Università degli Studi di Cagliari

DOTTORATO DI RICERCA

INGEGNERIA DEL TERRITORIO

Ciclo XXVIII

TITOLO TESI

On coastal flooding: from extreme offshore wave climate characterization to wave runup simulation with online-coupled numerical models

Settore/i scientifico disciplinari di afferenza

ICAR/02

Presentata da:	Dott. Massimiliano Ventroni
Coordinatore Dottorato	Prof. Roberto Deidda
Tutor	Prof. Andrea Balzano

Esame finale anno accademico 2015 – 2016

Tesi discussa nella sessione d'esame marzo – aprile 2017

Università degli Studi di Cagliari

DOTTORATO DI RICERCA

INGEGNERIA DEL TERRITORIO

Ciclo XXVIII

PhD Thesis

On coastal flooding: from extreme offshore wave climate characterization to wave runup simulation with online-coupled numerical models

Settore/i scientifico disciplinari di afferenza

ICAR/02

PhD candidate: Dott. Massimiliano Ventroni

Tutor Prof. Andrea Balzano

Massimiliano Ventroni gratefully acknowledges Sardinia Regional Government for the financial support of his PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Found 2007 – 2013 – Axis IV Human Resources Objective 1.3, Line of Activity 1.3.1).

Abstract

Coastal flooding is a topic of great relevance in the context of Coastal Engineering, particularly in the perspective of climate change and related sea level rise. Effective evaluation of coastal flooding at the *hindcast*, *nowcast* or *forecast* level, requires a high degree of interdisciplinary skills because of numerous aspects involved and different space and time scales usually considered. In particular, a peculiar perspective may be depicted by investigating the deep complementarity, as well as the close interrelation, coming up by the combined used of extreme value theory and advanced numerical modelling tools. As a matter of fact, numerical simulations are forced by offshore and/or nearshore boundary conditions, which, in the case of coastal flooding, are usually obtained by long-term statistical predictions of the sea state, related to the extreme regime of wave climate. The aim of this work is primarily focus on analysing and developing an operational methodology for the assessment of coastal flooding induced by extreme waves, with subsequent use of numerical modelling. In particular, the area of interest of wave climate analysis is related to the Sardinian coasts.

It is evident how a reliable specification of boundary conditions is a fundamental component in numerical modelling. First of all, each dataset must formally fulfil correctness, representativeness, reliability and homogeneity characteristics. Unfortunately, even a rigorous assessment of the offshore boundary conditions is affected by an inherent uncertainty, due to the limited spatial and temporal extension which often characterizes a wave-climate dataset source. The sample size largely hinders the effectiveness of the asymptotical hypothesis, which is one of the cornerstones of the extreme value theory. On account of the above limitations, it is therefore still an open question in the literature whether upper limited (e.g. bounded Generalized Pareto) or upper unlimited (e.g. Weibull) distributions are most suitable for engineering purposes. Operationally, an investigation of the adequateness of the dataset at disposal is preliminary carried out. Next, several metrics are used to assess the effectiveness of most popular extreme wave distributions. Finally, the extreme, offshore sea states over the region of interest were computed.

On the other hand, a plethora of numerical models have been implemented within the coastal community, with different levels of accuracy and complexity. A well-known and well-defined focused distinction is represented by the phase-averaged and phase-resolving approach. The main interest of this work is addressed to operational models, where outcomes' accuracy criteria and non-prohibitive computational burden requirements should be well-balanced and properly considered. With this purpose in mind, an online coupling implementation between a spectral phase-averaged model and a non-hydrostatic phase-resolving model is provided. Specifically, two open-source models have been adopted as model components: the spectral *SWAN* model and the non-hydrostatic *SWASH* model.

To sum up, the long term meteo-marine climate over the Sardinian coasts is evaluated first, by considering both mean and extreme characteristics of a heterogeneous dataset, previously deeply-corrected, ranging from waves to wind fields. From this, it is relatively straightforward to compute the offshore-boundary conditions. Second, a numerical coupled model involving two open source

state-of-the-art models, was implemented, in order to enable a seamless and accurate chain simulation of an extreme sea state from offshore up to the shore.

Contents

List of figures

List of tables

Introduction

In the context of coastal engineering, the probabilistic design of marine structures or sea defences is closely related to the statistical prediction of the sea state, such as defined by wave heights, periods, mean directions, etc. These wave climate elements are, furthermore, the data source for both coastal hazard analysis and evaluation of the safety level of coastal structures in the so-called Source-Pathway-Receptor concept (Oumeraci, 2004). In both cases, the extreme value theory is generally adopted to estimate the final values to be used either as design value or as the offshore boundary condition, which are, in turn, assumed to be equal to the input values of the coastal modelling chain.

Due to the fact that a common, standard and generally recognized approach has not been reached yet within the coastal community, but rather, remains a subject of continuous ongoing debate and active research, it was chosen to carried out the extreme value analysis with the intent of bringing together both an engineering and a statistical perspective. Specifically, the classical approach of statistical inference has been considered more suitable, tailored and easily implemented to the available dataset, compared to more advanced techniques, relatively recent but sometimes more complicated. A specific review of this last matter can be found, for example, in Hawkes (2008) and in all related papers, where one might find, among others, either the Bayesian framework, the cross-spectra and singular spectrum analysis, the multivariate extreme analysis and joint distribution, such as defined by copula tools, or even the non parametric modelling through neural network analysis.

The development of this work, as previously mentioned, is restricted to classical inference, which, despite all, represents the necessary starting point of several practical aspects and, most of all, is known to require a greater attention within the Mediterranean Sea in general, and over the Italian coasts in particular: basically, some concerns need to be either amplified or relaxed, due to the inherent lack of offshore data, as well as to the known underestimation of the wind field resulting from numerical models in that area (Cavaleri and Bertotti, 2004). Nonetheless, although hindcasting can be nowadays used to alleviate the lack of spatial density and to improve the temporal scale, it does neither solve completely the problem of underestimation of wind fields nor the corresponding underestimation of wave heights.

In the light of the above reasons and, moreover, because of the inherent small size of our dataset, the intent of bringing together both an engineering viewpoint and a statistical perspective, without chosen a priori one of them, becomes mandatory. In addition, compared to other works (Sartini et al. 2015, Li et al 2012, Mazas and Hamm 2011) some insights into parameter estimation methods, best fit selections and, last but clearly not least important, threshold selection methods (Mazas and Hamm 2011, Deidda 2010, Thompson et al. 2009) were added. Furthermore, the independency and homogeneity for waves and wind fields are determined on the basis of the sea-storm definition, according to the Wave Atlas (Corsini et al. 2004) and Boccotti (2000) recommendations. In addition, in order to enhance and improve the homogeneity, a combination between fetch distribution, sea storm polar plot, and wave climate characteristics, is considered so as to group data belonging to same meteorological events within homogeneous directional clusters (Corsini et al. 2004). From the viewpoint of quantile estimation, numerous parameter estimation methods are employed and the uncertainty associated to both parameters and quantile estimations (e.g. the

confidence intervals) is provided by implementing a Monte Carlo simulation technique. Finally, the best-fit model is assessed by analysing several parametric and non-parametric metrics.

Operationally, a long term wave and wind climate study is firstly provided by scrutinizing data from three wave buoys supplied by the Italian National Wave Measurement Network (Rete Ondametrica Nazionale, RON), and nine coastal wind measurement stations acquired by the Italian National Civil Aviation network (UGM-ENAV). Secondly, an extreme value model inter-comparison is carried out with respect to the most reliable directional wave buoy data.

On the whole, in this work the classical statistical approach of frequency analysis is followed, by using a multi-parent-distribution approach, as suggested by Goda (2010, 2011b), Mazas and Hamm (2011) and Hosking and Wallis 1997), among many others, because the need of taking into account for the small sample size effects.

The second part of this work addresses the development of a one-way, online, open-source coupled model between a spectral phase-averaged model (*SWAN*, Booij et al. 1999) and a non-hydrostatic model (*SWASH*, Zijlema et al. 2011b). The work aims at representing inherently all the relevant spatial scales involved in coastal flooding simulations (i.e. from shallow to deep water, say from 10^1 to 10^3 wave lengths). Phase-averaged and phase-resolving models are two class of wave-models traditionally used to simulate wave fields in coastal areas. Spectral phase-averaged models are based on the spectral wave action balance equation and can thus be conventional applied to larger areas (say from near-shore areas to regional scales of the order 10^2 to 10^3 wave lengths); on the other hand, phase-resolving (or time-domain) models resolve the wave field on the scale of the individual waves, within the wave period, and are based on the time-dependent mass and momentum balance equations. A variety of such kind of models has been presented in the literature, differing as to dimensionality (1D to 3D), kind of motion represented – whether rotational or irrotational –, vertical pressure distribution, dispersive properties, free surface conditions, wave breaking representation, possible inclusion of turbulence models, bottom friction specification, wetting and drying algorithm, and numerical solution technique. Models of very different ease of use and computational efficiency are thus included in this class. As a rule of thumb, most frequently used vertically-integrated models are suitable for simulating wave motion over narrower areas of the order of 10^1 to 10^2 wave lengths (say from very shallow to relative intermediate waters).

It is prominent the complementary characteristics on these modelling approaches, since weakness of each model component can be overcome by the strength of the other. On one hand, it seems appropriate to use phase-resolving models closed to the shore (i.e. in shallow water) in order to improve wave simulation, where phase-averaged wave action balance models are known to suffer from inaccurate prediction of the wave spectrum, due to, e.g., incomplete modelling of wave diffraction, reflection, or nonlinear interactions. The majority of these physical processes are inherently accounted for by time-domain models in a straightforward way. It is also to be pointed out that, for instance, triad wave-wave interactions are usually modelled in phase-averaged models through an explicit parametric formulation which comes from a phase-resolving model, i.e. the spectral version of the Boussinesq model proposed by Madsen and Sorensen (1993) is implemented in the Lumped Triad Approach (LTA) by Eldeberky et al (1996). Nonetheless, despite of the huge and numerous improvements achieved on the operational modelling in the last decades, so far it still

holds that "*to date there is no comprehensive model formulation for fully directional wave-wave interactions over two-dimensional bathymetry, applicable to arbitrary scales of propagation and suitable for operational wave forecasting problems*" (The *WISE* Group 2007, pg. 636).

On the other hand, extending phase-resolving models to the offshore would require using fully non-hydrostatic models with a large number of layers, resulting in excessive computational burden, not rewarded by a comparably better representation of waves at the outer limit of the nearshore. Relaxing the requirement for very accurate wave dispersion representation, which would be an issue in phase-resolving modelling up to offshore, an improved balance can be achieved between the limiting computational burden on the one hand, and the effectiveness, efficiency, and accuracy usually sought on the other hand, by restricting the use of phase-resolving models to the nearshore. In fact, the need for both wave dispersion and nonlinearity with models that are suitable for practical application is the main reason for the number of different existing formulations of phase-resolving models.

Therefore, it is often convenient to drive phase-resolving models running in the near-shore with boundary conditions obtained by phase-averaged wave models run over larger areas, with the aim of simulating, among others, wind-wave propagation, nonlinear shoaling, diffraction and refraction, wave runup, overtopping, inundation, infra-gravity waves, harbour entrances, and ship waves.

Objective and outline

The overarching objective of the present study is twofold. First, the attention is focused on the analysis of an evaluation methodology of the wave climate over the Sardinian coasts – at medium and extreme level –, within the framework of Regional mapping of Coastal Flooding Hazard. Second, an online, one-way coupled model implementation between a spectral wave model and a non-hydrostatic model has been developed, with the aim of modelling, seamlessly and efficiently, the wave evolution, from offshore up to the shore.

This thesis is composed by 6 Chapters. In Chapter 1 a literature review is provided with respect to both statistical and numerical modelling contexts. Chapter 2 describes the mathematical formulation of extreme value theory. Chapter 3 introduces the long-term wave climate of both wave and wind fields and the related data quality control procedures. Chapter 4 is focused on the application of the EV model inter-comparison on the most reliable wave-buoy dataset. Chapter 5 describes SWAN and SWASH governing equations and the implementation of the coupled SWAN + SWASH model, with a validation with laboratory data. Finally, conclusion are given in Chapter 6.

1 Literature Review

1.1 Statistical frequency and extreme value overview

Extreme value analysis may be generally addressed on either block-maxima (*BM*), peaks-over-threshold (*POT*), or Poisson point process (*PP*) methods. Synthetically, several choices should be considered within the frequency analysis framework, involving among others: de-clustering methodology, threshold selection, parent distribution selection, plotting position formula, parameter estimation method, confidence interval computing, and best-fit model selection. In the following paragraphs, the main framework followed in this work for each of these themes is briefly reported.

1.1.1 Parent distribution – multi-distribution approach

Among many cumulative distribution functions (CDF) employed in extreme data analysis, three distributions of the *Weibull* (*WBL*) family (Weibull 1951), the Generalized Pareto (*GP*) family (Pickands 1975, Hosking and Wallis 1987), and the Generalized Extreme-Value (*GEV*) family (Fisher and Tippet 1922, Gnedenko 1943, Gumbel 1958) are considered in this study. Following Goda (2011a), this set is assumed to be representative for any probable population in the field of extreme wave heights. Furthermore, this is consistent with the classical frequency analysis approach, where a set of parent distributions has to be chosen to statistically find and mimic the so-called true distribution. In addition, as underlined later, this is the methodology recommended by Hosking and Wallis (1997) in order to find out the best marginal distribution, and, for instance, highlighted by other authors (e.g. the multi-distribution approach suggested by Mazas and Hamm 2011).

From an engineering and historical point of view the *WBL* distribution has been one of the most recommended (Mathiesen et al 1994), used (Goda 2010) and accepted in several practical applications related to extreme waves (Goda et al. 1990, Goda 2010). From a theoretical and statistical viewpoint the extreme value (*EV*) theory provides appropriate recipes for developing the long term statistical analysis in each Earth science, including maritime and coastal fields.

Herein, some of the main *EV* theory concepts are going to be introduced, while a deeper review can be found, among others, in Kotz and Nadarajah (2000), Coles (2001), and Castillo (2004).
Since the pioneering work of Fisher and Tippet (1928), revised by Gnedenko (1943) and Gumbel (1958), it is known that if there exists a limiting distribution of the block maxima (typically with a yearly time scale) selected from an independent and identically distributed (*iid*) sample, this distribution belongs to the domain of attraction of the *GEV* family. Furthermore, under the same conceptual framework, the limit distribution of a sequence of excesses over a high enough threshold follows a corresponding approximate distribution within the *GP* family (Pickands 1975, Balkema and De Han 1974).

In addition, it can be shown that the *POT* approach may be used to indirectly fit the *GEV* distribution, with the aim of obtaining a more accurate extreme quantile estimates than fitting a *GEV* on annual maxima (Madsen et al, 1997a). The *GP-GEV* duality is consistent with the *PP* model representation (Davison and Smith 1990, Coles 2001, Katz et al. 2004), which, in turns, provides a formal theoretical justification to adopt the *POT* approach for fitting the *GEV* distribution; more concisely, a two-dimensional, inhomogeneous Poisson process is combined with

the dual-domain *POT* modelling. The *GP* distribution can therefore be used to fit the excesses over a threshold, while the Poisson distribution can be assumed to represent the occurrence time of exceedances. As a matter of fact, fitting *GP* on a reasonable number of exceedances over a proper threshold can give more reliable information concerning the whole phenomena than its *BM* counterpart and, consequently, a more representative insight into the maxima.

At this point, it seems useful to introduce another relevant issue within this context, that is the asymptotic behaviour required by the *EV* theory; on the whole, it can be assumed pertained to discriminate between the engineering perspective and the statistical viewpoint. It should be pointed out that inferences are critically dependent on sample size; unfortunately, small sample sizes are the routine in offshore and coastal engineering.

Goda (2011b) analysed sample size effect by using numerical simulations and comparison with real field data. First, it was emphasized by the author that the best-fitting distribution of a small size sample of yearly maxima (e.g. less than a few hundred) is not confined to *GEV* family, but scattered among the *WBL, GP* and *GEV* distributions. Second, an inherent negative bias of quantile estimates corresponding to return period longer than several dozen years and a tendency of increasing of the shape parameter for small sample size was underlined. Third, the superiority of the *GP-Poisson* model over both the *WBL-Poisson* model and *GEV* direct fit on annual maxima in case of large sample size of annual maxima (e.g. around 500) was confirmed. On balance, it might be said, as a crude obviousness, that the asymptotical behaviour is consistent with the large sample size only, while small sample size effects tend to hinder its reliability.

These results may be considered somewhat in common with other, as remarkably as questionable, aspects pointed out by some other authors (Castillo and Sanabria 1992, Jonathan and Ewans 2014), which emphasized the *GP* superiority. Actually, the *WBL* drawbacks and weakness are considered to be noticeably inferior with respect to the GP-flexibility. One of the driven reason to corroborate this thesis is inherent to the principle feature of the shape parameter, which governs the tail behaviour and the aptitude to originate shorter or broader, that is bounded or unbounded, or rather, using a common statistical terminology, light-tailed or heavy-tailed family distribution. This is, on the other hand, a straightforward means to recognize the domain of attraction for maxima of each distribution and, thus, to investigate the relative weight assigned to extreme values.

Specifically, Cunnane (1973) has shown that the *WBL* distribution becomes asymptotically of the Gumbel type (i.e. a *WBL* distribution with shape parameter equal to one). In other words, this implies that fitting using the *WBL* distribution implicitly constrains the solution to be exclusively unbounded on the right tail, whereas modelling using *GP* avoids this restriction. It is interesting to observe that this is one of the reason why fitting using *GP* and *WBL* are expected to give similar estimates for return values corresponding to return periods of the order of that of sample, but different estimates are likely to arise for return values especially at longer return periods. Most probably, this aspect addressed Goda et al. (2010) to strongly advise against of the *GP* distribution fit for coastal structure design; that is, quantile estimates would be greatly underestimated whenever *GP* distribution was used. Moreover, the same piece of advice was reported by Mazas and Hamm (2011) when the extreme analysis is applied to short dataset (say less than 10-20 years), such that

the *GP-Poisson* model is found to produce too small return values, in spite of a very good statistical fit.

On the other hand, for some authors (Jonathan and Ewans 2012, Mazas and Hamm 2011, Castillo and Sanabria 1992) unbounded and heavy tail distributions should be considered not appropriate for coastal engineering, whereas bounded and light tails should be more suitable. The reason is to be found in the fact that wave heights should be assumed bounded because of physical constraints on storm size, which, in turns, are determined by limitations on pressure and wind fields, storms intensity, as well as on water depth and fetch values. Nonetheless, they do not recommend the use of *GP* in the *Fréchet* domain, because unbounded, as a final suggestion.

Nevertheless, the above final assessment seems to be somewhat contrasting with the Goda (2010 pag 543) view, whose remarks of no physical ground in deep water lead to believe that extreme wave heights would not have an upper limit. Probably, it should also be considered contrasting with *GP-Poisson* model features, as well as with the aim of using the *POT* approach to obtain a better *GEV* estimates, so as to avoid less reliable quantile estimates than those provided with the single *GP* family. Moreover, they are contrasting with the Hosking and Wallis (1997) recommendation, which they do not give any priority to the asymptotic distributions, arguing that the number of storm events in a year is rarely large enough to justify the *EV* approximation.

Recapping, small sample size effect requires to focus the inference on a multi-distribution approach, without excluding any of the possible distributions, or, similarly, by not assuming a priori the *EV* family distributions only, because the asymptotical hypothesis may not completely be verified in practice.

On the whole, in this work the classical statistical approach of frequency analysis is followed, by using a multi-parent-distribution approach, as suggested by Goda (2010, 2011b), Mazas and Hamm (2011) and Hosking and Wallis 1997), among many others, because the need of taking into account for the small sample size effects.

1.1.2 De-clustering methodology

The *EV* theory is often based on sequences of independent and identically distributed (*iid*) random variables. However, in the Earth Science area, temporal independence is usually an unrealistic assumption. For instance, with particular reference to wave heights, it is widely known that waves travel in group and exhibit a tendency of being persistent for many hours; according to Goda (1979), an evident correlation between wave heights 24 hours apart is common in ocean waves. Moreover, the group of small wave heights is likely to constitute a population different from that of the group of large heights, thus violating the homogeneity assumption.

Although sometimes the independency is artificially assumed to be valid, it can also be relaxed by considering stationary series or, more easily, by utilising an adequate de-clustering method based on either physical or statistical assumptions. For example, a limiting condition is usually assumed in practical applications, where the events $X_i > u$ and $X_j > u$ (being u the threshold) can be considered approximately independent if u is high enough and the occurring times of i and j are far enough apart (Coles 2001, page 93).

In this work, the independency and homogeneity are determined on the basis of the sea-storm definition, according to the Wave Atlas (Corsini et al. 2004) and Boccotti (2000) recommendations. In addition, in order to enhance and improve the homogeneity, a combination between fetch distribution, sea storm polar plot, and wave climate characteristics, is considered so as to group data belonging to same meteorological events within homogeneous directional clusters (Corsini et al. 2004).

1.1.3 Threshold selection

Despite threshold selection may be defined one of the key aspects in each *POT* modelling, no final consensus has been achieved yet in the scientific community to identify a standard approach that may be clearly superior with respect to others. This is actually a still active research field. A first non-comprehensive review can be found in the work of Lang et al. (1999). Recently, Scarrot and MacDonals (2012) have proposed a wider overview of the so-called fixed-threshold approach, as well as about the mixture models implementation (e.g. Solari and Losada, 2012).

In general terms, a high enough threshold should be chosen, such that the problem, from a statistical perspective, can be seen as a classical balance between bias and variance. Generally, an extreme value model should be able to represent adequately the population tail, which is uniquely defined by the threshold; however, the threshold value must balance the bias arising from the asymptotic tail approximation, as well as the parameter estimation uncertainty due to the inherent scarcity of exceedances. Therefore, if the selected threshold is too low, bias causes distortions in the model because of violation on the model assumptions (e.g. values might not be independent or non-extreme threshold excess data might be included in the sample). If the threshold is too high, the variance is large because few data are included in the sample and, consequently, large differences in the quantile estimates might be obtained. Finally, the problem becomes further complicated whether other issues are included in the analysis; for instance, representation of the covariate relationship, uncertainty measurement related to the threshold choice, necessity of automated selection procedure, etc..

Commonly, a practical recommendation, having a general consensus, consists to set the threshold as low as possible in order to maximise the sample size, by using the diagnostic approach proposed by Coles 2001; although it is known that this classical procedure can neither be objective, being rather somewhat subjective, nor it is not uncommon the case where a threshold cannot be uniquely defined (see for example Figure 1 and the corresponding explanation reported by Scarrot and McDonalds 2012, paragraph 2.1). Basically, whenever the threshold is set as low as possible (at a reasonable value), the attention is focused on strengthening the variance at the expense of a higher bias. An equally valid alternative is, on the other hand, to favour bias by selecting a higher reasonable threshold which, in turn, can be more representative of extreme samples and, so, more tailored to the right tail. It must be further stressed that they are both consistent with the theory. In this work both alternatives were analysed, without assigning a priori a preference to one of them.

Operationally, a total of three threshold selection methods was chosen, by discriminating between some state-of-the-art fixed-threshold diagnostic procedures and a multiple-threshold-method (Deidda, 2010). First, the higher threshold is assumed to be represented by selecting a fixed number of sea storm per year (i.e. between two and five sea storms per year), as proposed by Mazas and

Hamm (2011), which might be seen as a rule of thumb, but it can also be assumed particularly tailored to the application of the *GP-Poisson* model. Second, the automated procedure proposed by Thompson et al. (2009), based on both the parameter-invariance characteristic of GP distribution and on the maximum likelihood property, has been considered. Finally, a multiple-threshold-method has been used and slightly adapted to our dataset. In this last case, in fact, only wave height within a significant directional sector should be chosen and, in addition, with the purpose of using a dataset as much as possible equal to the original method presented by the author, the independency should be slightly relaxed: that is, only daily maxima falling into the main directional sector were considered, at the expense of the de-clustering method used herein (i.e. the sea storm definition).

1.1.4 Parameter estimation and plotting position formula

Although one may argue that uncertainty associated with the choice of the parameters estimation method is less important than other uncertainty sources – such as threshold selection, independency, homogeneity –, it is also true that a certain degree of correlation exist between each of these elements. Moreover, the side effect of sample size is directly linked with the overall performance of each parameter estimation method. For instance, it is widely known the bad performance of the maximum likelihood (*ML*) method for small sample (e.g. $n \leq 25$, n being the sample size).

On the other hand, despite ML is the best unbiased, efficient and consistent estimator, other estimators can either denote comparable statistical performances or even outperform the *ML* results, for some of the properties above. Furthermore, some methods are usually computationally more tractable than the *ML* method as they require less frequent recourse to iterative procedures. Finally, neither the *GEV, GPD,* and *WBL* family appear to be insensitive to the parameter estimation method nor each parameter estimation method is completely valid on the whole domain of each cumulative distribution function adopted in this study.

Madsen et al (1997a) proposed a comparison between the L-moment (*LM*), method-of-moments (*MOM*), and probability-weighted-method (*PWM*), within the *AM-GEV* and *POT-GP* approach, with a total of six model combinations. It was advised against the use of the ML procedure for small sample size, while a selective applications of *MOM*, *ML* and *PWM* was recommended, depending on the value assumed by the shape parameter. Analogously, Akram and Hayat (2014) compared the small sample performance of several estimator – *LM*, *ML*, *MOM*, ordinary-least-squares (*OLS*), modified *ML*, modified *MOM*, and maximum-product-of-spacing (*MPS*) – of three-parameter *WBL* distribution. Albeit the best estimator is influenced by the shape parameter range, the *LM* method is found to be the lead in small samples as it was almost always close to the best method of estimation. These findings are also consistent with the numerical difficulties inherent in the *ML* method application within the three-parameter *WBL* distribution proven by other authors (Cousineau 2009 , Mazas et al. 2014).

Last but not least, it should be highlighted the importance of the plotting position formula. Unlike some authors severely criticized the plotting position usefulness (Makkonen 2005), in this work the view and recommendation of many others (Cunnane 1978, Goda 2010) were followed. The plotting-position rule is usually adopted to plot the sample on the probability diagram and is sometimes called the empirical distribution function. In addition, it is actively present in several best-fit model methods and, although not frequently done, it should be selectively assigned

depending on both the parameter estimation method and the cumulative distribution function, in order to ensure an appropriate balance with respect to other elements. With that purpose, useful choices might be addressed by considering unbiased plotting position formulas (Goda 2010, Goda 2011a).

To sum up, the parameter estimators are evaluated by considering the maximum likelihood method (Coles 2001), the L-moments method (Hosking and Wallis 1990, Goda 2010), the ordinary least-squares (*OLS*) method (Goda 2010), and the maximum product of spacing (*MPS*) method (Cheng and Amin 1983). In addition, the plotting position formulas considered in this study are: i) the Hazen formula (Hazen 1914) whether *ML* and *MPS* method are used, regardless the cumulative distribution function family; ii) the unbiased Weibull plotting position (Goda 1988) for the *OLS* method if referred to the *WBL* family; iii) the unbiased plotting position proposed by Goda (2011a) in case of the *LM* (and possibly *PWM*) method and for each cumulative distribution function family.

1.1.5 Best-fit model selection

It is important to check the adequacy of models upon which inferences are based. Models vary in the complexity and strengths of their assumptions, and model-checking needs vary correspondingly. Even when no covariate are present, rather large samples are often needed before the superiority of one model over another in terms of fit is indicated. In the simplest case a model may involve a single random variable, X, with cumulative distribution function, $F(x)$. The main problems are often to check whether some specific form, $F_o(x)$, is consistent with observed data, and whether assumptions about observations X_i being independent and identically distributed are satisfactory.

Informal methods of model checking emphasize graphical procedures such as probability and residual plots. These diagnostic tools are the probability-probability-plot (*PP-plot*) and quantile-quantile-plot (*QQ-plot*), as discussed and used in Coles (2001), or, for example, the probability plot correlation (*PPC*) method (Filliben 1973, 1975, Goda 2011a,b). In most cases, the variation inherent in graphical summaries is substantial, even when data are generated by the assumed model. Therefore, even if graphical methods and model expansion satisfy the majority of practical model-checking needs, other procedures may sometimes be more useful. For instance, formal tests are sometimes the only way to carry out the model checking.

In this work, both diagnostic, parametric and non-parametric best-fit metrics are used to assess the model performance. In particular, the so-called minimum residual of correlation coefficient (*MIR*) computed on both *pp-plot* and *qq-plot*, as well as the probability plot correlation coefficient (*PPC*) method, both proposed by Goda (2010), are devoted to synthesize the graphical (probabilistic) part of the model fit. On the other hand, the Akaike-information-criteria (*AIC*, Akaike 1973) and the Bayesian-information-criteria (*BIC*, Schwarz 1978) are used when the information matrix for maximum likelihood estimates can be computed (i.e. in case of using the *ML* and *MPS* estimator). In addition, the Monte Carlo based approach is used to compute confidence intervals, bias and root mean square error of parameter and quantile estimator, plus an improved version of the so-called error norm criteria, proposed by Li et al. (2008), which provides a weight factor to be used in order to emphasize the fit on the specific part of the distribution (i.e. the right tail if the analysis is mainly devoted to extreme events estimation). In this last case, it should also be stressed that the original version of the error norm criteria was improved, by assuming the empirical cumulative distribution

function equal to a consistent unbiased plotting position formula, Finally, if the *LM* estimator is considered, the diagram ratio plot is proved to be another useful diagnostic way to quickly analyse the adequateness of the fit (Hosking and Wallis, 1997).

1.2 Coastal numerical modelling overview

In the following paragraphs, some general characteristics of phase-averaged and phase-resolving models will be described, while the mathematic formulations may be found in standard textbooks, e.g. Komen et al. (1994), Vreugdenhil (1994), Dingemans (1997), Mei (2005), Svendsen (2007), Holthuijsen (2007). Moreover, some general reviews of wave propagation and surf-zone modelling in coastal engineering are described in in Svendsen and Putrevu (1996) and in Liu and Losada (2002), respectively, as well as a general historical review on the water wave theory is reported in Craik (2004).

1.2.1. Phase-averaged models

Spectral phase-averaged models are fundamentally based on describing the evolution of the energy density spectrum under the assumptions of zero-mean, quasi-homogeneous, quasi-stationary, Gaussian stochastic process possessing ergodicity (Komen et al. 1994, Holtuijsen 2007, Goda 2010). To a large extent, the wave field prediction simulated by phase-averaged models is assumed to be characterized by uniformly-distributed random phases, while the attention is addressed to simulating modifications of the energy density spectrum (Holtuijsen 2007).

A general spectral wave-model classification into first, second and third generation type can be represented by analysing the modelling of nonlinear wave-wave interactions. In first generation models, nonlinear energy transfers are treated implicitly through the wind-wave interaction and dissipation terms. Second generation models represent wave-wave interactions by means of parametric formulations, i.e. by applying a reference spectrum to redistribute the energy over the frequencies. Finally, third generation models are focused on modelling explicitly the nonlinear energy transfer, such that the spectral energy balance equation is solved without constraints on the shape evolution of the wave spectrum, although it is necessary for computational economy to make both analytic and numerical approximations.

Conventional third generation stochastic wave-models are based on the action balance equation, which is also known as the radiative transfer equation, the transport equation, the kinetic equation; also, the name Boltzmann equation is used when solely the quadruplets wave-wave interactions effect is considered (Komen et al. 1994). The equation describes propagation through a non-homogeneous medium (i.e. variable depth, currents, etc…), which conserves total wave action as an adiabatic invariant, balanced by a sum of source terms describing the generation, evolution and dissipation of the variance density.

Over the last two-three decades, numerous phase-averaged models have been developed, such as *WAM* (WAMDI group, 1988, Komen et al 1994), *TOMAWAC* (Benoit et al 1996), *WAVEWATCH III* (Tolman 2009), *SWAN* (Booij et al. 1999), and *MIKE 21-SW* (Sorensen et al 2004), among others. In this list, *SWAN* is particularly suitable to model near-shore areas, because it incorporates formulations for both deep water and shallow water zones, i.e. generation by wind, whitecapping, quadruplets and triad nonlinear wave-wave interactions, dissipation due to bottom friction, and

depth-induced wave breaking. A summary of the state-of the-art in third-generation wave modelling is proposed by The *WISE* Group (2007).

1.2.2. Phase-resolving models

Only a subset of the wide variety of phase-resolving models mentioned previously (i.e. see Introduction) is suitable for efficient use in practical applications, depending on the computational resources and CPU time needed to be run. A further problem can be represented by the amount of input/output data to manage. To date, an even smaller number of such models are available in commercial software packages, with graphical user interfaces (GUI) enabling a more efficient use.
Fundamentally, the following model taxonomy of phase-resolving models can be introduced:

- Non-hydrostatic models (Reynold-averaged Navier-Stokes equations, RANS, and related approximations)
  - 3D
  - 2D in the vertical plane (2DV);
  - 2D in the horizontal plane (2DH)
  - 1D
- Boussinesq type models (Boussinesq type equation, BTE)
  - 2D in the horizontal plane (2DH)
  - 1D
- Hydrostatic models (Non-Linear-Shallow-Water Equations, NLSWE)
  - 2DH
  - 1D

Obviously, three-dimensional models based on the RANS are expected to be the most performing Nonetheless, because only the simulation of inundated areas is of concern in coastal flooding problems, rather than the details of the velocity field in itself (which might be required and highly valuable in different types of problems), and because the computational burden required to run such models when applied to even a relatively narrow coastal area, approximations to the complete 3D equations according to the alternative formulations listed above are sought for. In particular, 2DH e 1DH formulations differ as to representation of dispersive and nonlinear properties, with the former related to non-hydrostatic pressure distribution.
In the following, strengths and weaknesses of the various type of most commonly used models – that is, those based on classical numerical integration methods as, fundamentally, finite differences, finite elements or finite volumes – are presented. Therefore, meshless methods as SPH (Monaghan 1992, Dalrymple et al. 2001, Gotoh et al. 2004, Shao 2010, Shao et al. 2006) or Lattice-Boltzmann methods are not analyzed in this work.

1.2.3. Non-hydrostatic models

Non-hydrostatic, 3D models, being based on the RANS equations, enable detailed representation of the velocity and pressure fields in turbulent, rotational flows in geometrically complex computational domains. The need for accurate representation of turbulence dynamics, particularly in the presence of wave breaking, requires representation of turbulence generation, transport and decay, thus requiring use of turbulence sub-models, mainly based on the k-ε formulation.

A crucial point in 3D models is related to the representation of the free surface. Models based on representation of the water elevation as a 1 to 1 function of the horizontal plane coordinates do not

allow for a detailed representation of plunging breakers, nor of the related air entrainment in the water mass. However, this can be accomplished via with proper free surface tracking techniques, such as the level-set method (Osher and Sethian 1988), the marker-and-cell (*MAC*) method (Harlow and Welch 1965), and the Volume Of Fluid (VOF, Hirt and Nichols 1981) technique, with this latter, that is based essentially on computing density variations in the cells of the computational mesh and subsequent reconstruction of the air-water interface, being the most currently used. Non-hydrostatic 3D models were used by Choi et al. (2007), Choi et al. (2008), Christensen (2006), Lynett e Liu (2005) for runup simulation on fixed bottom.

Models of this kind have been used for simulation of fluid interaction with a permeable structure, based on averaged equations on control volumes (VARANS, Lara et al. 2010) or on definition of porosity coefficients (Hur and Mizutani 2003, Hur et al 2004, Hur et al. 2008). Therefore, in principle, such models are able to take into account filtration in a permeable bottom when simulating runup on beaches.

Fully 3D models require using a large number of layers in order to provide correct representation of non-hydrostatic pressure distribution, laying in the range 20 to 120 in the cases examples above. The resulting huge dimensions of the computational problem require significant computing resources, preferably consisting of parallel machines or clusters of PCs, and large CPU times.

Depending on the specific problem at hand, 2DV non-hydrostatic models are often used (SKYLLA, Van Gent et al. 1994, Iwaa et al.1996; COBRAS, Lin and Liu 1998, Liu et al. 1999, Hsu et al. 2002; IH-2VOF, Losada et al. 2008, Lara et al. 2008, Guanche et al. 2009). In principle, such models are suitable to represent cylindrical waves only, as those generated in laboratory flumes, to which they are indeed typically applied.

Nonetheless, such models have been recently applied to simulate wave runup on natural beaches to produce flood hazard maps, e.g. Tomás et al. (2016) with application to the entire Spanish coastal perimeter. It is to be noted that the computational burden (10 to 40 hours of CPU time per hour of real time simulated on a single transect), has required to drastically reduce the number of simulations for such a large-scale application, resorting to a clustering procedure on dimensionless terrain profiles. The authors themselves indicate a limitation of the 2DV model to cases not including bays, estuaries or artificial basins. Moreover, it is plausible that the drawbacks of the 2DV formulation might be significant in the presence of bathymetries/topographies that are irregular in the longshore direction.

Aiming at increasing computational efficiency, 3D models have been proposed in the literature that are able to represent wave propagation with a reduced number of layers, using different and independent velocity profiles in each layer (Lynett 2002a,b, Hsiao et al. 2005) or a collocation of pressure in the mesh cell different than the usual, cell centered one (SWASH, Stelling and Zijlema 2003, Zijlema and Stelling 2005, 2008, Zijlema et al 2011b, Smit et al 2013). Such models assume a 1 to 1 water level-horizontal coordinate relation, do not use the VOF method and must therefore represent the shoreline advancement on a dry bed with a specific wetting & drying algorithm.

Whenever simulations aim at representing global features of wave propagation – namely, the extension of flooded areas – such models represent a valuable alternative to 3D models of the previous type. When run with a few layers, the dispersion properties of such models make them fundamentally equivalent to 2DH Boussinesq models.

2DH non-hydrostatic models are obtained as a special case of 3D models like SWASH, using only one layer. Beside SWASH, XBEACH is also based on the same 2DH formulation (Roelvink et al. 2009). Both models can be further used in 1D version.

Zijlema et al. (2011b) present accurate results obtained with SWASH using one layer for both schematic test cases with analytic solution and laboratory tests, with 1D or 2D geometry. In particular, 2D laboratory tests pertain to solitary wave and N-wave runup, these latter representing a typical initial condition for tsunami simulation. For the solitary wave, the accuracy of results slightly decays with increasing wave steepness. As to computational times, simulations of real time duration of 40 s on scale model on a 500×600 mesh required 1489 s with a single processor and 63 s on a cluster of 32 64-bit AMD processors (1.8 GHz, 4 MB L2 cache).

1.2.4. NLSWE models

When dealing with wave propagation in the nearshore area, as is the case for coastal flooding simulation, the concern is on the properties of long waves, that is waves with length L much larger than the water depth d ($\mu = d/L << 1$) – hence the equivalent term of shallow water waves. Models based on the potential flow hypothesis can be distinguished depending on the non-linearity parameter $\delta = H/d$, where H is the wave height. Following Svendsen (2007):

$$\frac{\delta}{\mu^2} = \frac{HL^2}{d^3} \begin{cases} \ll \mathcal{O}(1) & \text{linear shallow water waves} \\ = \mathcal{O}(1) & \text{Boussinesq waves} \\ \gg \mathcal{O}(1) & \text{nonlinear shallow water waves} \end{cases}$$

Models based on either the non-linear shallow water equations (*NLSWE*) or the Boussinesq-type equations (*BTE*) have been proposed in the literature for use in the nearshore area.

The *NLSWE* are mathematically equivalent to the Euler equations for compressible flows; they are strictly valid in shallow water because are non-dispersive models, or in other terms, are generally valid for the limit of nearly horizontal flows, where the vertical accelerations may be neglected and thus pressure distributions in verticals are hydrostatic (Falconer 1993). Nonetheless, near wave breaking, where the wave height H becomes of the same order as the water depth d, although the pressure distribution is locally highly non-hydrostatic, based on Svendsen (2007) classification, the *NLSWE* represent the long wave theory most suitable to describe the wave conditions. In particular, in the inner surf zone the waves tend to display a steep front similar to those in a moving hydraulic jump, which is consistently represented by the discontinuity in the *NLSWE* solution associated to the development of a shock wave, due to the hyperbolic property of the equations.

In fact the *NLSWE* are widely used in commercial codes due to their reliability, robustness, accuracy and proven effectiveness, but mainly for simulation of rapidly varied purely translation

waves, as those produced by, e.g., dam break. Models based on the *NLSWE* yield accurate reproduction of broken wave propagation and runup, including possible further inland propagation. Energy dissipation associated to wave breaking need not be simulated with ad hoc models, it being inherently represented in the solution by virtue of momentum conservation across the moving shock (in fact the energy dissipation in a bore represents the base for breaking models used in spectral models, e.g. Battjes & Janssen 1978). Godunov type, shock capturing methods are mainly used for the numerical solution of the *NLSWE* (Toro 2001, LeVeque 2002), based on exact or approximate solution of the Riemann problem for systems of conservation laws, in conjunction with flux- or slope-limiters near steep fronts to suppress spurious oscillations.

NLSWE models turn out to be computationally efficient and cheap when simulating a train of a large number of waves, which is needed to a reliable computation of the statistical indices, namely the usual 2% runup, and may represent at least a reasonable practical alternative to more complex models.

The capability of *NLSWE* models to simulate wave runup and overtopping of structures was first assessed by Kobayashi & Wurjanto (1989), Kobayashi & Wurjanto (1992) and Kobayashi & Raichle (1994) with the *RBREAK* model. That model widely underestimated the water depth above the top of the structure as well as the overtopping flow rate, as compared with laboratory results by Saville (1955). For the same test, Dodd (1998) also obtained markedly underestimated results as to the flow rate, particularly for steep waves. Inaccuracies on the wave setup were ascribed to poor representation of complex nonlinear interactions in the surf area. More recently, Kobayashi et al. (2010) have improved the model introducing a probabilistic representation of the *NLSWE*.

Titov and Sinolakis (1995) proposed a variable-space-step *NLSWE* model for simulation of solitary waves on beaches. The model yields results from good to excellent for both breaking and non-breaking waves, when compared to analytic solutions and laboratory tests, with exception for the case of plunging breaker, which is shifted offshore. Moreover, maximum runup is slightly underestimated for breaking waves. Results are equivalent to those of obtained by Zelt (1991) with a *BTE* model for non-breaking waves and slightly smaller for breaking waves, but at a remarkably smaller computational cost.

Verification of the *AMAZON* model (Hu et al. 2000) against analytical solutions and laboratory data of structure overtopping for both monochromatic and random waves showed inaccuracies compared to measurements, particularly for random waves. Moreover, sensitivity of results to specification of bottom friction was analyzed.
Puleo et al. (2002) compare results from 2DV e *NLSWE* (*Rbreak2*) models. Results for the water levels in the surf area are equally good for the two models, although with substantially different computational efforts, whereas larger differences can be observed in the velocity values, as could be expected.

Wave runup is the result of contributions from frequencies in the range of incident wave motion and in the infragravity range, which are generated in the surf area, with these latter being predominant in dissipative beaches, characterized by small values of the Iribarren number. The capability of the *NLSWE* (*Rbreak*) model of reproducing satisfactorily infragravity waves and energy decay at the

swell frequencies due to breaking was demonstrated by Raubenheimer et al. (1996) and Raubenheimer (2001). Larger discrepancies were observed between simulated and computed velocities.

Soldini et al. (2013) used Brocchini et al. (2001) *NLSWE* model to analyze the effect of different representation of the bathymetric/topographic profiles on computed runup of both random waves and wave packets associated to infragravity waves. It is shown that equivalent results are obtained assuming a natural profile, a theoretical equilibrium profile (Dean 1990), or an even simpler profile composed of two linear segments with different slopes in the submerged and emerging beach.

2DH models based on the *NLSWE* were proposed by Brocchini et al. (2001). Liu et al. (1995) and Hubbard & Dodd (2002). Liu et al. (1995) test their model on the laboratory test presented by Briggs et al. (1995) for wave runup of a solitary wave, obtaining good results for the measured levels and, in particular, for the maximum runup of nonbreaking wave.

The adaptive *OTT-2D* model by Hubbard & Dodd (2002) yields good results for the most demanding of the test cases presented by Thacker (1981), with asymmetric level and velocity distributions, and very good results for the Briggs et al. (1995) test case as to water levels and maximum runup and for a test case with non-orthogonal incidence on a structure slope, with exception for the overtopping flow rates. The use of adaptive mesh reduces the CPU time in a ratio 5:1 to 10:1 compared to non-adaptive meshes.

A general review of strengths and limits of the *NLSWE* in reproducing coastal flows can be found in Brocchini and Dodd (2008). The authors discuss the relations between physical phenomena to be represented, equations, numerical schemes and most representative test cases to be used for model validation.

As mentioned before, the suitability of *NLSWE* models for simulation of surf zone hydrodynamics and inland flooding suggests coupling with a spectral model for wave propagation from offshore. McCabe et al. (2011) have recognized the switching from the spectral model to the *NLSWE* model to be optimal for $H_{m0}/d \cong 0.65$, where H_{m0} is the spectral significant wave height, upon comparison of computed $R_{2\%}$ setup with the laboratory measurements by Mase (1989). Moreover, McCabe & Stansby (2010) obtained better results using a coupled spectral-*NLSWE* model than using a coupled Boussinesq-*NLSWE*.

1.2.5. Boussinesq models

According to Svendsen (2007) classification, the Boussinesq-type equations (*BTE*) represent, in their original formulation, models for weakly non-linear long waves (Peregrine 1967, Benjamin 1972). The *BTE* are deduced through a Taylor series development along the vertical of the velocity potential or of the horizontal component of velocity at a specific elevation from the bottom, and order of magnitude analysis of the resulting terms. In the classical *BTE* theory the curvature of streamlines in the vertical plane is described through a vertical velocity, the magnitude of which increases linearly from zero at the bed to a maximum at the free surface. More generally, the velocity profile is approximated by a polynomial expansion as a function of the vertical coordinate, which at lower order has a parabolic form. In this theory, therefore, the pressure is not longer

hydrostatic, but the vertical component of motion can be integrated out of the equations of motion to reduce the three-dimensional description to a two-dimensional one.

Even in low-order formulations, the *BTEs* are characterized by the presence of derivatives of order larger than two, resulting in a specific difficulty in their numerical solution related to both the discretization itself or such high-order derivatives and the need for high order discretization of low-order derivatives, so as to have the high-order derivatives appearing in the expression of the truncation error to vanish.

Since the pioneering work of Boussinesq (1872), much effort has been addressed on *BTE* improvement. After Peregrine's (1967) resuming to low-dispersive *BTEs* valid in uneven bottom, a huge number of *BTE* models have been defined so far, with the aim of enhancing both their dispersive and nonlinear characteristics, related to the μ and δ parameters, respectively. In-depth historical and relatively recent reviews may be found in Dingemans (1997), Kirby (1997, 2003), Madsen and Schäffer (1998, 1999), Madsen and Fuhrman (2010), and Brocchini (2013).

Main contributions to the improvements of dispersive and non-linear properties are due to Witting (1984), Madsen et al. (1991), Madsen and Sørensen (1992), Nwogu (1993), Schäffer and Madsen (1995), Madsen and Schäffer (1998, 1999), Green and Nahgdi (1976), Wei et al. (1995), Gobbi and Kirby (1999), Gobbi et al. (1999, 2000), Lynett et al. (2002), Madsen et al. (2003), Musumeci et al. (2005), Madsen et al. (2006), Chen (2006), Chazel et al. (2010), Bonneton et al. (2011), Zhang et al. (2013), Donahue et al. (2015, 2016), Beji and Nadaoka (1996), Kirby et al. (1997), Chen et al. (2001), Lynett and Liu (2004), Madsen et al.(2006), Shi et al.(2012a, b).

The derivation of the BTE through series development has naturally led to seeking for better properties increasing the order of formal accuracy of the equations, leading to increasing the number of equations and the maximum order of the derivatives therein (Agnon et al. 1999, Madsen et al. 2002,2003). One of the main downside to the *BTE* models is the generally need of prescribing a separate wetting and drying and moving shoreline algorithm (i.e. Madsen et al. 1997, Kennedy et al 2000, Brocchini et al 2002), as well as an ad hoc breaking model, the onset of breaking being represented via limiting values of the wave steepness, of the *H/d* ratio or of the vertical velocity at the free surface (Karambas and Koutitas 1992, Schäffer et al. 1993, Madsen et al 1997c, 1997d, Kennedy et al 2000, D'Alessandro and Tomasicchio, 2008, Tonelli and Petti 2009, 2010, among others).

For simulating broken wave propagation and related turbulence role, models based on turbulent viscosity (Zelt 1991; Karambas e Koutitas 1992; Wei et al 1995; Kennedy et al. 2000), surface roller (Brocchini et al. 2002, Schäffer et al. 1993, Madsen et al. 1997a,b, Sørensen et al. 1998) and on vorticity models (Veeramony e Svendsen 2000) have been proposed.

Hybrid models have been formulated more recently, based on low-order BTE up to the onset of breaking and on the NLSWE shoreward, to take advantage of the mentioned properties of these latter in representing broken waves and propagation over dry bed (Bonneton et al. 2011, Fang et al. 2014, Ha & Cho 2015, Orszaghova et al. 2012, 2014, Shi et al. 2012a, 2012b, 2013, Tehranirad et

al. 2011, Tonelli & Petti 2009, 2010, 2012, 2013, McCabe et al. 2013, Erduran 2005, 2007, Kazolea & Delys 2013).

Nevertheless, general achievements in *BTE* modelling of wave transformation have been nowadays widely recognized, since *BTE* research have undergone rapid developments, i.e. focusing on wave-current interaction (Chen et al. 1998), wave run-up (Fuhrman and Madsen 2008), tsunami simulation (Kirby et al. 2013), etc...

As per the BTE applied to wave runup simulation, computational results have been presented by Donahue et al. (2016), Dutykh et al. (2011), Furham & Madsen (2008), Geist et al. (2009), Liang et al. (2013), Lo Re et al. (2012, 2014), Lynett et al. (2002), Madsen et al. (1997c, 1997d), McCabe & Stansby (2010), McCabe et al. (2013), Park & Cox (2015), Roeber et al. (2010), Tsung et al. (2012), Zelt (1991), Zhang et al. (2014), Kennedy et al. (2000). Although some of the above models enjoy the most advanced dispersive and non-linear properties, ensuring very good results when compared to solutions in closed form and laboratory measurements, they suffer from the same limitations presented for the 2DV models, or even worse due to the only approximate non-hydrostatic behavior, when applied to transects extracted from markedly irregular topographies alongshore, as also pointed out by Geist et al. (2009).

2DH modelling with the BTE suffers from some more problems in the computation of propagation over dry bed. The fully nonlinear COULWAVE (Lynett et al. 2002) model suffers from instability in the test by Thacker (1981) for free oscillations of long waves; however, it yields good results for the laboratory test by Briggs et al. (1995) of solitary wave propagation and runup over a conical island. The fully non-linear model by Furham & Madsen (2008) is also unstable for the Thacker (1981) test, whilst performs well for the Briggs et al. (1995) test. Kazolea et al. (2012) obtain a good reproduction of this latter test with a weakly non-linear and weakly dispersive formulation on unstructured mesh. For the same test, Wang et al. (2006) and Zhang et al. (2016) do not show comparisons as to the extent of the flooded areas computed with, respectively, meshless and rotational Boussinesq-Green-Naghdi models.

It is to be noted, however, that tests with solitary waves can be considered representative (and not even completely representative) of tsunamis waves rather than wind waves, and that the Briggs et al. (1995) considered a relatively regular bathymetry/topography. Application of COULWAVE to simulate a natural event with irregular bottom (Cheung et al. 2003) has provided significant discrepancies between computed and measured limits of the flooded areas. The same model has been applied by Geist et al. (2009) to simulation of a landslide-generated tsunami, but only in its weakly nonlinear version; a simulation with the fully nonlinear 1D version with higher resolution revealed significant differences with the results of 2D simulations. However, no comparison with field data is presented. Results by Liu et al. (2010) with the meshless model by Wang & Liu (2006) in the simulation of a real tsunami in the Indian Ocean display a bad reproduction of measured water levels in the sea and a noticeable overestimation of flooded areas.

As to CPU times required for simulations on the Briggs et al. (1995), 20 s of real time simulation by Furham e Madsen (2008) on a 234×201 mesh with 0.15×0.15m squared cells took 11800 s (3.3 h) with a single 3.2 GHz Pentium 4 processor; 18 s of real time simulation by Kazolea et al. (2012) on

unstructured mesh with 52191 nodes, with cell sides in the range 0.07 m to 0.2 m, required 1980 s (33 min) on a single core of a single 2.4 GHz Intel Core 2 Quad Q6600 processor.

1.3 Discussion of BTE and NHE models applied to coastal flooding problems

From the consideration presented above, at the moment it seems that fully 3D models are unsuitable to be considered for application to real case studies due to the huge amount of computational resources and CPU time needed to run them. 2DV non-hydrostatic models might be an effective alternative for use in runup simulation on very regular bathymetries/topographies, due to their capability of inherently representing wave breaking with the VOF technique. Nonetheless they still remain very time consuming tools, which are not suitable for application to topographies that are irregular in the longshore direction – which is not at all unusual.
Therefore, one naturally looks at Boussinesq models and non-hydrostatic models using a small number of layers (whose equations are identified as NHE hereafter) as a suitable alternative for practical applications, at least for the case of narrow bays or artificial basins but also with an even wider scope. One of the fundamental objective of both model types, in the last two-three decades, has been to improve and to extend the phase-resolving modelling in (relatively) deep water, or, in other terms, to represent adequately a larger range, as well as better interactions of frequencies inside the wave spectrum. This is accomplished by increasing the order of derivatives in BTE models, while improvements in NHE models are obtained by (moderately) increasing the number of layers (Stelling and Zijlema 2003). However, neither the BTE nor the NHE are suitable for use up to the offshore for efficient computations, which calls for coupling one of the two kind of models with a spectral model cover the gap from the offshore to the nearshore, providing boundary conditions to the phase-resolving model. In addition, 2DH Boussinesq models suffer from problems in simulation of wave propagation over a dry bed, so that one has to resort to NLSWE in the nearshore, whereas NHE models are accurate and efficient in this respect.
The above considerations led us to formulate as an operational tool a model coupling a spectral model from the offshore to the outer edge of the nearshore and a NHE model in the nearshore.

In fact, *NHE* wave-flow models are widely accepted as equally valid and competitive alternative to *BTE* modelling (i.e., Zijlema et al 2011b, Ma et al 2012, Cui et al. 2012, 2014, and references quoted therein). As stated by Zijlema et al. (2011a) *NHE* models may be identified as a compromise between the capabilities of *BTE* modelling and operational-based requirements for numerical robustness, simplicity, ease of use and economy. A review of non-hydrostatic computing can be found in Zijlema and Stelling (2005). Stelling and Zijlema (2003) showed that the *NHE* models can be more efficient by representing the vertical domain with an edge based finite difference scheme. Additionally, Zijlema and Stelling (2008) showed that wave breaking can be accurately modelled by *NHE* model without using a separate breaking model. Moreover, Smit et al (2013) have further extended the efficiency of the breaking model by reducing locally the non-hydrostatic equations to the hydrostatic ones, once the rate of change of the free-surface exceeds a pre-determined threshold; hence, by mimicing the shock-capturing characteristic of the *NLSWE*, according to the hybrid *BTE-NLSWE* approach, and implementing a wave breaking initiation criterion similar to the criterion used in Kennedy et al. (2000).

2 Extreme value Theory

Following Goda (2011a), *GEV*, *GP* and *WBL* distribution family have been assumed to be representative for all possible marginal parent distributions within the statistical inference framework regarding coastal engineering studies.
Formally, the *GEV* and *WBL* distributions have three parameters of shape (denoted in this study by *k*), scale (*A*) and location (*B*), whereas the GP distribution must be considered a two-parameters distribution family. Although the GP distribution can be expressed as a function of the same three parameters (i.e. when L-moments method is used), strictly speaking only two of them are independent. As a matter of fact, in the so-called fixed-threshold approach, it is required to identify the exceedances by choosing an appropriate threshold *u*, which is not strictly properly a parameter of the distribution and cannot be considered a true distribution parameter.
In this chapter, the basic characteristics of *BM* and *POT* approach are presented, by introducing the cumulative distribution function, *F(x)*, the quantile function expressed in term of reduced variate, y_T, and the respective return value, x_T, for associated return period, *T*, for each of *GEV*, *GP* and *WBL* family distribution, as well as for the *GP-Poisson* model, and the *GP-MTM* model. Formally, two equivalent expressions of *GEV* and *GP* family may be considered, by changing the sign of the shape parameter; herein it is chosen to adopt the Hosking and Wallis (1997), Madsen et al. (1997a), Goda (2011 a, b) convention, among others.
Secondly, the fixed-threshold methods used in this study are discussed, namely the automated detection method of Thompson (2009) and the empirical approach based on the fixed mean rate value proposed by Mazas and Hamm (2011). In addition, the most known fixed-threshold procedures are included, i.e. the mean-residual-life and the diagnostic method based on the threshold stability property, both proposed by Coles (2001). Finally, the main characteristics of the multiple-threshold-method proposed by Deidda (2010) are reported.

2.1.Basics of Extreme Value theory

The extreme value theory (EV) provides similarities with the central limit theorem. Classical *EV* theory is concerned substantially with distributional properties of the maximum

$$M_n = max(X_1, \dots, X_n) \tag{2.1}$$

of *n* independent and identically distributed (*iid*) random variables, $\{X\}_{i=1}^{n}$, as *n* becomes large.

The *Extremal Types Theorem* (Coles 2001- Theorem 3.1) defines the possible limiting forms for the distribution of $M_n$under linear normalizations. Formally, one shall be concerned with conditions under which, for a suitable sequence of normalizing constants $\{a_n > 0\}$ and $\{b_n\}$,

$$Pr[a_n(M_n - b_n)] \xrightarrow{w} G(x). \tag{2.2}$$

Basically, it is stated that if for some sequences of normalizing constants $\{a_n > 0\}$ and $\{b_n\}$, the quantity $a_n(M_n - b_n)$ has a non-degenerate limiting distribution function $G(x)$, (e.g. for every max-stable distribution $G(x)$), then G must have one of just three possible forms – the well-known three extreme value distributions, *FT-I*, *FT-II*, and, *FT-III* (Fisher and Tippet 1928, Gnedenko 1942).

It can be proven (Jenkinson, 1955, Coles 2001 - Theorem 3.1.1) that the *GEV* distribution subsumes each of these three possible types; in fact, FT-I, FT-II, and FT-III are respectively equivalent to *GEV* distribution with shape parameter $k = 0$, $k < 0$, $k > 0$.

On the other hand, if the *BM* has approximate distribution *GEV*, then the threshold excesses have a corresponding approximate distribution within the *GP* family (see for example Coles 2001 - Theorem 4.1). Moreover, this duality *GP-GEV* may be stated by the parameters correlation, due to the fact that the *GP* parameters (A_{GP}, k, u) are uniquely determined by those of the associated *GEV* distribution of *BM* (A_{GEV}, k, B); as a matter of fact, the shape parameter is the same, whereas the scale parameters of both distributions are related by $A_{GP} = A_{GEV} + k(u - B)$.

2.2 The generalized extreme value (GEV) distribution

The *GEV* distribution was first introduced by Jenkinson (1955). The cumulative distribution function of the GEV is given by

$$GEV(x; k, A, B) = \begin{cases} exp\left[-exp\left(-\frac{x-B}{A}\right)\right] & k = 0 \\ exp\left[-\left(1 - k\frac{x-B}{A}\right)^{1/k}\right] & k \neq 0 \end{cases} \tag{2.3}$$

where the support is $x \leq B + A/k$, if $k > 0$, or $x \geq B + A/k$, if $k < 0$, whereas $-\infty < x < \infty$ results when $k = 0$. The shape parameter may be used to model a wide range of tail behaviours. The case $k = 0$ is that of an exponentially decreasing tail, while $k < 0$ is that of a polynomially decreasing tail function, corresponding therefore to a long-tailed parent distribution; finally, $k > 0$ is the case of a finite upper endpoint and hence is short-tailed.
It should be noted that there exists a number of non-regular situations associated with *k*, for instance the mean only exists for $k > -1$, while the variance for $k > -0.5$ and there may be problems when the second and higher moments do not exist to obtain the likelihood estimator. Nonetheless, the experience with real-world data suggests that the condition $-0.5 < k < 0.5$ is almost always satisfied in practical applications. Some of these problems may be circumvented by adopting other estimators and estimation methods different from the *ML* (see for example Hosking et al. 1985, Madsen et al 1997).

The *GEV* combines three simpler distributions into a single form, allowing a continuous range of possible shapes that includes all three of the simpler distribution. The three cases covered by the *GEV* distribution are often referred as the *Types I*, *II*, and *III*. Each type corresponds to the limiting distribution of block maxima from a different class of underlying distributions. Distributions whose tails decrease exponentially lead to the *Type I*; distributions whose tails decrease as a polynomial lead to the *Type II*; distributions whose tails are finite lead to *Type III*.

Types I, II, and *III* are sometimes also referred to as the *Gumbel*, *Fréchet*, and *Weibull* types, though this terminology can be slightly confusing. The *Type I* (*Gumbel*) and *Type III* (*Weibull*) cases actually correspond to the mirror images of the usual *Gumbel* and *WBL* distributions, respectively

reported in equations (2.6) and (2.7). Finally, the *Type II* (*Fréchet*) case is equivalent to taking the reciprocal of values from a standard *Weibull* distribution.

In terms of reduced variate, the quantile function is given by

$$y_p = \begin{cases} -ln[-lnP] & k = 0 \\ \frac{1}{k}[1-(-lnP)^k] & k \neq 0. \end{cases} \tag{2.4}$$

P being the non-exceedance probability.
The corresponding return value for an associate return period *T* is expressed by

$$x_T = \begin{cases} B - Aln\left[-ln\left(1-\frac{1}{T}\right)\right] & k = 0 \\ B + \frac{A}{k}\left\{1-\left[-ln\left(1-\frac{1}{T}\right)\right]^k\right\} & k \neq 0 \end{cases} \tag{2.5}$$

2.3 The Weibull (WBL) distribution

The *Weibull* distribution was first published in 1939 (Weibull 1939) for description of breaking strength of materials. It has been widely adopted in extreme waves analysis in many occasions (Mathiesen et al, 1994, Goda 2010).
The cumulative distribution function of the *WBL* distribution (Weibull 1951) is given by

$$WBL(x;k,A,B) = 1 - exp\left[-\left(\frac{x-B}{A}\right)^k\right] : \quad B < x < \infty. \tag{2.6}$$

It can be considered either a lower bounded or an upper unbounded distribution and it can be skewed either positively or negatively. The shape parameter governs the tail behaviour: in particular, the distribution becomes broader with decreasing *k*, with an attitude to generate heavier extremes. The case whit $k = 2$ demonstrates behaviour similar to the *Lognormal* and *Pearson-III* distributions, whereas the case with $k = 1$ is the *Exponential* distribution. Moreover, with this value of the shape parameter, it can be shown (Cunnane 1973) that the *WBL* distribution becomes asymptotically the *Gumbel* distribution

$$Gum(x;A,B) = exp\left\{-exp\left[-\frac{x-B^*}{A}\right]\right\}, \tag{2.7}$$

where

$$B^* = B + A\ln\lambda\,. \tag{2.8}$$

and λ is the mean rate of occurrences, counted as the averaged number of events per year, and is assumed to represent the Poisson-parameter.

In addition, the *WBL* family is related to the *GEV* family, because it can be considered a reverse *GEV* distribution with parameters

$$\zeta = 1/k \quad \eta = A/k \quad \delta = B - A. \tag{2.9}$$

Schematically

$$GEV(x;\zeta,\eta,\delta) = 1 - WBL(-x;k,A,B). \tag{2.10}$$

In terms of reduced variate, the quantile function is expressed by

$$y_p = F^{-1}(P) = \left[-ln(1-P)^{1/k}\right]. \tag{2.11}$$

Therefore, the return value associated with the return period *T* is equal to

$$x_T = F^{-1}\left(1 - \frac{1}{T}\right) = B + Aln(T)^{\frac{1}{k}}. \tag{2.12}$$

A detailed overview of the parameter estimation methods used for the *WBL* distribution may be found in Akram and Hayat (2014). Although the *ML* estimator is usually considered the best statistical estimator, in terms of unbiasedness, effectiveness and consistency, the authors have demonstrated that other estimators can be considered either adequately representative or even better than the maximum likelihood method. As a matter of fact, both for analytical, numerical and statistical reasons the *LM* method tends to outperform the *ML* estimator. The *LM* estimator for the *WBL* family are reported in Goda et al (2009), which added these formulations to the FORTRAN program "*lmoments*" proposed by Hosking and Wallis (1990), following the book of the same authors (Hosking and Wallis 1997). In addition, it may be demonstrated how the *MPS* method is further better than the *ML* estimator, because it has a wider range of validity. Finally, the Ordinary-Least-Square (*OLS*) method defined by Goda (2010) is found to have same unbiasedness of the *ML* method, when its use is associated with an unbiased plotting-position formula (Goda et al. 1988). Moreover, an Extended-Least-Square (*ELS*) method, reported in Goda (2010) as well, is found to be comparable in terms of effectiveness with respect to the *ML* method.

To sum up, for the *WBL* distribution several parameter estimation methods exists; hence, if the sample size effect has a great impact on the analysis, it will be better to carry out a comparison between their performances, without choosing a priori one of them.

2.4 The generalized Pareto (GP) distribution

The cumulative distribution function of the generalized Pareto (*GP*) distribution, introduced by Pickands (1975), is given by

$$GP(x;k,A,u) = \begin{cases} 1 - exp\left[-\frac{x-u}{A}\right] & k = 0, \xi = 0 \\ 1 - \left[1 - k\frac{x-u}{A}\right]^{1/k} = 1 - \left[1 + \xi\frac{x-u}{\alpha}\right]^{-1/\xi} & k \neq 0, \xi \neq 0 \end{cases} \tag{2.13}$$

where the support for x is $0 \leq x < \infty$ for $k \leq 0$ and $0 \leq x \leq u + A/k$ for $k > 0$. It must be stressed that the *GP* is a two-parameter distribution – being *A* the scale parameter, while *k* the shape parameter –, because the threshold value, *u*, has to be specified before the fitting (e.g., the fix-threshold approach). Moreover, even in the case of three-parameter-*GP* distribution, when a location parameter *B* is added to the usual aforementioned form (e.g. as proposed by Hosking and Wallis (1997) and largely adopted by Goda (2011a, b) or Mazas et al (2014)), it should still be considered as a two-parameter distribution, only two of them being independent from time to time.
Secondly, as previously denoted, it should be paid attention to the dual convention usually considered in literature, which is related to the shape parameter sign ($k = -\xi$), leading to two equivalent formulations. In the present work, the *k* parameter sign and, hence, the corresponding formulation, will be generally adopted, according, among others, to Goda (2011b).
The shape parameter *k* controls the tail behaviour of the distribution, hence the aptitude to originate heavier or weaker (i.e. flatter) tails or larger or smaller quantiles estimates. If $k < 0$ the GP distribution is unbounded and heavy-tailed and it leads to the so-called Pareto distributions, whereas it is bounded and light-tailed if $k > 0$, with an upper bound value equal to $(u + A/k)$. When $k = 0$ the distribution has the ordinary exponential form with mean *A*, while the case with $k = 1$ yields to the uniform distribution on $[u, u + A]$.

Two important features are relevant to the analysis of extreme events. First, a truncated *GP* distribution remains a *GP* distribution; that is, if the *G*P distribution is assumed to be applied for the threshold level u^*, then for a higher threshold, $u_1 > u^*$, the distribution of the exceedance given $X > u_1$ is

$$F(x|u > u_1) = \frac{F(x)-F(u_1)}{1-F(u_1)} = \begin{cases} 1 - exp\left[-\frac{x-u_1}{A^*}\right] & k = 0 \\ 1 - \left[1 - k\frac{x-u_1}{A^*-k(u_1-u^*)}\right]^{1/k} & k \neq 0 \end{cases} \quad (2.14)$$

which is a *GP* distribution with the same shape parameter *k* and a different scale parameter equal to $A^* - k(u_1 - u^*)$, if $k \neq 0$.

Second, if a set of *n* independent and identically distributed random variables, $\{X_i\}_{i=1}^n \sim GP(x; k, A)$,is considered, where *n* has a *Poisson* distribution with mean λ, than $Z = max\{X_i\}_{i=1}^n \sim GEV(z; k, A^*, B^*)$; that is the extreme events follow a *GEV* distribution, with the same shape parameter and with scale and location parameters linked by theoretical relationship (e.g. see §2.6). Actually, this is a fundamental connection between the *GP-Poisson* model and the *GEV* family, which may also be referred to as the *GP-GEV* model relationship.

In terms of reduced variate, the *P-quantile* of the *GP* is defined as

$$y_p = \begin{cases} -\ln(1-P) & if \quad k = 0 \\ \frac{1}{k}[1-(1-P)^k] & if \quad k \neq 0 \end{cases} \quad (2.15)$$

The return value corresponding to the return period *T* is defined as

$$x_T = F^{-1}\left(1-\frac{1}{T}\right) = \begin{cases} u - A\ln\left[1-\left(1-\frac{1}{T}\right)\right] & k = 0 \\ u + \frac{A}{k}\left\{\left[1-\left(1-\frac{1}{T}\right)\right]^k - 1\right\} & k \neq 0 \end{cases} \tag{2.16}$$

As mentioned before, only two parameters need to be estimated. Estimation of the parameters of the *GP* distribution is considered by Hosking and Wallis (1997, 1990, 1987, 1985). The *ML* estimation is determined by Grimshaw (2003), whereas the *MPS* estimator is defined by Cheng and Amin (1983).

2.5 The GP-Poisson model

Following Madsen et al. (1997) the relation between the *GP-Poisson* model and the *GEV* family will be derived. The starting point is represented by the relationship between the cumulative distribution function of *POT* data, *F(x)*, assumed both *GP*- and *Poisson*-distributed, and the cumulative distribution function of annual maxima, *G(x)*. Specifically, the *Poisson* distribution describes the probability of the event that occurs *t* times in a given time duration, *K*, when the mean rate of occurrence is λ:

$$P(t) = \frac{\lambda^t}{t!}exp(-\lambda), \quad t = 0, 1, 2, \cdots. \tag{2.17}$$

The probability that X_{max} among *t POT* data in a year does not exceed the value *x* is given by

$$G(x) = Pr[X_{max} < x] = \sum_{t=0}^{\infty} P(t)[F(x)]^t = exp\{-\lambda[1-F(x)]\}, \tag{2.18}$$

where $Pr[\mathcal{A}]$ denotes the probability of the event $\mathcal{A}$.
By substituting equations (2.17) and (2.13) into equation (2.18), Madsen et al. (1997a) proposed the cumulative distribution function of the *AM* data derived from the parent of *GP* distribution, showing its duality with the *GEV* distribution

$$G(x;k,A,B) = \begin{cases} exp\left[-exp\left(-\frac{x-B^*}{A^*}\right)\right] & k = 0 \\ exp\left[-\left(1-k\frac{x-B^*}{A^*}\right)^{1/k}\right] & k \neq 0 \end{cases}, \tag{2.19}$$

where the scale parameter of the *GEV* distribution is given by

$$A^* = A\lambda^{-k}, \tag{2.20}$$

while the location parameter B^* of the *GEV* distribution is equal to

$$B^* = \begin{cases} u + A^* \ln \lambda & k = 0 \\ u + \frac{A^*}{k}(1 - \lambda^{-k}) & k \neq 0 \end{cases}. \tag{2.21}$$

It is interesting to observe that this expression is equivalent to the *GP-GEV* relation stated by Coles (2001, pg. 75, eq. 4.3) and presented before (§2.2). In fact, by considering the sign convention, $(k = -\xi)$, and rearranging the terms in the above equation, it may be shown that

$$A^* = A + k(u - B^*) \Leftrightarrow \sigma_{GEV} = \sigma_{GP} - \xi(u - \mu), \tag{2.22}$$

where it is highlighted the equivalence between *GP* and *GEV* distributions.

Finally, the *T*-year event based on the *AM* data is defined as the return value

$$x_T = \begin{cases} B^* - A^* \ln\left[-ln\left(1 - \frac{1}{T}\right)\right] & k = 0 \\ B^* + \frac{A^*}{k}\left\{1 - \left[-\ln\left(1 - \frac{1}{T}\right)\right]^k\right\} & k \neq 0 \end{cases}. \tag{2.23}$$

2.6 The fixed-threshold methods

Several concerns have been raised about the threshold selection in the *GP-POT* modelling (i.e. Scarrot and MacDonald 2012), this still being indeed an open question. The difficulty arises because of the trade-off between making the threshold high enough to ensure the *GP* approximation valid, but not so high to result in too small number of exceedances for accurate estimation of the parameters. In this paragraph, two well-known methods, which are closely related to the *GP* characteristics presented before, are going to be introduced. They are both based on the fitting of models across a range of different thresholds and according to Coles (2001) may be defined either an exploratory technique or a diagnostic assessment of the stability of parameters estimates.

In more detail, the first method is the so-called mean residual life plot, which depends on the expectation of the *GP* excesses and on the property of parameters threshold-invariance. It is theoretically known that if the *GP* constitutes a valid model for the excesses over the optimal threshold, the expectation of the *GP* excesses is a linear function of the threshold. Assuming the sample mean of the threshold excesses as an empirical estimate of the expectation of the excesses and provided *GP* distribution is a valid model for these excesses, the mean residual life plot is the locus of points

$$mrl = \left\{\left[u, \frac{1}{n}\sum_{i=1}^{n}(x_i - u)\right] : u < x_{max}\right\}, \tag{2.24}$$

where, *n* represents the number of excesses over the threshold *u*, while x_{max} denotes the largest of the x_i random sample.

The second procedure implies to estimate the model at a range of thresholds, by analysing the concepts of parameters threshold-invariance; that is, if a *GP* distribution provides a reasonable

model for excesses of a threshold u^*, then excesses of a higher threshold *u* should follow a *GP* distribution, where the shape parameters of the two distribution are the same, whereas the scale parameters are analytically related. Formally, it follows that the scale parameter changes linearly with *u* unless *k = 0*, that is

$$A = A^* - k(u - u^*). \tag{2.25}$$

Coles (2001, pg 83) proposed a re-parameterization of the *GP* scale parameter by the expression

$$\tilde{A} = A + ku = A^* + ku^*, \tag{2.26}$$

in order to highlight the constancy of the *GP* parameters with respect to the threshold. Hence, if u^* is a valid threshold for excesses to follow the *GP* distribution, then both estimates of the shape and scale parameters, $\hat{k}$ and $\hat{A}$, should be constant above u^*. The optimal threshold, u^*, should be consequently selected as the lowest value of a range of thresholds for which the estimates remain near-constant (i.e. up to the sampling variability).

The third method that is going to be defined is still based on the *GP* characteristics, but it permits an automated selection of the optimal threshold over a range of valid values by using the maximum likelihood estimation and a goodness of fit test. In addition, an uncertainty quantification based on bootstrap technique and an extension to express the covariate influence on the threshold selection – for instance, the effect of the cosine of wave directions with respect to a sample of excesses constitutes by significant wave heights – can be included (Thompson et al. 2009).
The procedure presented by the authors is focused on the right tail of the distribution. They proposed to set a uniformly spaced sequence of possible threshold values $\{u_{i=1}^{n=100}: u_1 < u_2, \dots, u_n\}$ between the median and un upper threshold, which is assumed to be equal to the 98% empirical quantile or coincident with the 100$^{\text{th}}$-order threshold whether the number of excesses greater than u_n is lower than 100. For each potential threshold the *GP* is fitted using the *ML* technique and the differences in the modified scale parameters for neighbouring thresholds are calculated

$$\Delta\tau_j = \tilde{A}_i - \tilde{A}_{i-1}, \quad j = 2, \dots, n. \tag{2.27}$$

By assuming asymptotic normality and provided these scale differences being *GP*-distributed, that is the expected values of *ML* estimates is centred around zero, the procedure requires, simply, to define a sample of scale differences and test the consistency with a goodness-of-fit under the following null hypothesis

$$H_0: E[\Delta\hat{\tau}] \approx 0 \; and \; \Delta\tau \sim N(0, Var[\Delta\hat{\tau}]). \tag{2.28}$$

For each threshold, a sample of these scale differences can be obtained by increasing the threshold increment by increment and testing the H_0 hypothesis by a Pearson χ^2 test with a 0.2 significance level. The optimal threshold is found to be the first not rejected threshold.

2.7 The multiple-threshold-method

The multiple-threshold-method (*MTM*) is proposed by Deidda (2010) and, in general, it can be assumed as a mixture model because it is based on fitting the entire dataset source by a single distribution (i.e. the *GP* distribution). It is tailored to rain data and its performances are demonstrated to be flexible and robust in tackling problems involving samples constitutes by rounded-off data and, for these cases, even more suitable than the fixed-threshold approach. In this work, its applicability to wave data is tested, where some drawbacks may arise because of either the effect of directionality or the de-clustering procedure adopted herein. Therefore, its application is restricted to the total sample wave data within the main directional sector.

As done before for the *GP*, *GEV* and *WBL*, the purpose is to report, firstly, the expression of the cumulative distribution function and the corresponding quantile function, whereas, successively, the essential characteristics of the method are briefly defined.
The starting point is addressed to re-parameterize the *GP* distribution in equation (2.13) using estimates obtained with any threshold u, in order to obtain a simple threshold-invariant three-parameter distribution function which may assure a perfect overlapping with the *GP* fitted on the exceedances over any threshold larger than the presumed optimum one. For this purpose, according to the same symbols used by the author, but conforming to the sign convention used herein, the following formulation is assumed

$$GP\_MTM(x;\xi,\alpha,u) = \begin{cases} 1-\zeta_0 exp\left[-\frac{x}{\alpha}\right] & \xi = 0 \\ \\ 1-\zeta\left[1-\xi\frac{x}{\alpha}\right]^{1/\xi} & \xi \neq 0 \end{cases}, \tag{2.29}$$

where a modified shape parameter is found to be defined by the same relationship that are characteristic of the *GP-Poisson* model

$$\alpha_0 = \alpha_u + \xi_u. \tag{2.30}$$

On the other hand, the survival function (i.e. the probability to observe excesses of u) is determined through the following expression

$$\zeta_0 = \frac{N_0}{N} = \begin{cases} \zeta_u exp\left[\frac{u}{\alpha_0}\right] & \xi_u = 0 \\ \\ \zeta_u\left[1-\xi\frac{x}{\alpha_0}\right]^{1/\xi} & \xi_u \neq 0 \end{cases}, \tag{2.31}$$

where

$$\zeta_u = Pr[X > x | x \geq 0] = \frac{N_u}{N}, \tag{2.32}$$

represents the estimator of the probability to observe an exceedance of the threshold u, while N is the sample size (including the zeros) and N_u is the number of excesses above the threshold u.

On the same time, assuming *x* as an independent and identically distributed random variable, it is straightforward to obtain the relationship between the distribution function of annual maxima, *G(x)*, the threshold-invariant *GP* parameterization and the yearly return period, *T*; namely

$$G(x) = GP\_MTM(x)^n = 1 - \frac{1}{T} \tag{2.33}$$

where $n = 365.25$ is the average number of days in a year. The corresponding quantile function for the *T*-year return period is given by:

$$x_T = \begin{cases} -\alpha_0 ln\left[\frac{1-\left(1-\frac{1}{T}\right)^{1/n}}{\zeta_0}\right] & \xi = 0 \\ -\frac{\alpha_0}{\xi}\left\{\left[\frac{1-\left(1-\frac{1}{T}\right)^{1/n}}{\zeta_0}\right]^{\xi} - 1\right\} & \xi \neq 0 \end{cases} . \tag{2.34}$$

Finally, a hierarchical procedure is provided to evaluate the so-called Multiple-Threshold-Method (*MTM*) estimates, which are denoted as ξ^M, α_0^M, and ζ_0^M.
To a close extent, the shape, α, and scale, ξ, estimates on the excesses over a range of thresholds *u*, as well as the ζ_u estimates, are firstly computed; then, the α_0 and ζ_0 are evaluated by using equations (2.30) and (2.31). At this point, the author suggestion is to prescribe a set of thresholds values, ranging from the lowest to the largest value where the estimates can be assumed stable, in order to achieve an even more reliable *MTM* estimates. In fact, the *MTM* approach is subsequently carried out by formally considering the following set of steps:

i) the *MTM* estimate ξ^M of the shape parameter is assumed to be equal to the median of the ξ estimates within the previously fixed range of threshold;

ii) the *MTM* estimate α_0^M of the scale parameter is obtained as the median of modified scale parameters α_0^C estimates, function of both the ξ^M estimate and the new α_u estimates, conditioned to the same ξ^M estimate obtained at the previous step, on the same suggested range of thresholds;

iii) finally, the *MTM* estimate ζ_0^M is the median of new values of ζ_0, conditioned to the MTM estimates ξ^M and α_0^M obtained at previous steps, and denoted as ζ_0^C, within the same range of thresholds.

2.8 Best-fit metrics

Diagnostic, parametric and non parametric best-fit metrics are used to assess the model performance. Herein, three metrics used to assess the model performance are introduced: the Error Norm (EN, Li et al. 2008), the Akaike Information Criteria (AIC, Akaike 1973), the Bayesian Information Criteria (BIC, Schwarz 1978), and the bias and root-mean-square-error (RMSE) measures. Basically, the lower the criterion the better the fit.

According to Li et al. (2008) the Error Norm (EN) can be expressed by

$$EN_o(\hat{\theta}) = \max_{1\le j\le N}\left|\left(\frac{j}{1+N}\right)^h - F_{\hat{\theta}}(x_j)^h\right|, \tag{2.35}$$

where, $\{x_j\}_{j=1,N}$ are the ordered data $X=\{x_i\}_{i=1,N}$, in increasing magnitude, h is a weight-parameter, which controls the emphasis of the fitting algorithm to the sample. The equation (2.35) corresponds to the *Kolmogorov-Smirnov* test when h = 1.0, and equal emphasis is addressed to both halves of the rank ordered distribution. If h <1, emphasis is assigned to the lower tail, whereas for h > 1.0, it is focused on the upper tail, so on the extreme events. Practically, the equation (2.35) can be assumed as a measure of the maximum deviation between the empirical cdf $\tilde{F}$ (i.e. the plotting position), and the probabilistic model obtained with optimised parameters $\hat{\theta}$.

In this work, an improved version of this metric is assumed. Basically, it is to be noted that the first term on the equation is the Weibull formula, which is biased and not suitable for distribution used herein (e.g. Cunnane 1978). Therefore, unbiased estimator are employed, by using an ad hoc plotting position formula $\tilde{F}$ (e.g. §1.2.4), according to

$$EN(\hat{\theta}) = \max_{1\le j\le N}\left|\tilde{F}^h - \hat{F}(x_j)^h\right|. \tag{2.36}$$

where $\hat{F} = F_{\hat{\theta}}$ is the probabilistic model, while the other terms are identical to those expressed in the previous equation.

By assuming a series of B replicates, i.e. generated from a Monte Carlo procedure, or from a re-sampling technique, an improved estimator of the $EN_o(\hat{\theta})$ is introduced by the author, by computing the expected value of the error norm. Applying to equation (2.36).gives

$$E\left[EN(\hat{\theta}_{jr})\right] = E\left[\max_{1\le j\le N}\left|\tilde{F}^h - \hat{F}(x_{jr})^h\right|\right], \tag{2.37}$$

where $j = 1,2,\ldots,B$ and $E[\ \ ]$ is the expectation operator.

AIC is founded on information theory and gives the model providing the best compromise between bias and variance. Introducing the likelihood of the fit, L, the number of parameters of the distribution, k_p, and assuming the sample size N large enough (i.e. asymptotic condition), AIC is given by:

$$AIC = -2\ln L + 2k_p. \tag{2.38}$$

It can be interpreted as the sum of two terms, the first one measuring bias and the second one measuring variance.

Under the same assumptions as AIC, BIC is formally defined as:

$$BIC = -2\ln L + k_p \ln N, \tag{2.39}$$

Basically, AIC penalizes the number of parameters less strongly than does BIC.

Finally, bias and RMSE are computing by using a Monte Carlo procedure and are known to provide an overall measures of the performance of an estimator $\hat{\theta}$. In particular, the bias of an estimator $\hat{\theta}$ of θ is given by

$$Bias(\hat{\theta}) = E[\hat{\theta} - \theta] = E[\hat{\theta}] - \theta, \tag{2.40}$$

while the RMSE is obtained as

$$RMSE(\hat{\theta}) = \sqrt{E[\hat{\theta} - \theta]^2} = \sqrt{Bias(\hat{\theta})^2 + Var[\hat{\theta}]}. \tag{2.41}$$

3 Long-term climate analysis

Generally, wave climate analysis requires a large amount of data to ensure the statistical significance; unfortunately, spatial and temporal wave records are usually limited with respect to other Earth Science area, e.g. Hydrology. In the Mediterranean Sea and in our study area, the problem of having representative and effective data is amplified. Typically, three main data sources can be employed: (a) scalar or directional wave buoys, moored at specific locations; (b) satellites; (c) wave generation models. The best solution lies in the combined use of all sources In this work, the analysis are primarily restricted to available database, i.e. related to (a) and (c) types.

As a matter of fact, the orography characterizing the Northern Mediterranean side is very articulated and the presence of large island and protruding peninsulas constitutes a deep obstacle to an accurate and reliable simulations with wave generation models (Cavaleri and Bertotti 2004). On the other hand, one of the few existing wave atlas (e.g. the WW-Medatlas project, Medatlas Group, 2004) is not appropriate to perform extreme wave computation, since classical wave parameters are primarily reported on frequency tables of joint occurrences. Despite some solutions to this problem should be reasonable alleviated by adopting ad hoc chain of models, e.g. as one recently proposed by Mentaschi et al. (2013), the attention in this work is primarily focused on employing open source archives and data recorded by wave buoys. Notwithstanding, two consideration should be remarked. First, well-known global atmospheric reanalysis dataset, namely ERA-interim (provided by the European Centre of Medium Weather Forecast ECMWF) and NOAA WAVEWATCH III CFSR Reanalysis Hindcast (provided by the North Oceanic and Atmospheric Administration) are not totally suitable in that area, since wave fields are inherently underestimated due to known underestimation of the wind fields. Second, wave buoy data are clearly more difficult to correct, in spite of being more reliable.

In this chapter, a general overview of the study area, as well as the available data set are presented. After a short description of the dataset, the quality control check is synthetically introduced and the application to wave and wind dataset is described. Then, both long-term wave- and anemometric-climate are reported.

Pre- and post-processing checking operations have been carried out on each database, by implementing a comprehensive, semi-automated, graphical and numerical tool on a Matlab project. A complete automatic implementation is not achieved, since a semi-automatic graphical control is reputed necessary to adequately scrutinised anomalous data, according to WMO (2009) suggestions. In addition, long-term statistics are automatically classified either in graphical (e.g. wave climate-rose, wind-rose, sea storm polar plot, fetch distributions) or in tabular (e.g. bi-variate histograms, sea-storm classification, quality control report) form. Herein, only graphical results are included.

After this mandatory checking, the dataset correctness can be reasonable assumed valid to perform the EV analysis (Repetto and Solari 2011).

3.1 Study area

Fig. 3.1 depicts a general overview of the Sardinia Island, localized in the Western Mediterranean Sea, and a summary of the available dataset, which is constituted by wave, wind, and pressure data, having non-homogeneous spatial distribution and heterogeneous time scales.

Globally, the database at disposal is constituted by 12 data series. In detail:

- three wave data series are collected from The Italian National Wave Measurement Network (Rete Ondametrica Nazionale, RON), recorded by Alghero, Siniscola-Capo Comino, Cagliari-Capo Boi directional wave buoys, respectively moored on the Northern, Eastern and Western Sardinian deep water (light blue cross markers in Fig. 3.1, see also Tab. 3.1).

- nine wind and pressure data series are acquired from the Italian National Civil Aviation, meteo (Ufficio Generale per la Meteorologia - Ente Nazionale Aviazione Civile, UGM-ENAV) network (blue circle markers in Fig. 3.1, see also Tab. 3.2).

Wave RON data consists of 3- or 1/2-hourly bulk spectral parameters: significant wave height ($H_{m0}, (m)$), mean wave direction ($\theta_{wave}, (°N)$), peak wave period ($T_p, (s)$) and mean wave period ($T_m, (s)$).

Wind data are a collection of three-hourly (i.e. synop messages) wind speed ($U, (kn)$), mean wind direction ($\theta_{wind}, (°N)$) and air pressure ($p, (hPa)$) data.

Additionally, it is deemed useful to include open source data with the purpose of extending the whole dataset, by overcoming both the spatial lack of data and the limited temporal extension. Therefore, data computed by WAM and WAVEWATCH III models were downloaded; namely:

- an archive of three-integral wave parameters – the overall significant wave height (*hs, m*), the peak period (*tp, s*), and the average direction at the peak period (*dp, °N*) – extracted from the NOAA-NCEP's model, and covering about 30 years (1979-2008) at 3 hourly intervals on 1/6 degree resolution grid in the Mediterranean Sea (red triangle markers in Fig. 3.1);
- a longer but coarser dataset constituted by four integral wave parameters – significant wave height (*SWH, m*), mean wave direction (*MWD, °*), mean wave period (*MWP, s*), peak period of *1D* spectra (*P1PS, s*) – from the ECMWF's model, covering about 35 years (from 1979-2014) at 6 hourly intervals on 1 degree resolution grid in the Mediterranean (green disk markers in Fig. 3.1).

In Tabs. 3.1 and 3.2 a summary of RON wave buoy and wind station is reported. In particular, each station is identified by a WMO (World Meteorological Organization) code, and some terms related to time extension: i.e., time period, effective annual extension and effective duration. Finally, wind measurement stations are also classified according to the ICAO (Airport Code Aviation Organization) code.

Tab 3.1 Summary of RON wave buoys.

RON wave buoy	WMO CODE	Time Period	K_{AM} (year)	K_{POT} (year)
Alghero	61213	1989-2014	24.666	17.347
Cagliari	61221	2002-2011	4.225	3.286
Capo Comino	61212	2007-2014	4.740	3.604

Tab 3.2. Summary of wind stations.

UGM-ENAV Wind Station	WMO CODE	ICAO CODE	Time Period	K_{AM} (year)	K_{POT} (year)
Alghero	16520	LIEA	1951-2013	60.84	46.40
Capo Bellavista	16550	LIEB	1951-2013	61.99	55.82
Capo Caccia	16522	LIEH	1975-2013	28.62	24.48
Capo Carbonara	16564	LIEC	1951-2013	59.03	40.20
Capo Frasca	16539	LIEF	1962-2013	50.80	40.55
Capo San Lorenzo	16542	LIEL	1980-2013	24.10	4.92
Elmas	16560	LIEE	1980-2013	24.53	22.73
Guardiavecchia	16506	LIEG	1980-1997	13.90	9.32
Olbia	16531	LIEO	1969-2013	10.32	8.34

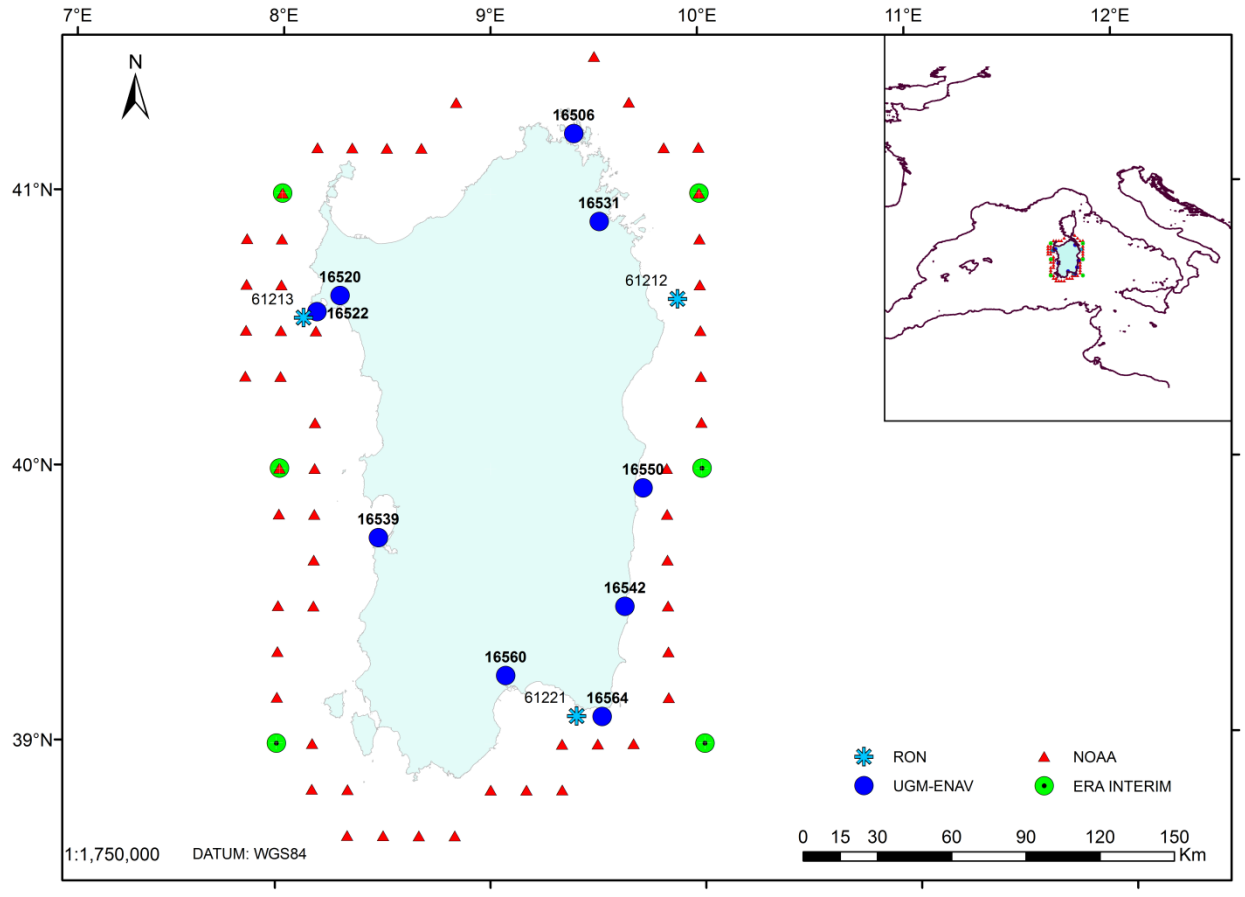

Fig. 3.1 Map of the study area and overview of data-source locations. In particular, these symbols are used: RON buoy data (light blue cross markers); ERA-interim grid points (green disc markers); NOAA grid points (red triangle markers); UGM-ENAV (blue circle markers).

Basically, K_{AM} and K_{POT} are computed as a function of the station efficiency, $\eta = K_{POT}/K_{AM}$, which, in turn, is the key result of the data quality control analysis (§3.3). According to Goda (2010), K_{AM} and K_{POT} are assumed to be tailored for AM or POT extreme analysis computing, respectively.

3.2 The RON Italian network

The RON network was constituted by 15 real-time directional buoys distributed along the Italian coasts (Bencivenga et al. 2012, The Italian Wave Atlas 2004). Shortly, data had been collected since 1989 by using 8 *pitch-roll Datawell Wavec buoys* dislocated at 8 measurement stations (Alghero included); in 1999 two more stations, equipped with *particle-following Datawell Waverider buoys*, were added and the remaining five buoys (*particle-following Axys Triaxis buoys*) had been displaced in between 2001-2002 (Capo Comino – Siniscola included) and 2007 (Cagliari – Capo Boi included); in addition, also the other 10 buoys were changed to *Triaxys* type. Finally, since 2009 *Watchkeeper* buoys were moored in each location. Lastly, the RON network was completely dismissed in December 2014.

It is to be remarked that until 1998 data were transmitted in semi-real time and were generally recorded in three-hourly series and every half-hour for limited time intervals, i.e. when storm peak wave heights exceeded a station-dependent-threshold value (e.g. $H_{m0} = 5.00\ m$ – Alghero buoy). Between 1999 and 2001 an improved data recording based on a real-time data transmission system was experimented. In the 2002 the network was restructured and upgraded and data were measured in real-time every thirty minutes. Finally, in 2009 a new network architecture was employed, based on a real-time transmission and half-hour data measurement.

Unfortunately, only three buoys are moored along the Sardinian coasts (i.e. light blue cross markers in Fig. 2.1). It is to be noted that spatial density of the wave buoys over the Sardinian coast are much less than sufficient, since the Northern part is totally un-detected, the Eastern and Southern part are almost under-detected, while the Western part is the best equipped, but still not sufficiently covered. Furthermore, an inherent limitation should be stressed, which is associated to the reduced temporal extension (see Tab. 3.1, Figs. 3.2-3.4, 3.7-3.9). Unlike Cagliari and Capo Comino wave buoys, temporal extension of Alghero wave buoy is assumed to be reasonable sufficient to perform extreme wave analysis computation (Chapter 4). Nevertheless, it is to be remarked that other difficulties arise, since the Alghero dataset is obtained by using three different instruments; hence, homogeneity need to be verified and eventually relaxed.

3.3 Data quality control

Data quality control is essential and becomes mandatory at the aim of obtaining validate dataset and rigorous statistical, numerical and physical results.

First of all, reliability, homogeneity, and representative criteria have been verified, to a large or less extent, for each dataset. A substantial part of the work has been addressed on analysing different data verification trustworthiness, on the basis of diversified procedures, due to the distinct physical behaviour of the considered variables.

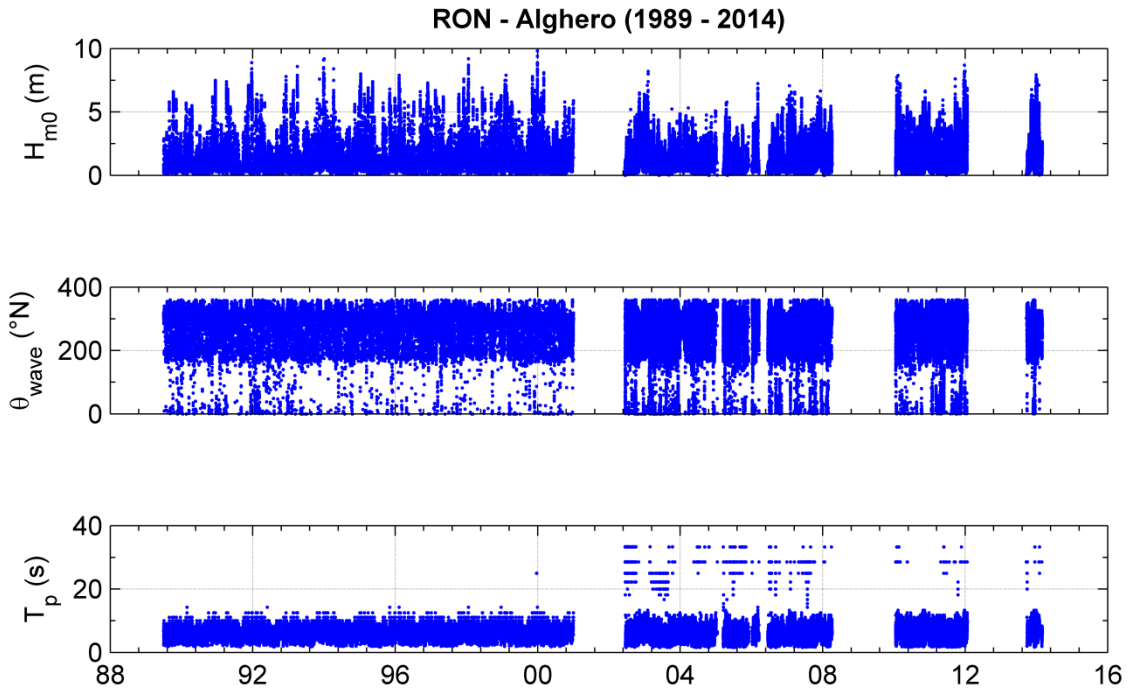

Fig. 3.2 General overview of wave data originally recorded by the RON –Alghero wave bouy.

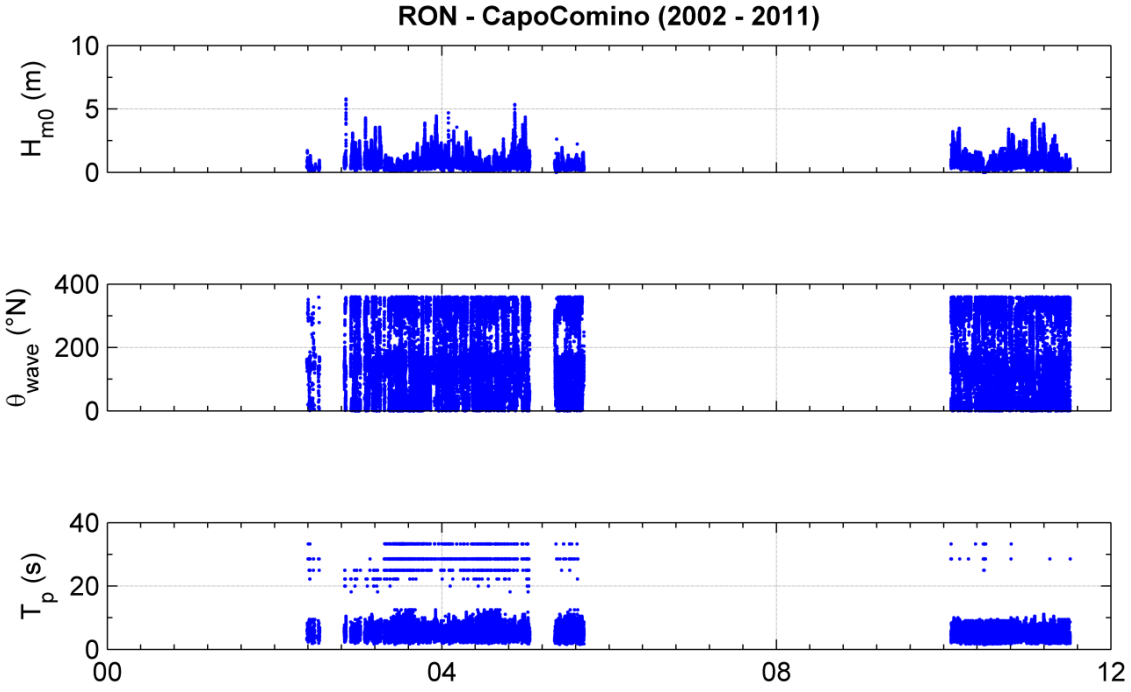

Fig. 3.3 General overview of wave data originally recorded by the RON – Capo Comino wave buoy.

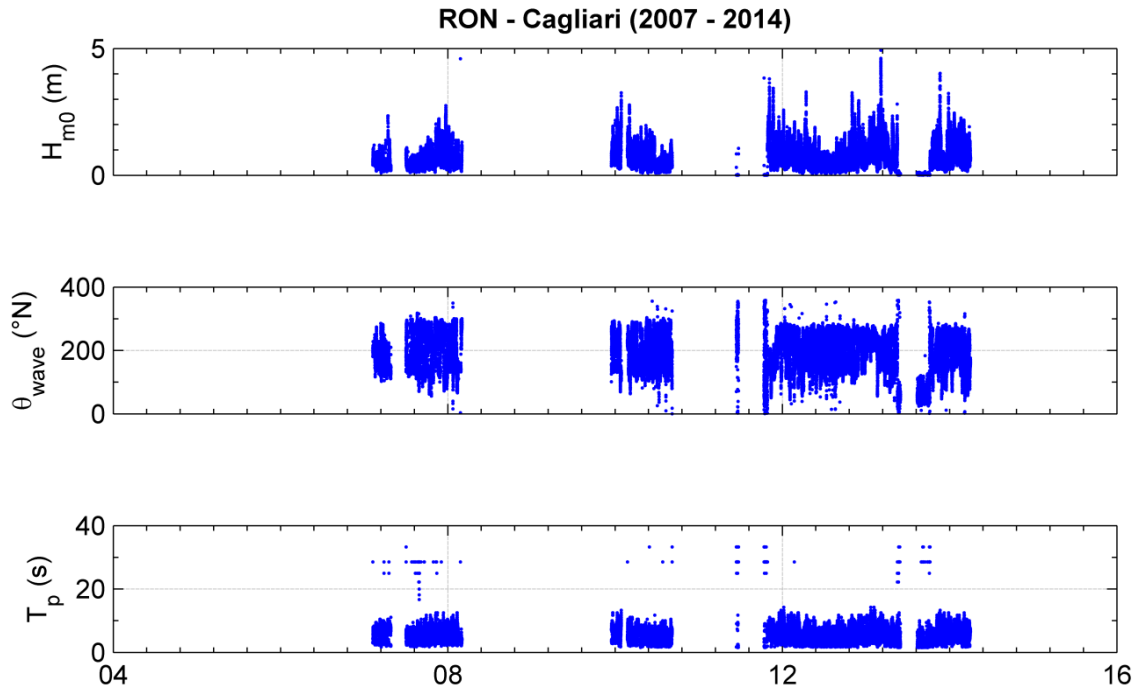

Fig. 3.4 General overview of wave data originally recorded by the RON – Cagliari wave buoy.

According to Baffo et al. (2005a) suggestions, as well as following the WMO (2009) guidelines and slightly adapting the procedure proposed by Repetto and Solari (2011) for wind data, at least two main classes of control can be employed:

a) a weak-climatic control, which is based on assigning minimum and maximum non-exceeding threshold values, in order to restrict each dataset field to a physically allowable level. In this way, clearly erroneous values can be quickly and simply deleted;
b) a consistency-climatic-control, finer and more elaborate than the previous one, which involves a mutual comparison of different variables at an adequate space-time scale; it is carried out for all doubt values either falling outside to the physical limits (previous defined) or denoting abnormal behaviour. In this way, the so-called outliers, or spikes, or unusual items, i.e. data characterized by suspicious values or significantly incorrect, are identified.
Furthermore, other typologies can be included as well; namely: i) an internal consistency control, based on comparison of two or more variables at the same instant time; ii) a temporal consistency control, where the time persistency between different climatic events is analyzed; iii) a spatial consistency control, which represents the natural spatial variation of the climatic data (Baffo et al 2005b).

Both type of controls are used to correct either wave or wind dataset. Furthermore, wave data were further scrutinised with respect to anomalous values occurring due to device transmission malfunctioning, measurement errors, un-moorings, missing information, according to ISPRA (2004) guidelines.

A MATLAB project has been implemented, with the aim of providing an automated computational procedure useful to the whole pre- and post-processing data-quality-control analysis. Specifically, the consistency-quality-checking requires a semi-automated-graphical approach, according to the WMO (2009) suggestions. As a matter of fact, the graphical control is the less subjective procedure to affirm whether a suspicious value is or not a true outlier. Obviously, is unthinkable to apply the graphical control to all points of the dataset, as for computational reasons as for practical convenience.

The station efficiency is evaluated as a function of: total number of expected observations (NDA), total number of observations actually recorded (NDP), total number of missing data (NDM), total number of temporal gaps (NDG), total number of errors (NDE) and total number of observations removed (NDR) in the studied time period. Basically, four efficiency parameters can be computed:

$$\eta_T = \frac{NDP}{NDA} = \left(1 - \frac{NDM}{NDA}\right), \tag{3.1}$$

$$\eta_G = \frac{NDG}{NDA}, \tag{3.2}$$

$$\eta_E = \frac{NDE}{NDA}, \tag{3.3}$$

$$\eta_R = \frac{NDR}{NDA}. \tag{3.4}$$

3.4 Correction of RON dataset

Two examples of graphical controls applied to the RON - Alghero wave dataset are reported in Fig. 3.5, 3.6. In Fig. 3.5 an internal climatic control is performed, by comparing the set of significant wave height, wave mean direction, peak and mean wave period, at a same instant time *t* on a 2-days-moving-window, i.e. ranging between $[t - 24, t + 24]$ hours. The controlled value (red circle point) is above the threshold value (red continuous line, H_{th}= 5.0 m) between three-hourly ($H_{m0} < H_{th}$) or half-hourly ($H_{m0} \geq H_{th}$) data. Despite some high oscillations on the mean wave direction and wave period recorded as null, the value is assumed to be valid; thus, a correction on the wave period is performed (a nearest neighbour value is assigned).

In Fig. 3.6 is shown an example of the internal consistency climatic control. In such situation, wave data is cross-controlled by employing other data sources. A comparison between wind field measured by three nearby anemometric stations (i.e. Alghero, Capo Caccia and Capo Frasca), as well as wave field simulated by the WAM model (ERA Interim archive) is provided. An erroneous sea storm the day after the effective sea storm may be easily noted.

A global synthesis of the whole correction is proposed graphically and analytically. In particular, splitting up the information in homogenous periods, a summary is provided in Tab. 3.3. Specifically, start and end time, actual and potential number of observations, total efficiency, effective annual extension and effective duration are shown for each period.

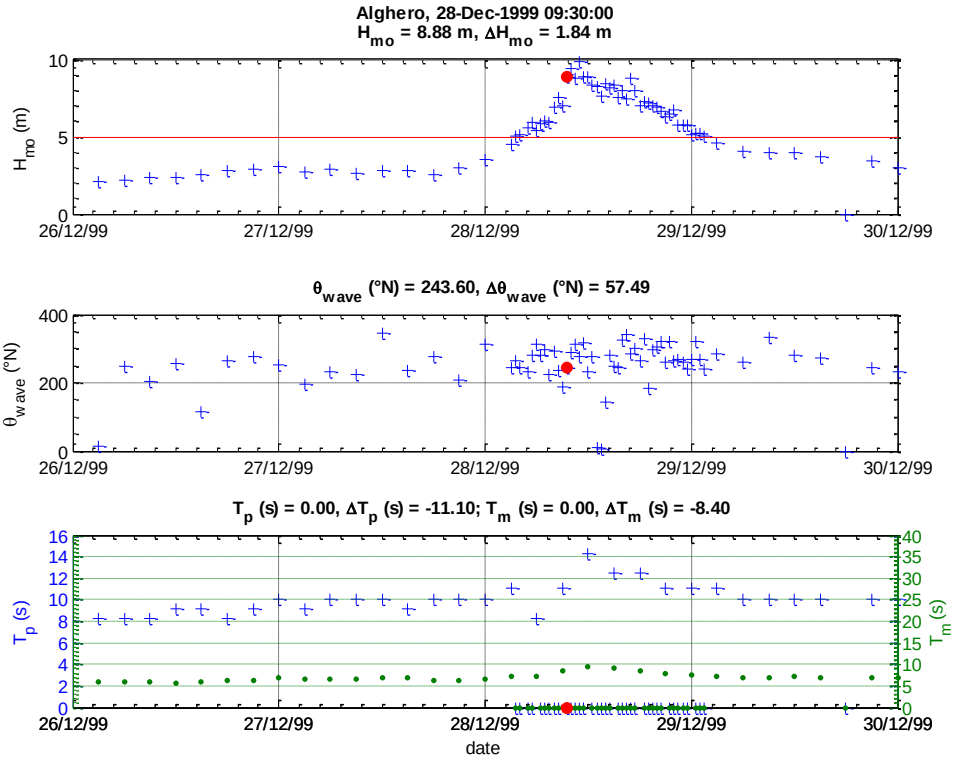

Fig. 3.5 RON Alghero - internal climatic control. From top to bottom are displayed the significant wave height, the mean wave direction and the peak and the mean period, on a (2 days)-moving window, focusing on a 1999 storm event (known as The Christmas Event). The red circle marker identifies the point under control.

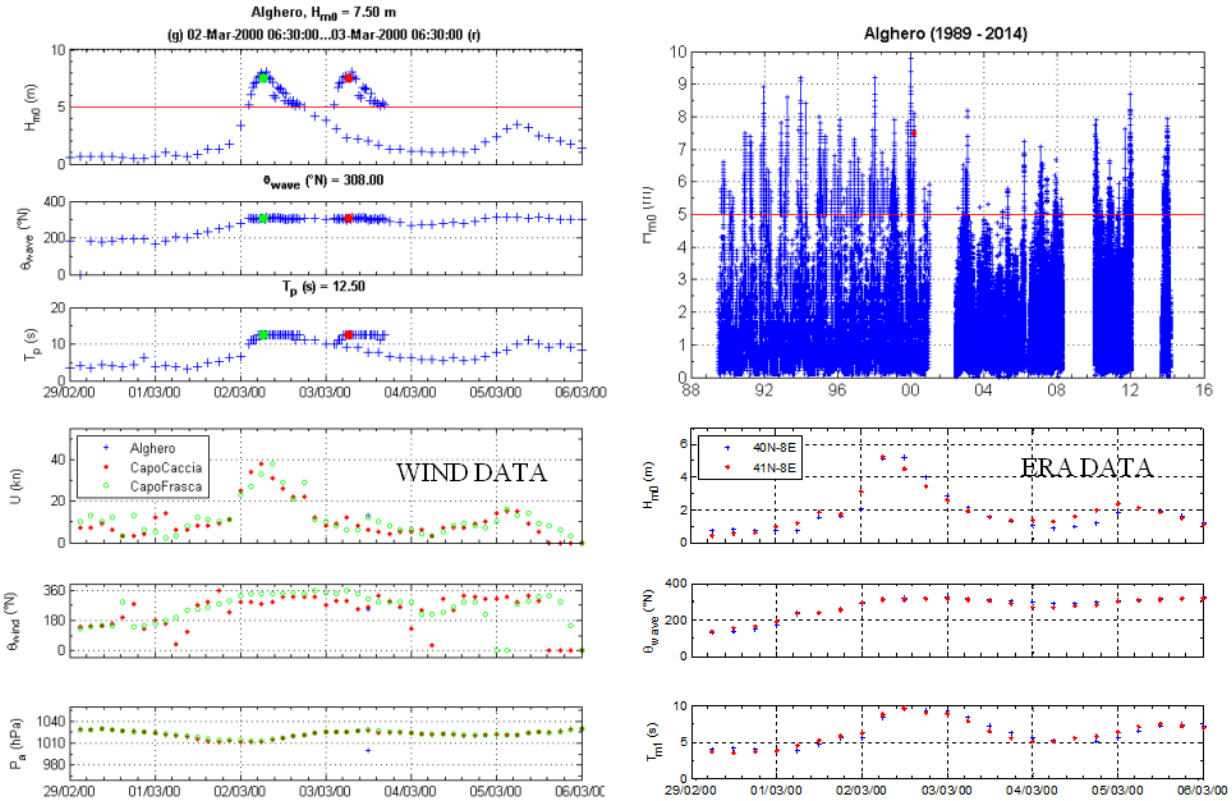

Fig. 3.6 RON Alghero – internal consistency climatic control. It is shown a comparison between RON-Alghero wave data (left and right upper panel), wind field measured by nearby anemometric stations (i.e. Alghero, Capo Caccia and Capo Frasca, left lower panel), and wave field simulated by the WAM model (ERA Interim archive, right lower panel).

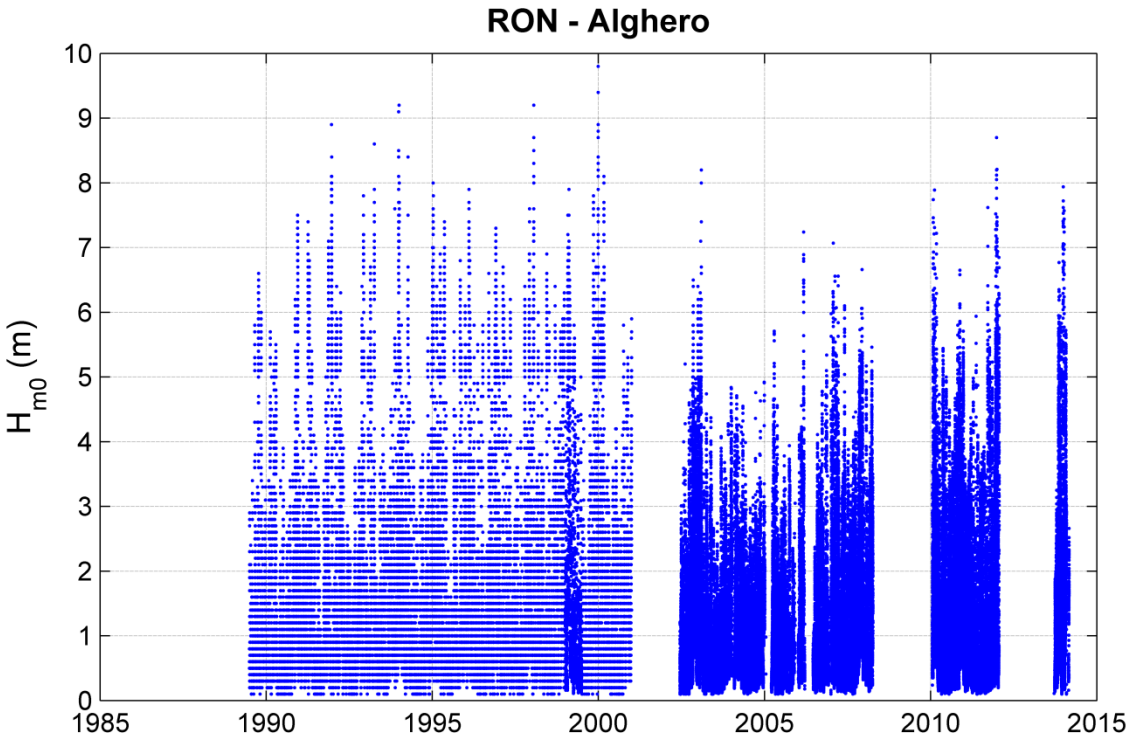

Fig. 3.7 Overview of corrected RON – Alghero-significant wave height dataset

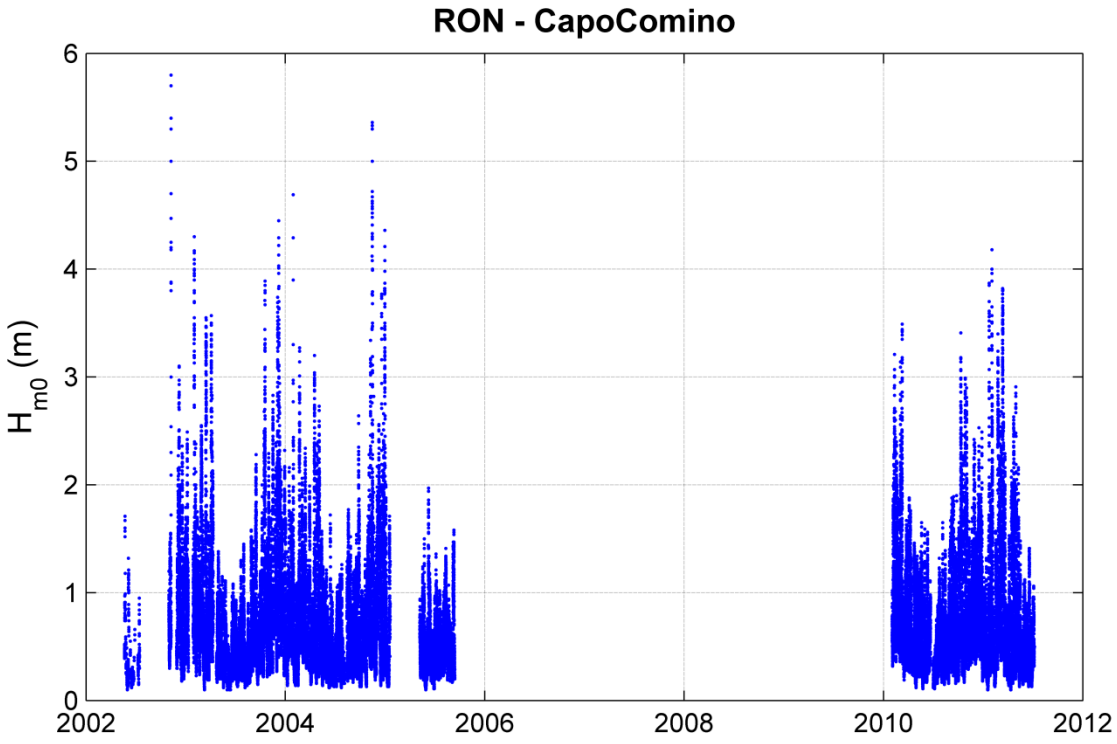

Fig. 3.8 Overview of corrected RON – Capo Comino-significant wave height dataset

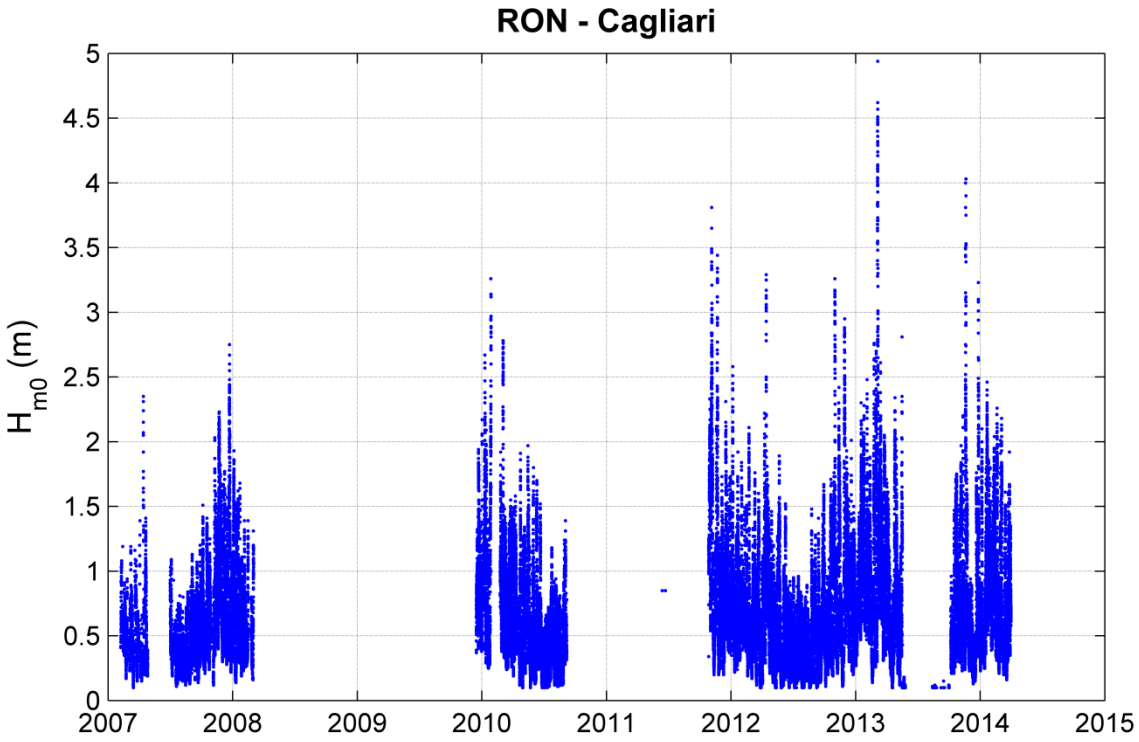

Fig. 3.9 Overview of corrected RON – Cagliari significant wave height dataset.

The global efficiency for each station is computed as weight-averaged of the total efficiency for each homogeneous period, assuming weights equal to the actual number of observation (NDA) and excluding long-term temporal gap. On the whole, values greater than 0.70 are obtained for each station.

Despite a fairly adequate efficiency is given by RON-Capo Comino and RON-Cagliari wave buoys, a drastic short duration is also evident, which is assumed not to be consistent for providing reliable extreme wave model inter-comparison. Therefore, a deeply application of the EV methodology, as described in the previous chapters, is carried out by considering exclusively the RON-Alghero data set.

Tab 3.3 Summary of RON dataset correction.

	Start time (UTC)	End time (UTC)	NDA	NDP	η_T	k_{AM} (year)	k_{POT} (year)
RON Alghero							
I period	01/07/1989 00:00	15/06/2002 09:30	37860	32346	0.8544	12.957	11.070
II period	15/06/2002 10:00	28/02/2014 23:30	205276	110044	0.5361	11.709	6.277
Tot	01/07/1989 00:00	28/02/2014 23:30	432436	304120	0.7033	24.666	17.347
RON Capo Comino							
I period	21/05/2002 17:30	12/09/2005 12:00	58069	39275	0.6764	3.312	2.24
II period	01/02/2010 09:00	07/07/2011 17:30	25025	23913	0.9556	1.427	1.364
Tot	21/05/2002 17:30	07/07/2011 17:30	83094	63188	0.7604	4.740	3.604
RON Cagliari							
I period	07/02/2007 08:30	02/03/2008 19:00	18693	10904	0.5833	1.066	0.622
II period	15/12/2009 11:30	06/09/2010 16:00	12729	10993	0.8636	0.726	0.627
III period	27/10/2011 14:00	31/03/2014 23:30	42643	35705	0.8373	2.432	2.037
Tot	07/02/2007 08:30	31/03/2014 23:30	74065	57602	0.7777	4.225	3.286

3.5 Sea storm definition, selection and classification

A sea storm is usually defined as a sequence of sea-states characterised by values of spectral wave parameters (e.g. $H_{m0,}\, \theta_{wave}$, $T_{p,}\, T_{m0*}$) that vary gradually within a given interval. The concept of sea storm is suitable to perform a preliminary de-clustering of wave data (e.g. Mathiesen et al 1994, Corsini et al. 2004, Mazas and Hamm 2011). In addition, it can be adequately used in view of the ascertainment of peak-storm-independency.
In this work a sea-storm is assumed to be represented according to the definition given by Corsini et al. (2004) and by slightly adapting some of the driven parameters. Thereby, a sea storm is a time series of sea states characterised by:

- wave height persistence over the threshold of $H_{thr} = 1.50\, m$ (which may be assumed representative of the background noise value on the Mediterranean Sea, e.g. Boccotti 2000), for more than 12 consecutive hours;
- wave height decay below the threshold H_{thr} for less than 12 consecutive hours;
- Original direction belonging to a determinate angular sector (i.e. ± 30° with respect to the initial direction).

In order to increment the sample size, it is chosen a threshold value ($H_{thr} = 1.00\, m$) to represent the Cagliari buoy. An example of sea storm definition and detection is depicted in Fig. 3.10.

In particular, once extracted, the sea storm sequence is classified, summarised through a set of values and parameters (i.e. starting and ending time, recorded values and statistical measures, gap percentage, ...), and stored into a file. In addition, another type of graphical control is implemented in order to check the goodness of anomalous sea storm values, i.e. outliers.

Once classified, the application of AM or POT approach require to identify each sea storm by its wave peak characteristics, i.e. $(H_{m0}, \theta_{wave}, T_p, T_m)_{peak}$. In particular, the peak employed to perform the extreme wave analysis is slightly different from the value recorded. Basically, the operation proposed by Goda, (2010, §10.1.2.C) to be used for wave profiles is adopted, by applying a parabolic fitting with a curve formed by three points $H_{peak-1}, H_{peak}, H_{peak+1}$. The aim is to eliminate the underestimation of the true maximum between two discrete sampling points.

Finally, according to the POT approach, a series of independent and identically distributed random peak is to be selected. First of all, herein, sample independency is assumed to be verified by imposing a fixed time interval between two consecutive sea storms. This value can be strictly calculated on the basis of the auto-correlation function of the observed time series; by imposing a weak independence between two consecutive storms (i.e. a correlation coefficient with a value equal to 0.4), a time interval of 48 hours is generally obtained. Therefore, it is chosen to assume a time interval equal to 48 hours, which is a common value recommended in literature (i.e. Mathiesen et al. 1994) and found to be valid on the Alghero dataset (for a shorter time series) also by other authors (e.g. Pisocopia et al. 2002b).

Tab 3.4 shows the number of independent sea storm identified for each RON wave buoy. Furthermore, initial (θ_{sec}^{in}) and final (θ_{sec}^{fin}) directions of homogeneous directional sector are

indicated. In particular, these last values are selected according to Corsini et al. (2004), as remarked in the next paragraph.

Tab 3.4 RON dataset. Homogeneous directional sector for each RON wave buoy: intial direction (θ_{sec}^{in} °N), final direction (θ_{sec}^{fin}°N), number of independent significant peak wave height (N, ev).

	RON Alghero				RON Capo Comino				RON Cagliari		
θ_{sec}^{in} (°N)	170	220	275	170	350	55	115	350	110	200	110
θ_{sec}^{fin} (°N)	225	275	335	335	55	115	170	170	170	265	265
N (ev)	17	95	632	744	30	13	36	79	53	55	114

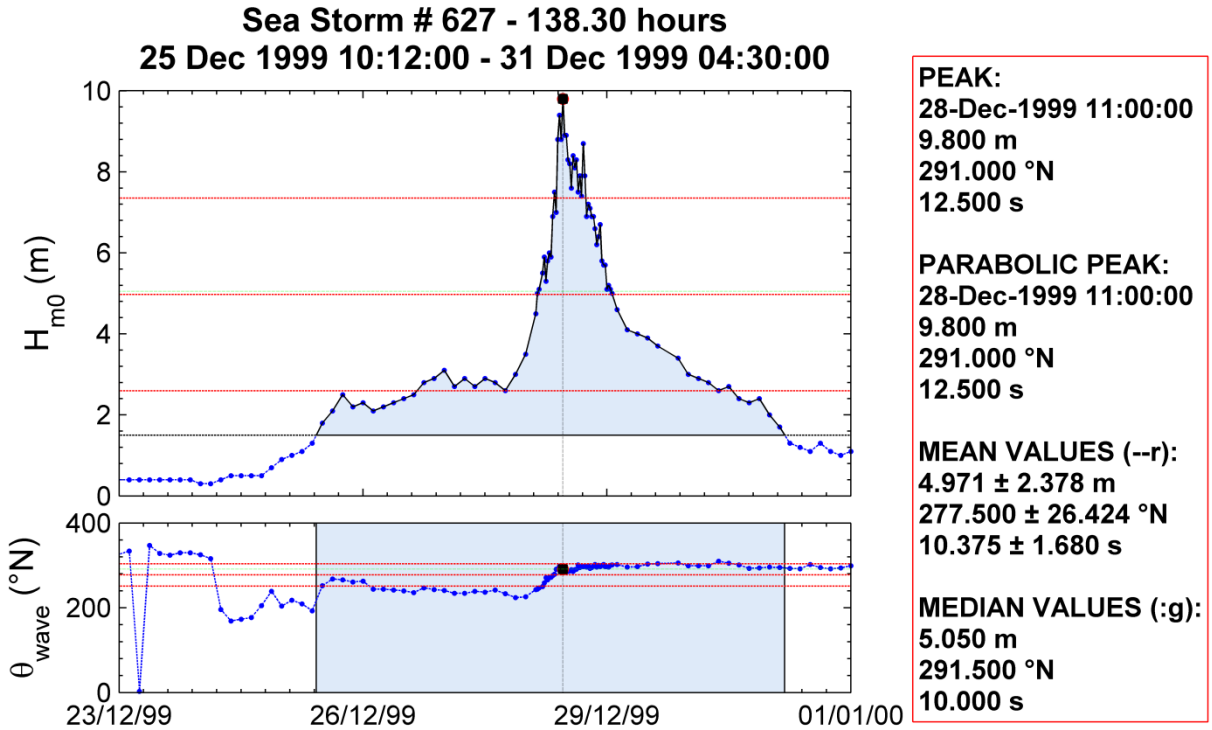

Fig. 3.10 RON Alghero: sea storm example. It is depicted a sea storm recorded during the so-called *The Christmas Event.* This representation is used to perform graphical control on the ascertainment of sea-storm validity

3.6 Long-term wave climate

In Fig. 3.11 a summary of the wave climate roses and sea-storm polar plots are depicted, while geographical and effective fetch distributions are displayed in Fig. 3.12 for each RON wave buoy. In particular, all buoys denote at least a bi-modal wave climate tendency, while the most energetic behaviour is evident on the North-West sector (Alghero buoy). On the other hand, it is to be noted that the limited sample size inherent to both Cagliari and Capo Comino RON buoys yield a lower number of sea storms and a less reliability than those provided by the Alghero RON wave buoy. According to Corsini et al. (2004), homogeneous directional sectors are identified by combining these three wave climate characteristics (Figs. 3.11 and 3.12). It is to be remarked that directional cluster values reported in Tab. 3.4 are subsequently used to perform the EV model inter-comparison (see Chapter 4).

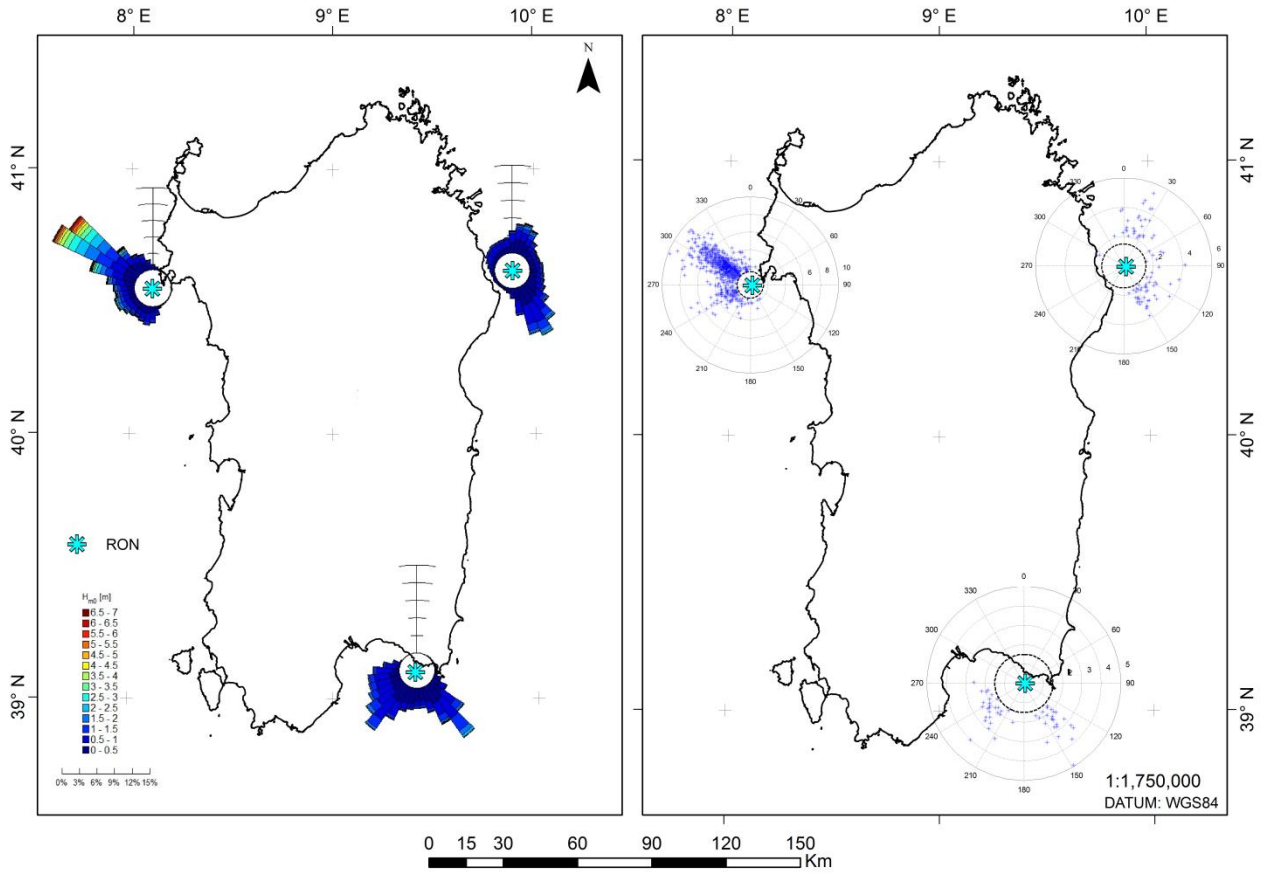

Fig. 3.11 Wave climate rose (on the left) and sea storm – polar plot (on the right). Alghero RON buoy, Capo Comino RON buoy, Cagliari – Capo Boi RON buoy are located on the Norhern, Eastern and Southern part, respectively.

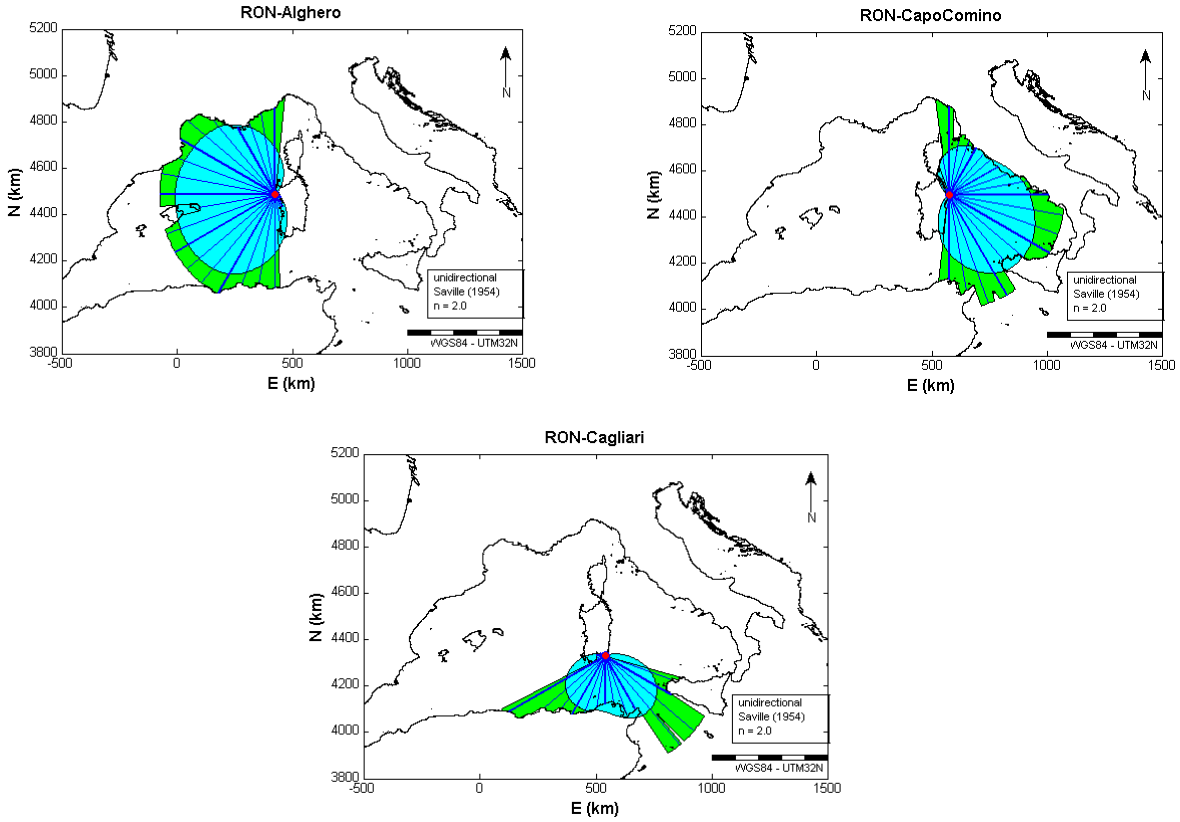

Fig. 3.12 Geographical- and effective-fetch distributions. Alghero RON buoy (upper right figure), Capo Comino RON buoy (upper left figure), Cagliari –Capo Boi RON buoy (lower figure).

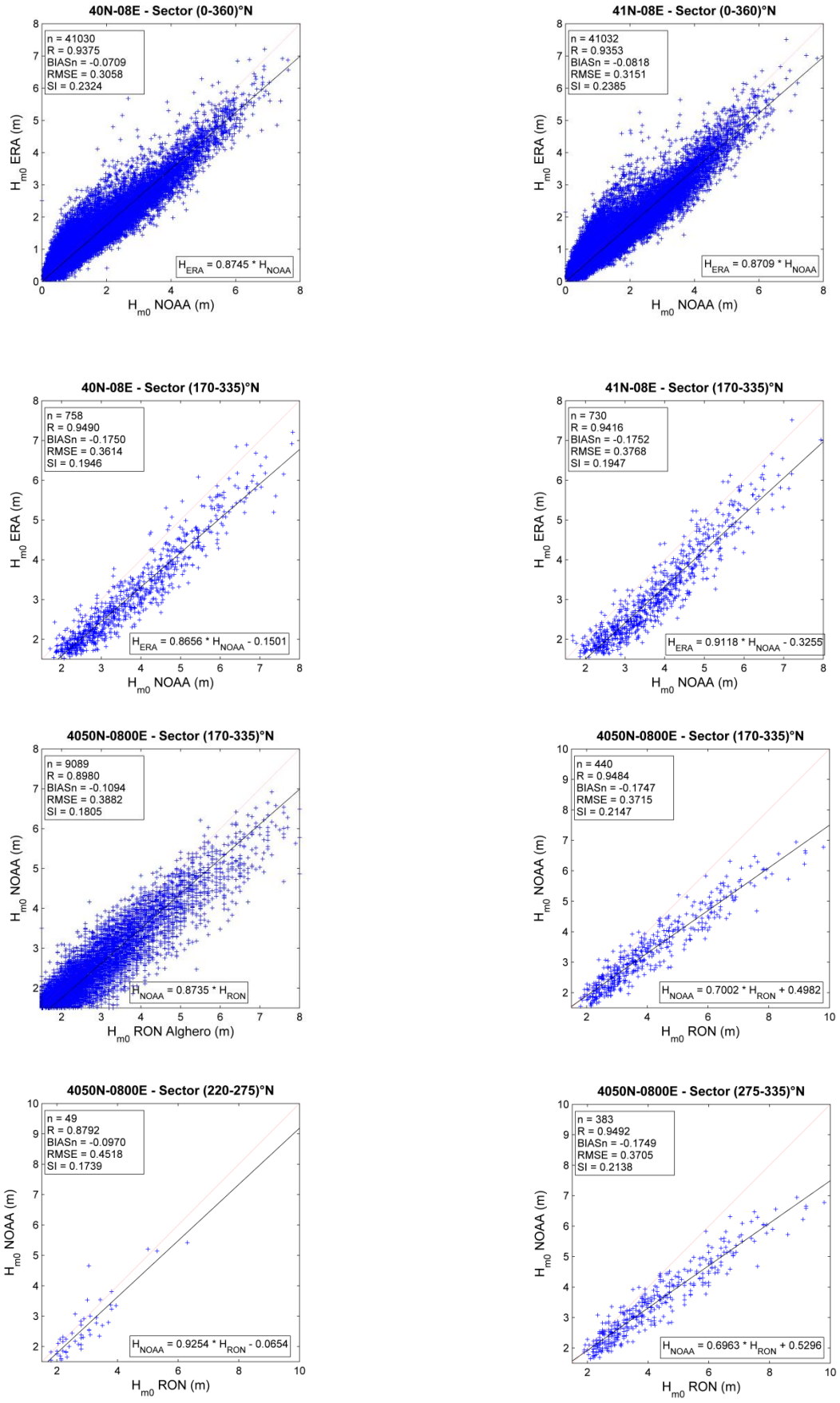

Fig. 3.13 Comparison between NOAA-ERA dataset for two of the common grid points 40 °N - 08 °E and 41 °N - 08 °E (first two rows of 4 plots) and between 40,50 °N - 08,00 °E NOAA grid point - RON Alghero wave buoy (last two rows of 4 plots).

3.7 Comparison between RON and NOAA wave dataset

In this paragraph, a brief comparison between RON and NOAA dataset is carried out and the correlation NOAA-RON in the Northern-Western (NW) part of the study area will be introduced, since it is going to be the subject of the EV analysis.

Formally, a deterministic time-domain analysis is performed at synchronous time instants between two grid points common to both NOAA and ERA dataset (40,00 °N 08,00 °E and 41,00 °N 08,00 °E grid points), as well as between the Alghero wave buoy and three of the most-closed NOAA grid points (40,50 °N 08,17 °E; 40,50 °N 08,00 °E; 40,67 °N 08,00 °E). Essentially, this approach is based on the comparison of the physical quantity of interest at specific time frames. In particular, the comparison is performed by applying a linear regression analysis and by using some classical statistical indexes – i.e. correlation coefficient (R), normalised bias (NBI), root-mean-square error (RMSE) and scatter index (SI) –, either for all data recorded/computed in the same specific instant time or for single peak storm events happened at most within a ±12 hours-time window.

In Figure 3.13 are shown some of the results obtained. The first two rows of 4 plots are referred to the NOAA-ERA interim comparison, while, for sake of economy, only the best fitted comparison between NOAA-RON Alghero is displayed in the other group of 4 plots.

Generally, NOAA dataset would seem to be more suitable with respect to ERA interim data for at least two reasons. First, it denotes a better spatial coverage along the Sardinian coasts, due to a finer spatial resolution (see Fig. 3.1), although it is to be pointed out that the overall temporal extension is shorter than that provided by the ERA interim dataset. Second, the peak storms are found to be higher than those recorded in the ERA interim database. As a matter of fact, in Fig. 3.13 the first two plots in the first row depict the comparison between all significant wave heights computed by each model (NOAA and ERA) in the same specific instant time and regardless of direction, while the other two plots in the second row display the results of a simple linear regression model between synchronous peak storms within the directional sector 170 – 335 °N (which is the primary directional sector for the RON-Alghero wave buoy), and detected over a ±12 hours-time window. From all of this plots it can be seen as the ERA interim database produces an underestimation of the significant wave height with respect to the NOAA dataset (the normalised bias is always negative).

On the other hand, the other four plots in Fig. 3.13 show an underestimation of the NOAA data (best fit grid point 40,50 °N–08,00 °E) with respect to the RON significant wave height. The comparison is carried out for all data within the directional sector 170-335 °N, as well as for peak storms enclosed in two homogenous (220-275 °N and 275-335 °N) and in the total directional sector (170-335°N). Although results confirm the well-known underestimation ($BIASn < 0$) of the hindcast data in the Northern Mediterranean (e.g. Cavaleri et al. 2004), it is interesting to observe a fairly good relationship between NOAA and RON data, which, on the other hand, would seem more than comparable (and even with a better correlation) with similar results obtained by Bosom and Jiménez (2011) on grid-points closed to the Catalan Coast (NW Mediterranean).

Finally, numerical results for the most closed two or three NOAA grid points with respect to each of the three RON wave buoys are presented in Tab. 3.5-3.7. In Tab. 3.5, it can be observed a high correlation between NOAA and RON-Alghero datasets ($R \cong 0.84 \div 0.95$), although it can also be appreciated the inherent underestimation ($NBI < 0$). In this case, the best fitting is obtained by considering the 40,50 °N-08,00 °E grid point. On the other hand, the limited temporal extension of

both RON-Capo Comino (Tab. 3.6) and RON-Cagliari (Tab. 3.7) wave buoys generally reduce the number of concurrent events, though a high and good correlation with NOAA data can still be noticed in the first and second buoy, respectively. In addition, some positive bias can be generally observed in Tab. 3.7, for two of the three NOAA grid points closed to the RON-Cagliari wave buoy (i.e. 39,00 °N-09,33 °E and 39,00 °N-09,50 °E). Nevertheless, it should also be pointed out as the RON-Cagliari data (see Fig. 3.1) are likely to be influenced by the shelter effect, hence the associated wave transformation, caused by the Capo Carbonara headland.

Tab 3.5 Numerical results of the comparison between peak storms computed/recorded by NOAA and RON-Alghero wave buoy, within a ±12 hours-time window.

	275-335 °N			220-275 °N			170-335 °N		
	40.50°N 08.17 °E	40.50 °N 08.00 °E	40.67 °N 08.00 °E	40.50°N 08.17 °E	40.50 °N 08.00 °E	40.67 °N 08.00 °E	40.50°N 08.17 °E	40.50 °N 08.00 °E	40.67 °N 08.00 °E
n	328	383	383	43	49	46	377	440	440
R	0.9388	0.9492	0.9420	0.8435	0.8792	0.8708	0.9293	0.9484	0.9404
NBI	-0.3137	-0.0175	-0.1761	-0.1550	-0.0970	-0.0979	-0.3056	-0.1747	-0.1762
RMSE (m)	0.3187	0.3705	0.3840	0.4581	0.4518	0.4636	0.3400	0.3715	0.3852
SI	0.3376	0.2138	0.2211	0.2112	0.1739	0.1759	0.3347	0.2147	0.2214

Tab 3.6 Numerical results of the comparison between peak storms computed/recorded by NOAA and RON-Capo Comino wave buoy, within a ±12 hours-time window.

	350-55 °N		55-115 °N		115-170 °N		350-170 °N	
	40.67°N 10.00 °E	40.50 °N 10.00 °E	40.67°N 10.00 °E	40.50 °N 10.00 °E	40.67°N 10.00 °E	40.50 °N 10.00 °E	40.67°N 10.00 °E	40.50 °N 10.00 °E
n	13	13	2	3	10	10	28	28
R	0.8371	0.8491	-	0.9809	0.9255	0.9365	0.8674	0.8587
NBI	-0.1446	-0.1868	-	-0.2136	-0.1057	-0.0961	-0.1363	-0.1553
RMSE (m)	0.5558	0.5037	-	0.1807	0.2826	0.2697	0.4244	0.4112
SI	0.2105	0.2344	-	0.2173	0.1415	0.1344	0.1917	0.2067

Tab 3.7 Numerical results of the comparison between peak storms computed/recorded by NOAA and RON-Cagliari wave buoy, within a ±12 hours-time window.

	110-170 °N			200-265 °N			110-265 °N		
	39.00°N 09.17 °E	39.00 °N 09.33 °E	39.00 °N 09.50 °E	39.00°N 09.17 °E	39.00 °N 09.33 °E	39.00 °N 09.50 °E	39.00°N 09.17 °E	39.00 °N 09.33 °E	39.00 °N 09.50 °E
n	83	85	88	9	20	24	94	107	115
R	0.8716	0.8834	0.8865	0.6797	0.5605	0.5255	0.8828	0.8881	0.8709
NBI	-0.0272	0.0165	0.0341	-0.0545	0.0509	0.1296	-0.0286	0.0238	0.0555
RMSE (m)	0.1885	0.1921	0.2004	0.0912	0.1617	0.1973	0.1812	0.1844	0.2022
SI	0.1222	0.1208	0.1286	0.1076	0.1404	0.2054	0.1209	0.1239	0.1449

3.8 Correction of wind dataset

In this paragraph a summary of the wind data set correction is given. In Figs. 3.14 and 3.15 are depicted graphical results related to the correction of the wind data set. In Fig. 3.15 are displayed: the total number of anomalous values removed (red plus marker) and held (green plus markers), as defined after the graphical control.

An overall summary is given in Fig. 3.15 and Tab 3.8, where each station is identified by both WMO and ICAO code (see Tab 3.2) and are included: number of absolute wind calms (LVO, i.e. $\{U = 0; \theta_{wind} = 0\}$); number of relative wind calms (LV1, i.e $\{U\epsilon[0, U_1 = 1 \div 4\ kn \approx 0.5 \div$

$2\ m/s]\}$); number of variable directions (LD99, $\theta_{wind} = 99$); number of potential data (NDA); number of effective data (NDP); number of missing data (NDM); number of temporal gap (NDG); number of suspicious absolute calm removed (NDR1); number of suspicious directions removed (NDR2); number of errors (NDE); global station efficiency η_T; percentage of gaps η_G; percentage of removed data η_R; percentage of errors η_E.

Finally, the wind roses are displayed for each station considered in Figs. 3.16-3.18.

Tab 3.8. Final summary of the wind data correction

	16520 LIEA	16550 LIEB	16522 LIEH	16564 LIEC	16539 LIEF	16542 LIEL	16560 LIEE	16506 NULL	16531 LIEO
LV0	12380	14823	3936	7315	7176	3092	4274	985	1058
LV1	5017	552	538	247	368	2	919	226	1049
LD99	1481	147	0	50	84	62	21	5	0
NDA	177788	181143	83618	172477	148423	70434	71664	40626	30162
NDP	135568	163113	71535	117460	118489	14364	66409	27227	24369
NDM	42220	18030	12083	55017	29934	56070	5255	13399	5793
NDG	13890	11818	7224	49609	23252	50273	224	11812	2995
NDR1	24460	5798	3790	4268	6239	1188	4940	397	2777
NDR2	1707	0	0	3	0	0	0	0	0
NDE	2163	415	1071	1137	443	4609	91	1190	21
η_T	76.25%	90.05%	85.55%	68.10%	79.83%	20.39%	92.67%	67.02%	80.79%
η_G	7.81%	6.52%	8.64%	28.76%	15.67%	71.38%	0.31%	29.08%	9.93%
η_R	14.72%	3.20%	4.53%	2.48%	4.20%	1.69%	6.89%	0.98%	9.21%
η_E	1.22%	0.23%	1.28%	0.66%	0.30%	6.54%	0.13%	2.93%	0.07%

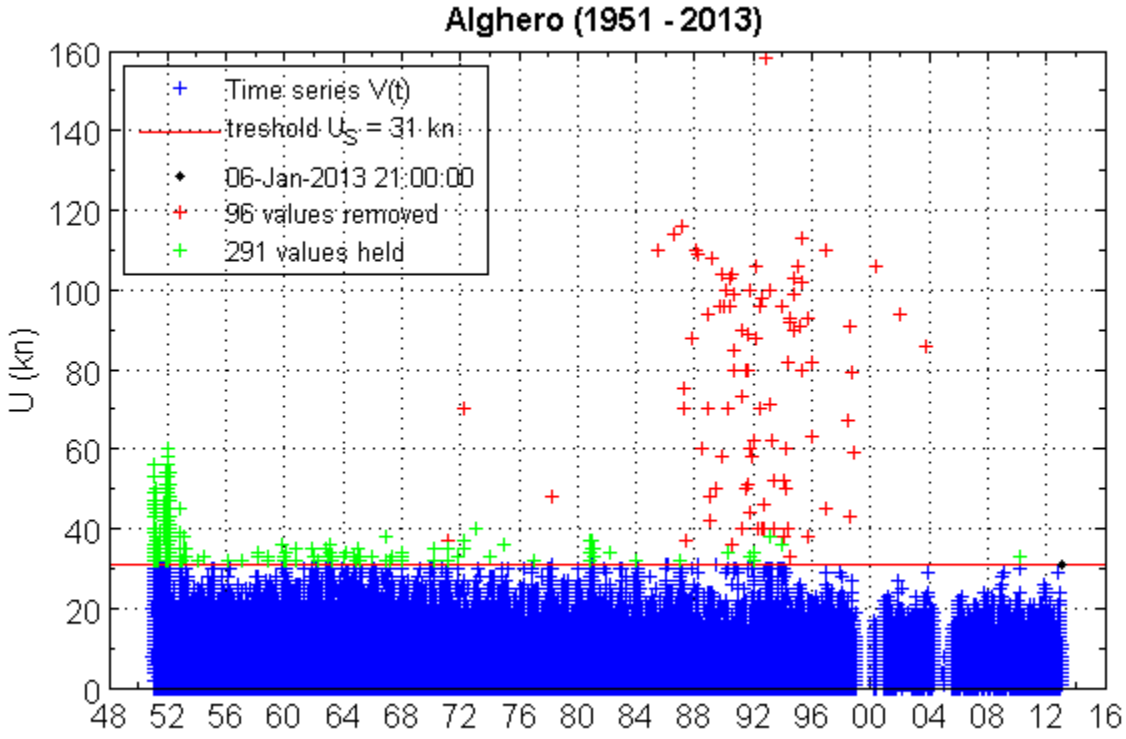

Fig. 3.14 UGM-ENAV Alghero. Global synthesis of the correction phase.

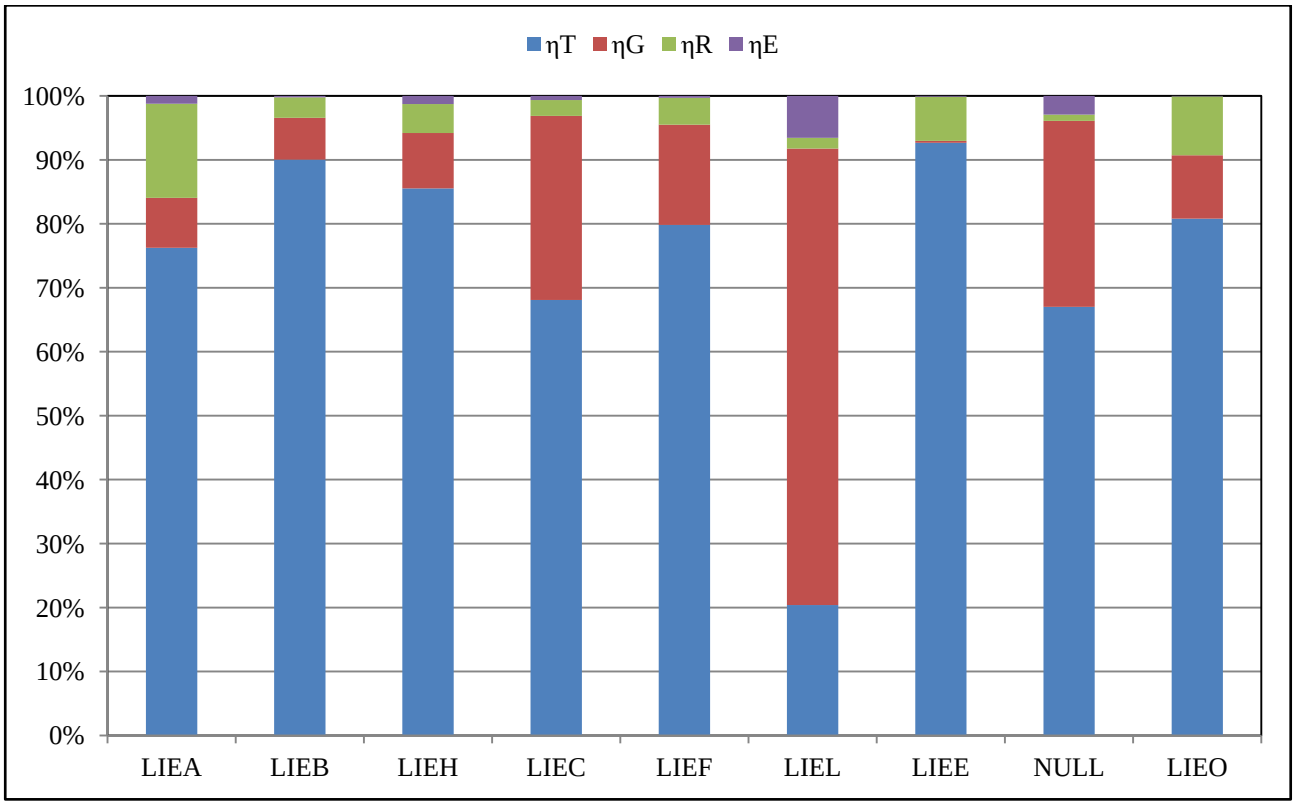

Fig. 3.15 Summary of wind data correction. Symbols on the horizontal axis are referred to the wind station-ICAO code (see Tab. 3.2).

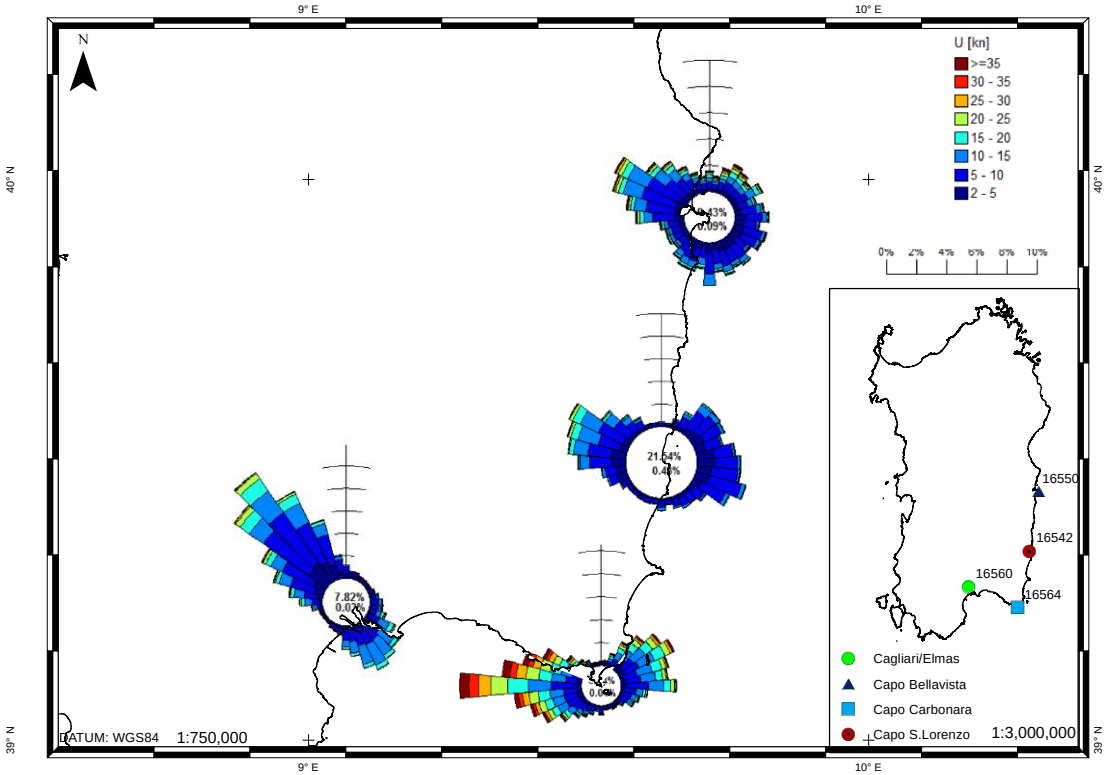

Fig. 3.16 Wind rose – South-East coast: Elmas, Capo Carbonara, Capo S. Lorenzo, Capo Bellavista

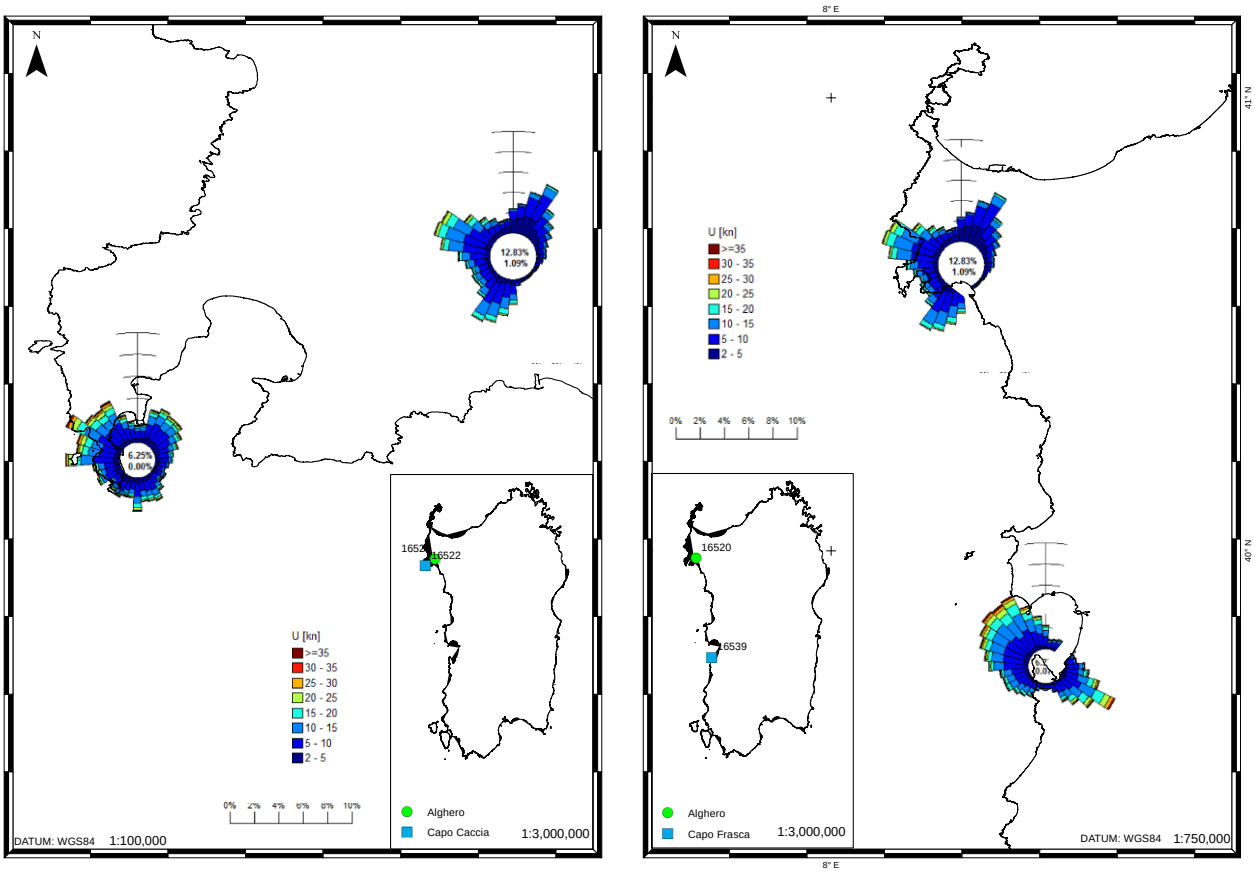

Fig. 3.17 Wind rose – Northern coast (left panel): Alghero, Capo Caccia; Northern – West coast (right panel): Alghero, Capo Frasca

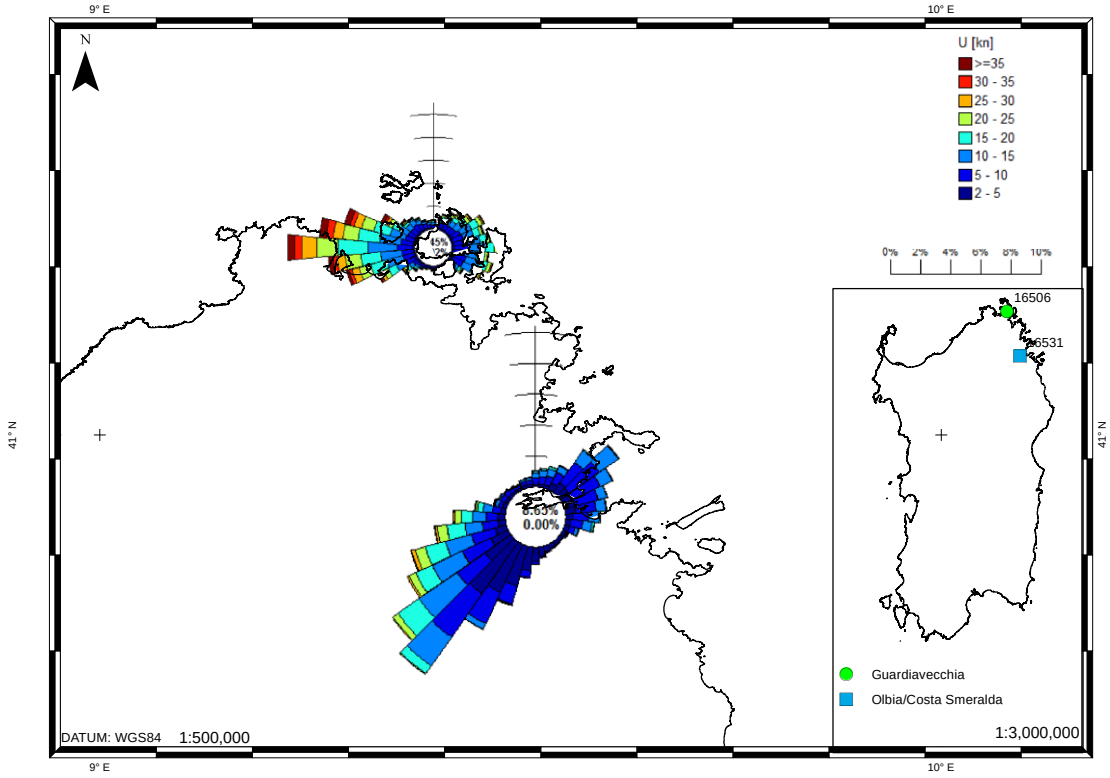

Fig. 3.18 Wind rose – Nort-East coast: Olbia, Guardiavecchia

4 Extreme Value model intercomparison

In this Chapter, the extreme wave analysis on data recorded by the RON Alghero buoy is going to be discussed. A comparison of principal best-known extreme value methodologies (as introduced in Chapter 2 and by extending the work of Ventroni 2016a) is carried out, with the cross-purpose of analysing the GP and WBL performances. In particular, both graphical and numerical results are provided, with the first one mainly focused on the significant wave height within the total directional sector (170-335) °N, while the second one are presented for each homogeneous sector.

4.1 Threshold selection methods

In Fig. 4.1, both the fixed-threshold selection methods and the MTM approach, with values computing by means of the ML method, are presented. The first group of subplots on the left, displays the fixed-threshold selection procedure adopted herein (§2.7). Starting from the upper panel, the following elements are shown: (i) the mean residual life plot, (ii) the shape parameter and (iii) the modified scale parameter (Eq. 2.26) as requested by the GP-parameter-stability; (iv) the Thompson et al. (2009)'s parameter (Eq. 2.27); (v) the mean rate, λ. All steps are depicted over a uniformly distributed range of thresholds, as suggested to compute (iv) (i.e. from the median, 3.30 m, till un upper value, 7.0 m). In each subplot, the vertical red line shows the threshold value ($u = 5.91$ m) corresponding to $\lambda = 5$ *ev/year* (Mazas and Hamm 2011). In addition, the vertical blue line represents the threshold value ($u = 3.30$ m) according to the automated threshold method (Thompson et al. 2009).

On the other hand, the right part of Fig. 4.1 depicts the MTM-hierarchical procedure (Deidda 2010) for a greater range of threshold values, i.e. u ∈ (0, 7), performed on a sample constituted by daily maximum wave heights. The first subplot from the top, shows the shape-parameter (ξ) estimates, where the ξ^M MTM estimate (=0.23) is the median value within a range of thresholds, which are selected according to a parameter-stability criterion (e.g. an interval from 2.9 till 6.1m seems to be the most suitable). The second and third subplots display the α_0 and ζ_0 estimates (Eqs. 2.30 and 2.31). In the fourth subplot the α_0^C estimates conditioned to the ξ^MMTM estimates is reported, where the horizontal line represents the median value α_0^MMTM (= 2.56) estimates. In the fifth subplot the ζ_0^Cestimates conditioned to both $\xi^M$and α_0^M estimates is depicted, where the ζ_0^M MTM estimate is the median of ζ_0^Cvalues. It is worthwhile underlying two main considerations.

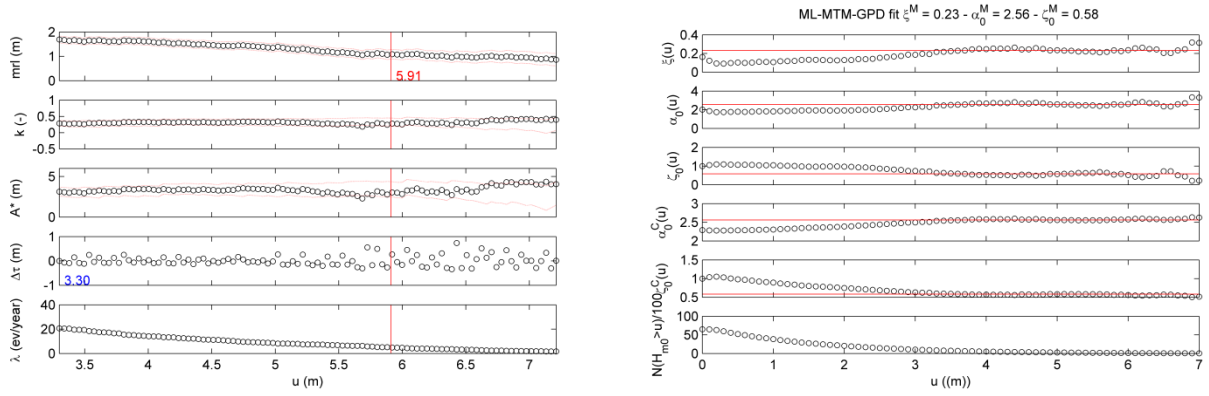

Fig. 4.1 RON Alghero buoy, sector (170-335) °N – ML method: example of threshold selection methods. The set of subplots on the left displays the fixed-threshold selection procedure adopted herein (§2.6), while the group of subplots on the right represents the MTM hierarchical procedure (§2.7).

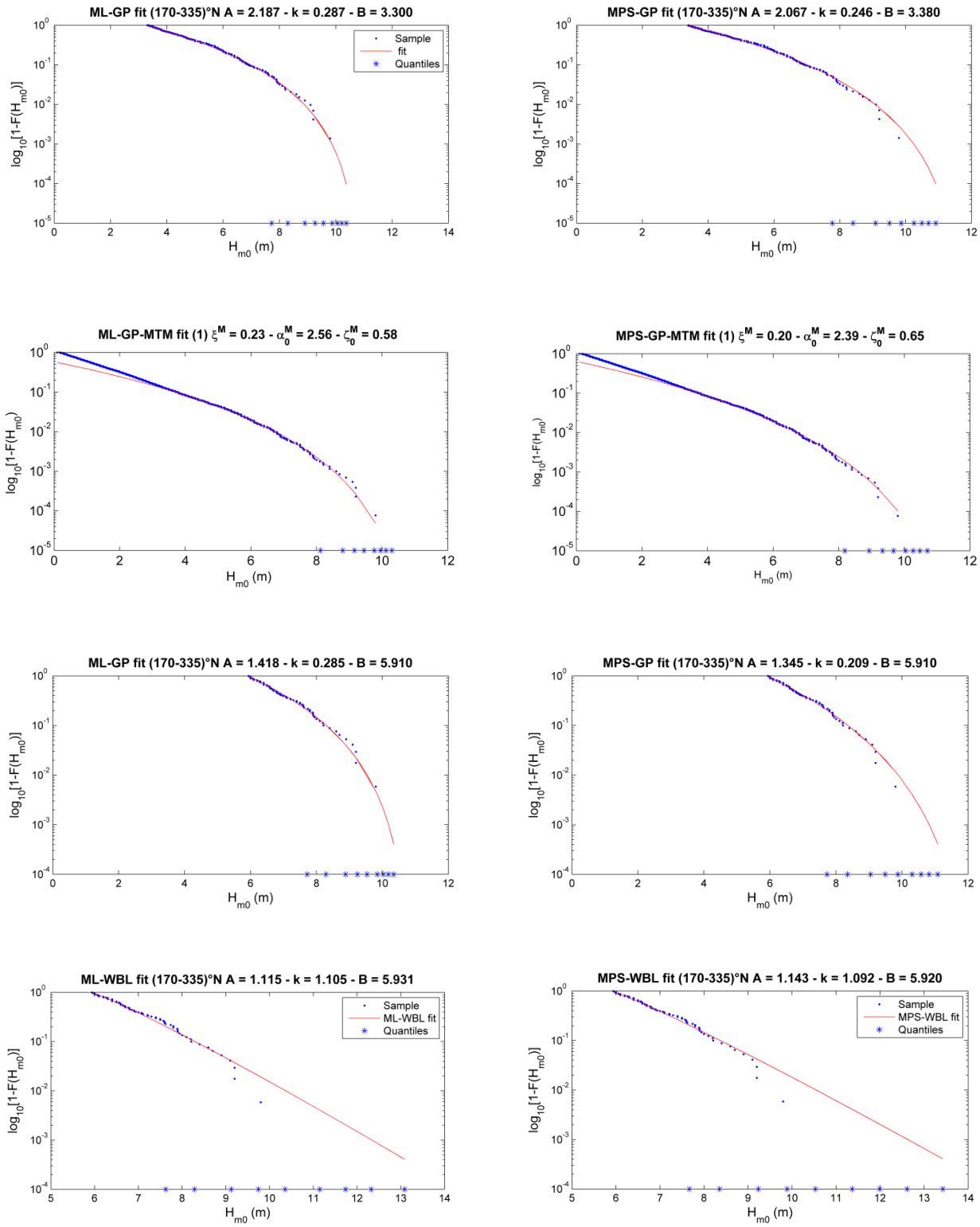

Fig. 4.2 RON Alghero buoy, sector (170-335) °N: summary of survival functions. Each figure depicts the sample (blue points), the best-fit model (red continuous line) and the values of return values for a set of 2, 5, 10, 20, 50, 100, 200, 500-yearly-return periods (blue cross markers) The group of figures on the left is related to the ML method, while the MPS method is considered on the right. The first row is obtained by applying the automated threshold selection method proposed by Thompson et al. (2008). The second row is obtained by using the MTM approach. The third row is related to the a threshold value equal to 5 storms per year, as well as the last row but with respect to WBL.

First, the MTM procedure requires to determine a practical range of thresholds to be used for computing the median values of the MTM estimators; usually, a different range must be prescribed for each homogeneous directional cluster. Second, a much higher CPU time is required in order to generate a set of (quasi-) random sample through the Monte Carlo method.

Fig. 4.2 displays a group of survival functions (i.e. the exceeding probability) obtained by using either the ML (group of figures on the left) or the MPS (group of figures on the right) estimators, for each fixed-threshold method and for the MTM approach. Specifically, the first row from the top shows the GP-ML and GP-MPS pair obtained by assuming a fixed threshold $u = 3.30$ m, i.e. according to (iv); the second group of figures represents the results of the MTM approach, namely GP-MTM-ML and GP-MTM-MPS; the third row depicts the GP-ML and GP-MTM results by fixing $u = 5.91$ m, whereby according to $\lambda = 5$ *ev/year*. Finally, the WBL-ML and WBL-MPS results are reported in the last row, from the sample selected with the same threshold as in the previous case. Basically, it should be stressed the different behaviour of the GP- and WBL-upper limit, which leads to obtain lower quantiles for the GP distribution.

4.2 Confidence bounds and return values.

Generally, results reported in this paragraph, which are mainly referred to the excesses above the threshold $u = 5.91$ m, may be split up in three parts. First of all, it is assumed that the threshold value selected is the most suitable to perform the comparison between several parameter estimation methods. As a matter of fact, it should be stressed that this may be carried out whether the bias inherent to the threshold selection is assumed to be negligible for each parameter estimation method, since the value of the threshold is selected over an ML- (or MPS-) estimators range. On the other hand, such fixed-threshold approach is assumed to be suitable to perform the GP-POI modelling (e.g. according to Mazas and Hamm 2011), as well as the comparison with the AM approach.

A second general overview of some fixed threshold values is depicted in Fig. 4.3. Basically, five values are shown, even if only two of them are used in this study; namely: the dispersion index (i.e. the ratio between the variance and the expectation of the number of peaks, u = 6.31 *m*); an automated mean residual life threshold (u = 5.80 m) – obtained by means of a simple algorithm taking into account the slope change - variance combination, i.e. similar to that proposed by Sartini et al. (2015) –; the highest threshold value corresponding to a mean rate $\lambda = 2$ *ev/year* (u = 6.95 m), whilst the other two are related to the algorithm of Thompson et al. (2009) (u = 3.30 m) and to $\lambda = 5$ *ev/year* (u = 5.91 m).

Consequently, a set of 8 combinations between parameter estimation methods – cumulative distribution functions are considered, namely: ML-GP, LM-GP, MPS-GP, PWM-GP and ML-WBL, LM-WBL, MPS-WBL, OLS-WBL, whether the threshold is fixed equal to the value, i.e. u = 5.91 *m*.

Return values corresponding to return period (*T*) and relative confidence intervals, about the total directional sector (170-335) °N, are shown in Fig. 4.4. Additionally, two type of confidence bounds are reported; both are based on 1000 synthetic samples provided by the Monte Carlo procedure. In the first method, the 90% confidence intervals are obtained by using percentile estimates adjusted

according to the recommendations of Coles and Simiu (2003), i.e. an ad hoc bias correction is applied. The second method represents the RMSE-curves computed from the same set of empirical samples. Principally, some general aspects can be qualitatively identified by analysing these results.

First, it is evident how the quantile functions depend on the parameter estimation methods, as well as on the plotting position formula adopted. It is also interesting to observe the bad performances of ML-WBL and MPS-WBL, which are probably due to either small sample size or because of the plotting position used. Evidently, the Hazen plotting position can be assumed a biased estimator for the WBL distribution. Notwithstanding, the Weibull plotting position is found to provide unsatisfactory results as well. On the whole, the WBL-OLS seems to be reasonable adequate to describe this data set, probably because of the unbiased plotting position used.

Second, a same order of uncertainty appears to be consistent with the pair of LM estimators, since comparable confidence bounds are provided by both GP-LM and WBL-LM.

Third, the GP distribution is upper limited in all cases, and a quantile-saturation feature can be noted on larger return periods. It is evident, consequently, that the GP-related values are lower than the values provided by the WBL distribution.

Additionally, the GP-POI model performance can be evaluated by giving a comparison with the GEV-AM approach. It should be remarked that the confidence intervals are computed according to Madsen et al. (1997), Goda (2011b) and Silva et al. (2012), since for this case it should be taken into account the λ-Poisson parameter uncertainty. A total of 8 plots is reported in Fig. 4.5, namely 4 combinations – ML-GP-GEV, LM-GP-GEV, MPS-GP-GEV, PWM-GP-GEV – for two directional sectors – (170-335) °N and (275-335) °N. Moreover, other 6 combinations are reported in Fig. 4.6 by applying the AM approach; i.e. ML-GEV, LM-GEV, MPS-GEV and ML-WBL, LM-WBL, OLS-WBL.

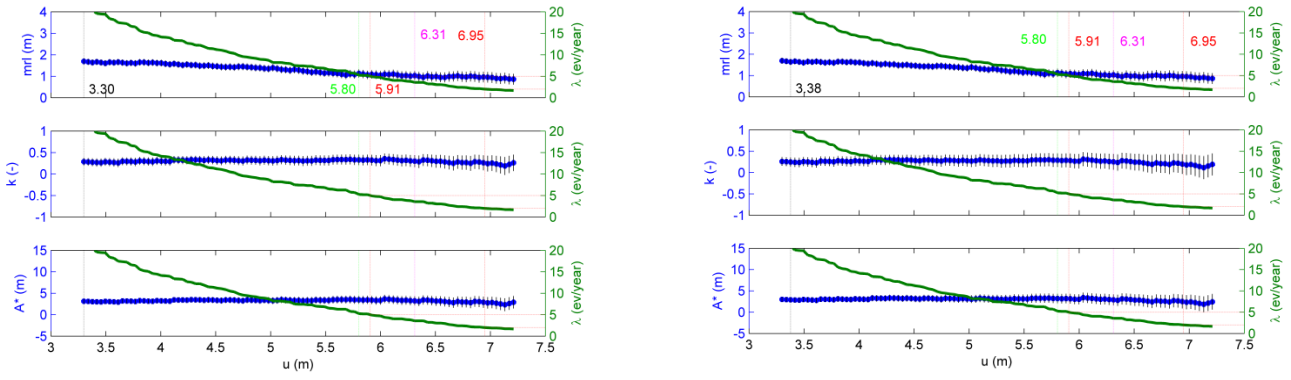

Fig. 4.3 Fixed-threshold methods - RON Alghero, Sector (170-335) °N. From left to right are shown values computed by ML and MPS method, respectively. Each group of figures depicts the mean residual life plot (upper panel), the modified scale parameter (middle panel) and the shape parameter (lower panel) with the associated 95% confidence interval, as well as the mean rate (continuous green line), measured by the vertical left axes values. Five vertical lines are shown in each panel, with values representing: i) the automated threshold selection of Thompson et al (2009) (dotted black colour); ii) an automated mrl value (dotted green colour); iii) the Poisson dispersion index (dotted magenta colour); iv) the mean rate equal to 5 and v) 2 (dotted red colours) ev/year.

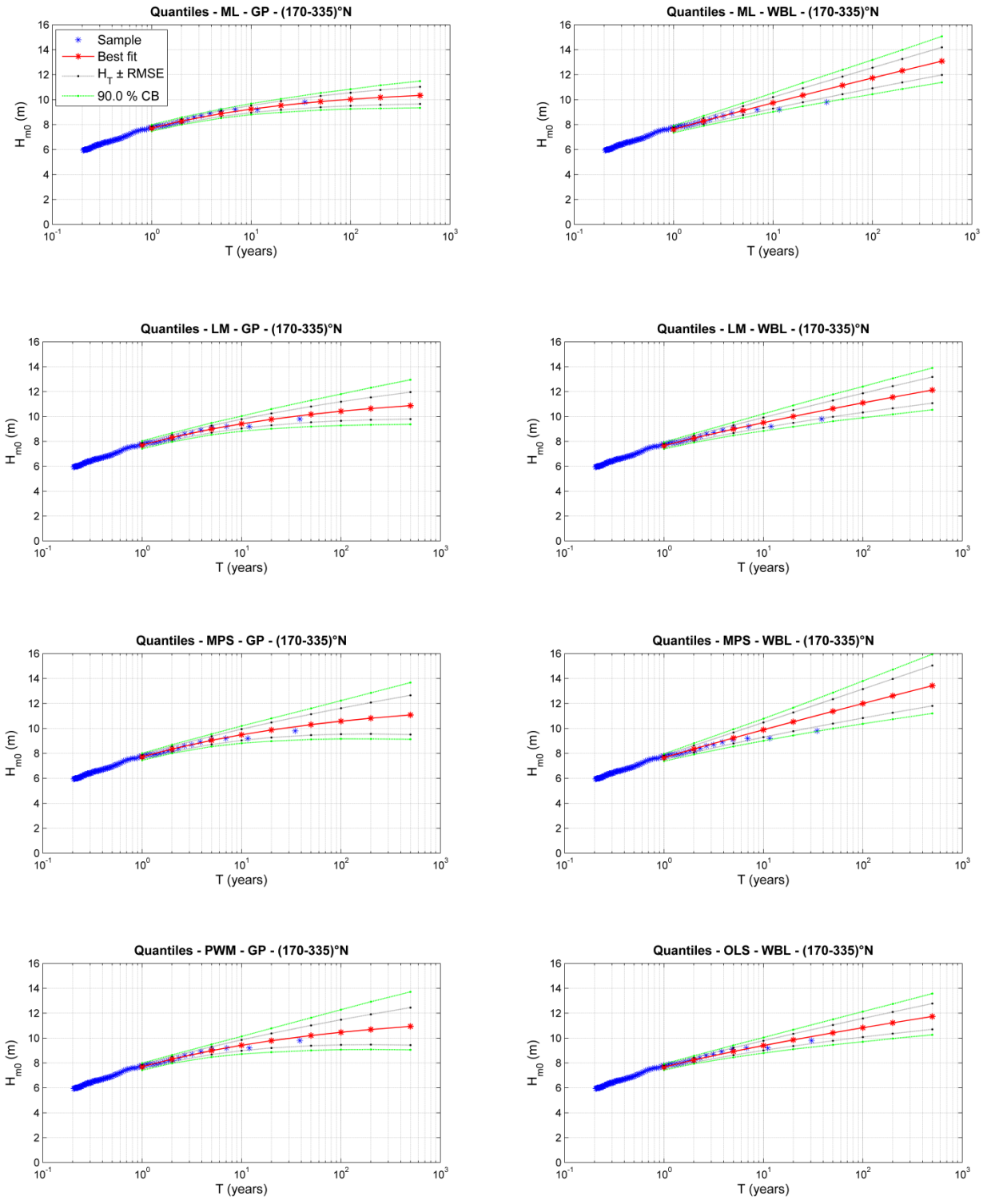

Fig. 4.4 RON Alghero buoy, sector (170-335) °N: summary of quantiles and confidence bounds for some estimators. The group of figure on the left is related to the GP results, whilst the WBL outcomes are displayed on the right.

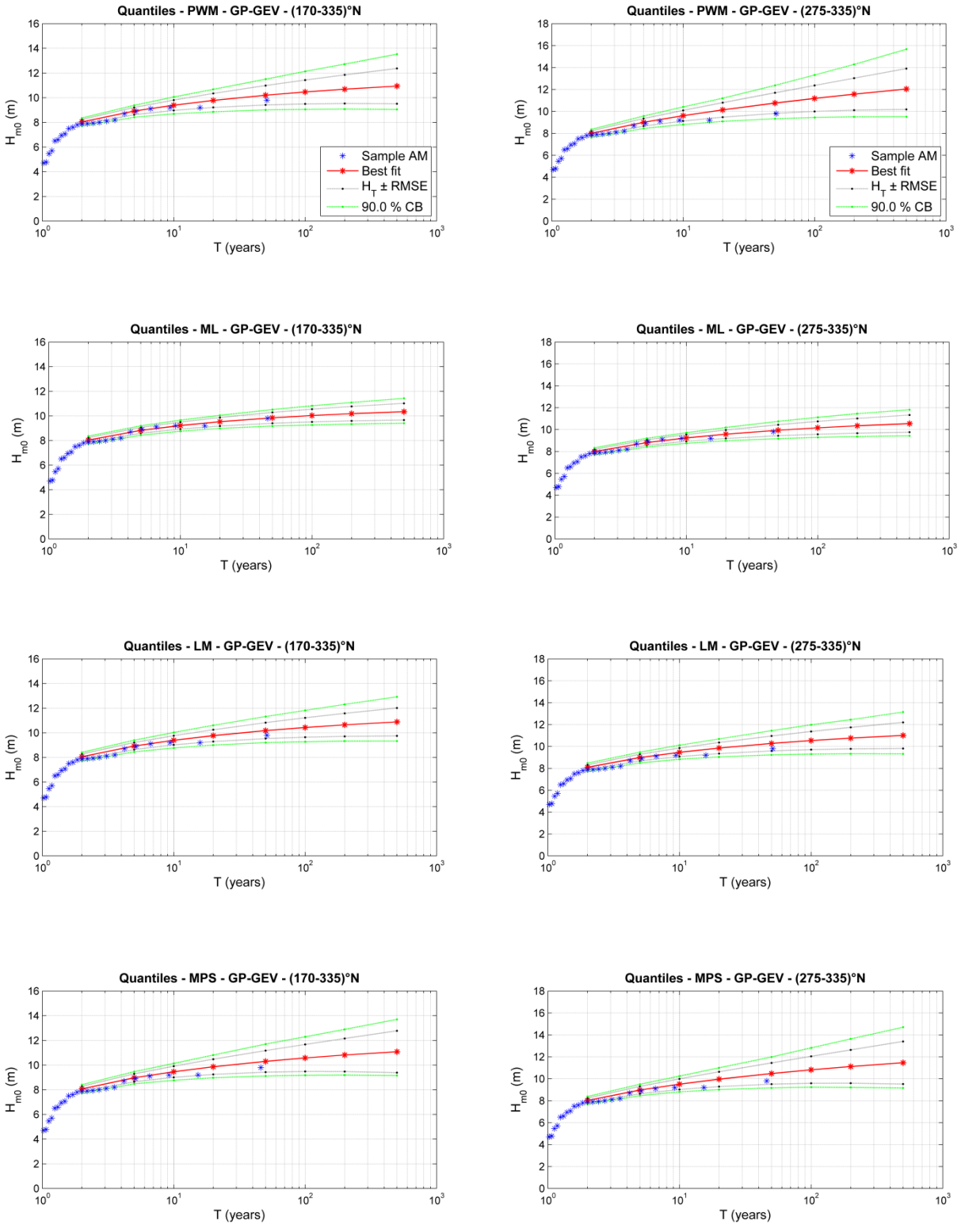

Fig. 4.5 RON Alghero buoy: summary of quantiles and confidence bounds for some estimators within the GP-POI model. On the left sector (170-335) °N results are reported, while the (275-335) °N outcomes are displayed on the right.

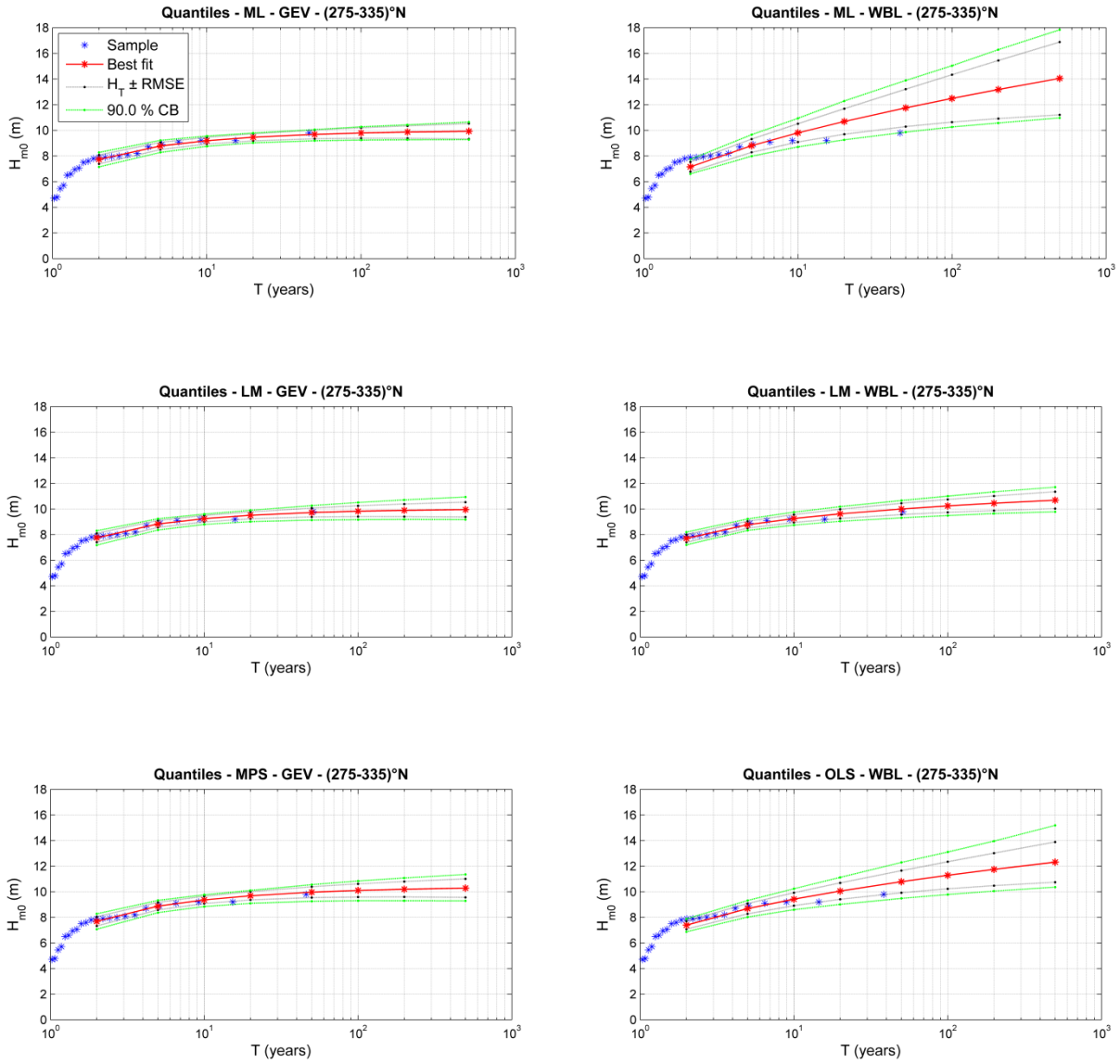

Fig. 4.6 RON Alghero buoy, AM approach: summary of quantile functions and confidence intervals. It is to be noted that the AM sample of the total sector (170-335) °N are coincident with the most energetic sector values, i.e.(270-335) °N. On the left are shown the GEV-AM results, while WBL outcomes are displayed on the right.

Starting from Fig. 4.5, it can be noted that the ML and LM methods are the best for each sector, since they yield narrower confidence intervals than those provided by the MPS-GP, PWM-GP cases. On the other hand, from Fig. 4.6, it may be observed as narrower confidence intervals are computed with the GEV-AM models. As in the previous case, the LM estimator denotes the most suitable set-up between GP and WBL CDFs.

Lastly, some further issues can be expressed with respect to estimator performances. Specifically, LM can be assumed the best reliable estimator, since is proved to be the less sensitive to the presence of outliers and to the sample size effect, as well as the one with the best numerical characteristics, due to its lowest CPU time and ease of implementation.

The OLS method is intended to represent the so-called Goda method (Goda 2010). Herein, a comparison between several methods can be defined, as given by Li et al. (2012) and by Sartini et al. (2015). However, in the present work, the Goda method is extended by including an optimization procedure based on the diagnostic Probability Plot Correlation Coefficient (Filliben 1973, 1975; PPCC), as applied by (Goda 2011a), in order to find the best shape parameter between a set of prescribed values (ranging from 0.5 till 2). Moreover, the confidence intervals are computed based on the Monte Carlo method, instead of using empirical formulas (e.g. Goda 2010). On the whole, the quantiles computed are found to be reasonable, as well as the ease of implementation is acceptable. However, the required CPU time to perform the Monte Carlo procedure is much higher than the LM case.

4.3 Comparison of best-fit methods

For sake of economy, only figures related to the LM estimator are reported, although same type of graphics are obtained for each combination. In Figs 4.7, 4.8 a summary of the PP-plot and QQ-plot, is shown. It can be seen as LM-GP and WBL-GM are equally valid and reasonable adequate to describe the RON-Alghero data set.

A quantitative summary is given in Tabs. 4.1, 4.2 for each CDF-estimator combination. Specifically, it is assumed to primarily focus the results of the Monte Carlo procedure, represented in terms of bias (Eq. 2.40) and root mean square error (Eq. 2.41), on the 50-year and 100-year quantile estimates, while minor attention should be attributed for greater quantiles because of low sample size and due to the clearly different behaviour of the CDF right tails.

In Tab. 4.1, GP-ML turns out to be the best for both AIC and BIC metrics, although slightly sensible difference may be observed with respect to the WBL values. Both MIR-QQ and MIR-PP diagnostic procedures lead to assume GP-LM as the best-fit model. Furthermore, the EN metrics provides the best-fit values to WBL-MBS ($h = 1.00$), and to WBL-LM ($h = 1.25$ and $h = 1.50$). Finally, GP-LM denotes the lowest bias, while RMSE values are the tightest for GP-POI-ML.

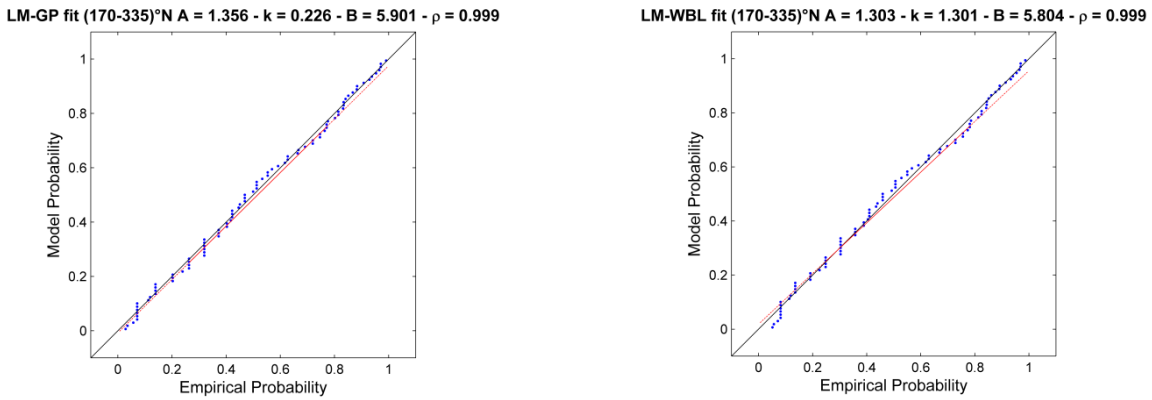

Fig. 4.7 RON Alghero buoy, Sector (170-335) °N: pp-plot.

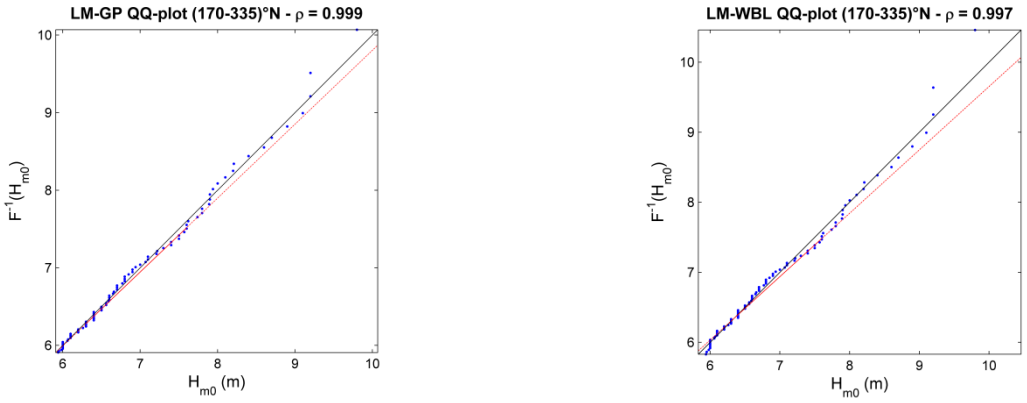

Fig. 4.8 RON Alghero buoy, Sector (170-335) °N: qq-plot.

In Tab. 4.2., GEV-ML outperforms the other two combinations according to AIC and BIC metrics. The diagnostic QQ-plot is in favour of GEV-LM, while both PP-plot and all EN values prescribe WBL-LM as the best-fit model. Finally, bias is lower in case of WBL-LM, whilst RMSE is in favour of GP-POI-ML.

In the light of these results, both models, GP and WBL, seems to be reasonable adequate to describe the extreme wave behaviour of this dataset. On balance, it may be assumed that: (a) LM is the best estimator; (b) WBL-LM and GP-LM are proved to be the best-fit model whether the EN metrics or the BIAS values are used to quantify the goodness of the analysis.

Tab 4.1 RON Alghero buoy, Sector (170-335) °N: summary of diagnostic and parametric best-fit metrics. Results obtained from a sample of excesses above the threshold u = 5.91 m.

u=5.91 m		GP-ML	GP-MPS	GP-LM	WBL-ML	WBL-MPS	WBL-LM	WBL-OLS
	AIC	186.851	187.474		187.252	188.408		
	BIC	194.179	194.802		194.580	195.736		
	MIR-QQ	0.0011	0.0014	0.0013	0.0061	0.0064	0.0033	0.0024
	MIR-PP	0.0011	0.0009	0.0009	0.0013	0.0012	0.0014	0.0019
EN, h=	1.00	0.0705	0.0690	0.0662	0.0579	0.0595	0.0602	0.0687
	1.25	0.0669	0.0640	0.0531	0.0599	0.0604	0.0523	0.0593
	1.50	0.0668	0.0636	0.0544	0.0637	0.0643	0.0542	0.0605
BIAS	H_{50}	-0.1398	0.3503	0.0154	-0.1398	0.3380	0.0491	0.0417
	H_{100}	-0.1553	0.4423	0.0335	-0.1553	0.4086	0.0600	0.0585
RMSE	H_{50}	0.4515	0.8385	0.6346	0.7067	0.9783	0.6559	0.6342
	H_{100}	0.5258	1.0401	0.7653	0.8223	1.1617	0.7707	0.7502

Tab 4.2 RON Alghero buoy, Sector (170-335) °N: summary of diagnostic and parametric best-fit metrics. Results obtained from a sample of excesses above the threshold u = 5.91 m (GP-POI) and from a sample of annual maxima (GEV and WBL).

AM		GP-POI-ML	GEV-ML	WBL-ML	GP-POI-MPS	GEV-MPS	GP-POI-LM	GEV-LM	WBL-LM	WBL-OLS
	AIC	186.851	55.349	58.097	187.474	56.204				
	BIC	194.179	58.183	60.931	194.802	59.037				
	MIR-QQ	0.0120	0.0066	0.0195	0.0168	0.0065	0.0164	0.0065	0.0063	0.0093
	MIR-PP	0.0078	0.0066	0.0145	0.0070	0.0063	0.0071	0.0063	0.0064	0.0074
EN, h=	1.00	0.5656	0.1018	0.1197	0.5543	0.0989	0.5475	0.0917	0.0912	0.1040
	1.25	0.5761	0.1074	0.1239	0.5889	0.1044	0.5760	0.0968	0.0952	0.1083
	1.50	0.5901	0.1137	0.1271	0.6042	0.1115	0.5905	0.1027	0.1006	0.1121
BIAS	H_{50}	-0.1373	-0.140	0.4751	0.3891	0.3045	0.0624	0.0563	0.0450	0.1351
	H_{100}	-0.1544	-0.141	0.6749	0.4924	0.4298	0.0973	0.1088	0.0676	0.2016
RMSE	H_{50}	0.4420	0.4608	1.0599	0.8738	0.8776	0.6482	0.5375	0.4916	0.6204
	H_{100}	0.5150	0.5978	1.3443	1.0961	1.2412	0.7879	0.7017	0.5760	0.7609

4.4 Final summary

In this paragraph, a final quantitative summary is reported in Tabs. $4.3 \div 4.16$. The set of parameters and quantile estimations, as well as bias and RMSE are given for each combination herein considered and for each homogenous sector.

A summary of the best-fit metrics for both the homogeneous sectors (275-335) °N and (220.275) °N is provided in Tabs. $4.9 \div 4.16$.

It is found that both GP-LM and WBL-LM are reasonable valid as the best-fit model for the most energetic sector (Tab. 4.9). Generally, GP-LM can be considered the best model for both MIR-PP and MIR-QQ parameters and, in addition, is characterized by the less RMSE values, while WBL-LM leads to the less biased quantiles.

On the other hand, a different behaviour is evident on the other directional sector (Tab. 4.13), since WBL-OLS may be assumed as the best-fit model for this case. In addition, it is to be noted that an exponential distribution is probably the true distribution for this sector (i.e. the WBL shape parameter is approximately equal to one, while the GP shape parameter is almost equal to or less than zero).

Tab 4.3 RON Alghero buoy, Sector (170-335) °N: summary of parameters and quantiles estimations. POT sample with u = 5.91m.

u=5.91m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	8.29	8.28	8.35	8.35	8.32	8.26	8.25
5	8.88	9.12	9.05	9.24	8.99	8.99	8.92
10	9.24	9.74	9.49	9.89	9.41	9.51	9.40
20	9.54	10.35	9.88	10.53	9.77	10.01	9.85
50	9.85	11.15	10.31	11.37	10.17	10.64	10.42
100	10.03	11.74	10.58	11.99	10.42	11.10	10.83
200	10.18	12.32	10.82	12.61	10.64	11.55	11.23
500	10.34	13.09	11.08	13.42	10.88	12.13	11.74
A	1.42	1.11	1.34	1.14	1.36	1.30	1.40
k	0.29	1.10	0.21	1.09	0.23	1.30	1.41
B	5.91	5.93	5.91	5.92	5.90	5.80	5.74

Tab 4.4 RON Alghero buoy, Sector (170-335) °N: summary of quantiles and parameters bias. POT sample with u = 5.91m.

u=5.91m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	-0.0213	-0.0049	0.0497	0.0688	0.0018	0.0147	-0.009
5	-0.0596	0.0017	0.1104	0.1333	-0.0047	0.0214	-0.0005
10	-0.0877	0.0094	0.1705	0.1893	-0.0050	0.0282	0.0096
20	-0.1128	0.0189	0.2415	0.2504	0.0000	0.0364	0.0221
50	-0.1398	0.0339	0.3503	0.3380	0.0154	0.0491	0.0417
100	-0.1553	0.0469	0.4423	0.4086	0.0335	0.0600	0.0585
200	-0.1667	0.0611	0.5415	0.4826	0.0567	0.0718	0.0767
500	-0.1763	0.0816	0.6822	0.5850	0.0945	0.0888	0.1029
A	0.0530	-0.0143	-0.0094	0.0042	0.0300	0.0267	0.0027
k	0.0404	-0.0034	-0.0295	-0.0227	0.0146	0.0327	0.0250
B		0.0095		0.0007		-0.0163	-0.0114

Tab 4.5 RON Alghero buoy, Sector (170-335) °N: summary of quantiles and parameters RMSE. POT sample with u = 5.91m.

u=5.91m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	0.1807	0.2428	0.2197	0.2777	0.2179	0.2163	0.1972
5	0.2290	0.3590	0.3285	0.4440	0.2881	0.3186	0.2959
10	0.2859	0.4566	0.4505	0.5903	0.3710	0.4108	0.3876
20	0.3541	0.5607	0.6020	0.7501	0.4752	0.5120	0.4891
50	0.4515	0.7067	0.8385	0.9783	0.6346	0.6559	0.6342
100	0.5258	0.8223	1.0401	1.1617	0.7653	0.7707	0.7502
200	0.5982	0.9418	1.2588	1.3532	0.9012	0.8897	0.8704
500	0.6893	1.1049	1.5725	1.6171	1.0853	1.0521	1.0347
A	0.2175	0.1154	0.1910	0.1261	0.2260	0.1897	0.2144
k	0.1143	0.0848	0.1138	0.1134	0.1180	0.2014	0.2467
B		0.0200		0.0211		0.1049	0.1414

Tab 4.6 RON Alghero buoy, Sector (170-335) °N: summary of parameter and quantiles estimations. AM sample (GEV and WBL) and POT sample (GP-ML) with u = 5.91m, , while the interval Δu=(2.9,6.1) m is used for the GP-MTM model.

	ML				MPS			LM			OLS
AM	GP-POI	GEV	GP-MTM	WBL	GP-POI	GEV	GP-MTM	GP-POI	GEV	WBL	WBL
2	8.04	8.05	8.12	7.79	8.07	8.02	8.19	8.05	7.63	7.63	7.57
5	8.82	8.81	8.80	8.83	8.97	8.93	8.93	8.93	8.65	8.66	8.60
10	9.22	9.20	9.16	9.49	9.46	9.40	9.33	9.39	9.14	9.13	9.14
20	9.53	9.49	9.45	10.10	9.86	9.79	9.67	9.77	9.52	9.50	9.60
50	9.84	9.79	9.76	10.84	10.30	10.20	10.04	10.18	9.92	9.91	10.12
100	10.03	9.97	9.95	11.37	10.58	10.45	10.27	10.43	10.16	10.17	10.47
200	10.18	10.11	10.12	11.87	10.82	10.66	10.47	10.65	10.37	10.41	10.80
500	10.34	10.25	10.30	12.51	11.08	10.90	10.70	10.88	10.59	10.69	11.19
A	0.90	0.889		1.676	0.96	0.993		0.95	0.927	2.997	2.027
k	0.29	0.298		1.429	0.21	0.238		0.23	0.247	3.109	2.000
B	7.72	7.740		6.491	7.73	7.675		7.72	7.701	5.368	6.254

Tab 4.7 RON Alghero buoy, Sector (170-335) °N: summary of parameter and quantiles bias. AM sample (GEV and WBL) and POT sample (GP-ML) with u = 5.91m, while the interval Δu=(2.9,6.1) m is used for the GP-MTM model.

	ML				MPS			LM			OLS
AM	GP-POI	GEV	GP-MTM	WBL	GP-POI	GEV	GP-MTM	GP-POI	GEV	WBL	WBL
2	-0.004	0.035	-0.035	-0.090	0.035	-0.013	-0.036	-0.014	-0.002	-0.021	-0.042
5	-0.049	-0.036	-0.057	-0.031	0.113	0.053	-0.060	-0.008	-0.016	-0.016	-0.034
10	-0.081	-0.083	-0.071	0.086	0.184	0.107	-0.076	0.005	-0.011	-0.001	0.005
20	-0.108	-0.117	-0.084	0.238	0.266	0.178	-0.090	0.025	0.009	0.017	0.055
50	-0.137	-0.140	-0.099	0.475	0.389	0.305	-0.107	0.062	0.056	0.045	0.135
100	-0.154	-0.141	-0.109	0.675	0.492	0.430	-0.118	0.097	0.109	0.068	0.202
200	-0.167	-0.128	-0.118	0.888	0.604	0.584	-0.128	0.137	0.175	0.091	0.272
500	-0.178	-0.089	-0.128	1.188	0.762	0.842	-0.140	0.198	0.282	0.123	0.370
A	-0.026	-0.022	0.012	-0.275	0.040	0.074	0.012	-0.003	-0.015	0.688	-0.251
k	0.043	0.078	0.066	-0.263	-0.034	0.004	0.064	0.001	-0.001	0.942	-0.234
B	0.007	0.048	-0.013	0.187	0.018	-0.037	-0.014	-0.014	0.004	-0.710	0.203

Tab 4.8 RON Alghero buoy, Sector (170-335) °N: summary of parameter and quantiles RMSE. AM sample (GEV and WBL) and POT sample (GP-ML) with u = 5.91m, while the interval Δu=(2.9,6.1) m is used for the GP-MTM model.

	ML				MPS			LM			OLS
AM	GP-POI	GEV	GP-MTM	WBL	GP-POI	GEV	GP-MTM	GP-POI	GEV	WBL	WBL
2	0.198	0.255	0.146	0.237	0.217	0.263	0.1593	0.224	0.247	0.254	0.228
5	0.234	0.252	0.203	0.365	0.323	0.285	0.2270	0.288	0.264	0.278	0.298
10	0.284	0.270	0.247	0.519	0.452	0.363	0.2790	0.372	0.297	0.324	0.367
20	0.348	0.328	0.290	0.726	0.616	0.530	0.3311	0.480	0.373	0.389	0.461
50	0.442	0.461	0.344	1.060	0.874	0.878	0.3986	0.648	0.538	0.492	0.620
100	0.515	0.598	0.384	1.344	1.096	1.241	0.4480	0.788	0.702	0.576	0.761
200	0.587	0.768	0.420	1.650	1.340	1.712	0.4954	0.935	0.895	0.663	0.915
500	0.678	1.059	0.465	2.079	1.696	2.567	0.5544	1.135	1.190	0.781	1.135
A	0.085	0.177	0.038	0.381	0.111	0.215	0.0385	0.100	0.168	2.523	0.486
k	0.115	0.255	0.232	0.315	0.115	0.234	0.2320	0.117	0.192	3.377	0.407
B	0.191	0.249	0.077	0.235	0.199	0.258	0.0942	0.211	0.240	2.457	0.378

Tab 4.9 RON Alghero buoy, Sector (275-335) °N: summary of diagnostic and parametric best-fit metrics. Results are obtained from a sample of excesses above the threshold u = 5.88 m.

	u=5.88 m	GP-ML	GP-MPS	GP-LM	WBL-ML	WBL-MPS	WBL-LM	WBL-OLS
	AIC	189.62	189.04		192.09	198.86		
	BIC	196.43	197.01		199.49	200.26		
	MIR-QQ	0.0204	0.0024	0.0015	0.0086	0.0096	0.0036	0.0027
	MIR-PP	0.0062	0.0010	0.0010	0.0017	0.0016	0.0016	0.0023
EN, h=	1.00	0.5612	0.0622	0.0527	0.0588	0.0587	0.0526	0.0584
	1.25	0.5995	0.0617	0.0540	0.0621	0.0626	0.0544	0.0596
	1.50	0.6333	0.0635	0.0570	0.0656	0.0668	0.0574	0.0618
BIAS	H_{50}	-0.1562	0.4274	0.0476	-0.0345	0.3635	-0.0127	0.035
	H_{100}	-0.1769	0.5476	0.0806	-0.0370	0.4436	-0.0076	0.048
RMSE	H_{50}	0.4823	0.9482	0.0476	0.8067	1.1035	0.6678	0.630
	H_{100}	0.5650	1.2045	0.0806	0.9435	1.3193	0.7871	0.744

Tab 4.10 RON Alghero buoy, Sector (275-335) °N: summary of quantiles and parameters estimations. POT sample with u = 5.88m.

u=5.88m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	8.25	8.31	8.32	8.39	8.32	8.26	8.24
5	8.88	9.23	9.06	9.38	9.01	9.01	8.93
10	9.26	9.93	9.55	10.12	9.44	9.54	9.41
20	9.59	10.61	9.98	10.86	9.81	10.05	9.87
50	9.94	11.51	10.49	11.84	10.22	10.70	10.45
100	10.16	12.19	10.82	12.57	10.48	11.18	10.87
200	10.35	12.86	11.12	13.31	10.71	11.64	11.27
500	10.55	13.75	11.47	14.28	10.95	12.24	11.79
A	1.35	1.09	1.27	1.11	1.38	1.33	1.44
k	0.25	1.04	0.17	1.01	0.22	1.30	1.42
B	5.88	5.88	5.88	5.88	5.82	5.72	5.64

Tab 4.11 RON Alghero buoy, Sector (275-335) °N: summary of quantiles and parameters bias. POT sample with u = 5.88m.

u=5.88m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	-0.0205	-0.0132	0.0579	0.0681	-0.0150	-0.0171	-0.003
5	-0.0616	-0.0210	0.1296	0.1367	-0.0126	-0.0197	0.003
10	-0.0932	-0.0260	0.2020	0.1977	-0.0029	-0.0194	0.011
20	-0.1227	-0.0301	0.2896	0.2653	0.0141	-0.0175	0.020
50	-0.1562	-0.0345	0.4274	0.3635	0.0476	-0.0127	0.035
100	-0.1769	-0.0370	0.5476	0.4436	0.0806	-0.0076	0.048
200	-0.1934	-0.0388	0.6808	0.5284	0.1195	-0.0013	0.062
500	-0.2092	-0.0402	0.8757	0.6469	0.1787	0.0086	0.082
A	0.0532	-0.0036	-0.0111	0.0019	0.0043	0.0138	0.0164
k	0.0412	0.0111	-0.0337	-0.0206	0.0033	0.0349	0.0331
B	0.0000	0.0036	0.0000	0.0007	0.0000	-0.0206	-0.0179

Tab 4.12RON Alghero buoy, Sector (275-335) °N: summary of quantiles and parameters RMSE. POT sample with u = 5.88m.

u=5.88m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	0.1834	0.2640	0.2245	0.2994	0.8277	0.2137	0.203
5	0.2377	0.3987	0.3421	0.4868	-0.0126	0.3185	0.300
10	0.3001	0.5126	0.4804	0.6536	-0.0029	0.4137	0.389
20	0.3748	0.6347	0.6586	0.8377	0.0141	0.5184	0.488
50	0.4823	0.8067	0.9482	1.1035	0.0476	0.6678	0.630
100	0.5650	0.9435	1.2045	1.3193	0.0806	0.7871	0.744
200	0.6463	1.0855	1.4910	1.5465	0.1195	0.9108	0.862
500	0.7494	1.2801	1.9158	1.8624	0.1787	1.0799	1.024
A	0.2069	0.1162	0.1847	0.1241	0.2214	0.1916	0.2288
k	0.1115	0.0806	0.1117	0.0973	0.1176	0.2036	0.2413
B	0.0000	0.0141	0.0000	0.0158	0.0000	0.1112	0.1542

Tab 4.13RON Alghero buoy, Sector (220-275) °N: summary of diagnostic and parametric best-fit metrics. Results obtained from a sample of excesses above the threshold u = 3.05 m.

	u=3.05 m	GP-ML	GP-MPS	GP-LM	WBL-ML	WBL-MPS	WBL-LM	WBL-OLS
	AIC	79.79	80.38		79.52	80.46		
	BIC	84.37	84.10		84.96	85.04		
	MIR-QQ	0.0035	0.0105	0.0054	0.0059	0.0079	0.0052	0.0035
	MIR-PP	0.0031	0.0029	0.0030	0.0032	0.0035	0.0030	0.0043
EN, h=	1.00	0.7092	0.0887	0.0790	0.0866	0.0866	0.0823	0.0899
	1.25	0.6127	0.0892	0.0793	0.0920	0.0919	0.0818	0.0882
	1.50	0.5384	0.0933	0.0827	0.0974	0.0977	0.0843	0.0898
BIAS	H_{50}	-0.279	2.725	0.105	-0.048	1.266	0.0215	-0.016
	H_{100}	-0.309	4.708	0.237	-0.042	1.658	0.0581	0.000
RMSE	H_{50}	1.351	5.854	1.620	1.298	2.613	1.4704	1.2286
	H_{100}	1.854	10.356	2.303	1.586	3.394	1.8756	1.5380

Tab 4.14 RON Alghero buoy, Sector (220-275) °N: summary of parameters and quantiles estimations. POT sample with u = 3.05 m.

u=3.05m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	4.56	4.53	4.56	4.60	4.55	4.55	4.61
5	5.52	5.62	5.76	5.87	5.61	5.61	5.61
10	6.22	6.48	6.77	6.89	6.41	6.42	6.34
20	6.91	7.36	7.88	7.94	7.23	7.23	7.05
50	7.78	8.53	9.53	9.39	8.31	8.30	7.97
100	8.42	9.44	10.92	10.51	9.14	9.11	8.65
200	9.04	10.35	12.46	11.66	9.98	9.93	9.33
500	9.83	11.57	14.74	13.21	11.09	11.01	10.21
A	1.14	1.05	1.01	1.09	1.13	1.13	1.28
k	0.04	0.92	-0.14	0.86	-0.01	0.99	1.11
B	3.05	3.05	3.05	3.04	3.00	3.00	2.91

Tab 4.15 RON Alghero buoy, Sector (220-275) °N: summary of parameters and quantiles bias. POT sample with u = 3.05m.

u=3.05m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	0.017	-0.016	0.056	0.089	0.020	-0.0047	0.005
5	-0.057	-0.036	0.262	0.303	0.004	-0.0206	-0.017
10	-0.130	-0.045	0.607	0.533	0.000	-0.0209	-0.025
20	-0.202	-0.049	1.223	0.816	0.019	-0.0100	-0.026
50	-0.279	-0.048	2.725	1.266	0.105	0.0215	-0.016
100	-0.309	-0.042	4.708	1.658	0.237	0.0581	0.000
200	-0.305	-0.032	7.879	2.093	0.450	0.1053	0.023
500	-0.227	-0.013	15.103	2.731	0.895	0.1833	0.062
A	0.103	-0.020		0.034		0.0456	0.076
k	0.090	0.013		-0.038	0.031	0.083	0.110
B		0.013		0.0019		-0.041	-0.061

Tab 4.16RON Alghero buoy, Sector (220-275) °N: summary of parameters and quantiles RMSE. POT sample with u = 3.05m.

u=3.05m	GP-ML	WBL-ML	GP-MPS	WBL-MPS	GP-LM	WBL-LM	WBL-OLS
2	0.266	0.283	0.311	0.332	0.312	0.278	0.275
5	0.398	0.508	0.699	0.727	0.483	0.479	0.442
10	0.579	0.715	1.380	1.171	0.685	0.711	0.631
20	0.849	0.950	2.643	1.725	0.994	1.004	0.865
50	1.351	1.298	5.854	2.613	1.620	1.470	1.229
100	1.854	1.586	10.356	3.394	2.303	1.876	1.538
200	2.487	1.893	18.108	4.263	3.218	2.321	1.872
500	3.587	2.326	37.893	5.542	4.901	2.968	2.348
A	0.346	0.209	0.288	0.239	0.336	0.323	0.370
k	0.233	0.110	0.265	0.143	0.209	0.314	0.342
B		0.028		0.028		0.161	0.213

4.5 Conclusions

Firstly, some general issues can be expressed with respect to the threshold selection methods adopted herein. Since the shape parameter of the GP distribution is generally always positive, i.e. the distribution is upper limited, no significant differences are found on quantile estimation provided by each methodology; although an exception is given by the (220-275) °N sector, where an exponential distribution seems to be fairly adequate to describe the associated extreme behaviour. Consequently, the comparison may be focused on describing some related numerical features, e.g. the ease of implementation or either the computational burden, as well as on practical characteristic, such as the ability of being automated. In particular, it is found that the MTM method is the less favourable on the side of computational time, at least with respect to the MPS and ML methods. Probably, an optimization to reduce CPU time may be achieved by implementing the MTM approach with other estimators (e.g. the LM method, as done by Deidda 2010 and Hellies 2016). On the other hand, even from the practical perspective, related to the coastal dataset-modelling process, it is found that a range of thresholds have to be fixed for each directional sector; hence, the applicability would seem relatively reduced in case of multiple applications on large dataset.

Secondly, regarding the parameter estimation method, it seems to be more convenient to choose the LM estimator since it provides the better statistical justification and the less differences in the quantile estimation between the two probabilistic models (GP or WBL). Basically, it is a confirm of the low sample size peculiarity, as reported in Akram and Hayat (2014) and previously stated (in §1.1.4). In addition, it is proved to be more robust than the ML or MPS methods with respect to the presence of potential outliers. Last but not least, it is relatively simple to implement with respect to others methods, and, most of all, it requires less CPU time to carry out the Monte Carlo approach.

In the light of the results obtained, it seems not to be straightforward providing a final answer about the most suitable probabilistic model representation (i.e. GP or WBL), or better, it should not be expressed a priori. The GP-POT approach is clearly more attractive on several perspectives, ranging from theoretical to analytical one; nevertheless, the lack of the asymptotic validity, the wider sensitivity of parameter estimation, and, last but not least, the different behaviour of the right tail inherent to both models, required to address the answer into the practical, or engineering, area: namely, choosing the safest results is more favourable than assuming the best-fit model exclusively on statistical ground. Surely, the lower the simple size, the greater is the uncertainty and the statistical metrics are not sufficient to achieve a final answer. Since the sample size are lower, much more a multi-distribution approach is favourable.

In conclusion, if one had to choose a larger quantile, it would be more close to the engineering practice, where a safety reason should be followed, because neither the best-fit model nor the statistical test have demonstrated an evident outperform of one model with respect to the other (i.e. according to the advice of Mazas and Hamm 2011 and Goda et al 2010).

5 Coupled numerical model

In this chapter, an overview of the governing equations that can be solved by the *SWAN* and *SWASH* programs will be briefly given; then, the main elements of the coupled model will be introduced and the validation with respect to laboratory data will be provided.

5.1 SWAN model

Simulating WAves Nearshore (*SWAN*, Holtuijsen 2007) is a third-generation spectral model that solves the spectral density evolution equation (Booij et al. 1999, Ris et al. 1999). *SWAN* computes random, short-crested, wind-generated waves in coastal regions and it is driven by imposing boundary condition provided by local winds and wave input.

The model describes the evolution of the action density, defined as the ratio of the energy density spectrum to the relative angular frequency ($N(\omega,\theta) = E(\omega,\theta)/\omega$), numerically solving with an implicit scheme the spectral action balance equation, which in Cartesian co-ordinates reads (Booij et al. 1999):

$$\frac{\partial N}{\partial t} + \nabla\cdot\left[\left(\vec{c}_g+\vec{V}\right)N\right] + \frac{\partial c_\omega N}{\partial \omega} + \frac{\partial c_\theta N}{\partial \theta} = S_{tot}, \tag{5.1}$$

where ω and θ are the wave angular frequency and the wave propagation direction, respectively. The left-hand side is the kinematic part of this equation. The first term in the left-hand side is the local rate of change of wave action density spectrum in time; the second term denotes the propagation of wave energy in two-dimensional physical space, with c_g the group velocity and V the ambient current; the other two terms represent the energy shift due to depth-induced and current-induced refraction with propagation velocity c_ω, c_θ in the spectral ω- and θ-space, respectively. The right-hand side contains the dynamic part of the equation, where S_{tot} is a source/sink term representing numerous physical processes that generate, dissipate and redistribute wave energy. Specifically, six processes contribute to S_{tot}:

$$S_{tot} = S_{in} + S_{nl3} + S_{nl4} + S_{wc} + S_{bot} + S_{db}. \tag{5.2}$$

These terms denote, respectively, energy input by wind, nonlinear transfer of wave energy through three-wave and four-wave interactions, wave decay due to whitecapping, bottom friction and depth-induced wave breaking. These physical processes in *SWAN* are implemented following various formulations; i.e., wave generation is described as the sum of linear and exponential growth term according to the reasonance mechanism of Phillips (1957) and a feed-back mechanism of Miles (Miles 1957); dissipation of energy due to white-capping is computed with the pulse-based model of Hasselman et al. (1974), or by the saturation-based model proposed by Van der Westhuysen et al. (2007); bottom-dissipation is represented with the empirical *JONSWAP* model of Hasselmann et al. (1973), the drag-law model of Collins (1972), or the eddy-viscosity model of Madsen et al. (1988); depth-induced wave breaking is defined by the bore model of Battjes and Janssen (1978), or by the model proposed by Thornton and Guza (1983); the quadruplet wave-wave interactions are computed with the *DIA* of Hasselman et al. (1985), or with the *XNL* approach proposed by Van Vledder and Bottema (2003); the triad wave-wave interactions with the Lumped Triad Approximation (*LTA*) of Eldeberky (1996) or by the Stochastic Parametric model based on Boussinesq equations (*SPB*) method proposed by Becq-Girard et al. (1999). Additionally, *SWAN*

takes into account diffraction in restrict sense, by considering a phase-decoupled approach, as described in Holthuijsen et al. (2003). Finally, the wave-induced set-up of the mean surface can be computed, based on exact shallow water equations in *1D* mode, whilst in *2D* cases an approximate formulation is adopted since the effects of wave-induced currents are ignored (*SWAN* Technical documentation, page 79). Recently, new formulation have been added, with respect to triad computations, wave-breaking models, and new default calibrated parameters values (see Salmon et al. 2014, 2015 and Zijlema et al. 2012).
In order to obtain a unique solution of equation (5.1), boundary conditions should be provided. The incoming wave components at the seaward boundaries are specified by a directional spectrum, where the shape of spectrum could be defined as *JONSWAP* spectrum, Pierson-Moskowitz spectrum, Gaussian-shaped frequency spectrum, or simulated results provided by external model predictions or other data sources (Holthuijsen 2007). The land boundaries are fully absorbing for wave energy leaving the geographical domain. The frequency spectrum is split up into deterministic prognostic and analytical diagnostic parts. The prognostic part is limited by a fully absorbing boundaries, in the range between minimum, ω_{min}, and maximum, ω_{max}, frequencies, whereas a ω^m $(m = 4 \div 5)$ diagnostic tail is added above the high-frequency cut-off, which is necessary to simulate nonlinear wave-wave interactions and for computing integral wave parameters. No boundary condition are needed in the directional space.

SWAN is provided in stationary and optionally non-stationary mode, and can be applied in Cartesian, curvilinear or spherical co-ordinates (*SWAN* Technical Manual).
The numerical scheme is an implicit scheme, unconditionally stable and not subject to Courant criteria (both in geographical and spectral space). Mainly, three type of schemes can be used: first order *BSBT*, second order *SORDUP* and second order Stelling and Leendertse (S&L, Stelling and Leendertse 1992) schemes; a detailed overview can be found in Rogers et al. (2002) and in the *SWAN* Technical Documentation. In particular, *SORDUP* and S&L are the default scheme for stationary and non-stationary computations, respectively.
The overall algorithm of solution is based on a four-direction Gauss-Seidel iteration technique (Booij et al. 1999, Zijlema and Van der Westhuysen, 2005). The code is provided either in structured or unstructured mode (Zijlema 2009), as well as it is extended to the parallel framework (Zijlema 2010).

Lastly, *SWAN* can be further improved by considering the coupling with other models, such as the circulation *ADCIRC* (ADvanced CIRCulation model; Luettich and Westerink 2004) model, the morphological *COHERENS* model, or, the Boussinesq *TRITON* (Groeneweg et al. 2004) model. The *SWAN + SWASH* coupling is developed by using these types as general guidelines.

5.2 SWASH model

SWASH (an acronym of Simulating WAves till SHore, Zijlema et al. 2011b) is an operational, general-purpose, public domain model intended to be used for simulating wave fields, rapidly varied shallow water flows, as well as large-scale flows and transport phenomena driven by tidal, wind and buoyancy forces, in coastal waters. The model has been developed based on the work of Stelling and Zijlema (2003), Stelling and Duinmeijer (2003), Zijlema and Stelling (2005, 2008). It can be assumed as a multi-layered, non-hydrostatic, free-surface, rotational wave-flow model, which solves numerically the Reynolds-averaged Navier-Stokes equations for an incompressible

fluid, as well as, optionally, the equations for conservative transport of salinity, temperature and suspended load for both cohesive and non-cohesive sediment.

The governing three-dimensional, primitive variable equations describing constant density, free surface flows can be derived from the Navier-Stokes equations after turbulent averaging and under the assumption that the pressure is composed by hydrostatic and non-hydrostatic terms. The resulting equations are the nonlinear shallow water equations with the addition of a vertical momentum equation and non-hydrostatic pressures in horizontal momentum equations. In lieu of simplify the exposition, although either curvilinear or spherical coordinates are implemented, the equations are introduced in terms of Cartesian coordinates; moreover, a *2D* vertical plan is considered, while the extension to full *3D* is straightforward (i.e. Zijlema and Stelling 2005, 2008). The set of continuity and momentum equations reads

$$\frac{\partial \mathrm{u}}{\partial \mathrm{x}}+\frac{\partial \mathrm{w}}{\partial \mathrm{z}}=0, \tag{5.3}$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho}\frac{\partial(p_h+p_{nh})}{\partial x_i}+\frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xz}}{\partial z}, \tag{5.4}$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+w\frac{\partial w}{\partial z}=-\frac{1}{\rho}\frac{\partial(p_h+p_{nh})}{\partial z}+\frac{\partial \tau_{zx}}{\partial x}+\frac{\partial \tau_{zz}}{\partial z}-g, \tag{5.5}$$

where t is time, x is a horizontal coordinate, located at the still water level, and z is the vertical coordinate with the z-axis pointing upwards; $u(x,z,t)$ and $w(x,z,t)$ are, respectively, the horizontal and vertical velocity; ρ is fluid density, p_h and p_{nh} are the hydrostatic and non-hydrostatic pressure components, respectively, and τ_{xx}, τ_{xz}, τ_{zx}, τ_{zz} are turbulent stresses. The domain is bounded vertically by the free surface at $z=\eta$ and the bottom at $z=-d$. The hydrostatic pressure is expressed in terms of the free surface elevation as $p_h=\rho g(\eta-z)$ so that $\partial_x p_h=\rho g\partial_x\eta$ and $\partial_z p_h=-\rho g$,with $(\partial_x,\partial_z)\equiv(\partial/\partial_x,\partial/\partial_z)$, and g being the gravitational acceleration. A detailed description of the non-hydrostatic term, p_{nh}, is given in Zijlema et al. (2011) and Zijlema and Stelling (2003).
Kinematic boundary conditions are prescribed at the free surface and bottom; given by

$$w|_{z=\eta}=\frac{\partial \eta}{\partial t}+u\frac{\partial \eta}{\partial x}, \tag{5.6}$$

$$w|_{z=-d}=-u\frac{\partial d}{\partial x}. \tag{5.7}$$

These boundary conditions ensure that particles laying on the free surface or on the fixed bottom do not leave those surfaces.

By integrating equation (5.3) over the water depth, $h=\eta+d$, and using the kinematic conditions, equations (5.6-5.7), the free surface condition is obtained

$$\frac{\partial \eta}{\partial t}+\frac{\partial}{\partial x}\int_{-d}^{\eta}u\,dz=0. \tag{5.8}$$

At the free surface the dynamic boundary condition is constant atmospheric pressure (i.e. zero relative pressure, $p_h = p_{nh} = 0$) and no shear stresses. Since the vertical grid spacing usually does not allow for solving the dynamics of the boundary layer, a bottom friction term is added to the near bottom layer, and included to the horizontal momentum equation (5.4), based on a quadratic friction law

$$\tau_b = c_f \frac{U|U|}{h}, \tag{5.9}$$

where U is the depth averaged velocity and c_f is a dimensionless friction coefficient. For typically depth-averaged calculations, the drag coefficient can be determined from the Chézy number, C, or the Manning number, n, or assuming a constant value

$$c_f = \frac{g}{C^2}, \tag{5.10}$$

$$c_f = \frac{gn^2}{h^{1/3}}. \tag{5.11}$$

Alternatively, the logarithmic wall law based on a Colebrook-White formulation may be applied, where a distinction should be made between smooth and rough beds.

The turbulence is modelled with several turbulence models, by prescribing separately the eddy viscosity for the vertical (ν^v) and the horizontal (ν^h) transport. The turbulent stresses are represented in a diffusive manner by assuming the Boussinesq approximation, whereby in the multi-layered case is written

$$\tau_{ij} = \tau_{ji} = \nu^h \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad \tau_{jz} = \nu^h \frac{\partial w}{\partial x_j}, \quad \tau_{zi} = \nu^v \frac{\partial u_i}{\partial z}, \quad \tau_{zz} = \nu^v \frac{\partial w}{\partial z}, \tag{5.12}$$

where i, j are indices that represent the horizontal x, y directions and z is the vertical direction.
In particular, three different horizontal eddy viscosity models are available: a constant viscosity, the Smagorinsky model (Smagorinsky, 1963) and the Prandtl mixing length hypothesis; whilst vertical mixing can be modelled by using either a constant value (at low resolutions) or the standard $k - \varepsilon$ closure model approximation (Launder and Spalding 1974) whether high vertical resolution is used.

Appropriate boundary conditions need to be imposed at the open and closed boundaries of the computational domain. At the offshore boundary incoming, regular and irregular waves are introduced by prescribing a vertical profile of horizontal velocities (Zijlema et al. 2011b, Smit et al. 2013 – Appendix B), eventually including a second-order weakly-nonlinear contribution (Rijnsdorp et al. 2014). These velocities are obtained from time-or Fourier series for each point on the boundary, or by imposing a wave spectrum, which is synthesized by using the single-summation method (Miles 1989). Additionally, to simulate entering waves and to allow frequency energy leave the domain, a weakly-reflective condition acting on outgoing waves is adopted (Blayo and Debreu 2005), assuming that incoming and outgoing waves are perpendicular to the boundary.

5.2.1 Numerical implementation

The physical domain is discretized on a structured, Arakawa C-staggered grid with constant grid size in the horizontal planes; whilst, either a depth-averaged or multi-layered mode can be employed in vertical, by splitting-up the computational domain in *K* fixed, terrain-following layers with a spatially varying thickness (sigma plane), between the free-surface and the fixed bottom, or by imposing a distribution of vertical layers with fixed thickness (z-grid). The numerical implementation is typically based on an explicit, second order (in space and time) finite difference method that conserves both mass and momentum at numerical level (according to Stelling and Duinmeijer 2003).

Basically, *SWASH* provides numerous options for space discretization and time integration of the governing equations. For a complete overview reference is made to the *SWASH* manual.

The time integration of equations (5.3-5.5) are typically based on an explicit leapfrog scheme (or better, on a slightly modified version which employs staggering in time) in conjunction with a second order explicit time step for advection, a first order implicit time step for the non-hydrostatic part, and a first-order explicit time step for the viscosity term (Zijlema et al. 2011b). Furthermore, the bottom friction term expressed by equation (5.9) is approximated implicitly, to enhance the robustness of the model. Alternatively, unconditional stability can be achieved and beneficial to large-scale applications (i.e. tidal flows), since time discretization may be carried out using an explicit time stepping for horizontal advective and viscosity terms, semi-implicit time stepping using the θ-method for both surface levels, pressure gradients and free-surface condition (Zijlema and Stelling 2005).

In the surf zone, *SWASH* intrinsically accounts for the energy dissipation of a breaking wave, at a rate analogous to that of a bore, and is able to reproduce accurately the actual location of incipient wave breaking (Zijlema and Stelling 2008). This is due to the shock-capturing property rooted in the momentum-conservative advection scheme used in *SWASH* (Stelling and Duinmeijer 2003). However, high vertical resolution ($10 \div 20$ layers) is required to achieve such accurate results, whereas at low resolution ($1 \div 6$ layers) wave breaking is generally delayed. Therefore, the so-called Hydrostatic front Approximation (*HFA*) may be alternatively used to capture wave breaking with only a few vertical layers, where a (local) reduction to the *NLSWE* equation is performed (Smit et al. 2013). In this approach, wave breaking is initiated when the vertical velocity of the free surface exceed a threshold ratio of the shallow water wave celerity (i.e. $\partial_t \eta > \alpha\sqrt{gh}$). The pressure at the wave front is then forced to be hydrostatic. Wave breaking is terminated when $\partial_t \eta / \sqrt{gh} < \beta < \alpha$. The lower value β is called persistence parameter, while α is called the maximum steepness parameter.

Acceptable frequency dispersion can be achieved by using only a few layers, due to the Keller-box scheme employed for the approximation of the vertical gradient of the non-hydrostatic pressure (Stelling and Zijlema 2003). Introducing the wave number *k*, the angular frequency ω, it may be shown that with *K* vertical equidistant layers the frequency dispersion is represented with a $[2K-2,2K]$ Padé expansion in $\mu = kd$ of the expression for the phase velocity ω/k (Zijlema and Stelling 2008). This implies that one, two and three layers are sufficient to compute linear

dispersive waves, respectively, up to $\mu \cong 0.5$, 8 and 16, with a relative error of at most 1% in phase velocity. Alternatively, a classical explicit, central differencing scheme may be adopted for computing the vertical pressure gradient if a high vertical resolution is needed. It should be noted that these approximations contain the most computational intensive part of the code. Briefly, a Poisson equation linking the non-hydrostatic pressure correction at a grid point to its neighbors is obtained as a final stage of the pressure correction technique employed. The solution of this Poisson equation is carried out by employing either the *SIP*, in the case of depth-averaged mode, or the *BiGGSTAB* preconditioned with an *ILU*, in the case of multi-layered mode. Further details can be found in Zijlema and Stelling (2005). In spite of reducing the computational effort devoted to inverting the Poisson pressure matrix, since one of the last version of the model (*SWASH 3.14AB*) the stencil of the pressure equation (i.e. the number of pressure unknowns per water column) can be diminished, with the purpose of reducing the rank of the Poisson matrix. Nevertheless, a non-equidistant layer distribution needs to be sought to balance the less accuracy concerning the wave dispersion, in spite of having a more efficiency in terms of *CPU* time.

Finally, the reliable and simple wet-dry approach presented in Stelling and Duinmeijer (2003) is implemented. This method tracks the shoreline motion by ensuring non-negative water depths and using the upwind water depths in the momentum flux approximations.

Lastly, the following CFL condition based on the Courant number C_r is given

$$C_r = \frac{\Delta t\left(\sqrt{gd}+|u|\right)}{\Delta x} \leq 1, \tag{5.13}$$

since the time integration is of explicit type and strict conformity of stability criteria for a stable solution is required. In particular, a dynamically adjusted time step controlled by the Courant number is implemented in *SWASH*, i.e. $C_{r,min} \leq C_r \leq C_{r,max}$. Therefore, the time step is either half or doubled so that the new value always lies within this range. Moreover, an additional stability condition is required due to the explicit treatment of viscosity terms; such that a second time step restriction should be imposed, according to

$$\Delta t \leq \frac{\Delta x^2}{2\nu_t}, \tag{5.14}$$

Since this time step restriction is typically very local but more severe than the CFL condition (Eq. 5.13), a constraints on the maximum of the eddy viscosity is adopted at each time step by the model., i.e.

$$\nu_t = min\left(\nu_t, \frac{\Delta x^2}{2\Delta t}\right). \tag{5.15}$$

5.3 SWAN + SWASH model

5.3.1 Coupling implementation

A code producing a single executable is implemented, where *SWAN* (version 41.10) is the master model and *SWASH* (version 4.01) subroutines are packed into a library (either static or dynamic) called inside the *SWAN* main time loop. It is a one-way coupling, not taking into account the seaward-directed reflected waves computed by the phase resolving model, which, consequently, cannot affect the action balance equation model.

The model components run sequentially in time on structured, Cartesian (or curvilinear) meshes, with *SWAN* acting on a larger domain extended up to offshore, whereas the *SWASH* domain is typically located in the nearshore, possibly including the emerging beach terrain. However, it is to be pointed out that the *SWASH* domain can be properly modified according to the type of application, such that an extension of the computational domain at most equal to the *SWAN* domain may be adopted.

The coupling takes place at the seaward (offshore and/or lateral) *SWASH* open boundary, where: (i) the spectra computed by *SWAN* are prescribed; (ii) wave trains are synthesized by a single summation method (Miles 1989), which produces a quasi-homogeneous wave variance (Miles and Funke 1989); (iii) the horizontal velocity normal to the boundary, computed using the linear theory, is prescribed (Zijlema et al 2011b, Smit et al 2013); (iv) a weakly reflective boundary condition is added to prevent spurious outgoing oscillations (Verboom and Slob 1984; Verboom and Segal 1986; Vreugdenhil 1994, Section 5.4; Blayo and Debreu 2005; Smit et al. 2013, Appendix B). Eventually, the weak-nonlinear, second-order velocity may be prescribed; hence, the contribution of incident bound ig-waves can be added to the *SWAN* spectra (Rijnsdorp et al 2014). In addition, in the last version of the coupled model (*SWAN-4110 + SWASH-4.01*), *SWASH* may also be forced by a band-pass filtered (i.e. pass band $\omega = [0.5 \div 3]\omega_p$, $\omega_p$being the incident peak frequency) Fourier series of levels directly obtained from the *SWAN* spectra, and imposing, subsequently, the horizontal velocity profile as uniformly-distributed. This option is found to be fairly adequate for simulations in shallow and very shallow waters, where it appears insufficient to adopt exclusively the hyperbolic profile; moreover, some spurious low-frequency motion are most likely to be generated because of the combined effect between the breaking process, which governs the energy dissipation in this zone, and the presence of sub-harmonic triad interactions, either reflected off the beach and eventually freely re-reflected off at the inner coupling point between the two model components. Finally, the wave-induced setup calculated as option by *SWAN* either in a (geographic) 1D or 2D case, as briefly reported before (§5.1, see *SWAN* Technical documentation, page 79), may be easily included in the coupled model and so passed to the *SWASH* domain.

5.3.2 SWASH wave-maker

In his paragraph the attention will be focused on the main characteristics of the wave generating boundary condition within the coupled *SWAN-SWASH* model.

Incident irregular waves at the open boundaries are, usually, represented by realisations of zero-mean, stationary, random Gaussian stochastic process. The surface-elevation time series, $\eta_{wm}(x,t)$

at the wavemaker $\boldsymbol{x}_{wm} = (x_{wm}, y_{wm})$ can be obtained as a linear superposition of *N* regular wave components

$$\eta_{wm}(z,t) = \sum_{j=1}^{N} a_j cos\left[\boldsymbol{k}_j \cdot \boldsymbol{x}_{wm} - \omega_j + \varepsilon_j\right], \tag{5.16}$$

where (a_j, ε_j) are the couple amplitude – phase of the j^{th}- wave component having frequencies w_j and directions θ_j, while $\boldsymbol{k}_j = \left(k_j cos\theta_j, k_j sin\theta_j\right)$ is the wave number, related to the wave frequency by the dispersion relationship $\omega^2 = gktanh(kh)$. In particular, it is to be noted as *SWASH* uses its own dispersion relation when at most three equidistant layers with variable thickness (i.e. sigma plane) and the Keller-Box scheme are used; this approximate relation is known to be more consistent with the model, particularly for relatively high frequencies.
The target wave spectrum (i.e. the *SWAN* spectrum with a logarithmic frequency distribution in the coupled model) is divided into N frequency bands with uniform spacing, i.e. $\omega_j = j\Delta\omega$, from $\omega_{min} = 1/2\,\omega_p$ till $\omega_{max} = max\left(3\omega_p, \omega_N = 1/2\Delta t, \omega_{cf} = 2K\sqrt{g/d}\right)$, where ω_p is the peak frequency, $\omega_N$is the Nyquist frequency and ω_{cf} is the cut-off frequency. The Nyquist frequency represents the higher frequency value when the incident spectrum is defined by one of the classical parametric form (*JONSWAP*, Pearson Moskowitz or *TMA*, e.g. see Holthuijsen 2007), whereas the cut-off frequency is the frequency at which the so-called evanescent modes are generated. Specifically, these evanescent modes will be removed by *SWASH* when the offshore weakly-reflective boundary condition and the hyperbolic velocity profile are imposed and if and only if the dispersive properties of the model (governed by the number of layers *K*) are not sufficient. To avoid repetition of the incident wave conditions, the frequency band is set equal to the reciprocal of the duration of the simulation, *T*, i.e. $\Delta\omega = 1/T$ (*T* is assumed to be equal to the *SWAN* time step in the coupled model). Obviously, the more *T* is increased or $\Delta\omega$ is reduced, the more the waves computed with this method are consistent with the continuous target wave spectrum, but the more computational intensive will be the simulation. Furthermore, values of these parameter are limited by the angular and the frequency resolution of the wave-maker.

The amplitudes and directions of each wave component can be selected in various ways according to the particular synthesis method used. Two methodologies are usually available and are denoted as the single and double summation methods. The main drawback of the double summation model is that a spatially inhomogeneous wave field is produced when finite duration time records are considered. In practice, the phase difference between wave components with same frequencies but propagating in different directions is no longer random but locked. In other words, for each of the frequencies w_j, a number of wave fronts (equal to the number of directional bins) are generated, with different directions θ_j, but having the same frequency and leading to a "phase locking", i.e. there exists (at least) couple of waves having the same frequency w_j but different phase ε_j, and, eventually different directions of propagation θ_j. According to Jeffreys (1987), the resulting wave field is therefore spatially non-homogenous and non-ergodic since it presents nodes and antinodes. This means, in other words, that the wave field will have different statistical properties from one point to another.

In both cases, the amplitudes and directions are chosen to approximate a target directional spectrum, according to

$$S(\omega,\theta) = S(\omega)D(\omega,\theta), \tag{5.17}$$

where $S(\omega)$ is the desired frequency spectrum and $D(\omega,\theta)$ is the directional spreading function (Kuik et al. 1988, Goda 2010), eventually expressed as function of the wave frequency. The frequency spectra can be assigned by the classical *JONSWAP*, Pierson-Moscowitz (*PM*) and *TMA* spectrum (see Holthuijsen 2007) as pointed out before, while the directional spreading function is of the type:

$$D(\theta) = D_0 cos^m\left(\theta - \theta_p\right) \tag{5.18}$$

where, θ is the azimuth measured counter-clockwise from the principal wave direction, assumed to be represented by the peak wave direction θ_p; m (-) is the spreading coefficient given by $m = max\{DSPR^{-2} - 2;\ 1\}$ and related to the directional standard deviation, *DSPR (rad*), as proposed by Kuik et al. (1988) – see also Appendix A of the *SWAN* User Manual; D_0 is a constant given by

$$D_0 = \left[\int_{\theta_{min}}^{\theta_{max}} cos^m\left(\theta - \theta_p\right)\right]^{-1} = \begin{cases} \frac{1}{\pi} 2^m \frac{\Gamma\left(1+\frac{m}{2}\right)^2}{\Gamma(1+m)}, & m < 10 \\ \left(1 - \frac{1}{4m}\right)\sqrt{\frac{m}{2\pi}}, & m \geq 10 \end{cases} \tag{5.19}$$

which satisfy the normalization condition, and Γ denotes the Gamma function. For example, by setting m = 10 the directional spreading function is calculated as shown by the solid line in Fig. 5.1. The cumulative value of $D(\theta)$ is also shown in the same figure as the dashed line and it is used by the single-summation method as specified later. From this cumulative distribution of $D(\theta)$, is observed that the wave energy is principally contained in the angular range and the directional spreading of wave energy is narrower around the spectral peak frequency.

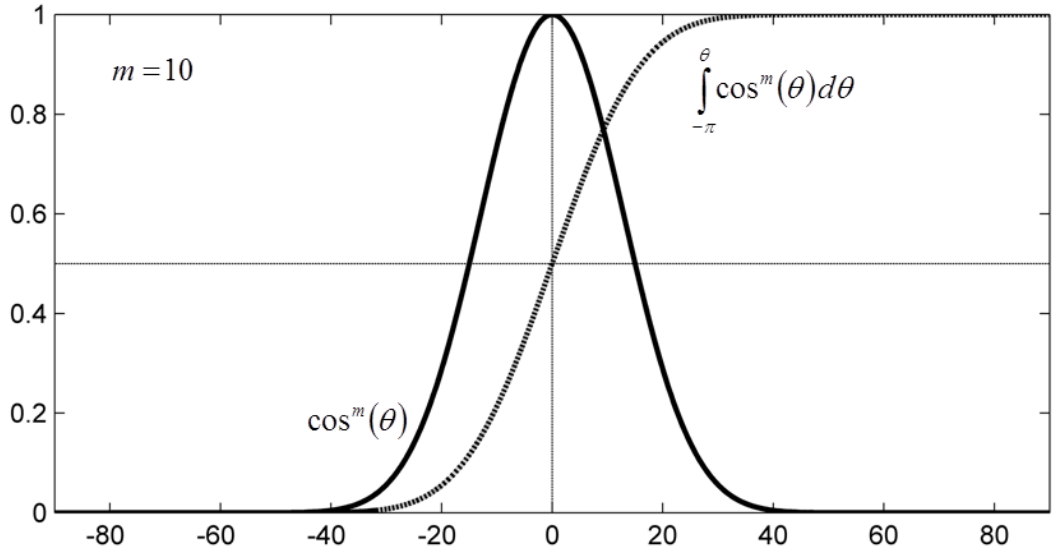

Fig. 5.1 Example of directional spreading function.

Each frequency band corresponds to a long crested wave with direction θ_j, and spatially varying amplitude, obtained deterministically from the wave spectrum $S(\omega_j)$, which reads

$$a_j = \sqrt{2S(w_j)\Delta\omega}, \tag{5.20}$$

so that, the target spectrum is match exactly, while the propagation angles are selected at random according to the cumulative distribution of $D(\omega,\theta)$ (Miles 1989). Usually, within a narrow frequency band, all components have approximately the same amplitude but the probability of obtaining a given propagation angle will be the greatest at the peak of the spreading function. The wave angle is selected by a random number R_n being uniformly distributed between *0* and *1* through the cumulative distribution of the directional distribution function, i.e.

$$\theta_n = F^{-1}(R_n) \quad : \quad F(\theta|f_n) = \int_0^\theta D(\theta|f_n)d\theta. \tag{5.21}$$

Goda (2010) reported a numerical procedure for solving this problem, since the inverse solution of the cumulative distribution cannot be obtained analytically. Firstly, the random number is assigned to each frequency and the spreading function $D(\theta|f_n)$ is initially integrated for the full range of wave angles so as to evaluate the normalization constant G_0. Then, the integration of the second equation is made step by step until the value $F(\theta|f_n)$ equals R_n, which is assigned to each frequency, and finally, the value of θ_n is assumed to be equal to the angle obtained by this procedure.

Generally, waves are generated in *SWASH* by prescribing a local velocity profile normal to the boundary, either by imposing a wave spectrum along the boundary or starting from time series directly assigned for each point on the boundary.
In the first case, *SWASH* allows to specify either unidirectional (Zijlema et al. 2011b) or multidirectional (Smit et al. 2013, Appendix B) waves, in both depth-averaged or multi-layered mode (namely, by computing the layered-averaged velocity through integration over each layer). The expression of the velocity and flux boundary condition along the cross-shore direction can properly be obtained from the surface elevation using the linear transfer function (i.e. Goda 2010, page 364) and by assuming the total local horizontal velocity u_b equal to a superposition of the incident (or target) velocity signal u_t and a velocity signal of the reflected waves u_r. In particular, all of these velocities are assumed to be perpendicular to the boundary.
For instance, the expression of the depth averaged case for unidirectional waves is given by

$$u_b(z,t) = \pm\sum_{n=1}^{N} a_n \left[\omega_n \frac{cosh k_n(z+d)}{sinh k_n h} + \sqrt{\frac{g}{h}}\right] cos(\omega_n t - \varepsilon_n) - \sqrt{\frac{g}{h}}\eta, \tag{5.22}$$

where $\varepsilon_n \sim U[0,2\pi]$ is the random phase uniformly distributed between 0 and 2π. The sign in the equation (5.22) – also used in the equations (5.23) and (5.26) – depends on the location of the boundary. The minus sign refers to an inflow velocity at the eastern and northern boundaries, whereas the plus sign is related to an inflow velocity at the western and southern boundaries.

Furthermore, in the 2D case and for short-crested waves with a prescribed directional distribution (assumed constant along the boundary), the normal velocity at the boundary is given as

$$u_b(s,z,t,) = \pm \sum_{n=1}^{N} a_n \left[\omega_n \frac{coshk_n(z+d)}{sinhk_n h}\right] cos[k_n cos(\theta_n)s - \omega_n t + \varepsilon_n] cos(\theta_n) + u_r, \quad (5.23)$$

where it is assumed that incoming and outgoing waves are perpendicular to the boundary and that reflected waves are small amplitude shallow water waves, which implies, essentially, that the high frequency energy has dissipated in the domain to the extent that it may be neglected when it leaves the domain. The depth averaged horizontal velocity of reflected waves is then computed based on the surface elevation of outgoing waves, which is obtained by subtracting the surface elevation due to the incident waves η_t from the instantaneous surface elevation η computed by *SWASH*. The reflected wave contribution is detected by using the radiation condition in combination with the mass conservation (Vreugdenhil (1994, Section 5.2; Blayo and Debreu 2005)

$$u_r = \frac{\sqrt{gd}}{d}(\eta - \eta_t), \quad (5.24)$$

where $c = \sqrt{gd}$ is taken as the shallow water phase velocity.

On the other hand, the incident signal for a time series or a Fourier series can be given by

$$\eta_b = a_0 + \sum_{n=1}^{N} a_n cos(\omega_n t - \delta_n) \quad (5.25)$$

where N is the number of Fourier components, a_0 is the mean water level, and a_n, ω_n, δ_n are the amplitude, the angular frequency and local random phase lag of the n^{th}-component. Hence, *SWASH* can alternatively be forced by imposing the following weakly-reflective boundary condition (Zijlema et al. 2011b)

$$u_b = \pm\sqrt{\frac{g}{h}}(2\eta_b - \eta) \quad (5.26)$$

It should be noted that for highly nonlinear irregular waves in shallow water, the velocity boundary conditions based on the hyperbolic profile (Equations 5.22, 5.23) would have to be modified to take into account the presence of lower and higher frequency wave components induced by nonlinear interactions between the primary wave components. It is known the Second-order Stokes theory cannot accurately describe the shape of cnoidal-type waves in shallow water when the Ursell parameter is large. Moreover, the weakly-nonlinear (and weakly-reflective) second-order boundary condition is neither valid in the surf zone nor in deep water, since it is generally derived under the assumption of small amplitude waves (i.e. $a/h \ll 1$), and because the computation of the depth-averaged second-order velocity amplitude is obtained under the hypothesis of long wave response restricted in shallow water ($k_{mn}d \ll 1$), where $k_{mn} = k_m - k_n$ is the difference wave number of two primary waves (f_m, f_n) (see Rijnsdorp et al. 2014 for a detailed derivation). In addition, the weakly-reflective condition is not more valid, since the dissipation produced by the high frequency energy inside the domain cannot be assumed sufficient to be neglected.

On balance, shallower is the coupling point, the less valid are the hypothesis of hyperbolic profile at the wave-maker, and more spurious waves are created on the *SWASH* domain. For all of these reasons, a combination between a Fourier series and uniform velocity profile may be expressly and alternatively used (i.e. Equations 5.25, 5.26) in the coupled model.

5.3.3 Parallelization strategy

Both *SWAN* and *SWASH* have been already extended to the parallel framework (i.e. Zijlema 2010, Zijlema et al. 2011b), by using the domain decomposition on distributed computer system with each processor having its own private memory. In both models a balanced, static strip-wise partitioning is employed, where the whole computational domain is split up into a number of contiguous, non-overlapping strips, along the *x*- or *y*-axis, with each of them being assigned to a different processor. This is a case of single program multiple domain (*SPMD*) programming model, since the same algorithm performs an all available processors. In the *SWASH* model, also an orthogonal recursive bisection (*ORB*) is available in case of *2D* multi-layered mode. Message passing is implemented by a high level communication library *MPI*, where either point-to-point or collective communications have been employed.

Although an emerging practice is to couple heterogeneous models through a generic framework – for instance, the Modelling Coupling Toolkit (*MCT*, Larson et al. 2005, Jacob et al. 2005, Warner et al. 2008) –, it seemed to be more convenient acting on the source code of both models, because it is likely to produce more effective results. Specifically, it was chosen to develop the parallel implementation according to *ADCIRC + SWAN* coupled model approach (Dietrich et al 2011). Highly-localized and efficient communication were achieved in *ADCIRC + SWAN*, because parallel simulations utilize identical sub-meshes; fundamentally, intra-model communication is inter-core and occurs merely on adjacent sub-meshes edges, while inter-model communication is intra-core. Although its proven highly-scalable performances, this scheme cannot be completely followed in *SWAN + SWASH*, because, to save computational time, it is necessary to have grids with different dimensions and resolutions. Nevertheless, adequate local characteristic and efficiency can be reached as well. Communication between models (intra-model) can be passed through local memory or cache, without the need for any network-based inter-core communication.

In order to compile and link the coupled code, some conflicts between *SWAN* and *SWASH*-related routines were solved. As a matter of fact, both models share the same general structure, including: (i) subroutines and functions that are perfectly identical as to code name; (ii) slightly different subs and functions performing essentially the same tasks having the same name; (iii) many variables with the same names representing different items that must be kept distinct, e.g. arrays mapping the domains of each of the two models. To cope with this problem, it was decided to have the main program of one of the two codes as the main program of the coupled code, treating the second one as a system of subroutines of the first. Namely, it was chosen to have *SWAN* as the main program, whilst *SWASH* was packed into a library, static or dynamic, that can be called inside the *SWAN* main time loop. To this aim, editing the name of the routines having same meaning but slightly different codes was mandatory in order to avoid conflicts inside the common executable.

Once the action density is computed by *SWAN*, the sending and receiving processes within the coupled model may be carried out by using either an off-line or an on-line approach (it is worthwhile underlying that it is a double option implemented into the single executable of the coupled model). Formally, new modules and subroutines were added to the original source code of both models. As a rule, the spectra calculated by *SWAN* are sent to *SWASH* through file read/write, in the off-line case. On the other hand, a shared data structure (*Fortran 90* derived data type) is employed to assure the online data exchanging.

From the parallel coupled model perspective, it is chosen to adopt a multiple program multiple data (*MPMD*) paradigm, based on the modular construction provided by *MPI*, and supported via the related-communicator mechanism. According to Foster (2003), three general forms of composition are available: sequential, parallel and concurrent. Specifically: i) in sequential composition, two model components execute in sequence on the same set of processors; ii) in parallel composition, two program components execute concurrently on disjoint set of processors; iii) in concurrent composition, two model components execute on potentially non-disjoint set of processors.

Basically, the *MPI*'s *MPMD* programming model means that the full generality of concurrent component is not generally available.

The sequential composition is supported by the *MPI_COMM_DUP* function, which is able to create a new communicator comprising the same process group but a new context to ensure that the communication performed for different purposes are not confused. On the other hand, the *MPI_COMM_SPLIT* supports the parallel compositions, where new communicator is created with just a subset of a given group of processors, while the *MPI_INTERCOMM_CREATE* constructs an intercommunicator, which links processes in two groups. In this work, it is chosen to employ the first type of modular construction, since it seems appropriate to leave the whole group of processors to each program. In this way, the communication is localized, because inter-model communication happens on the same core and it is kept on the serial part of the code, while intra-model happens intra-communicator.

Further work is needed to be done with respect to the mapping of both grids. Up to now, domain decomposition of both models is carried out separately, and one model does not know the sub-domain managed by the other. Therefore, the action density computed by some cores inside *SWAN*, need to be passed globally to all sub-domains of *SWASH* and some improvements are needed. For instance, it seems useful implementing something similar to the intra-coupler communication, as the corresponding data storage and decomposition, employed by the *MCT* coupler (Jacob et al. 2005); namely, it should be implemented a common mapping of decomposition grids of both models, in order to perform the action density computation and the wave synthesized generation between the solely sub-domains actually involved in the sending and receiving tasks. Unfortunately, the task is not so trivial, because a substantial modifications of the *SWASH* reading boundary conditions subroutine is required. Nonetheless, the *ADCIRC+SWAN* parallel structure cannot be adapted as well, because it is tailored to grids having the same size and dimensions. On the other hand, some improvements may be achieved by focusing the implementation on the virtual topology feature, which is an inherent MPI communicator characteristic.

Notwithstanding the above-mentioned improvements will be necessary, preliminary analysis have been carried out on a simplified test case with a single 2.60 GHz Intel Core i7 – 4720 HQ 8 processor, having 16 GB of memory, showing that the online/offline coupled model extension (*SWAN+SWASH*) with one executable is capable to reduce the *CPU* time with respect to the "manual" procedure of using two separate executable (i.e. separated *SWAN* and *SWASH* models). One of the next tasks will be to measure the parallel performance of the coupled model on a dedicated cluster.

Hence, to sum up, higher efficiency is expected to be achieved by further optimizing the online coupling in the parallel mode.

5.3.4 Non-stationary implementation

A further improvement is implemented by adding the option for computing a non-stationary simulation. The coupling procedure, between two models having substantially different computational time steps, is generally performed admitting that the results obtained by the model with the larger time step, and used to force the other model, are kept constant during the computations in between the coupling interval. An exemption may be considered the two-way *ADCIRC + SWAN* coupled model, where *ADCIRC* extrapolates the radiation stress from *SWAN* both at the beginning of the current interval and at the beginning of the previous interval for each time step of the whole simulation (Dietrich et al. 2011).

Herein, the idea is to force *SWASH* open boundaries with cyclic-(quasi-stationary) spectra computed by *SWAN*, either assuming an averaged status or a final sea state, with the aim of representing better the mean wave conditions simulated on the whole coupling period. Basically, the averaged or the final status may be assessed by requiring an interpolation of the *SWAN* spectra computed at the beginning and at the end of the coupling interval (excluding the first coupling interval, which can be assumed as a spin-up time of the coupled model). This can be done using the procedure already implemented in *SWAN* (SWAN Technical Documentation page 66). In practice, the interpolation between two spectra is done (i) normalizing the spectra by averaged frequency and direction, (ii) interpolating and transforming back; thereby, no reduction of the spectral peaks would happened.

Figure 5.2 depicts a schematic overview of the non-stationary procedure implemented in the coupled-code. Specifically, the *SWAN* model is run first, and the coupling interval is taken to be equal to the *SWAN* time step, since the *SWAN* sweeping solution can take much larger time steps than *SWASH*, which is Courant-time-step limited due to its explicit formulation.

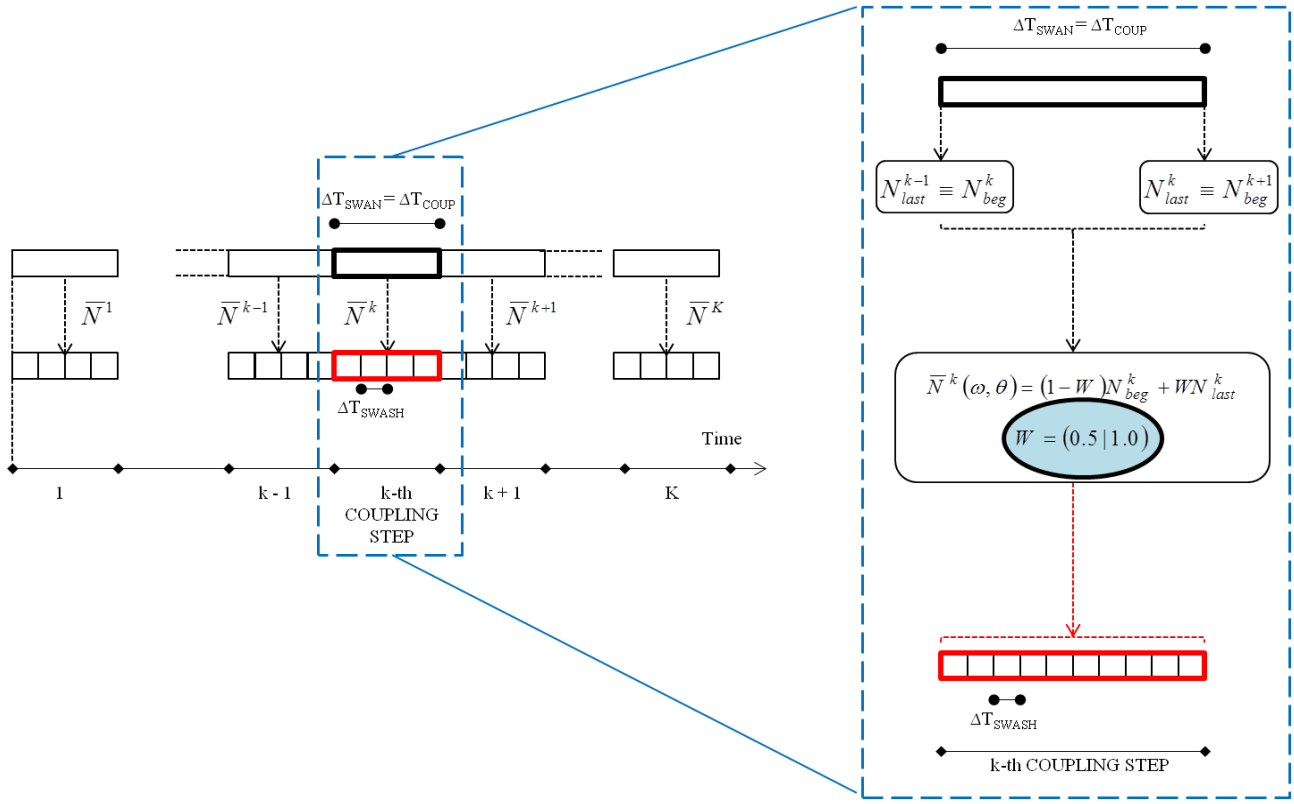

Fig. 5.2 Schematic representation of the non-stationary, online coupling and detail overview of the k^{th}coupling interval. The coupling interval (ΔT_{COUP}) is set equal to the SWAN time step (ΔT_{SWAN}). Scales of the SWASH time step (ΔT_{SWASH}) are enlarged, since its value is much more lower than the SWAN time step. N^k_{last} and N^k_{beg} stand for the action density at the end and at the beginning of the k^{th}- coupling interval, where $k = 1, \cdots, K$. Finally, W is a weight-factor, which can assume only values equal to 0.5 or 1.0, to determine whether SWASH is forced by the final sea state or by an averaged status of the action densities computed by SWAN in the k-th coupling interval, according to the interpolation $\bar{N}^k = (1 - W)N^k_{beg} + WN^k_{last}$.

5.3.5 Validation of the coupled numerical model

In the following paragraphs, a validation of the coupled SWAN+SWASH model for nearshore application is provided, by using laboratory data of unidirectional random wave runup on a gentle, smooth and impermeable slope (Mase 1989). Two main groups of results are going to be presented, which are related to simulations carried out with two different versions of the coupled model. The first is linked with the issues presented in Ventroni et al. (2016b,c), which is based on a previous version of the coupled model (i.e. *SWAN 4101+SWASH 3.14*), whilst the second one is obtained by employing an improvement of the coupling, as well as a more recent version of the two model components (i.e. *SWAN 4110 + SWASH 4.01*). Anyhow, the purpose of both group of simulations is the same and at least twofold.

First, given the heavy computational requirements of phase-resolving modelling, an interesting solution in setting up the coupled model is represented by the optimal location of the inner boundary, so as to reduce as much as possible the *SWASH* domain and related computational burden without losing the outcome's accuracy and reliability. Namely, the trade-off is between having the inner boundary far enough from the shore to represent the wave nonlinearity wherever significant, but not as much as to include areas of large wave dispersion, requiring too many layers to be represented.

Second, it appears interesting comparing the results of the online coupling with those obtained by McCabe et al. (2010, 2011), which used the same tests to investigate the best location within an offline–coupling between SWAN and a NLSWE model, as well as between SWAN and a BTE model.

The above procedure was carried out manually, with use of distinct codes for the spectral and the phase-resolving models. Higher efficiency is expected to be achieved by the online coupling, resulting in one code and one executable, for seamlessly simulating wave evolution from generation to runup and land inundation. To this aim, SWAN, SWASH, and SWAN + SWASH model runs were carried out.

5.3.6 Mase's (1989) test description

Mase (1989) studied random wave runup for a range of wave conditions and four test slopes of a model beach. The experimental setup can be seen in Figure 1 of Mase & Iwagaki (1984), while a schematic view is depicted in Fig. 5.3 (last subplot of each panel). Model results have been compared with four of the random wave runup laboratoty tests, with bed slope 1:20 and water depth at the wavemaker d_{wm}=0.45 m. A Pierson-Moscowitz spectrum was imposed at the wavemaker, with values of significant wave height H_{wm}, and peak period T_p as reported in Tab. 5.1, where representative wave length at the wavemaker, L_{wm}, related depth-to-wavelength ratio, $(d/L)_{wm}$, and equivalent deepwater significant wave height H_{dw} and wavelength L_{dw} are also shown. The *SWASH* numerical model is set-up along a one-dimensional flume of length 25 m, with horizontal resolution and time step shown in Tab. 5.1. In particular, at least 20 points per wave length for waves up to $3f_p$ are chosen, f_p=1/T_p being the peak frequency. *SWASH* was run for a duration of 1200 s, with spin-up time of 300 s.

Tab 5.1 Parameters of incident waves for the Mase (1989) tests. K_s is the shoaling coefficient based on linear wave theory

	H_{dw} (m)	H_{wm} (m)	K_s (-)	T_p (s)	L_{dw} (m)	L_{wm} (m)	$(d/L)_{wm}$ (-)	Δx (m)	Δt (s)	Breaker type
TEST A	0.0477	0.0495	1.037	2.50	9.758	4.998	0.111	0.020	0.002	Plunging
TEST B	0.0639	0.0618	0.968	2.00	6.245	3.884	0.086	0.020	0.002	Spilling
TEST C	0.0793	0.0734	0.930	1.67	4.337	3.120	0.069	0.015	0.001	Spilling
TEST D	0.0990	0.0914	0.915	1.25	2.440	2.122	0.047	0.010	0.001	Spilling

5.3.7 SWAN4101, SWASH 3.14 and SWAN 4101 + SWAS H3.14 set-up and results

SWAN, SWASH and SWAN + SWASH model runs were carried out. Following McCabe et al. (2011), the SWAN model is run without nonlinear interactions (quadruplets and triads) and bottom friction, but activating the depth-induced wave breaking (*Battjes & Janssen,* 1978) and including the wave-induced setup. The lower and upper boundaries in frequency space, subdivided in 200 frequencies, are generally chosen equal to $0.5f_p$ and $3f_p$, respectively. However, in the coupled model simulations using SWASH with one layer, a lower value for the upper limit is chosen, in the range $(2.4 - 2.5)f_p$, in order to introduce only accurate harmonics in the phase-resolving model. In fact, based on an approximate dispersion relation, using one layer, *SWASH* is accurate up to a kd=2.9 for primary waves (k = wave number), with a relative error in the normalized wave celerity $c/(gd)^{1/2}$ of at most 3%. Finally, a $cos^{800}(\theta)$ directional distribution is used.

The SWASH model is run throughout the entire domain using four different configurations for each test cases: (i) one vertical layer, (ii) the same as (i) including incident bound waves (*Rijnsdorp et al.* 2014) at the wavemaker boundary, (iii) two and (iv) three vertical layers. The default values for the maximum steepness parameter (α=0.6) and the persistence parameter (β=0.3), found after calibration by *Smit et al.* (2013), were set in the, so called, hydrostatic front approximation, used to simulate depth-induced wave breaking in accurate way even with few layers. Furthermore, a threshold value of 0.1 mm was set in each run to represent the minimum inundation depth, while turbulence and bottom friction are neglected.

Finally, the *SWAN+SWASH* model is run by using five different random seeds for the generation of time series at the *SWASH* open boundary, in order to take account for sensitivity of runup to the random input phases, and choosing six or seven coupling points, with values of H_{m0}/d (as computed by SWAN) in the range 0.1–0.6. In these cases, the SWASH model is run with one and two layers, with a total of 250 runs.

In Figure 5.3 the significant wave height and the wave-induced setup are shown, normalized with respect to the depth at the wave-maker, as well as the fraction of breaking waves as computed by SWAN. This latter is helpful in choosing the coupling points, considered within the coupled model, which are shown in the last subplot.

Comparison between the measured and calculated runup statistics $R_{2\%}$, $R_{1/10}$ and $R_{1/3}$ will be shown by using the relative error, $(R_{Model} - R_{Mase})/R_{Mase}$.

Figure 5.4 shows the relative errors in runup statistics $R_{2\%}$, $R_{1/3}$ and $R_{1/10}$ with respect to Mase's experimental data, as a function of the surf similarity parameter $\xi_{wm} = m/(H_{wm}/L_0)^{1/2}$ calculated with *SWASH* throughout the computational domain in three different configurations: with a single layer, a single layer accounting for bound waves at the wavemaker, and with three layers. Note that tests A to D are in the order of decreasing ξ_{wm}.

Figure 5.5 depicts the relative errors with respect to the experimental data, calculated with the coupled *SWAN + SWASH* model as a function of the nonlinearity parameter at the coupling point, H_{m0}/d, to investigate how the location of the coupling point influences the accuracy of results with respect to the complete phase-resolving run. The errors of the coupled model were computed as average values over five runs with randomly phased wave trains prescribed as boundary condition of *SWASH*, generated from the same action density spectrum computed by *SWAN*. Each run was performed using the *SWASH* model with a single layer, in order to test the accuracy of the coupled model with the less computationally expensive configuration.

From comparison between Figures 5.4-5.5, it can be seen that results in the coupled model can be improved varying the location of the coupling point. For example, the model performance in Test D is enhanced either using more layers, but increasing the computational time, or choosing a coupling point closer to the shore, with a significantly lower computational effort.
On the other hand, it is apparent that the choice of the coupling point is not unique for the wave conditions analyzed, nor the error–H_{m0}/d curves are monotone.

The optimal coupling point depends also on the accuracy of the *SWAN* spectral modelling. However, it seems that a common value of $H_{m0}/d \approx 0.55$ might be chosen for test cases *B, C, D,* while a lower value of $H_{m0}/d \approx 0.3$ appears to be the best in Test A, differing from the other tests as to breaker type (see Tab. 5.1). These values are lower than McCabe et al. (2011) results, where a value of $H_{m0}/d \approx 0.65$ was found to represent the optimal switching point in their *SWAN+NLSWE* model. This is probably due to *SWASH* being still able to represent some degree of wave dispersion even with one layer compared to the fully non-dispersive *NLSWE* formulation.

Figure 5.6 depicts a summary of relative errors between experimental data and *SWASH* runs (panel P1), *SWAN + SWASH* run using *SWASH* with one (panel P2) and two layers (panel P3).

Although the effect of increasing the number of layers is apparent, more layers should be chosen to ensure adequate modelling of the phase differences between the representative wave components, including shorter waves. Furthermore, results using one layer are most affected by the so-called evanescent modes, especially for the lower peak period. On the other hand, the largest errors appear when two layers are used.

To sum up, simulations are shown to be reasonably effective with the single SWASH model used throughout for sufficiently large peak periods, and in a broader peak period range with the coupled model, for the wave conditions studied.

5.3.8 SWAN4110, SWASH 4.01 and SWAN 4110 + SWAS H4.01 set-up and results

In this paragraph, a different set-up for the single SWASH model is used. First of all, only the depth-averaged-NHE mode is considered and the bottom friction and both vertical and horizontal turbulence are included in order to enhance the stability of the model. The best results in terms of relative errors with respect to experimental runup data was obtained with a bed friction coefficient, c_f, calculated with $n = 0.011$ $m^{1/3}s^{-1}$. In addition, both a background viscosity of 1×10^{-4} m^2/s and a horizontal viscosity 2.5×10^{-4} m^2/s were added as a fixed value.

Computations with *SWAN* are carried out with two main configurations, by activating or not the Lumped Triad Approximation (LTA of Eldeberky 1996) for computing triad wave-wave interactions. Furthermore, the depth-induced wave breaking model of Battjes and Janssen (1978) and also the source term for bottom friction – the eddy-viscosity model of Madsen et al. (1988) with a constant equivalent roughness length scale of the bottom ($kn = 0.00015$ m) –, are used. The last value is consistent with the Manning coefficient assumed for the *SWASH* model.

Finally, the coupled model is run by using the best *SWASH* set-up, and by using both *SWAN* configurations, i.e. with and without nonlinear interactions. In addition, *SWASH* is run both in non-hydrostatic depth-averaged (*SWAN-NHE*) and hydrostatic depth-averaged (*SWAN-NLSWE*) mode and is forced with the pass-band Fourier levels computed online from the *SWAN* spectra (Equation 5.25) and by imposing an uniform profile (Equation 5.26). In particular, results are found to be reasonable accurate and more stable than the previous case.

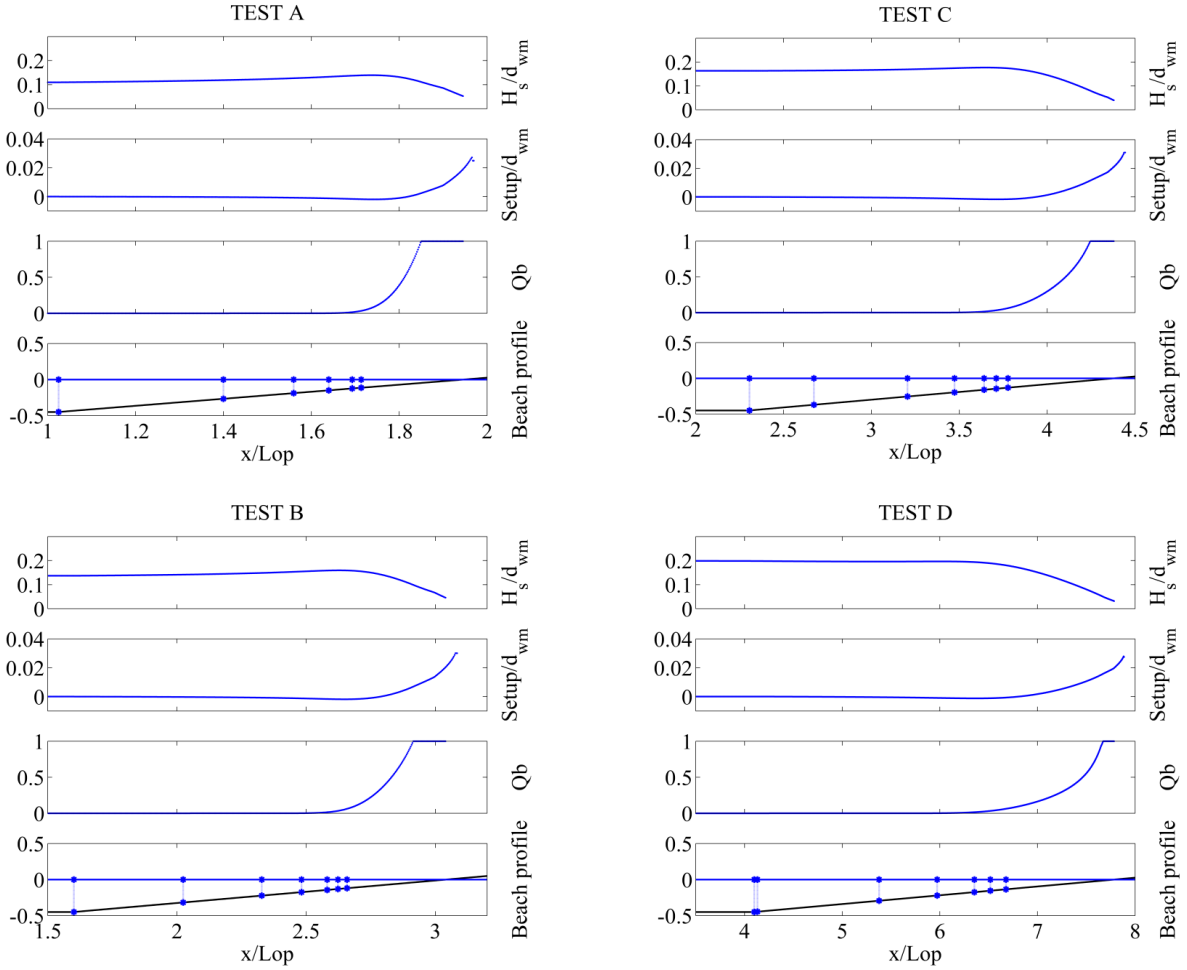

Fig. 5.3 SWAN results. Computed wave height (first panel) and setup (second panel), normalized with respect to the wavemaker depth, fraction of breaking waves (third panel) and coupling points along the bottom profile (last panel), for each test case.

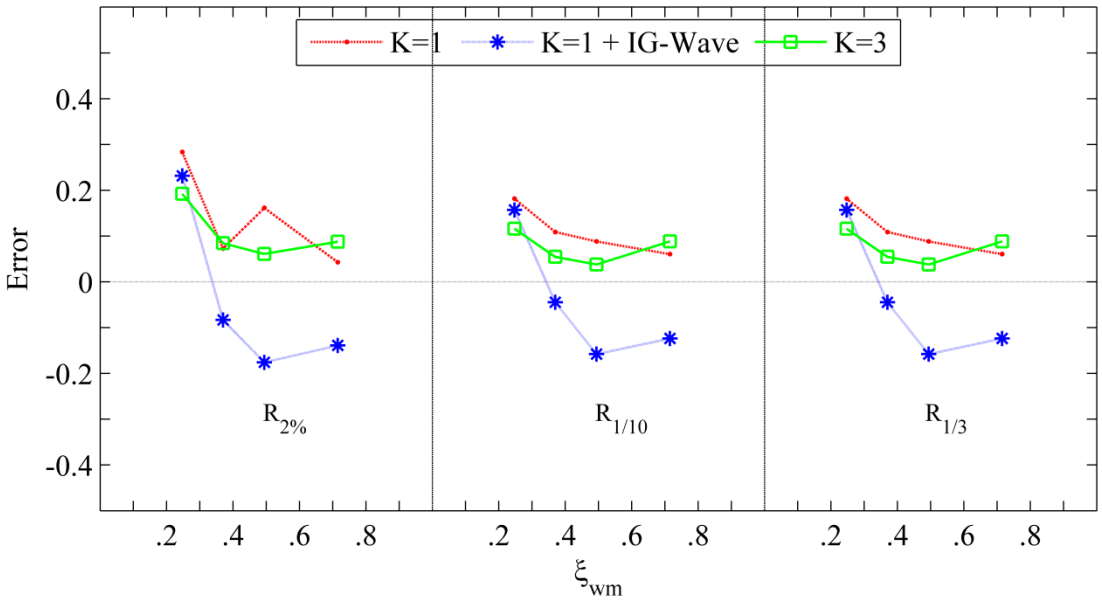

Fig. 5.4 Errors in runup statistics, compared with Mase's (1989) results, with SWASH model used throughout the domain, with one, or three layers, and with one layer plus bound waves at the wavemaker, as a function of the surf similarity parameter, ξ_{wm}

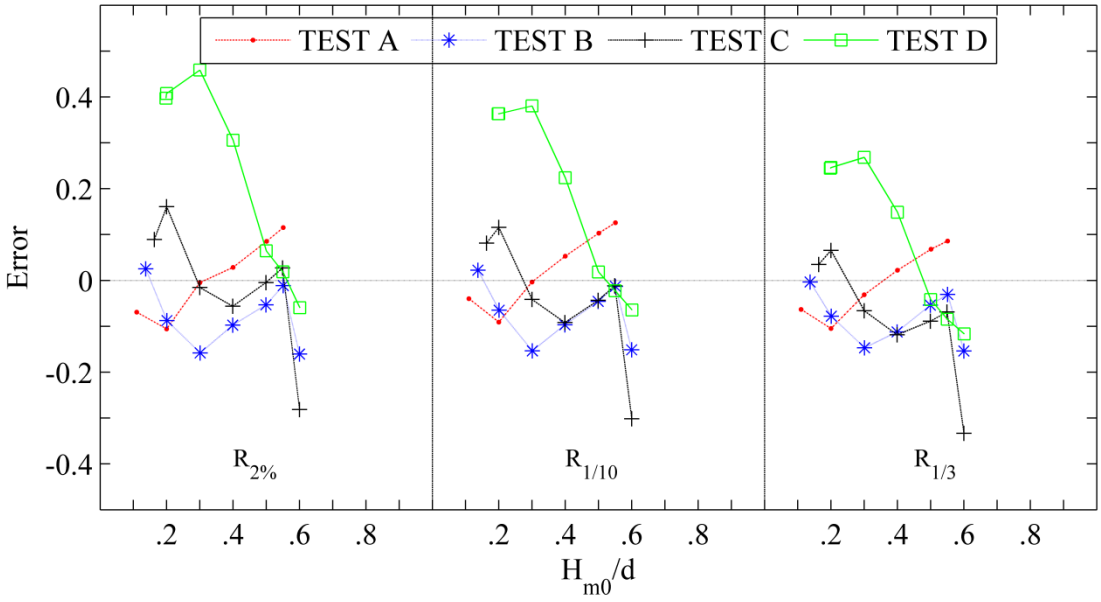

Fig. 5.5 Errors in runup statistics for the coupled SWAN+SWASH model, compared with Mase's (1989) laboratory results, as a function of H_{m0}/d at coupling location. Errors are mean values of 5 randomly phased wave trains prescribed as boundary conditions of SWASH, generated from the same action density spectrum computed by SWAN.

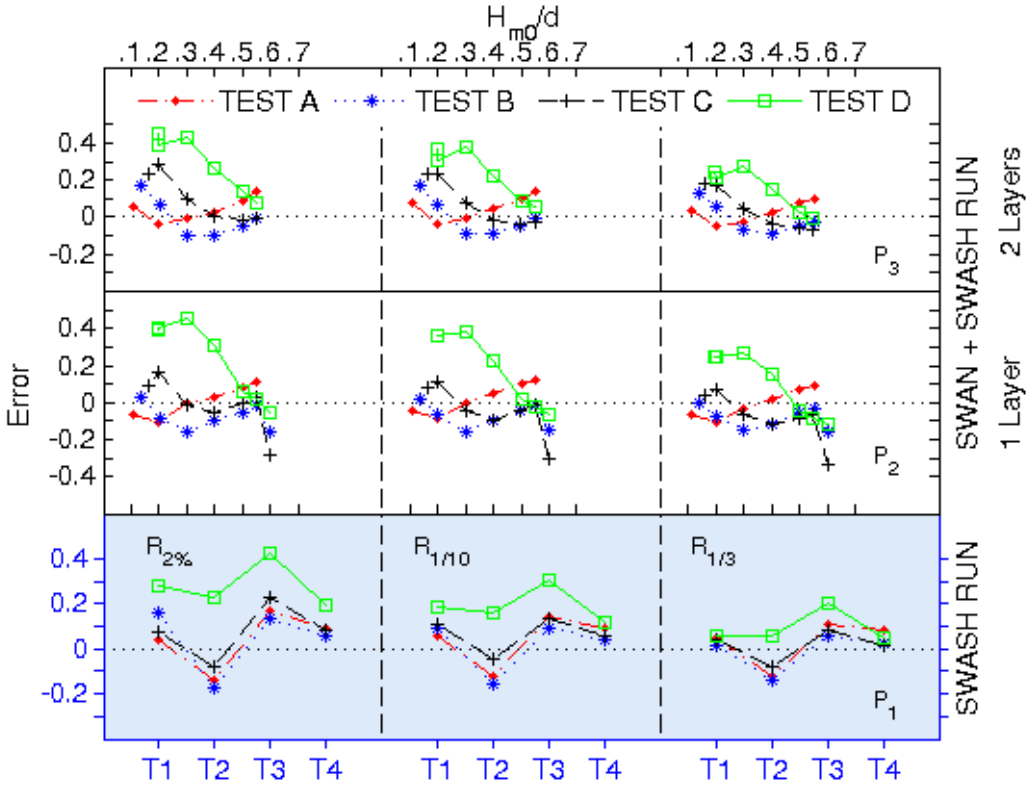

Fig. 5.6 Error in runup statistics (*R2‰*, *$R_{1/10}$*, *$R_{1/3}$*), compared with Mase's (1989) laboratory results: SWASH run throughout the domain (panel P_1); SWAN + SWASH with 1 layer (panel P_2) and 2 layers (panel P_3). In panel P_1: T1 = run with single layer, T2 = single layer plus bound waves at wavemaker; T3 = two layers, and T4 = three layers. Panel P_2 and panel P_3 represent errors as a function of the nonlinear parameter H_{m0}/d. Errors are computed as mean values resulting from 5 randomly phased wave trains prescribed as boundary conditions of SWASH, generated from the same action density spectrum computed by SWAN.

The results are presented in Figures 5.7 5.10, in terms of curves of the relative error for runup statistics $R_{2\%}$, $R_{1/3}$ and $R_{1/10}$ with respect to Mase's experimental data, calculated, as in the previous case, as average values over five runs with different random seeds. In particular four different model combinations are considered: (i) *SWAN+SWASH-NHE*, i.e. without triads and with non-hydrostatic version of *SWASH* (Fig. 5.7); (ii) *SWAN-LTA+SWASH-NHE*, i.e. with triads computed and with non-hydrostatic version of *SWASH* (Fig. 5.8); (iii) *SWAN+SWASH-NLSWE*, i.e. without triads in *SWAN* and with the hydrostatic non linear shallow water version of *SWASH* (Fig. 5.9); (iv) *SWAN-LTA+SWASH-NLSWE*, with triads computed in *SWAN* and non linear shallow water version of *SWASH* (Fig. 5.10).

In general, from comparison with respect to the previous results (i.e. Fig. 5.6), it can be immediately seen as the behavior of the errors is much more regular and also more closed to the shallow water, i.e. very low errors ($|Error|_{R_{2\%}} < 0.05$) can be appreciated within the interval $H_{m0}/d = 0.6 \div 0.7$ for each of the four (i)÷(iv) combinations. Furthermore, it is interesting to observe as the wave conditions denote a consistent feature, since the optimal coupling point is approached with the same regular trend for each combination. In addition, it is to be pointed out as the overall stability of the simulation is highly improved with respect to the previous case, the reason being the different type of open boundary condition adopted.

On the other hand, it is interesting to compare the behavior of the errors with respect to the computation of nonlinear interactions with the *SWAN* model, as well as with regard to the use of *NLSWE* or *NHE* versions. First of all, from comparison between (i)-(ii) and (iii)-(iv) it appears that the errors have a regular trend for both cases without triads, since a better tendency and lower error values appear with respect to the LTA computations. Secondly, the optimal coupling point appears to be located shallower in the *NLSWE* case (iii), whereas is more outward for the *NHE* option (i). These results are consistent with the different performance of these models, since *SWASH* being still able to represent some degree of wave dispersion even with one layer compared to the fully non-dispersive *NLSWE* formulation.

5.3.9 Final discussion

In the present study, an online, one-way coupling between the *SWAN* and *SWASH* numerical models has been introduced. Results shown in this chapter appear to be reasonably in good agreement with the set of runup statistics of random wave experimental data, either using *SWASH* or the coupled model with the *NHE* or *NLSWE* configuration.

It is shown that the *SWASH* domain can be reduced shifting the coupling point shorewards, consequently reducing the computational effort, while still obtaining reliable results.
The choice of the optimal coupling point is found to be dependent on the wave conditions analyzed, with a common location expressed by the value $H_{m0}/d \approx 0.3\text{-}0.55$, whether a previous coupling is adopted, whereas a much more regular tendency can be appreciated by the last improvements and by using a newer different versions of the model components. At this purpose, it should be emphasized the importance of both nonlinear effects and spectral wave model accuracy in very shallow water, where instabilities might arise due to wave-maker algorithm characteristic, which is based on linear wave theory and horizontal bottom, and because of increased wave reflection.

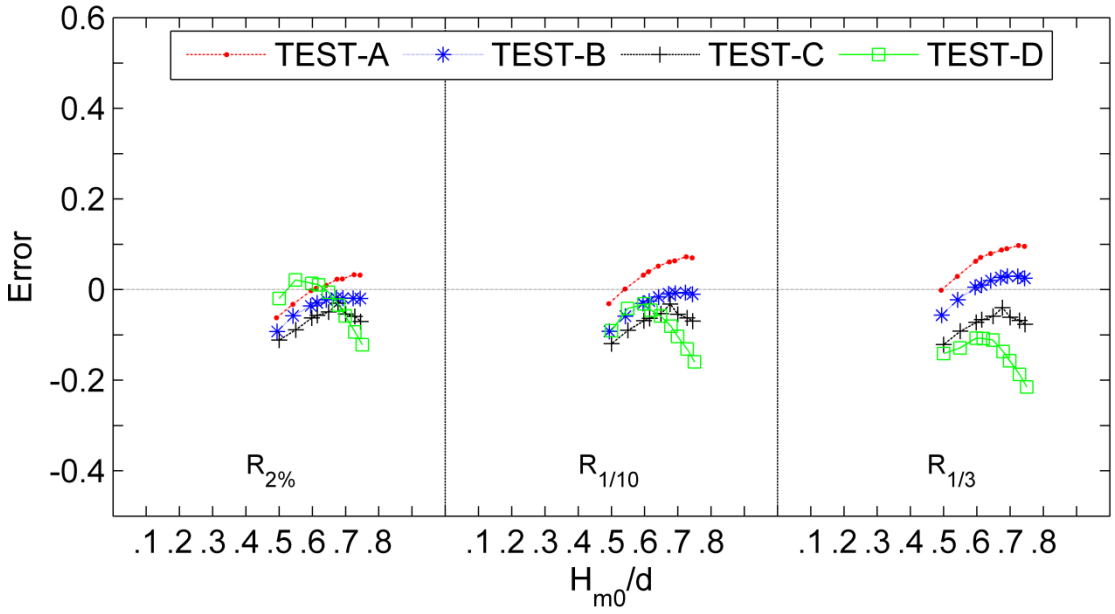

Fig. 5.7 Errors in runup statistics for the coupled SWAN 41.10 + SWASH 4.01 model, compared with Mase's (1989) laboratory results, as a function of H_{m0}/d at coupling location. Errors are mean values of 5 randomly phased wave trains prescribed as boundary conditions of SWASH. SWAN + SWASH-NHE combination, where SWAN is run without triads, whereas SWASH is run in non-hydrostatic mode with one layer.

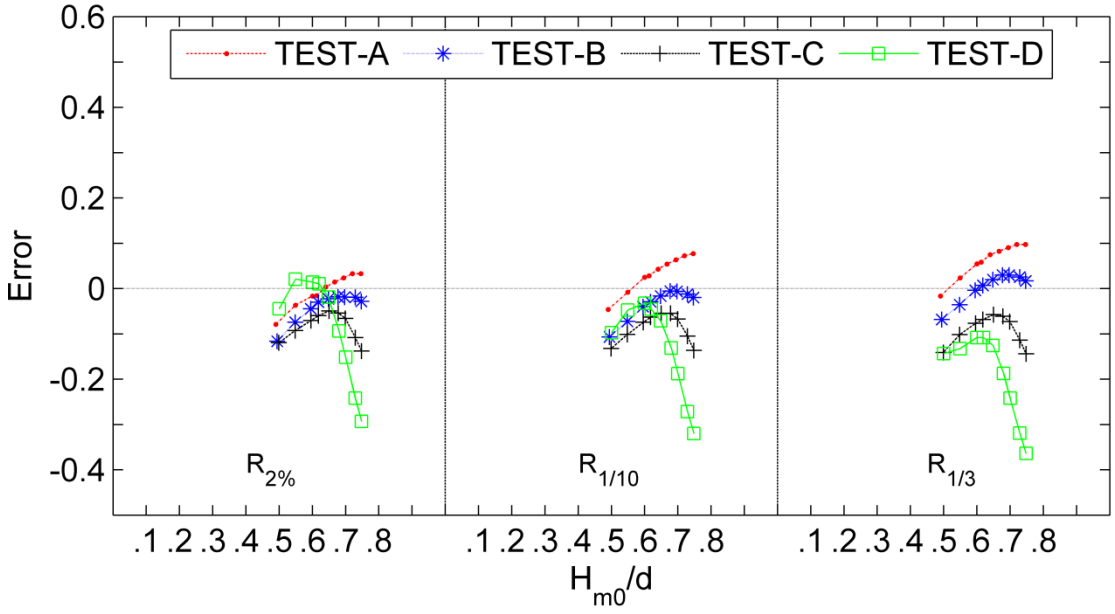

Fig. 5.8 Errors in runup statistics for the coupled SWAN 41.10 + SWASH 4.01 model, compared with Mase's (1989) laboratory results, as a function of H_{m0}/d at coupling location. Errors are mean values of 5 randomly phased wave trains prescribes as boundary condition of SWASH. SWAN-LTA + SWASH-NHE combination. SWAN is run by computing triads, whereas SWASH is run in non-hydrostatic mode with one layer.

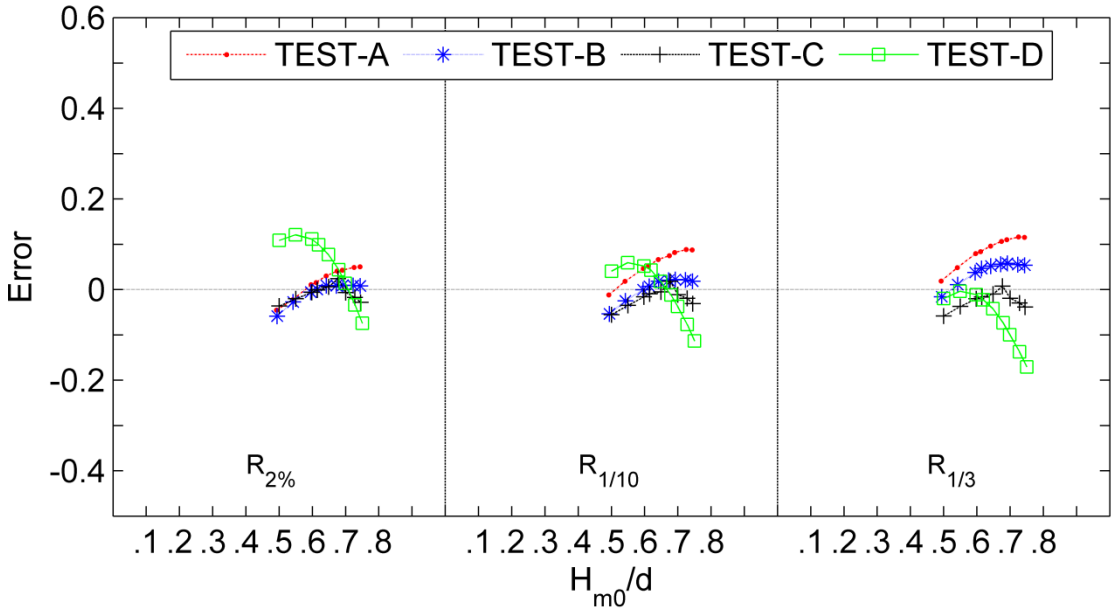

Fig. 5.9 Errors in runup statistics for the coupled SWAN 41.10 + SWASH 4.01 model, compared with Mase's (1989) laboratory results, as a function of H_{m0}/d at coupling location. Errors are mean values of 5 randomly phased wave trains prescribed as boundary conditions of SWASH. SWAN + SWASH-NLSWE. SWAN is run without triads, whereas SWASH is run in hydrostatic mode.

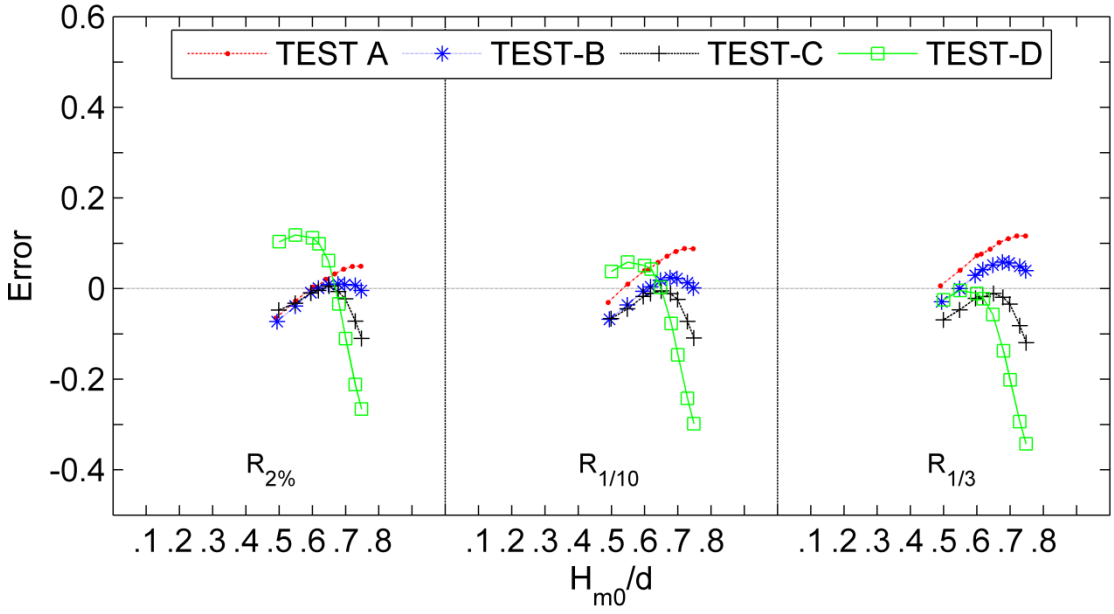

Fig. 5.10 Errors in runup statistics for the coupled SWAN 41.10 + SWASH 4.01 model, compared with Mase's (1989) laboratory results, as a function of H_{m0}/d at coupling location. Errors are mean values of 5 randomly phased wave trains prescribed as boundary conditions of SWASH. SWAN-LTA+ SWASH-NLSWE. SWAN is run with triads, whereas SWASH is run in hydrostatic mode.

Further evaluations may be made including the contribution of incident (bound) infragravity-waves at the inner coupling point. On the other hand, different SWAN configurations as to triplete interactions and depth-induced breaking modelling (Salmon et al. 2015), might be used to validate the results.

6 Conclusions

In this study, both peculiarities and main characteristics associated with two fundamental coastal flooding perspectives have been analysed. In particular, the extreme value analysis is an essential starting point to compute the offshore boundary conditions, as well as numerical modelling is required to simulate wave propagation, transformation and wave inundation from offshore up to the shore.

First of all, the set of procedures for assessing the extreme offshore sea state is addressed by combining both the statistical inference and the EV theory. Secondly, an online coupled model is implemented, with the purpose of obtaining an operational model chain for computing, seamlessly and accurately, an extreme sea state from offshore up to the shore.

From the statistical viewpoint, the general validity of EV theory is partially governed by the combination between sample size and asymptotical hypothesis. According to the classical inference approach, a multi distribution analysis is performed, by assuming three main distributions (GEV, GP, and WBL) to be representative for all possible marginal parent distributions within either the POT, AM or PP methods. Within the POT context, a de-clustering methodology based on the sea storm definition is employed to assure sample independency, which is enhanced by selecting only peak wave heights 48 hours apart; then, a combination between long-term wave climate and fetch distribution is considered in order to identify homogeneous directional clusters. Three different threshold selection methods are used within the POT framework, namely: the automated procedure proposed by Thompson et al. (2009), the approach of Mazas and Hamm (2011) based on assuming a fixed value of storm peaks per year; the multiple threshold method (Deidda, 2010) tailored to rain data, which is slightly adapted to wave sample, with respect to directionality and to independency elements. Furthermore, three main parameter estimation methods are employed – maximum likelihood (ML), maximum product of spacing (MPS), L-moments (LM), for each distribution –; in addition, the ordinary least square (OLS) method proposed by Goda (2010) is considered and associated to the WBL distribution. Finally, the best-fit model is determined by either diagnostic, parametric or non-parametric metrics, represented by the minimum residual of correlation coefficient (*MIR*) computed on both *pp-plot* and *qq-plot*, the Akaike-information-criteria (*AIC*) and the Bayesian-information-criteria (*BIC*), an improved version of the so-called error norm criteria, proposed by Li et al. (2008), and, finally, bias and root mean square error of parameter and quantile estimations obtained by the Monte Carlo technique. Lastly, the uncertainty associated with parameters and quantiles estimation (i.e. confidence intervals) is evaluated by using the Monte Carlo approach previously mentioned.

Comparison between different theoretical methodologies are provided and applied to the most reliable wave buoy dataset, with the primarily cross-purpose of analysing GP and WBL performances. In fact, the driving reason may be linked to the considerable problem posed by authoritative authors. In practice, it is demonstrated that the GP distribution can lead to dangerously low return values when the analysis is carried out for a very short period (Mazas and Hamm 2011), such that the GP distribution is not recommended for engineering applications (Goda et al. 2010).

As a general results of this work, it is assumed that one would be more close to the engineering criterion whenever a larger quantile has to be chosen for a very short period (which, unfortunately, is almost always the routine in Coastal Engineering), where safe design criteria rather than perfect statistical fits must be considered, since neither the best-fit model nor the statistical test have demonstrated an evident outperform of one model with respect to the other.

On the other hand, analysis of some intermediate, but critical, steps is interesting as well. Three-fold considerations can be defined. First, a comparison between mono-threshold and the MTM threshold selection methods is carried out. Second, a comparison between different estimators is performed with respect to both parameters and quantiles estimations. Finally, some parametric and non-parametric best-fit methodologies are compared.

With respect to the threshold selection methods, some general issues can be focused on numerical and practical features. In particular, it is found that the MTM method is the most computational expensive and, even from the practical perspective related to the coastal dataset modelling, may be recognized to be the less favourable, since the applicability appears to be hindered for multiple applications on large dataset. As a matter of fact, the automated threshold method proposed by Thompson et al. (2009), as well as the Mazas and Hamm (2011)'s advice, are assumed to be a pragmatic automated, simple and computationally inexpensive choice.

Regarding the parameter estimation methods, it seems to be more convenient to choose the LM estimator since it generally provides both the best statistical justification and the less differences in the quantile estimations among GP and WBL models. In addition, it is assumed to be the most reliable and robust in tackling problems involving low sample size and outliers. Last but not least, it is found to be relatively simple to implement with respect to others methods, and, most of all, it requires less CPU time to carry out the Monte Carlo technique.

With reference to best-fit metrics, it is found that they sometimes are mutually conflicting, therefore, as a general rule, it appears always more appropriate to carry out a comparison among different typologies instead of using a single metric.

Finally, it is worthwhile underlying that a long term analysis is provided, as well as a deep correction of the original dataset, before the extreme value application. In particular, a comprehensive semi-automated, graphical and numerical methodology is implemented on Matlab routines to aid these preliminary steps. In addition, the extreme value computation is also linked to such previously implemented wave climate analysis and quality control tools. For this purpose, some external facilities, such as the WAFO toolbox (WAFO-group, 2011), have been added to the whole project, in order to compute different parameter estimation methods for either GP, GEV, or WBL distributions. Nonetheless, the fixed-threshold methods (adapted from the work of Thompson et al. 2009, Coles 2001, and Mazas and Hamm 2011), the MTM threshold methods (based on the procedure proposed by Deidda 2010), the LM method (based on the LMOMENT Fortran routine implemented by Hosking and Wallis 1997, as well as on the work of Goda 2010, 2011 with respect to the WBL LM estimators), the best-fit model selection (principally Coles 2001 and Li et al 2008), the Monte Carlo approach, as well as the set of pre- and post-processing functions have been specifically implemented as a final result of this work.

The second purpose of the present study is to introduce an online coupling mechanism between *SWAN* and *SWASH* models, with the aim of obtaining a unique operational model to be used for computing, seamlessly and efficiently, the wave evolution, from offshore up to the shore, including wind-wave generation, nonlinear wave-wave interactions, offshore and near-shore dissipations.

Despite excellent wave-dispersion and noticeable nonlinearity have been nowadays reached, the computational burden of phase-resolving models is still excessive to consider application outside the near-shore area as a routine in operational models. Therefore, a phase-averaged approach might be adopted there, at expense of an overall decreasing of accuracy, but with still good provisional results at a lower computational price. A reasonable solution would be to carry out practical application using phase-resolving models in synergy with spectral models, where the first is reduced to nearshore zones, while the second is extended to larger areas. To a large extent, the computational cost required to propagate wave from offshore to near-shore would be largely reduced, while the accuracy in surf zone should be well-described by the phase-resolving model.

From the practical viewpoint, the coupling is obtained forcing the open offshore and lateral boundaries of the *SWASH* domain by directional random wave fields generated through the embedded wave-maker algorithm included in the *SWASH* code, based on the action density spectra computed by *SWAN*. This assures the continuity of the information, while preventing spurious reflection of outgoing waves.

Both models can be run either on Cartesian or curvilinear structured grids with different spatial resolutions, with the *SWAN* domain being typically larger than the *SWASH* domain. Stationary and non-stationary versions have been implemented, where *SWAN* and *SWASH* run sequentially in time. The *SWAN* model is run first, and the coupling interval is taken to be equal to the *SWAN* time step, since the *SWAN* sweeping solution algorithm can take much larger time steps than *SWASH*, which is Courant-time-step limited due to its explicit formulation. In addition, in the non-stationary case, the *SWASH* model can be forced either with the final sea state or through an averaged status, which might be able to ensure mean wave conditions for the whole coupling period.

A validation of the coupled model with laboratory data is provided. Numerical results indicate a fairly good agreement of computed runup statistics with data from unidirectional random wave laboratory tests. The coupled model proved effective in reducing the *SWASH* domain extent, thus reducing the overall modelling effort, while retaining outcomes' accuracy, suggesting that it can be a comprehensive and valuable tool for both engineering and scientific purposes. The choice of the optimal coupling point is found to be dependent on both the wave conditions and the model configuration, as well as on the type of open boundary condition used to force the *SWASH* model at the inner boundary. In particular, a common value of $H_{m0}/d \approx 0.55$ is identified to be the best for three test cases (characterized by a spilling breaker), while a lower value of $H_{m0}/d \approx 0.30$ is recognized for the case with a plunging breaker, whether a hyperbolic profile is chosen to force the SWASH model. On the other hand, better results are obtained with an improvement of the coupling, by using a more recent version of the model components (i.e. *SWAN 41.10* and *SWASH 4.01*) and by forcing *SWASH* through a band-pass filtered Fourier series of levels, automated obtained from the *SWAN* spectra, and imposing, subsequently, the horizontal velocity profile as uniformly-distributed. This option is found to be fairly adequate for simulations in shallow and very shallow

waters. In addition, a high stability is achieved in the coupling and a high degree of flexibility is obtained since *SWASH* may be run both in non-hydrostatic depth-averaged (*SWAN-NHE*) and hydrostatic depth-averaged (*SWAN-NLSWE*) mode. In both cases, very low errors (less than 5 %) are found within the interval $H_{m0}/d = 0.6 \div 0.7$ and, most of all, a consistent trend is noted for the wave conditions analyzed, which is independent on the model combination adopted. Moreover, results are found to be consistent with the different performance of *SWASH-NHE* and *SWASH-NLSWE*. Finally, it is to be noted that these values are comparable with results obtained by McCabe et al. (2011), where a value of $H_{m0}/d \approx 0.65$ was found to represent the optimal switching point in their *SWAN+NLSWE* model. Nonetheless, it is worthwhile underlying that neither ad hoc smoothing of the bathymetry is pursued herein nor a manual procedure is needed to pass the *SWAN* spectra at the coupling point, in contrast to McCabe et al. (2011). This is reasonable due to the different type of wave-maker and boundary conditions adopted, as well as to the different type of coupling (i.e. offline *SWAN + NLSWE*).

Running a wider variety of test cases with different model configurations is required to fully validate the proposed approach. For instance, further evaluations can be made by including the contribution of incident (bound) infragravity waves at the coupling point. On the other hand, different *SWAN* configurations as to different depth-induced wave breaking and triplet interactions, might be used to further validate the results.

Furthermore, it is to be remarked that *SWASH* domain may be selectively assigned and extended up to the *SWAN* domain. In this sense, a similar procedure as stated by Diaz-Hernandez et al. (2015) may be achieved. Shortly, the authors have modified analytically the offshore *SWAN* spectra to include the energy input associated with infra-gravity waves, in order to obtain a full spectrum of short and ig-waves; then, a hybrid clustering-numerical wave propagation downscaling technique is forced with the computed full spectra to produce a set of nearshore boundary conditions. As a matter of fact, such approach may be automatically carried out by the coupled-model, whether the coupling point is located in intermediate water and the weakly-nonlinear second order contribution is included at the inner boundary between both models.

On the other hand, the *MPI* (Message Passing Interface) modularity has provided key aspects to the parallel implementation of the coupled model. Although an emerging practice is to couple heterogeneous models through a generic framework – for instance, the Modelling Coupling Toolkit (*MCT*, Larson et al. 2005, Jacob et al. 2005, Warner et al. 2008) –, it seemed to be more convenient working on the source code of both models. Specifically, it was chosen to develop the parallel implementation according to *ADCIRC + SWAN* coupled model approach (Dietrich et al 2011). Although its proven highly-scalable performances, this scheme cannot be completely followed in *SWAN + SWASH*, because, to save computational time, it is necessary to have grids with different dimensions and resolutions. Nevertheless, adequate local characteristic and efficiency can be reached as well. Communication between models (intra-model) can be passed through local memory or cache, without the need for any network-based inter-core communication.

On the whole, it is assumed to be more appropriate to duplicate the standard *MPI* communicator, without adding other inter-communicators. Basically, each model component was already developed within the parallel framework, so the *MPI*'s *MPMD* (Multiple Program Multiple Data)

parallel paradigm can be generally adopted to implement a unique parallel code, while ensuring safety and local characteristics. It is also to be noted that a series of tests on a specific cluster should be carried out to better evaluate the parallel performance.

Last but not least, the phase-averaged – phase-resolving paradigm obtained by the *SWAN+SWASH* coupling, can be conveniently employed to simulate large-scale flow and transport phenomena driven by tidal, wind and buoyancy forces. In principle, this can be done due to the meteorological, baroclinic forcing and solute transport extensions included in the *SWASH* code. At this purpose, a tight coupling may be further implemented, where also levels and velocities computed by *SWASH* can be passed to *SWAN* at each time step.

Alternatively, other extensions can be obtained by adding *SWAN + SWASH* to other classes of coupled models, i.e. the *SWAN + ADCIRC* model (Dietrich et al. 2011), the *SWASH-SPH* model (i.e. Altomare et al. 2015, Crespo et al. 2015), in order to obtain a more efficient and comprehensive modelling of coastal areas.

References

Medatlas Group. (2004). *Wind and Wave Atlas for the Mediterranean Sea.* W.E.A.O: Research Cell.

Agnon, Y., Madsen, P. A., & Schäffer, H. A. (1999). A new approach to high-order Boussinesq models. *J. Fluid Mech. , 399*, 319-333.

Akaike, H. (1974). Information theory and an extension of the maximum likelihood principle. *Second international symposium on information theory* (pp. 267-281). Budapest: Akademiai Kiado.

Akram, M., & Hayat, A. (2014). Comparison of estimators of the Weibull distribution. *Journal of Statistical Theory and Practice , 8* (2), 238-259.

Altomare, C., Oshima, Y., Chen, X., Crespo, A., & Suzuki, T. (2015). Study of the overtopping flow impacts on multifunctional sea dikes in shallow foreshores with an hybrid numerical model. *E-Proc. 36th IAHR World Congress.* The Hauge, the Netheralnds.

Baffo, F., Desiato, F., Lena, F., Suatoni, B., Toreti, A., Bider, M., et al. (2005a). Criteri di calcolo degli indicatori meteo climatici. Sistema SCIA, ISPRA.

Baffo, F., Suatoni, F., & Desiato, F. (2005b). Indicatori climatici: i controlli di validità e la ricerca dei valori errati. 31-41. Sistema SCIA, ISPRA: Bollettino Geofisico, anno XXVIII, n. 1-2.

Balkema, A. A., & De Haan, L. (1974). Residual life time at great age. *The Annals of Probability , 2* (5), 792-804.

Balzano, A., De Falco, G., Simeone, S., Sulis, A., Antognarelli, F., Massaro, G., et al. (2013). Linea B - Morfodinamica della Spiaggia di S. Giovanni del Sinis. In A. Abis, C. Campo, M. Careddu, & M. Deriu (Eds.), *La rete per il monitoraggio delle spiaggie, applicazioni ed esperienze sviluppate nel progetto Resmar (azione di sistema A)* (pp. 98-136). Olbia: Taphros.

Battjes, J., & Janssen, J. (1978). Energy loss and set-up due to breaking of random waves. *Proc. 16th Int. Conf. Coastal Engng* (pp. 569-587). ASCE.

Becq-Girard, F., Forget, P., & Benoit, M. (1999). Non-linear propagation of unidirectional wave fields over varying topography. *Coastal Engng. , 38*, 91-113.

Beji, S., & Nadaoka, K. (1996). A formal derivation and numerical modelling of the improved Boussinesq equations for varying depth. *Ocean Engng , 23* (8), 691-704.

Bencivenga, M., Nardone, G., Ruggiero, F., & Calore, D. (2012). The Italian data buoy network (RON). (pp. 121-133). Cardiff (UK): Proc. 28th ICCE.

Benjamin, T. B., Bona, F. R., & Mahony, J. J. (1972). Model equations for long waves in nonlinear dispersive systems. *Phil. Trans. R. Soc. Lond. A , 272*, 47-78.

Benjamin, T. (1972). The stability of solitary waves. *Proc. Roy. Soc. Lond. A* .

Benoit, M., Marcos, F., & Becq, F. (1996). Development of a third generation shallow-water wave model with unstructured spatial meshong. *Proc. 25th Int. Conf. Coastal Engng.* (pp. 465-478). ASCE.

Blayo, E., & Debreu, L. (2005). Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean Modelling , 9*, 231-252.

Boccotti, P. (2000). Wave mechanics for ocean engineering. Netherlands: Elsevier Oceanograohy Series, 64.

Bonnetton, P., Barthelemy, E., Chazel, F., Cienfuegos, R., Lannes, D., Marchef, F., et al. (2011). Recent advances in Serre-Green Naghdi modelling for wave transformation, breaking and runup processes. *European J. Mechanics B/Fluids , 30*, 589-597.

Booij, N. R. (1999). A third generation wave model for coastal regions. 1. model description and validation. *Journal of Geophysical Research , 104* (C4), 7649-7666.

Bosom, E., & Jiménez, J. (2011). Probabilistic coastal vulnerability assessment to storms at regional scale - application to Catalan beaches (NW Mediterranean). *Nat. Hazards and Earth Syst. Sc. ,* 475-484.

Boussinesq, J. (1872). Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *J. Mathématiques Pures et Appliquées , 17* (2), 55-108.

Briggs, M., ynolakis, C., Harkins, G., & Green, D. (1995). Laboratory experiments of tsunami run-up on circular island. *Pur Appl. Geophys. , 114* (3/4), 569-593.

Brocchini, M. (2013). A reasoned overview on Boussinesq-type-models: the interplay between phisics, mathematics and numerics. *Proc. R. Soc. A , 469* (2160), 1-27.

Brocchini, M., & Dodd, N. (2008). Nonlinear shallow water equation modelling for coastal engineering. *J. Waterway, Port, Coastal, Ocean Eng. , 134*, 104-120.

Brocchini, M., Bernetti, R., Mancineeli, A., & Albertini, G. (2001). An efficient solver for nearshore flows based on the WAF method. *Coast. Engng , 43* (2), 105-129.

Brocchini, M., Svendsen, I., Prasad, R., & Bellotti, G. (2002). A comparison of two different types of shoreline boundary conditions. *Comp. Methods. Appl. Mech. Eng. , 191*, 4475-4496.

Caires, S. (2007). *Extreme wave statistics. Methodology and applications to North Sea wave data.* Delft: WL Delft Hydraulics.

Castillo, E., & M., S. J. (1992). Engineering analysis of extreme value data: selection of models. *Journal of Waterway, Port, Coastal, and Ocean Engineering , 118:2* (129), 129-146.

Castillo, E., Hadi, A. S., Balakrishnan, N., & Sarabia, J. M. (2004). *Extreme value and related models with applications in engineering and science.* Canada: Wiley Series on Probability and Statistics.

Cavaleri, L., & Bertotti, L. (2004). Accuracy of the modelled wind and wave fields in enclosed seas. *Tellus , 56A*, 167-175.

Chazel, F., Lannes, D., & Marche, F. (2010). Numerical simulation of strongly nonlinear and dispersive waves using a Green-Naghdi model. *J. Sci. Comput.*

Chazell, F., Benoit, M., Ern, A., & Piperno, S. (2009). A double-layer Boussinesq-type model for highly nonlinear and disperive waves. *Royal Society .*

Chen, Q. (2006). Fully nonlinear Boussinesq-type equations for waves and currents over porous beds. *J. of Engng. Mech.* , 219-230.

Chen, Q., Kirby, J., Dalrymple, R., Kennedy, A., & Chawla, A. (2000). Boussinesq modeling of wave transformation, breaking, and runup II: 2D. *J. Wtrwy, Port, Coastal and Ocean Engng , 126* (1), 48-56.

Chen, Q., Kirby, J., Dalrymple, R., Kennedy, A., Thornton, E., & Shi, F. (2001). Boussinesq modelling of waves and alongshore currents under field conditions. *Proc. 27th ICCE* (pp. 651-663). Reston, Va: ASCE.

Cheng, R. C., & Amin, N. A. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society , 45* (3), 394-403.

Cheung, K., Phadke, A., Wei, Y., Rojas, R., Douyer, Y., Martino, M., et al. (2003). Modeling of storm induced coastal flooding for emergency management. *Ocean Engng , 30*, 1353-1386.

Choi, B. H., Hong, S. J., & Pelinovsky, E. (2006). Distribution of runup heights of the December 26, 2004 tsunami in the Indian Ocean. *Geophys. Res. Lett. , 33*, L13601.

Choi, B. H., Kim, D. C., Pelinovsky, E., & Woo, S. B. (2007). Three-dimensional simulation of tsunami run-up around conical island. *Coast. Eng.* , 618-629.

Choi, B. H., Pelinovsky, E., Kim, D. C., Didenkulova, I., & Woo, S. B. (2008). Two- and three-dimensional computation of solitary wave runup on non-plane beach. *Nonlin. Processes Geophys , 15*, 489-502.

Christensen, E. D. (2006). Large eddy simulation of spilling and plunging breakers. *Coastal Eng. , 53*, 463-485.

Coles, S. (2001). An introduction to statistical modeling of extreme values. London: Springer-Verlag.

Coles, S., & Simiu, E. (2003). Estimating uncertainty in the extreme value analysis of data generated by a hurricane simulation model. *J. of Engn. Mech , 129* (11), 1288-1294.

Collins, J. (1972). Prediction of shallow water spectra. *J. Geophys. Res. , 15*, 2693-2707.

Corsini, S., Inghilesi, R., Franco, L., & Pisocopia, R. (2004). *Atlante delle onde nei mari italiani. Italian wave atlas.* Roma: ISPRA, Università degli Studi Roma Tre.

Cousineau, D. (2009). Fitting the three-parameter weibull distribution: review and evaluation of existing and new methods. *IEEE Dieletrics and Electrical Insulation Society , 16* (1), 281-288.

Craik, A. (2004). The origins of water wave theory. *Annu. Rev. Fluid Mech. , 36*, 1-28.

Crspo, A., Altomare, C., Dominguez, J., Suzuki, T., Verwaest, T., & Gomez-Gesteira, M. (2015). SPH modelling in coastal engineering. *36th IAHR World Congress.* The Hauge, the Netherlands.

Cui, H., Pietrzak, J., & Stelling, G. (2012). Inproved efficiency of a non-hydrostatic, unstructured grid, finite volume model. *Ocean Modell. , 54-55*, 55-67.

Cui, H., Pietrzak, J., & Stelling, G. (2014). Optimal dispersion with minimized Poisson equations for non-hydrostatic free surface flows. *Ocean Modell. , 81*, 1-12.

Cunnane, C. (1973). A particular comparison of annual maxima and partial duration series methods of flood frequency predictions. *Journal of Hydrology , 18*, 257-271.

Cunnane, C. (1978). Unbiased plotting position - A review. *Journal of Hydrology , 37*, 205-222.

D'Alessandro, F., & Tomasicchio, G. (2008). The BCI criterion for the initiation of breaking process in Boussinesq-type equations wave models. *Coastal Engng , 55*, 1174-1184.

Dalrymple, R. A., Knio, O., Cox, D. T., Gomez-Gesteira, M., & Zou, S. (2001). Using a Lagrangina particle method for deck overtopping. *Proc. Waves, American Society of Civil Engineers* , 1082-1091.

Davison, A. &. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Sociery. Series B (Methodological) , 52* (3), 339-442.

Dean, R. (1990). Equilibrium beach profiles: charateristics and applications. *J. Coast. Res. , 7* (1), 53-84.

Deidda, R. (2010). A multiple threshold method for fitting the generalized Pareto distribution and a simple representation of the rainfall process. *Hydrology and Earth System Sciences , 14*, 2559-2575.

Diaz-Hernandez, G., Mendez, F., Losada, I. J., Camus, P., & Median, R. (2015). A nearshore long-term infragravity wave analysis for open harbours. *Coastal Engng , 97*, 78-90.

Dietrich, J. C., Zijlema, M., Westerink, J. J., Holthuijsen, L. H., Dawson, C., Luettich Jr., R. A., et al. (2011). Modeling hurricane waves and storm surge using integrally-coupled, scalable computations. *Coastal Engineering , 58*, 45-65.

Dingemans, M.-W. (1997). *Water wave propagation over uneven bottoms.* Delft.

Dodd, N. (1998). Numerical model of wave run-up, overtopping and regeneration. *J. of Waterways, Port, Coastal and Ocean Engng , 124* (2), 73-81.

Donahue, A., Kennedy, A., Westerinka, J., Zhang, Y., & Dawson, C. (2016). Simulation of wave phenomena in the nearshore through application of of O(μ2) and O(μ4) pressure-Poisson Boussinesq type models. *Coastal Engng , 114*, 61-76.

Donahue, A., Zhang, Y., Kennedy, A., Westerink, J., Panda, N., & Dawson, C. (2015). A Boussinesq-scaled, pressure-poisson water wave model. *Ocean Modelling , 86*, 36-57.

Dutykh, D., Katsaounis, T., & Mitsotakis, D. (2011). Finite volume schemes for dispersive wave propagation and runup. *J. of computational Physics , 231*, 3035-3061.

Eldeberky, Y. (1996). *Nonlinear transformation of wave spectra in the nearshore zone.* Ph.D. thesis, Delft University of Technology, Department of Civil Engineering, Delft, The Netherlands.

Erduran, K. S. (2007). Further application of hybrid solution to another form of Boussinesq equations and comparisons. *Int. J. Numer. Meth. Fluids , 53* (5), 827-849.

Erduran, K. S., Ilic, S., & Kutija, V. (2005). Hybrid finite-volume finite-difference scheme for the solution of Boussinesq equation. *Int. J. Numer. Meth. Fluids , 49*, 1213-1232.

Falconer, R. A. (1993). An introduction to nearly-horizontal flows. In E. &. Ltd (Ed.), *Coastal, Estuarial and Harbour Engineers' Reference Book* (pp. 27-36). London: Abbott, M. B & Price, W. A., eds.

Fang, K., Zhang, Z., Zou, Z., & Liu, Z. (2014). Modelling of 2D extended Boussinesq equations using a hybrid muerical scheme. *J. of Hydrodynamics , 26* (2), 187-198.

Filliben, J. J. (1973). Techniques for tail length analysis. *Proceedings of the Conference on the design of experiments in army research development and testing (18th)* (pp. 425-450). Aberdeen: U. S. Army Research Office.

Filliben, J. (1975). The Probability Plot Correlation Coefficient Test for Normality. *Technometrics , 17* (1), 111-117.

Fisher, R. A., & Tippet, L. H. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society , 24*, 180-190.

Foster, I. (2003). *Designing and building parallel programs: concepts and tools for parallel software engineering.* Boston: Addison Wesley Longman Publishing.

Fuhrman, D., & Madsen, P. (2008). Simulation of nonlinear wave run-up with a high-order Boussinesq model. *Coastal Engng. , 55*, 139-154.

Geist, E., Lynett, P., & Chaytor, J. (2009). Hydrodynamic modeling of tsunamis from the Currituck landslide. *Marine Geology , 264*, 41-52.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatorie. *Annals of Mathematics. Second series , 44* (3), 423-453.

Gobbi, M., & Kirby, J. (1999). Wave evolution over submerged sills, tests of a high order Boussinesq model. *Coastal Engng , 37* (1), 57-96.

Gobbi, M., Kirby, J., & Wei, G. (2000). A fully nonlinear Boussinesq model for surface waves. Part 2: extension to O(kh)4. *J. Fluid Mech , 405*, 181-210.

Goda, Y. (1979). *A review on statistical interpretation of wave data.* Japan: Port and Harbour Research Institute.

Goda, Y. (2011b). Inherent negative bias of quantile estimates of annual maximum data to sample size effect: a numerical simulation study. *Coastal Engineering Journal , 53* (4), 397-429.

Goda, Y. (1988). On the methodology of selecting design wave height. *1*, pp. 899-913. Costa del Sol-Malaga, Spain: Proceedings, 21st International Conference on Coastal Engineering.

Goda, Y. (2011a). Plotting.position estimator for the L-Moment method and quantile confidence interval for the GEV, GPA, and Weibull distributions applied for extreme wave analysis. *Coastal Engineering Journal , 53* (2), 111-149.

Goda, Y. (2010). Random seas and design of maritime structures. Advanced Series on Ocean Engineering. *33* . Singapore: World Scientific.

Goda, Y., & Kobune, K. (1990). Distribution function fitting for storm wave data. (pp. 18-31). Delft - The Netherlands: Conference of Coastal Engineering.

Goda, Y., Kudaka, M., & Kawai, H. (2010). Incorporation of Weibull distribution in L-Moments method for regional frequency analysis of Peaks-Over-Treshold wave heights. Proceedings of 32nd International Conference on Coastal Engineering.

Gotoh, H., Shao, S., & Memita, T. (2004). SPH-LES model for numerical investigation of wave interaction with partially immersed breakwater. *Coastal Engineering of Japan , 46* (1), 39-63.

Green, A. E., & Naghdi, P. M. (1976). A derivation of equations for wave propagation in water of variable depth. *J. Fluid Mech. , 78*, 237-246.

Groeneweg, J., Doorn, N., Borsboom, M., & Van Dongeren, A. (2004). Near-shore wave modeling with two coupled models: SWAN and TRITON. *Proc. 29th ICCCE* (pp. 868-880). ASCE.

Group, T. W. (1988). The WAM model - A third generation ocean wave prediction model. *Journal of Physical Oceanography , 18*, 1775-1810.

Guanche, R., Losada, I., & Lara, J. (2009). Numerical analysis of wave loads for coastal structure stability. *Ocean Engng , 56*, 543-558.

Guimaraes, P., Farina, L., Toldo-Jt, E., Diaz-Hernandez, G., & Akhmatskaya, E. (2015). Numerical simulation of extreme wave runup during storm events in Tramandai Beach, Rio Grande Do Sul, Brazil. *Coastal Engng , 95*, 171-180.

Gumbel, E. (1958). *Statistics of extremes.* New York: Columbia University Press.

Ha, T., & Cho, Y. (2015). Tsunami propagation over varying water depths. *Ocean Engng , 101*, 67-77.

Harlow, F., & Welch, J. (1965). Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids , 8*, 2182-2189.

Hasselmann, K. (1974). On the spectral dissipation of ocean waves due to whitecapping. *Bound. Layer Meteor. , 1-2*, 107-127.

Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., et al. (1973). Measurement of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Dtsch. Hydrogr. Z. Suppl. , 12* (A8).

Hasselmann, S., Hasselmann, K., Allender, J., & Barnett, T. (1985). Computations and parameterizations of the nonlinear transfer for application in wave models. *J. Phys. Oceanogr. , 15* (11), 1378-1391.

Hawkes, P. J. (2008). Joint probability analysis for estimation of extremes. *Journal of Hydraulic Research , 46* (Extra Issue 2), 246-256.

Hazen, A. (1914). Storage to be provided in impounding reservoirs. *Trans. Amer. Soc. Civ. Eng. Pap. , 1308*, 1547-1550.

Hellies, M. (2016). *Extreme rainfall regime characterization in Sardinia using daily rainfall data.* PhD Thesis: University of Cagliari, DICAAR.

Hirt, C. W., & Nichols, B. D. (1981). Volume of fluid (VOF) methods for the dynamics of free boundaries. *J. Comput. Phys , 39*, 201-225.

Holthuijsen, L. H. (2007). *Waves in Oceanic and Coastal Waters.* Cambridge: Cambridge University Press.

Holthuijsen, L., Herman, A., & Booij, N. (2003). Phase-decoupled refraction-diffraction for spectral wave models. *Coastal Engng , 49*, 291-305.

Hosking, J. R. (1990). L-Moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological) , 52* (1), 105-124.

Hosking, J. R., & Wallis, J. R. (1987). Parameter and quantile estimation for the Generalized Pareto Distribution. *Technometrics , 29* (3), 339-349.

Hosking, J. R., & Wallis, J. R. (1997). Regional frequency analysis: an approach based on L-Moments. Cambridge: Cambridge University Press.

Hsiao, S. C., Lynett, P., Hwung, H. H., & Liu, P. L. (2005). Numerical simulation of nonlinear short waves using a multilayered model. *J. of Eng. Mech. , 131* (3).

Hsu, T., Sakakiyama, T., & Liu, P. (2002). A numerical model for wave motions and turbulence flow in front of a composite breakwater. *Coastal Engng , 46*, 25-50.

Hu, C. G., Mingham, & Causon, D. (2000). Numerical simulation of wave overtopping of coastal structures using the nonlinear shallow water equations. *Coast. Eng, , 41*, 433-465.

Hubbard, M., & Dodd, N. (2002). A 2-D numerical model of waver run-up and overtopping. *Coast. Engng , 47*, 1-26.

Hur, D. S., & Mizutani, N. K. (2004). Direct 3-D numerical simulation of wave forces on asymmetric structures. *Coast. Eng. , 51*, 407-420.

Hur, D., & Mizutani, N. (2003). Numerical estimaton of the wave forces acting on a three-dimnesional body on submerged breakwater. *Coastal Engng , 47*, 329-345.

Hur, D., Kim, C., Kim, D., & Yoon, J. (2008). Simulation of the nonlinear dynamic interactions between waves, a submerged breakwater and the seabed. *Ocean Engng , 35* (5-6), 511-522.

ISPRA (formerly APAT). (2004). *Atlante delle coste - il moto ondoso al largo delle coste italiane.* Technical Report: In Italian.

Iwata, K., Kawasaki, K., & Kim, D. (1996). Breaking limit, breaking and post breaking wave deformation due to submerged breakwaters. *Proc. 25th Int. Coastal Eng. Conf.*, (pp. 2338-2351).

Jacob, R., Larson, J., & Ong, E. (2005). M x N communication and parallel interpolation in community climate system model version 3 using the model coupling toolkit. *Intern. J. of High Performance Comp. Applications , 19* (3), 293-307.

Jeffreys, E. (1987). Directional seas should be ergodic. *Applied Ocean Res. , 9* (4), 186-191.

Jenkinson, F. (1954). The frequency distribution of the annual maxima (or minimum) values of meterorolgical elements. *Q.J,R. Meteorol. Soc. , 81*, 158-171.

Jonathan, P., & Ewans, K. (2013). Statistical modelling of extreme ocean environments for marine design: a review. *Ocean Engineering , 62*, 91-109.

Karambas, T., & Koutitas, C. (1992). A breaking wave propagation model based on the Boussinesq equations. *Coastal Engineering , 18* (1-2), 1-19.

Katz, R. W., Brush, G. S., & B., P. M. (2005). Statistics of extremes: modeling ecological disturbances. *Ecology , 86* (5), 1124-1134.

Katz, R. W., Parlange, M. B., & P., N. (2002). Statistics of extremes in hydrology. *Advances in Water Resources , 25*, 1287-1304.

Kazolea, M., & Delis, A. I. (2013). A well-balanced shock-capturing hybrid finite volume-finite difference numerical scheme for extended 1D Boussinesq models. *Applied Numerical Math. , 67*, 167-186.

Kennedy, A., Chen, Q., Kirby, J., & Dalrymple, R. (2000). Boussinesq modeling of wave transformation, breaking and runup I: 1D. *J. Wtrwy, Port, Coastal and Ocean Engng , 126*, 39-47.

Kirby, J. T. (2003). Boussinesq models and applications to nearshore wave propagation, surf zone processes and wave-induced currents. In V. C. Lakhan (Ed.), *Advances in coastal modeling* (pp. 1-41). Amsterdam, AE: Elsevier.

Kirby, J. T. (1997). Nonlinear, dispersive long waves in water of variable depth. In J. N. Hunt (Ed.), *Gravity waves in water of finite depth* (pp. 55-125). Southampton, UK: Computational mechanics pubblications.

Kirby, J., Shi, F., Tehranirad, B., Harris, J., & Grilli, S. (2013). Dispersive tsunami waves in the ocean: model equations and sensitivity to dispersion and Coriolis effects. *Ocean Modelling , 62*, 39-55.

Kobayashi, N., & Raichle, A. (1994). Irregular wave overtopping of revetments in surf zones. *J. of Waterway, Port, Coastal and Ocean Engng , 120* (1), 56-73.

Kobayashi, N., & Wurjanto, A. (1992). Irregular wave setup and runup on beaches. *J. Wtrwy., Port, Coast. and Ocean Engng , 111* (4), 368.

Kobayashi, N., & Wurjanto, A. (1989). Wave transmission over submerged breakwaters. *J. of Waterway, Port, Coastal and Ocean Engng , 115*, 662-680.

Kobayashi, N., Farhadsadeh, A., & Melby, J. A. (2010). Wave overtopping and damage progression of stone armor layer. *J. of Waterway, Port, Coastal and Ocean Engng , 136* (5), 257-265.

Komen, G., Cavaleri, L., Donelan, M., Hasselmann, K., Hsselmann, S., & Janssen, P. (2004). *Dynamics and modelling of ocean waves.* Cambridge: Cambridge University Press.

Kotz, S., & Nadarajah, S. (2000). *Extreme value distributions. Theory and applications.* Singapore: Imperial College Press.

Kuik, A. V. (1988). A method for the routine analysis of pitch-and-roll buoy wave data. *Journal of Physical Oceanography , 18*, 1020-1034.

Lang, M., Ouarda, T. B., & Bobee, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology , 225*, 103-117.

Lara, J. L., Losada, I., Del Jesus, M., Barajas, G., & Guanche, R. (2010). IH-3VOF: a three-dimensional Navier-Stokes model for wave and structure interaction. *Proc. of the 32nd Int. Conf on Coastal Engng.* ASCE.

Lara, J., Losada, I., & Guanche, R. (2008). Wave interaction with low mound breakwaters using a RANS model. *Ocean Engng , 35*, 1388-1400.

Larson, J., Jacob, R., & Ong, E. (2005). The model coupling toolkit: a new fortran90 toolkit for building multiphysics parallel coupled models. *Intern. J. of High Performance Comput. Applications , 8* (19), 277-292.

Launder, B., & Spalding, D. (1974). The numerical computation of turbulent flows. *Computer Meth. in Applied Mech. and Engng. , 3* (2), 269-289.

Leveque, R. J. (2002). *Finite volume methods for hyperbolic problems.* Cambridge: Cambridge University Press.

Li, F., Bicknell, C., Lowry, R., & Li, Y. (2012). A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset. *Coastal Engineering , 69*, 1-11.

Li, Y., Simmonds, D., & Reeve, D. (2008). Quantifying uncertainty in extreme values of design parameters with resampling techniques. *Ocean Engng. , 35*, 1029-1038.

Liang, D., Gotoh, H., Khayyer, A., & Chen, J. (2013). Boussinesq modelling of solitary wave and N-wave runup on coast. *Applied Ocean Research , 42*, 144-154.

Lin, P., & Liu, P. (1998). A numerical study of breaking waves in the surf zone. *J. of Fluid Mechanics , 359*, 239-264.

Liu, P. F., & Losada, I. J. (2002). Wave propagation modeling in coastal engineering. *J. of Hydr. Res. , 40* (3), 229-240.

Liu, P., Cho, Y., Briggs, M., Kanoglu, U., & Synolakis, C. (1995). Run-up of solitary waves on a circular island. *J. Fluid Mech, , 302*, 259-285.

Liu, P., Lin, P., Chang, K., & Sakakiyama, T. (1999). Numerical modeling of wave interaction with porous structures. *J. Waterway, Port, Coastal and Ocean Engng , 125*, 322-330.

Lo Re, C., Musumeci, R., & Foti, E. (2012). A shoreline boundary condition for a higlhy nonlinear Boussinesq model for breaking waves. *Coastal Engng , 60*, 41-52.

Lo Re, C., Musumeci, R., Foti, E., & Ferreri, G. (2014). Random wave run-up with a physically-based-Lagrangian shoreline model. *Procedia Engng , 70*, 1046-1054.

Losada, I., Lara, J.-L., Guanche, R., & Gonzalez-Ondina, J. (2008). Numerical analysis of wave overtopping of rubble mound breakwaters. *Coastal Engng , 55* (1), 47-62.

Luettich, R., & Westerink, J. (2004). *Formulation and numerical implementation of the 2D/3D ADCIRC Finite Element Model Version 44.XX*.

Lynett, P. (2002a). A multi-layer approach to modeling nonlinear, dispersive waves from deep water to the shore. *PhD Thesis* . Ithaca, NY USA: Cornell University.

Lynett, P., & Liu, P. L. (2005). A numerical study of the run-up generated by three-dimensional landslides. *J. Geophys. Res. , 110*, C03006.

Lynett, P., Wu, T., & Liu, P. (2002b). Modeling wave runup with depth-integrated equations. *Coast. Engng. , 46*, 89-107.

Ma, G., Shi, F., & Kirby, J. T. (2012). Shock-capturing non-hydrostatic model for fully dispersive surface wave processes. *Ocan Modelling , 43-44*, 22-35.

Madsen, H., Pearson, C. P., & Rosbjerg, D. (1997b). Comparison of annual maximum series and partial duration series methods for modelling extreme hydrologic events. 2. regional modeling. *Water Resourches Research , 33* (4), 759-769.

Madsen, H., Rasmussen, P. F., & Rosbjerg, D. (1997a). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modeling. *Water Resourches Research , 33* (4), 747-757.

Madsen, O., Poon, Y., & Graber, H. (1988). Spectral wave attenuation by bottom friction: Theory. *Proc. 21th Int. Conf. Coast- Engng.* (pp. 492-504). ASCE.

Madsen, P. A., & Fuhrnam, D. R. (2010). High-order Boussinesq-type modelling of nonlinear wave phenomena in deep and shallow water. In M. Qingwei (Ed.), *Advances in numerical simulation of nonlinear water waves* (Vol. 11, pp. 245-285). Singapore: World Scientific.

Madsen, P. A., & Schäffer, H. A. (1998). Higher-order Boussinesq-type equations for surface gravity waves: derivation and analysis. *Phil. Trans. R. Soc. Lond. A , 356*, 3123-3184.

Madsen, P. A., & Sørensen, O. R. (1992). A new form of the Boussinesq equations with improved linear dispersion characteristics. Part 2. A slowly-varying bathymetry. *Coastal Engineering , 18* (3-4), 183-204.

Madsen, P. A., Bingham, H. B., & Liu, H. (2002). A new Boussinesq method for fully nonlinear waves from shallow to deep water. *J. Fluid Mech. , 462*, 1-30.

Madsen, P. A., Bingham, H. B., & Schäffer, H. A. (2003). Boussinesq-type formulations for fully nonlinear and extremely dispersive water waves: derivation and analysis. *Proc. R. Soc. Lond. A , 459*, 1075-1104.

Madsen, P. A., Fuhrman, R., & Wang, B. (2006). A Boussinesq-type method for fully nonlinear waves interacting with a rapidly varying bathymetry. *Coastal Engineering , 53*, 487-504.

Madsen, P. A., Murray, R., & Sørensen, O. R. (1991). A new form of the Boussinesq equations with improved linear dispersion characteristics. *Coastal Engineering , 15*, 371-388.

Madsen, P., & Schäffer, H. A. (1999). A review of Boussinesq-type equations for surface gravity waves. In *Advances in coastal and ocean engineering* (Vol. 5, pp. 1-94). Singapore: World Scientific.

Madsen, P., & Sørensen, O. (1993). Bound waves and triad interactions in shallow water. *Ocean Engng. , 20*, 359-388.

Madsen, P., Sørensen, O., & Schäffer, H. (1997c). Surf zone dynamics simulated by a Boussinesq type model. Part I. Model description and cross-shore motion of regular waves. *Coastal Engineering , 32* (4), 255-287.

Madsen, P., Sørensen, O., & Schäffer, H. (1997d). Surf zone dynamics simulated by a Boussinesq type model. Part II: surf beat and swash oscillations for wave groups and irregular waves. *Coastal Engineering , 32* (4), 289-319.

Makkonen, L. (2006). Plotting position in extreme value analysis. *45*, 334-340.

Makkonen, L. (2005). Plotting positions in extreme value analysis. *Journal of Applied Meteorology and Climatology , 45*, 334-340.

Mase, H. &. (1984). Run-up of random waves on gentle slopes. *Proc. 19th Coastal Engrg. Conf.* (pp. 593-609). Houston: ASCE.

Mase, H. (1989). Random wave runup height on gentle slope. *Journal of Waterway, Port, Coastal and Ocean Engineering , 115* (5), 649-661.

Mathiesen, M. (1994). Estimation of wave height duration statistics. *Coastal Engineering , 23* (1-2), 167-181.

Mazas, F., & Hamm, L. (2011). A multi-distribution approach to POT methods for determining extreme wave heights. *Coastal Engineering , 58*, 385-394.

Mazas, F., Garat, P., & Hamm, L. (2014). Questioning MLE for the estimation of environmental extreme distributions. *Ocean Engng , 92*, 44-54.

McCabe, M. V., Stansby, P., & Apsley, D. (2011). Coupled wave action and shallow water modelling for random wave runup on a slope. *J. Hydr. Research , 49*, 515-522.

McCabe, M., & Stansby, P. (2010). An investigation on the coupling of spectral energy modelling and Boussinesq type modelling in the nearshore. *Proc. 1st European IAHR Congress.* Edinburgh.

McCabe, M., & Stansby, P. (2010). An investigation on the coupling of spectral energy modelling and Boussinesq-type modelling in the nearshore. *Proc. 1st European IAHR Congress.* Edinburgh.

Mei, C. C., Stiassnie, M., & Yue, D. K.-P. (2005). *Theory and applications of ocean surface waves Part1: linear aspects. Part 2: nonlinear aspects* (Advanced series on ocean engineering ed.). Singapore: World scientific.

Mentaschi, L. B. (2013). Developing and validating a forecast/hindcast system for the Mediterranean Sea. *Journal of Coastal Research , Special Issue 65* (International Coastal Symposium Volume 2), 1551-1556.

Miles, J. (1957). On the generation of surface waves by shear flows. *J. Fluid Mech, , 3*, 185-204.

Miles, M. (1989). A note on directional random wave synthesis by the single-summation method. *Proc. 23rd IAHR Congress* (pp. C-243-C-250). Ottawa: The National Research Council Canada.

Miles, M., & Funke, E. (1989). A comparison of methods for synthesis of directional seas. *Journal of Offshore Mechanics and Arctic Engineering , 1110* (1), 43-48.

Monaghan, J. (1992). Smoothed particle hydrodynamics. *Annual Review of Astronomy and Astrophysics , 30*, 543-574.

Musumeci, R. E., Svendsen, I. A., & Veeramony, J. (2005). The flow in the surf zone: a fully nonlinear Boussinesq-type approach. *Coastal Engineering , 52*, 565-598.

Nwogu, O. (1993). Alternative form of Boussinesq Equation for nearshore wave propagation. *J. Waterway, Port, Coastal, Ocean Eng. , 119* (6), 618-638.

Orszaghova, J., Taylor, P., Borthwick, A., & Raby, A. (2012). Importance of second-order wave generation for focused wave-group run-up and overtopping. *Coast. Engng. , 94*, 63-79.

Orszaghova, O., Borthwick, A., & Taylor, P. (2014). From the paddle to the beach: a Boussinesq shallow water numerical wave tank based on Madsen and Sorensen's equations. *J. Comput. Phys. ,* 328-344.

Osher, A., & Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: algorithms based on Hamiltonian-Jacobi formulations. *J. Comput. Phys. , 79*, 12-49.

Oumeraci, H. (2004). Suistainable coastal flood defences: scientific and modelling challenges towards an integrated risk-based design concept. *Proc. First IMA International Conference on Flood Risk Assessment* (pp. 9-24). Bath (UK): IMA - Institute of Mathematics and its Application.

Park, H., & Cox, D. (2015). Empirical wave run-up formula for wave, storm surge and berm width. *Coastal Engng* .

Peregrine, D. H. (1967). Long waves on a beach. *Journal of Fluid Mechanics , 27* (4), 815-827.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics , 3* (1), 119-131.

Pisocopia, R., Panizzo, A., & Noli, A. (2000). Sul calcolo dello spettro direzionale di moto ondoso. *IDRA 2000 - XXVII Convegno di Idraulica e Costruzioni Idrauliche.* Genova.

Puleo, J. A., Holland, T. K., Slinn, D. N., Smith, E., & Webb, B. M. (2002). Numerical modelling of swash zone hydrodynamics. *Proc. ICCE 2002.* ASCE.

Raubenheimer, B. (2002). Observations and predictions of fluid velocities in the surf and swash zones. *J. Geophys. Res. , 107* (C11), 11-1-11-7.

Raubenheimer, B., Guza, R., & Elgar, S. (1996). Wave transformation across the inner surf zone. *J. Geophys. Res. , 101* (C10), 25589-25597.

Repetto, M., & Solari, G. (2011). *Analisi statistica dei dati storici di vento registrati dalle stazioni anemometriche.*

Rijnsdorp, D. P., Smit, P. B., & Zijlema, M. (2014). Non-hydrostatic modelling of infragravity waves under laboratory conditions. *Coastal Engng , 85*, 30-42.

Rijnsdorp, D., Ruessink, G., & Zijlema, M. (2015). Infragravity-wave dynamics in a barred coastal region, a numerical study. *J. of Geophysical Research: Oceans , 1202*, 4068-4089.

Roeber, V., Cheung, K. F., & Kobayashi, M. H. (2010). Shock-capturing Boussinesq-type model for nearshore wave processes. *Coastal Engineering , 57*, 407-423.

Roelvink, D., & al., e. (2009). *XBeach model description and manual.* Delft: Unesco-IHE Institute for Water Edcation Deltares and Delft niversity of Technology.

Rogers, W., Kaihatu, J., Petit, H., Booij, N., & Holthuijsen, L. (2002). Diffusion reduction in a arbitrary scale third generation wind wave model. *Ocean Engng. , 29*, 1357-1390.

Salmon, J., Holthuijsen, L., Sijlema, M., Van Vledder, G., & Pietrzak, J. (2015). Scaling depth-induced wave-breaking in two dimensional spectral wave models. *Ocean Modelling , 87*, 30-47.

Salmon, J., Holthuijsen, L., Smit, P., Van Vledder, G., & Zijlema, M. (2014). Alternative source terms for SWAN in the coastal region. *Proc. 34th Int. Conf. Coastal Engng.* Seoul, Korea.

Sartini, L. M. (2015). Comparing different extreme wave analysis models for wave climate assessment along the Italian coast. *Coastal Engineering , 100*, 37-47.

Saville, T. J. (1955). *Laboratory data in wave run-up and overtopping on 21 shore structures.* Dayton, Ohio: U.S. Army, Beach Erosion Board, Document Service Center.

Scarrot, C., & MacDonald, A. (2012). A review of extreme value treshold estimation and uncertainty quantification. *Statistical Journal , 10* (1), 33-60.

Schäffer, H. A., Madsen, P. A., & Deigaard, R. (1993). A Boussinesq model for waves breaking in shallow water. *Coastal Engineering , 20* (3-4), 185-202.

Schäffer, H., & Madsen, P. (1995). Further enhancements of Boussinesq-type equations. *Coast. Engng. , 26*, 1-15.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics , 6* (2), 461-464.

Shao, S. (2010). Incompressible SPH flow model for wave interactions with porous media. *Coastal Engng , 57* (3), 304-316.

Shao, S., Ji, C., Graham, D., Reeve, D., James, P., & Chadwick, A. (2006). Simulation of wave overtopping by an incompressible SPH model. *Coastal Engng , 53*, 723-735.

Shi, F., Kirby, J. T., Harris, J. C., Geiman, J. D., & Grilli, S. T. (2012b). A high-order adaptive time-stepping TVD solver for Boussinesq modeling of breaking waves and coastal inundation. *Ocean modelling , 43-44*, 36-51.

Shi, F., Ma, G., Kirby, J. T., & Hsu, T. (2012a). Application of a TVD solver in a suite of coastal engineering models. ICCE.

Shi, F., Tehranirad, B., & Kirby, T. (2013). *FUNWAVE-TVD - Fully nonlinear Boussinesq wave model with TVD solver, documentation and user's manual (version 2.1).* Center for applied coastal research - University of Delaware.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Mon. Weather Rev. , 91* (3), 99-164.

Smit, P., Zijlema, M., & Stelling, G. (2013). Depth-induced wave breaking in a non-hydrostatic, near-shore wave model. *Coastal Engineering , 76*, 1-16.

Smith, R., & Weissman, I. (1994). Estimating the extremal index. *Journal of the Royal Statistical Society. Series B (Methodological) , 56* (3), 515-528.

Solari, S., & A., L. M. (2012). A unified statistical model for hydrological variables inclufing the selection of threshold for the peak over threshold method. *Water Resources Research , 48*, W10541.

Soldini, L., Antuono, M., & Brocchini, M. (2013). Numerical modeling of the influence of the beach profile on wave run-up. *J. of Wtrwy, Port, Coastal and Ocean Engng , 139* (1).

Sorensen, O., Kofoed-Hansen, H., Rugbjerg, M., & Sorensen, L. (2004). A third-generation spectral wave model using an unstructured finite volume technique. *Proc. 29th Int. Conf. on Coastal Engng.*

Sørensen, O., Kofoed-Hansen, H., Rugbjerg, M., & Sørensen, L. (1998). A third-generation spectral wave model using an untructured finite volume technique. *Proc. 29th Int. Conf. Coastal Engng.* (pp. 894-906). ASCE.

Stelling, G. S., & Duinmeijer, S. P. (2003). A staggered conservative scheme for every Froude number in rapidly varied shallow water flows. *Int. J. Numer. Meth. Fluids , 43*, 1329-1354.

Stelling, G., & J.J., L. (1992). Approximation of convective processes by ciclic AOI methods. *Proc. 2nd Intern. Conf. on Estuarine and Coast. Modelling* (pp. 771-782). Tampa, Florida: ASCE.

Stelling, G., & Zijlema, M. (2003). An accurate and efficient finite-difference algorithm for non-hydrostatic free-surface flow with application to wave propagation. *Int. J. Numer. Meth. Fluids , 43*, 1-23.

Svendsen, I. A. (2007). *Introduction to nearshore hydrodynamics.* University of Delaware, USA: World Scientific.

Svendsen, I., & Putrevu, U. (1996). *Surf-zone modeling.* Center for Applied Coastal Research: Research Report No. CACR-96-08.

SWAN Team. (2015). *SWAN - Scientific and technical documentation.* Delft University of Technology.

Tehranirad, B., Shi, F., Kirby, J., arris, J., & Grilli, S. (2011). *Tsunami benchmark results for fully nonlinear Boussinesq wave model FUNWAVE-TVD, version 1.0.* Research report N.CACR-11-02.

Thacker, W. (1981). Some exact solutions to the nonlinear shallow-water-wave equations. *J. Fluid. Mech , 107*, 499-508.

The SWASH Team. (2016). *SWASH user manual.* Delft University of Technology.

The WISE Group, A. V. (2007). Wave modelling - The state of the art. *Progress in Oceanography , 75*, 603-674.

Thompson, P., Cai, Y., Reeve, D., & Stander, J. (2009). Automated threshold selection methods for extreme wave analysis. *Coastal Engineering , 56*, 1013-1021.

Thornton, E., & Guza, R. (1983). Transformation of wave height distribution. *J. Geophys. Res. , 88* (C10), 5925-5938.

Titov, V., & Synolakis, C. (1995). Modeling of breaking and nonbreaking long-wave evolution and runup using vtcs-2. *J. Wtrwy, Port, Coastal and Ocean Engng , 121* (6).

Tolman, H. L. (2009). *User manual and system documentation of WAVEWATCH III version 3.14.* NOOA/NWS/NCEP/MMAB Technical Note 276.

Tomás, A., Méndez, F. J., Medina, R., Jaime, F. F., Higuera, P., Lara, J. L., et al. (2016). A methodology to estimate wave-induced coastal flooding hazard maps in Spain. *J. Flood Risk Manga , 9* (3), 289-305.

Tonelli, M., & Petti, M. (2010). Finite volume Scheme for the solution of 2D extended Boussinesq equations in the surf zone. *Ocean Engineering , 37*, 567-582.

Tonelli, M., & Petti, M. (2009). Hybrid finite volume – finite difference scheme for 2DH improved Boussinesq equations. *Coastal Engineering , 56*, 609-620.

Tonelli, M., & Petti, M. (2013). Numerical simulation of wave overtopping at coastal dikes and low-crested by means of a schock-capturing Boussinesq model. *Coastal Engineering , 79*, 75-88.

Tonelli, M., & Petti, M. (2012). Shock-capturing Bousinesq model for irregular wave propagation. *Coastal Engineering , 61*, 8-19.

Toro, E. F. (2001). *Shock-capturing methods for free-surface shallow flows* (1st ed.). New York, NY: John Wiley & Sons, LTD.

Tsung, W., Hsiao, S., & Lin, T. (2012). Numerical simulation of solitary wave run-up and overtopping using Boussinesq-type model. *J. of Hydrodynamics , 24* (6), 299-913.

Van der Westhuysen, A., Zijlema, M., & Battjes, J. (2007). Nonlinear saturation-based whitecapping dissipation in SWAN for deep and shallow water. *Coast. Engng. , 54*, 151-170.

Van Gent, M. R., Tonjes, P., Petit, H. A., & Van den Bosh, P. (1994). Wave action on and in permeable structures. *Proc. of 24th Int. Conf. on Coastal Engineering* (pp. 1739-1753). ASCE.

Van Vledder, G. P., & Bottema, M. (2003). Improved modelling of nonlinear four-wave interactions in hallow water. *Proc. 28th ICCE* (pp. 459-471). ASCE.

Veeramony, J., & Svendsen, I. (2000). The flow in surf-zone waves. *Coastal Engineering , 39* (2-4), 93-122.

Ventroni, M. (2016a). Confronto fra le distribuzioni di Weibull e generalizzata di Pareto nell'analisi delle altezze d'onda estreme. *Studi Costieri , 23*, 91-92.

Ventroni, M., Balzano, A., & Zijlema, M. (2016c). Online coupling of spectral and non-hydrostatic models for wave simulation from offshore to nearshore. *XXXV Convegno di Idraulica e Costruzioni Idrauliche* (pp. 313-316). Bologna: http://amsacta.unibo.it/id/eprint/5400.

Ventroni, M., Balzano, A., & Zijlema, M. (2016b). Online coupling of SWAN ans SWASH for nearshore applications. *4th European Congress of IAHR,* (pp. 323-329). Liege.

Verboom, G. K., & Segal, A. (1986). Weakly-reflective boundary conditions for shallow water equations. *25th Meeting Dutch Working group on Numerical Flow Simulations.* Delft, The Netherlands: Delft University of Technology.

Verboom, G. K., & Slob, A. (1984). Weakly-reflective boundary conditions for two-dimensional water flow problems. *Advances in Water Resources , 7* (4), 192-197.

Vreugdenhil, C. B. (1994). *Numerical methods for shallow water flow.* Springer Science+Business Media.

WAFO-group. (2011). WAFO - A MATLAB toolbox for analysis of random waves and loads - A tutorial. Lund, Sweden: Math. Stat., Center for Math. Sci., Lund Univ.

Wang, B., & Liu, H. (2006). Solving a fully nonlinear highly dispersive Boussinesq model with mesh-less least square based finite difference method. *Intern. J. for Numerical Methods in Fluids , 52*, 213-235.

Wang, B., Zhu, Y., Song, Z., & Liu, A. (2006). Boussinesq-type modeling in surf-zone using mesh-less least square based finite difference method. *Conference of Global Chinese Scholars on Hydrodynaimcs* .

Warner, J., Sherwood, C., Signell, R., Harris, C. K., & Arango, H. (2008). Development of a three-dimensional, regional, coupled wave, current, and sediment transport model. *Computer Geosciences , 36*, 1284-1306.

Wei, G., Kirby, J. T., Grilli, S. T., & Subramanya, R. (1995b). A fully nonlinear Boussinesq model for surface waves. Part 1. Highly nonlinear unsteady waves. *J. Fluid Mech. , 294*, 77-92.

Weibull, W. (1951). A statistical distribution function of wide applicability. *ASME Journal of Applied Mechanics* , 293-297.

Weibull, W. (1939). A statistical theory of the strength of materials. *Proc. R. Swedish Inst. Eng. Res. , 151* (1), -.

Witting, J. M. (1984). A unified model for the evolution of nonlinear water waves. *J. Comp. Phys. , 56*, 203-236.

WMO. (2009). *Climate Data and Monitoring - WCDMP-No. 72.* Geneva, Switzerland: WMO.

Wong, T. &. (2006). A note on the estimation of extreme value distribution using maximum product of spacing. *IMS Lecture Notes-Monograph Series , 52*, 272-283.

Zelt, J. A. (1991). The runup of nonbreaking and breaking solitary waves. *Coast. Engng , 15*, 205-246.

Zhang, Y., Kennedy, A., Donahue, A., Westerink, J., Panda, N., & Dawson, C. (2014). Rotational surf zone modeling for O(μ4) Boussinesq–Green–Naghdi systems. *Oean Modelling , 79*, 43-53.

Zhang, Y., Kennedy, A., Panda, N., Dawson, C., & Westerink, J. (2013). Boussinesq-Green-Naghdi rotational water wave theory. *Coast. Engng , 73*, 13-27.

Zhang, Y., Kennedy, A., Tomiczek, T., Donahue, A., & Westerink, J. (2016). Validation of Boussinesq-Green-Naghdi modeling for surf zone hydrodynamics. *Ocean Engng , 111*, 299-309.

Zijlema, M. (2010). Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coastal Engng , 57*, 267-277.

Zijlema, M. (2009). Parallel, unstructured mesh implementation for SWAN. *Proc. 31th Int. Conf. Coastal Engng.* (pp. 470-482). ASCE.

Zijlema, M., & Stelling, G. (2008). Efficient computation of surf zone waves using the nonlinea shallow water equations with non-hydrostatic pressure. *Coastal Engineering , 55*, 780-790.

Zijlema, M., & Stelling, G. (2005). Further experiences with computing non-hydrostatic free-surface flows involving water waves. *International Journal for Numerical Methods in Fluids , 48*, 169-197.

Zijlema, M., & Van der Westhuysen, A. (2005). On convergence behaviour and numerical accuracy in stationary SWAN simulations of nearshore wind wave spectra. *Coastal Engng. , 52*, 237-256.

Zijlema, M., Stelling, G., & Smit, P. (2011a). Simulating nearshore wave transformation with non-hydrostatic wave-flow modelling. *Proceed. 12th Int. Workshop on Wave hindcasting and Forecasting.* Hawaii'i, USA.

Zijlema, M., Stelling, G., & Smit, P. (2011b). SWASH: an operational public domain code for simulating wave fields and rapidly varied flows in coastal waters. *Coastal Engineering , 58*, 992-1012.

Zijlema, M., Van Vledder, G., & Holthuijsen, L. (2012). Bottom friction and wind drag for wave models. *Coast. Engng. , 65*, 19-26.

Zou, Z. (2000). A new form of higher order Boussinesq equations. *Ocean Engng. , 27*, 557-575.