hemer and Trenham (2016)). Climate Change would also intensify wave-storms, conditioning on several coastal hazards: flooding (Hinkel et al. (2014); Wahl et al. (2016)), erosion (Hinkel et al. (2013); Casas-Prat et al. (2016); Li et al. (2014); Valchev et al. (2010)), harbour agitation (Sánchez-Arcilla et al. (2016); Sierra et al. (2015)) and overtopping (Sierra et al. (2016)). The stationary assumption, key pillar in the last decades for designing infrastructures, does not hold valid for Climate Change impact studies. Hence, there is a pressing urge for methodologies that consider non-stationarity in wave storm variables. Not only to model average values, but also the variability of such variables and the joint probability structure should be addressed.

1.2. Objectives

The main objective of this thesis is to characterize the wave-storm intensity variables in two fetch-limited, micro-tidal environments: the Catalan Coast (northwestern Mediterranean Sea) and the northwestern Black Sea. This characterization leads to building multivariate models that consider the non-stationary joint probability structure of the wave-storm intensity variables. The non-stationarity help include changes in the climatic trend. The non-stationary models also establish

the relationship between the weave-storm intensity variables and three atmospheric climate-patterns. These models can be used to define hydrodynamic loads for the design of coastal infrastructures, under present and future climate.

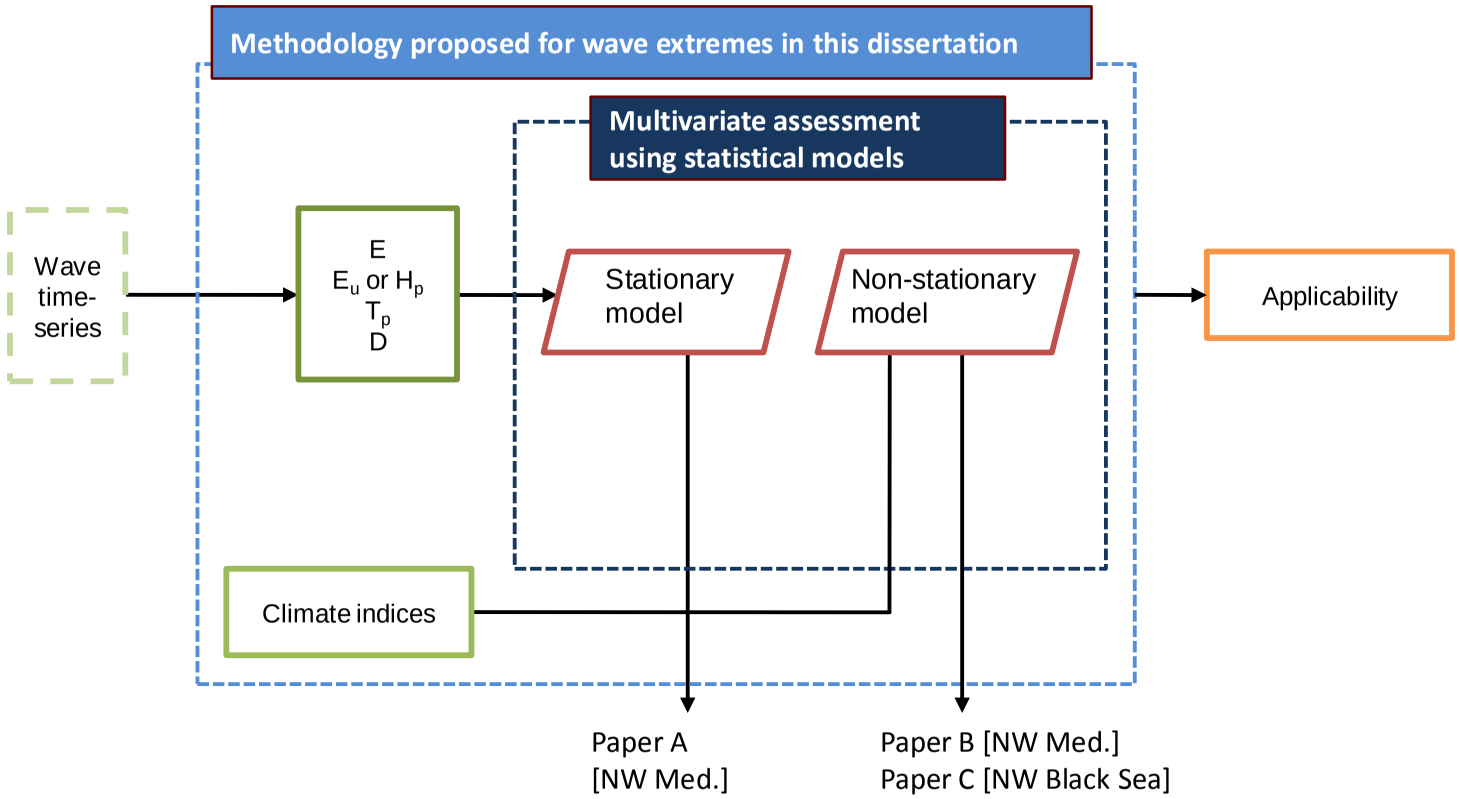
The specific objectives, per study area, are as follows:

1. In the Catalan Coast, there are two main groups of tasks.
    a) The stationary approach deals with wave-storms in the present climate of the Catalan Coast. The detailed tasks are:
        i. To characterize wave-storm intensity variables with probability distribution function.
        ii. To propose a statistical tool to characterize their joint probability structure.
        iii. To characterize the wave-directionality of the wave-storm and to relate it to the wave-storm intensity variables.
        iv. To relate wave-storm intensity variables to the intra-time variables of the wave-storm.
    b) The non-stationary approach deals with future wave-storms in the Catalan Coast. It uses wave projections for the years 1950-2100, considering a severe greenhouse gas emission scenario. The detailed tasks are:
        i. To characterize wave-storm intensity variables through non-stationary probability distribution function.
        ii. To propose a pseudo-non-stationary characterization of their joint probability structure.
        iii. To establishes the relationship of the wave-storm intensity variables to climate-indices.
2. In the northwestern Black Sea, the non-stationary approach is repeated . It deals with projections in the same period of 1950-2100. It considers both a mild and a severe emission scenario.

1.3. Outline of the thesis

This thesis is the presentation of the work reflected in the flow-chart of Fig. 1.1. The thesis consists of an introduction (chapter 1), a presentation of the study areas (chapter 2), a presentation of the methodology (chapter 3), a compendium of three published scientific papers (chapter 4 through chapter 7), a general results and discussion (chapter 8), a conclusion (chapter 9) and a list of proposed future works (chapter 10).

The compendium of papers comprises a total of three articles: chapter 4 proposes a stationary characterization of wave-storms in the Catalan Coast; chapter 5 and chapter 6 propose a non-stationary characterization of wave-storms in the Catalan Coast; and chapter 7 applies the methodology in chapter 5 to the northwestern Black Sea. The bibliographical information of the three papers is:



(a)

Figure 1.1.: Flow-chart of the thesis. The green rectangular boxes represent data. The green discontinuous box stands for data produced outside of the thesis but which is used for the model. The red rhombuses are the models built for the thesis, the orange rectangle is the output and the blue discontinuous boxes integrate several parts of the thesis.

[Lin-Ye et al. 2016] J. Lin-Ye, M. García-León, V. Gràcia, A. Sánchez-Arcilla (2016). Multivariate statistical model of extreme events: An application to the Catalan coast, *Coastal Engineering*, 117, 138–156

[Lin-Ye et al. 2017] J. Lin-Ye, M. García-León, V. Gràcia, M.I. Ortego, P. Lionello, A. Sánchez-Arcilla (2017). Multivariate statistical modelling of future marine storms, *Applied Ocean Research*. 65, 192–205

[Lin-Ye et al. 2018] J. Lin-Ye, M. García-León, V. Gràcia, M.I. Ortego, A. Stanica, A. Sánchez-Arcilla (2017). Multivariate hybrid modelling of future wavestorms at the northwestern Black Sea, *Water*. 10 (2)

2. Study areas

The proposed analysis via multivariate non-stationary methodologies has been applied both to the Catalan Coast and the northwestern Black Sea. These are two fetch-limited and micro-tidal environments. This common feature set ground to characterize their wave-storms and to compare them to each other.

2.1. Catalan Coast

The Mediterranean Sea (see Fig. 2.1) is of semi-enclosed nature and is dominated by its orographic patterns, the air-sea temperature differences and the passage of low pressure centres from the Atlantic (Lionello et al. (2012)). The main morphological features at the Catalan Coast are the existence of mountain chains parallel and close to the coast, the Pyrenees Mountains to the north, and the Ebre river valley to the south; these orographic discontinuities, along with the major river valleys, let strong winds flow towards the coast (Grifoll et al. (2015)).

The northwestern Mediterranean coastal winds are typically low to medium, on average, ranging up to $11.05^{m}/_{s}$ (Sánchez-Arcilla et al. (2008c)). The most frequent and intense wind is the northerly Tramuntana, appearing from November to March. It is the major forcing for the northern and central northwestern Mediterranean coast waves. From latitude $41°N$ southward, the principal wind direction is the northwesterly Mistral. It is channelled by the western Pyrenees and the Ebre valley. The Mistral are formed by the superposition of gap and downhill flows from the Pyrenees. A secondary wind mass, the westerly Ponent, comes from the depressions in northern Europe and sweeps across the entire Iberian Peninsula from west to east.

Easterly winds are frequent during the summer. They are commonly triggered by an intense high-pressure area on the British Islands. A different source for this wind is a high level of cold air pool deepening over the Mediterranean Sea, which leads to cyclo-genesis, resulting in the passage of a low off the northwestern Mediterranean coast (Bolaños et al. (2009); Lionello et al. (2012)). Winds are more variable in higher intensities. Thus, some relatively large wind modulus variability is generated during wave-storms (Bolaños (2004)). Winds are the main forcing for waves at the northwestern Mediterranean. Wave directions are directly correlated with wind direction, except the waves that form 50° with the coast, which can be generated by any wind direction. This might be explained by the orientation of the Catalan coast-line. Also, all winds, at some point, are capable of creating an alongshore wave train in the Catalan Coast.

The largest waves come from the east, caused by the joint action of the most significant fetches and winds (Bolaños et al. (2009)). This is especially evident at the southern Catalan Coast. Such directionality is evident, also, for the mean peak wave-period, but not for wave-storm durations. Appart from the fact that the northwestern Mediterranean coast was a micro-tidal environment (Lionello et al. (2012)), the slope of the bathymetry is relatively steep in the north, while it widens in the south. This has a direct impact on how waves behave when approaching the coast, as the bathymetry has an effect on the type of impacting waves and the beach slope determines the vulnerability to floods. Waves on the northwestern Mediterranean coast also have an important effect on sediment transport, as short wave lengths do not help beach sediments to restore itself during summer time. The maximum significant wave-height, $H_{s,max}$, is 0.85m in the nothern part of the Catalan coast, and is 5.48m in the southern part. The maximum peak wave-period, $T_{p,max}$, is 15.87s in the northern half of the Catalan Coast and is 14.1s in the southern half.

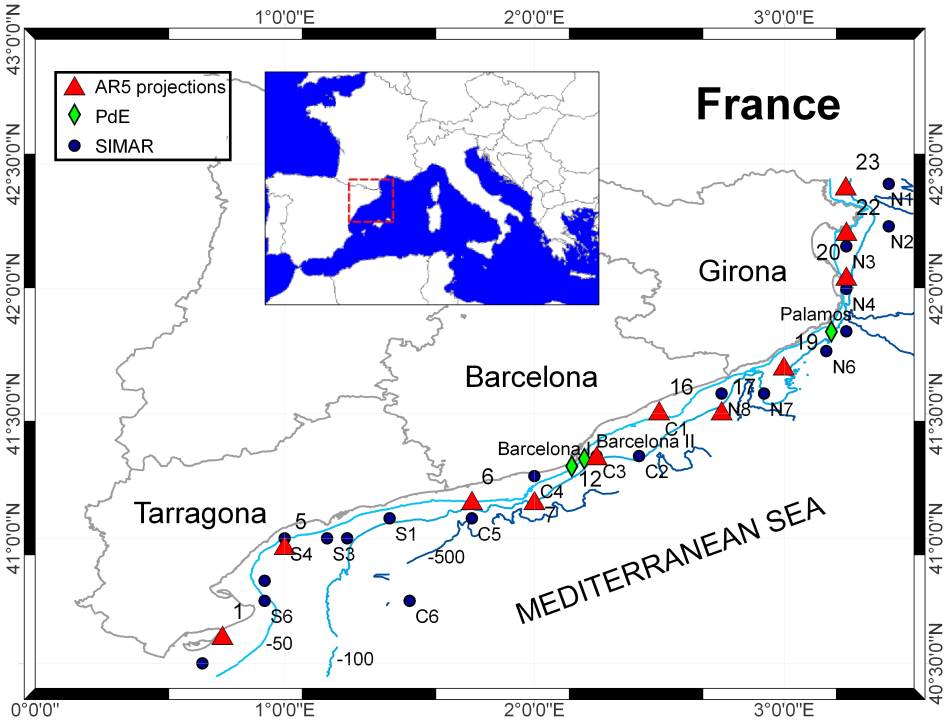


Figure 2.1.: Map of the Catalan Coast (northwestern Mediterranean Sea). The bathymetry is in meters, AR5 nodes are represented by red triangles; Puertos del Estado buoy nodes, by green rhombes; and SIMAR wave model nodes, by solid black points. Source: Lin-Ye et al. (2016)

2.2. Northwestern Black Sea

The Black Sea is located between 41 and 46°N and 27 and 42°E (see Fig. 2.2). It is connected to the Mediterranean Sea through the Turkish straits, formed by the Bosporus strait, the Sea of Marmara, and the Dardanelles strait. The greater part of the sea is a basin with depths exceeding 2000m. However, the western shelf slope is considerably smooth, sparing great depths. The northwestern Black Sea bathes the coasts of Romania and Ukraine, among others, and one important region located here is the Danube Delta, which is notorious for the preservation of its natural resources.

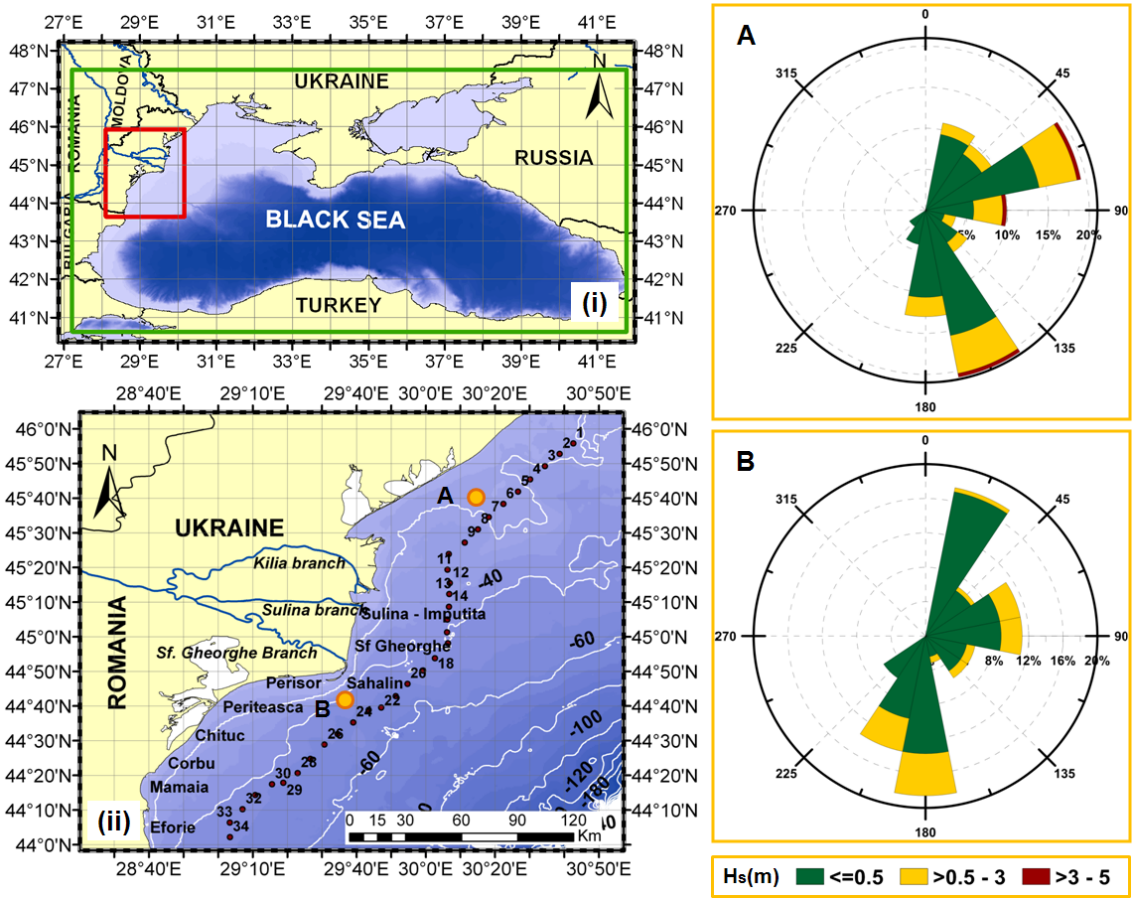


Figure 2.2.: (**i**) Map of the Black Sea. The northwestern Black Sea is enclosed by a red rectangle. (**ii**) Map of the northwestern Black Sea. Nodes from the statistical model are in red and are numbered from north to south. The right figure show wave-roses at points **A** and **B** (orange dots) from the map of the northwestern Black Sea. The bar on the right-bottom shows the wave height-ranges at the wave-rose. Source: Lin-Ye et al. (2017)

A great part of the Black Sea's coast is surrounded by mountains, which are the Balkans, the Pontic Mountains, the Caucasus and the Crimean mountains. This

feature marks the wind patterns in the coastal areas. The general atmospheric circulation in the Black Sea is influenced by the configuration of the Azores and Siberian high-pressure areas and the Asian low-pressure area. Local winds, such as sea breezes, mountain–valley circulation, slope winds, foehns and bora have a considerable impact on the atmospheric circulation pattern, as well.

The most marked feature of wind and wave climate in the northwestern Black Sea is a significant seasonal variability. During cold seasons, the most relevant configuration is determined by the relative position, displacement and resulting interactions between the Mediterranean cyclones and the Eastern European (Siberian) anticyclone (Valchev et al. (2010)). The most intense and frequent winds affecting the coast are those from the northeast, east and southeast. They have the largest fetch, thus leading to the most severe wave-storms.

The average maximum wave heights in the western Black sea is 6.2m, whereas the average mean wave heights is 0.2m. Following the wind pattern, waves propagate most frequently from the east, northeast and southeast. The easterly waves are predominant within the entire shelf zone. Their probability of occurrence ranging between 30% and 40%. The fraction of northeastern waves have a probability of occurrence of 30% and the fraction of southeastern waves have a probability of occurrence of over 10% (Valchev et al. (2010)).

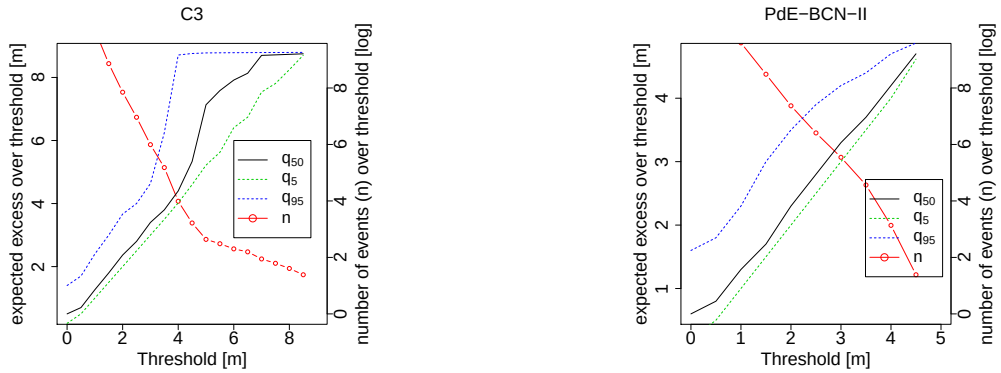
3. Methodology: main steps

This chapter is organized as follows: 1) storm definition and dataset (sec. 3.1) and 2) model formulation (sec. 3.2).

3.1. Storm definition and dataset

This section is organized as follows: a) stationary model at the Catalan Coast, b) General circulation models and atmospheric climate-patterns, c) non-stationary model at the Catalan Coast, d) non-stationary model at the northwestern Black Sea and e) observational data used for validation.

3.1.1. Storm definition and dataset for the stationary model in the Catalan Coast



(a) **(b)**

Figure 3.1.: Mean-excess-plot of H_{m0} for the C3 and PdE-BCN-II nodes, which are a pair of model and buoy nodes, central to the Catalan Coast. The red line represents the number of events over a given threshold, while q_{50}, q_5 and q_{95} are the H_{m0} quantiles. The horizontal and left-vertical axes are in log-scale (in meters). Source: Lin-Ye et al. (2016)

The data used are SIMAR hindcasted waves (Gomez and Carretero (2005)) for the years 1996-2013. The data consists of wave hindcast simulations by WAM (WAMDI Group et al. (1988)) and WAVEWATCH3 (Tolman (2009)), fed by HIRLAM

wind fields (Unden et al. (2002)). SIMAR provides consistent, gapless and spatially dense time-series. A wave-height threshold is selected to separate the wave-storms from the wave regimes in regular conditions. One of the characteristics of wave-storms is that their occurrence in time, for a given geographical location, follows a Poisson distribution. An implication of this statement is that the time lapse between wave-storms must follow an exponential probability distribution function; if else, these events would not be extreme. The wave-height used here is the spectral significant wave-height (H_{m0}), which is equal to H_s divided by 0.95 (Holthuijsen (2007)). The candidates for wave-height threshold range from 1.5m to 3m. 1.5m is the double of the mean wave height in the Catalan Coast (CIIRC (2010)). The selected threshold is also used as the location parameter x_0 of the Generalized Pareto distribution employed to fit the H_{m0}. Thus, the threshold should belong to the linear part of a excess-over-threshold plot (Ortego et al. (2012), see an example of such plot in Fig. 3.1). On the same mean excess wave-height plot, the statistical significance of storms is assessed as well.

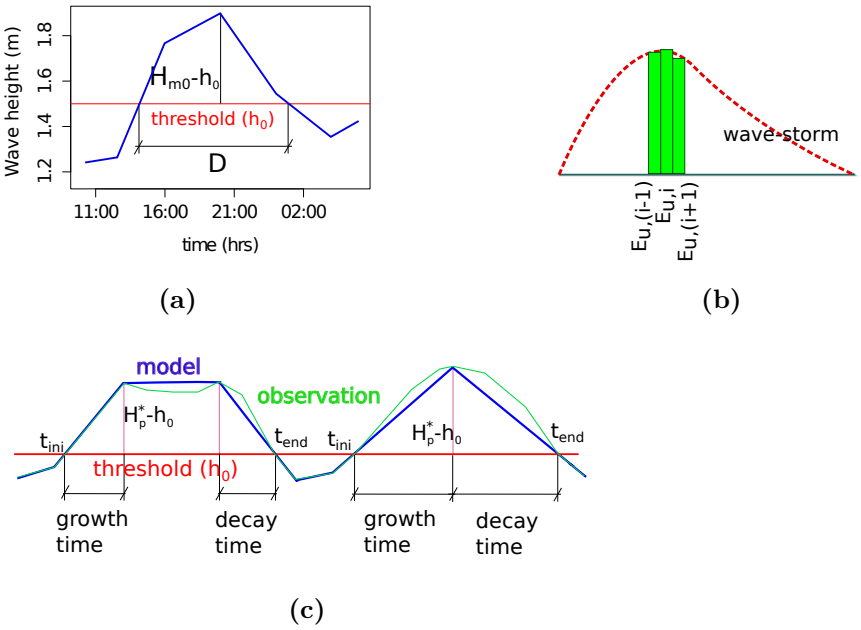


Figure 3.2.: a) Definition of variables for a single wave-storm, b) definition of the unitary energy at the peak of the wave-storm, E_u (the red dashed line is the wave-storm), c) proposed storm shapes (irregular-trapezoid and triangular) and wave-storm variables required for their definition. Source: Lin-Ye et al. (2016)

From these wave-storms, the following variables are modelled: its total energy (E), the unitary energy at the peak of the wave-storm (E_u, which is equivalent to $(H_{m0})^2$), the peak wave-period at the peak of the wave-storm (T_p), the total duration (D), the ratio of time from the start of the wave-storm ($iniT$) to its peak and from the

peak to the end ($endT$) (growth-decay rates, see Fig. 3.2c) and the wave direction at the peak of the wave-storm (written as $\theta_{H_{max}}$ or θ_p^*). E is defined as

$$E = \int_{iniT}^{endT} H_{m0}^2 dt$$

where t is time.
The definition of E_u is:

$$E_u = mean\left(\sum_{i=1}^{k} \int_{iniT_i}^{endT_i} H_{m0}^2 dt\right) \approx mean\left(\sum_{i=1}^{k} \left(E_{u,(i-1)} + E_{u,i} + E_{u,(i+1)}\right)\right)$$

where k is the number of occurrences of maximum H_{m0} and $iniT_i - endT_i$ encloses 3hrs of storm duration. E_u synthesizes the energy shortly before and after the peak of the wave-storm. The subset (see Fig. 3.2b) presents a) point $(t-1)$ where the wave-height is growing to reach the peak, b) point (t) is the storm peak and c) point $(t+1)$ where the wave-height is decreasing or maintaining. The differential energy at $(t+1)$ in decreasing or maintaining the energy is an important assumption for point t. It is what reflects the skewness of the wave-storm. It is considered in the stationary model for the Catalan Coast that the variables E and E_u provide more complete metrics for the storm hazard rather than a representative wave-height, as they describe the behaviour of the entire wave-storm, rather than a snapshot. In the non-stationary characterizations of wave-storms, this variable is dropped in favor of the significant wave-height at the peak of the wave-storm, which is more widely used in Literature. E, E_u, T_p and D take positive real values; consequently, they are log-transformed to avoid scale effects (Egozcue et al. (2006)).

The growth rate is the percentage of time from $iniT$ to the maximum H_{m0}. The decay rate is the percentage of time from the maximum H_{m0} to $endT$. The storm-wave's evolution over time is modelled with either a irregular-trapezoidal or a triangular shape (see Fig. 3.2c). The residuals associated to both candidates are computed as the area below the hindcasted function of H_{m0} minus the area below the corresponding theoretical geometry, as defined in Fig. 3.2c. The geometric shape that provides the lowest residual is chosen for the stationary wave-storm characterization model.

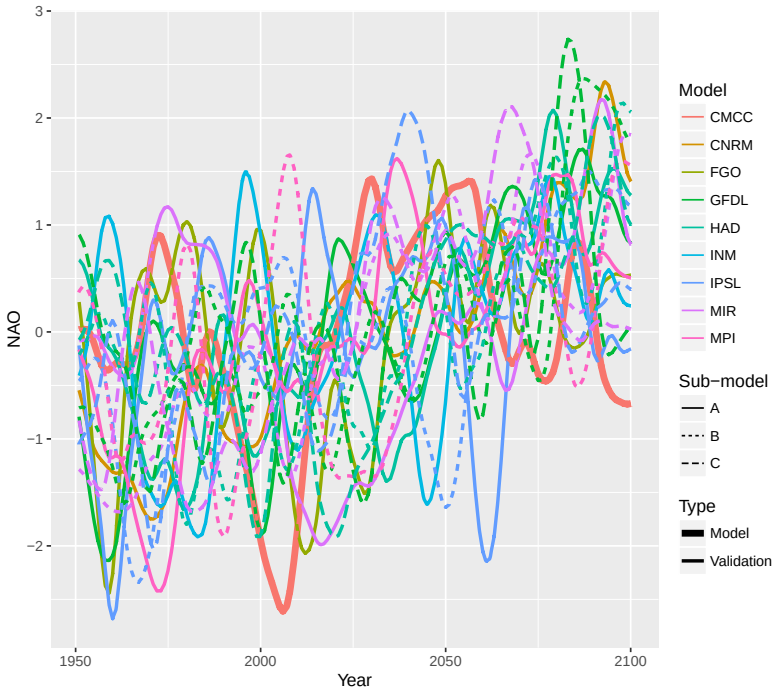
3.1.2. General Circulation Models (GCM) and atmospheric climate-patterns

GCMs provide the general circulation of the Earth's atmosphere. They contain information on atmospheric dynamics, also under the effects of Climate Change (Voldoire et al. (2013); Kwak et al. (2017); Luo et al. (2017)). Different GCMs

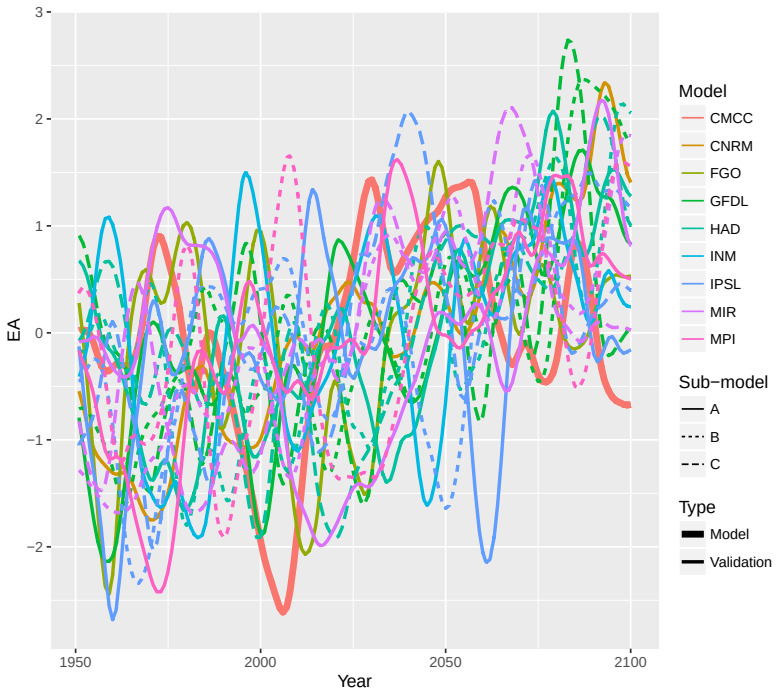
downscalings can be obtained from the CMIP5 experiment. Some of the GCMs created in the CMIP5 experiment are shown on Tab. 3.1. A selected GCM is used in each study area to generate, after a series of steps, the projected wave data to be characterized. The CMCC-CM GCM is used in the Catalan Coast and the CNRM-CM5 GCM is used in the northwestern Black Sea. In each study area, 18 other GCMs (see Tab. 3.1) are valued to bound the uncertainty. The proposed atmospheric climate-patterns are the North Atlantic Oscillation (NAO, Barnston and Livezey (1987)), the East Atlantic pattern (EA) and the Scandinavian pattern (SC). Their climate-indices are used as variables. For instance, the climate-index of NAO is *NAO* (in italic). These climate-indices can be derived from monthly sea-level pressure-fields that can be downloaded from the CMIP5 Project's website. They have been scaled to have a mean value equal to zero and a variance equal to unity.

Table 3.1.: General Circulation Models from the CMIP5 experiment (Taylor et al. (2012)).

GCM model	**focused on Country/ Continent**
CMCC-CM	Europe
CMCC-CMS	Europe
CNRM-CM5	Europe
FGOALS-G2	Europe
GFDL-CM3	United States
GFDL-ESM2G	United States
GFDL-ESM2M	United States
HadGEM2-AO	Europe
HadGEM2-CC	Europe
HadGEM2-ES	Europe
INM-CM4	Europe
IPSL-CM5A-LR	Europe
IPSL-CM5B-LR	Europe
IPSL-CM5A-MR	Europe
MIROC-ESM	Japan
MIROC-ESM-CHEM	Japan
MIROC5	Japan
MPI-ESM-LR	Europe
MPI-ESM-MR	Europe



(a)



(b)

Figure 3.3.: Monthly outputs of a) North Atlantic oscillation and b) East Atlantic pattern, from differen GCMs (Tab. 3.1). Source: Lin-Ye et al. (2017)

The *NAO* index (Barnston and Livezey (1987)) is based on the surface sea-level pressure difference between the Azores High and the Subpolar Low. The positive phase of the NAO reflects below-normal heights and pressure across the high latitudes of the North Atlantic and above-normal heights and pressure over the central North Atlantic, the eastern United States and western Europe. The negative phase reflects an opposite pattern of height and pressure anomalies over these regions. The NAO exhibits considerable inter-seasonal and inter-annual variability, and prolonged periods, of several months, of both positive and negative phases of the pattern are common (see Fig. 3.3a).

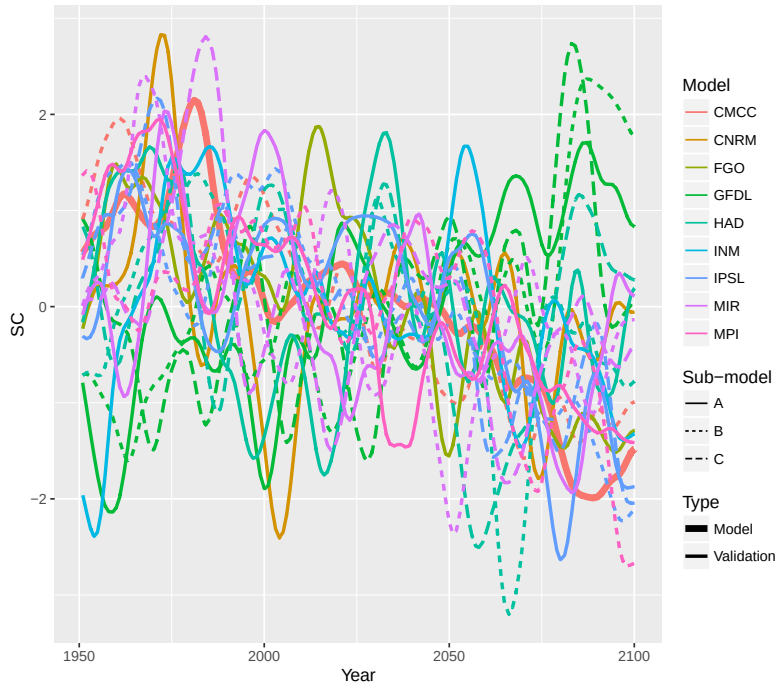


Figure 3.4.: Monthly outputs of Scandinavian pattern, from differen GCMs (Tab. 3.1). Source: Lin-Ye et al. (2017)

The EA pattern (see Fig. 3.3b) is structurally similar to the NAO, and consists of a north-south dipole of anomaly centers spanning the North Atlantic from east to west. The EA pattern could be interpreted as a southward shifted NAO pattern if not because the lower-latitude centre contained a strong subtropical link in association with modulations in the subtropical ridge intensity and location. The EA pattern exhibits very strong multi-decadal variability in the years 1950-2004, with the negative phase prevailing during much of 1950-1976, and the positive phase occurring during much of 1977-2004.

The SC pattern (see Fig. 3.4) consists of a primary circulation center over Scandinavia, with weaker centers of opposite sign over western Europe and eastern Russia/western Mongolia. The positive phase of this pattern is associated with positive height anomalies, sometimes reflecting major blocking anticyclones, over Scandinavia and western Russia, while the negative phase of the pattern is associated with negative height anomalies in these regions.

3.1.3. Storm definition and dataset for the non-stationary model in the Catalan Coast

The non-stationary statistical model to characterize wave-storms in the Catalan Coast uses wave projections for the years 1950-2100. These projections are for a Representative Concentration Pathway 8.5 Climate-Change scenario (RCP8.5, i.e. an increase of the radiative forcing values in the year 2100 relative to pre-industrial values of 8.5W/m^2, Stocker et al. (2013)). This scenario considers a CO_2 concentration in the atmosphere near 1250ppm in 2100, which is double that of any other scenario in the Fifth Assessment Report (Stocker et al. (2013)). The modelling chain is as follows. The CMCC-CM (Sánchez-Arcilla et al. (2008c)) GCM provides boundary conditions for the Regional Circulation Model (RCM) COSMO-CLM (Rockel et al. (2008)). The COSMO-CLM grid, that has a resolution of $0.125° \times 0.125°$, spans the whole Mediterranean region. The wind field that derives from the COSMO-CLM RCM feeds the WAM wave model, which generate wave conditions for the same domain and spatial resolution than the RCM.

A statistical model assuming stationarity in time-frames of 50 years is used to obtain initial values to estimate non-stationary wave-storm thresholds. These time-frames are: past (PT, 1950-2000), present and near-future (PRNF, 2001-2050) and far future (FF, 2051-2100). The wave-storm threshold, h_0, for the H_s is $h_{0,RCP4.5} = 2.09m$. It is based on the methodology in sec. 3.1.1, where the threshold for H_{m0} is 2.2m.

Four wave-storm variables are characterized: E, H_s at the peak of the wave-storm (H_p), T_p and D. Here, E reflects the potential total wave energy within a single event (i.e. analogous to the total volume, in run-off studies), whereas H_p links to the pulsing event that can trigger phenomena, infrastructure damage, flooding or erosion. E, H_p, T_p and D take positive real values; consequently, they are log-transformed to avoid scale effects (Egozcue et al. (2006)).

3.1.4. Storm definition and dataset for the non-stationary model in the northwestern Black Sea

The non-stationary statistical model in the northwestern Black Sea is fed with output from a process-based model. Two Climate-Change scenarios are considered: RCP 4.5 and RCP 8.5. RCP4.5 supposes an increase of the radiative forcing values in the

year 2100, relative to pre-industrial values, of 4.5W/m^2 (Stocker et al. (2013)). A general Circulation Model (GCM) provides atmospheric conditions to be downscaled with a Regional Circulation Model. Here, the GCM is CNRM-CM5, and the RCM is ALADIN. Wind-fields from the ALADIN model have been downloaded from the website of the Mediterranean Coordinated Regional Downscaling Experiment (MedCORDEX) initiative. These wind fields span the whole Europe with a spatial and temporal resolution of 12km × 12km and 3h, respectively. They have served as input for the SWAN spectral wave-model. The computational domain spans the whole Black Sea, with a regular grid of 9km × 9km and a time-step of 20min. The time period modelled spans 1950 to 2100.

The wave-storm threshold, h_0, in this stationary model is $h_{0,RCP4.5} = 1.8m$ for RCP4.5, and 2.0m = $h_{0,RCP8.5}$ for RCP8.5, based on excess-over-threshold graphs of H_s. The methodology in sec. 3.1.1 to select a wave-storm threshold is repeated in the northwestern Black Sea. The difference in value of the storm thresholds for different RCPs is because it has been observed that the storms generated under the RCP4.5 is slightly less extreme than the ones generated under the RCP8.5. Therefore, while the $h_{0,RCP8.5}$ is considered to be 2.0m, the $h_{0,RCP4.5}$ is defined as 0.2m lower than $h_{0,RCP8.5}$. h_0 serves as an initial value for the non-stationary wave-storm thresholds. The energy and hydrodynamics of the storms are also characterized by the same wave-storm intensity variables as in the Catalan Coast. They take positive real values and are log-transformed to avoid scale effects (Egozcue et al. (2006)). θ_p^* is analyzed in the northwestern Black Sea.

3.1.5. Observational data used for validation

Buoy data from the Puertos del Estado buoy network is used to validate the stationary model in the Catalan Coast, in the years 1996-2013. The same buoy data and the SIMAR hindcasted wave data are used to validate the non-stationary model in the Catalan Coast, in the years 1996-2013. ERA-interim reanalysis (Hemer et al. (2013); Dee et al. (2011); Wang et al. (2014)) is used to to validate the non-stationary model in the northwestern Black Sea, in the years 1979-2016. Despite having a spatial resolution of 80km, this global reanalysis presents gapless, bias corrected information from 1979 to present day.

3.2. Model formulation

This section is organized as follows: a) wave-storm intensity sub-model, b) wave-storm intra-time sub-model and wave-storm directionality sub-model, c) generation of data from stationary synthetic wave-storms d) validation and e) estimation of the uncertainty from the GCM.

3.2.1. Wave-storm intensity sub-model

First, the stationary Generalized Pareto distribution (GPD) function is used in the Catalan Coast. Then, a non-stationary GPD is developed, by combining the techniques from a stationary GPD and a Vectorial generalized additive model. A stationary hierarchical Archimedean copula (HAC) is first applied to the stationary model. Then, a series of stationary HACs produce time-series of dependence parameters for the non-stationary approaches. This sub-section is organized as follows: i) Generalized Pareto distribution, ii) Vectorial generalized additive models and iii) the joint dependence structure.

3.2.1.1. Generalized Pareto distribution function (GPD)

E, E_u, H_s, T_p and D are wave-storm intensity variables, and are characterized by GPDs (Coles (2001)). Stationary GPDs are used in the stationary model and non-stationary GPDs are used in the non-stationary models. The definition of a stationary GPD is as follows. It is assumed that the wave-storms are time points which have an associated random magnitude. Variable probability patterns must be independent and identically distributed (Coles (2001); Tolosana-Delgado et al. (2010)). If X is the magnitude of a wave-storm variable and x_0 is, at the same time, a value of the support of X and a location parameter x_0. The excess over the location parameter x_0 is $Y = X - x_0$, conditioned to $X > x_0$. Therefore, the support of Y can be either $[0\,,\, y_{sup}]$ or a positive real line. y_{sup} is the upper bound. The GPD cumulative function is

$$F_Y\left(y|\beta,\xi\right) = 1 - \left(1 + \frac{\xi}{\beta}y\right)^{-\frac{1}{\xi}} \;,\; 0 \leq y \leq y_{sup}\,, \beta \geq 0\,, \xi \in \mathbb{R}$$

and the associated probability density function is

$$f_Y\left(y|\beta,\xi\right) = \frac{1}{\beta}\left(1 + \frac{\xi}{\beta}y\right)^{-\frac{1}{\xi}-1} \;,\; 0 \leq y < y_{sup}\,, \beta \geq 0\,, \xi \in \mathbb{R}$$

where β is the scale parameter and ξ is the shape parameter. ξ determines the domain of attraction of the distribution. The selection of a physically justified threshold for each variable can promote convergence of the tail of the probability distribution function. Given the location-parameter x_0 and the scale-parameter β, the mean value of a wave-storm variable is given by:

$$E\left(X\right) = x_0 + \frac{\beta}{1-\xi} \approx x_0 \qquad \xi < 1,$$

and the variance of a wave-storm variable is:

$$Var\left(X\right) = \frac{\beta^2}{\left(1-\xi\right)^2\left(1-2\xi\right)} \approx \beta^2 \qquad \xi < \frac{1}{2},$$

The location parameter x_0 of the GPD, in the stationary model of the Catalan Coast, is: the square of the storm threshold, h_0^2, for E_u; the peak wave-period related to the H_0 (CIIRC (2010)), for T_p; the minimum storm duration D_{min} = 6hrs for D and $h_0^2 \cdot D_{min}$ for E. In the non-stationary models, the initial values of the location parameters x_0 are: the storm threshold, h_0, for H_p; the peak wave-period related to the h_0 (CIIRC (2010)), for T_p; the minimum storm duration D_{min} = 6hrs for D and $h_0^2 \cdot D_{min}$ for E. Also, the shape parameter ξ is supposed to be constant.

3.2.1.2. Vectorial generalized additive models (VGAM)

The non-stationary models use VGAM to determine the parameters of a non-stationary GPD, as well as the non-stationary wave-storm threshold and the storminess. The initial values of location parameters x_0 used to estimate the GPD are as defined above. The VGAM presents the linear function (Fessler (1991)):

$$\eta_{i(j)} = \beta_{1(j)} + \mathrm{f_{2(j)}}\left(x_{i2}\right) + \ldots + \mathrm{f_{p(j)}}\left(x_{ip}\right)$$

where $\eta_{i(j)}$ is the jth dependent variable, x_i is the ith independent variable that generates η_i. η_i is a sum of smooth functions of the individual covariates. Additive models do all smoothing in $\mathbb{R}$, avoiding large bias introduced in defining areas in $\mathbb{R}^n$.

The assumptions for regression models are: 1) independence of residuals, 2) residuals$\sim N\left(0,\sigma^2\right)$ and 3) homoscedasticity of residuals. The 1) is tested with a AutoCorrelation Function plot, the 2) can be solved with a Quantile-Quantile plot against a $N\left(0,\sigma^2\right)$ distribution (where the sample standard deviation is used as σ^2). 3) can simply be analyzed through a scatterplot of the fitted values vs. the residuals.

The storm-threshold is estimated by fitting its relationship with any factor by a Laplace function. The factors for the non-stationary models can either be time, a climate-index, or either one of the first two time-derivative of a climate-index. The wave-storminess can be estimated by approximating its relationship with any factor by a Poisson probability-distribution-function. As the wave-storminess is a counting variable, a vectorial generalized linear model (VGLM, Yee and Wild (1996)) can be employed in this case. The VGLM is a particular case of VGAM.

The specific technique employed to obtain the GPD location-parameters x_0 is called quantile regression (Koenker (2005); Muraleedharan et al. (2016)). The quantile regression is a type of VGAM, it estimates the $100\hat{\tau}\%$ conditional quantile $y_{\hat{\tau}}\left(x\right)$ of a response variable Y as a function $u\left(x,\tau\right)$ of covariates x. $l_u^* = l_u + \varrho_u R_u$ is minimized. Where $l_u = \hat{\tau}\sum_{i:r_i\geq 0}\left|r_i\right| + \left(1-\hat{\tau}\right)\sum_{i:r_i<0}\left|r_i\right|$ for residuals $r_i = y_i - u\left(x_i,\hat{\tau}\right)$. ϱ

is a roughness coefficient that controls the trade-off between quality of fit to the data and roughness of the regression function. R is a roughness penalty (Northrop and Jonathan (2011); Jonathan et al. (2013)). It is assumed that ξ remains constant, while β can depend on co-variates. β estimated through a VGLM.

The Climatic-indexes NAO, EA and SC, as well as their first two time-derivative, are used as covariates to the VGAM in the prediction of the storm-threshold, the storminess and the GPD parameters. The two main tools to test for the influence of each climate-pattern or its dynamics on the predicted parameters are the Akaike information criterion (AIC, Akaike (1987)) and the Bayesian information criterion (BIC, Tamura et al. (1991)). In the non-stationary model for the northwestern Black Sea, this influence is also quantified by comparing VGAM coefficient to the relative order of magnitude between the predictand and the predictor.

3.2.1.3. Joint dependence structure: the hierarchical Archimedean copula

There is dependence among wave-storm intensity variables (Corbella and Stretch (2012, 2013); Salvadori et al. (2016)). A dependogram can aid visualizing this (see Fig. 3.5). A Dependogram is an independence test based on empirical copula processes (Genest and Remillard (2004)). It provides insight into joint dependencies of any subsets of the variables, displaying the subsets on the horizontal axis and the statistic per subset (the departure from independence) on the vertical axis. A statistic (vertical line) below the threshold value (bullets) means a totally independent subset, whereas the length of the vertical line above the bullet reflects the degree of co-dependence of the variables in the subset.

Copulas (Sklar (1959)) simplify the modelling of the multivariate joint probability structure by linking it to the cummulative probability distribution of the variables. For any multivariate distribution function H with margins F_j, $j \in \{1, ..., d\}$, a copula C can be defined such that

$$\mathrm{H}\left(\mathbf{x}_1, ..., \mathbf{x}_d\right) = \mathrm{C}\left(\mathrm{F}_1\left(\mathbf{x}_1\right), ..., \mathrm{F}_d\left(\mathbf{x}_d\right)\right) \qquad , \mathbf{x} \in \mathbb{R} \tag{3.2.1}$$

Inversely, given a copula C and univariate distribution functions F_j, $j \in \{1, ..., d\}$, an H defined by eq. 3.2.1 is a distribution function with marginals F_j, $j \in \{1, ..., d\}$. Being $u_j = F_j$, a d-dimensional copula is Archimedean if it is of the form

$$\mathrm{C}\left(\mathbf{u}; \phi\right) = \phi\left(\phi^{-1}\left(u_1\right) + \cdots + \phi^{-1}\left(u_d\right)\right), \quad \mathbf{u} \in [0,1]^d,$$

for a given generator function ϕ. An example of a generator function is the Gumbel generator function

$$\left(-\log\left(u\right)\right)^{\theta} \;\; \theta \in [1, \infty).$$

u is the wave-storm variable, and θ is the dependence parameter which indicates independence when $\theta = 1$ and total dependence when $\theta \to \infty$. θ in a copula

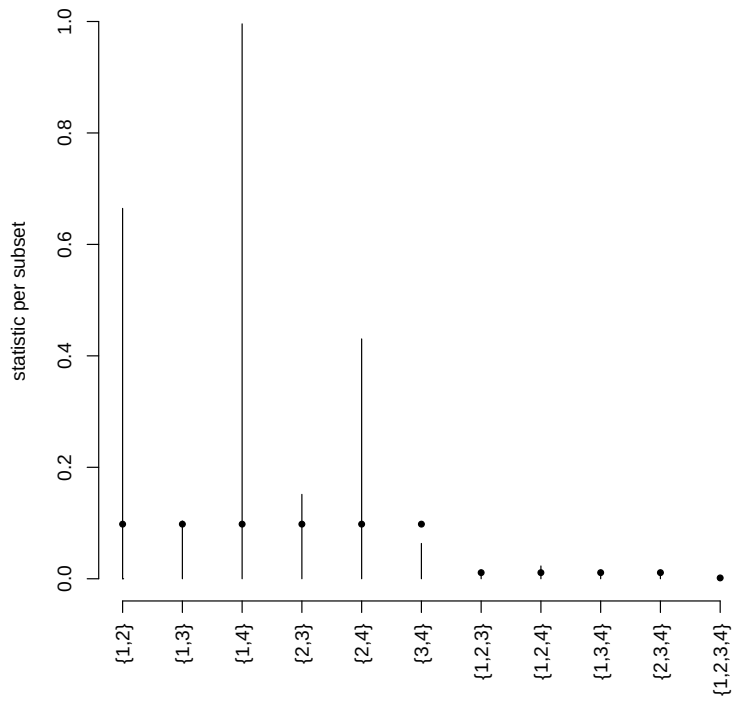


Figure 3.5.: Example of a dependogram. The numbers represent the following wave-storm intensity variables: 1) total energy (E), 2) unitary energy at the peak of the wave-storm (E_u) or significant wave-height at the peak of the wave-storm (H_p), 3) peak wave-period at the peak of the wave-storm (T_p) and 4) total duration (D).

should not be mistaken for a parameter of the movMF distribution. Other types of Archimedean copula generator functions, such as Clayton and Frank can be referred to in Wahl et al. (2011).

Here, θ of different Gumbel copulas are not easily comparable, as the support of θ is semi-infinite. Thus, θ are transformed into Kendall's Tau, or Kendall's rank correlation coefficient (Kendall (1937)), whose support is $[0, 1)$. τ reflects the concordance of sets of observations $\{(x_1, y_1), \ldots, (x_i, y_i), \ldots, (x_n, y_n)\}$:

$$\tau = \frac{(number\ of\ concordant\ pairs) - (number\ of\ discordant\ pairs)}{\frac{1}{2}n\,(n-1)}$$

$\forall i, j \in \mathbb{N},\ i \neq j$, a pair of observations is concordant if $(x_i > x_j) \bigcap (y_i > y_j)$ or $(x_i < x_j) \bigcap (y_i < y_j)$; a pair is discordant if $(x_i > x_j) \bigcap (y_i < y_j)$ or $(x_i < x_j) \bigcap (y_i > y_j)$; if $x_i = x_j$ of $y_i = y_j$, the pair is neither concordant nor discordant.

Archimedean copulas are insensitive to variable permutation and often have con-

strained multivariate dependence structures, as they depend on a single parameter. Hierarchical Archimedean copulas can be a useful tool to overcome these problems by nesting simple 2D-Archimedean copulas into multilayer tree structures (see Fig. 3.6) that are fitted in a recursive way (Okhrin et al. (2013)). Several generator functions can be used for each level of a HAC. For simplicity, in this thesis, only one generator function is used per HAC.

Being $\phi_1, \ldots, \phi_d$ Laplace-Transform generators, for $1 \leq k \leq d-2$, $k+1 < j \leq d$ and $v, \tilde{v} > 0$, multi-dimensional variables can be nested, as the following proves true:

$$\exp\left(-v\phi_{k+1}^{(k)-1}\left(\phi_j^{(k)}\left(\cdot;\tilde{v}\right);\tilde{v}\right)\right) = \phi_j^{(k+1)}\left(\cdot;v\right).$$

Thus, although HACs build structures based on 2-D distributions, they can still preserve the properties of copulas. A HAC is aggregated in the following manner. The couple with the strongest dependence is aggregated and substituted by a joint pseudo-variable (Okhrin et al. (2013)). For example, let E and D share a dependence parameter $\theta_{(E,D)}$. Let it be the highest valued dependence parameter among all the pairs of variables. The pair of variables (E, D) can be substituted by the pseudo-variable

$$\mathbf{Z}_{(E,D)} \overset{def}{=} \phi_{\hat{\theta}_{(E,D)}}\left[\phi_{\hat{\theta}_{(E,D)}}^{-1}\left\{\hat{\mathrm{F}}_D\left(E\right)\right\} + \phi_{\left(\hat{\theta}_{E,D}\right)}^{-1}\left\{\hat{\mathrm{F}}_E\left(D\right)\right\}\right]$$

At the next level, the parameter of all the pairs of variables and pseudo-variables are again evaluated. This procedure is continued until the highest hierarchical level (i.e. the root) is reached.

In order to test the goodness-of-fit of a nested copula, each of its levels must be tested. The model probability distribution is compared to the empirical probability distribution. An empirical copula is constructed as

$$\hat{\mathrm{C}}\left(\mathbf{u}_1, ..., \mathbf{u}_d\right) = n^{-1}\sum_{i=1}^{n}\prod_{j=1}^{d}\mathbf{I}\left\{\hat{\mathrm{F}}_j\left(\mathbf{X}_{ij}\right) \leq \mathbf{u}_j\right\}$$

where n is the sample size, d is the number of variables, $\mathbf{u}_j$ is a vector belonging to the interval $[0, 1]$ and $\mathbf{I}$ is an identity matrix where the diagonal elements $\hat{\mathrm{F}}_j$ are the empirical marginal distribution function of variables $\mathbf{X}_j$. The k^2 of Kendall (Gan et al. (1991)) can be employed to provide a quantitative indicator for such a comparison.

There are three types of aggregation method: «minimum», «mean» and «maximum». If the absolute difference of the dependence parameters of two subsequent nodes is smaller than a constant ε,

$$|\theta_1 - \theta_2| < \varepsilon$$

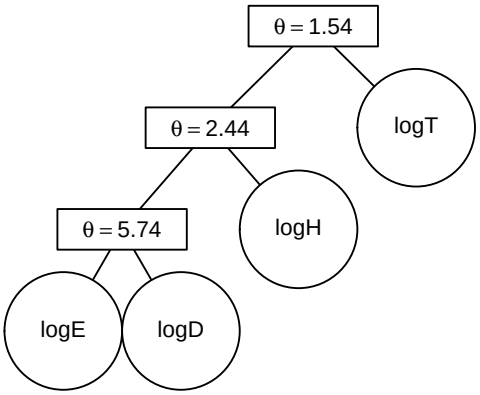


Figure 3.6.: Example of a Gumbel type HAC structure. In this case, it characterizes the joint probability structure of the following four wave-storm intensity variables: total energy and duration (E and D, respectively), significant wave-height and peak wave-period at the peak of the wave-storm (H_p and T_p, respectively). It characterizes the joint probabilities in pairs of variables or pseudo-variables. The θ are the dependence parameters.

«minimum» means that the smallest dependence-parameter of the two is chosen, «maximum» selects the largest dependence parameter, and «mean» calculates their average value. The aggregation method is not attained to any theoretical restrictions, so the one that maximizes goodness-of-fit is the one selected.

Presumably, wave-storms might present a typical pattern of producing extreme values for most wave-storm variables. A Gumbel type HAC is used to include these possible upper extreme dependences (Salvadori et al. (2007)). The HAC have a series of advantages (Okhrin et al. (2013)): a) they are more flexible and hold a more intuitive dependence structure than the simple Archimedean copulas; b) they require less parameters than other kinds of copulas (e.g. elliptical copula); c) because of the hierarchical nature of the structure, there is a marginal cumulative distribution function at each node of the HAC structure, which is easier to interpret; and d) when basing each copula on a single generator function, dependence analyses are simplified.

In order to build a joint probability structure for the non-stationary models, a series of stationary HAC has been obtained for periods of 15 years. The time-windows of 15 years overlap with the previous and the latter time-windows, in order to reproduce a non-stationary effect. Then, the corresponding dependence parameters are linked to form time-series of dependence parameters. A time-window of 15 years has been selected because a shorter time-window would provide an insufficient number of storms to be fitted by a HAC. On the other hand, the smaller the blocks, the more non-stationary effect can be represented with this quasi-stationary HAC fitting. Therefore, 15 years is the optimum number in this case as a tradeoff between the lack of storms for fitting and the maximum size of the blocks. As in the stationary model,

the Gumbel HAC is used, while the fitting method is the Maximum Likelihood method. The goodness-of-fit of a Gumbel type HAC is tested using stationary HACs, built on time-frames of 50 years (PT, PRNF and FF).

3.2.2. The wave-storm intra-time sub-model and the wave-storm directionality sub-model

The growth and decay rates in the stationary model are related to D through a polynomial function. A suitable relationship is a third degree polynomial function, where the independent variable is D: $f(D) = a_0 + a_1 D + a_2 D^2 + a_3 D^3$.

The wave directions in a stationary model are modelled with von Mises-Fisher distributions (vMF), $f(x|\theta)$. The definition of the parameter θ is $\theta = (\mu, \kappa)$, where μ is the mean and κ is the "standard deviation". This type of probability distribution function allow for a more flexible definition of the wave direction contingency, as its support is on $[0, 2\pi)$. The von Mises-Fisher distributions can help transform θ_p^* into categories of principal wave-directions (PD), instead of fitting them into predefined bins of cardinal directions. A mixture of vMF (movMF) distributions (Barnerjee et al. (2005); Mardia and Jupp (2009)) is defined as:

$$f(x|\Theta) = \sum_{h=1}^{k} \alpha_h f_h(x|\theta_h), \tag{3.2.2}$$

where $\Theta = \{\alpha_1, \ldots, \alpha_k, \theta_1, \ldots, \theta_k\}$. The α_h are the h-th mixture probabilities ($1 \leq h \leq k$), they are non-negative and sum to one; by definition, the mode with the largest α_h is the principal direction.

An Expectation maximization (EM) approach is used for maximizing the expectation of eq. 3.2.2 with the constraints $\mu_h^T \mu_h = 1$ and $\kappa_h \geq 0$:

$$\alpha_h = \frac{1}{n} \sum_{i=1}^{n} p(h|x_i, \Theta)$$

From the soft EM framework used here, the distribution of the hidden variables is given by

$$p(h|x_i, \Theta) = \frac{\alpha_h f_h(x_i|\Theta)}{\sum_{l=1}^{k} \alpha_l f_l(x_i|\Theta)}$$

The θ_p^* is decomposed into the sine and cosine of the angle, and these two elements are fit by the movMF distributions. The means μ_k of each movMF within the mixture are considered as a principal direction (PD_k). These PD_k directions are considered categories. All θ_p^* fall into the "influence area" of one of the PD_k and engross the probability of this principal direction. Each PD_k is represented by a μ_k.

The predicted PD_k categories are linked to $\log E$, $\log T$, $\log D$ through a multinomial logistic model (Hosmer et al. (2013)).

3.2.3. Generation of data from (stationary) synthetic wave-storms

In the case of the stationary model, synthetic wave-storms can be generated, given a desired return period (Tr). Here, a single return period should represent the univariate values of the four wave-storm intensity variables E, E_u, T_p and D. For example, if the Tr of the wave-storm was set to be 5 years, the wave-storm intensity variables should be the typical ones of a 5 year return-period wave-storm. θ_p^* and the growth-decay rates are assumed to be independent from the Tr.

The proposed definition of Tr is:

$$Tr = \frac{1}{n}\sum_{i=1}^{n} Tr_i\left(x\right) \ , x \in \mathbb{R}$$

where Tr_i is the return period of a wave-storm intensity variable x. It is calculated by means of the Kendall return period Salvadori et al. (2007):

$$Tr_k = \frac{1}{\lambda \cdot \left(1 - F\left(\mathbf{x}\right)\right)} \ , \quad \lambda \in \mathbb{R} \, , \, \mathbf{x} = \left(x_1, \ldots, x_i, \ldots, x_n\right) \in \mathbb{R}^n \tag{3.2.3}$$

where λ is the annual occurrence of storms and $F\left(\mathbf{x}\right)$ is

$$F\left(\mathbf{x}\right) = \frac{1}{n}\sum_{i=1}^{n} F\left(X_i < x_i\right)$$

where u_i is the cumulative probability of a 1-D variable, $\mathbf{I}$ is the unit interval $[0, 1]$, the critical threshold $t \in \mathbf{I}$ is given by $t = \inf\left\{s \in \mathbf{I} : K_C\left(s\right) = p\right\} = K_C^{[-1]}\left(p\right)$, where K_C is the Kendall coefficient. The selected tolerance of error, imposed on the joint and marginal return periods of elements to be classified within a wave-storm of a given Tr, is 20%. This is the observed limit of errors in the reproduction of wave-storm intensity variables in the northwestern Mediterranean (Sánchez-Arcilla et al. (2008a, 2014)).

3.2.4. Validation

The comparison of the model with the observations is relatively straightforward for the stationary model. The wave-storm variables in the stationary model: E, E_u, T_p, D, θ_p^*, growth-decay rates, can be compared to observational data through a Quantile-Quantile plot.

In a non-stationary model, the wave-storm intensity variables (E, H_p, T_p and D) of the model are compared to the observational data. The methodology of validation for a non-stationary model is as follows. Being the observational data:

$$\{H_{p,1}, \ldots, H_{p,i}, \ldots, H_{p,n}\}, \quad i = 1 \div n, \ n \in \mathbb{R},$$

and the model data (written as H_p^*, just for this explanation):

$$\left\{H_{p,1}^*, \ldots, H_{p,j}^*, \ldots, H_{p,m}^*\right\}, \quad j = 1 \div n, \ m \in \mathbb{R},$$

they are combined to form a joint set of data

$$\left\{H_{p,1}, \ldots, H_{p,i}, \ldots, H_{p,n}, H_{p,1}^*, \ldots, H_{p,j}^*, \ldots, H_{p,m}^*\right\}.$$

For the Catalan Coast, such set of data is partitioned into four intervals, separated by the quartiles of the set. For the northwestern Black Sea, the set of data is partitioned into quantiles that are multiples of 10.

To illustrate the methodology, the one used in the Catalan Coast is described below. It is easily adaptable to the northwestern Black Sea. Being the quartiles $\{q_0, q_{25}, q_{50}, q_{75}, q_{100}\}$, there are elements from both the observed H_p and the model H_p^*, in each interval. Two vectors can be defined as

$$vec_{obs} = \left(\sum_{q_0}^{q_{25}} p\left(H_{p,i}\right), \sum_{q_{25}}^{q_{50}} p\left(H_{p,i}\right), \sum_{q_{50}}^{q_{75}} p\left(H_{p,i}\right), \sum_{q_{75}}^{q_{100}} p\left(H_{p,i}\right)\right),$$

and

$$vec_{model} = \left(\sum_{q_0}^{q_{25}} p\left(H_{p,j}^*\right), \sum_{q_{25}}^{q_{50}} p\left(H_{p,j}^*\right), \sum_{q_{50}}^{q_{75}} p\left(H_{p,j}^*\right), \sum_{q_{75}}^{q_{100}} p\left(H_{p,j}^*\right)\right),$$

where vec_{obs} is the vector for the observation, and vec_{model} is the one for the model. Each element of the vector is the summation between two quartiles of the probability distribution function. vec_{obs} and vec_{model} are compositional data, as their elements are parts of a whole (Egozcue and Pawlowsky-Glahn (2011)), appart of fulfilling other properties defined in Aitchison (1982) and Egozcue et al. (2003). The distance between these two vectors can be determined with an Aitchison measure (Aitchison (1992); Pawlowsky-Glahn and Egozcue (2001)),

$$d\left(\mathbf{x}, \mathbf{y}\right) = \left|\ln \frac{\mathbf{x}\left(\mathbf{1} - \mathbf{y}\right)}{\mathbf{y}\left(\mathbf{1} - \mathbf{x}\right)}\right|, \quad \mathbf{x}, \mathbf{y} \in (0, 1) \in \mathbb{R}, \tag{3.2.4}$$

Where $\mathbf{x}$ and $\mathbf{y}$ are the vectors compared. Another metric is the Kullback-Leibler

divergence (Kullback (1997)):

$$D_{KL}\left(P \parallel Q\right) = \sum_{i} P\left(i\right) \log \frac{P\left(i\right)}{Q\left(i\right)}. \tag{3.2.5}$$

This function measures the extra entropy of the probability distribution Q of the model, respect to the probability distribution P of the observations. Note that for any i, $Q\left(i\right) = 0$, must imply $P\left(i\right) = 0$, or there would be a indertemination. That is, the model should consider all the values that the observations show. Also, whenever $P\left(i\right) = 0$, the the contribution of the i-th term is null, as $\lim_{x \to 0} x \log\left(x\right) = 0$.

Both eq. 3.2.4 and 3.2.5 take values in $\mathbb{R}_0^+$. The module of the vector is a particular case of both measures, thus both can be compared to the vectorial module in the Euclidean space of $\mathbf{x}$ and $\mathbf{y}$, which are of the order of 1 (Egozcue and Pawlowsky-Glahn (2011)). Measures below 1 denote similarity of the model to the observations.

3.2.5. Estimation of the uncertainty from the GCM

For the non-stationary models, different GCM can be compared to each other by measuring the distance between the outcoming wave-storm variables. A Partial AutoCorrelation Function-based distance can measure the autocorrelation between two time series by providing comparison information of lower dimensions than the time-series themselves (Montero and Vilar (2014)). It is a model-free approach. That is, the two time-series compared, do not have to belong to specific models. The PACF coefficient measures the correlation of pairs of elements from each time series at all shorter lags of time.

Let $\hat{\rho}_{X_T} = \left(\hat{\rho}_{1,X_T}, \ldots, \hat{\rho}_{L,X_T}\right)^T$ and $\hat{\rho}_{Y_T} = \left(\hat{\rho}_{1,Y_T}, \ldots, \hat{\rho}_{L,Y_T}\right)^T$ be the estimated autocorrelation vectors of the two time-series, X_T and Y_T respectively, for some L such that $\hat{\rho}_{i,X_T} \approx 0$ and $\hat{\rho}_{i,Y_T} \approx 0$ for $i > L$. Then, an autocorrelation distance is defined as $\mathrm{d}_{ACF}\left(X_T, Y_T\right) = \sqrt{\left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)^T \Omega \left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)}$, where Ω is a matrix of weights of value ≤ 1, and $0 \leq \mathrm{d}_{ACF}\left(X_T, Y_T\right) \leq 2$, where 0 corresponds to total coincidence, and 2, to total discordance.

4. A multivariate statistical model of extreme events: an application to the Catalan coast

J. Lin-Ye, M. García-León, V. Gràcia, A. Sánchez-Arcilla
Coastal Engineering, **117**, 138–156, 2016.
doi.org/10.1016/j.coastaleng.2016.08.002

Attention ¡i
Pages 42 to 62 of the thesis, containing the text mentioned above, are available at the editor's web
https://www.sciencedirect.com/science/article/pii/S0378383916301661

5. Multivariate statistical modelling of future marine storms

J. Lin-Ye, M. García-León, V. Gràcia, M.I. Ortego, P. Lionello, A. Sánchez-Arcilla

Applied Ocean Research, **65**, 192–205, 2017.
doi.org/10.1016/j.apor.2017.04.009

Attention ¡i
Pages 64 to 80 of the thesis, containing the text mentioned above, are available at the editor's web
https://www.sciencedirect.com/science/article/pii/S0141118717300743

6. Suplementary material to "Multivariate statistical modelling of future marine storms "

J. Lin-Ye, M. García-León, V. Gràcia, M.I. Ortego, P. Lionello, A. Sánchez-Arcilla

Applied Ocean Research, **65**, 192–205, 2017.

Multivariate statistical modelling of future marine storms

J. Lin-Ye[a,1,*], M. García-León[a], V. Gràcia[a], M.I. Ortego[b], P. Lionello[c], A. Sánchez-Arcilla[a]

[a] *Laboratory of Maritime Engineering, Barcelona Tech, D1 Campus Nord, Jordi Girona 1-3, 08034, Barcelona, Spain*

[b] *Department of Civil and Environmental Engineering, Barcelona Tech, C2 Campus Nord, Jordi Girona 1-3, 08034, Barcelona, Spain*

[c] *Centro Euro-Mediterraneo sui Cambiamenti Climatici, Via Augusto Imperatore, 16, Lecce, Italy*

Abstract

Extreme events, such as wave-storms, need to be characterized for coastal infrastructure design purposes. Such description should contain information on both the univariate behaviour and the joint-dependence of storm-variables. These two aspects have been here addressed through generalized Pareto distributions and hierarchical Archimedean copulas. A non-stationary model has been used to highlight the relationship between these extreme events and non-stationary climate. It has been applied to a Representative Concentration pathway 8.5 Climate-Change scenario, for a fetch-limited environment (Catalan Coast). In the non-stationary model, all considered variables decrease in time, except for storm-duration at the northern part of the Catalan Coast. The joint distribution of storm variables presents cyclical fluctuations, with a stronger influence of climate dynamics than of climate itself.

Keywords: wave storm, Catalan Coast, hierarchical Archimedean copula, generalized Pareto distribution, non-stationarity, generalized additive model

*Corresponding author
Email address: jue.lin@upc.edu (J. Lin-Ye)

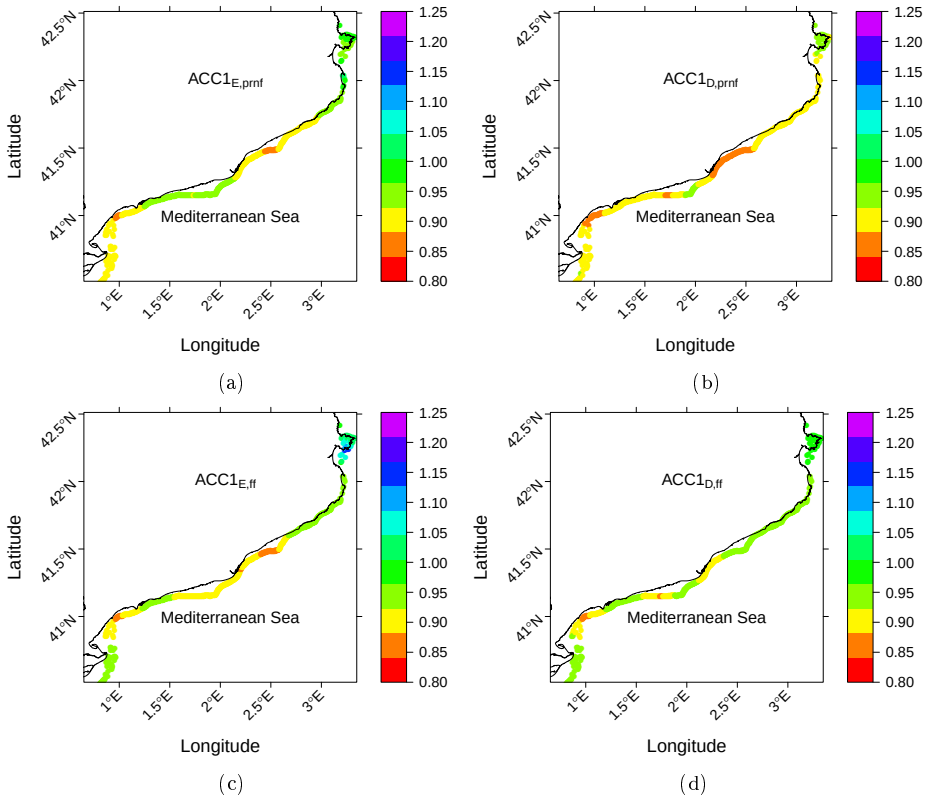


Figure 1: ACC1 in different time-frames for wave-energy (E) (subfigures a and c) and duration (D) (subfigures b and d). Results for H_p and T_p are not shown because their ACC1 values are close to 1.

1

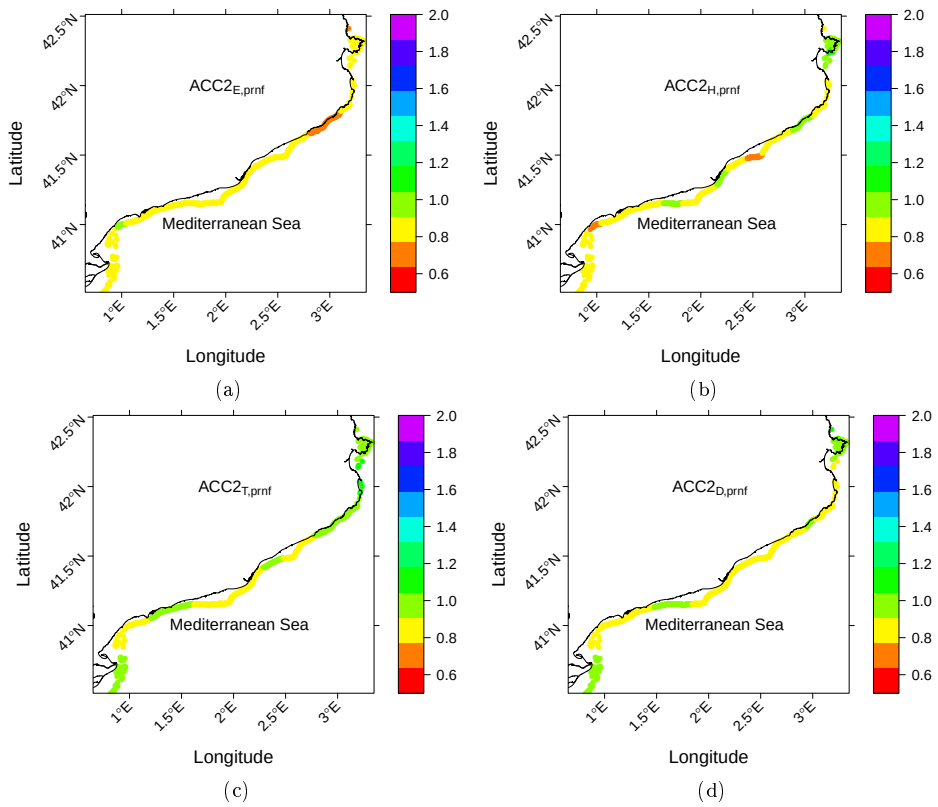


Figure 2: ACC2 at present-near-future (PRNF, 2000-2050) for a) E, b) H_p, c) T_p and d) D.

2

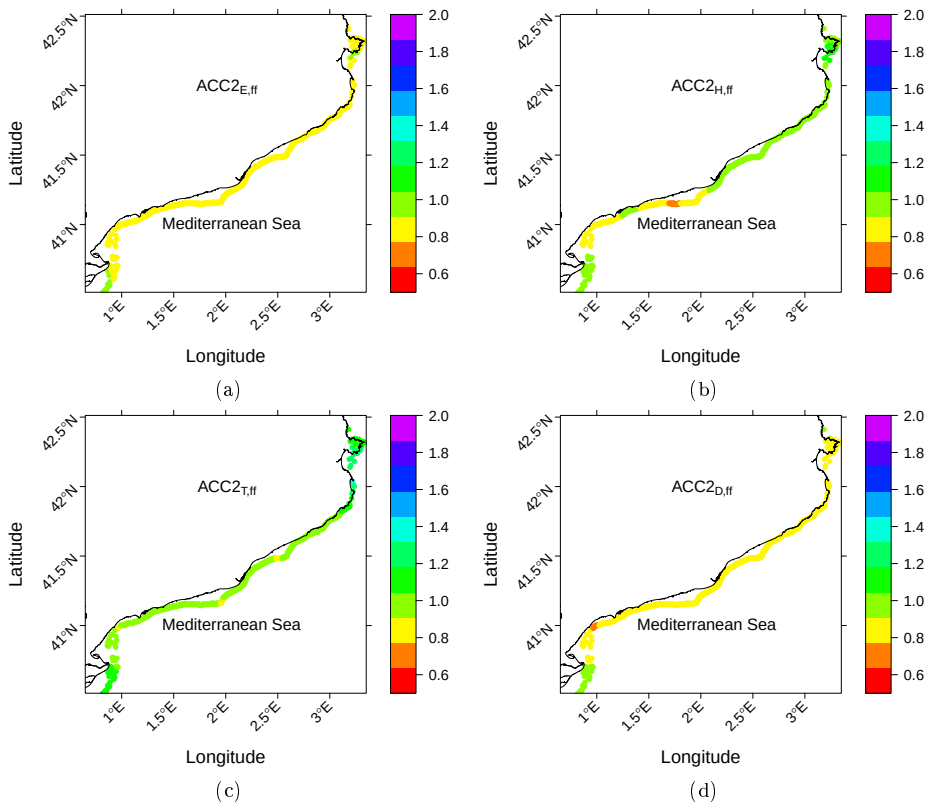
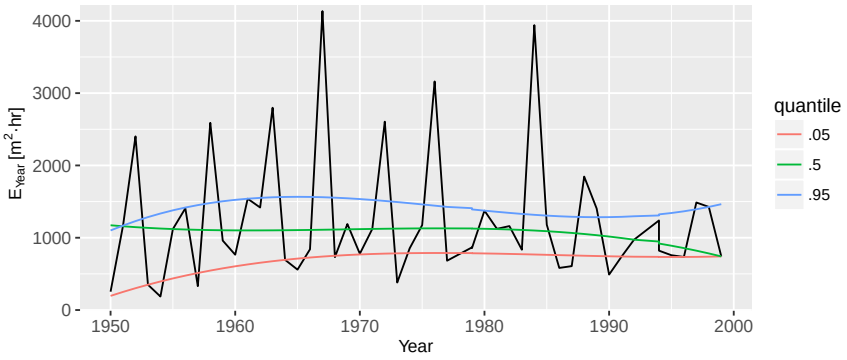
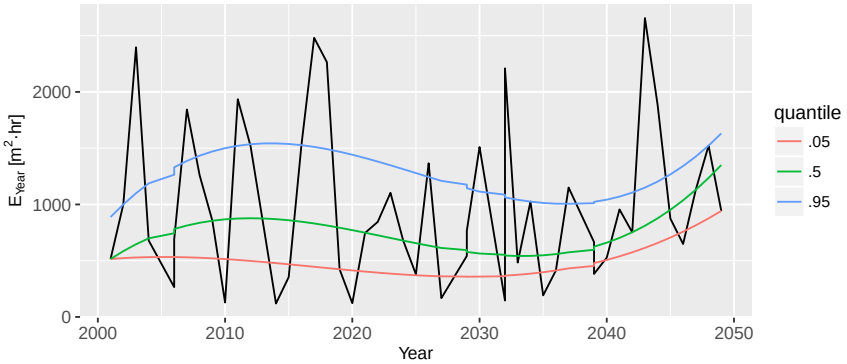


Figure 3: ACC2 at far-future (FF, 2050-2100) for a) E, b) H_p, c) T_p and d) D.

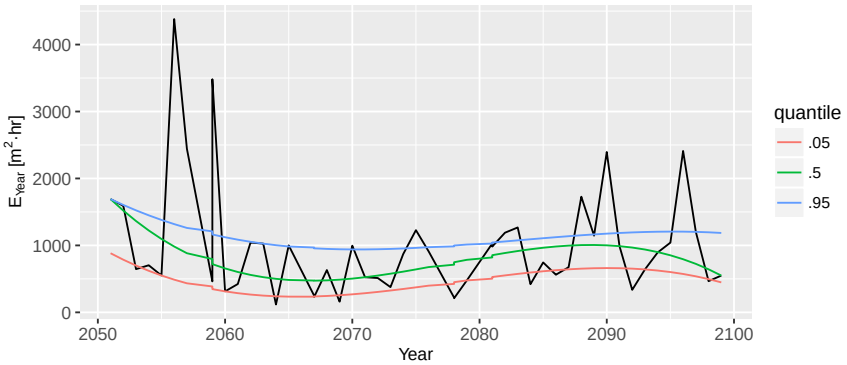
3



(a)



(b)



(c)

Figure 4: Yearly averaged storm-energy (E_{year}) during a) past (PT), b) present-near-future (PRNF) and c) far-future (FF), at node 12.

4

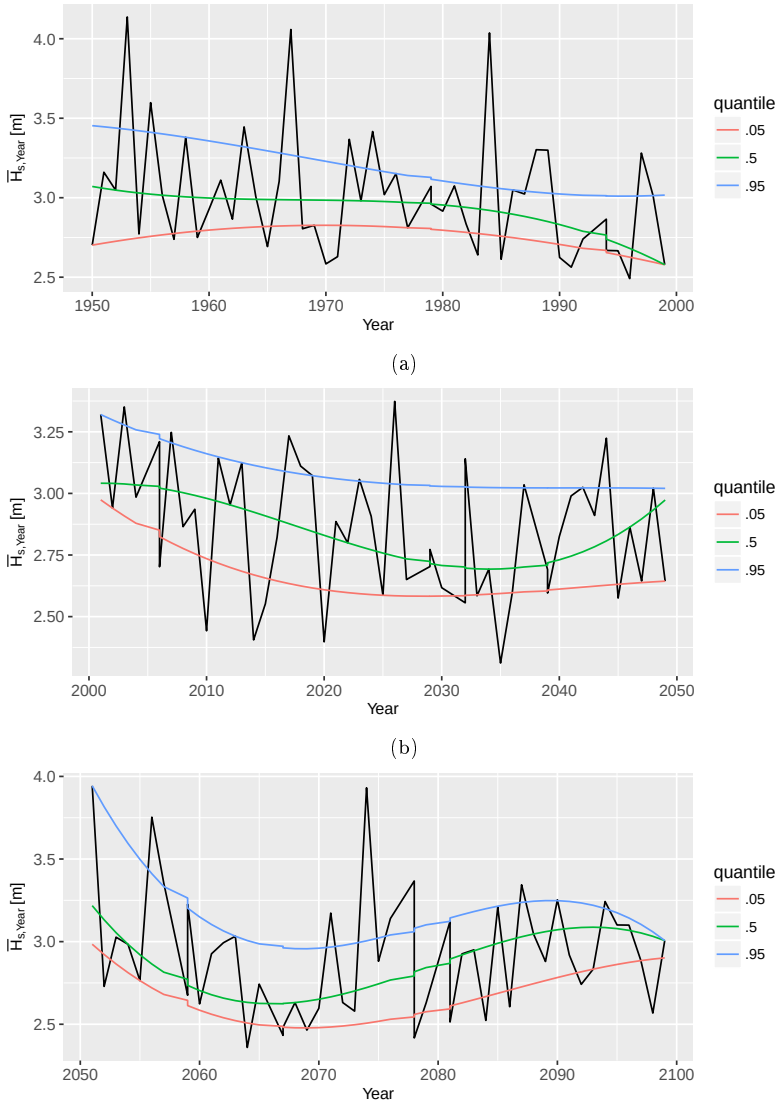
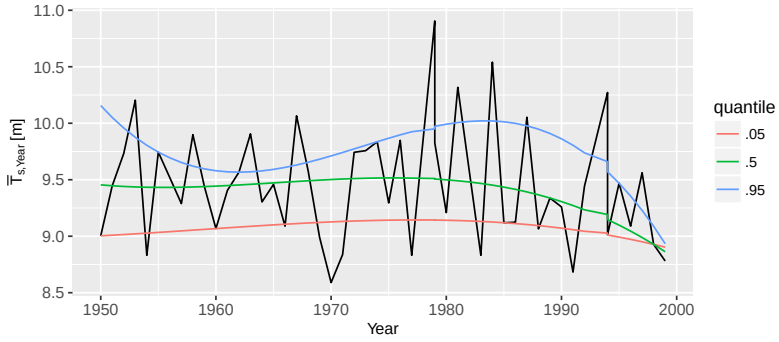
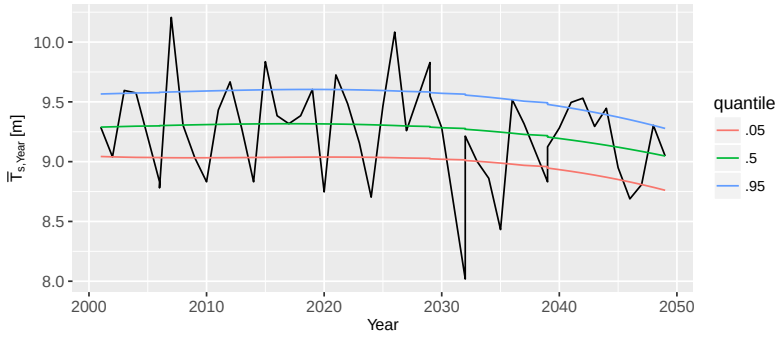


Figure 5: Yearly averaged storm-duration ($\overline{H}_{s,year}$) during a) past (PT), b) present-near-future (PRNF) and c) far-future (FF), at node 12.

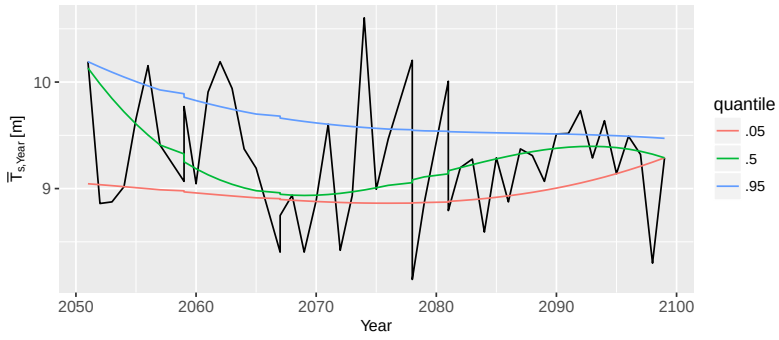
5



(a)



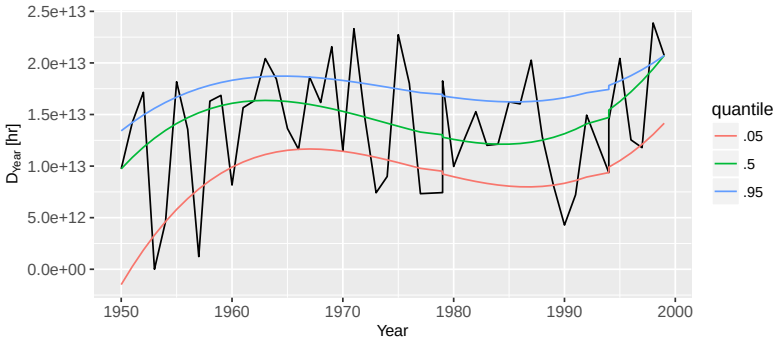
(b)



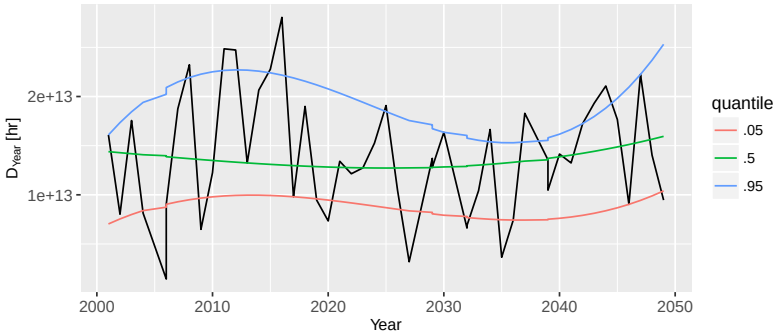
(c)

Figure 6: Yearly averaged storm-duration ($\overline{T}_{p,year}$) during a) past (PT), b) present-near-future (PRNF) and c) far-future (FF), at node 12.

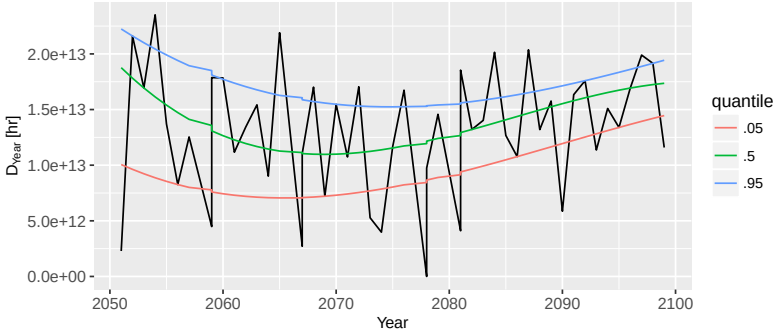
6



(a)



(b)



(c)

Figure 7: Yearly averaged storm-duration (D_{year}) during a) past (PT), b) present-near-future (PRNF) and c) far-future (FF), at node 12.

7

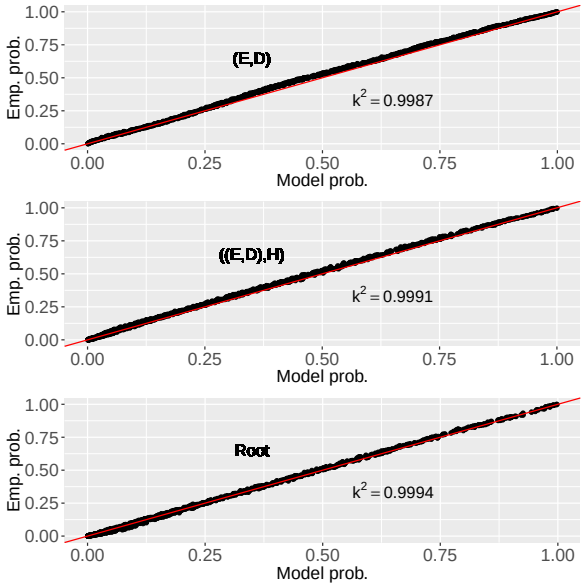


Figure 8: Goodness-of-fit of the dependence structure at node 12, for past (PT). The nesting level of the top image corresponds to the lowest in Fig.4 (main text), the image in the middle corresponds to the second lowest nesting level in Fig. 4 (main text), and the last figure corresponds to the highest (root) nesting level. The k^2 parameter (Gan et al., 1991) quantifies the goodness-of-fit.

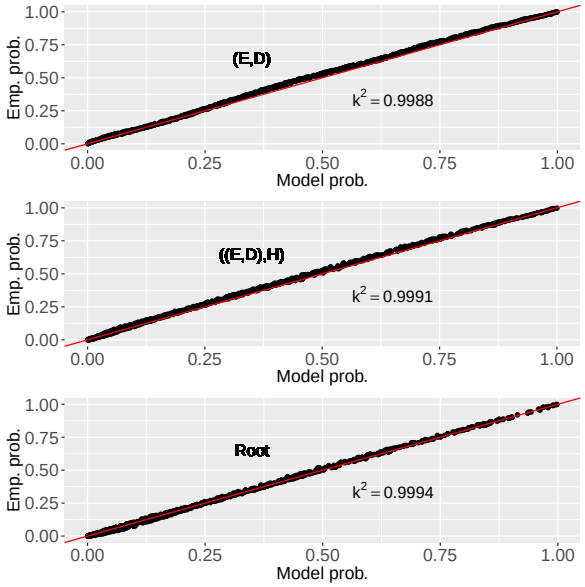


Figure 9: GoF of the dependence structure at node 12, for present-near-future (PRNF).

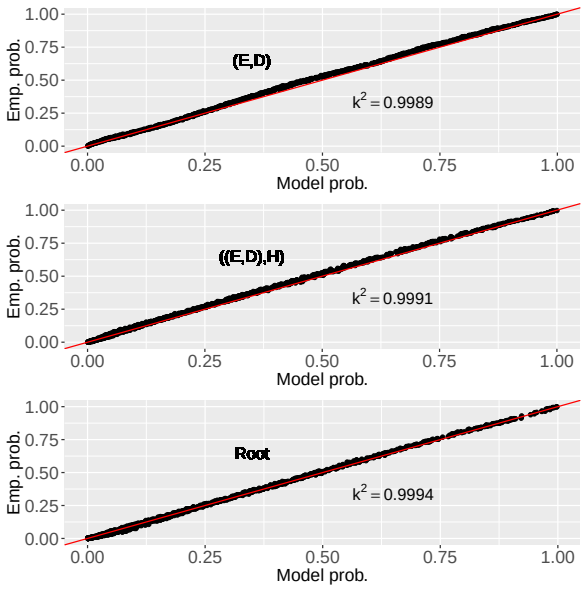


Figure 10: GoF of the dependence structure at node 12, for far-future (FF).

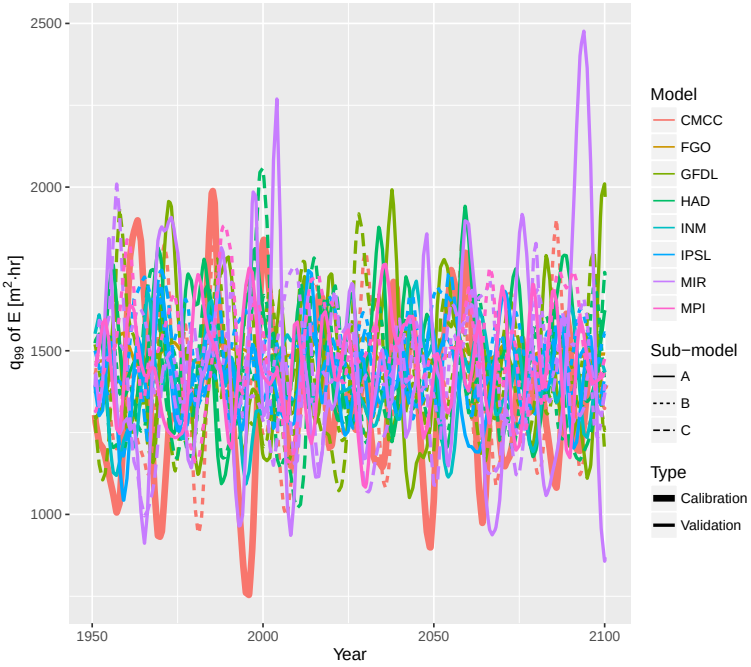


Figure 11: 99^{th} quantile of E at node 12, using VGAM (GPD distribution) with climate-indices as covariates: $E \sim \left(GPD\left(\mu\left(dSC\right), \sigma\left(d^2EA\right), \xi\right)\right)$.

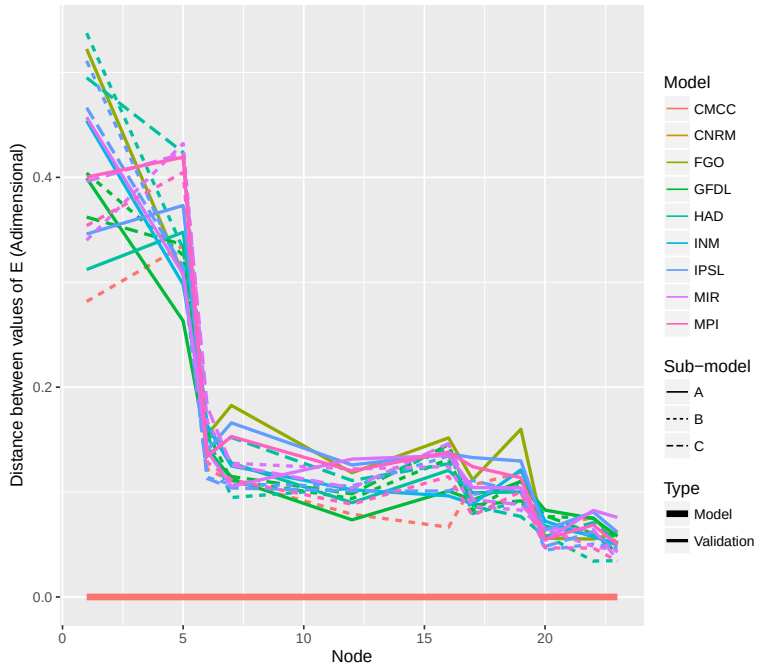


Figure 12: Partial autocorrelation distance between 99^{th} quantiles of E from the CMCC-CM compared to the other global circulation-models. This is related to Fig. 11, where the same colour categories and linetypes are used. A null distance means perfect agreement with the reference (CMCC-CM), whereas a 1 is the opposite. The x-axis represent different nodes (see Fig. 1 of the main text).

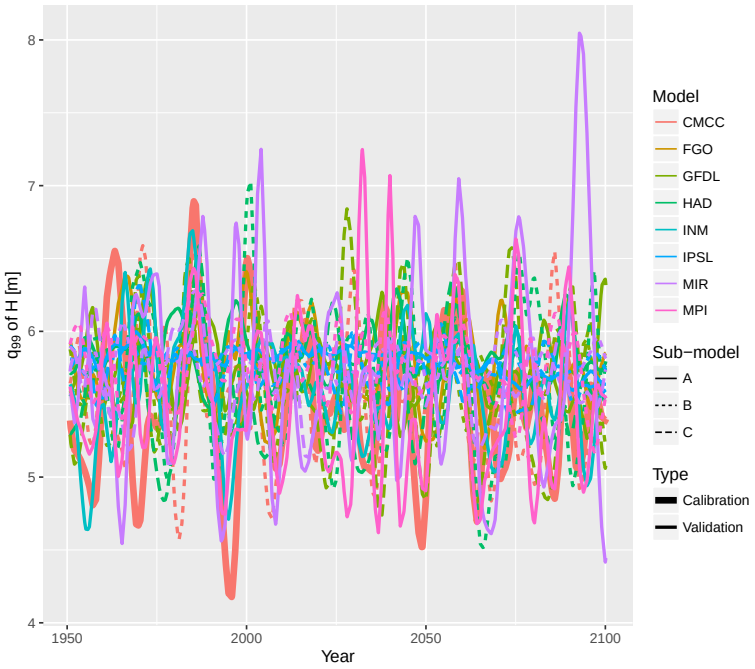


Figure 13: 99^{th} quantile of H_p at node 12, using VGAM (GPD distribution) with climate-indices as covariates: $H_p \sim \left(GPD\left(\mu\left(SC\right), \sigma\left(d^2EA, d^2SC\right), \xi\right)\right)$.

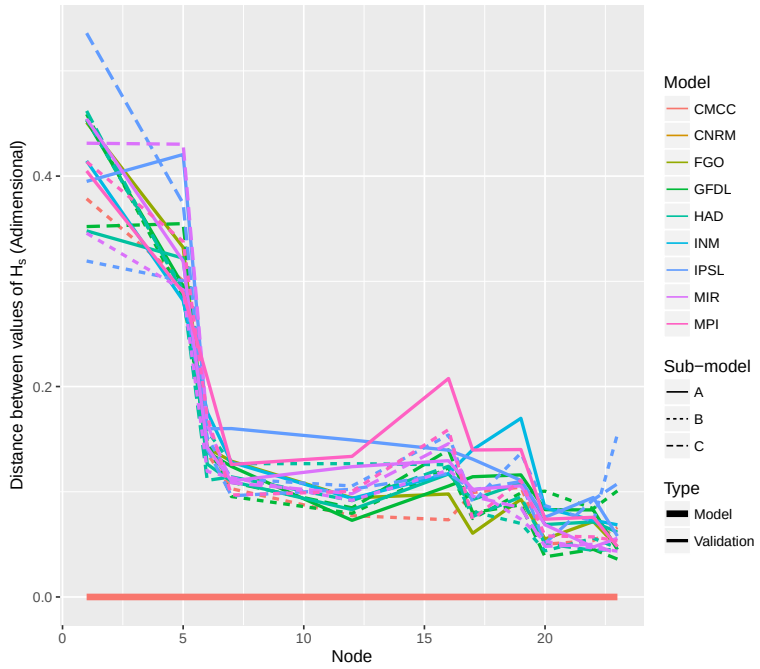


Figure 14: Partial autocorrelation distance between 99^{th} quantiles of H_p from the CMCC-CM compared to the other global circulation-models. This is related to Fig. 13. The x-axis represent different nodes.

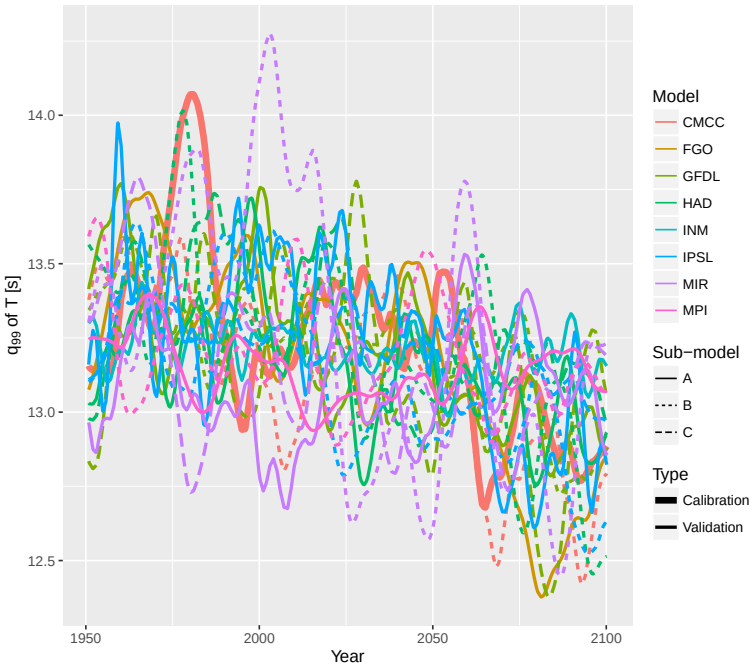


Figure 15: 99^{th} quantile of T_p at node 12, using VGAM (GPD distribution) with time series climate-indices as covariates: $T_p \sim (GPD\left(\mu\left(SC\right), \sigma\left(NAO\right), \xi\right))$.

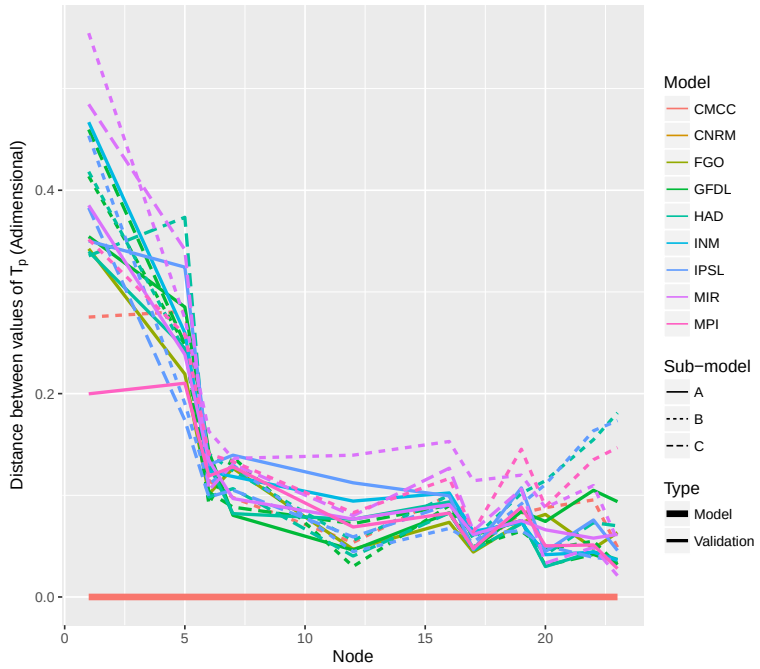


Figure 16: Partial autocorrelation distance between 99^{th} quantiles of T_p from the CMCC-CM compared to the other global circulation-models. This is related to Fig. 15. The x-axis represent different nodes.

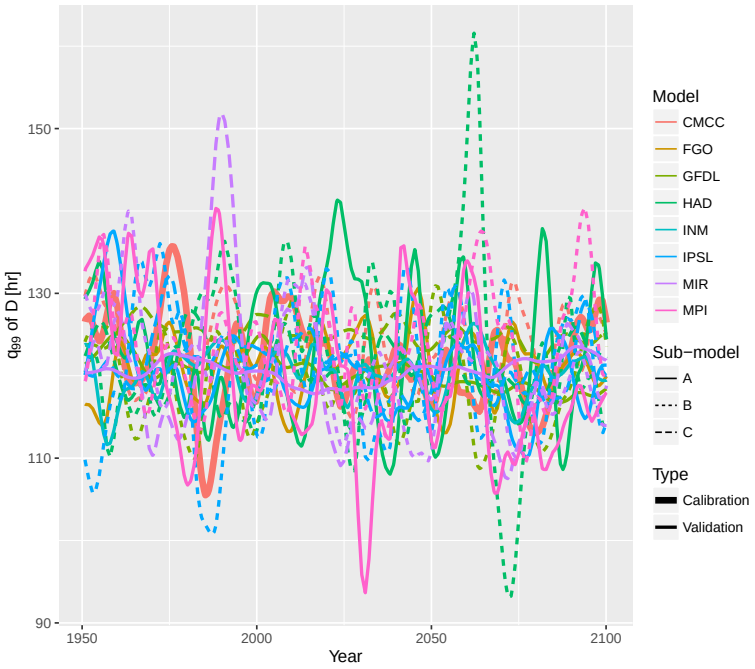


Figure 17: 99^{th} quantile of D at node 12, using VGAM (GPD distribution) with time series climate-indices as covariates: $D \sim (GPD\left(\mu\left(EA\right), \sigma\left(dSC\right), \xi\right))$.

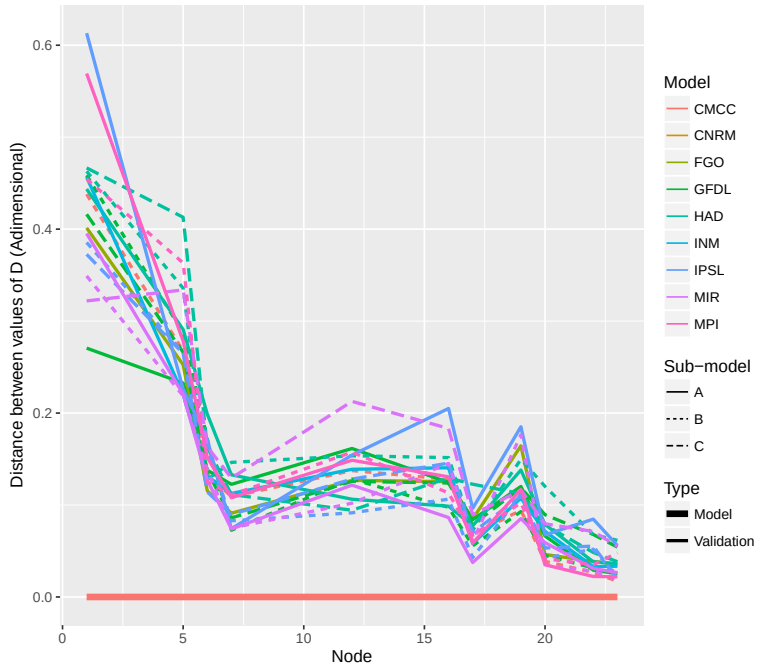


Figure 18: Partial autocorrelation distance between 99^{th} quantiles of D from the CMCC-CM compared to the other global circulation-models. This is related to Fig. 17. The x-axis represent different nodes.

Gan, F., K. Koehler, and J. Thompson: 1991, 'Probability Plots and Distribution Curves for Assessing the Fit of Probability Models'. *The American Statistician* **45**(1), 14–21.

7. Multivariate hybrid modelling of future wave-storms at the northwestern Black Sea

J. Lin-Ye, M. García-León, V. Gràcia, M.I. Ortego, A. Stanica, A. Sánchez-Arcilla

Water, **10**(2), 2018.

water

MDPI

Article

Multivariate Hybrid Modelling of Future Wave-Storms at the Northwestern Black Sea

Jue Lin-Ye [1,*], Manuel García-León [1], Vicente Gràcia [1], M. Isabel Ortego [2], Adrian Stanica [3] and Agustín Sánchez-Arcilla [1]

1 Laboratory of Maritime Engineering, Universitat Politècnica de Catalunya, D1 Campus Nord, Jordi Girona 1-3, 08034 Barcelona, Spain; manuel.garcia-leon@upc.edu (M.G.-L.); Vicente.gracia@upc.edu (V.G.); agustin.arcilla@upc.edu (A.S.-A.)
2 Department of Civil and Environmental Engineering, Universitat Politècnica de Catalunya, C2 Campus Nord, Jordi Girona 1-3, 08034 Barcelona, Spain; ma.isabel.ortego@upc.edu
3 GeoEcoMar, GeoEcoMar, Bucharest, 23-25, Dimitrie Onciul, Bucharest 024053, Romania; astanica@geoecomar.ro
* Correspondence: jue.lin@upc.edu

Received: 16 November 2017; Accepted: 13 February 2018; Published: 19 February 2018

Abstract: The characterization of future wave-storms and their relationship to large-scale climate can provide useful information for environmental or urban planning at coastal areas. A hybrid methodology (process-based and statistical) was used to characterize the extreme wave-climate at the northwestern Black Sea. The Simulating WAve Nearshore spectral wave-model was employed to produce wave-climate projections, forced with wind-fields projections for two climate change scenarios: Representative Concentration Pathways (RCPs) 4.5 and 8.5. A non-stationary multivariate statistical model was built, considering significant wave-height and peak-wave-period at the peak of the wave-storm, as well as storm total energy and storm-duration. The climate indices of the North Atlantic Oscillation, East Atlantic Pattern, and Scandinavian Pattern have been used as covariates to link to storminess, wave-storm threshold, and wave-storm components in the statistical model. The results show that, first, under both RCP scenarios, the mean values of significant wave-height and peak-wave-period at the peak of the wave-storm remain fairly constant over the 21st century. Second, the mean value of storm total energy is more markedly increasing in the RCP4.5 scenario than in the RCP8.5 scenario. Third, the mean value of storm-duration is increasing in the RCP4.5 scenario, as opposed to the constant trend in the RCP8.5 scenario. The variance of each wave-storm component increases when the corresponding mean value increases under both RCP scenarios. During the 21st century, the East Atlantic Pattern and changes in its pattern have a special influence on wave-storm conditions. Apart from the individual characteristics of each wave-storm component, wave-storms with both extreme energy and duration can be expected in the 21st century. The dependence between all the wave-storm components is moderate, but grows with time and, in general, the severe emission scenario of RCP8.5 presents less dependence between storm total energy and storm-duration and among wave-storm components.

Keywords: SWAN; storminess; climate change; climate patterns; Black Sea; copula; generalized additive model

1. Introduction

The hydrosphere presents several types of extreme events, such as droughts [1], floods [2], and wave-storms [3]. Coastal areas are one of the most active environments, which can lead to conflicts and incompatibility of uses. Wave-action dynamics drives these changes, affecting infrastructure stability, sediment dynamics, and the resilience of coastal systems [4–6]. Such variability reaches hazardous

rates under extreme wave regimes (i.e., wave-storms) [7]. Hence, wave-storm characterization can provide useful information regarding their destructive potential.

Wave-storm characterization can be performed through two approaches: process-based [8,9] and statistical [10,11] models. Process-based models can include complex physical phenomena, but they can be computationally expensive. Statistical models are easier to interpret and computationally cheaper, but they cannot reproduce local phenomena. A hybrid strategy (i.e., statistical models built using process-based outputs as input) is a better method for both approaches, and recent works have addressed this methodology [10,12,13]. In order to tackle multiple wave-storm components at once, a multivariate statistical model can serve to characterize individual storm components, as well as the dependence structure. The significant-wave-height (H_s) is the most frequently used wave-storm component. It is usually regarded as being independent of other storm components, such as peak-period (T_p) or storm-duration (D). However, this assumption has been questioned in [14,15], among others. Similarly, it was discussed in Salvadori *et al.* [16] that univariate analyses lead to an inaccurate estimation of marine drivers, so these cannot describe coastal processes adequately. Fully-nested Archimedean copulas have previously been successfully applied to characterize semi-dependence among variables [15,17], so this hypothesis was adopted in the proposed statistical model.

The effects of wave-storms may be aggravated as a consequence of climate change [11,18,19]. Changes in extreme wave-climate add a layer of complexity, and the often-used stationary methods are limited when addressing climate trends. Non-stationary models can better capture the variation introduced by climate change by handling the changing trends of the storm components better [20]. Extreme value distributions of wave-storm variables can be modelled as linear or smooth functions of covariates [21], and a generalized additive model can be used to estimate the location and scale parameters of a generalized Pareto distribution fitted to wave-storm variables [22–24]. Indices related to atmospheric climate patterns can serve as covariates in these regression models. Thus, the relationship between a changing atmosphere and the wave-storms can be tackled.

Hybrid approaches that address the main wave-storm physical processes and the nature of the storm components are suitable for managing coastal areas such as the coast of the northwestern Black Sea. However, there is a lack of future wave projections in this area. The aim of this paper is to characterize the extreme wave-climate at the northwestern Black Sea with a hybrid strategy, under two climate change scenarios (Representative Concentration Pathways (RCPs) 4.5 and 8.5). These scenarios represent an increase of the radiative forcing-values in the year 2100 relative to pre-industrial values of 4.5 and 8.5 W/m^2 [25,26], respectively. Wave projections were obtained from a process-based model (Simulating WAves Nearshore, SWAN). SWAN outputs have served to build a multivariate non-stationary statistical model that characterizes the probability distributions and the joint probability structure of the wave-storm variables. It also relates these variables to climate indices. This characterization can help to assess the level of change in wave-storm characteristics under the effects of climate change. The wave-storms under the two proposed emission scenarios are compared to each other. Once the relationship between wave-storm components and climate indices is determined, a set of different Global Circulation Models are used in order to bound their own uncertainty. The paper is structured as follows. Section 2 describes the study-area and its climate. Section 3 states the methodology to project and to build the hybrid framework. Results are listed in Section 4, discussed in Section 5, and concluded in Section 6.

2. Study Area

The Black Sea is a fetch-limited, wave-dominated, and micro-tidal basin (see Figure 1), located between 41.0 and 46.0°N and 27.0 and 42.0°E [27]. In fetch-limited basins, waves do not have enough length of fetch to reach the fully arisen sea condition, and fetch-limited sea states are generated. In these situations, there may be two effects: (i) for a fixed wind-speed, the maximum significant wave-height is limited by the fetch-length; (ii) the time-duration of swell-waves may be shorter than in non-limited

fetch condition. A micro-tidal basin has a small tidal range, on the order of decimetres. Hence, the sea variability depends on wind and waves, partly contributed by the Danube river and to a lesser degree by the Mediterranean Sea. The Black Sea is connected to the Mediterranean Sea through the Sea of Marmara and the Bosporus and Dardanelles straits to the southwest, and to the Sea of Azov through the Kerch Strait on the opposite side. The greater part of the Black Sea is a basin with a relatively flat bottom relief and depths exceeding 2000 m. However, its western shelf slope is considerably gentle. The proposed non-stationary statistical model is built at the northwestern area, where 34 nodes are used (see Figure 1).

The general large-scale atmospheric circulation over the Black Sea is influenced by the configuration of the Azores and Siberian high-pressure areas and the Asian low-pressure area. Additionally, a great part of the Black Sea's coast is surrounded by mountains, which are the Balkans, the Pontic, the Caucasus, and the Crimean mountains. This feature generates specific wind patterns in the inner shelf-area. Local winds such as sea breezes, mountain-valley circulation, and slope winds also have a considerable impact on the atmospheric circulation pattern of the study-area [28]. The most remarkable feature of wind and wave-climate in the northwestern Black Sea is their seasonal variability [4].

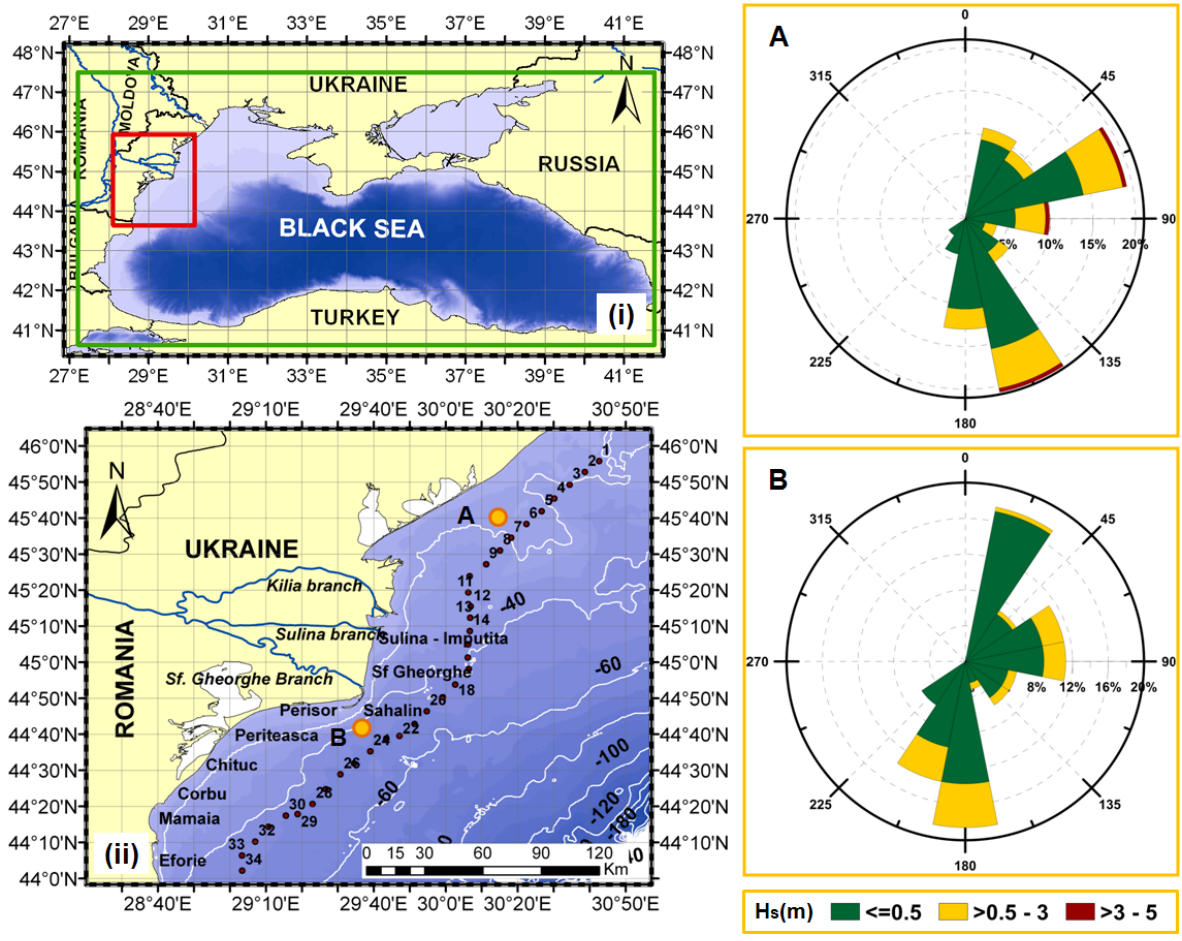


Figure 1. (**i**) Map of the Black Sea. The computational domain for the Simulating WAves Nearshore (SWAN) model is enclosed by a green rectangle (see first step of the proposed methodology), whereas the northwestern Black Sea is enclosed by a red rectangle (see the second step of the same methodology). (**ii**) Map of the northwestern Black Sea. Nodes from the statistical model are in red and are numbered from north to south. The right figure shows wave-roses at points **A** and **B** (orange dots) from the map of the northwestern Black Sea [29]. The bar on the right-bottom shows the wave height-ranges at the wave-rose.

A wave-storm year is a year-long period of intense wave-storm activity, which ranges in the present approach from July of the previous year to June of the next year. During a wave-storm year, the most relevant pattern is determined by the relative position, displacement, and resulting interactions between the Mediterranean cyclones and the Eastern European (Siberian) anticyclone. The most intense and frequent winds affecting the coast are those from the northeast, east and southeast. They have the longest fetch and produce the most severe wave-storms. However, as this is a fetch-limited environment, little energy is absorbed from the wind forcing, and wave-periods tend to be shorter than in large water bodies such as the Atlantic Ocean.

The average significant wave-height at deep waters in the northwestern Black Sea range between a minimum of 0.35 m in spring (March–May) to a maximum of 0.75 m in winter (from December–February) [28]. The average peak wave-period ranges from 1.8 s in spring to 2.4 s in winter. Due to the influence of the wind pattern, waves propagate most frequently from the east, northeast and southeast [30]. Eastern waves are predominant within the entire shelf zone (see Figure 1). Their directional sector frequency ranges between 30% and 40%. The fraction of northeastern waves has a frequency of 30% and the frequency of southeastern waves is over 10% [4]. The most energetic months are December to February [28,31]. Winter wave-storms are much more frequent than summer ones [4]. The average wave-storm duration between 1980 and 1993 was 30 h, whereas the maximum wave-storm duration was about 130 h. The most energetic wave directions were northerly, whereas the average wave-heights during wave-storms were 1.5–4.5 m.

3. Methods

The proposed hybrid methodology (See Figure 2) consists of four stages: (i) Wave-projections at the Black Sea generated with a process-based model; (ii) set-up of a multivariate non-stationary statistical model; (iii) validation of the non-stationary statistical model; (iv) comparison of the obtained wave-storm components with those derived from other general circulation models. The same methodology was applied for each climate change scenario (RCP 4.5 and RCP 8.5).

3.1. First Step: Process-Based Dynamical Modelling

The temporal coverage in this study spans from 1950 to 2100. Two climate change scenarios were considered in this study: RCP 4.5 and RCP 8.5 scenarios. A general circulation model (GCM), with a basis in climate-assessment studies [32–34], provides the general circulation of the Earth's atmosphere. Although the spatial coverage is global, the grid size is too coarse for modelling wave forcings (i.e. wind-fields) at a relatively small basin such as the Black Sea. In this case, spatial and temporal resolution (plus the addition of physical processes that need to be considered at a regional scale) can be downscaled with a regional circulation model (RCM).

The process to build the projections starts by dynamical downscaling of the CRNM-CM5 GCM (see Figure 2), through the ALADIN 5.2 RCM [35–37], and then projecting ocean-waves with the SWAN model. ALADIN 5.2 is a bi-spectral RCM that uses a double Fourier representation for spectral fields. It uses a semi-implicit, semi-Lagrangian advection scheme. The main parameterizations employed are: a mass-flux scheme with convergence of humidity closure for the convection [38]; the statistical cloud scheme by [39]; and the large-scale precipitation scheme by [40]. The RCM lateral boundary conditions were obtained from the CRNM-CM5 GCM outputs.

RCP4.5 and RCP8.5 wind-fields from the ALADIN model were downloaded from the Mediterranean Coordinated Regional Downscaling Experiment (Med-CORDEX) initiative [41]. Wind fields coincided for both RCP scenarios at the 1950–2005 period (historical time slice), whereas the fields for the future period (2005–2100) differed at each RCP. These wind fields span the whole of Europe with a spatial and temporal resolution of 12 km$\times$12 km and 3 h, respectively. This spatio-temporal resolution for the wave forcings follows the state-of-the-art in future wave climate projections [42–44].

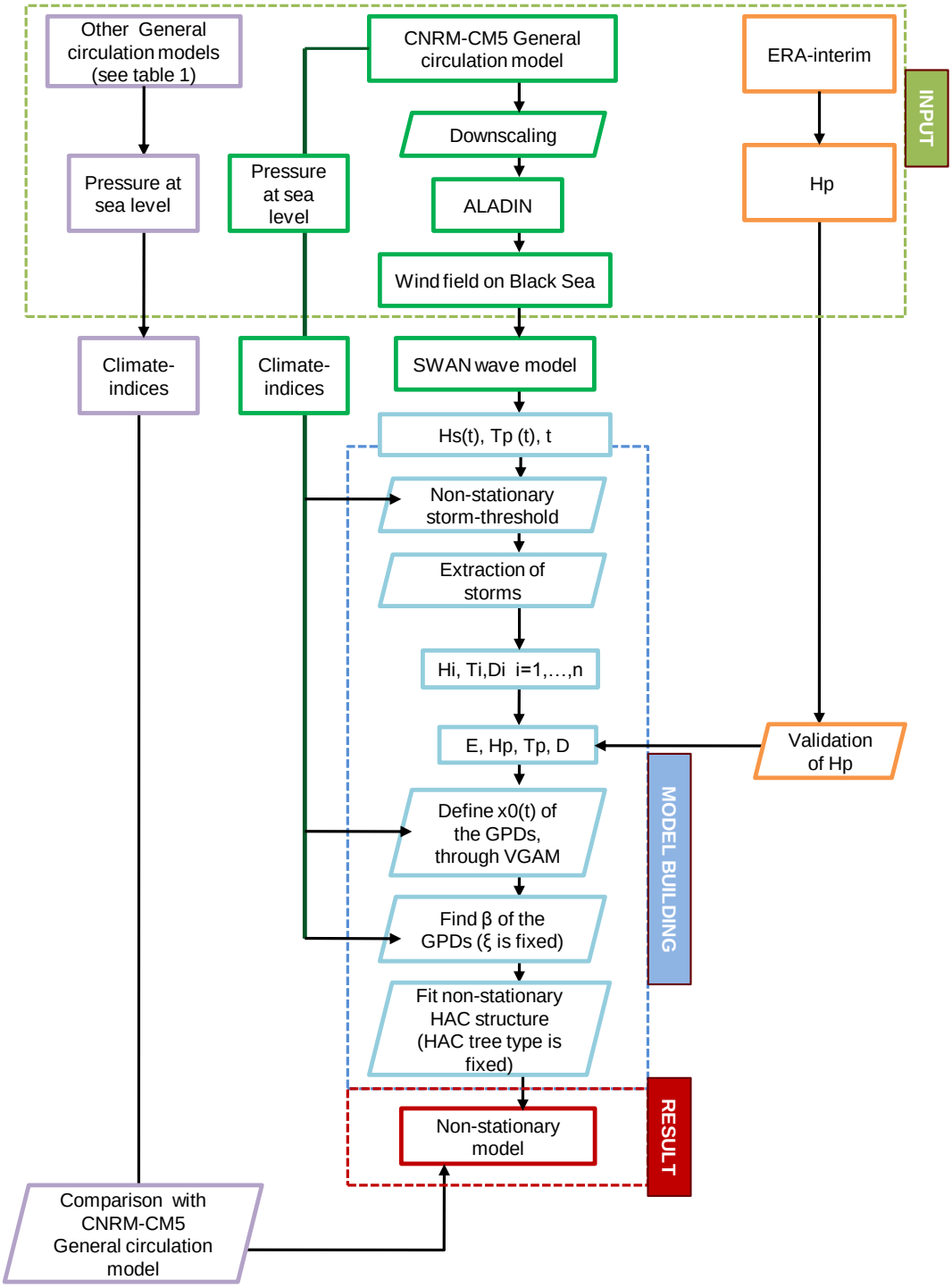


Figure 2. Flow-chart of the different steps in the proposed hybrid methodology. This methodology is analogous for each climate change scenario (Representative Concentration Pathways, RCP, 4.5 and 8.5). CNRM-CM5 is the General (Atmospheric) Circulation Model used to build the statistical model. ERA-interim is the reanalysis used to validate the statistical model. Each variable is fit by a generalized Pareto distribution (GPD) function and their joint probability structure is characterized by a series of hierarchical Archimedean copulas (HAC). Rectangles are the results, rhombuses are the methods. Different colors are only intended to separate different stages of the process. Elements inside the "input" box have been obtained from external sources. The side analyses are not included in the graph, for the sake of clarity.

These winds serve as input for the SWAN spectral wave-model [45]. The computational domain spans the whole Black Sea with a regular grid of 9 km×9 km (see the domain marked with a green solid line in Figure 1). The bathymetry comes from the GEBCO dataset (GEB-2008), which has a spatial resolution of approximately one arc-minute. The spatial resolution taken can reproduce the wave generation and the deep-water wave propagation phenomena in the study area. Hence, wind wave-growth, quadruplet interaction, whitecapping dissipation, and bottom friction terms are activated, whereas triad interaction and wave breaking are deactivated.

The SWAN model was run in a non-stationary mode with a time-step of 20 min for the whole 1950–2100 period. Wind fields are updated every 3 h with a spatial resolution of 12 km×12 km. Wave outputs are saved hourly at a subset of computational nodes. These outputs mainly consist of time series of integrated wave-spectra parameters: significant wave-height, peak and mean wave-period, wave-direction, among others. Once these time-series are obtained , the second stage of the proposed hybrid methodology is applied at the 34 selected nodes of the northwestern Black Sea (see Figure 1).

The second stage in the proposed methodology (see the next section) consists of fitting a non-stationary multivariate model that has as response variables a set of storm components that use large-scale climate indices as predictors (covariates). These storm components come from the SWAN time series obtained in this first step.

3.2. Second Step: The Statistical Model

3.2.1. Definition of Wave-Storms and Their Components

The second stage in the proposed analysis deals with the construction of a multivariate non-stationary statistical model. This model assesses the changes in the wave-storm components during 1950–2100 in the northwestern Black Sea (see Figure 1). Although models have long been built considering stationarity, there is a series of drawbacks with this approach. Extreme events are rare, and in the cases where several time-windows were considered (e.g., a window of less than 15 years), samples of high extreme events in each time window would not be statistically significant. Consequently, the estimated upper tail of the probability distribution function would not yield reliable results. Further, climate change has a non-negligible effect on extremes, threatening the foundations of assumptions such as stationary wave-storm thresholds. A non-stationary model can overcome these shortcomings because it can include the effects of climate change on extreme wave-climate.

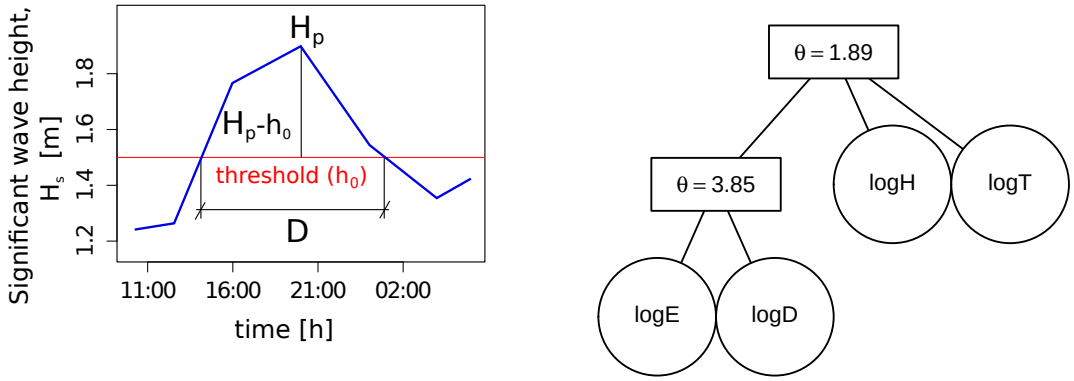
The sample of wave-storms were extracted from the projections (see Figure 2) by splitting the projections into mean sea-conditions and extremal conditions by means of a wave-storm threshold (see Figure 3a). As a preliminary analysis, a model considering stationarity in time-windows of 50 years was built separately, in order to obtain an initial approximation for the parameters of the non-stationary statistical model built here. Wave-height thresholds in the stationary model were $h_{0,RCP4.5} = 1.8$ m for RCP4.5 and 2.0 m $= h_{0,RCP8.5}$ for RCP8.5, which are the corresponding 90th quantile [46] at node 19. Wave-storms in the sample are independent. An independent wave-storm is defined as having a minimum duration of 6 h and a minimum time-interval between wave-storms of 72 h [11]. A sensitivity analysis on the minimum time-interval was carried out, testing for two possible minimum time-intervals: the proposed one and 12 h.

The storm-threshold is obtained through vectorial generalized additive models (VGAMs, [47,48]), using climate indices as covariates. All the nodes use the same climate index as covariate for a given RCP scenario, following the principle of parsimony. The covariates of a VGAM are predictors, whereas the predictand is the wave-storm threshold. VGAM presents the linear function [49]:

$$\eta_{i(j)} = \beta_{1(j)} + f_{2(j)}\left(x_{i2}\right) + \ldots + f_{p(j)}\left(x_{ip}\right), \quad (1)$$

where $\eta_{i(j)}$ is the jth dependent variable, x_i is the ith independent variable that generates η_i. η_i is a sum of smooth functions of the individual covariates. Additive models do all smoothing in $\mathbb{R}$,

avoiding large bias introduced in defining areas in $\mathbb{R}^n$. The assumptions for regression models are: (1) independence of residuals; (2) residuals follow a Gaussian distribution of the form $N\left(0,\sigma^2\right)$; and (3) residuals are homoscedastic. Assumption (1) is tested with an autocorrelation function plot [50]. Assumption (2) can be tackled with a Q-Q plot comparing the empirical distribution of the residuals to a $N\left(0,\sigma^2\right)$ distribution, where the sample standard deviation is used as σ^2. Assumption (3) can simply be visually analysed through a scatter-plot of the fitted values vs. the residuals.



(**a**) (**b**)

Figure 3. (**a**) Schematic view of the definition of a wave-storm. The blue thick line represents a wave, the red thin line represents a wave-storm threshold of value h_0. (**b**) Hierarchical Archimedean copula (HAC)-structure-tree at node 13, in the period 2001–2050, under the RCP4.5 scenario. The circles contain the names of the storm components, which are bound together by a dependence-parameter, written in the rectangles ($\theta \in [0,\infty)$). This HAC-structure-tree comes from the stationary statistical model. It is representative of all nodes, at each time-window of 50 years and in both emission scenarios. Then, it is selected to represent the dependence-structure of the whole non-stationary statistical model.

The VGAM of the statistical model built in this paper uses as covariates the climate indices in the GCMs that represent the large-scale climate patterns of the North Atlantic Oscillation (NAO), the Eastern Atlantic Pattern (EA) and the Scandinavian Pattern (SC) [11,51]. The first and second time-derivatives of these climate indices are employed as well. This adoption of climate indices as covariates was previously discussed in [11]. These large-scale indices can be derived from monthly sea-level pressure-fields that can be downloaded from the CMIP5 Project's website. They have been scaled to have a mean value equal to zero and a variance equal to unity. In order to avoid sudden oscillations that would hinder interpretation, they have been filtered with a second-order low-pass Butterworth filter [52], whose low-pass period was 10 years. The Akaike information criterion [53] and the Bayesian information criterion [54] are applied on the VGAM to test the sensitivity of the wave storm threshold to each climate index.

Storms are clustered by storm-years (referred to hereafter as "years"). Storminess can be estimated by approximating its relationship with the selected climate index and time derivatives by a Poisson probability distribution function. As it is a counting variable, a vectorial generalized linear model (VGLM, a particular case of VGAM [47]) can be adopted. A storm-threshold is estimated by approximating its relationship with the climate index by a Laplace function. The Akaike information criterion and the Bayesian information criterion can also be applied to the VGLM to test the sensitivity of storminess to each climate index. The averaged value of the regression coefficient of the VGAM at nodes 13 and 29 ($\overline{C}_r$) is used to quantify the influence of the climate index chosen. C_r is the coefficient that multiplies the predictor x_{i2} in the function $f_{2,(i)}\left(x_{i2}\right) = C_r \cdot x_{i2}$ from Equation (1), the predictor

being a climate index. $\overline{C}_r$ is back-transformed from its logarithmic form to its natural form in Sections 4 and 5 ("Results" and "Discussion") for ease of understanding. If Equation (1) was rewritten as

$$\eta_{i(j)} \approx \beta_{1(j)} + C_r \cdot x_{i2},$$

for $x_{i2} \neq 0$, both sides of the equation could be divided by x_{i2} and become

$$\eta_{i(j)} / x_{i2} \approx \beta_{1(j)} / x_{i2} + C_r.$$

The presence of x_{i2} close to zero—which would make such division an indetermination—can be avoided by translating x_{i2} from the interval $(-2, 2]$ to the interval $(0, 4]$. Therefore, the two terms on the right side—$\beta_{1(j)} / x_{i2}$ and C_r—could be comparable. $\beta_{1(j)} / x_{i2}$ can be herein defined as the ratio between the wave-storm threshold and the climate index. This ratio is approximated by the mean of the natural value of the wave-storm threshold divided by double the maximum climate index value.

The wave-storms are represented by the following storm components: storm-energy (E), significant wave-height and peak-wave-period at the peak of wave-storm (H_p and T_p, respectively), and storm-duration (D). The definition of E is

$$E = \frac{1}{D} \int_{t_i}^{t_f} H_s^2 (t)\, dt, \tag{2}$$

where t_i and t_f are the starting and ending time of a wave-storm. The inclusion of E and D are thought to provide more information on the general behaviour of a wave-storm, whereas H_p and T_p represent upper limit wave conditions of a wave-storm. E, H_p, T_p, and D take positive real values. Consequently, they are log-transformed to avoid scale effects when building the statistical models [55], but are still referred to as E, H_p, T_p, and D in the interpretation of the models, for ease of understanding. The mean wave-storm wave-directions have also been extracted for a complementary side analysis, which serves to describe the evolution of wave directionality in 1950–2100.

3.2.2. Generalized Pareto Distribution: Univariate Distribution-Function

Storm-thresholds from the stationary statistical model are used as starting values to iterate for non-stationary location parameters x_0 of H_p. The starting value for a non-stationary location parameter x_0 of T_p is 5.8 s for both RCP4.5 and RCP8.5 scenarios. The starting value for the non-stationary location parameter x_0 of D is defined as the minimum storm duration (6 h) and the starting value of the non-stationary location parameter x_0 of E is $E_{0,RCP4.5} = h^2_{0,RCP4.5} \cdot 6 = 19.4\ \mathrm{m}^2/\mathrm{h}$ ($E_{0,RCP8.5} = 24.0\ \mathrm{m}^2/\mathrm{h}$). All these starting values for non-stationary location parameters x_0 fall on the linear part of corresponding excess-over-threshold functions, so their probability distributions can be modelled by generalized Pareto distributions (GPDs, [11,14,55,56]).

The definition of a GPD is as follows. $Y = X - x_0$ is the excess of a magnitude X over a location parameter x_0, conditioned to $X > x_0$, the support of Y is $\left[0\,,\, y_{sup}\right]$, where y_{sup} is the upper bound of y, if it exists [57]. The GPD cumulative function is

$$\mathrm{F_Y}\left(y|\beta,\xi\right) = 1 - \left(1 + \frac{\xi}{\beta} y\right)^{-\frac{1}{\xi}},\ 0 \leq y \leq y_{sup}, \tag{3}$$

where $\beta \geq 0$ is the scale parameter and $\xi \in \mathbb{R}$ is the shape parameter. Given the location parameter x_0 and the scale parameter β, the mean value of a wave-storm component is given by:

$$\mathrm{E}\left(X\right) = x_0 + \frac{\beta}{1-\xi} \approx x_0, \qquad \xi < 1, \tag{4}$$

and the variance of a wave-storm component is:

$$\mathrm{Var}\,(X) = \frac{\beta^2}{(1-\xi)^2\,(1-2\xi)} \approx \beta^2, \qquad \xi < \frac{1}{2}, \tag{5}$$

where the parameters are the same ones as in Equation (3). Therefore, the location parameter x_0 and the scale parameter β provide information on the mean value and the variance, respectively.

A quantile regression—a particular type of VGAM [58,59]—is selected to estimate the location parameters x_0 of the GPD in the non-stationary statistical model. It estimates the $100\hat{\tau}$% conditional quantile $y_{\hat{\tau}}\,(x)$ of a response variable Y as a function $u\,(x,\tau)$ of covariates x. $l_u^* = l_u + \varrho_u R_u$ is minimized, where $l_u = \hat{\tau} \sum_{i:r_i \geq 0} |r_i|\,(1-\hat{\tau}) \sum_{i:r_i<0} |r_i|$ for residuals $r_i = y_i - u\,(x_i,\hat{\tau})$. ϱ is a roughness coefficient that controls the trade-off between quality of fit to the data and roughness of the regression function. R is a roughness penalty [60,61]. As a side analysis, a VGAM that uses time as the single covariate helps visualize the non-stationarity of the projected location parameters x_0.

An assumption in this paper for the non-stationary GPDs is that the shape-parameter ξ remains constant, while the scale parameter β can depend on co-variates. β is estimated through a VGLM. The sensitivity of the GPD parameters to the proposed climate indices and time-derivatives is tested by applying the same Akaike and Bayesian information criterion on the VGAM or VGLM. The averaged value of the regression parameters of the VGLM and the quantile regression at nodes 13 and 29, $\overline{C}_r$, is used again to quantify the influence of the predominant climate index. It is compared to the ratio between the wave-storm component and the climate index.

3.2.3. Copulas: The Joint-Dependence Structure

Wave-storm components are semi-dependent, according to graphical dependence-tests (not shown). Because of this semi-dependence, the non-stationary joint-dependence-structure of the storm components may be parametrized by hierarchical Archimedean copulas (HACs, [62]). A copula is a multivariate distribution function with standard univariate margins [63–65]. A d-dimensional copula is Archimedean [66] if it is of the form

$$\mathrm{C}\,(\mathbf{F};\boldsymbol{\phi}) = \boldsymbol{\phi}^{-1}\,(\boldsymbol{\phi}\,(F_1) + \cdots + \boldsymbol{\phi}\,(F_d))\,, \quad \mathbf{F} \in [0,1]^d\,,$$

for a given generator function $\boldsymbol{\phi}$. Being $\phi_1, \ldots, \phi_d$ the Laplace-Transform generators, for $1 \leq k \leq d-2$, $k+1 < j \leq d$ and $v, \tilde{v} > 0$, multi-dimensional variables can be nested, as the following proves true:

$$\exp\left(-v\phi_{k+1}^{(k)-1}\left(\phi_j^{(k)}\,(\cdot;\tilde{v})\,;\tilde{v}\right)\right) = \phi_j^{(k+1)}\,(\cdot;v)\,.$$

A HAC can be a useful tool to nest simple Archimedean copulas into larger and more complex ones (see Figure 3b). A HAC provides a dependence parameter, θ, at each nesting level. θ can be transformed into other correlation parameters, such as Kendall's τ [66–68], the interpretation of which is more straightforward: $\tau = 0$ indicates independence, whereas $\tau \to 1$ means total dependence. This τ should not be mistaken by the $\hat{\tau}$ in Section 3.2.2.

To model the non-stationary dependence, the storm components are clustered into periods of 15 years, in order to compute the dependence parameters for each block and node. This creates a pseudo-non-stationary HAC [11] per node. The clusters are of the size 15 years, because a smaller size would provide an insufficient number of wave-storms to characterize the joint probability structure. On the other hand, the smaller the blocks, the more non-stationary the HAC would be. Therefore, 15 years is the optimum number in the trade-off between the lack of wave-storms for fitting and the maximum size of the blocks. Additionally, each time-window overlaps with the previous and the later time-window, in order to reflect the non-stationarity.

The HAC type and structure have been suggested by the stationary model built apart. The HAC tree in Figure 3b—which is a Gumbel-type HAC—represents the HACs used in all computational nodes and for both RCPs in the non-stationary statistical model. Wave-storms typically produce extreme values for several storm components (i.e., H_p and E) at the same time. If extremal dependence [69,70] among the storm components is evident, Gumbel-type Archimedean copulas can include such upper-tail dependence [15,17], as discussed previously in [14] and [11]. The HAC can be tested by a goodness-of-fit test in the stationary statistical model [71]. The κ^2 statistic [72] can serve to quantify this goodness-of-fit. κ^2 takes values in the interval $[0, 1]$, where a perfect fit corresponds to $\kappa^2 = 1$. The Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test [73] is applied on the dependence parameters of the pseudo-non-stationary HACs to confirm that these are non-stationary. The p-value in a KPSS test gives the level of significance at which the null test (the necessity of a non-stationary model to characterize the HAC parameters) cannot be rejected.

3.3. Third Step: Validation of the Non-Stationary Statistical Model

The data used to validate the statistical model in both emission scenarios is the ERA-interim reanalysis [43,74–77]. "ERA" stands for "European Reanalysis". This global reanalysis, despite having a spatial resolution of 80 km, presents gapless, bias corrected information from 1979 to present day. The ERA-interim data available for the northwestern Black Sea is at 43.5°N, 31.5°E. The nearest node from the wave-model is node number 19, at 44.8°N, 30.0°E (see Figure 1). Node 19 is central to the northwestern Black Sea, and is thus considered to be able to represent the study area. The validation period is 1979–2016, which is shorter than the 1950–2100 (control period) of the wave-climate projections. Therefore, it is only intended to validate the years 1979–2016.

ERA-interim (validation) wave-storms are extracted by using the same non-stationary thresholds as the projections. The ERA-interim wave-storms do not coincide in timing with the projected wave-storms, as the latter reflect a climate signal and do not put emphasis on the prediction of the exact timing. With this in mind, [11] proposed a method to estimate the likelihood of the projected probability density function to the probability density function of measurements. The H_p of ERA-interim data is:

$$\left\{H_{p,1}, \ldots, H_{p,i}, \ldots, H_{p,n}\right\}, \quad i = 1 \div n,\ n \in \mathbb{R},$$

and the H_p of the model data (written as H_p^*, here)

$$\left\{H^*_{p,1}, \ldots, H^*_{p,j}, \ldots, H^*_{p,m}\right\}, \quad j = 1 \div n,\ m \in \mathbb{R}.$$

H_p from both sources are combined to form a joint dataset:

$$\left\{H_{p,1}, \ldots, H_{p,i}, \ldots, H_{p,n}, H^*_{p,1}, \ldots, H^*_{p,j}, \ldots, H^*_{p,m}\right\}.$$

Such a set is partitioned into ten intervals, separated by the quantiles that are multiples of 10. There are elements from both ERA-interim H_p and the projections in each interval. The probability of falling into each one of these intervals must be similar for the model and the reanalysis, so that the model can be validated.

Two vectors are defined as

$$vec_{obs,k} = \sum_{q_k}^{q_{k+1}} p\left(H_{p,i}\right), \qquad k = 1, \ldots, 11$$

and

$$vec_{model,k} = \sum_{q_k}^{q_{k+1}} p\left(H^*_{p,j}\right), \qquad k = 1, \ldots, 11,$$

where vec_{obs} is the vector for ERA-interim reanalyses and vec_{model} is the one for projections. vec_{obs} and vec_{model} are compositional data, their elements being parts of a whole [78] and can be defined in the Aitchison space [79,80]. The distance between these two vectors can be measured with an Aitchison distance [81,82],

$$d(\mathbf{x},\mathbf{y}) = \left| \ln \frac{\mathbf{x}(\mathbf{1}-\mathbf{y})}{\mathbf{y}(\mathbf{1}-\mathbf{x})} \right|, \quad \mathbf{x},\mathbf{y} \in (0,1) \in \mathbb{R}, \tag{6}$$

where $\mathbf{x}$ and $\mathbf{y}$ are vec_{obs} and vec_{model}, respectively, in this case. Another measure is the Kullback–Leibler divergence [83]:

$$D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}. \tag{7}$$

This function measures the extra entropy of the probability distribution Q of the model, with respect to the probability distribution P of the observations. Note that for any i, $Q(i) = 0$ must imply $P(i) = 0$, to avoid indetermination, thus ensuring that the model considers all the values that the observations show. Additionally, whenever $P(i) = 0$, the contribution of the i-th term is null, as $\lim_{x \to 0} x \log(x) = 0$. Both Equations (6) and (7) are measures, and thus take values in $\mathbb{R}_0^+$. The module of a vector is a particular case of both measures [78], and thus both measures can be compared to the vectorial modules, in Euclidean space, of $\mathbf{x}$ and $\mathbf{y}$, which are of order 1.

3.4. Fourth Step: Comparison of the Different GCMs

The non-stationary statistical model built in the second step uses large-scale climate indices derived from GCM outputs (pressure-fields at sea-level). The SWAN projections (first step) were built from MedCORDEX ALADIN wind-fields. These wind-fields result from a dynamical downscaling of the CNRM-CM5 GCM. Hence, to put it clearly, the fitting of the statistical model (second step) consists of using CNRM-CM5 projections of large-scale climate indices as predictors to estimate probability distribution functions of the wave-storm components (response variables) derived from the SWAN time-series.

One of the main advantages of the proposed statistical model is that it can predict how the probability-distribution-functions of the storm components would vary if other projections of large-scale climate indices were used as input. The first step of the proposed methodology requires a computational cost that hampers the simulation of waves with different GCMs wind forcings. The proposed statistical model can partly alleviate this computational burden via the establishment of relationships between storm-wave components and climate indices. However, there is significant uncertainty in climate change projections, and it is required to consider the results from several GCMs prior to extracting conclusions. In this regard, a comprehensive list of GCMs that address climate change scenarios can be found at the CMIP5 (Coupled Model Intercomparison Project Phase 5, [84]) Project site. The GCMs in Table 1—other than CNRM-CM5—have been employed to compare how much they can differ from this GCM in extreme wave-climates at the Northwestern Black Sea. Note that most of these models are centred on the European continent, except the MIROC branch (MIROC-ESM, MIROC-ESM-CHEM and MIROC5), which focuses on the area of the Japanese Isles [85], and the GFDL branch (GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M), which focuses on the United States [86].

The comparison of different GCMs can be carried out by contrasting the 99 th quantiles of the storm components projected with the CNRM-CM5 GCM to the 99 th quantiles of the storm components projected with other GCMs (see Table 1). Partial autocorrelation Function (PACF, [87]) based distances are used as a metric [87]. It measures the distance among time-series based on the corresponding PACF, and is a model-free approach. That is, the two time-series compared—X_a and Y_b—do not have to belong to specific time-series models. The PACF coefficient measures the correlation of pairs of elements from each time series at all shorter time lags (i.e., the relationship between two points at one hour apart, two hours apart, and so on). Let $\hat{\rho}_{X_T} = \left(\hat{\rho}_{1,X_T}, \ldots, \hat{\rho}_{L,X_T}\right)^T$

and $\hat{\rho}_{Y_T} = \left(\hat{\rho}_{1,Y_T}, \ldots, \hat{\rho}_{L,Y_T}\right)^T$ be the estimated autocorrelation vectors of X_T and Y_T, respectively, for some L such that $\hat{\rho}_{i,X_T} \approx 0$ and $\hat{\rho}_{i,Y_T} \approx 0$ for $i > L$. Then, an autocorrelation distance is defined as $\mathrm{d}_{ACF}\left(X_T, Y_T\right) = \sqrt{\left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)^T \Omega \left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)}$, where Ω is a matrix of weights of value ≤ 1, and $0 \leq \mathrm{d}_{ACF}\left(X_T, Y_T\right) \leq 2$, where 0 corresponds to total coincidence, and 2 to total discordance.

Table 1. List of general circulation models (GCM) employed. The GCM CNRM-CM5 is used to build the non-stationary statistical model, whereas the other GCMs are compared to CNRM-CM5. CNRM-CM5 is shown in boldface to ease its search in the list.

GCM	Latitude Grid Size (°)	Longitude Grid Size (°)
CMCC-CM	0.7484	0.7500
CMCC-CMS	3.7111	3.7500
CNRM-CM5	**1.4008**	**1.4063**
FGOALS-G2	2.7906	2.8125
GFDL-CM3	2.0000	2.5000
GFDL-ESM2G	2.0225	2.0000
GFDL-ESM2M	2.0225	2.5000
HadGEM2-AO	1.2500	1.8750
HadGEM2-CC	1.2500	1.8750
HadGEM2-ES	1.2500	1.8750
INM-CM4	1.5000	2.0000
IPSL-CM5A-LR	1.8947	3.7500
IPSL-CM5B-LR	1.8947	3.7500
IPSL-CM5A-MR	1.2676	2.5000
MIROC-ESM	2.7906	2.8125
MIROC-ESM-CHEM	2.7906	2.8125
MIROC5	1.4008	1.4063
MPI-ESM-LR	1.8653	1.8750
MPI-ESM-MR	1.8653	1.8750

4. Results

This section is organized as follows. For both emission scenarios, it addresses: (a) a sensitivity test on the time-interval between wave-storms; (b) the evolution of wave directionality in 1950–2100; and (c) a test of the assumptions of the VGAM. For each emission scenario, it addresses: (i) bounding of the uncertainty of the GCM; (ii) value of wave-storminess and its relationship to the selected climate indices; (iii) values of location parameters x_0 of GPDs and their relationship to the selected climate indices; (iv) values of scale parameters β of GPDs and their relationship to the selected climate indices; (v) values of the parameters in the dependence structure; and (vi) validation of the non-stationary statistical model.

A sensitivity analysis on the minimum time-interval between storms equal to 12 h provides a storminess of approximately 50 storms/year at some nodes, in both RCP scenarios, which is unrealistic. Therefore, a minimum time-interval between storms equal to 72 h is used. The wave-storm mean wave-directions for the whole area of study and in both RCP scenarios are the same ones as in historical records (see Figure 1). They stay the same over the period of study. Therefore, the wave-directionality can be considered stationary, and emphasis can be put only on the wave-storm components E, H_p, T_p, and D. Neither independence nor normality assumptions of the residuals of the data used in the VGAM can be rejected. Therefore, jointly with the compliance of the wave-storm thresholds with the mean-excess plots (see Section 3), the wave-storm components can be modelled by non-stationary GPDs.

4.1. RCP4.5

The GCMs coincide mainly at wave-storm components E and D, with PACF based distances below 0.1. In the case of H_p and T_p, however, the maximum PACF-based distance can be 0.3 and 0.2, respectively. The GCMs presenting the largest PACF-based distance in the case of H_p and T_p can come from any of the branches of GCMs in Table 1. Employing the CNRM-CM5 GCM, and assuming non-stationarity, the estimated average number of wave-storms per year ranges from 27 storms/year at node 1 to 35 storms/year at node 29, then decreases to 34 at node 33. Although SC is the most influential climate index on storminess, the estimated number of storms per year is affected by neither time nor climate indices. The storm-threshold is most sensitive to a combination of the first and the second time-derivatives of SC.

The location parameters x_0 of the GPD under RCP4.5 are shown in Figure 4. A selection of nine out of thirty-four nodes are represented. There is an upward trend of the location parameter x_0 of E and D, while the location parameter x_0 of H_p and T_p has a constant value. The averaged value (in natural values) at nodes 13 and 29 of the mean of E, H_p, T_p, and D are 144.2 m^2/h, 2.7 m, 6.6 s, and 33.0 h. The assumption in Equation (4) is fulfilled, because the shape parameter ξ is negative. The most influential covariates on the GPD location parameter x_0 are EA for E ($\overline{C}_r = 1.02$) and D ($\overline{C}_r = 1.02$), the first time-derivative of NAO for H_p ($\overline{C}_r = -1.03$) and SC for T_p ($\overline{C}_r - 1.01$). The ratio between the wave-storm component and the climate index is 35 in the case of E and of the order of magnitude of 1 in the rest of the wave-storm components.

The temporal evolution of the scale parameter β is shown in Figure 5, where the trend of the location parameters x_0 are reflected in the scale parameters β. The averaged value at nodes 13 and 29 of the variance of E, H_p, T_p, and D are 8.4 m^2/h, 1.1 m, 1.0 s and 7.6 h. The assumption in Equation (5) is fulfilled, as the shape parameter ξ is negative. The most influential covariates on the GPD scale parameters β are the second time-derivative of SC for E ($\overline{C}_r = 1.17$), EA for H_p ($\overline{C}_r = -1.00$), and D ($\overline{C}_r = 1.01$), and the first time-derivative of EA for T_p ($\overline{C}_r = -1.42$). The ratio between the wave-storm component and the climate index is approximately 80 in the case of E; and approximately of the order 1 in the rest of the wave-storm components. These trends and relationships to the selected climate indices will be further discussed in Section 5.

The Gumbel-type HAC is suitable for the proposed statistical model under RCP4.5. It has significantly good graphical fit between empirical and theoretical HAC, while κ^2 is over 0.90 at each level of the HAC. The stationarity of the τs cannot be rejected, because the p-value from the test is 1.0%. This proves that the HACs used in the proposed models are indeed non-stationary. The dependence among variables, reflected by Kendall's τ, is shown in Figure 6. τ_{root} ranges between 0.55 and 0.65, and has a positive trend at all nodes. $\tau_{(E,D)}$ ranges between 0.82 and 0.87, and apparently presents cyclical fluctuations. The test of the assumptions for VGAM shows conformity for the non-stationary threshold of the ERA-interim data, which serves to validate projections in both RCP scenarios. The non-stationary statistical model is validated for 1979–2016, as both types of measures (Equations 6 and 7) are invariably below 1.0.

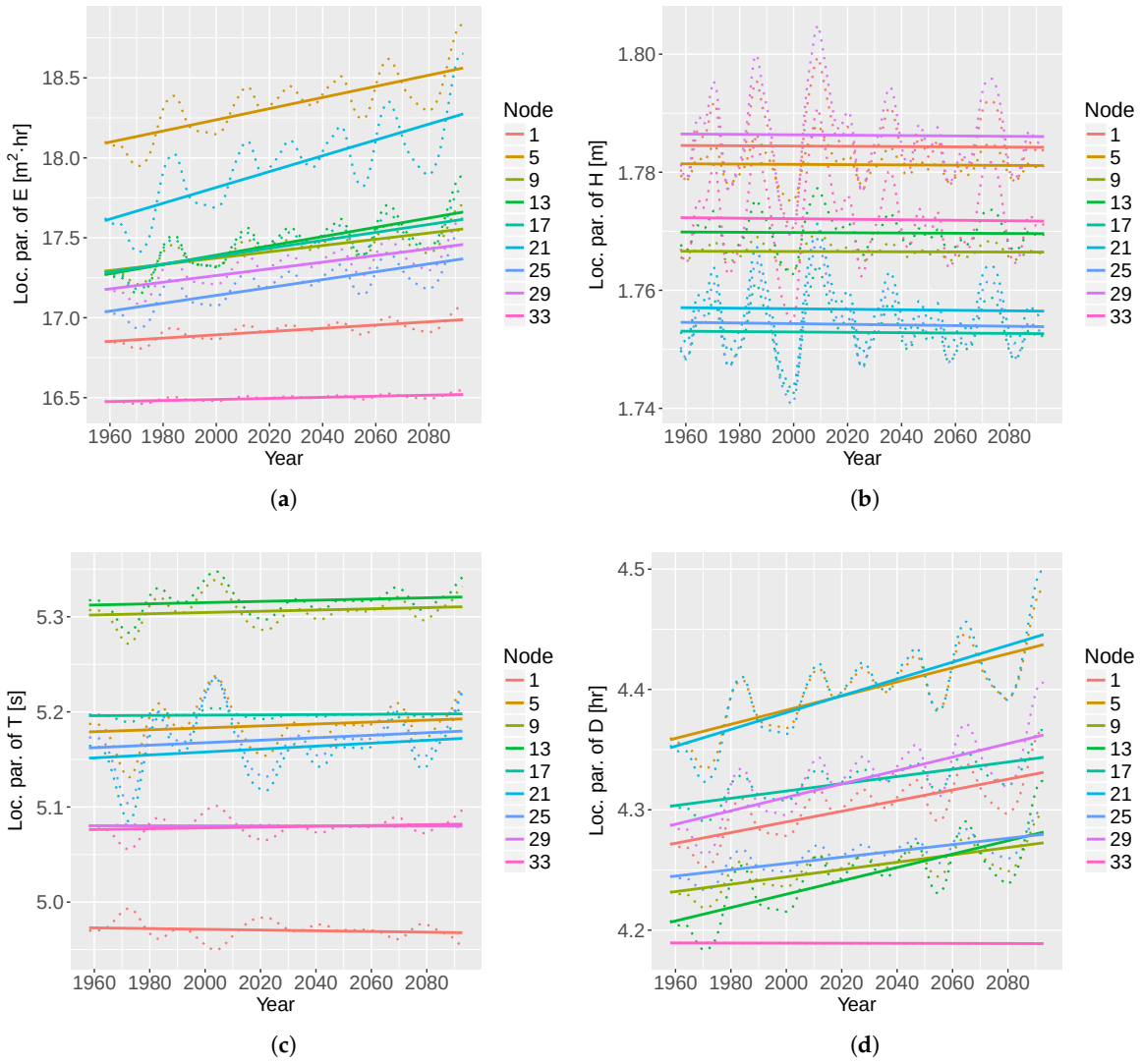


Figure 4. Location parameter x_0 of the generalized Pareto distribution (GPD) that models (**a**) E, (**b**) H_p, (**c**) T_p, and (**d**) D at selected nodes, under the RCP4.5 scenario. The Eastern Atlantic Pattern (EA) was used in the vectorial generalized additive model (VGAM) to predict the location parameter x_0 of E and D. The first time-derivative of the North Atlantic Oscillation (NAO) was used in the VGAM to predict the location parameter x_0 of H_p. The Scandinavian Pattern (SC) was used in the VGAM to predict the location parameter x_0 of T_p. The time series are approximated by a straight line, which helps to see the trend of the location parameters x_0 in each case.

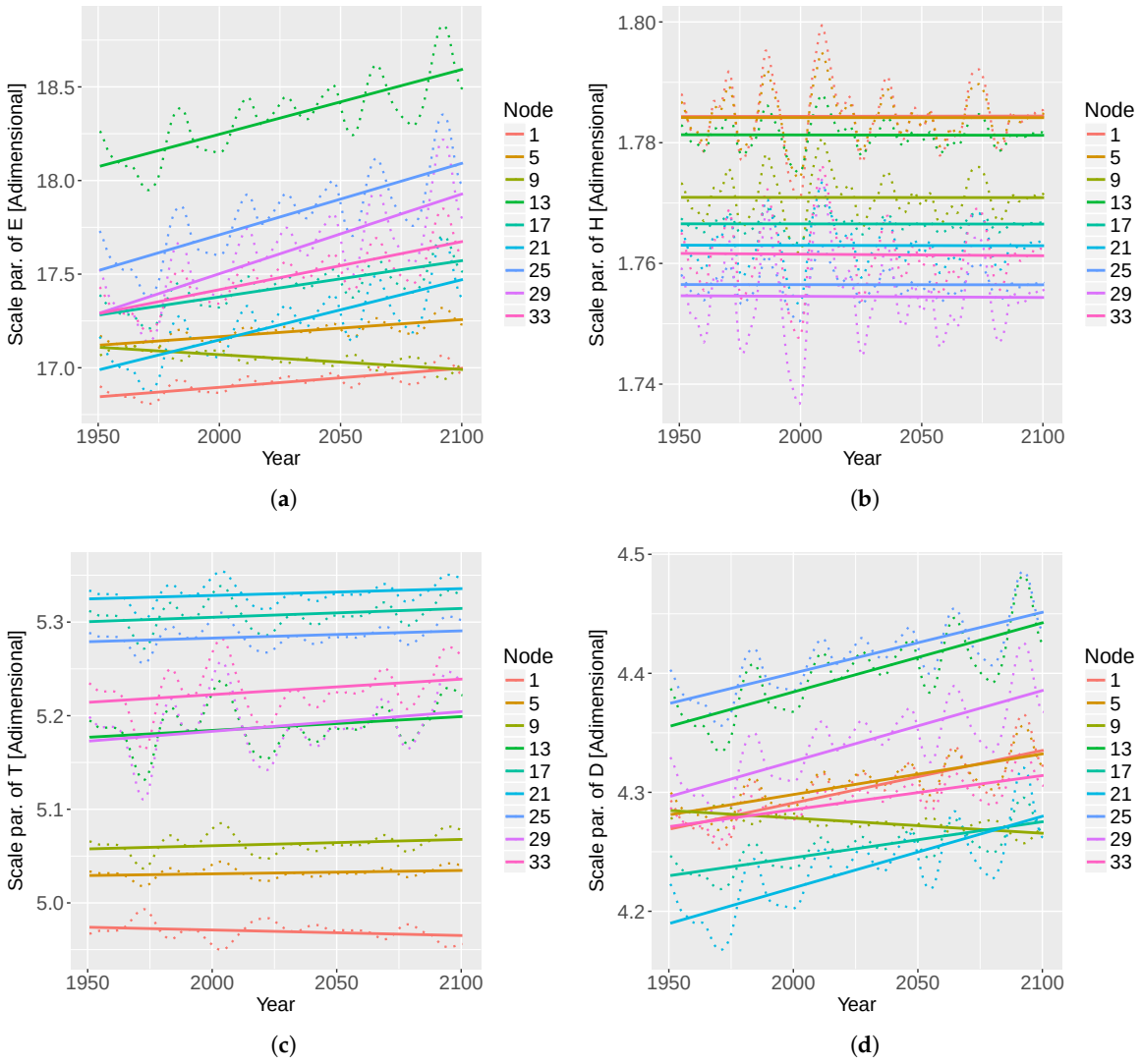


Figure 5. Scale-parameter β of the GPD that models (**a**) E, (**b**) H_p, (**c**) T_p, and (**d**) D at selected nodes, under the RCP4.5 scenario. The EA was used in the VGAM to predict the scale parameter β of H_p and D. The first time-derivative of EA was used in the VGAM to predict the scale parameter β of T_p. The scale parameter β of E is not affected by the selected climate indices. The time series are approximated by a straight line, which helps to see the trend of the scale parameter β in each case.

(**a**) (**b**)

Figure 6. The dependence-parameter τ of (**a**) all the storm components and (**b**) the subset *E*-*D*, under the RCP4.5 scenario. Selected nodes serve to represent the results. Note that the scales for τ are different for (a) and (b).

4.2. RCP8.5

PACF-based distances for E and D from different GCMs are below 0.1, except for models GFDL, MIROC, and MPI, which present higher distances in the case of E. In the case of H_p, the maximum PACF-based distance is 0.2, except for GFDL, IPSL, and MIROC GCMs. T_p is the least coincident wave-storm component among the GCMs, with maximum PACF-based distance equal to 0.6 for the IPSL and MIROC branches of the GCMs, whereas the second maximum PACF-based distance is 0.4 for the CMCC branch. Employing the CNRM-CM5 GCM, and assuming non-stationarity, the estimated average number of wave-storms per year range from the 23 storms/year at node 1 to the 30–32 storms/year at nodes 24–34. Additionally, although NAO is the most influential climate index on storminess, the estimated number of storms per year is not significantly affected by time nor climate index. The wave-storm threshold is mostly influenced by the second time-derivative of EA.

The temporal evolution of the location parameters x_0 is shown in Figure 7. The trends of the location parameters x_0 of H_0, T_p, and D are constant. The location parameter x_0 of E, however, does not have a clear trend. The averaged value at nodes 13 and 29 of the mean of E, H_p, T_p, and D are 149.1 m^2/h, 2.9 m, 6.8 s, and 28.4 h. The assumption in Equation (4) is fulfilled, as the shape parameter ξ is negative. The most influential covariates on the GPD location parameter x_0 are EA for E ($\overline{C}_r = -1.00$), the second time-derivative of EA for H_p ($\overline{C}_r = -1.32$) and D ($\overline{C}_r = 1.40$), and NAO for T_p ($\overline{C}_r = 1.00$). The ratio between the wave-storm component and the climate index is approximately 50 in the case of E, 13 in the case of H_p, 1 in the case of T_p, and 140 in the case of D.

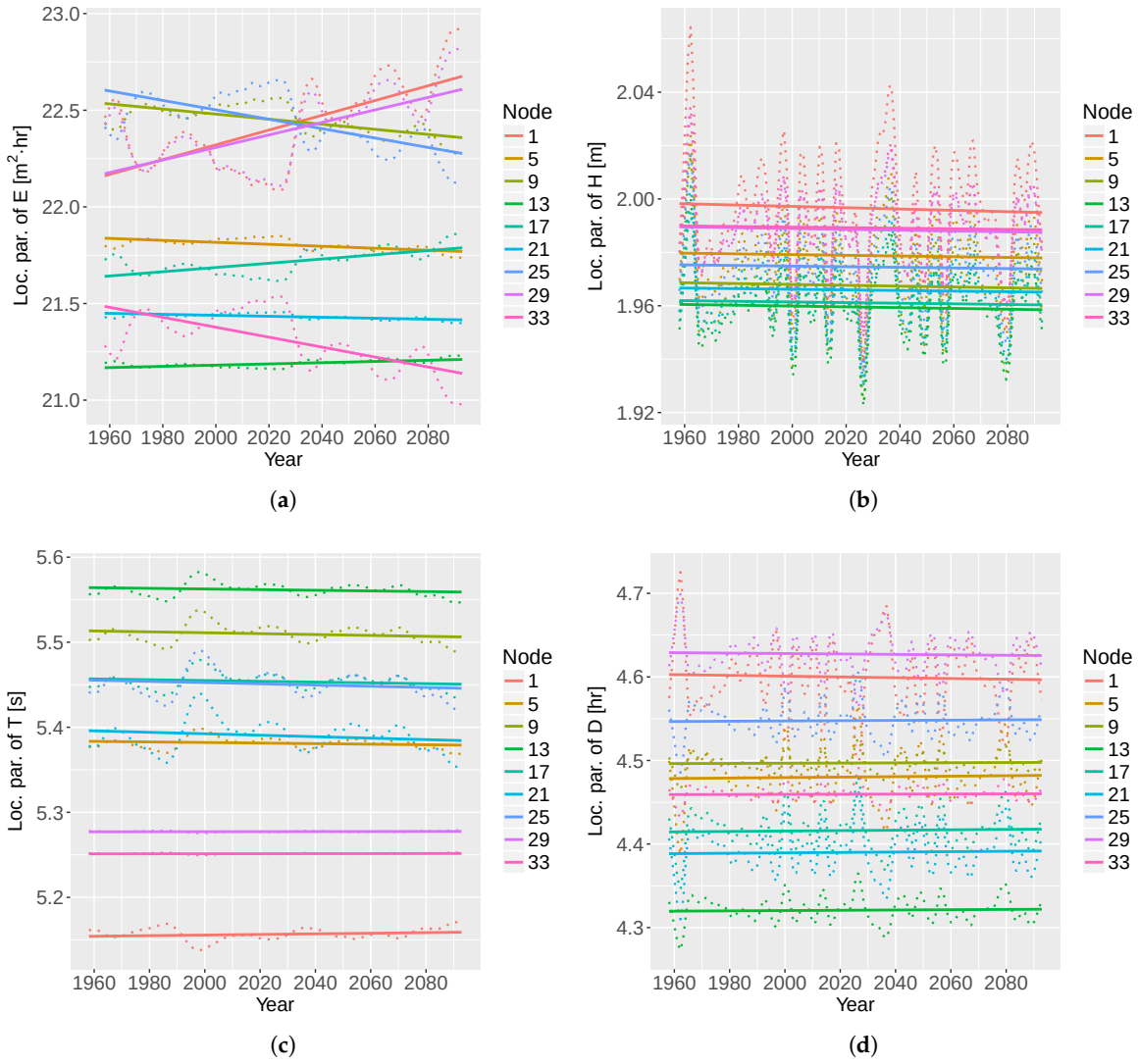


Figure 7. Location parameter x_0 of the GPD that models (**a**) E, (**b**) H_p, (**c**) T_p, and (**d**) D at selected nodes, under the RCP8.5 scenario. The second time-derivative of EA was used in the VGAM to predict the location parameter x_0 of H_p. The NAO was used in the VGAM to predict the location parameter x_0 of T_p. The location parameter x_0 of E and D are not affected by the selected climate indices. The time series are approximated by a straight line, which helps to see the trend of the location parameters x_0 at each case.

The temporal evolution of the scale parameter β is shown in Figure 8, where the scale parameters β show similar trends to their corresponding location parameters x_0. The averaged value at nodes 13 and 29 of the variance of E, H_p, T_p, and D are 6.4 m^2/h, 1.1 m, 1.0 s, and 5.5 h. The assumption in Equation (5) is fulfilled, as the shape parameter ξ is negative. The most influential covariates on the GPD scale parameter β are the second time-derivative of EA for E ($\overline{C}_r = -1.23$), SC for H_p ($\overline{C}_r = -1.03$), EA for T_p ($\overline{C}_r = 1.04$), and the first time-derivative of SC for D ($\overline{C}_r = 1.05$). The ratio between the wave-storm component and the climate index is approximately 800 in the case of E, 1 in the case of H_p and T_p, and 35 in the case of D.

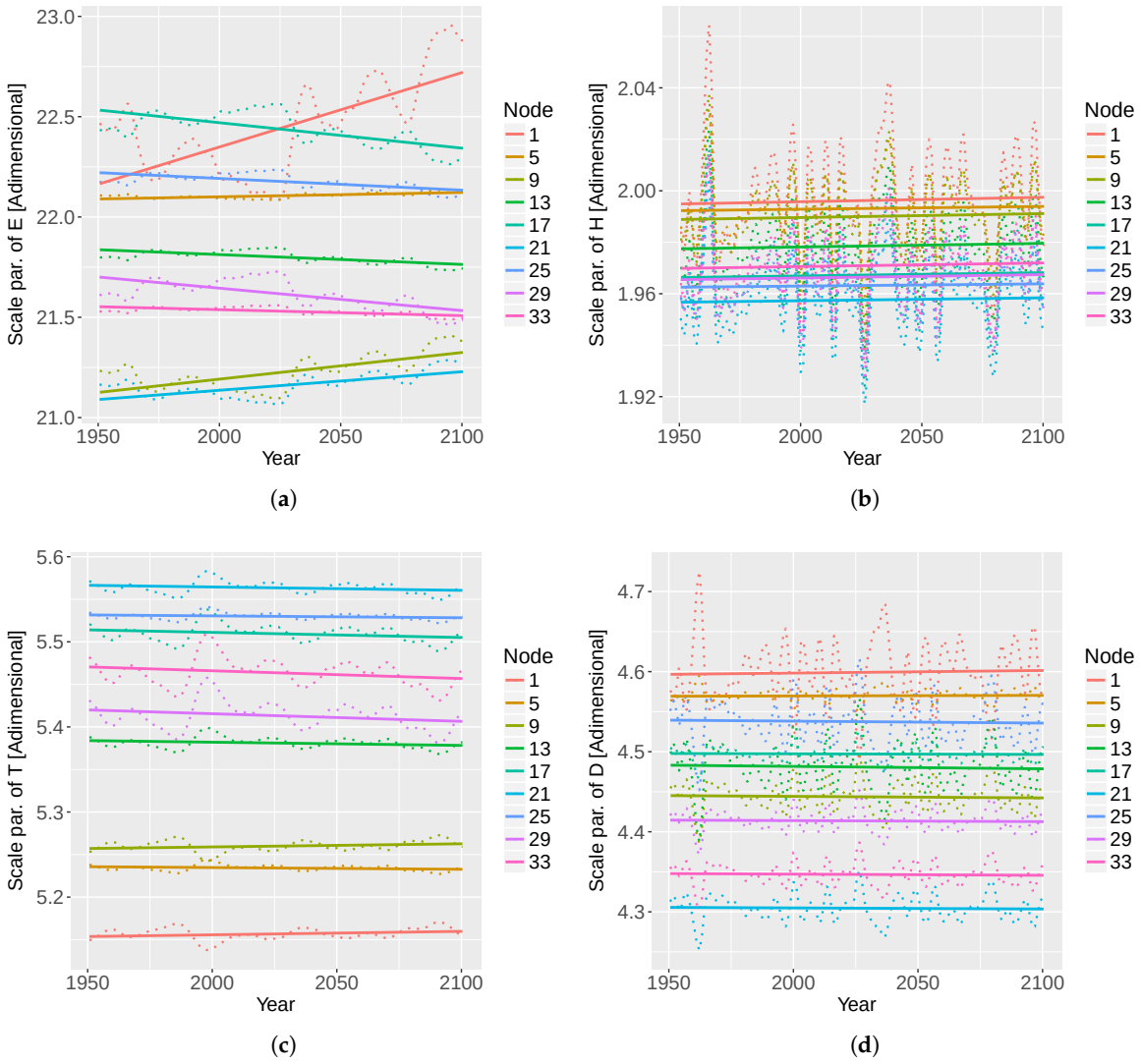


Figure 8. Scale parameter β of the GPD that models (**a**) E, (**b**) H_p, (**c**) T_p, and (**d**) D at selected nodes, under the RCP8.5 scenario. The SC was used in the VGAM to predict the scale parameter β of H_p. The EA was used in the VGAM to predict the scale parameter β T_p. The first time-derivative of SC was used in the VGAM to predict the scale parameter β of D. The scale parameter β of E is not affected by the selected climate indices. The time series are approximated by a straight line, which helps to see the trend of the scale parameter β at each case.

The Gumbel-type HAC is suitable for the proposed statistical model under RCP8.5. It has significantly good graphical fit between empirical and theoretical HAC, while κ^2 is over 0.90 at each level of the HAC. The null hypothesis of stationarity of the θs cannot be rejected in 1.0% of the cases, in each HAC, according to the KPSS test. Therefore, the use of pseudo-non-stationary HACs is suitable. The values of Kendall's τ are shown in Figure 9. τ_{root} ranges between 0.45 and 0.65, and has a positive trend at all nodes. $\tau_{(E,D)}$ ranges between 0.82 and 0.86, and presents cyclical fluctuations. The non-stationary statistical model is validated for 1979–2016, as both the Aitchison and Kullback–Leibler measures (Equations 6 and 7) are below 1.0.

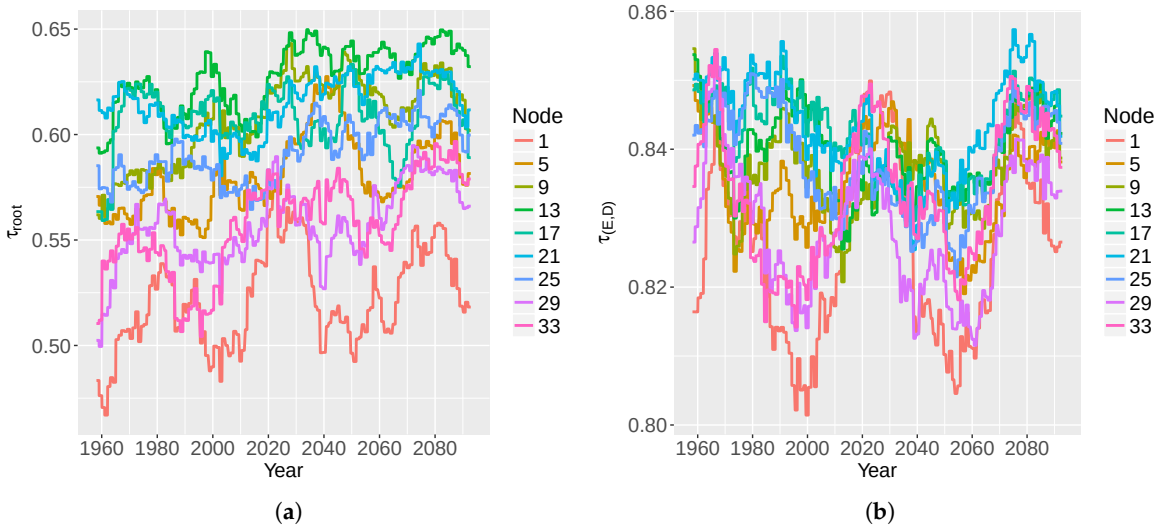


Figure 9. The dependence-parameter τ of (**a**) all the storm components and (**b**) the subset *E*-*D*, under the RCP8.5 scenario. Selected nodes serve to represent the results. Note that the scales for τ are different for (a,b).

5. Discussion

This section is organized as follows. For both emission scenarios, it addresses: (a) the process-based modelling chain; and (b) a comparison of the stationary and the non-stationary approach to the statistical model. For each emission scenario, it addresses: (i) bounding of the uncertainty of the GCM; (ii) value of wave-storminess and its relationship to the selected climate indices; (iii) values of location parameters x_0 of GPDs and their relationship to the selected climate indices; (iv) values of scale parameters β of GPDs and their relationship to the selected climate indices; and (v) values of the parameters in the dependence structure. As a final remark, the applicability of this research is envisaged for both emission scenarios.

Present and future wave conditions under two climate change projections (RCP4.5 and RCP8.5) were modelled with SWAN in the Black Sea, providing possibly the first future extreme wave climate projections in the area. These wave projections are a gapless dataset of long temporal coverage (from 1950 to 2100). Their spatial resolution is enough to represent the main wave generation and propagation phenomena, with an affordable computational cost. Their generation required a computational time of two and a half months, using two four-core workstations. For the control period, the results agree with former works at the Black Sea [4,28,31]. Additionally, the model validation shows that the SWAN outputs agree with the ERA-interim reanalysis for 1979–2016.

Similar agreement between the ALADIN52-RCM and ERA-INTERIM atmospheric fields has been reported in the state-of-the-art [41,88–90]. Additionally, the performance of the GCM-RCM modelling chain (CNRM-CM5 and ALADIN) has recently been compared with other pairs of GCM and RCM, under the same RCP 4.5 and 8.5 scenarios. For example, changes in Mistral and Tramontane wind patterns in the northwestern Mediterranean Sea have been analysed by [89] for the same time interval as this study. Their results noted that there would be fewer Tramontane events under the RCP8.5 scenario than those under the RCP4.5. However, small frequency changes were found for the Mistral at both RCPs. On the other hand, [90] estimated an increase in extreme precipitation event intensity for the Lez, Aude (France), and Muga (Spain) basins. These works reinforce the usefulness of the CNRM-CM5 GCM-RCM dataset for assessing changes in extreme weather.

Apart from the main non-stationary statistical model, the stationary statistical model in the side analysis provides its own GPD estimation. The parameters of the GPD of the stationary statistical model are similar to the parameters of the GPD of the non-stationary statistical one in this study. Therefore,

this assures that the initial values for the parameters of the GPD of the non-stationary statistical model should be maintained. The mean values and the variances from the GPD (see Equations 4 and 5) in both RCP scenarios coincide in order of magnitude with the statistics of the input wave-storm components. Then, the GPDs reflect the characterized wave-storm components correctly. Additionally, in agreement with Equations (4) and (5), parameters of the non-stationary GPDs can represent mean values and variances of wave-storm components. The Gumbel-type HAC structure from the stationary statistical model proves suitable according to the related goodness-of-fit test. Another important advantage is that the non-stationary statistical model characterized the SWAN data without having to cluster the wave-storms into time-windows. A clustering would certainly lead to excessively dependent results in the chosen time-window. Below, the outcomes of the non-stationary statistical model for each emission scenario are discussed in detail (see a summary in Table 2).

Table 2. Summary table of the Discussion section. $d(\cdot)$ and $d^2(\cdot)$ are the first and second time-derivatives, respectively.

Variable or Test	Parameter	RCP4.5 Main Covariate	RCP8.5 Main Covariate
Estimated storminess		27–35 storms/year	23–32 storms/year
Storminess	x_0	None	None
Wave-storm threshold	x_0	dSC + d^2SC	d^2EA
E	x_0 β	EA None	None None
H_p	x_0 β	dNAO EA	d^2EA SC
T_p	x_0 β	SC dEA	NAO EA
D	x_0 β	EA EA	None dSC
τ_{root}		0.55–0.65	0.45–0.65
$\tau_{(E,D)}$		0.82–0.87	0.82–0.86
HAC is non-stationary?		Yes	Yes
Validated for 1979–2016?		Yes	Yes

5.1. RCP4.5

The CNRM-CM5 GCM provides similar E and D to the rest of the GCMs. However, the selection of this model affects the outcomes in H_p and T_p. As the maximum discordance between GCMs corresponds to a PACF-based distance of 2, a distance of 0.3 is still relatively small. However, an explanation for this production of different H_p and T_p might be that some GCMs, like the ones from the MIROC branch (MIROC-ESM, MIROC-ESM-CHEM and MIROC5) and the GFDL branch (GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M), are centred on areas away from the European Continent.

The number of storms in the area should range between 15 and 20 [31]. An estimated average of 27 storms/year might be due to the fact that a non-stationary threshold was significantly lower at some points than the initial values, so more events were qualified as extreme. However, the minimum time-interval between storms chosen (72 h) has helped prevent the consideration of replicating storms as independent ones, thus reducing the modelled storminess to a reasonable quantity. There is a relatively higher storminess at nodes 17–34 (see Figure 1) compared to nodes 1–16, which might be due to more frequent extreme winds in this area. However, storminess is constant under the effects of climate change in the case of RCP4.5. Despite this, the first and the second time-derivatives of SC present special relevance to the storm-thresholds in the northwestern Black Sea. The sensitivity of storm-threshold to SC means that SC does affect the combination of storminess and intensity of

wave-storms. The lack of relationship between storminess and climate patterns is not common in the present [91] or future climates. For instance, storminess is significantly related to negative NAO in the Catalan Coast, another micro-tidal environment in the northwestern Mediterranean Sea [11].

The mean of H_p and T_p remain constant in 1950–2100, whereas E and D rise under the effects of climate change (See Figure 4). D seems sensible to the temperature energetic input into the atmosphere by climate change, as well as being the only contributor to the rise in E in the second half of the 21st century. A higher E, combined with an equally higher D, is often related to greater destructive forces, whether on the natural environment or urban areas. A constant mean T_p is related to a constant fetch, following the stationarity of the wave-storm mean wave-directions. Additionally, a constant mean H_p means that during this century the joint effect of sea-level-rise and H_p would not worsen in the form of, for instance, more flooding. The location parameters x_0 of E, H_p, T_p, and D are significantly influenced by EA, the first time-derivative of NAO, SC, and EA, respectively. According to the regression coefficients $\overline{C}_r$ of the VGAM used to obtain the location parameters x_0, it seems plausible that EA > 0 contributed to higher mean of E and D. The location parameter x_0 of H_p is in turn significantly influenced by a fast increase of negative NAO. The effect of NAO on wave-height under extreme wave-climate could be of concern, considering that this relationship was also observed by [92] in a hindcasted model from the Liverpool Bay. The location parameter x_0 of T_p is significantly influenced by a negative SC.

Figure 5 shows that the variances of E, H_p, T_p, and D increase in the 1950–2100. The figure also reflects how greater mean values are followed by greater variances, which is common in field observation. There seems to be a peak of scale parameter β in the period 2001–2025, which would explain that in the northern Aegean Sea, very near to the Black Sea [93], extreme wave-heights were higher during 2001–2050. The scenario considered in [93] is the A1B scenario of the Assessment Report 4 of the Intergovernmental Panel on Climate Change 2007. It is equivalent to RCP6.0 of the Assessment Report 5 of the Intergovernmental Panel on Climate Change 2013 [94]. The variabilities of H_p, T_p, and D are strongly dependent on EA, the second time-derivative of EA, and EA, respectively. The variance of E is not especially sensible to any climate index, nor their accelerations. However, the variabilities of H_p and D increase with high values of negative and positive EA, respectively. Additionally, the variability of T_p increases with a fast approach towards negative EA.

As for the joint statistical structure of the storm components at the northwestern Black Sea (see Figures 3b and 6), there is a lower dependence between the H_p and the E-D, unlike in the Catalan Coast [11]. The maximum value of τ_{root} is approximately 0.67, so the dependence among H_p, T_p, E, and D are not especially strong and wave-storms with extreme values of all of these components are not probable. This is corroborated by historical wave-storms observations, where it is common to have extreme conditions in wave-height, but a combination of extreme wave-height, wave-period, wave-storm duration, and energy is extremely rare. Nevertheless, the increasing trend in τ_{root} suggests that all four storm components present a greater common semi-dependence with time. $\tau_{(E,D)}$ takes values of approximately 0.85, so there is a strong association of E and D. This should be true, by the definition of E (see Equation 2), but here the relatively lower role of H_p in E is verified. This also implies that a wave-storm of maximum E and D is more feasible in this century. Note that this semi-dependence is lower in the period 2000–2050 than during the rest of the century.

5.2. RCP8.5

The CNRM-CM5 GCM presents higher PACF-based distances in the RCP8.5, compared to the RCP4.5 scenario. The wave-storm components E and D are still the most coincident wave-storm components between GCMs. Additionally, as in the RCP4.5 scenario, the selection of the GCM affects the outcomes of H_p and T_p, mainly the models from the GFDL, IPSL, and MIROC branches, in comparison to the CNRM branch. Therefore, while the PACF-based distances in these GCMs are still relatively small, like in the RCP4.5 scenario, a selection of the models from the GFDL, IPSL, and MIROC branches would project slightly different wave-storms.

Estimated storminess is similar between the RCP8.5 scenario and the RCP4.5 scenario, presenting more estimated storminess at nodes 17–34 (see Figure 1). Again, the estimated average number of storms exceeds the usual 15–20 storms/year. This is possibly for the same reasons as in the RCP4.5 scenario. However, the minimum time-interval between storms of 72 h has prevented replicants of the same storm from mistakenly elevating storminess. Additionally, similar to the RCP4.5 scenario, climate change does not affect storminess, as the latter stays constant during 1950–2100 and is unaffected by climate patterns. EA, the second time-derivative of which is the most influential on the wave-storm threshold, is mostly active between the months of December and February, as the winter average EA brings a low surface pressure anomaly centre over the Caspian region, east of the Black Sea. In the positive phase of EA, associated with the anticyclonic activity over the Caspian Sea, the Black Sea region is exposed to cold and dry air masses from the northeast-to-northwest sector. On the contrary, in its negative phase associated with the anticyclonic activity over the Caspian Sea, the Black Sea region is affected by air masses flowing from the southwest to the southeast [95]. The sensitivity of the storm-threshold to the second time-derivative of EA indicates that a sudden appearance of EA during a winter season foreshadows a worsening of the sea state during that period.

The location parameters x_0 of H_p show a constant trend in 1950–2100 (see Figure 7), but decrease slightly. This phenomenon coincides with [77], which stated that under the RCP8.5 scenario, the maximum wave-height in the Black Sea decreases in the period comprising 2080–2099. The mean value of D has a constant trend in 1950–2100, and it is not higher for the more severe emission scenario RCP8.5. The trend of the mean value of E should be constant, as in the cases of H_p and D. However, the presence of different trends at each node suggests that it depends on the relative dominance of H_p or D at each node. The location parameter x_0 of T_p is constant over the 21st century, so the statistical mean of this storm component might be constant over time, like under RCP4.5. According to the regression parameters $\overline{C}_r$ of the VGAMs used to predict the storm components from the corresponding climate indices, the mean values of E and D are not significantly influenced by any climate index. However, the mean value of H_p is strongly influenced by the second time-derivative of EA. The regression parameter $\overline{C}_r$ of a VGAM that predicts the location parameter x_0 of H_p from EA shows no strong relationship between the location parameter x_0 of H_p and EA, so it can be inferred from the results that the mean H_p can be produced by an acceleration towards negative EA. T_p is strongly influenced by NAO, and it increases with a positive phase of this climate pattern.

Additionally, like in the RCP4.5 scenario, the trend of the scale parameters β for each storm component (see Figure 8) reflects the trends in the location parameter x_0 of the respective storm components (see Figure 7). That is, the degree of variability parallels the mean value of each storm component. A possible explanation is that a greater mean value of each storm component was related to a greater variance of the storm component. The variability of H_p, T_p, and D are significantly influenced by the respective climate indices, while the variability of E is not sensible to any climate index. From the results, it can be deduced that the variability of H_p increases with negative SC, the variability of T_p increases with positive EA, whereas the variability of D increases with a rapid increase of positive SC. At this point, it should be noted that EA has a relatively greater influence on the GPD parameters than the other two selected climate-patterns. All the EA climate-index and its first and second time-derivatives have this effect. Therefore, EA can be crucial in future mean values and variabilities of wave-storm components.

As for the joint probability of occurrence of different storm components, there is also a lower dependence between the H_p and the E-D than in the Catalan Coast [11]. The maximum τ_{root} is 0.65 (see Figure 9), similar to the RCP4.5 scenario (see Figure 6). It denotes similar moderate general dependence among wave-storm components. However, there is again a positive trend for this semi-dependence among the four storm components. The mean $\tau_{(E,D)}$ is 0.83, similar to the RCP4.5 scenario, meaning that wave-storms of maximum E and D are possible in this century. This high dependence between E-D is also present in the also micro-tidal environment of the Catalan Coast, under the RCP8.5 scenario [11]. Like in the RCP4.5 scenario, there is also a period of lower semi-dependence

among E and D, this time from 1980 to 2015 and from 2040 to 2070. It should be noted that τ are generally 2 centesimals lower in the RCP8.5 scenario than in the RCP4.5 scenario. This suggests a lower dependence among wave-storm components in higher emission scenarios, possibly due to the presence of extremes of single wave-storm components during wave-storms.

5.3. Applicability of the Results

Water infrastructure tends to be designed with pre-defined loads that aims to fulfill a particular need (i.e., water storage and distribution, ocean wave sheltering, storm flooding, etc.) Climate change adds a further layer of complexity: (i) the stationary assumption of the loads cannot be assumed; and (ii) there is a relevant uncertainty in how extreme climate forcings will change [96]. An option for bounding this uncertainty is to analyse GCMs' climate projections derived from IPCC climate scenarios [25], such as RCP 8.5 or 4.5. Dynamical downscaling (the first step in our methodology) joint with non-stationary multivariate statistics (second step) provides an affordable solution, without requiring prohibitive computational cost. Once the statistical model is fitted with process-based model outputs and after establishing relationships between the loads and related climate indices [21,97,98], uncertainty can be bounded by comparing climate indices from several publicly available GCMs (fourth step).

These statistical models can deal with level 3 (probabilistic) designs, considering random simulation of the loads, while presenting a pre-defined dependence structure (for example, via Archimedean copulas). Recent methodologies for coastal infrastructure design consider the joint action of wave-height, wave-period, and storm-duration [16,99]. Beach management could also benefit from the additional robustness, due to consistency in climatic co-factors. For instance, stakeholders could thus pro-actively protect a beach for a cruder winter season which has much more destructive waves or more unpredictable wave-storms [29,100] than other projected milder winters. More lives could be saved if preparedness campaigns were made for people to avoid navigation under high probability of especially hazardous wave-storms. These improvements can be extended to other fields, such as river hydrology or the preservation of coastal biosphere. In any of these fields, predictability is strongly linked to variability and joint probability-structure of extreme components. Hence, stakeholders can benefit from these hybrid methodologies for disaster risk management.

6. Conclusions

Present and future wave conditions under two climate change projections (RCP4.5 and RCP8.5) have been modelled with SWAN in the Black Sea. These projections have been characterized with a multivariate non-stationary statistical model that deals with the storm components and the semi-dependence among them. The individual probability of the storm components are characterized by non-stationary GPD functions. The parameters of such GPDs have been estimated through VGAM, where climate indices derived from the CNRM-CM5 general circulation model (NAO, EA, and SC) have been employed as covariates. The joint probability structure is characterized by a pseudo-non-stationary HAC.

The proposed hybrid methodology shows the following features. Similar trends are found when using climate indices derived from other GCMs as predictors. Hence, loads due to future extreme climate have been characterized in the statistical model. Estimated storminess is affected by neither time nor climate indices, in any RCP scenario. The wave-storm threshold, however, is strongly influenced by the first and second time-derivatives of SC in the RCP4.5 scenario, and by the second time-derivative of EA in the RCP8.5 scenario. Under both RCP scenarios, the mean values of H_p and T_p remain fairly constant over the 21st century. The mean value of E is more markedly increasing in the RCP4.5 scenario than in the RCP8.5 scenario. The mean value of D is increasing in the RCP4.5 scenario, as opposed to the constant trend in the RCP8.5 scenario. The variability of each storm component increases with increasing mean values, under both RCP scenarios. All three climate indices

can influence the mean values and variances of the wave-storm components, but EA and its dynamics play a special role.

Under both scenarios, wave-storms of maximum E, H_p, T_p, and D are not excessively common, but wave-storms of concurrent extreme E and D can be expected. A positive trend in the semi-dependence among all wave-storm components can be observed in both scenarios. There are also periods of lower semi-dependence among E and D; e.g., 2000–2050 in the RCP4.5 scenario, and 1980–2015 and 2040–2070 in the RCP8.5 scenario. Additionally, the more severe emission scenario of RCP8.5 presents lower dependence among wave-storm components. The joint knowledge of the effects of climate indices, the trend of each storm component in this century, and the probability of joint occurrence can help to improve performance in the design of coastal infrastructure and the management of natural and urban resources.

Acknowledgments: This paper has been supported by the European project CEASELESS (H2020-730030-CEASELESS), the Spanish national project PLAN-WAVE (CTM2013-45141-R), and MINECO FEDER Funds co-funding CODA-RETOS (MTM2015-65016-C2-2-R). As a group, we would like to thank the Secretary of Universities and Research of the department of Economics of the Catalan Generalitat (Ref. 2014SGR1253, 2014SGR551). The second author acknowledges the Ph.D. scholarship from the Generalitat de Catalunya (DGR FI-AGAUR-14). The authors duly acknowledge the wind field projections from the ALADIN model, used in this work, that were part of the Med-CORDEX initiative (www.medcordex.eu) supported by the HyMEX programme (www.hymex.org). ECMWF ERA-INTERIM data used in this study have been obtained from the ECMWF data server. Information and expertise from the Black Sea area have been provided via the Romanian National Core Programme PN 16450103 and the RI EUXINUS. Special thanks are acknowledged to the academic editor and to three anonymous peer-reviewers whose genuinely constructive comments have enriched this paper.

Author Contributions: Jue Lin-Ye analysed the model outputs and led the paper. Together with Manuel García-León, they conceived the proposed hybrid framework. M. Isabel Ortego and Agustín Sánchez-Arcilla contributed in the multivariate non-stationary statistical model. Vicente Gràcia, Adrian Stanica and Agustín Sánchez-Arcilla provided expertise on Oceanography and contributed in the process-based modelling. Monthly proof-readings have been held by all co-authors via presential and videoconference meetings.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

D	total wave-storm duration
E	total wave-storm energy
EA	East Atlantic Pattern
GCM	general (atmospheric) circulation model
GPD	generalized Pareto distribution
HAC	hierarchical Archimedean copula
H_p	significant wave-height at the peak of the wave-storm
NAO	North Atlantic Oscillation
PACF	partial autocorrelation function
RCM	regional (atmospheric) circulation model
SC	Scandinavian Pattern
SWAN	Simulating WAves Nearshore (spectral wave-model)
T_p	peak wave-period at the peak of the wave-storm
VGAM	vectorial generalized additive model

References

1. Rivera, J.A.; Penalba, O.C.; Villalba, R.; Araneo, D.C. Spatio-temporal patterns of the 2010–2015 extreme hydrological drought across the Central Andes, Argentina. *Water* **2017**, *9*, 652.
2. Nguyen, H.Q.; Radhakrishnan, M.; Huynh, T.T.N.; Baino-Salingay, M.L.; Ho, L.P.; Van der Steen, P.; Pathirana, A. Water Quality Dynamics of Urban Water Bodies during Flooding in Can Tho City, Vietnam. *Water* **2017**, *9*, 260.

3. Thompson, D.A.; Karunarathna, H.; Reeve, D.E. Modelling extreme wave overtopping at Aberystwyth Promenade. *Water* **2017**, *9*, 663.
4. Valchev, N.; Davidan, I.; Belberov, Z.; Palazov, A.; Valcheva, N.; Chin, D. Hindcasting and assessment of the western Black Sea wind and wave climate. *J. of Environ. Prot. Ecol.* **2010**, *11*, 1001–1012.
5. Zacharioudaki, A.; Pan, S.Q.; Simmonds, D.; Magar, V.; Reeve, D.E. Future wave climate over the west-European shelf seas. *Ocean Dyn.* **2011**, *61*, 807–827.
6. Sierra, J.P.; García-León, M.; Gràcia, V.; Sánchez-Arcilla, A. Green measures for Mediterranean harbours under a changing climate. *Proc. inst. civ. Eng.-marit. Eng.* **2017**, *170*, 55–66.
7. Guo, L.L.; Sheng, J.Y. Statistical estimation of extreme ocean waves over the eastern Canadian shelf from 30-year numerical wave simulation. *Ocean Dyn.* **2015**, *65*, 1489–1507.
8. Vledder, G.; Akpinar, A. Wave model predictions in the Black Sea: Sensitivity to wind fields. *Appl. Ocean Res.* **2015**, *53*, 161–178.
9. Sánchez-Arcilla, A.; García, M.; Gràcia, V. Hydro-morphodynamic modelling in Mediterranean storms - errors and uncertainties under sharp gradients. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 2993–3004.
10. Camus, P.; Losada, I.J.; Izaguirre, C.; Espejo, A.; Menéndez, M.; Pérez, J. Statistical wave climate projections for coastal impact assessments. *Earth's Future* **2017**, *5*, 918–933.
11. Lin-Ye, J.; García-León, M.; Gràcia, V.; Ortego, M.I.; Lionello, P.; Sánchez-Arcilla, A. Multivariate statistical modelling of future marine storms. *Appl. Ocean Res.* **2017**, *65*, 192–205.
12. Kumar, P.; Min, S.K.; Weller, E.; Lee, H.S.; Wang, X.L. Influence of climate variability on extreme ocean surface wave heights assessed from ERA-Interim and ERA-20C. *Am. Meteorol. Soc.* **2016**, *29*, 4031–4046.
13. Camus, P.; Rueda, A.; Méndez, F.J.; Losada, I.J. An atmospheric-to-marine synoptic classification for statistical downscaling marine climate. *Ocean Dyn.* **2016**, *66*, 1589–1601.
14. Lin-Ye, J.; García-León, M.; Gràcia, V.; Sánchez-Arcilla, A. A multivariate statistical model of extreme events: An application to the Catalan coast. *Coast. Eng.* **2016**, *117*, 138 – 156.
15. Wahl, T.; Jensen, J.; Mudersbach, C. A multivariate statistical model for advanced storm surge analyses in the North Sea. *Coast. Eng. Proc.* **2011**, *1*, 19.
16. Salvadori, G.; Tomasicchi, G.; d'Alessandro, F. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. *Coast. Eng.* **2014**, *88*, 1–14.
17. Wahl, T.; Mudersbach, C.; Jensen, J. Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas: a multivariate statistical approach based on Copula functions. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 495–510.
18. Wang, X.L.; Feng, Y.; Swail, V.R. Climate change signal and uncertainty in CMIP5-based projections of global ocean surface wave heights. *J. Geophys. Res. Oceans* **2015**, *120*, 3859–3871.
19. Hemer, M.A.; Trenham, C.E. Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble. *Ocean Model.* **2016**, *103*, 190 – 203.
20. Vanem, E. Long-term time-dependent stochastic modelling of extreme waves. *Stoc. Environ. Res. Risk Assess.* **2011**, *25*, 185–209.
21. Yee, T.W.; Stephenson, A.G. Vector generalized linear and additive extreme value models. *Extremes* **2007**, *10*, 1–19.
22. Du, T.; Xiong, L.H.; Xu, C.Y.; Gippel, C.J.; Guo, S.; Liu, P. Return period and risk analysis of nonstationary low-flow series under climate change. *J. Hydrol.* **2015**, *527*, 234 – 250.
23. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2005**, *54*, 507–554.
24. Karim, F.; Hasan, M.; Marvanek, S. Evaluating annual maximum and partial duration series for estimating frequency of small magnitude floods. *Water* **2017**, *9*, 481.
25. Stocker, T.F.; Qin, D.; Plattner, G.K.; tignor, M.; Allen, S.K.; Boschung, J.; Nauels, A.; Xia, Y.; Bex, V.; Midgley, P.M. IPCC, 2013: Summary for Policymakers. Climate Change 2013: The physical science basis. In *Contribution of working group I to the Fifth Assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.
26. Liu, J.; Luo, M.; Liu, T.; Bao, A.M.; De Maeyer, P.; Feng, X.W.; Chen, X. Local Climate Change and the impacts on hydrological processes in an arid Alpine catchment in Karakoram. *Water* **2017**, *9*, 344.
27. Panin, N. Impact of global changes on geoenvironmental and coastal zone state of the Black Sea. *Geo-Eco-Marina* **1996**, pp. 7–23.

28. Arkhipkin, V.S.; Gippius, F.N.; Koltermann, K.P.; Surkova, G.V. Wind waves on the Black Sea: results of a hindcast study. *Nat. Hazards Earth Syst. Sci. Discussions* **2014**, *2*, 1193–1221.
29. Sánchez-Arcilla, A.; García-León, M.; Gràcia, V.; Devoy, R.; Stanica, A.; Gault, J. Managing coastal environments under climate change: Pathways to adaptation. *Sci. Total Environ.* **2016**, *572*, 1336–1352.
30. Halcrow team. *Masterplan for the Protection against erosion of the Romanian Black Sea coast*. Technical report, Halcrow: London, UK, 2011.
31. Rusu, E. Wave energy assessments in the Black Sea. *J. Mar. Sci. Technol.* **2009**, *14*, 359–372.
32. Voldoire, A.; Sanchez-Gomez, E.; Salas y Mélia, D.; Decharme, B.; Cassou, C.; Sénési, S.; Valcke, S.; Beau, I.; Alias, A.; others. The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim. Dyn.* **2013**, *40*, 2091–2121.
33. Kwak, J.; St-Hilaire, A.; Chebana, F.; Kim, G. Summer season water temperature modeling under the Climate Change: Case study for Fourchue River, Quebec, Canada. *Water* **2017**, *9*, 346.
34. Luo, M.; Meng, F.H.; Liu, T.; Duan, Y.C.; Frankl, A.; Kurban, A.; De Maeyer, P. Multi–model ensemble approaches to assessment of effects of local Climate Change on water resources of the Hotan River Basin in Xinjiang, China. *Water* **2017**, *9*, 584.
35. Farda, A.; Déué, M.; Somot, S.; Horányi, A.; Spiridonov, V.; Tóth, H. Model ALADIN as regional climate model for Central and Eastern Europe. *Stud. Geophys. Geod.* **2010**, *54*, 313–332.
36. Colin, J.; Déqué, M.; Radu, R.; Somot, S. Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. *Tellus A* **2010**, *62*, 591–604.
37. Herrmann, M.; Somot, S.; Calmanti, S.; Dubois, C.; Sevault, F. Representation of spatial and temporal variability of daily wind speed and of intense wind events over the Mediterranean Sea using dynamical downscaling: impact of the regional climate model configuration. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 1983–2001.
38. Bougeault, P. A simple parameterization of the large-scale effects of cumulus convection. *Mon. Weather Rev.* **1985**, *113*, 2108–2121.
39. Ricard, J.L.; Royer, J.F. A statistical cloud scheme for use in an AGCM. *Annu. Geophys.* **1993**, *11*, 1095–1115.
40. Smith, R.N.B. A scheme for predicting layer clouds and their water content in a general circulation model. *Q. J.R. Meteorol. Soc.* **1990**, *116*, 435–460.
41. Ruti, P.M.; Somot, S.; Giorgi, F.; Dubois, C.; Flaounas, E.; Obermann, A.; Dell'Aquila, A.; Pisacane, G.; Harzallah, A.; Lombardi, E.; others. Med-CORDEX Initiative for Mediterranean Climate Studies. *Bull. Am. Meteorol. Soc.* **2016**, *97*, 1187–1208.
42. Mori, N.; Yasuda, T.; Mase, H.; Tom, T.; Oku, Y. Projection of extreme wave climate change under global warming. *Hydrol. Res. Lett.* **2010**, *4*, 15–19.
43. Hemer, M.A.; Fan, Y.L.; Mori, N.; Semedo, A.; Wang, X.L. Projected changes in wave climate from a multi-model ensemble. *Nat. Clim. Change* **2013**, *3*, 471–476.
44. Casas-Prat, M.; Sierra, J.P. Projected future wave climate in the NW Mediterranean Sea. *J. Geophys. Res. Oceans* **2013**, *118*, 3548–3568.
45. Booij, N.; Ris, R.C.; Holthuijsen, L.H. A third-generation wave model for coastal regions: 1. Model description and validation. *J. Geophys. Res. Oceans* **1999**, *104*, 7649–7666.
46. Goda, Y. *Random Seas and Design of Maritime Structures*, 3rd ed.; Advanced Series on Ocean Engineering; World Scientific: Singapore, 2010; Volume 33.
47. Yee, T.W.; Wild, C.J. Vector generalized additive models. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58*, 481–493.
48. Lin, Y.P.; Lin, W.C.; Wu, W.Y. Uncertainty in various habitat suitability models and its impact on habitat suitability estimates for fish. *Water* **2015**, *7*, 4088–4107.
49. Fessler, J.A. Nonparametric fixed-interval smoothing with vector splines. *IEEE Trans. Signal Process.* **1991**, *39*, 852–859.
50. Wei, W.W.S. *Time series analysis*; Addison-Wesley: Boston, MA, USA, 1994.
51. Barnston, A.G.; Livezey, R.E. Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns. *Mon. Weather Rev.* **1987**, *115*, 1083–1126.
52. Butterworth, S. On the theory of filter amplifiers. *Wirel. Eng.* **1930**, *7*, 536–541.
53. Akaike, H. Factor analysis and AIC. *Psychometrika* **1987**, *52*, 317–332.

54. Tamura, Y.; Sato, T.; Ooe, M.; Ishiguro, M. A procedure for tidal analysis with a Bayesian information criterion. *Geophys. J. Int.* **1991**, *104*, 507–516.
55. Egozcue, J.J.; Pawlowsky-Glahn, V.; Ortego, M.I.; Tolosana-Delgado, R. The effect of scale in daily precipitation hazard assessment. *Nat. Hazards Earth Syst. Sci.* **2006**, *6*, 459–470.
56. Tolosana-Delgado, R.; Ortego, M.I.; Egozcue, J.J.; Sánchez-Arcilla, A. Climate change in a Point-over-threshold model: an example on ocean-wave-storm hazard in NE Spain. *Adv. Geosci.* **2010**, *26*, 113–117.
57. Coles, S. *An introduction to Statistical modeling of extreme values*; Springer: Berlin, Germany, 2001; pp. 801,804.
58. Koenker, R. *Quantile regression*; Econometric Society Monographs, Cambridge University Press: Cambridge, UK, 2005.
59. Muraleedharan, G.; Lucas, C.; Guedes Soares, C. Regression quantile models for estimating trends in extreme significant wave heights. *Ocean Eng.* **2016**, *118*, 204 – 215.
60. Northrop, P.J.; Jonathan, P. Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics* **2011**, *22*, 799–809.
61. Jonathan, P.; Ewans, K.; Randell, D. Joint modelling of extreme ocean environments incorporating covariate effects. *Coast. Eng.* **2013**, *79*, 22 – 31.
62. Okhrin, O.; Okhrin, Y.; Schmid, W. On the structure and estimation of hierarchical Archimedean copulas. *J. Econom.* **2013**, *173*, 189–204.
63. Sklar, A. *Fonctions dé repartition à n dimension et leurs marges*; Université Paris 8: Saint-Denis, France, 1959.
64. Nelsen, R.B. *An introduction to copulas*; Springer Science & Business Media: Berlin, Germany, 2007.
65. Bezak, N.; Rusjan, S.; Fijavž, M.K.; Mikoš, M.; Šraj, M. Estimation of Suspended Sediment Loads Using Copula Functions. *Water* **2017**, *9*, 628.
66. Wang, Y.; Li, C.Z.; Liu, J.; Yu, F.L.; Qiu, Q.T.; Tian, J.Y.; Zhang, M.J. Multivariate analysis of joint probability of different rainfall frequencies based on copulas. *Water* **2017**, *9*, 198.
67. Kendall, M.G. A new measure of rank correlation. *Biometrika* **1937**, *6*, 83–93.
68. Salvadori, G.; De Michele, C.; Durante, F. On the return period and design in a multivariate framework. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3293–3305.
69. Eastoe, E.; Koukoulas, S.; Jonathan, P. Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. *Ocean Eng.* **2013**, *62*, 68–77.
70. Kereszturi, M.; Tawn, J.; Jonathan, P. Assessing extremal dependence of North Sea storm severity. *Ocean Eng.* **2016**, *118*, 242 – 259.
71. Okhrin, O.; Ristig, A. Hierarchical Archimedean copulae: the HAC package. *J. Stat. Softw.* **2014**, *58*, doi:http://dx.doi.org/10.18637/jss.v058.i04.
72. Gan, F.F.; Koehler, K.J.; Thompson, J.C. Probability Plots and Distribution Curves for Assessing the Fit of Probability Models. *Am. Stat.* **1991**, *45*, 14–21.
73. Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econom.* **1992**, *54*, 159 – 178.
74. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; others. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J.R. Meteorol. Soc.* **2011**, *137*, 553–597.
75. ECMWF., Part VII: ECMWF Wave Model. In *IFS Documentation CY31R1*; IFS Documentation; Operational implementation 12 September 2006; ECMWF: Reading, UK, 2007.
76. Sadio, M.; Anthony, E.J.; Diaw, A.T.; Dussouillez, P.; Fleury, J.T.; Kane, A.; Almar, R.; Kestenare, E. Shoreline changes on the wave-Influenced Senegal River Delta, West Africa: The roles of natural processes and human Interventions. *Water* **2017**, *9*, 357.
77. Wang, X.L.; Feng, Y.; Swail, V.R. Changes in global ocean wave heights as projected using multimodel CMIP5 simulations. *Geophys. Res. Lett.* **2014**.
78. Egozcue, J.J.; Pawlowsky-Glahn, V. Evidence information in Bayesian updating. Proceedings of the 4th International Workshop on Compositional Data Analysis. Sant Feliu de Guixols, Spain, 9–13 May 2011; pp. 1–13.
79. Aitchison, J. The statistical analysis of compositional data. *J. R. Stat. Soc. Ser. B Methodol.* **1982**, *44*, 139–177.
80. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric logratio transformations for compositional data analysis. *Math. Geol.* **2003**, *35*, 279–300.

81. Aitchison, J. On criteria for measures of compositional difference. *Math. Geol.* **1992**, *24*, 365–379.
82. Pawlowsky-Glahn, V.; Egozcue, J.J. Geometric approach to statistical analysis on the simplex. *Stoc. Environ. Res. Risk Assess.* **2001**, *15*, 384–398.
83. Kullback, S. *Information theory and statistics*; Courier corporation: North Chelmsford, MA, USA, 1997.
84. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498.
85. Hirota, N.; Takayabu, Y.N.; Watanabe, M.; Kimoto, M. Precipitation reproducibility over tropical oceans and its relationship to the double ITCZ problem in CMIP3 and MIROC5 climate models. *J. Clim.* **2011**, *24*, 4859–4873.
86. Ahlström, A.; Schurgers, G.; Arneth, A.; Smith, B. Robustness and uncertainty in terrestrial ecosystem carbon response to CMIP5 climate change projections. *Environ. Res. Lett.* **2012**, *7*, 044008.
87. Montero, P.; Vilar, J.A. Tsclust: An R package for time series clustering. *J. Stat. Softw.* **2014**, *62*, doi:http://dx.doi.org/10.18637/jss.v062.i01.
88. Panthou, G.; Vrac, M.; Drobinski, P.; Bastin, S.; Li, L. Impact of model resolution and Mediterranean sea coupling on hydrometeorological extremes in RCMs in the frame of HyMeX and MED-CORDEX. *Clim. Dyn.* **2016**, doi:https://doi.org/10.1007/s00382-016-3374-2.
89. Obermann-Hellhund, A.; Conte, D.; Somot, S.; Torma, C.Z.; Ahrens, B. Mistral and Tramontane wind systems in climate simulations from 1950 to 2100. *Clim. Dyn.* **2018**, *50*, 693–703.
90. Colmet-Daage, A.; Sanchez-Gomez, E.; Ricci, S.; Llovel, C.; Borrell Estupina, V.; Quintana-Seguí, P.; Llasat, M.C.; Servat, E. Evaluation of uncertainties in mean and extreme precipitation under climate change for northwestern Mediterranean watersheds from high-resolution Med and Euro-CORDEX ensembles. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 673–687.
91. Zăinescu, F.I.; Tătui, F.; Valchev, N.N.; Vespremeanu-Stroe, A. Storm climate on the Danube delta coast: evidence of recent storminess change and links with large-scale teleconnection patterns. *Nat. Hazards* **2017**, *87*, 599–621.
92. Wolf, J.; Brown, J.M.; Howarth, M.J. The wave climate of Liverpool Bay-observations and modelling. *Ocean Dyn.* **2011**, *61*, 639–655.
93. Galiatsatou, P.; Prinos, P. Modeling non-stationary extreme waves using a point process approach and wavelets. *Stoc. Environ. Res. Risk Assess.* **2011**, *25*, 165–183.
94. Kim, D.S.; Wang, Y.D. Economic analysis of rural green-village planning with solar energy considering Climate Change. *J. Korean Soc.Rural Plan.* **2013**, *19*, 25–36.
95. Oguz, T.; Dippner, J.W.; Kaymaz, Z. Climatic regulation of the Black Sea hydro-meteorological and ecological properties at interannual-to-decadal time scales. *J. Mar. Syst.* **2006**, *60*, 235–254.
96. Trenberth, K.E.; Fasullo, J.T.; Shepherd, T.G. Attribution of climate extreme events. *Nat. Clim. Change* **2015**, *5*, 725–730.
97. Zhang, Q.; Gu, X.H.; Singh, V.P.; Xiao, M.Z.; Chen, X.H. Evaluation of flood frequency under non-stationarity resulting from climate indices and reservoir indices in the East River basin, China. *J. Hydrol.* **2015**, *527*, 565 – 575.
98. Gao, M.; Mo, D.Y.; Wu, X.Q. Nonstationary modeling of extreme precipitation in China. *Atmos. Res.* **2016**, *182*, 1 – 9.
99. Martín-Hidalgo, M.; Martín-Soldevilla, M.J.; Negro, V.; Aberturas, P.; López-Gutiérrez, J.S. Storm evolution characterization for analysing stone armour damage progression. *Coast. Eng.* **2014**, *85*, 1 – 11.
100. Salvadori, G.; Durante, F.; Tomasicchio, G.R.; D'Alessandro, F. Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering. *Coast. Eng.* **2015**, *95*, 77 – 83.

8. General results and discussion

This chapter first exposes and discusses (sec. 8.1) the main results from the two study areas in this compendium.

The stationary characterization of wave-storm variables in the Catalan Coast (sec. 8.1.1.1) adresses: i) the wave storminess, ii) the storm shape, iii) the growth-decay rates, iv) the joint probability structure and the values of its parameters and v) the wave-directionality at the peak of the wave-storm, θ_p^*.

The non-stationary characterization of wave-storm intensity variables in the Catalan Coast (sec. 8.1.1.2) addresses: i) the bounding of the uncertainty coming from the GCM, ii) the wave-storm threshold and its relationship to climate-indices, iii) the wave-storminess and its relationship to climate-indices, iv) the wave-storm intensity variables and their relationship to climate-indices and v) the joint probability structure and the values of its parameters.

The issues addressed in the non-stationary characterization of wave-storm intensity variables in the northwestern Black Sea (sec. 8.1.2) is the same as in sec. 8.1.1.2, only that the θ_p^* is discussed before the item iv). Hereafter, sec. 8.2 addresses: a) relationship between the papers, b) limitations of the models, c) main similarities and differences, d) improvement that this new methodology provides, e) specific improvements and new knowledge derived from this dissertation and f) applicability of the results of the thesis.

8.1. The two study areas

8.1.1. The Catalan Coast

8.1.1.1. The stationary approach

The stationary characterization of wave-storms in the Catalan Coast deals with a fetch-limited, micro-tidal environment, in the present climate (years 1996-2013). The statistical model employed to do this characterization is divided into three sub-models: wave-storm intensity, wave directionality and intra-time distribution. In the wave-storm intensity sub-model, the wave-storm intensity variables E, E_u, T_p and D are characterized through GPDs. Their joint probability structure is fit by a stationary HAC. In the wave directionality sub-model, the θ_p^* is fitted to a movMF, and then related to E, T_p and D through a multinomial logistic model. In the intra-time distribution sub-model, D is related to the growth-decay rates of the

wave-storm through a third-degree polynomial. The whole model is validated with Puertos del Estado buoy records from the same period of time.

The wave storminess in the Catalan Coast, in the present wave-climate, is of around $1 - 15$storms/year, which is what can be expected in this study area. The most appropriate geometric figure to describe the evolution of the wave-height is an irregular-trapezoid. This shape is considerably versatile in modelling wave-storms, since it includes the triangular types of wave-storms. For D over 100h, the peak of the wave-storm is located farther from the start of the wave-storm than from the end. Hence, it can be deduced that wave-storms grow slowly and fall sharply after the height of the wave-storm.

Two types of HAC structures are considered in the stationary characterization of wave-storms in the Catalan Coast. The configuration of such HAC structures can be seen in Fig. 8.1. Type A HAC-structures shows more dependence of E_u to E and D. In the present and future wave-climate of the Catalan Coast and the future wave-climate of northwestern Black Sea, wave-storms not always present maximum values of all the wave-storm variables. However, the probability of extreme wave-storm energy and duration can often be considerably high. The principal θ_p^* on the northern Catalan coast are North and East, whereas eastern and southern directions are predominant in central and southern Catalan coast. The intensity of eastern waves is corroborated by Bolaños et al. (2009). The θ_p^* depends on the directionality of the waves that are energetic enough to be part of a wave-storm. In the northern Catalan Coast, Tramuntana winds are prone to create large waves that end up forming part of a wave-storm. Also, in the central and southern Catalan Coast, wave-storms are mainly originated in the east, due to larger fetch from this direction.

8.1.1.2. The non-stationary approach

The non-stationary characterization of wave-storms in the Catalan Coast uses a WAM wave projection for the years 1950-2100, considering a RCP8.5 greenhouse gas emission scenario for the projection. The target wave-storm intensity variables are E, H_p, T_p and D. The parameters of the non-stationary GPDs, used to characterize these variables, are estimated through VGAM. Their joint dependence structure is characterized through a pseudo-non-stationary HAC. The non-stationary statistical model establishes the relationship between the wave-storm threshold, the wave-storminess, the wave-storm intensity variables and the climate-indices NAO, EA and SC, as well as the first two time-derivatives of these climate-indices. This relationship is newly introduced in the non-stationary models. The statistical model is validated with Puertos del Estado buoy records and SIMAR hincast in the period 1996-2013.

In the Catalan Coast, different GCMs have been compared to the CMCC-CM GCM. This can help bound the uncertanty coming from GCMs. In this study area, all GCMs provide similar wave-storm conditions. Thus, this uncertainty is small. The Catalan Coast might be related to little uncertainty coming from the GCMs because

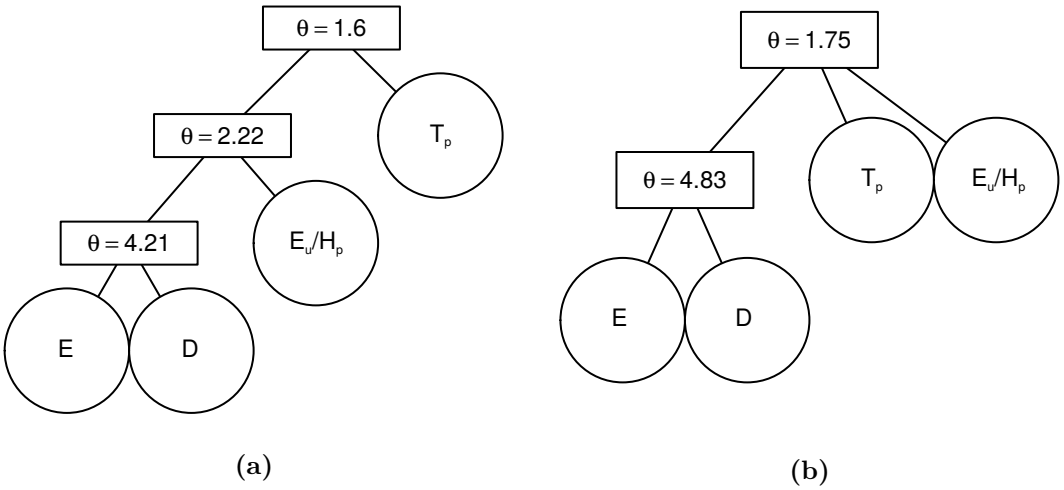


Figure 8.1.: Examples of HAC structures that appear in the three papers. The variable E_u is substituted by H_p in the non-stationary models. Both HAC-structure types are present in the stationary model. Type A HAC-structure (a) is representative of the non-stationary HAC-structure in the Catalan Coast. type B HAC-structure (b) is representative of the non-stationary HAC-structure in the northwestern Black Sea.

of the following reasons. First, the Catalan coast is located in Europe, thus it might be modelled the same way in all the European centred GCMs. Second, it is at the same distance from the United States than from central Europe, thus it might be modelled by the GFDL in a similar way than the European centred GCMs. The MIROC branch of the GCMs does not present significantly more uncertainty. What is more, the elements in this branch are relatively few in number, compared to all the GCM that present similar wave-climate results to the CMCC-CM GCM. Therefore, it can be said that the uncertainty coming from the CMCC-CM GCM was considerably limited.

In 21^{st} century wave-climate of the Catalan Coast, the wave-storm threshold is mostly influenced by NAO, the second time-derivative of EA and SC. The wave storminess is $5-25$storms/year. It is boosted by negative NAO, which might create temperature changes in a favourable way to originate stronger winds. This, in turn, can generate energetic waves that finally increased the wave storminess in the area. However, it is not certan to what extend this was related to natural variablity of atmospheric climate-patterns and how it was affected by the antrhopogenic footprint (Trenberth et al. (2015)). The mean values of all the wave-storm intensity variables decrease in time, except for D at the northern part of the Catalan coast. The decline in the mean values of E, H_p and T_p means milder wave conditions in the future. It could be possible in the Mediterranean Sea, which is much more sheltered than the major oceans such as the Atlantic or the Pacific Oceans (Hemer et al. (2013)). However, the increase in D may have a relevant impact on harbours, which would

require adaptive engineering (Burcharth et al. (2014); Sánchez-Arcilla et al. (2016)). The mean and variance of the storm variables in the Catalan Coast depend mainly on the dynamics of the climate-indices, such as the second time-derivative of EA and SC. Usually, what used to be discussed was the effect of the nominal value of the climate-indices. The relevance of the time-derivatives is evaluated in sub-section sec. 8.2.

The selected non-stationary HAC structure is the type A HAC-structure (see Fig. 8.1). The reason is a strong dependence of H_p to E and D. The HACs of the wave-storms present an increasingly strong dependence between E and D over 1950-2100. As the mean values of E and D are approximately constant in this period of time, there seems to be a fixed configuration of the energy and the duration of the wave-storms.

8.1.2. The non-stationary approach in the northwestern Black Sea

The northwestern Black Sea is also fetch-limited and micro-tidal. A process-based approach provide data for the statistical characterization. These wave-data projections were made for the greenhouse gas emission scenarios RCP4.5 and RCP8.5, in the years 1950-2100. The non-stationary characterization of wave-storms in the northwestern Black Sea uses the same methodology as in the Catalan Coast.

The uncertainty due to the GCMs is small. However, unlike in the Catalan Coast, some GCMs show different wave-storm conditions. They are the GCMs from the MIROC and the GFDL branches. The reason might be that these GCMs were not centred over Europe, where the northwestern Black Sea is located. What is more, unlinke in the Catalan Coast, The MIROC GFDL branch of GCM might not produce the same atmospheric climate conditions as the European centred GCMs.

The non-stationary wave-storm threshold in the northwestern Black Sea is mostly influenced by the first and the second time-derivatives of SC (in RCP4.5) and the second time-derivative of EA (in RCP8.5). The wave-storminess is of around $27 - 35$storms/year and is not affected by any one of the proposed climate-pattern. Therefore, the wave-storminess in the northwestern Black Sea is more independent from the proposed climate patterns, despite the sensitivity of the wave-storm threshold to them. It is possible that the wave-storm threshold depended mildly on them, but that the final wave-storminess did not.

In the northwestern Black Sea, one of the main forcings for waves are winds. The extreme winds are highly influenced by orographic effects and atmospheric climate patterns. Besides, despite that Climate Change leads to important modifications in the variability and intensity of winds, it seems that the wave-direction should not change significantly. The movMF that are fitted to the wave-directions present constant mean and the variance. Thus, only wave conditions are modelled.

The detailed evolution of the mean values and the variance of the wave-storm intensity variables are diverse, showing different trends. The mean H_p and T_p remain

fairly constant over the 21^{st} century, in both RCP scenarios. The average E is more markedly increasing in the RCP4.5 scenario than in the RCP8.5 scenario. The mean D increases in the RCP4.5 scenario, opposed to being constant under the RCP8.5 scenario. As for the variances of the four wave-storm variables, they show the same behaviour for both RCP scenarios: the rising average values induce an increase in the variances.

In both emission scenarios, the mean of the wave-storm variables have a more positive trend than in the Catalan Coast. The Black Sea has a similar fetch to the northwestern Mediterranean Sea. What is a more positive trend is actually a more constant mean value of the wave-storm variables. Considering also the independence of wave-storminess to climate-patterns, it could be suggested that the effects of Climate Change were milder in the northwestern Black Sea than in the northwestern Mediterranean Sea. The mean and variance of the storm variables in the northwestern Black Sea depend on all three climate-patterns: NAO, EA and SC, but they are most strongly influenced by the EA. The importance of each one of the climate-pattern on the parameters of the GDPs are different for the Catalan Coast and the northwtestern Black Sea. This might be due to the different effect that each climate-pattern has for distinct parts of the Planet.

The non-stationary HAC structure in the northwestern Black Sea is represented by the type B HAC-structure (see Fig. 8.1). This HAC structure denotes a weaker dependence of H_p to E and D. In both RCP scenario, the general dependence among the wave-storm intensity variables increase with time. It should be noted that in the Catalan Coast there is not such an upward general trend among the wave-storm intensity variables. It seems that, despite a HAC structure that reflects a weaker dependence of H_p to E and D, this dependence was recovered during the 21^{st} century. Therefore, the wave-storms in the northwestern Black Sea might present a similar joint probability structure to the wave-storms in the Catalan Coast, later in this century.

8.2. Integrated discussion

In comparison, the SIMAR wave hindcasts characterized in the stationary approach to the Catalan Coast only spans years 1996-2013, whereas the wave-projections are for 1950-2100. The statistical significance of wave-storms in the non-stationary approaches is much higher. This advantage has been profitted precisely by considering non-stationarity. In any case, in the presence of the Climate Change, the assumption of stationarity of the wave-climate would not have been realistic, since one of the basic assumptions of Climate Change is that the climatic trend would vary.

The wave-storm threshold does have one limitation. Wave-storminess seems to be systematically higher in non-stationary models. The wave-storm threshold in the non-stationary approach can be lower at some moments in time, thus providing a lower threshold for events to be classified as storms. This does not happen to the

stationary model in the Catalan Coast, which has a constant wave-storm threshold for the whole time-frame, and thus should be kept in mind.

A main similarity between the two study areas is that the E and D, have a high dependence. Is is so because of the definition of E. Another similarity is that there is a mild general dependence of the wave-storm intensity variables. A main divergence is that the average value of the wave-storm intensity variables seem to have a more negative trend in the Catalan coast than in the northwestern Black Sea.

θ_p^* has been modelled in the two study areas because the wave-storm intensity might depend on the directionality of the wave-storm at its peak. Nevertheless, this has not been of concern in the northwestern Black Sea, where there has not been a need to model θ_p^*, as it stays constant from 1950 to 2100.

There are a series of discoveries in this thesis that might help improve knowledge on Coastal Engineering. For instance, there is a larger effect of the dynamics of the climate-patterns on the wave-storm variables in the Catalan Coast. This shows how the dynamics of the climate-patterns can also have a significant effect on the wave climate in one such region. This phenomenon might be extensible to other ocean environments and it poses a certain risk. There might be periods of time where a region presented similar characteristics of the atmosphere, but this part of the atmosphere might be accelerating to other configurations. According to the discovery above, the regional wave-storms might acquire new characteristics under these circumstances. A second possible new discovery is that the configuration of E and D in a wave-storm in the Catalan Coast is increasingly fixed in the 21^{st} century. Similarly, the configuration of all the wave-storm intensity variables is increasingly fixed in the northwestern Black Sea, in the 21^{st} century. This simplify matters as the future configurations of wave-storms are expected to be similar to the ones at the beginning of the century.

These statistical models have an applicability into Civil Engineering. Specifically, the stationary statistical model can serve to generate data of synthetic wave-storms. This data can be used for several purposes, one of which is as input for numerical models to simulate wave conditions in a given coastal area. Alternatively, the non-stationary approaches provides statistical information on wave-storm variables and their joint probability structure. This can enhance Level 3 structure designs (Martín-Hidalgo et al. (2014)), a level of design that consider the uncertain nature of the hydrodynamic forcings with probabilistic distributions. Pre-defining a dependence structure among the wave-storm variables avoids to model unfeasible or extremely rare combinations of the wave-storm intensity variables (for instance, those with high E, H_p, T_p and D at the same time). Moreover, it can help stakeholders take decisions involving risk management, by knowing better the most probable combinations of wave-storm characteristics, how these evolve in time, and how they relate to NAO, EA and SC.

9. Conclusions

9.1. Main goals and summary of the methodology

Wave-storms are responsible for major changes in the coast. The probability of occurrence of certain configurations of a wave-storm, like extreme H_p and D, has been a growing concern in Coastal Engineering. Moreover, Climate Change can lead to non-stationarity of the wave-storm intensity variables. Also, atmospheric climate-patterns might present a connection to wave-storm characteristics. It is expected that research on these aspects would help avoid an inappropriate risk assessment of the wave-storms.

The first aim of this thesis is to characterize the dependence structure of the wave-storm intensity variables in a fetch-limited, micro-tidal enviroment. Hence, a stationary approach is to characterize wave-storms variables in the present climate of the Catalan Coast. The proposed stationary model is formed by three sub-models: wave-storm intensity, wave directionality and intra-time. A stationary wave-storm threshold is used to define wave-storms. Stationary GPDs are used to characterize the wave-storm intensity variables of E, E_u, T_p and D. Stationary HACs are employed to characterize the joint probability structure of these variables. The θ_p^* is modelled with movMF and related to E, T_p and D through a multinomial logistic model. Growth-decay rates are related to D through a third-degree polynomial. A geometric shape is proposed to describe the time-series of the wave-height during wave-storms. This figure can be either a triangle or an irregular-trapezoid.

The second aim is to incorporate the non-stationarity, caused by Climate Change, in the characterization. The same methodology for a non-stationary approach is applied to the Catalan Coast and the northwestern Black Sea. The E, H_p, T_p and D of wave-projects in the years 1950-2100 are characterized. The projections are for the RCP8.5 emission scenario, in the Catalan Coast, whereas it is for both the RCP4.5 and the RCP8.5 scenarios, for the northwestern Black Sea. Furthermore, the relationship of the wave-storm intensity variables to the climate-indices NAO, EA, SC and their first two time-derivatives is established. A non-stationary wave-storm threshold and non-stationary GPDs are employed. A series of HACs for short time periods of 15 years provide time-dependent parameters of HAC. The temporal evolution of θ_p^* is fitted to movMF and evaluated. The uncertainty from the GCMs that are used in each study area is bound.

9.2. Summary of the results

This section concludes the main results from the two study areas in this compendium. For Lin-Ye et al. (2016), chapter 4 adresses: i) the wave storminess, ii) the intra-time sub-model, iii) the wave directionality sub-model and iv) the wave-storm intensity sub-model. For Lin-Ye et al. (2017), chapter 5 addresses: i) the bounding of the uncertainty from the GCMs, ii) the wave-storm threshold and its relationship to climate-indices, iii) the wave-storminess and its relationship to climate-indices, iv) the wave-storm intensity variables and their relationship to climate-indices and v) the joint probability structure and the values of its parameters. For Lin-Ye et al. (2018) (chapter 7), the issues addressed are the same as for chapter 5, only that the θ_p^* is discussed before the item iv).

The stationary approach shows a wave-storminess in the Catalan Coast of approximately $1-15$storms/year. The joint probability structure of the wave-storm intensity variables (E, E_u, T_p and D) can be characterized by two types of HAC structures. The dependence of E_u to E and D is stronger in one type of HAC structure than in the other one. In both the Catalan Coast and the northwestern Black Sea, extreme events with maximum values of all the storm variables is unlikely, while the probability of storms of extreme storm energy and duration is considerably high. The principal wave directions on the northern Catalan coast are North and East, whereas eastern and southern directions are predominant in the central and southern Catalan coast. The most appropriate geometric figure to describe the evolution of the wave-height is a irregular-trapezoid. For D over 100h, the peak of the wave-storm is located farther from the start of the wave-storm than from the end.

In the non-stationary characterization of wave-storm intensity variables of the Catalan Coast, the wave-storm threshold is mostly influenced by NAO, the second time-derivative of EA and SC. The wave-storminess is $5-25$storms/year. It is boosted by negative NAO. All storm variables (E, H_p, T_p and D) decrease in time, except for the storm duration at the northern part of the Catalan coast. The mean and variance of the storm variables in the Catalan Coast depend mainly on the dynamics of the climate-indices, such as the second derivative of EA and SC. The selected HAC structure presents a stronger dependence of H_p with E and D, while the wave-storms have an increasingly strong dependence between E and D. The uncertainty due to the GCM is small, as all GCMs provide similar wave-storm conditions in this region.

In the non-stationary approach on the northwestern Black Sea, the non-stationary wave-storm threshold in the northwestern Black Sea is mostly affected by the first and the second time-derivatives of SC (RCP4.5) and the second time-derivative of EA (RCP8.5). The wave-storminess approximately $27-35$storms/year and is not affected by any of the proposed climate-pattern. The wave directionality stays the same during 1950-2100. Thus, only wave conditions are modelled. The uncertainty due to the GCM is small. The only GCMs that show different wave-storm conditions are the ones from the MIROC and the GFDL branches.

At the northwestern Black Sea, under RCP4.5 and RCP8.5 scenario, the mean H_p and T_p remain fairly constant over the 21^{st} century. The average storm energy rise more steadily in the RCP4.5 scenario than in the RCP8.5 scenario. The mean D increases in the RCP4.5 scenario. On the contrary, it is constant under the RCP8.5 scenario. The variance of the four wave-storm variables increase with increasing average values, under both RCP scenarios.

In comparison, the average value of the wave-storm variables seem to have a more positive trend in the northwestern Black Sea than in the Catalan Coast. The mean and variance of the storm variables in the northwestern Black Sea depend on the three climate-patterns (NAO, EA and SC), but they are most frequently influenced by the EA. The pseudo-non-stationary HAC structure in the northwestern Black Sea shows a weaker dependence of H_p with E and D. Under both RCP scenario, the dependence among E, H_p, T_p and D increase with time.

9.3. Applicability of the results

The characterization of wave-storm intensity, directionality and intra-time variables, their joint probability structure and the effect of atmospheric climate-patterns on them have an applicability in Civil Engineering. They provide statistical information about the ocean environment, which can enhance Level 3 structure designs, a level of design that considers the uncertain nature of the hydrodynamic forcings with probabilistic distributions. Pre-defining a dependence structure among the wave-storm variables avoids to model unfeasible or extremely rare combinations of wave-storm intensity variables. Also, this characterization can help stakeholders take decisions involving risk management.

10. Future works

(A) Evaluate the uncertainty in the future projections via Bayesian statistics

One of the reviewers of "Multivariate statistical modelling of future marine storms" suggested to evaluate the uncertainty of the future projections in the Catalan coast. The uncertainty from the GCM has been evaluated and the one from other sources could be tackled in future works.

(B) Introducing asymmetric copulas in the storm-model

In "Multivariate statistical modelling of future marine storms", we have been suggested by one of the reviewers to use the asymmetric copulas detailed in Vanem (2016). Vanem argues that the copula structure does not have a symmetric tail, a symmetric relationship between the extreme values. That is, Gumbel, Frank, Joe, and all the common Archimedean copulas do not seem suitable to him to model the joint probability structure of wave-storm intensity variables. Vanem proposes the Marshal-Olkin copula, which is a parametric copula that is asymmetric. The further complication in this approach is to integrate an asymmetric copula to a non-stationary model. Therefore, the non-stationary model has been first built with a symmetric copula. The asymmetric copula awaits to be introduced in future works.

(C) Fully non-stationary model

The HAC structure used in the non-stationary models are chains of stationary HACs. A fully non-stationary model, with a non-stationary copula structure, could be bivariate first. Then, it could include more variables, if possible.

(D) Relate the growth-decay rates to the storm energy, duration, wave-height, and wave-period

The growth-decay rates have been related to D in the stationary model, via a polynomial. A skewed normal bi-variate probability distribution function has been suggested, later. However, it has not yet been possible to implement it, as the methodology to build a non-stationary storm model had to be consolidated, before looking for a joint probability structure that included the growth-decay rates.

(E) Partition of wind and swell waves

Several attempts have been made to separate wind and swell waves into different subsets. One statistical model could be proposed for each one of these two types of waves. Then, it would be possible to compare the wave-storm intensity variables and their joint probability structures, for the different types of waves.

(F) Neural networks, deep learning and other techniques

Some hot topics right now in science are neural networks and deep learning. These might improve the knowledge in fields like remote sensing data analysis of wave-storm intensity variables. Also, they can help improve our non-stationary models.

The proposed ideas for future work are generally not dramatically different from the work introduced in this thesis. They are small improvements to the existing statistical models. If possible, several decades of research work should help improve many other aspects of ocean physical sciences, which integrated with advanced mathematics and computer sciences, should help existing methods of Civil Engineering to evolve.

A. Errata

A.1. A multivariate statistical model of extreme events: an application to the Catalan coast

Fig. 7a shows the 95^{th} quantile, not the 59^{th}.

A.2. Multivariate statistical modelling of future marine storms

Eq. 8 should be written as $\left|\ln \frac{x(1-y)}{y(1-x)}\right|$. This has first been pointed out, after publication, by Prof. J.J. Egozcue.

Bibliography

Aitchison, J.: 1982, 'The statistical analysis of compositional data'. *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 139–177. 3.2.4

Aitchison, J.: 1992, 'On criteria for measures of compositional difference'. *Mathematical Geology* **24**(4), 365–379. 3.2.4

Akaike, H.: 1987, 'Factor analysis and AIC'. *Psychometrika* **52**(3), 317–332. 3.2.1.2

Barnerjee, A., I. Dhillon, J. Ghosh, and S. Sra: 2005, 'Clustering on the Unit Hypersphere using von Mises-Fisher Distributions'. *Journal of Machine Learning Research* **6**, 1345–1382. 3.2.2

Barnston, A. G. and R. E. Livezey: 1987, 'Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns'. *Monthly Weather Review* **115**(6), 1083–1126. 3.1.2, 3.1.2

Bolaños, R.: 2004, 'Tormentas de oleaje en el Mediterraneo: Física y predicción'. Ph.D. thesis, BarcelonaTech. 2.1

Bolaños, R., G. Jordà, J. Cateura, J. López, J. Puigdefàbregas, J. Gómez, and M. Espino: 2009, 'The XIOM: 20 years of a regional coastal observatory in the Spanish Catalan coast'. *Journal of Marine Systems* **77**, 237–260. 2.1, 8.1.1.1

Burcharth, H. F., T. L. Andersen, and J. L. Lara: 2014, 'Upgrade of coastal defence structures against increased loadings caused by climate change: A first methodological approach'. *Coastal Engineering* **87**, 112 – 121. 8.1.1.2

Casas-Prat, M., K. L. McInnes, M. A. Hemer, and J. P. Sierra: 2016, 'Future wave-driven coastal sediment transport along the Catalan coast (NW Mediterranean)'. *Regional Environmental Change* pp. 1–12. 1.1

CIIRC: 2010, 'State of the Catalan coastline'. Technical report. 3.1.1, 3.2.1.1

Coles, S.: 2001, *An introduction to Statistical modeling of extreme values*. Springer. 3.2.1.1

Corbella, S. and D. Stretch: 2012, 'Multivariate return periods of sea storms for coastal erosion risk assessment'. *Natural Hazards and Earth System Sciences* **12**, 2699–2708. 1.1, 3.2.1.3

Corbella, S. and D. D. Stretch: 2013, 'Simulating a multivariate sea storm using Archimedean copulas'. *Coastal Engineering* **76**, 68–78. 3.2.1.3

de Waal, D. J. and P. H. A. J. M. van Gelder: 2005, 'Modelling of extreme wave heights and periods through copulas'. *Extremes* **8**, 345–356. 1.1

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, et al.: 2011, 'The ERA-Interim reanalysis:

configuration and performance of the data assimilation system'. *Quarterly Journal of the Royal Meteorological Society* **137**(656), 553–597. 3.1.5

Egozcue, J. J. and V. Pawlowsky-Glahn: 2011, 'Evidence information in Bayesian updating'. In: *Proceedings of the 4th International Workshop on Compositional Data Analysis*. pp. 1–13. 3.2.4, 3.2.4

Egozcue, J. J., V. Pawlowsky-Glahn, G. Mateu-Figueras, and C. Barceló-Vidal: 2003, 'Isometric logratio transformations for compositional data analysis'. *Mathematical Geology* **35**(3), 279–300. 3.2.4

Egozcue, J. J., V. Pawlowsky-Glahn, M. I. Ortego, and R. Tolosana-Delgado: 2006, 'The effect of scale in daily precipitation hazard assessment'. *Natural Hazards and Earth System Sciences* **6**(3), 459–470. 3.1.1, 3.1.3, 3.1.4

Embrechts, P., C. Kluppelberg, and T. Mikosch: 1997, *Modeling extreme events for insurance and finance*. Berlin: Springer. 1

Fessler, J. A.: 1991, 'Nonparametric fixed-interval smoothing with vector splines'. *Signal Processing, IEEE Transactions on* **39**(4), 852–859. 3.2.1.2

Gan, F., K. Koehler, and J. Thompson: 1991, 'Probability Plots and Distribution Curves for Assessing the Fit of Probability Models'. *The American Statistician* **45**(1), 14–21. 3.2.1.3

Genest, C. and B. Remillard: 2004, 'Test of independence and randomness based on the empirical copula process'. *Test* **13**, 335–369. 3.2.1.3

Goda, Y.: 2010, *Random Seas and Design of Maritime Structures*, Vol. Advanced Series on Ocean Engineering Vol. 33. World Scientific, 3rd edition. 1.1

Gomez, M. and J. Carretero: 2005, 'Wave forecasting at the Spanish coasts'. *Journal of Atmospheric and Ocean Science* **10**(4), 389–405. 3.1.1

Gràcia, V., M. García, M. Grifoll, and A. Sánchez-Arcilla: 2013, 'Breaching of a barrier under extreme events. The role of morphodynamic simulations'. *Journal of Coastal Research* **65**, 951–956. 1.1

Grifoll, M., A. Aretxabaleta, and M. Espino: 2015, 'Shelf response to intense offshore wind'. *Journal of Geophysical Research: Oceans* **120**(9), 6564–6580. 2.1

Hemer, M. A., Y.-L. Fan, N. Mori, A. Semedo, and X.-L. Wang: 2013, 'Projected changes in wave climate from a multi-model ensemble'. *Nature climate change* **3**(5), 471–476. 3.1.5, 8.1.1.2

Hemer, M. A. and C. E. Trenham: 2016, 'Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble'. *Ocean Modelling* **103**, 190 – 203. 1.1

Hinkel, J., D. Lincke, A. T. Vafeidis, M. Perrette, R. J. Nicholls, R. S. J. Tol, B. Marzeion, X. Fettweis, C. Ionescu, and A. Levermann: 2014, 'Coastal flood damage and adaptation costs under 21st century sea-level rise'. *Proceedings of the National Academy of Sciences* **111**(9), 3292–3297. 1.1

Hinkel, J., R. Nicholls, R. Tol, Z. Wang, J. Hamilton, G. Boot, A. Vafeidis, L. McFadden, A. Ganopolski, and R. Klein: 2013, 'A global analysis of erosion of

sandy beaches and sea-level rise: An application of DIVA'. *Global and Planetary Change* **111**, 150–158. 1.1

Holthuijsen, L.: 2007, *Waves in oceanic and coastal waters*. Cambridge University Press. 3.1.1

Hosmer, D., S. Lemeshow, and R. Sturdivant: 2013, *Applied Logistic Regression, 3rd Edition*. Wiley. 528 pp. 3.2.2

Jonathan, P., K. Ewans, and D. Randell: 2013, 'Joint modelling of extreme ocean environments incorporating covariate effects'. *Coastal Engineering* **79**, 22 – 31. 3.2.1.2

Kendall, M.: 1937, 'A new measure of rank correlation'. *Biometrika* **6**, 83–93. 3.2.1.3

Koenker, R.: 2005, *Quantile regression*, Econometric Society Monographs. Cambridge University Press. 3.2.1.2

Kullback, S.: 1997, *Information theory and statistics*. Courier corporation. 3.2.4

Kwak, J., A. St-Hilaire, F. Chebana, and G. Kim: 2017, 'Summer season water temperature modeling under the Climate Change: Case study for Fourchue River, Quebec, Canada'. *Water* **9**(5), 346. 3.1.2

Li, F., P. van Gelder, R. Ranasinghe, D. Callaghan, and R. Jongejan: 2014, 'Probability modelling of extreme storms along the Dutch coast'. *Coastal Engineering* **86**, 1–13. 1.1

Lin-Ye, J., M. García-León, V. Gràcia, M. I. Ortego, P. Lionello, and A. Sánchez-Arcilla: 2017, 'Multivariate statistical modelling of future marine storms'. *Applied Ocean Research* **65**, 192–205. 2.2, 3.3, 3.4, 9.2

Lin-Ye, J., M. García-León, V. Gràcia, M. I. Ortego, A. Stanica, and A. Sánchez-Archilla: 2018, 'Multivariate hybrid modelling of future wave-storms at the northwestern Black Sea'. *Water* **10(2)**. 9.2

Lin-Ye, J., M. García-León, V. Gràcia, and A. Sánchez-Arcilla: 2016, 'A multivariate statistical model of extreme events: An application to the Catalan coast'. *Coastal Engineering* **117**, 138 – 156. 2.1, 3.1, 3.2, 9.2

Lionello, P., F. Abrantes, L. Congedi, F. Dulac, M. Gacic, D. Gomis, C. Goodess, H. Hoff, H. Kutiel, J. Luterbacher, S. Planton, M. Reale, K. Schroder, M. V. Struglia, A. Toreti, M. Tsimplis, U. Ulbrich, and E. Xoplaki: 2012, *The climate of the Mediterranean region*. Elsevier. 2.1

Luo, M., F.-H. Meng, T. Liu, Y.-C. Duan, A. Frankl, A. Kurban, and P. De Maeyer: 2017, 'Multi–model ensemble approaches to assessment of effects of local Climate Change on water resources of the Hotan River Basin in Xinjiang, China'. *Water* **9**(8), 584. 3.1.2

Mardia, K. and P. Jupp: 2009, *Directional statistics*, Vol. 494. John Wiley & Sons. 3.2.2

Martín-Hidalgo, M., M. J. Martín-Soldevilla, V. Negro, P. Aberturas, and J. S. López-Gutiérrez: 2014, 'Storm evolution characterization for analysing stone armour damage progression'. *Coastal Engineering* **85**(Supplement C), 1 – 11. 8.2

Martín-Soldevilla, M. J., M. Martín-Hidalgo, V. Negro, J. S. López-Gutiérrez, and P. Aberturas: 2015, 'Improvement of theoretical storm characterization for different climate conditions'. *Coastal Engineering* **96**, 71–80. 1.1

Melby, J. and N. Kobayashi: 2011, 'Stone armor damage initiation and progression based on the maximum wave momentum flux'. *Journal of Coastal Research* **27**, 110–119. 1.1

Merkens, J.-L., L. Reimann, J. Hinkel, and A. T. Vafeidis: 2016, 'Extended Shared Socioeconomic Pathways for Coastal Impact Assessment: Spatial Coastal Population Scenarios'. In: *EGU General Assembly Conference Abstracts*. 1

Montero, P. and J. A. Vilar: 2014, 'Tsclust: An R package for time series clustering'. *Journal of Statistical Software*. 3.2.5

Muraleedharan, G., C. Lucas, and C. Guedes Soares: 2016, 'Regression quantile models for estimating trends in extreme significant wave heights'. *Ocean Engineering* **118**, 204 – 215. 3.2.1.2

Neumann, B., A. T. Vafeidis, J. Zimmermann, and R. J. Nicholls: 2015, 'Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment'. *PLoS ONE* **10**(3), e0118571. 1

Northrop, P. J. and P. Jonathan: 2011, 'Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights'. *Environmetrics* **22**(7), 799–809. 3.2.1.2

Okhrin, O., Y. Okhrin, and W. Schmid: 2013, 'On the structure and estimation of hierarchical Archimedean copulas'. *Journal of Econometrics* **173**, 189–204. 3.2.1.3, 3.2.1.3

Ortego, M., R. Tolosana-Delgado, J. Gibergans-Báguena, J. Egozcue, and A. Sánchez-Arcilla: 2012, 'Assessing wavestorm hazard evolution in the NW Mediterranean with hindcast and buoy data'. *Climatic Change* **113**, 713–731. 3.1.1

Pawlowsky-Glahn, V. and J. J. Egozcue: 2001, 'Geometric approach to statistical analysis on the simplex'. *Stochastic Environmental Research and Risk Assessment* **15**(5), 384–398. 3.2.4

Rockel, B., A. Will, and A. Hense: 2008, 'The Regional Climate Model COSMO-CLM (CCLM)'. *Meteorologische Zeitschrift* **17**, 347–348. 3.1.3

Salvadori, G., C. De Michele, N. Kottegoda, and R. Rosso: 2007, *Extremes in nature: An approach using copulas*. Springer. 1.1, 3.2.1.3, 3.2.3

Salvadori, G., F. Durante, C. De Michele, M. Bernardi, and L. Petrella: 2016, 'A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities'. *Water Resources Research* pp. n/a–n/a. 3.2.1.3

Sánchez-Arcilla, A., M. García, and V. Gràcia: 2014, 'Hydro-morphodynamic modelling in Mediterranean storms - errors and uncertainties under sharp gradients'. *Nat. Hazards Earth Syst. Sci.* **14**, 2993–3004. 1.1, 3.2.3

Sánchez-Arcilla, A., J. Gómez-Aguar, J. Egozcue, M. Ortego, P. Galiatsatou, and P. Prinos: 2008a, 'Extremes from scarce data: The role of Bayesian and scaling

techniques in reducing uncertainty'. *Journal of Hydraulic Research* **46**(sup2), 224–234. 3.2.3

Sánchez-Arcilla, A., D. González-Marco, and R. Bolaños: 2008b, 'A review of wave climate and prediction along the Spanish Mediterranean coast'. *Nat. Hazard. Earth Sys.* **8(6)**, 1217–1228. 1

Sánchez-Arcilla, A., D. González-Marco, and R. Bolaños: 2008c, 'A review of wave climate and prediction along the Spanish Mediterranean coast'. *Nat. Hazard. Earth Sys.* **8(6)**, 1217–1228. 2.1, 3.1.3

Sánchez-Arcilla, A., J. P. Sierra, S. Brown, M. Casas-Prat, R. J. Nicholls, P. Lionello, and D. Conte: 2016, 'A review of potential physical impacts on harbours in the Mediterranean Sea under climate change'. *Regional Environmental Change* pp. 1–14. 1.1, 8.1.1.2

Sierra, J. P., I. Casanovas, C. Mösso, M. Mestres, and A. Sánchez-Arcilla: 2016, 'Vulnerability of Catalan (NW Mediterranean) ports to wave overtopping due to different scenarios of sea level rise'. *Regional Environmental Change* **16**(5), 1457–1468. 1.1

Sierra, J. P., M. Casas-Prat, M. Virgili, C. Mösso, and A. Sánchez-Arcilla: 2015, 'Impacts on wave-driven harbour agitation due to climate change in Catalan ports'. *Natural Hazards and Earth System Sciences* **15**(8), 1695–1709. 1.1

Sklar, A.: 1959, *Fonctions dé repartition à n dimension et leurs marges*. Université Paris 8. 3.2.1.3

Small, C. and R. J. Nicholls: 2003, 'A global analysis of human settlement in coastal zones'. *Journal of Coastal Research* pp. 584–599. 1

Stocker, T., D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley: 2013, 'IPCC, 2013: Summary for policymakers'. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. 3.1.3, 3.1.4

Tamura, Y., T. Sato, M. Ooe, and M. Ishiguro: 1991, 'A procedure for tidal analysis with a Bayesian information criterion'. *Geophysical Journal International* **104**(3), 507–516. 3.2.1.2

Taylor, K. E., R. J. Stouffer, and G. A. Meehl: 2012, 'An overview of CMIP5 and the experiment design'. *Bulletin of the American Meteorological Society* **93**(4), 485–498. (document), 3.1

Tolman, H.: 2009, 'User manual and system documentation of WAVEWATCH III version 3.14.'. Technical report, NOAA / NWS / NCEP / MMAB. 3.1.1

Tolosana-Delgado, R., M. Ortego, J. Egozcue, and A. Sánchez-Arcilla: 2010, 'Climate change in a Point-over-threshold model: an example on ocean-wave-storm hazard in NE Spain'. *Advances in Geosciences* **26**, 113–117. 3.2.1.1

Trenberth, K. E., J. T. Fasullo, and T. G. Shepherd: 2015, 'Attribution of climate extreme events'. *Nature Climate Change* **5**(8), 725–730. 8.1.1.2

Unden, P., L. Rontu, H. Jarvinen, P. Lynch, J. Calvo, G. Cats, J. Cuixart, K. Eerola, C. Fortelius, J. A. Garcia-Moya, C. Jones, G. Lenderlink, A. McDonald, R. McGrath, B. Navascues, N. W. Nielsen, V. Odergaard, E. Rodrigues, M. Rummukainen, R. Room, K. Shattler, B. H. S. nd H. Savijarvi, B. Schreur, R. Sigg, H. The, and A. Tijm: 2002, 'HIRLAM-5 Scientific Documentation'. Technical report, HIRLAM-5 Project, c/o Per Unden SMHI, S-601 76 Norrkoping, SWEDEN. 144 p. 3.1.1

Valchev, N., I. Davidan, Z. Belberov, A. Palazov, N. Valcheva, and D. Chin: 2010, 'Hindcasting and assessment of the western Black Sea wind and wave climate'. *Journal of Environmental Protection and Ecology* **11**(3), 1001–1012. 1.1, 2.2

Vanem, E.: 2016, 'Joint statistical models for significant wave height and wave period in a changing climate'. *Marine Structures* **49**, 180–205. 10

Voldoire, A., E. Sanchez-Gomez, D. Salas y Mélia, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, et al.: 2013, 'The CNRM-CM5.1 global climate model: description and basic evaluation'. *Climate Dynamics* **40**(9), 2091–2121. 3.1.2

Wahl, T., J. Jensen, and C. Mudersbach: 2011, 'A multivariate statistical model for advanced storm surge analyses in the North Sea'. *Coastal Engineering Proceedings* **1**(32), 19. 3.2.1.3

Wahl, T., N. G. Plant, and J. W. Long: 2016, 'Probabilistic assessment of erosion and flooding risk in the northern Gulf of Mexico'. *Journal of Geophysical Research: Oceans* **121**(5), 3029–3043. 1.1

WAMDI Group, S. H., K. Hasselmann, P. A. E. M. Janssen, G. J. Komen, L. Bertotti, P. Lionello, A. Guillaume, V. C. Cardone, J. A. Greenwood, M. Reistad, L. Zambresky, and J. A. Ewing: 1988, 'The WAM model: a third-generation ocean wave predicition model'. *Journal of Physical Oceanography* **18**, 1775–1810. 3.1.1

Wang, X. L., Y. Feng, and V. R. Swail: 2014, 'Changes in global ocean wave heights as projected using multimodel CMIP5 simulations'. *Geophysical Research Letters*. 3.1.5

Wang, X. L., Y. Feng, and V. R. Swail: 2015a, 'Climate change signal and uncertainty in CMIP5-based projections of global ocean surface wave heights'. *Journal of Geophysical Research: Oceans* **120**(5), 3859–3871. 1

Wang, X. L., Y. Feng, and V. R. Swail: 2015b, 'Climate change signal and uncertainty in CMIP5-based projections of global ocean surface wave heights'. *Journal of Geophysical Research: Oceans* **120**(5), 3859–3871. 1.1

Yee, T. W. and C. J. Wild: 1996, 'Vector generalized additive models'. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**(3), 481–493. 3.2.1.2