Final Draft
of the original manuscript

Benetazzo, A.; Barbariol, F.; Pezzutto, P.; Staneva, J.; Behrens, A.; Davison, S.; Bergamasco, F.; Sclavo, M.; Cavalieri, L:
Towards a unified framework for extreme sea waves from spectral models: rationale and applications.

First published online by Elsevier: 08.12.2020

https://dx.doi.org/10.1016/j.oceaneng.2020.108263
Towards a unified framework for extreme sea waves from spectral models: rationale and applications

Alvise Benetazzo1*, Francesco Barbariol1, Paolo Pezzutto1, Joanna Staneva2, Arno Behrens2, Silvio Davison1, Filippo Bergamasco3, Mauro Sclavo4, Luigi Cavaleri1

1 Istituto di Scienze Marine (ISMAR), Consiglio Nazionale delle Ricerche (CNR), Venice, Italy.
2 Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research, Geesthacht, Germany
3 DAIS – Università Ca’ Foscari, Venice, Italy.
4 Istituto di Scienze Polari (ISP), Consiglio Nazionale delle Ricerche (CNR), Venice, Italy.

*Corresponding author address: Istituto di Scienze Marine (ISMAR) – Consiglio Nazionale delle Ricerche (CNR), Arsenale Tesa 104, Castello 2737/f, 30122, Venice (Italy); e-mail: alvise.benetazzo@cnr.it

Abstract

Reliable predictions of oceanic waves during storms have always been foremost for offshore design and operation, coastal hazards, and navigation safety. Indeed, many accidents that occurred during storms were ascribed to the impact with unforeseen large waves. In this context, the purpose of this study is to improve the present state extreme wave estimate from spectral wave models. We describe an implementation for the WAM model, and we investigate the use of WAM and WAVEWATCH III fed with common routines designed to evaluate the short-term/range maximum wave statistics. An extensive assessment of models’ results in the Adriatic and North Sea is performed using time and space-time wave measurements, and through an intercomparison between WAM and WAVEWATCH III applied with three different input/dissipation source term parametrizations (ST3/4/6). Further, models’ capabilities are investigated, and extreme waves characterized, in the Mediterranean Sea, aiming also at disentangling the wave spectrum bulk parameters that may point to favourable conditions for the generation of high waves. Based on the comparisons between model results and measurements, we conclude that for the model characterization of extremes, the accuracy of the significant wave height is pivotal; differences between models of other spectral parameters seem to have a minor effect.

Keywords: Sea wave extremes; Maximum crest and wave height; WAM; WAVEWATCH III; Stereo imaging; Mediterranean Sea.
1 Introduction

The characterization of extreme waves during marine storms has been an active topic of research for decades because of its importance for shipping, offshore design and operations, and coastal hazards. Theoretical progress has been fast, but it has not been accompanied by an improvement of numerical predictions and of field measurement strategies. Significant efforts are being undertaken to better understand the occurrence and magnitude of extreme waves, including rogue waves (Benetazzo et al., 2017a; Cavaleri et al., 2016, 2012; Dematteis et al., 2019; Donelan and Magnusson, 2017; Dysthe et al., 2008; Fedele et al., 2017; Gemmrich and Garrett, 2011; Janssen, 2003; Onorato et al., 2001, 2013; Slunyaev et al., 2005; Toffoli et al., 2005; Waseda et al., 2011), but modeling strategies resulted ineffective in warning seafarers or avoiding structural damage to offshore facilities (see e.g. Bitner-Gregersen and Gramstad, 2015; Fedele et al., 2017). Although there are different theoretical approaches to tackle the extreme and rogue wave problem (Kharif et al., 2009), what mainly limited their high-resolution characterization at large spatial and temporal scales is the lack of widespread modeling tools and standardized observational data.

From a wave model point of view, the extreme value problem can in principle be addressed by numerically solving in detail the complete fluid mechanics equations describing the sea surface elevation field (Mei, 1983). This direct approach has a high computational cost and requires accurate initial conditions that are generally not easy to achieve. It is generally limited to simulations of small space and time scales (a hundred wavelengths or periods), typical, for instance, of coastal surf zones or laboratory conditions. Still, recent efforts using High Order Spectral simulations of the Euler equations for directional sea states (Dommermuth and Yue, 1987) proved to be a powerful tool for investigating the oceanic rogue wave behavior, focusing, however, on individual rogue wave events (e.g. Fedele et al., 2016). In this paper, this deterministic approach and any model examination of wave-by-wave heights are set aside, and we look instead at the statistical properties of the wave field. These are customarily described by the directional wave energy spectrum, which is the variable integrated by the phase-averaged wave models solving the wave energy balance equation (Gelci et al., 1957). This approach is based on a spectral decomposition of the sea surface elevation variance across frequencies or wavenumbers and directions, while wave phases are not resolved.

Phase-averaged spectral wave models are the primary objects of our study. The focus is on WAM (Janssen, 2008; Komen et al., 1994) and WAVEWATCH III® (hereinafter WW3; Tolman. H.L. and the WAVEWATCH III
Development Group, 2014), which are the two most widely used open-source community spectral wave models applied to forecast and hindcast studies. Spectral wave models provide an essential part of marine weather analysis and made significant improvements in simulating wave spectrum bulk parameters, such as the significant wave height, the peak period and the zero crossing mean wave period (see e.g. http://www.ecmwf.int/en/forecasts/charts/). This means that model parameterizations describe the air-sea interaction processes reasonably well. Little information, on the other hand, is known about their performance with focus on extremes, namely the maximum wave and crest heights.

In this respect, the motivation for the present model development arises from the extreme wave analysis that was implemented by Janssen and Bidlot (2009) in the European Centre for Medium-Range Weather Forecasts (ECMWF) version of the WAM model (called ECWAM). In ECWAM, the distribution of extremes is based on the Mori and Janssen (2006) formulation for the wave envelope height and is used to estimate the expected value of the maximum envelope height over a three hour duration. There are no other operational forecasting centers that release maximum wave products on a regional or oceanic scale. However, information on the maximum wave heights are required by users and the offshore industry for the definition of the environmental loads over the lifetime of a ship or structure (DNV GL, 2017), but the common strategy used so far has been based on the off-line processing of model spectra to derive meaningful parameters at certain sea locations. To define extreme wave features and their spatial-temporal variability, WW3 was equipped with on-line routines for maximum waves (Barbariol et al., 2017), and the Simulating WAves Nearshore (SWAN) spectral model (Booij et al., 1999) was developed with the same purpose (Barbariol et al., 2015; Sclavo et al., 2015).

In this study, an implementation for extreme wave computation has been developed for the calculation in WAM (cycle 4.5.4 or later versions; https://github.com/mywave/WAM) of the maximum crest height (i.e. the maximum surface elevation) and the maximum crest-to-trough wave height (i.e. the maximum vertical distance between a crest and the closest trough) for each sea state. These two variables complement the envelope height released by ECMWF, and, at the same time, are consistent with the standard outputs of wave analysis, e.g. obtained via zero-crossing (Holthuijsen, 2007). This study concentrates on documenting the code modification and its optimization for operational purposes. The goal is to upgrade WAM and to endow it with the same extreme wave outputs as WW3. This allows the assessment of their relative performance with focus on extremes and documenting how well they compare against observations. To this end, sea surface elevation records from two
wave buoys and one wave probe are used in this study to establish differences between models. These temporal records are complemented with stereo-video data that are used to analyze maxima of space-time wave trains. We also discuss the way field observations ought to be managed so as to provide meaningful extreme wave information that may prove useful for model validation. Further, we point out that since extreme wave statistics depend on the higher-order spectral moments, the assessment of maximum waves is an indirect way to test how well models reproduce the shape of the directional wave spectrum.

The combined use of WAM and WW3 is used in this study to provide for the first time a characterization of the magnitude and spatial variability of extreme waves one can encounter in the Mediterranean Sea. The atmospheric and wave pattern in this basin is rather complex with strong offshore winds, such as the Mistral in the western Mediterranean region, the Etesian (or Meltem) in the Aegean Sea, and the Bora and Sirocco in the Adriatic Sea. In the innermost areas, the wave growth is limited by the available fetch. The Mediterranean Sea is also exposed to site-specific storm events called Medicanes (Mediterranean Hurricanes) that are tropical-like cyclones with winds well exceeding 100 km/h. Additionally, this basin is among the world's busiest waterways, hence information about the intensity and distribution of maximum wave heights is crucial. There are many reports on very large waves hitting ships, for instance the Louis Majesty accident that caused two fatalities (Cavaleri et al., 2012) while en route from Barcelona to Genoa, or the Jean Nicoli ferry accident (in 2017) that en route from Ajaccio to Marseille was hit by what was considered a rogue wave. So far, however, the wave climate characterization has been limited to the extremes of the significant wave height (Sartini et al., 2017), while it lacks any information on the maximum waves, which are instead considered in this study.

This manuscript will proceed as follows. Section 2 is dedicated to explaining the theoretical formulations implemented in WAM for the estimate of maximum crest and wave heights. The theories are presented for temporal (short-term) extremes and spatio-temporal (short-term/range) extremes. In the same section, the observations of time extreme and space-time extreme waves and the WAM and WW3 setups in the Adriatic, Mediterranean, and North Sea are described. Section 3 presents details of the WAM implementation and assessment of model outputs against observations (collected in the Adriatic and North Seas). A model analysis of extreme waves in the Mediterranean Sea completes the section. A discussion and summary of the main conclusions of the study are presented in Section 4.
2 Materials and Methods

2.1 Short-term/range extreme value statistics

In this section we are going to describe theories that are adopted for the estimation of extreme wave events in WAM. Technical details of this implementation are given afterwards in section 3.1, while here we focus on the theoretical aspects and present key formulae. The equivalent WW3 implementation was extensively described in a previous study (Barbariol et al., 2017), so only what it is necessary to make a straight comparison with WAM and with newly collected observations will be recalled in the present work.

The basic concept supporting the extreme value analysis is the steady-state sea condition, and the derived statistics is referred to as short-term (time interval from minutes to 1 hour or so). Later in this section we shall include the condition of homogeneity of the wave field and the joined short-term/range derived statistics. In this context, it is therefore necessary to motivate the two approaches we have adopted for extreme wave evaluation. Indeed, till recently, researchers have concentrated on the probability of extreme events from a given sea surface elevation $\eta(t)$ time series (i.e. point statistics in the temporal domain $t$) at a fixed point $x_0 = (x_0, y_0)$ in space, being the two Cartesian horizontal $x$ and $y$ axes lying on the still water surface. In this context, starting from the Rayleigh distribution, valid for linear wave trains with a narrow spectrum (Longuet-Higgins, 1952), it has been established that nonlinear effects, like second-order bound waves and four-wave interactions, may have a profound impact on the statistics of extreme events (Fedele and Tayfun, 2009; Janssen, 2003; Mori and Janssen, 2006; Waseda et al., 2011). Also including nonlinearities, however, in the point statistics very high and even rogue waves remain very unlikely (Christou and Ewans, 2014; Gemmrich and Thomson, 2017; Magnusson et al., 1999).

A development of these models is due to advances started by Adler (1981) and Piterbarg (1996) who studied the extreme value problem in multi-dimensional manifolds and introduced the concept of excursion probability. For oceanic waves crossing a 2D sea surface region, being the surface elevation $\eta(x, y, t)$ the variable of interest, the problem was reduced to three dimensions (Fedele, 2012; Fedele et al., 2013; Krogstad et al., 2004), two in space ($x$ and $y$) and one in time ($t$). In the 3D spatio-temporal domain ($x, y, t$), the number of extreme events is larger than what could be expected in the 1D temporal domain ($t$), regardless the average number of waves (Benetazzo et al., 2015). This achievement was used to explain the high likelihood of rogue waves when extreme events are studied in a 3D spatio-temporal region (Benetazzo et al., 2017a), so that the very notion of rogue wave in stormy seas was questioned (Cavaleri et al., 2016). For consistency, in the degenerated case of null space (e.g. in...
long-crested sea states, for which the areal effects become negligible), the spatio-temporal statistics coincides with the temporal one.

Having said this, and in line with the WW3 extreme wave implementation (WW3DG, 2019), for WAM we have chosen to estimate two variables for a given sea state: the expected value (or expectation) of the random variable maximum crest height ($C_m$), and the expected value of the random variable maximum crest-to-trough wave height ($H_m$). Both variables are specified in the 1D temporal domain (hereinafter time extremes) and in the 3D spatio-temporal domain (hereinafter space-time extremes).

2.1.1 Time extreme waves

In this study, to approximate the actual distribution function of $C_m$ at a fixed point $x_0$ on the sea surface, we considered two time extreme statistical models of the wave elevation $\eta(t) = \eta(x_0, t)$ that include the effects of second-order bound nonlinear modes (Longuet-Higgins, 1963). The nonlinear effects on the geometry of ocean waves is to display higher and sharper crests, a feature predicted by the Gram-Charlier solution of Longuet-Higgins. We then rely on the dispersive focusing of second-order non-resonant harmonic waves as the leading physical mechanism for large wave formation. This mechanism proved to suffice to explain most of the real oceanic large and rogue wave behaviour (Benetazzo et al., 2015; Fedele et al., 2016), still limiting, from a numerical model point of view, the computational burden that is required to assess extremes with higher-order corrections to the linear first-order representation (Janssen, 2017). We have hence not considered the role of four-wave nonlinear interactions and any effect such as the modulation instability for narrow-band wave packets in deep water (Janssen, 2003; Mori and Janssen, 2006; Onorato et al., 2013). Still, those higher-order nonlinearities are considered, via the kurtosis of the surface elevation probability density functions (pdf), in the freak wave warning system that is included in the wave model ECWAM (ECMWF, 2019). However, granted that difference in the order of nonlinearities and the choice of defining in ECWAM the wave height as twice the envelope, the ECWAM theoretical approach for extremes and the methods used in WW3 implementation (and here adopted for WAM) proved to have similar statistical scores when compared against open-sea estimates of average values of maxima (Barbarel et al., 2019). This suggests that four-wave nonlinear interactions may have a minor effect on expected extreme waves occurring in directionally spread stormy seas, as it was also proposed in other studies (Fedele et al., 2019, 2016).
The first model we shall consider for $C_m$ is based upon the Tayfun (1980) nonlinear formulation and it will be detailed further as a particular case (i.e. null $xy$-space) of the space-time extremes (section 2.1.2). The second nonlinear model taken into consideration relies on the Forristall (2000) exceedance distribution of the crest height ($C$). For an elevation threshold $h$ and a sea state with significant wave height $H_s$, the Forristall distribution $\Pr_{FO}$ is expressed by a Weibull-like function given by:

$$\Pr_{FO}\{C > h\} = \exp \left[ -\left( \frac{h}{\alpha_F H_s} \right)^{\beta_F} \right] \quad (1)$$

where $\alpha_F H_s > 0$ is the scale parameter and $\beta_F > 0$ is the shape parameter of the distribution. The effects of second-order bound nonlinearities are included as a function of the water depth $d$ and the wave spectrum via the steepness parameter $S = 2\pi H_s (g T_1^2)^{-1}$ and the Ursell number $U = H_s (k_1^2 d^3)^{-1}$, which both contribute to the values of $\alpha_F$ and $\beta_F$, as follows:

$$\alpha_F = 0.3536 + 0.2568 S + 0.0800 U$$
$$\beta_F = 2 - 1.7912 S - 0.5302 U + 0.284 U^2 \quad (2)$$

Here, $g$ is the gravitational acceleration, $T_1$ is the mean wave period calculated from the ratio of the first two moments of the frequency spectrum, and $k_1$ is the deep-water mean wavenumber for a frequency $T_1^{-1}$. Forristall obtained $\alpha_F$ and $\beta_F$ by empirical fitting, using simulations of short-crested nonlinear waves, typical of stormy sea states.

Following Gumbel (1958), the distribution of extremes can be written down as a function of the initial distribution and of the sample size $N$. Accordingly, the probability that the “maximum crest height” $C_m$ in a sea state of duration $D$ (from minutes to few hours) and holding $N$ waves on average is more than $h$ can be expressed as follows:

$$\Pr\{C_m > h \mid D\} = 1 - (1 - \Pr_{FO})^N \quad (3)$$

The previous formula, for large values of $h$ and $N$, is fairly well approximated by a double exponential Gumbel-type distribution of crest heights. This allows computing the expected value (indicated with an overbar) of the maximum crest height at a fixed point in the following form

$$\overline{C}_m = 4\sigma \alpha_F (\ln N)^{1/\beta_F} \left( 1 + \frac{\gamma}{\beta_F \ln N} \right) \quad (4)$$
being $\gamma \approx 0.5772$ the Euler-Mascheroni constant, and $\sigma^2$ the sea surface elevation variance. The average number of waves is computed as $N = DT_z^{-1}$, where and $T_z$ is the average zero-crossing wave period. The value of $c_m$ in Eq. (4), as well as the extremal values defined below, increases with the sample size.

The time extreme model for the maximum crest-to-trough height $H_m$ is based on the Naess (1985) theory of crest-to-trough heights $H$ that generalizes the Rayleigh distribution to account for the spectral bandwidth effects on the wave heights as follows:

$$
Pr_{Na}[H > h] = \exp \left[ -\frac{1}{4(1-\psi^*)} \left( \frac{h}{\sigma} \right)^2 \right] \quad (5)
$$

where $\psi^* \in [-1, 0]$ is a bandwidth parameter (to be defined in section 3.1), which in typical wind-sea conditions assumes values in the [-0.75, -0.65] range and it is a function, according to Naess, of the dominant wave period. The inclusion of the parameter $\psi^*$ decreases the probability levels with respect to Rayleigh (for which $\psi^* = -1$ since the bandwidth of the spectral density function approaches zero), which hence represents an upper bound for the wave height probabilities. The choice of a linear model for individual wave heights was dictated by two factors: the effect on $H$ of second-order nonlinearities is small (a few percent), much smaller than that on crest heights (Tayfun and Fedele, 2007), and the requirement of limiting the computational effort needed to model higher-order wave heights. We shall verify this assumption later on in the paper by comparing model and observed extremes.

The extreme value exceedance distribution of the “maximum wave height” $H_m$ in a sea state with an average number $N$ of waves can be expressed as follows:

$$
Pr[H_m > h \mid D] = 1 - (1 - Pr_{Na})^N \quad (6)
$$

For large values of $h$ and $N$, the Gumbel-type approximation provides the expected value of the maximum wave height at a fixed point, that is:

$$
\bar{H}_m = 2\sigma \sqrt{1 - \psi^*} (\ln N)^{1/2} \left( 1 + \frac{\gamma}{2\ln N} \right) \quad (7)
$$

The same result for $\bar{H}_m$ can be obtained in a similar manner by means of the Quasi-Determinism asymptotic theory (Boccotti, 2000) that assumes a scale relationship between the shape of the highest waves and the expected value of the maximum crest height. Boccotti theory will be used in this study to characterize the value and the meaning of $\psi^*$ for a given wave model spectrum.
2.1.2 Space-time extreme waves

When the domain of interest for extremes includes a 2D sea surface region (i.e. the physical \( xy \)-space crossed by wave trains), the extreme value statistics incorporates the condition of short-range homogeneous sea state, which is the analogous condition for random variables that are ordered in space instead of time. Therefore, if the averages and the variances of the variables are constant in time and space, the process is called weakly stationary and homogeneous, and the derived extreme statistics is defined as space-time. The so-called space-time extremes are here theoretically approached within the framework of the Space-Time Quasi Determinism theory (Benetazzo et al., 2017b, 2015; Boccotti, 2000; Fedele, 2012). Accordingly, the excursion probability \( \Pr\{\cdot\} \) of second-order nonlinear maximum crest heights \( C_m \) exceeding a large elevation threshold \( h \) over a 2D sea surface region (with sides \( X \) and \( Y \) and area \( A = XY \)) and a 1D time interval (of duration \( D \)) is expressed as follows (Benetazzo et al., 2015):

\[
\Pr\{C_m > h \mid DA\} = \left[ N_3 \left( \frac{h_0^2}{\sigma^2} - 1 \right) + N_2 \frac{h_0}{\sigma} + N_1 \right] \exp \left( -\frac{h_0^2}{2\sigma^2} \right) \tag{8}
\]

where

\[
h_0 = \sigma \left( -1 + \sqrt{1 + 2\mu h \sigma^{-1}} \right) \mu^{-1} \tag{9}
\]

is the Tayfun formula that relates the linear \( (h_0) \) and the second-order nonlinear \( (h) \) crest heights via the steepness parameter \( \mu \) (see section 3.1.1 for its characterization). In this respect, Fedele (2015) confirmed that in the space-time extreme models the second-order nonlinearities cannot be neglected, while it is expected a modest contribution from the third-order nonlinearities.

The exceedance probability in Eq. (8) is the sum of three terms (Fedele, 2012): \( i \) the probability (proportional to the number of waves \( N_3 \)) that the threshold is exceeded within the space-time region \( \Gamma \) of 3D volume \( V = XYD \); \( ii \) the probability (proportional to the number of waves \( N_2 \)) that the threshold is exceeded on the 2D faces of \( \Gamma \); and \( iii \) the probability (proportional to the number of waves \( N_1 \)) that the threshold is exceeded on the 1D edges of \( \Gamma \). In analogy with time extreme models, in space and time the average numbers of 3D, 2D, and 1D waves are computed including the region sides and characteristic wavelengths and period; in particular, we have

\[
N_{3D} = \frac{N_3}{2\pi} = \frac{X Y D}{L_x L_y T_z} \sqrt{1 - \alpha_{xt}^2 - \alpha_{xy}^2 - \alpha_{yt}^2 + 2\alpha_{xt}\alpha_{xy}\alpha_{yt}}
\]

\[
N_{2D} = \frac{N_2}{\sqrt{2\pi}} = \frac{X D}{L_x T_z} \sqrt{1 - \alpha_{xt}^2} + \frac{Y D}{L_y T_z} \sqrt{1 - \alpha_{yt}^2} + \frac{XY}{L_x L_y} \sqrt{1 - \alpha_{xy}^2} \tag{10}
\]
The term \( \frac{X \, Y \, D}{L_x \, L_y \, T_z} \) in the formula for \( N_{3D} \) represents the average number of 3D waves (with mean crest length \( L_x \) and mean crest length \( L_y \)) crossing the cuboid region \( \Gamma \) assuming that the sea surface is “confused” with no organized movement of waves. In that case the covariance matrix of sea surface elevation gradients is diagonal, meaning that the space-time surface elevation gradients are linearly uncorrelated, which is consistent with them being independent (Baxevani and Rychlik, 2004). The degree of organization of the 3D wave motion is expressed by the additional term \( \sqrt{1 - \alpha_{xt}^2 - \alpha_{xy}^2 - \alpha_{yt}^2 + 2\alpha_{xt}\alpha_{xy}\alpha_{yt}} \) that becomes equal to zero in the limit of directionally narrow sea, for which then \( N_{3D} = 0 \) (see section 3.1.1 for the spectral definition of the parameters \( \alpha_{xt}, \alpha_{xy} \), and \( \alpha_{yt} \)).

For large values of the crest height, the excursion probability can be approximated with a Gumbel-type distribution, from which we determine the expected value of the maximum crest height in space and time (space-time extreme) as follows:

\[
\bar{c}_m = \left( \xi_0 + \frac{\mu \xi_0^2}{2 \sigma} \right) + \sigma \gamma \left( 1 + \mu \xi_0 \sigma^{-1} \right) \left( \xi_0 \sigma^{-1} - \frac{2N_3 \xi_0 \sigma^{-1} + N_2}{N_3 \xi_0^2 \sigma^{-2} + N_2 \xi_0 \sigma^{-1} + N_1} \right)^{-1}
\] (11)

In Eq. (11), \( \xi_0 \sigma^{-1} \) is the mode of the pdf of linear space-time extremes (Fedele et al., 2012). As a particular case, assuming \( X = 0 \) and \( Y = 0 \), the number of waves \( N_{3D} \) and \( N_{2D} \) vanish, while \( N_1 = D / T_z \), and Eq. (11) provides the prediction based on the Tayfun (1980) model, which is then incorporated in the WAM (and WW3) implementations of time extremes as particular case of space-time extremes.

Finally, maximum crest-to-trough wave heights \( H_m \) over a space-time region are estimated by means of the Quasi-Determinism theory (Boccotti, 2000), which establishes a relationship, via the surface elevation autocovariance function, between the expected maximum crest and crest-to-trough wave heights, in the context of the deterministic behaviour of the largest waves. Letting \( \mu = 0 \) in Eq. (11), and hence assessing the linear part of \( \bar{c}_m \), the space-time extreme value of \( \bar{H}_m \) is computed as follows:

\[
\bar{H}_m = \left( \xi_0 + \sigma \gamma \left( \xi_0 \sigma^{-1} - \frac{2N_3 \xi_0 \sigma^{-1} + N_2}{N_3 \xi_0^2 \sigma^{-2} + N_2 \xi_0 \sigma^{-1} + N_1} \right)^{-1} \right) \sqrt{2(1 - \psi^*)}
\] (12)

To complete this section, it is worth clarifying the meaning of the sea surface area \( A \) with orthogonal sides of length \( X \) and \( Y \) that are used for space-time extremes. We start by assuming the ergodicity of the random wave
process, such that the elevation probability density function over the 2D space give the same results as over an
ensemble of realisations. However, the space must satisfy the requirement that statistical properties (such as
moments of the distribution) are invariant under translation of coordinates. This condition of homogeneity is
preserved if the sea surface area is small enough that the wave spectrum and its moments do not change within it.
The short-range condition is satisfied by keeping both X- and Y-range within $O(10^2)$ m, plainly depending also on
the spatial variability of external forcing and other wave-related processes. For numerical models, the area $A$ is
supposed to be drawn around each grid node where the wave spectrum is firmly defined at each time step. As a
consequence, $X$ and $Y$ should be approximately smaller than the usual model grid size, through which spectral
properties are expected to change.

### 2.2 Time and space-time extreme wave observations

The assessment of the theoretical formulations aimed at predicting the short-term/range distribution of maximum
wave parameters is a challenging task. Indeed, for a given sea state, maximum heights in time and space-time wave
trains are randomly distributed and cannot be deterministically predicted, but only statistically characterised by
their pdf. Therefore, nothing can be said about the worthiness of a theoretical model for extremes just having, for
instance from a record $\eta(t)$ of given duration $\Delta$ (say, 20 minutes for buoys), only one realization of the random
variable $H_m = \max \{ H \mid t \in \Delta \}$ or $C_m = \max \{ C \mid t \in \Delta \}$. A set of independent realizations of $C_m$ and $H_m$ must be
produced so as to simulate the stochastic process and to provide an assessment of the theoretical formulations.

For fixed point oceanic wave observations, this requires that, within a time interval $\Delta$ where $\eta(t)$ can be
assumed stationary (in a statistical sense), the elevation series is split into independent wave records (say, $n$ of
equal duration $D = \Delta / n$) where maxima are gauged (Benetazzo et al., 2018). The samples of the random variables
$\{C_m\}$ and $\{H_m\}$ (with $i = 1, 2, \ldots n$) have an empirical distribution function, of which we can compute the sample
means defined as

$$\bar{C}_{m,obs} = <C_m>, \bar{H}_{m,obs} = <H_m>$$

(13)

to be then compared with the theoretical predictions $\bar{C}_m$ and $\bar{H}_m$, respectively. As a corollary, because of the
inherent randomness of the process, very little can be said about the relationship between the sea state
characteristics and the likelihood or intensity of wave maxima taking only one observation $H_{mi}$ or $C_{mi}$, unless the
pdf of maximum heights is fairly narrow. All this said, a strict requirement for model assessment is that the
observed sea surface elevation series, such as $\eta(t)$, is logged and made available for processing. This data is rarely
available, which implies that modeled extremes can only be assessed for short time periods and on a few sparse places, requesting that in general ad hoc experiments with dedicated technology must be planned (see e.g. Filipot et al., 2019).

To be compared with modeled results, in this study dedicated in-situ observations of maximum waves from one Adriatic Sea and two North Sea stations were taken into account. In the Adriatic Sea, instruments were mounted on the Acqua Alta oceanographic platform (longitude 12.51 °E; latitude 45.31 °N; 15 km off the Venice coast, Italy). Although the Adriatic Sea is on average a calm sea, the Acqua Alta platform is a favorable site for wave observations during storms, since it is exposed to a variety of directional sea states (Cavaleri, 1999), spanning short-fetched northeasterly wave conditions (generated by Bora winds), well-developed southeasterly wave conditions (generated by Sirocco winds), and, occasionally, southeasterly swells crossing northeasterly wind waves. Since our purpose is to assess both time and space-time extreme model predictions, two distinct datasets of observed waves are considered: the first one includes space-time elevation data $\eta(x, y, t)$ gathered with a stereo imaging system, while the second one contains time data $\eta(t)$ collected with a vertical-looking point-like wave radar (model VEGAPULS 42) that sampled the surface elevation at 2 Hz with accuracy of 0.05 % and 1-mm resolution. It is worth reminding that Acqua Alta was the test site for previous experiments dedicated to the observation and modeling of the largest waves during storms (Benetazzo et al., 2017a, 2015), including the test used to assess the WW3 implementation for extremes (Barbariol et al., 2017).

As mentioned before, for space-time extreme assessment we have used a stereo wave imaging system based on the technology of WASS (Wave Acquisition Stereo System; Benetazzo et al., 2012; Bergamasco et al., 2017) that allows to observe 3D wave fields while they evolve on the space-time domain. The WASS setup at Acqua Alta is similar to that adopted in previous applications (Benetazzo et al., 2012; Leckler et al., 2015) and it used a pair of 5-Mpixel digital cameras (2456 columns by 2048 rows array of 3.45 $\mu$m square active elements), mounting 5-mm distortion-less lenses, and placed 3.23 m apart and 14.5 m above the mean sea level. The cameras’ internal parameters were estimated using the Bouguet’s Matlab calibration toolbox (http://www.vision.caltech.edu/bouguetj/calib_doc/), whereas the spatial configuration of cameras (i.e. the extrinsic parameters) was computed using an auto-calibration approach based on the photometric consistency between stereo images (Bergamasco et al., 2017). With the device calibrated, each image pair acquired was stereo rectified and processed by a modified version (Bergamasco et al., 2017) of the dense stereo algorithm proposed by Hirschmüller.
(2008) available in the OpenCV library (http://opencv.org) by Bradski and Kaehler (2008). The semi-global nature of the approach has the great advantage that it can relate the photometric consistency of several matching pixels to improve the reliability of the disparity map, especially for areas with loosely distinctive features. As a consequence, we can keep a relatively small window size (13 x 13 pixels) while still obtaining a precise localization of the matches. A discussion of the errors associated with the WASS measurements is reported in the study of Benetazzo et al. (2012), from which we recall the mean absolute quantization error of the adopted WASS setup, which is in the order of 3 cm along the three orthogonal directions.

For each image pair, output of the stereo process is a cloud of 3D points of the sea surface field. For each point cloud, after being transformed to a common earth reference frame (Benetazzo, 2006), a patch-wise planar surface was constructed by means of 2D Delaunay triangulation over the xy-space. Then, the wave elevations \( \eta \) were resampled over a regular xy-grid (Figure 1) at a resolution of 0.3 m to span the space \( x \in [-65 \text{ m, 65 m}] \) and \( y \in [20 \text{ m, 100 m}] \). To limit the influence of high-frequency noise, time series \( \eta(t) \) taken at each xy-position of the space–time wave burst \( \eta(x,y,t) \) were low-pass filtered at 1.2 Hz, thus removing from the field the spectral components shorter than about 1 m. For each burst, the standard deviation \( \sigma \) has been estimated from the second-order moments of the surface displacement \( \eta(x, y, t) \) probability density function as follows:

\[
\sigma = \sqrt{ \langle \eta(x, y, t)^2 \rangle - \langle \eta(x, y, t) \rangle^2 } \tag{14}
\]

and, by definition, the observed significant wave height was taken as \( H_s = 4\sigma \).
Figure 1 – Example of instantaneous 3D wave field $\eta(x, y)$ observed with WASS from the Acqua Alta platform (Adriatic Sea) during a Bora-driven sea condition. The direction of the horizontal coordinated $x$- and $y$-axis is sketched on the left. The positive $y$-axis points to 46 °N. The coloured surface overlaid the grey-scale image is mapped proportionally to the sea surface elevation (reddish for wave crests and bluish for wave troughs).

<table>
<thead>
<tr>
<th>Date (UTC)</th>
<th>$H_s$ (m)</th>
<th>$T_p$ (s)</th>
<th>$U_{10}$ (m/s), $Dir$</th>
<th>$&lt; C_{mi} &gt; / H_s$</th>
<th>$&lt; H_{mi} &gt; / H_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018/03/21, 08:11</td>
<td>2.14</td>
<td>6.83</td>
<td>15, NE (Bora)</td>
<td>1.07</td>
<td>1.76</td>
</tr>
<tr>
<td>2018/03/22, 13:37</td>
<td>1.25</td>
<td>5.53</td>
<td>8, NE (Bora)</td>
<td>1.26</td>
<td>1.97</td>
</tr>
<tr>
<td>2018/04/05, 09:44</td>
<td>0.92</td>
<td>6.02</td>
<td>3, SE (swell)</td>
<td>1.02</td>
<td>1.73</td>
</tr>
<tr>
<td>2018/04/09, 13:07</td>
<td>1.11</td>
<td>4.36</td>
<td>11, SE (Sirocco)</td>
<td>1.30</td>
<td>1.99</td>
</tr>
<tr>
<td>2018/09/24, 09:57</td>
<td>1.81</td>
<td>5.02</td>
<td>19, NE (Bora)</td>
<td>1.24</td>
<td>1.93</td>
</tr>
<tr>
<td>2018/09/24, 10:56</td>
<td>1.82</td>
<td>5.53</td>
<td>17, NE (Bora)</td>
<td>1.20</td>
<td>1.89</td>
</tr>
<tr>
<td>2018/09/24, 11:44</td>
<td>1.98</td>
<td>6.10</td>
<td>16, NE (Bora)</td>
<td>1.18</td>
<td>1.86</td>
</tr>
<tr>
<td>2018/10/29, 13:00</td>
<td>2.75</td>
<td>7.88</td>
<td>17, SE (Sirocco)</td>
<td>1.01</td>
<td>1.65</td>
</tr>
</tbody>
</table>

Table 1 – Space-time wave bursts collected from the Acqua Alta oceanographic research platform using the stereo wave imaging system.

Starting date and hour (UTC), wave parameters (significant wave height $H_s$ and peak period $T_p$) and mean wind conditions (10-m height speed $U_{10}$ and direction $Dir$). The expected value of the observed maximum crest $C_{mi}$ and wave heights $H_{mi}$ are gathered within a sea surface 3D region $\Gamma$ with sides $X = 30$ m and $Y = 30$ m, and a time interval spanning $D = 180$ s.

For the investigation of the model capability to estimate space-time maximum wave parameters, we have focused on metocean conditions that include wind-sea (dominant NE Bora and SE Sirocco winds) and one case of SE swell, for a total of eight bursts that were collected during the stormy periods March-April and September-October 2018 (Table 1). We have then considered the maximum waves within the eight space-time elevation records $\eta(x, y, t)$. To this end, we have followed the same approach as adopted by Benetazzo et al. (2015) and Benetazzo et al. (2018b), and we have split the entire time interval $\Delta$ in adjacent and non-overlapping space-time subrecords. Here, we choose space-time regions $\Gamma_i$ of equal square area $A = XY = 30$ m x 30 m and duration $D = \ldots$
180 s, in order to yield a set of independent observations of the random variables maximum crest and wave heights. 

For each $I_i$, the former is defined as follows:

$$C_{mi} = \max \{ \eta(x, y, t) \mid (x, y, t) \in I_i \} \quad (15)$$

while the individual maximum wave height $H_{mi}$ for $(x, y, t) \in I_i$ was calculated via zero-crossing analysis of each time series of duration $D$ defined in the sampled $xy$-space. The statistical independence of these realizations was verified by checking that maxima did not belong to the same 3D wave group.

A similar strategy and same time interval were used to extract the time extreme statistics using point-wise radar data collected in the Adriatic Sea. In that case the continuous time series of surface elevations $\eta(t)$ has been divided in consecutive time intervals $\Delta = 3600$ s, each of them split in $D = 180$ s long non-overlapping records $R_i$ (20 time per hour). Then, for the $i$th record $R_i$, we defined the maximum crest height as follows

$$C_{mi} = \max \{ \eta(t) \mid t \in R_i \} \quad (16)$$

and, accordingly, $H_{mi}$ was computed by adopting a zero-crossing analysis within each interval. This analysis produced a sample, at hourly intervals, of 20 individual values of $C_{mi}$ and $H_{mi}$, whose averages $<C_{mi}>$ and $<H_{mi}>$ were compared with model predictions centered at the same hour. Observed extremes have been up-scaled to account for the underestimation of wave height statistics due to the sampling-rate (Tayfun, 1993). Correction factors depend on the sampling interval and mean wave period; they vary in the range [1.00,1.08], with most probable values around 1.03. The histogram of observed values is shown in Figure 2. Mean and maximum values for $<C_{mi}>$ and $<H_{mi}>$ are, respectively, 0.85 m and 1.59 m, and 4.34 m and 6.88 m.
Figure 2 – Empirical density estimate (Histogram) of time extreme waves (point statistics) observed at the Acqua Alta platform in the Adriatic Sea (March-April and September-October 2018). Mean values of the maximum crest height $C_{\text{mi}}$ and maximum wave height $H_{\text{mi}}$ are shown with a blue and red marker, respectively.

To complement Adriatic Sea data, in the North Sea raw sea surface elevation data $\eta(t)$ were collected by two Datawell Directional Waverider buoys, moored at the stations Westerland ($8^\circ 13' 18.12''$ E, $54^\circ 55' 01.92''$ N, depth 13.3 m; years spanning 2013-2017) and Helgoland ($7^\circ 52' 04.80''$ E, $54^\circ 09' 36.00''$ N, depth 15.7 m; years spanning 2011-2017). In order to obtain empirical estimates of $\overline{C}_m$ and $\overline{H}_m$, we averaged twelve values of maximum crest height $C_{\text{mi}}$ and crest-to-trough height $H_{\text{mi}}$ (from zero-crossing analysis) that were obtained by splitting the 30-minute records into two equally long bursts ($D = 900$ s) and then assuming a stationary sea state for a duration of 3 hours. The Tayfun correction factors due to the buoy sampling-rate varied in the range $[1.0, 1.1]$, with most probable values around 1.05. Histograms of the observed maximum wave parameters from the two buoys are shown in Figure 3. Results highlight the different characteristics of the wave climate at the two stations in the North Sea; in particular, the Helgoland station displays wave extremes that are around 20 %, on average, and 50 %, at most, higher than those at the Westerland station.
Figure 3 – Empirical density estimate (Histogram) of time extreme waves (point statistics) observed at the Westerland buoy (dashed lines) and at the Helgoland buoy (solid lines) in the North Sea. Mean values of the maximum crest height $C_{\text{mi}}$ and maximum wave height $H_{\text{mi}}$ are shown with a blue and red marker, respectively.

2.3 Wave modeling: WAM and WW3 setups and forcings

Wave model datasets have been based on WAM cycle 6 (http://mywave.github.io/WAM/) and WW3 version 6.07 (https://polar.ncep.noaa.gov/waves/wavewatch/). Both models integrate the spectral wave action equation in space and time, with discretized wavenumbers, or frequencies, and directions, and balance conservative wave processes by the non-conservative sources and sinks. For the purposes of the study, WAM and WW3 were set up in the Adriatic Sea and Mediterranean Sea sharing the same domains and computational grids, whilst in the North Sea only WAM runs were taken into account. Models were prepared, on the one hand, to assess the models’ capability of reproducing the maximum wave statistics derived from the Adriatic Sea (time and space-time extreme data) and North Sea (time extreme data) observations, and, on the other hand, to make a WAM versus WW3 comparison and a climate assessment of the maximum waves in the Mediterranean Sea.

At first, to evaluate the extreme wave predictions at the position of the Acqua Alta platform (section 3.2), WAM and WW3 used a model grid covering the Adriatic Sea and Ionian Sea (the southern boundary is placed along the 39th parallel north where no wave forcing was imposed) with 0.05° uniform resolution in longitude and latitude, and a spectral grid composed of 36 evenly spaced directions and 32 frequencies exponentially spaced from 0.0500 to 0.9597 Hz at an increment of 10%. In WAM, the wind-wave interaction source term is based on the wave growth theory of Miles (1957) that was applied to wave forecasting by Janssen (1991). In our study, wind input + dissipation source terms (ST) were switched to the Bidlot et al. (2005) configuration, with parameters...
chosen to match the operational ECWAM contained in the forecasting system of ECMWF since version CY38R1 (ECMWF, 2012). Dissipation owing to bottom friction is taken from the JONSWAP parameterization (Hasselmann et al., 1973), with friction coefficient equal to -0.038 m²s⁻³ (default in WAM and ECWAM).

WW3-based Adriatic Sea runs were implemented setting three alternative ST parameterizations, namely ST3, ST4, and ST6 (WW3DG, 2019). They describe the wind input, wave dissipation due to whitecapping, and swell dissipation differently. With the nonlinear four-wave interactions (the same for all ST setups), those are the main processes in deep-water conditions. The objective is to focus on the ST performance especially on the predictions of maximum waves.

The ST3 parameterization (hereinafter WW3-ST3) follows ECMWF (2012) to make the source term implementation as close as possible to the one in the WAM model. The ST4 setup (hereinafter WW3-ST4) is described by Ardhuin et al. (2010), and, as it was originally implemented, it allows a set of parameters to be defined by the user. We focused on two of them, namely $\beta_{\text{max}}$ and $z_{0,\text{max}}$ that were accordingly modified using three different combinations: i) WW3 default parameters of TEST471; ii) parameters of TEST405 of Ardhuin et al. (2010); iii) parameters of TEST471, but setting $\beta_{\text{max}} = 1.55$ and $z_{0,\text{max}} = 0.002$ as in TEST405, which was thought to perform well for younger seas. From our tests, the last setup provided the best performance (not shown here) compared with observations of $H_s$ at Acqua Alta and least scatter with respect to WAM and WW3-ST3 outputs, and it was then used for all ST4-based implementations of this study. Finally, ST6 (hereinafter WW3-ST6) package for parameterizations of wind input, wave breaking, and swell dissipation terms was set using the default configuration (Babanin et al., 2010; Liu et al., 2019; Rogers et al., 2012; WW3DG, 2019; Zieger et al., 2015). For all three above-mentioned WW3 setups, as for WAM, the bottom friction was switched to the JONSWAP formulation with default value of -0.067 m²s⁻³ for the friction coefficient. This definition is different from the standard in WAM.

Further, numerical schemes for WAM (first-order upwind) and WW3 (third-order with garden-sprinkler effect alleviation) were kept as default. Although different numerics may lead to different results (ECMWF, 2019), we have chosen to test and intercompare the models in their standard configuration for wave action equation solution, aware that differences in model results might be partially due to numerics. The purpose of testing three physical formulations in WW3 is to make an analysis, albeit preliminary, of their relative performances, with focus on the estimation of extremes and related spectral parameters, and to verify, on the one hand, if the WW3 behaviour compared to WAM is influenced by the source terms, which, on the other hand, would be an unimportant factor for
wave extremes, obviously excluding the correct computation of $H_s$. In this respect, Stopa et al. (2016) made a thorough comparison of the three source term parameterizations. They concluded that ST3 is a good predictor of $H_s$ and higher-order moments of the wave spectrum and can be used under the majority of wave conditions; for all sea states ST4 has minimal $H_s$ biases and the higher-order wave parameters perform reasonably well; ST6 performs similarly to ST4, with major differences in the higher-order wave parameters and low wave heights. A recent recalibration of ST6 improved its performance and it was made available in WW3 since version 6.07 (Liu et al., 2019). However, differences between ST4 and ST6 still remain with respect to the directionality of the wave field.

In the Adriatic Sea runs, WAM and WW3 were forced with high-resolution 10-meter height wind fields produced by the COSMO-I2 forecasting local model. The COSMO-I2 model is run operationally at the Regional Environmental Agency (ARPA-SIMC) of the Emilia Romagna region (Italy) twice a day (at 00UTC and 12UTC), producing hourly wind fields with horizontal spatial resolution of 0.025° x 0.025°. COSMO-I2 winds performed reasonably well in the area of interest (Figure 4). At the 95th percentile level the model minus observation bias is -0.2 m/s (left panel), while wind model results are less accurate for rarer events. Figure 4 depicts on the right panel a time series comparison between modeled and observed wind speeds during the Sirocco storms of the 28-29 October 2018. There is good agreement between the two data, being the Cross-correlation Coefficient $CC = 0.85$, the Root Mean Square Difference $RMSD = 2.48$ m/s, and the slope of the best-fit line between modeled and observed winds $p = 1.03$

![COSMO-I2 wind performance at the Acqua Alta platform (March-April and September-October 2018).](image)

*Figure 4 – COSMO-I2 wind performance at the Acqua Alta platform (March-April and September-October 2018). (left) Scatter diagram between COSMO-I2 (MOD) and observed (OBS) values of $U_{10}$. Colours are given as number of entries. Circle markers show the quantile-quantile plot (the vertical line shows the 95th percentile). (right) Time series comparison during the 28-29 October 2018 Sirocco storm in the northern Adriatic.*
The Mediterranean Sea wave model setup covers the whole basin with 0.05° uniform resolution in longitude and latitude and with a spectral grid composed of 36 evenly spaced directions and 32 frequencies exponentially spaced from 0.0500 to 0.9597 Hz at an increment of 10%. WAM was run with the same setup as for the Adriatic Sea implementation, while WW3 was formulated using the ST4 source term configuration, with input parameters as in the Adriatic Sea run (i.e. ST4 default value configuration, but with adjusted coefficients $\beta_{\text{max}} = 1.55$ and $z_{0,\text{max}} = 0.002$). Wave models were forced with hourly winds from the ECMWF reanalysis ERA5 (https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation) that were downloaded at the nominal resolution of 0.25° (the native ERA5 resolution is 0.28125°). Simulations spanned 10 years, from 01 January 2001 to 31 December 2010.

Despite partially lacking the high resolution recommended for the wind forcing in closed and semi-enclosed sea areas (Cavaleri and Bertotti, 2004), the worthiness of the ERA5 forcing wind field in the Mediterranean Sea was demonstrated taking advantage of the derived measurements of $U_{10}$ from the ENVISAT-1 satellite altimeter for the period 2002-2010, which partially spans the numerical simulations period. Typical error metrics of the ERA5 performance clearly indicate that model winds are overall well reproduced in the basin (Figure 5): the model minus observation bias is -0.02 m/s, the RMSD is 1.44 m/s and best fit linear slope $p$ is slightly smaller than 1 ($p = 0.97$). Notwithstanding these good marks, there are features that are worth being mentioned. At first (not shown here) performances are not spatially homogenous, since the ERA5 resolution penalizes the coastal areas and the sea regions sheltered by small islands. Moreover, there is clearly a $U_{10}$-dependent performance: the quantile-quantile plot (circles in Figure 5) and its residual (right panel) exhibit a negative model bias that grows with the wind intensity, i.e. the bias is larger for the more intense storm. This residual is equal to -0.1 m/s, -0.4 m/s, and -1.2 m/s for the 90th, 95th, and 99th percentile, respectively. Given this score result, the 99th percentile threshold is adopted as the highest confidence level for sea states and wave extreme model assessment in the Mediterranean Sea (section 3.3).
Figure 5 – ERA5 10-m height wind performance in the Mediterranean Sea (years 2002-2010). ERA5 model and ENVISAT-1 satellite altimeter $U_{10}$ data (600101 collocations). (left) Scatter diagram and quantile-quantile plot (circles). (right) Model minus observation Bias versus model wind speed threshold.

In the North Sea, WAM hindcast spanned eight years from 2010 to 2017 and was applied to the GCOAST area (Geesthacht Coupled cOAstal model SysTem) that includes the North Sea and the Baltic Sea (Staneva et al., 2018, 2015). WAM was run on a model grid located between [40°04’00” N to 65°56’15” N] and [-19°53’20” E to 30°10’00” E] with a spatial resolution of about 3.5 km (318745 active model grid points). The solution of the energy balance equation was provided for 24 directional bands at 15° each, starting at 7.5° and measured clockwise with respect to true North, and 30 frequencies logarithmically spaced from 0.042 Hz to 0.66 Hz at intervals of $\Delta f/f = 0.1$. The wave model takes into account depth refraction and depth-induced wave breaking and is driven by hourly ERA5 wind fields. At its open boundaries, it receives values generated by a coarse grid model (spatial resolution 0.25° x 0.25°) for the North Atlantic (situated from 5° S to 75° N and from -100° W to 40° E). The integrated wave parameters for all active model grid points and spectral data at locations where measurements are available are stored at hourly intervals.
3 Results

3.1 The WAM implementation for maximum waves

The novel WAM code reproduces the structures that can be encountered in WW3, with some tweaks to adapt them to the WAM philosophy and some optimization addenda. In this section, we focus on how the parameters for maximum waves have been included in WAM (new subroutine WAMAX; details in Appendix A) and how the computational time that is required to evaluate the new variables has been optimized. WAM input files and namelists have been adapted. In particular, modifications dealt with the following items:

- Three new input variables: $X$ (in metres), $Y$ (in metres), $D$ (in seconds). For the computations of space-time extremes all three values must be larger than zero.
- Four new output variables: $\mathbf{c}_m$ (in metres) and $\mathbf{r}_m$ (in metres), for both time extremes and space-time extremes.

3.1.1 Spectral moments and derived parameters

Focusing that leads to energy concentration in wave trains is a phase-governed process. The use of phase-averaged models imposes to set aside a deterministic approach for the description of maximum waves and, instead, to use a statistical approach as a basis of the methods used in this study. We then exclude any examination of wave-by-wave heights and, rather, we look at the properties of the wave field as it is customarily described by the directional wave energy spectrum $S(\omega, \theta)$. Hence, consistently with the WW3 implementation, to estimate all parameters included in the extreme wave pdfs, we rely on the $ijl$-moments $m_{ijl}$ of $S(\omega, \theta)$ defined as follows:

$$m_{ijl} = \int \int k_x^i k_y^j \omega^l S(\omega, \theta) d\theta d\omega \quad (17)$$

where $\omega$ and $\theta$ are intrinsic angular wave frequency and wave direction, $(k_x, k_y)$ are the components of the wavenumber vector $k$. No contribution from the frequency spectral tail is taken into account, and the full spectrum is integrated without partition for multiple wave systems. Moments $m_{ijl}$ with $i \neq 0$ and $j \neq 0$ depend on the orientation of the coordinate system and it should be a natural choice to select the $x$- and $y$-axis orientation in such a way that the variance of spatial derivatives along one axis (say $x$) is maximized (Baxevani and Rychlik, 2004). This condition is satisfied assuming the peak direction of wave propagation coincident with the $x$-axis. After this axis orientation, below we describe the way the new variables are used in WAM and their meaning.
To estimate the value of the bandwidth $\psi^*$ that is used in the distribution of maximum wave heights, we rely on Boccotti (2000) who defined $\psi^*$ for general bandwidth processes as the normalized minimum of the temporal autocovariance function $\psi(\tau)$ of the sea surface elevation field $\eta(t)$. That is,

$$\psi^* = \psi(\tau^*)/\psi(0) \quad (18)$$

where $\tau^*$ is the abscissa of the absolute minimum of $\psi(\tau)$, which is defined as follows:

$$\psi(\tau) = <\eta(t)\eta(t+\tau)> \quad (19)$$

where $\tau$ is the time lag. The autocovariance function $\psi(\tau)$ is estimated from the directional spectrum as follows:

$$\psi(\tau) = \int \int S(\omega, \theta) d\theta \cos(\omega\tau) d\omega \quad (20)$$

In regard to the nonlinear, large crest height distribution, an integral measure of the steepness $\mu$ that accounts for bandwidth effects has been proposed by Fedele and Tayfun (2009):

$$\mu = \mu_a (1 - v + v^2) \quad (21)$$

where

$$\mu_a = \sigma k_a = \sigma \omega_k^2 / g = \sigma (m_{001} / m_{000})^2 / g \quad (22)$$

is an integral measure of the steepness for narrow band waves in deep waters, and $v = \sqrt{m_{000}m_{002}/m_{001}^2} - 1$ is the spectral bandwidth (Longuet-Higgins, 1975).

As for the space-time extremes formulations, it is pivotal to obtain the three values of average number of waves $N_1, N_2$ and $N_3$ that are defined in Eq. (10). They include the irregularity parameters, which are given by:

$$\alpha_{xt} = \frac{m_{101}}{\sqrt{m_{200}m_{002}}}, \alpha_{yt} = \frac{m_{011}}{\sqrt{m_{020}m_{002}}}, \alpha_{xy} = \frac{m_{110}}{\sqrt{m_{200}m_{020}}} \quad (23)$$

and are equal to the normalized cross-correlation coefficient between the components of the sea surface gradient. It is worth noting that while some moments $m_{ijl}$ change with the orientation of the axes (Baxevani and Rychlik, 2004), the determinant of the sea surface gradients covariance matrix which enter the extreme estimate is unaffected by an axes rotation. With the selected axes, moments $m_{200}$, $m_{020}$ and $m_{002}$ are related to the mean zero-crossing period ($T_z$), the mean zero-crossing wavelength ($L_x$) and the mean wave crest length ($L_y$) by the following equations:

$$T_z = 2\pi \sqrt{\frac{m_{000}}{m_{002}}}, L_x = 2\pi \sqrt{\frac{m_{000}}{m_{200}}}, L_y = 2\pi \sqrt{\frac{m_{000}}{m_{020}}} \quad (24)$$
where $m_{000} = \sigma^2$ is the variance of the wave field. Such mean wave parameters have a direct connection with the mean square slope ($m_{s}$) of the sea surface elevation field; indeed, we can write $m_{s}$ as follows:

$$m_{s} = (2\pi)^2 m_{000} \left( \frac{1}{L_x} + \frac{1}{L_y} \right)$$  \hspace{1cm} (25)

The so defined wave lengths are proportional to the fourth-order moment of the wave spectrum (in deep water $k^2$ is proportional to $f^4$), and therefore, as for $m_{s}$, they are strongly influenced by the high-frequency wave components. This leads to the important result that extremes depend on the higher-order spectral moments, which were verified to be sensitive to some extent to the source-term parameterizations in spectral models (Rascle and Ardhuin, 2013; Stopa et al., 2016). In this respect, assessment of model extremes is also an indirect way for the evaluation of the directional spectrum shape.

In Eq. (11), the computation of the mode of the probability density function of space-time maximum linear crest heights requires the solution of an implicit equation in $\xi_0$ (Fedele, 2012). Therefore, to optimize numerical performance, in WW3 we adopted a strategy that gives an estimate of $\xi_0$ as follows (Barbariol et al., 2017):

$$\xi_0 \approx \sqrt{2\ln(N_3) + 2\ln[2\ln(N_3) + 2\ln[2\ln(N_3)]]}$$  \hspace{1cm} (26)

where $N_3$ is given by Eq. (10). The above approximation strictly works when the number of 3D waves is large enough. Given the customary duration of a sea state, including $O(10^7)$ waves on average over the duration $D$, error in using Eq. (26) is smaller than 3% when the sea surface area $XY > L_x L_y$. In WAM, we have adopted the same strategy.

### 3.1.2 Numerical performance optimization

The most time-demanding task in WAMAX involves the computation of the auto-covariance function $\psi(\tau)$; this stage was then optimized in order to reduce the time required for its calculation. Clearly, there is no need for a full computation of $\psi(\tau)$ for $\tau \in [0, \infty)$, as only the normalized minimum is required, being $\psi^* = \min \{\psi(\tau)\} / \max \{\psi(\tau)\}$. The maximum value $\max \{\psi(\tau)\}$ occurs clearly for $\tau = 0$ and it equals the elevation variance $m_{000}$. Hence only the time lag and value of $\min \{\psi(\tau)\}$ must be sought. This led Barbariol et al. (2017) to develop an ad-hoc strategy, which was implemented in WW3. This implementation computes twenty points around the time position where the minimum is assumed to occur. This method is robust, and its computational time is known, which is an advantage for operational purposes. However, this strategy can not control the accuracy of the result,
even though the actual error is less than a few percent of the actual value. The point is that the assumption that the minimum is located inside a region (which is statistically determined on simple spectra) may fall.

The use of a search method to estimate the second zero of the $\psi(\tau)$ derivative can overcome this aspect and help improve the accuracy and reduce the computational cost. The optimal alternatives we tested are the Newton method, the secant method, the bi-section method, and the golden-section search. For all strategies, a reliable starting point/region is needed. This is especially crucial for the Newton method. In order to determine a consistent statistic for the estimate, we evaluated the time $\tau_{\text{min}1}$ of the first minimum of $\psi(\tau)$ using one month (January 2014) of global ERA-Interim directional spectra. Afterwards we normalized $\tau_{\text{min}1}$ and the half-period $\tau_{\text{max}2} - \tau_{\text{min}1}$ (where $\tau_{\text{max}2}$ is the time of the second maximum of $\psi$) with respect to three characteristic wave periods: $T_{m01}$, $T_{m02}$ and $T_e = T_{m-10}$. Table 2 indicates the median values. The distribution of $T_e$ appears sharper than the other two, thus the value $\tau = 0.47T_e$ will be taken as the starting point for the search.

<table>
<thead>
<tr>
<th>$T_m = T_{m01}$</th>
<th>$T_m = T_{m02}$</th>
<th>$T_m = T_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{\text{min}1}/T_m$</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>$(\tau_{\text{max}2} - \tau_{\text{min}1})/T_m$</td>
<td>0.69</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 2 – Median values of the first minimum $\tau_{\text{min}1}$ and half-period $\tau_{\text{max}2} - \tau_{\text{min}1}$ with respect to characteristic periods (data from global hourly ERA-Interim directional wave spectra for the month of January 2014).

Moreover, using the ERA-Interim spectra we found that the best computational option is represented by the golden-section search. The first option, i.e. Newton method is the fastest (see Table 3), even if it had to be degraded to a quasi-Newton method, with a search path strategy. Problems arise from the second derivative of the autocovariance function which, besides doubling the computational load, often shows problematic behavior leading to non-convergence. The Secant method happened to be slower than the quasi-Newton strategy, even though the number of iterations could be controlled by accepting a sufficiently high tolerance. The most robust method has proven to be the bi-section, despite being the slowest. The golden-section search showed a stable and trustworthy behaviour, while maintaining a certain speed.

The computational time required to produce extreme wave outputs was found to have a small impact on the total CPU load used by the model. Using the Mediterranean Sea setup described in section 2.3 the time spent in WAMAX is less than 0.5% of the total model time.
<table>
<thead>
<tr>
<th></th>
<th>GS</th>
<th>QN</th>
<th>S</th>
<th>B</th>
<th>WW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{iter}}$</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>$N_{\text{int}}$</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>% fail</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3 - Median number of iterations ($N_{\text{iter}}$) and total number of integrals ($N_{\text{int}}$) to be computed in success conditions. Last row indicates the failure rate. Headers indicate Golden-section (GS), Quasi-Newton (QN), Secant (S), Bisection (B) and the WW3 implementation.

We have also tested cases where the convergence may fail in providing a good estimate of $\psi^*$. Indeed, it may happen that the first minimum of $\psi(\tau)$ is not the absolute one, and in very few cases it lies above zero. Those situations occur in mixed sea states, when a swell is super-imposed to a wind sea of similar energy (see for instance Fig. 4.9 in Boccotti, 2000). In this respect, Figure 6 shows the scatter between the values of the first minimum $\psi_{\text{min1}}$ and of the second minimum $\psi_{\text{min2}}$. In most of the cases (99 %), the first minimum is deeper than the second one, but there are few wave conditions for which the first local minimum is not the absolute one. These, however, are represented by non-severe sea conditions (median value of $H_s$ about 0.4 m, and maximum below 2 m), for which, as a consequence, the prediction of $\overline{H}_m$ might be slightly underestimated.

Figure 6 – Scatter diagram of $\psi_{\text{min2}}$ vs. $\psi_{\text{min1}}$ normalized with $\psi_{\text{max}}$. Colours indicate the number of entries in logarithmic scale (data from global hourly ERA-Interim directional wave spectra for the month of January 2014).

3.2 WAM and WW3 maximum wave assessment (Adriatic Sea and North Sea)

Preliminary to any analysis of model outputs is the assessment of model scores using reference observations of targeted parameters. In this respect, we must recall that maximum crest and wave heights derived from numerical runs were compared against in situ observations in a few previous studies. Barbariol et al. (2019) used ERA-Interim and ERA5 wave spectra in the Pacific Ocean (Station Papa) to assess time and space-time extreme wave
statistics against stereo (two experiments in a single day near the station) and buoy data (at the station), while in ECWAM envelope maximum heights were assessed against Canadian and Norwegian buoys (ECMWF, 2019). However, in most cases, model extremes were characterized off-line by processing 2D wave spectra, either computed from the data (Benetazzo et al., 2015) or modeled at the specific location of wave observations (Fedele et al., 2016).

Having wave models with inbuilt routines for extremes allows bypassing this strategy since extremes may be computed on-line by wave models whose overall performance can be hence assessed. To support this evaluation, in this study, 2D spectra were still processed off-line to calculate ancillary variables, such as the wave steepness and the spectral bandwidth that enter in the theoretical formulations. In-situ observations from stereo cameras, wave radar, and wave buoys are hence compared with the modeled results. For this purpose, data collected in the Adriatic Sea are here used to assess the WAM and WW3 capabilities of correctly estimating the statistics of observed maximum waves, whilst in the North Sea only WAM outputs were considered. We distinguish between space-time and time extremes, as we show below.

### 3.2.1 Space-time extremes

In regard to the space-time statistics, observed wave parameters are summarized in Table 1, which encloses details of the experiments we performed with the stereo imaging system. The reference 3D space-time region \( \Gamma \) used to produce the extreme data sample has orthogonal sides \( X = 30 \) m, \( Y = 30 \) m, and \( D = 180 \) s. Same values were then set in WAM and WW3 model runs. The observational wave sample is composed by a variety of sea conditions that, accordingly, produced a large spread of extreme values: \( \bar{C}_m \) is equal to \( 1.01H_s \) at least and \( 1.30H_s \) at most, while \( \bar{H}_m \) ranges between \( 1.65H_s \) and \( 1.99H_s \). It is understood that the proportional relation in Eqs. (11) and (12) between extremes and the number of waves \( N_3, N_2 \) and \( N_1 \) provides for a fixed space-time region \( XYD \) and given \( H_s \) (that is the principal scale factor for maximum wave parameters) a path for the maximum waves whose expectations \( \bar{C}_m \) and \( \bar{H}_m \) decrease (increase) for sea states that comprises, on average, longer (shorter) waves (Mori and Janssen, 2006). We shall partially account for this behaviour in the following analysis.

Before quantitatively evaluating the errors on extreme values for each model parameterization, we firstly proceed assessing the performance in the reproduction of the \( H_s \) behaviour at the Acqua Alta platform for the four months 01 March – 30 April and 01 September – 31 October 2018 when stereo data were collected (Figure 7, top-
We observe that all model configurations show a degree of scatter around the observations, which however are well modeled on the whole, with absolute bias for $H_s$ smaller than 0.1 m. Despite having a somewhat limited prospect (both temporal, spanning four months, and spatial, comprising a single station), this result is consistent with that of Stopa et al. (2016), who noted that at a global scale the three WW3 source term parameterizations behave reasonably well for most wave conditions. Going into details, we note that WAM generally overestimates WW3 with relative bias larger than 0.07 m. Instead, there are small differences among the three WW3 parameterizations, being the relative bias at most 0.06 m, and, in particular, ST4 and ST6 that perform similarly up to $H_s = 3$ m. The largest differences between WAM and WW3 (all parameterizations) occur when the sea states are more energetic, as can be observed by comparing the Q-Q (quantile-quantile) plot of $H_s$ depicted in Figure 7 (bottom-left): the probability distributions of $H_s$ deviate above the 95th percentile, where the WW3 minus WAM residual reaches on average about -10 %. The largest discrepancy visible at $H_s > 4$ m is associated to the peak condition of a single storm (occurred on 28-29 October 2018), which, hence, has small statistical significance. We note, however, that differences between WAM and WW3 may be partially explained based on the friction coefficient default value which makes the bottom friction of WW3 stronger.

Examining the space-time extreme values for the same period, scatter and Q-Q plots of $\overline{C}_m$ and $\overline{\eta}_m$ in Figure 7 (central and right panels, respectively) show that, as for $H_s$, WAM slightly overestimates WW3. The WW3 minus WAM bias is small and larger than -0.17 m for $\overline{\eta}_m$ and -0.11 m for $\overline{C}_m$. WW3 runs with different source term parameterizations show similar distribution (maximum bias is about 0.1 m for both variables), in particular ST4 and ST6, with the only exception of the October 29 storm when $\overline{C}_m \approx 4$ m and $\overline{\eta}_m \approx 6$ m. WW3 underestimation with respect to WAM grows with the order of the quantile percentage: at the 95th percentile (shown with a vertical dashed line in the QQ-plots) the residuals are, on average, about -8 % and grow for the rarer conditions.
Figure 7 - Scatter diagram (top; in the legend the bias with respect the observations, OBS) and quantile-quantile plot (bottom) of model (MODEL) significant wave height $H_s$ and space-time extremes $\bar{c}_m$ and $\bar{r}_m$. Acqua Alta platform data (01 March – 30 April 2018 and 01 September – 31 October 2018). The size of the 3D space-time region used for extreme value evaluation is 30 m x 30 m x 180 s. The dashed vertical black lines on the quantile-quantile plots show the 95th percentile, of which the normalized bias is shown in the legend. For error metrics, only wave data with $H_s > 0.5$ m is considered.

The differences on extreme values we have pointed out stem mostly from the performance of $H_s$, which is then the pivotal variable for maximum wave accuracy. To complement this view, we examine the spectral parameters that contribute to extreme estimates from model spectra. We concentrate on the nonlinear steepness parameter $\mu$, used in the statistics of $c_m$, and on the bandwidth parameter $\psi^*$, used in the statistics of $H_m$. Results are shown in Figure 8, where WW3 is compared with WAM (top panels) and model outputs with observations (bottom panels). We note that WW3 spectra are on average steeper than WAM spectra (the normalized bias for $\mu$ is at most $+6\%$ for ST6 and at least $+2\%$ for ST3), a fact that would favour, other parameters being equal (e.g. $H_s$ and the average number of waves within the space-time region), that WW3 maximum crest heights are slightly larger than WAM heights. Differences of a smaller degree, albeit pointing to the same direction (i.e. WW3 values larger than WAM ones), are observed for $|\psi^*|$: WW3 spectra are generally narrower, and hence the probability density function of maximum wave heights $H_m$ is expected to shift towards larger values. When compared to observations (bottom panels), WAM and WW3-ST3 have a good match with the observed $\mu$, while WW3-ST4 and
WW3-ST6 produce slightly steeper conditions (difference of a few percent). Inspecting the distribution of the variable $\sqrt{1 + |\psi'|}$ that is used in the computation of $H_m$, we note a large spreading of model data around the observation, which are, however, well reproduced on average by all models, being the normalized bias smaller than 1%. We shall return to the analysis of the WAM and WW3 values of $\mu$ and $\psi^*$ in the next section, where we shall assess extremes in the Mediterranean Sea.

Figure 8 – Scatter diagram of the steepness parameter $\mu$, absolute value of the minimum of the autocovariance function $|\psi'|$, and maximum wave height parameter $\sqrt{1 + |\psi'|}$. On top panels, in the legend, the normalized bias of WW3 (ST3, ST4 and ST6) with respect to WAM is shown, while on bottom panels, the normalized WAM and WW3 bias with respect to observations (OBS) is shown. Acqua Alta platform model (MODEL) and observed (OBS) data (01 March - 30 April 2018 and 01 September - 31 October 2018).

Next, we focus our attention on the comparison between modeled and stereo derived space-time extremes. Granted the possible difference between modeled and observed $H_s$, in order to remove the contribution of the sea state severity, the assessment is provided for values of $\overline{C}_m$ and $H_m$ that are normalized with $H_s$ (indicated as $\overline{C}_m$ and $H_m'$ in Figure 9). We observe for WAM a small negative bias (-0.03$H_s$ and -0.02$H_s$ for $\overline{C}_m$ and $H_m'$, respectively), suggesting that model frequency spectra are slightly less steep and broader than observations. The opposite is true for WW3, with a positive bias for both extreme values, and with ST4 and ST6 that typically produce similar values and with smaller variability than ST3. As it was stated before, the normalization with $H_s$ removes only part of the difference between observed and modeled values, since the extreme distributions
incorporate also the sample size (i.e. the average number of waves $N_1$, $N_2$ and $N_3$) over the space-time domain. For example, for a fixed number of waves, the maximum wave height distribution in the time domain according to Rayleigh theory is constant. We shall examine how to remove the effect of the sample size in the analysis provided in section 3.3. For the time being, we point out that, after normalization, more (less) energetic sea states produce longer (shorter) waves and, hence, for a fixed space-time or time region, a smaller (larger) number of waves, which imply smaller (larger) wave extremes. It is easily understood, therefore, that the probability of extreme waves exceeding a specified value depends on the number of waves within the record.

To account for this variability, we have analyzed the relative error between model and observed normalized extremes as a function of the difference between mean wave periods, that we use as a proxy for the number of waves (Figure 9, right panels). As expected, in confirmation of the theoretical approaches, positive errors (i.e. model $>$ observation) imply that the model period is smaller than the observed period, and vice versa. However, overall, model errors are small for all model runs. Indeed, model estimates of normalized space-time maximum wave parameters compared fairly with stereo observations: for $\bar{C}'_m$, the model minus observations mean difference is 1.1% with a root-mean-square-difference (RMSD) of 4.9 %, while for $\bar{H}'_m$ the mean difference is 1.9 % and the RMSD of 4.8 %.
Figure 9 – Space-time extreme assessment. Comparison between normalized (with $H_s$ indicated with prime marks) space-time maximum crest height $\bar{C}_{m,obs}$ (top) and maximum wave height $\bar{H}_{m,obs}$ (bottom) observed (OBS) using the stereo wave imaging system and modeled (MOD) using different configurations. In the legend of the left panels the bias is shown in brackets. On the right panels, $T_z$ is the zero-crossing mean period. The size of the 3D space-time region $I$ used for extreme value evaluation is 30 m x 30 m x 180 s.

### 3.2.2 Time extremes

As for the time extreme assessment, WAM and WW3 have been firstly tested against the observed values obtained processing the radar probe data from Acqua Alta in the same period (March-April and September-October 2018) during which we run models in the space-time extreme configuration. We have thus set $X = 0, Y = 0$ and $D = 180$ s in order to switch on the computation of the maximum wave parameters that are expected in the temporal domain (point statistics). Scatter diagrams in Figure 10 represent the comparison between observed and WAM values of $\bar{C}_m$ (Forristall distribution, left panel) and $\bar{H}_m$ (Naess distribution, right panel). Error metrics of the model performance (text in the panels) clearly indicates that WAM performs well in estimating both variables: indeed, the absolute bias is smaller than 3 cm, and the RMSD is 0.25 m and 0.41 m for crest and wave heights, respectively. Observed and modeled datasets are well correlated, with cross-correlation coefficient $CC = 0.90$ and best fit linear slope $\sim 1$ for both variables. When $\bar{C}_m > 3$ m and $\bar{H}_m > 5$ m, WAM overestimates the observations, although those data refer to the 29 October 2018 storm, which also challenged the model simulations of $H_s$, as we pointed out before. In the context of Forristall and Naess models for extremes, results presented here have similar errors as those presented
by Barbariol et al. (2019), who performed a comprehensive study of extremes using ERA-Interim and ERA5 wave spectra in the Pacific Ocean.

Figure 10 - Time extreme assessment (Adriatic Sea). (left) Comparison between WAM model estimate of $C_m$ based on the Forristall model (MOD) and observations (OBS). (right) Comparison between model estimate of $H_m$ based on the Naess model (MOD) and observations (OBS). On both panels the solid black curve shows the best-fit line with slope $p$. The colorbar is given on a linear scale representing data entries at 0.25-m bin.

Finally, WW3 and WAM time extreme model scores are summarized in Table 4 and show a good overall agreement between modeled and observed values of $C_m$ and $H_m$. As we have pointed out for the space-time extremes, the WW3 source term parameterization has a minor impact on extremes; in particular those obtained setting ST4 and ST6 show good and almost identical statistical scores. ST3 results are slightly less satisfactory, although with differences of a few centimetres that we do not consider as statistically relevant. Averaging model scores, we can conclude that $C_m$ and $H_m$ are estimated, respectively, with absolute bias of 3 cm and 10 cm, and with RMSD of 24 cm and 40 cm.
In the North Sea, surface elevation buoy data at the two stations Westerland and Helgoland provided an opportunity to compare the WAM time extremes with an extensive dataset covering a few years. Model performance for $H_s$, $\overline{c}_m$ and $\overline{H}_m$ are summarized in Table 5. On both stations, WAM performs well in reproducing $H_s$, with under/overestimation within 5%, and absolute model-minus-observation bias smaller than 0.08 m. Likewise, WAM extremes show similar scores, with departures from observations consistent with those of $H_s$. Quantitatively, performance at the two North Sea buoys is similar to that found at the Acqua Alta platform. This result confirms the validity of the theoretical and model approaches that are used to estimate extremes in WAM and WW3 and gives it generality.

**Table 4** – Time extreme assessment (Adriatic Sea). Comparison between WAM and WW3 model estimates and observations at the Acqua Alta platform (Adriatic Sea). Error metrics is given by the cross-correlation coefficient (CC), the root-mean-square-difference (RMSD), the model-minus-observation bias (Bias), and the slope $p$ of the best-fit line.

<table>
<thead>
<tr>
<th></th>
<th>$\overline{c}_m$, $\overline{H}_m$</th>
<th>CC</th>
<th>RMSD (m)</th>
<th>Bias (m)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAM</td>
<td></td>
<td>0.90, 0.90</td>
<td>0.25, 0.41</td>
<td>0.03, -0.03</td>
<td>1.02, 0.97</td>
</tr>
<tr>
<td>WW3 – ST3</td>
<td></td>
<td>0.90, 0.90</td>
<td>0.23, 0.40</td>
<td>-0.03, -0.16</td>
<td>0.93, 0.87</td>
</tr>
<tr>
<td>WW3 – ST4</td>
<td></td>
<td>0.91, 0.90</td>
<td>0.23, 0.39</td>
<td>0.02, -0.10</td>
<td>0.97, 0.90</td>
</tr>
<tr>
<td>WW3 – ST6</td>
<td></td>
<td>0.91, 0.90</td>
<td>0.24, 0.39</td>
<td>0.02, -0.10</td>
<td>1.00, 0.93</td>
</tr>
</tbody>
</table>

In summary, the evaluation of maximum wave parameters shows that model data are statistically consistent with spatio-temporal stereo and temporal probe and buoy data. In particular, the relevant role of $H_s$ in reproducing correct maximum wave heights must be acknowledged. With this in mind, the following analysis focuses on WAM and WW3 model results in the Mediterranean Sea, where we proceed with the aim of disentangling the principal spectral parameters responsible for the generation of high waves within the theoretical framework adopted in this study.

**Table 5** – Time extreme assessment (North Sea). Comparison between WAM model estimates and observations at the Westerland (2013-2017) and Helgoland (2011-2017) North Sea buoys. The value of $\overline{c}_m$ is computed with theoretical predictions based on the Forristall model, and $\overline{H}_m$ based on the Naess model. Error metrics is given by the slope $p$ of the best-fit line and the model-minus-observation bias (Bias).

<table>
<thead>
<tr>
<th></th>
<th>Westerland ($p$, Bias)</th>
<th>Helgoland ($p$, Bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_s$</td>
<td>1.05, 0.05 m</td>
<td>0.94, -0.08 m</td>
</tr>
<tr>
<td>$\overline{c}_m$</td>
<td>1.07, 0.06 m</td>
<td>0.96, -0.04 m</td>
</tr>
<tr>
<td>$\overline{H}_m$</td>
<td>1.03, 0.04 m</td>
<td>0.91, -0.20 m</td>
</tr>
</tbody>
</table>
3.3 Short-term/range extreme wave analysis in the Mediterranean Sea

The advantage of having a set of wave models equipped with routines capable of providing estimates of wave maxima with high spatial-temporal resolution is that predictions of their intensity can be made at large spatial and temporal scale. This type of analysis would help identify the sea regions exposed with high probability to very large waves and discern the physical processes responsible for their occurrence. For the analysis made in this study we have used WAM and WW3-ST4 results in the Mediterranean Sea, with the following goals: i) the assessment of the maximum wave climate in the Mediterranean Sea; ii) the comparison between WAM and WW3-ST4 predictions of wave maxima at basin scale; and iii) the understanding of the principal processes yielding maximum crest and wave heights. To pursue these objectives, we used the 10-year wave hindcast dataset using the space-time extreme formulations that we have applied with two strategies (a first one with fixed term and range, and a second one with variable term and range), as we describe in the following sections.

3.3.1 Pattern and size of maximum crest and wave heights

We aim at examining the intensity and spatial distribution of space-time maximum waves (crest and wave heights) that can be encountered in the Mediterranean Sea. We have adopted a fixed-volume space-time region with sides $X = 100$ m, $Y = 100$ m, and $D = 1200$ seconds, hence targeting the typical horizontal footprint of an offshore platform ($XY = 100 \times 100$ m$^2$) and the typical wave buoy record length (20 minutes), still satisfying the requirements of spatial homogeneity (short-range) and temporal stationarity (short-term) of the sea state. The latter, in particular, is fulfilled since 20-minute is a duration smaller than the timescale related to the changes in the synoptic conditions in the Mediterranean Sea. Due to its nature of semi-enclosed basin, previous global estimates of wave extremes (Barbariol et al., 2019), which relied on ERA-Interim spectra at 1° and 6-h resolution, lacked the possibility of describing the complex atmospheric pattern and the maximum wave conditions in the Mediterranean Sea. Using ERA5 reanalysis wind forcings, which are more resolved in space and time than the ERA-Interim ones, we focus on the spatial distribution of the 50th (the median value) and 99th percentile (a rare value with 1% chance of being exceeded in ten years) of the two variables $\bar{C}_m$ (Figure 11) and $\bar{H}_m$ (Figure 12). As for the spatial variability of $H_s$ (not shown here), individual waves with maximum heights are triggered in regions with strongest prevailing and dominant offshore winds (Cavaleri and Sclavo, 2006; Sartini et al., 2017; Soukissian et al., 2018), which blow in the northwestern Mediterranean basin, near the Gulf of Lion and Balearic Sea (north-westerly Mistral winds steadily present through the year), and in the Aegean Sea (north-easterly Etesian seasonal winds).
The typical value (50th percentile) of the probability distribution of $\overline{C}_m (\overline{H}_m)$ do not exceed about 1.8 m (2.8 m) east of Crete Island, and 1.6 m (2.6 m) south of the Gulf of Lion, and are as low as 0.8 m (1.2 m) in the enclosed sub-basins (such as the Adriatic Sea and the Tyrrenian Sea surrounding the Italian peninsula). The rest of the Mediterranean Sea has a median value of $\overline{C}_m (\overline{H}_m)$ smaller than 1.3 m (2.0 m). We do observe large regional differences with maxima arranged in patterns following the variety of meteorological features. We observe differences between WAM and WW3-ST4 maxima. In all the areas with highest median conditions, WW3-ST4 overestimates WAM up to 15 cm (25 cm), while differences are smaller (about 10 cm) in the calmest sub-basins.

The spatial variability of extreme waves for more intense storms (99th percentile) is indeed different. It is relevant where the Mistral wind jet drives very intense sea states in the Gulf of Lion, which experiences $\overline{C}_m = 8$ m and $\overline{H}_m = 12$ m, whilst the rest of the Mediterranean Sea does not exceed about $\overline{C}_m = 5.5$ m ($\overline{H}_m = 8$ m). For the strongest sea states, basin-wide differences between WAM and WW3-ST4 are more articulated, with WAM overestimating WW3-ST4 up to (Gulf of Lion) $\overline{C}_m = 0.5$ m ($\overline{H}_m = 1.0$ m) in the northernmost part of the Sea, while an underestimation of about 0.2 m is visible in the southern regions.
3.3.2 Extreme wave sensitivity to spectral parameters

The previous analysis was conducted using a fixed-volume $XYD$ for extremes, since this is the way the maximum wave evaluation was implemented in WAM and WW3. This choice, albeit being suitable for most practical applications (DNV GL, 2017), partially masks the capability of assessing the dependence of the occurrence of high waves on the characteristics of the sea state (such as nonlinearity parameters or spectral spreading), since the number of waves (i.e. the sample size) is as important as other factors for determining the probability of maximum crest and wave heights in linear and nonlinear wave trains (Mori and Janssen, 2006; Ochi, 1998). As we have made clear above in the comparison with observations, in order to gain insight into the principal features of the wave field in influencing maxima, the strategy that must be adopted would keep the number of waves as constant as possible and hence use, for each model directional spectrum, a sea-state dependent space-time region $I$ of variable volume $XYD$. We have thus loosened the variability of the number of waves $N_{3D}$ by normalizing the 3D space sides $X, Y$ and $D$ with the local and instantaneous spectral values of the mean zero-crossing wavelength $L_x$, crest length $L_y$, and wave period $T_z$, respectively. We have chosen the following values $X = L_x, Y = L_y, D = 100T_z$, and selected two key locations exposed to different sea states for a detailed analysis: the Acqua Alta platform position and the grid point of coordinates $[4.65^\circ$E, $42.06^\circ$N], just southeast of the Gulf of Lion.

We discuss the general behavior of the pdf of maximum crest and wave heights in the wave field, by showing the histograms of $\bar{C}_m$ and $\bar{H}_m$ normalized with $H_s$ as a function of $\mu$ and $\psi^*$, respectively, at the Acqua
Alta (Figure 13) and the Gulf of Lion station (Figure 14). Data are highly correlated, and the dependencies are clear, as it was expected from Eqs. (11) and (12). At Acqua Alta, WW3-ST4 produces sea states marginally steeper and less variable than WAM, which generate slightly larger (+2%) crest heights (Figure 13, left panels). However, differences with WAM are very small (<2%), also when the variability of wave heights (equal on average to about 1.95) with the bandwidth is taken into account (Figure 13, right panels). Normalized extreme values of \( \bar{C}_m (H_m) \) from both models vary (difference between 95th and 5th percentile values) of 0.18 (0.20), about 16% and 10% of the mean of \( \bar{C}_m \) and \( H_m \), respectively. At the Gulf of Lion station, spectra are +10% steeper and as narrow as those at Acqua Alta. This reflects proportionally in the values of wave maxima. It is interesting that, albeit the wind and wave regime at the two locations are indeed different (in variety and intensity), normalized maxima at the two locations have very similar distributions (even though variation of extremes in the Gulf of Lion is smaller) and values (<2% on average for both extremes), as well as WAM and WW3-ST4 not showing large differences, which instead exist for non-normalized extremes for which \( H_s \) scales the extreme values.

Figure 13 – Normalized (with \( H_s \)) space-time maximum crest and wave heights at the Acqua Alta platform (period January 2001-December 2010). Size of the space-time region: \( X = L_x \), \( Y = L_y \), \( D = 100T_z \). Variation of the expected value of the maximum crest height \( \bar{C}_m \) (left) and maximum crest height \( H_m \) (right) with mean wave steepness \( \mu \) and absolute value of the first minimum of the autocovariance function \( \psi^* \), respectively. WAM (top) and WW3-ST4 (bottom) run outputs. Data is given in a density scatter plot. For each panel, the title gives the [average value ± standard deviation] of the two variables, the text reports the Cross-correlation Coefficient (CC) between the variables, and the dashed black curve shows the best-fit line.
Figure 14 – Normalized (with $H_s$) space-time maximum crest and wave heights at the Gulf of Lion station (period January 2001-December 2010). Size of the space-time region: $X = L_x$, $Y = L_y$, $D = 100T_z$. Variation of the expected value of the maximum crest height $\mathcal{C}_m$ (left) and maximum crest height $\mathcal{R}_m$ (right) with mean wave steepness $\mu$ and absolute value of the first minimum of the autocovariance function $\psi^*$, respectively. WAM (top) and WW3-ST4 (bottom) run outputs. Data is given in a density scatter plot. For each panel, the title gives the [average value ± standard deviation] of the two variables, the text reports the Cross-correlation Coefficient (CC) between the variables, and the dashed black curve shows the best-fit line.

4 Discussion and Conclusions

Our present study opens a research direction on the extreme value prediction with spectral wave models. The variables of interest are the maximum crest height and the maximum crest-to-trough wave height in a given sea state. All parameters used in the time and space-time extreme theoretical distributions have been computed from the wave energy directional spectrum $S(f, \theta)$ or $S(k, \theta)$, which is then the key quantity used for the evaluation. As a consequence, because of the phase-averaging approach of spectral models, only statistical quantities of extreme waves, such as the mean, are defined and provided. In this study, the wave model WAM has been targeted for the new implementation, in order to comply with the existing WW3 community wave modeling framework. Extreme wave outputs have been assessed against observations, and differences between WAM and WW3 results have been considered. Ad-hoc runs have been performed to assess the principal features of the maximum wave climate in the Mediterranean Sea.

The estimate of extreme wave heights presented in this paper deals with wave conditions in short-term/range steady-state and homogenous seas (duration ~ 1 hour, extent in deep water ~1 km, far less in near-coast
areas). The basic assumption is that the wave spectrum as well as the significant wave height and spectral moments are constant in a given sea state. This is the way spectral models compute and process wave spectra, and, as a consequence, the short-term/range statistics is the unique option for the evaluation of extremes within such a model code. It may be of interest, however, to estimate the maximum waves in the short-range but long-term (from hours to days), when the sea severity (namely \( H_s \)) is changing, for instance during a storm, including its growing and decaying phases. The method to determine the long-term statistics of individual maximum wave heights in a storm relies on the accumulation of the statistics for all short-term stationary conditions, taking into account the occurrence probability and the number of waves for each piecewise steady-state sea state (Battjes, 1972; Borgman, 1973). For space-time extremes during storms the long-term exceedance probability of \( C_m \) was outlined by Fedele (2012), who assumed that the sea surface elevation field is spatially homogeneous over the area \( A \) and piecewise, non-stationary in a time interval covering a storm. Using those methods, the long-term analysis of extreme waves can be easily achieved, for instance, by post-processing, the wave spectra at a certain location in the ocean.

The distribution function of maximum wave heights has been used by computing the expected value (probability-weighted average) of maximum crest height and maximum wave height. Other parameters would be possible, such as the standard deviation of the extremes (that is included in WW3, but not in WAM), or the probability that the maximum wave heights exceed a given level (for instance the rogue wave threshold \( 1.25H_s \) for \( C_m \)). Our choice to release the expected values is based on the consideration that it can be easily compared with the arithmetic mean from observations, either from stereo systems or traditional wave probes. Theoretical statistics for extremes have been considered in both their temporal and spatio-temporal formulations, the latter allowing the description of the statistics of large wave groups crossing sea surface regions, which have a practical application, for instance, at the air gap problem of offshore structures (Forristall, 2011). In general, due to the larger sample size of space-time formulations, the likelihood of extreme event is increased.

The choice of the space-time region sides \( X, Y \) and \( D \) is constrained to the homogeneity and stationary conditions and it is arbitrary, depending only on practical applications (e.g. \( X \) and \( Y \) may correspond to the sides of a marine structure). We mention that the magnitude of extreme waves increase substantially with \( N_{1D}, N_{2D}, \) and \( N_{3D} \) during the first 30-60 minutes (Ochi, 1998) and few tens of meters (Benetazzo et al., 2017b), and thereafter increases slowly with time and space, irrespective of sea severity. It is important to note that the extreme value pdfs (both time and space-time) lack the saturation induced by wave breaking (Babanin, 2011) or nonlinear dispersion...
(Fedele, 2014), which should modify the far tail of the pdf by restricting the wave heights (Fedele et al., 2017).

This would suggest that some of the largest waves in the pdf might be at the onset of incipient breaking or already breaking. This effect is more pronounced for space-time extreme theoretical models with large sample size, which may overestimate maximum waves over large areas and time duration. In this respect, the study by Benetazzo et al. (2020) showed that space-time estimates are reliable over surface areas with a side $O(10^2)$ m for all sea states and time interval shorter than one hour.

Formulae for space-time extreme wave estimates depend on sea state parameters that are obtained from higher-order moments of the directional spectrum. As pointed out by Stopa et al. (2016) and Liu et al. (2019), wave models may have difficulty in describing the directional properties of wave spectra, and differences exist among the WW3 source term parameterizations as far as the wave spectrum moments are concerned. By comparing modeled space-time extremes against stereo observations, we have shown, however, that model performances and differences between models mostly stem from the accurate simulation of the significant wave height $H_s$, which represents the leading scale factor for maximum waves. We note also that models' relative performance is better for extremes than for $H_s$; the reason for that is that the underestimation (overestimation) of $H_s$ implies a larger (smaller) number of waves (i.e. the sample size over a fixed duration / region), which in turn gives rise to higher (smaller) maximum heights (compared to $H_s$).

With the WAMAX routine implementation, both WAM and WW3 are equipped with procedures to compute maximum waves parameters, though with some differences between models, as shown in Table 6. In particular, in WW3 the nonlinear formulation for the time ($X = 0$ and $Y = 0$) and space-time ($X \neq 0$ and $Y \neq 0$) crest height is based on the Tayfun model only, while in WAM the nonlinear crest height for time extremes is computed also using the Forristall model. Moreover, we choose to produce time extremes in WAM (Forristall and Naess models) also when setting $X \neq 0$ and $Y \neq 0$. This option is not available for WW3, where a strict and distinct set of variables is computed for time and space-time extremes. In this sense, WAM allows additional time extreme outputs to be computed compared to WW3, whereas the space-time extreme implementation is based on identical formulations. Some differences between WAM and WW3 concern also the numerical implementation and its optimization.
Table 6 – Maximum crest height $\bar{c}_m$ and crest-to-trough wave height $\bar{r}_m$. WW3 and WAM estimate of space-time extremes ($STE: X, Y \neq 0$ and $D \neq 0$) and time extremes ($TE: X, Y = 0$ and $D \neq 0$).

Finally, the capability of assessing modeled extremes is based on the capacity of processing observed data with ad-hoc strategies. Strictly speaking, the model response requires continuous time or space-time elevation records at high temporal and spatial resolution to be processed to isolate independent high wave trains. Such series of data are rarely available (see for instance https://cdip.ucsd.edu/m/) and this limits the possibility of extensive model output assessment, though spotty validations can still be carried out, as was established in this study.

To summarize some of the key achievements and findings of our study:

- WAM code was upgraded to generate outputs of maximum crest and wave heights, both in time and in space-time. Theoretical formulations were based on second-order nonlinear models for crest height and linear model for crest-to-trough wave heights. We relied on the existing WW3 implementation for extremes. Tests at the global scale with the extreme wave routines implemented on WAM are in progress on a research version of ECWAM.

- Comparison with observations showed that theories used for extremes provide good model performance, provided that the wave model and its forcings have sufficient accuracy to properly reproduce the total energy (say $H_s$) of the sea states.

- We observed differences between extreme wave outputs from WAM and WW3, and for the latter we documented the sensitivity of extremes on the three source term configurations ST3, ST4, and ST6. We noted that the differences mainly concern the value of $H_s$, while the variability of other parameters, such as the steepness, the bandwidth, or the directional properties of wave spectra, seems to have a smaller effect on extremes.
Main features of the maximum wave climate in the Mediterranean Sea have been highlighted. Although limited to a 10-year period (2001-2010), the analysis disclosed the magnitude of maximum waves and helped reveal the relevant process concurring in generating high waves during marine storms.

Because of its importance for marine safety, a better understanding of wave extremes is critical for a variety of applications. There is a consensus that high-quality predictions of extreme events caused by storms could substantially contribute to avoiding or minimizing human and material damages and losses. Therefore, reliable wave forecasts, together with long-term statistics of extreme conditions are of utmost importance for marine areas. This work strengthens the capability of spectral wave models to be a powerful tool for studying the intensity and geographical distribution of extreme waves. More work is needed to determine how the proposed approach apply to a variety of sea conditions that have not been focused in this study, such as the oceanic swells or the strongly forced wind-wave conditions (for instance, hurricanes and typhoons). Future measurements and numerical experiments will be valuable to further test the models’ performance.

**Appendix A: Changes in the WAM source file code (new subroutine WAMAX)**

The new WAM code that includes the WAMAX routine was tested, debugged, and transferred to the code maintainer Helmholtz-Zentrum Geesthacht (HZG; contact persons: Dr. Joanna Staneva and Dr. Arno Behrens). Main changes in the WAM source file code are described in Table A.1.

<table>
<thead>
<tr>
<th>Source File</th>
<th>Purpose</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>chief/read_wam_user.f90</td>
<td>Reads user options from input file WAM_User</td>
<td>WAMAX user options reader</td>
</tr>
</tbody>
</table>
| mod/wam_interface_module.f90 | Procedures used in WAM to compute parameters from spectra and to interpolate spectra | Added subroutine WAMAX which computes:
  - \( CMAX_F \): Time maximum crest height (Forristall)
  - \( HMAX_N \): Time maximum wave height (Naess)
  - \( CMAX_ST \): Space-Time maximum crest height (STQD)
  - \( HMAX_ST \): Space-time maximum wave (STQD)
  Given:
  - \( F \): Block of spectra
  - \( DEPTH \): Depth
  - \( THMAX \): Peak direction |
| mod/wam_output_parameter_module.f90 | Output parameter names and scales for printing | Extended output to include the four WAMAX output variables |
| mod/wam_output_set_up_module.f90 | Output times and flags | Added a subroutine SET_WAMMAX_OPTIONS |
Table A.1 - Summary of the WAM source code changes made to implement the WAMAX routine for the computation of the two extreme wave parameters: maximum crest height and crest-to-trough maximum wave height.

| mod/wam_print_user_module.f90 | Prints set up options | Modified number of output variables |
| mod/wam_user_module.f90      | Default settings      | Adapted for WAMAX user options      |
| print/wam_netcdf_module.f90  | Creates NetCDF files from binary format | Adapted to include WAMAX outputs |

Declaration of competing interest

The author declares that there is no competition financial interests or personal relationships that could be appeared to influence the work reported in this paper.

CRediT authorship contribution statement


Acknowledgments

This work has been carried out as part of the Copernicus Marine Environment Monitoring Service (CMEMS) LATEMAR project (https://www.mercator-ocean.fr/en/portfolio/latemar/). CMEMS is implemented by Mercator Ocean in the framework of a delegation agreement with the European Union. Dr. Tiziana Paccagnella and Dr. Andrea Valentini from the Regional Agency ARPAE-SIMC (Italy) are tanked for providing COSMO-I2 wind data. We are grateful to Dr. Mauro Bastianini (CNR-ISMAR) for his work on the Acqua Alta platform, and to Dr. Jean Bidlot (ECMWF) for useful discussions on wave model source terms.

References


Ardhuin, F., Rogers, E., Babanin, A., Filipot, J.-F., Magne, R., Roland, A., Van Der Westhuysen, A., Queffeleulou, P.,


ECMWF, 2019. IFS DOCUMENTATION – Cy46r1 Operational implementation 6 June 2019 PART VII: ECMWF WAVE MODEL.

ECMWF, 2012. IFS DOCUMENTATION – Cy38r1 Operational implementation 19 June 2012 PART VII: ECMWF WAVE MODEL.


doi:10.1175/JPO-D-11-0179.1

measurements of oceanic sea states. Ocean Model. 70, 103–115. doi:10.1016/j.ocemod.2013.01.001

Fedele, F., Brennan, J., Sonia Ponce de León, Dudley, J., Dias, F., 2016. Real world ocean rogue waves explained
without the modulational instability. Sci. Rep. 6, 1–11.

doi:10.1016/j.matcom.2011.05.009

Fedele, F., Herterich, J., Tayfun, A., Dias, F., 2019. Large nearshore storm waves off the Irish coast. Sci. Rep. 9,
15406. doi:10.1038/s41598-019-51706-8


doi:10.1017/S0022112008004424

Filipot, J.-F., Guimaraes, P., Leckler, F., Hortsmann, J., Carrasco, R., Leroy, E., Fady, N., Accensi, M., Prevosto,
Jument lighthouse: a real-scale laboratory for the study of giant waves and their loading on marine structures.


1931–1943.


mechanisms in different physical contexts. Phys. Rep. 528, 47–89. doi:10.1016/j.physrep.2013.03.001


Figure 02