Genoa, 20/01/2020

Dear Editor,
we would like to thank the reviewer for further reading and commenting the paper. We went through the comments and we answered in detail to all of them. An item-by-item reply follows for the revision.
The new manuscript has been modified according to the suggestions of the Reviewer. The parts modified are highlighted in red, while the new ones are highlighted in blue. The line numbers recalled in this document refer to the marked manuscript attached to the present document.

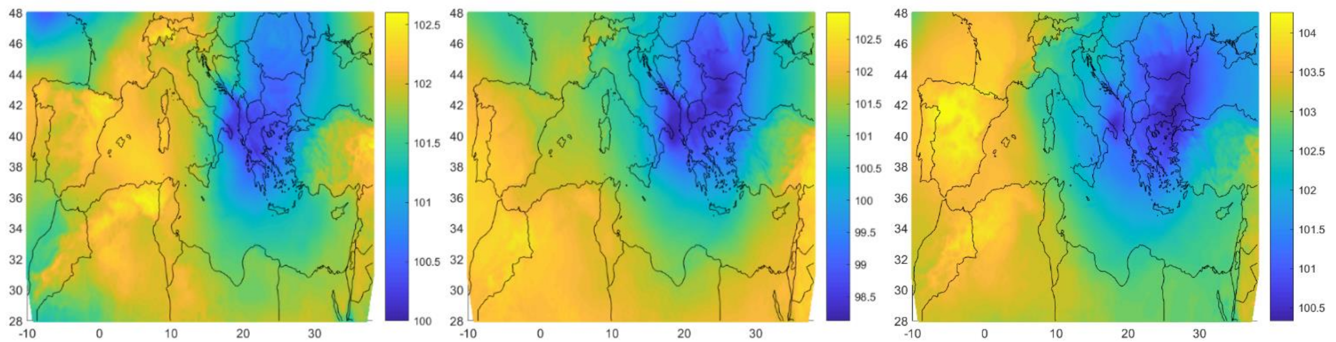
Reviewer #1

In the manuscript "Extreme waves analysis based on atmospheric patterns classification: an application along the Italian coast" the authors propose a methodology for classifying data of a physical quantity, to be applied prior to perform Extreme Value Analysis, for complying with three key requirements of the data samples: independent and identically distributed and, in addition for directional variables, grouped in homogeneous subsets. This last requirement is the principal objective of the manuscript, applied to significant wave height peaks along the Italian coast. Following previous works, they propose to use the atmospheric processes producing extreme wave conditions in a given location, (1) to select the homogeneous subset based on the weather patterns (WPs), and (2) to estimate the overall extreme values distribution starting from the distribution fitted to each subset.
The method relies on the physical connection between the atmospheric processes, spatial and temporal evolution of the surface pressure and wind fields, behind the occurrence of the extreme wave conditions at a given location. Consequently, the classification of extreme events is based on observed surface wind fields during the hours before and concomitant to the time of the peaks and on the correlation maps between the wind velocities and significant wave height peaks. Once the wind fields producing the peak wave conditions are identified the wind fields were used for clustering and classifying the extreme events. Then, the classification of the peaks, once the threshold of the wave height is chosen, depends only on the normalized wind fields.
The manuscript addresses relevant scientific and technical questions within the scope of NHESS, wave climate extreme value problem, presents new data and some novel concepts, with well-developed tools and very interesting results. Then it is worthy to be published in NHESS. However, before to recommend the manuscript for publication, there are, in my opinion, two important questions which the authors should clarify:
(a) The range of validity of the classification of extreme events, only based on surface wind fields.
(b) The "quality" of the obtained homogeneous data sets resulting from the feeding the k-means algorithm with the normalized wind fields producing the wave height peak conditions.

As regards comment a), it should be pointed out that this study focuses on waves driven by wind; therefore, other physical quantities that may concur in generating sea waves (tides, soil vibrations etc.) are not considered. In case of wind waves, the variables that may affect their characteristics, beyond the wind, are the local bathymetry (geometry of the basin), sea level variations (i.e. water depth) and currents. However, in this case the bottom depth is not affecting the waves propagation, since all the locations investigated lie in deep water (see Table 1). As for the current, this was not considered in the wave generation model; nevertheless, the hindcast data were widely validated and proved to be reliable. Therefore, wind data are reasonable expected to be sufficient for the characterization of the extreme waves in our case. This has been strengthened at line 85. As such, the key point is how to define the time and spatial domain of the wind to be considered: since the Mediterranean Sea is an enclosed basin with limited fetches, 12 hours were found to be enough; other areas may require different domains, depending on the characteristics of the basin.

As for comment b) (the “quality" of the subsets), the events belonging to different clusters are remarkably homogeneous in terms of wave bulk parameters, even though the latter were not used for clustering the peaks (this can also be appreciated with respect to Tp, as shown in the Figure at the end of this document). Moreover, in the framework of Extreme Value Analysis, data are homogeneous in term of the parent distribution, since thresholds were selected in order for the peaks to come from a Generalized Pareto Distribution by following the method proposed in Solari et al. (2017).

Finally, the homogeneity of the wave fields can be validated by looking at the mslp related to single events with respect to the overall mslp average. An example for location B7 is attached below, showing the mslp fields for the three highest events belonging to WP#2. These events were selected because they are characterized by Hs significantly higher than those of the point cloud, and are further recalled by the Reviewer in the next comment.



From the example it can be noticed how the mslp fields related to the single events are very similar between each other, and in turn to the average mslp field reported in Figure 4 (panel D) of the manuscript. The location of the low pressure of the cyclone is similar in both events, though there are differences in terms of the absolute value of the pressure and slightly in the shape of the cyclone too; something to be expected as it partially explains the different intensities of Hs (of course, local effects have to be taken into account as well).

Developed Questions

The main difference between previous published research and present work is that the variable wind wave is generated under fetch and time limited conditions. Then, the correlation between observed wind velocity and wave height fields depends on the generation process quantified through the non-dimensional variables. Significant wave height and peak period, Hs, Tp* depend on the nondimensional fetch, F* and the non-dimensional time t**

$Hs^* = gHs/U^2 = f(gF/U^2, gt/U)$ $Tp^* = gTp/U^2 = f(gF/U^2, gt/U)$ $F^* = gF/U^2$ $t^* = gt/U$

U is the mean wind velocity, at a certain height, over the fetch F. Fetch, and consequently U are defined based on the mean wind direction (well identified by the authors, figure 10).
These functional relationships between the non-dimensional quantities should be used to link the weather patterns, wind velocity and significant wave height time series. Based on that, F and t* should be relevant quantities (bring in the physics of the wave generation process) to classify the extreme events and the correlation maps for different lags. For that purpose, the procedure defined by eqs (1) -(3) should be applied, not with wind velocity but with the non-dimensional quantities, and as well as to feed the k-means algorithm for finding the homogeneous subset.*

The relationships proposed by the Reviewer can be related to the so called Significant Wave Method. This model allows to predict the wave field at a given location by looking at non-dimensional ratios of the wind fetch (Bretschneider, 1959, among others), thus quantifying the wave generation process through non-dimensional variables. Although this theory is well known and established in the scientific literature, it refers to empirical formulations that may happen to be too simplified in case of random environmental conditions and complex coastline geometries.

Furthermore, the use of non-dimensional quantities introduces an additional problem: as shown by the correlation maps (Figure 2 of the manuscript), it is not possible to detect a prevailing fetch for the extremes at punctual location, and the wind velocities may vary dramatically along different fetches and time lags. Therefore, it would be difficult to define F and U to feed the non-dimensional relations with. On the other hand, this problem does not arise when raw wind data are used. Moreover, the wave model used to build the hindcast was fed with high-resolution wind data, allowing to describe in great detail the wave generation physic. As such, is Authors' belief that hindcast wind velocity data over the whole Mediterranean Sea provides the best information to be plugged into the k-means algorithm for classifying the Hs peaks.

As regard the homogeneity of the subsets, Reviewer is referred to the example reported in the previous comment: even though the highest two data points depart of approximately 15% from the third largest value of Hs, the mslp fields related to such events are very similar to the one corresponding to the third largest event (and to the average mslp field in turn, reported in Figure, 4 panel D).

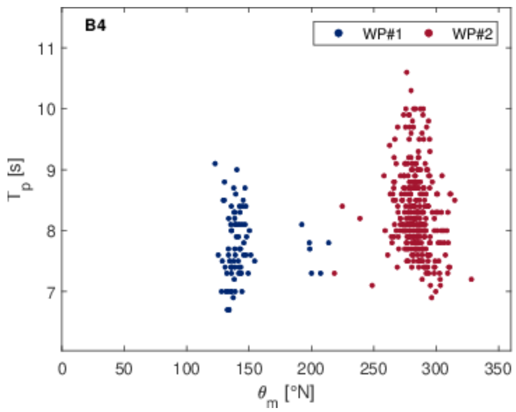
In addition, but very important issue, the non-dimensional quantities for each direction will help to define asymptotic values of the extreme distributions, as seems to occur figure 11. In location B4, it seems that there is an upper limit around Hs=8 m. This trend can be checked working with the non-dimensional quantities for bounded values of F and t for that direction. Similarly, for location B7, the two highest data points (Hs > 6m) depart approximately a 15% from the third largest value of Hs. Belong to the same subset of data? Please use the non-dimensional quantities for checking the homogeneity of the subsets.

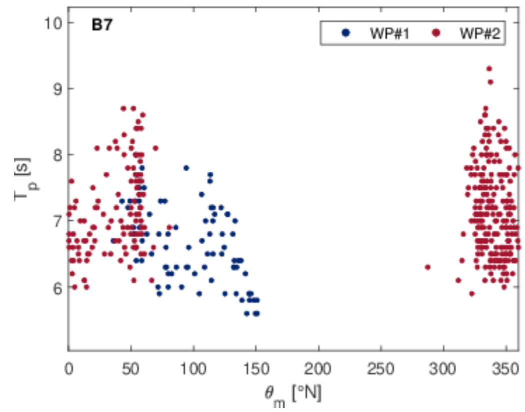
Regarding the definition of asymptotic values, we do agree in that some information could be extracted from the analysis of the dimensionless variables, but the task is not as straightforward as suggested by the Reviewer and, in our opinion, the value of the information obtained this way would not be exhaustive. As an example, we took buoy B7, as suggested by the Reviewer. For this buoy most extreme waves are associated with WP#2 (see Figure 10 in the manuscript) and come from WNW (approx. 290º). Dimensional fetch in this case is approx. 820 km (limited by the Balearic Islands); if wind speed of 22 m/s (number 9 in Beaufort scale) is considered and dimensionless relations between F and H are used as presented by Holthuijsen (2007), a value of Hs of approx. 9.9 m would be obtained. This is significantly larger than the observed values and is similar to the upper limit of the confidence interval shown in figure 11 of the manuscript.

As informative as this would be in terms of understanding the conditions generating an extreme event, there are many uncertainties in this kind of approximation that in our opinion prevent its use as an upper bound for the extreme distribution of Hs, namely: what wind speed should be considered? does it blow along the whole available fetch? and for how long? All these variables would significantly affect the computation of the Hs upper bound, and their definition would imply a high degree of subjectivity.

Finally, while working with the wave height and neglecting the wave period, any risk analysis on the coastline or relevant maritime structures would be not complete. By using non-dimensional quantities, the values Hs and Tp are computed simultaneously because they depend on F and t*. Please, if possible include the peak period to complete the necessary information for developing a risk analysis.*

We agree with the Reviewer that Tp is essential for developing risk analysis. Actually, classifying the wave fields according to the wind speed lead to clusters of Tp consistent with those of Hs, as the latter parameter is closely tied to the former one (especially in case of extreme sea states). This can be appreciated in the figure below, which shows the classification of Tp peaks according to θm for B4 and B7, similarly to what has been done in Figure 10 of the manuscript.





Once the series of Hs and Tp are defined, Extreme Value Analysis can be performed on the joint distribution of the two parameters (as an instance following Haver, 2015). This has been commented at line 301 of the manuscript.

REFERENCES

Bretschneider, C. L. (1959). *Wave variability and wave spectra for wind-generated gravity waves* (No. TM-118). CORPS OF ENGINEERS WASHINGTON DC BEACH EROSION BOARD.

Solari, S., Egüen, M., Polo, M. J., & Losada, M. A. (2017). Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value. *Water Resources Research*, *53*(4), 2833-2849.

Haver, S., & Winterstein, S. R. (2009). Environmental contour lines: A method for estimating long term extremes by a short term analysis. Transactions of the Society of Naval Architects and Marine Engineers, 116, 116-127.

Holthuijsen, L. (2007) WAVES IN OCEANIC AND COASTAL WATERS. Cambridge University Press.

Extreme waves analysis based on atmospheric patterns classification: an application along the Italian coast

Francesco De Leo[1], Sebastián Solari[2], and Giovanni Besio[1]

[1]Dept. of Civil, Chemical and Environmental Engineering - University of Genoa, Genoa, 16145, Italy
[2]Instituto de Mecánica de los Fluidos e Ingenieria Ambiental - Universidad de la República, Montevideo, 11300, Uruguay

Correspondence: Francesco De Leo (francesco.deleo@edu.unige.it)

Abstract.

The identification of homogeneous populations of data prior to perform Extreme Value Analysis (EVA) is advisable in all fields of sciences. When performing EVA on sea storms, it is also recommended to have an insight on the atmospheric processes behind the occurrence of the extremes, as this might facilitate the interpretation and ultimately use of the results. In this work, a "bottom-up" approach for the identification and classification of the atmospheric processes producing extreme wave conditions is revisited, and applied to several locations among the Italian buoy network. A methodology is given for classifying samples of significant wave height peaks in homogeneous subsets, and for the computation of the overall extreme values distribution starting from the distributions fitted to each single subset. From the obtained results, it is concluded that the proposed methodology is capable of identifying clearly differentiated subsets driven by homogeneous atmospheric processes, and it allows to estimate high return-period quantiles consistent with those resulting from the usual EVA. Two well-known cyclonic systems are identified as most likely responsible of the extreme conditions detected in the investigated locations. These systems are analysed in the context of the Mediterranean sea atmospheric climatology, and compared with those figured out by previous researches developed in similar frameworks.

1 Introduction

The extreme value theory is widely used for the analysis of extreme data in most of the geophysical applications. It allows to estimate extreme (un-observed) values, starting from available records or modelled data which are assumed to be independent and identically distributed (Coles, 2001). It is therefore crucial to identify homogeneous datasets complying with the above mentioned requirement before performing the EVA of a given physical quantity.

When dealing with directional variables, it is common to group the data according to different directional sectors (Cook and Miller, 1999; Forristall, 2004), being such approach recommended as well in many regulations (API, 2002; ISO, 2005; DNV, 2010, among others). However, the use of directional sectors involves certain drawbacks. First, it cannot be employed for variables being not characterized by incoming directions (such as storm surge or rainfall). Second, data showing the same direction may be due to different forcing; in the frame of wave climate, an example is that of waves propagating in shallow waters, affected by refraction and/or diffraction. Finally, the borders of the directional sectors are often subjectively defined, without verifying if the data belonging to each subset are homogeneous and independent with respect to those of the other sectors (see Folgueras et al., 2019, where they tackled this issue and proposed a methodology to overcome it).

An alternative approach to classify the extremes implies resorting to the atmospheric circulation conditions they are driven by, associating each extreme event to a particular weather pattern (referred to as WP). Such approach has been already deep-seated in atmospheric sciences for the analysis of precipitations, snowfalls, temperature, air quality and winds (Yarnal et al., 2001; Huth et al., 2008, among others). Nevertheless, there are few studies linking weather circulation patterns with the most likely induced sea states (e.g., wave climate and storm surge). Holt (1999) classified WPs leading to extreme storm surges in the Irish Sea and the North Sea. Guanche et al. (2013) simulated multivariate hourly sea state time series in a location in the northwester Spanish coast, starting from the simulation of weather pattern time series. Dangendorf et al. (2013) linked the atmospheric pressure fields with the sea level in the German Bight (southeastern North Sea). Pringle et al. (2014, 2015) investigated how extreme wave events may be tied to synoptic-scale circulation patterns in the east coast of South Africa. Camus et al. (2014) proposed a statistical downscaling of sea states based on weather types, then applied to a couple of locations in the Atlantic coast of Europe to hindcast the wave climate during the twentieth century plus modelling it under different climate change scenarios. The latter methodology was improved by Camus et al. (2016), and further used by Rueda et al. (2016) for analysing significant wave height maxima. Solari and Alonso (2017) used WP classification to perform EVA of significant wave heights in the south-east coast of South America.

Except for Rueda et al. (2016) and Solari and Alonso (2017), none of the previous works focused on exploiting WP classification methodologies for defining homogeneous datasets to be further employed for EVA. However, the two methodologies differ in several aspects. Rueda et al. (2016) dealt with daily maxima significant wave heights along with surface pressure fields and pressure gradients, averaging over different time periods and applying a regression-guided classification to define 100 WPs. They subsequently fit a Generalized Extreme Value distribution, estimating an Extremal Index from the daily maxima significant wave height of each WP, from which they rebuilt the overall distribution of annual maxima (referred to as AM H_s). Despite the proposed methodology is able to reproduce the AM H_s distribution, with such a large number of WPs it may be difficult to detect the most relevant physical processes behind the occurrence of extreme wave conditions. Furthermore, as shown in Rueda et al. (2016), even though a large number of WPs was considered, only a few happened to significantly affect the EVA, as most of the WPs resulted to be associated to mild wave conditions. Finally, to retain daily maxima does not ensure the data to be independent, thus implying the need to use the Extremal Index. Solari and Alonso (2017) introduced instead a “bottom-up" scheme: they first selected a series of independent extreme sea states; then, they identified a reduced number of WPs that allows to group the selected data into homogeneous populations. A small number of WPs makes it easier to link the different subsets of extremes with known climate forcing. Above all, to work with independent peaks allows to rely on the classic and well known extreme value theory, with no need to refer to additional indexes and/or more complex models that may be unfamiliar for many analysts.

In this paper, the methodology of Solari and Alonso (2017) is revisited and applied to several wave datasets along the Italian coastline. The objective of this research is twofold: (i) to explore how the definition of homogeneous subsets, based on WP, affects the estimation of H_s extreme values; (ii) to characterize the identified WPs in the framework of the Mediterranean Region (MR) cyclones climatology.

The paper is structured as follows: in Sect. 2 we introduce the data and describe the methodology developed; results are presented and discussed in Sect. 3; finally, in Sect. 4 conclusions are summarized and further developments are introduced.

2 Data & Methods

2.1 Wave and atmospheric data

This work takes advantage of eight hindcast points located in the Italian seas, as shown in Fig.1. This choice allowed to test the reliability of the proposed methodology under different local wave climates. In fact, the selected points are differently located along the Italian coastline, and, being exposed to different fetches, they are characterized by peculiar wave conditions. The same locations were taken into account by Sartini et al. (2015), where they performed an overall assessment of the different frequency of occurrence of the extreme waves affecting the Italian coasts. Table 1 reports the names, depths, and coordinates of the selected locations.

The points correspond to as many buoys belonging to the Italian Data Buoy Network (Rete Ondametrica Nazionale or "RON", Bencivenga et al., 2012), which collected directional wave parameters over different periods between 1989 and 2012. Unfortunately, most of the buoys are characterized by significant lacks of data due to malfunctions and maintenances of the devices. We therefore referred to hindcast data, since such a widespread lack of data would imply a loss of reliability for the following analysis. We relied to the hindcast of the Department of Civil, Chemical and Environmental Engineering of the University of Genoa (http://www3.dicca.unige.it/meteocean/hindcast.html; Mentaschi et al., 2013, 2015)~~, providing~~ It now provides wave parameters on a hourly base from 1979 to 2018 over the whole Mediterranean sea, with a spatial resolution of 0.1° in both longitude and latitude (however, at the time the study was developed the series were defined up to 2016). Data were validated against the records of the buoys (when available); more details can be found in Mentaschi et al. (2013, 2015). The wind data used to drive the wave generation model were derived from the NCEP Climate Forecast System Reanalysis for the period from January 1979 to December 2010 and CFSv2 for the period from January 2011 to December 2018, downscaled over the MR at the same resolution of the hindcast, along with the pressure fields through the model Weather Research and Forecast (WRF-ARW version 3.3.1, see Skamarock, 2009; Cassola et al., 2015, 2016). These wind velocities data were used here to feed the cluster analysis of the wave peaks~~, while the pressure data were used for the analysis of the climatology related to the identified WPs (as explained in Sect.~~ 2.3 ~~and Sect.~~ 2.4~~, respectively).~~ It should be pointed out that, in case of sea waves, other variables may concur to affect their propagation (and therefore the bulk parameters), i.e. the local bathymetry and the currents. However, the bottom depth is reasonably expected not to be relevant, since all the locations investigated lie in deep water (see. Table 1), while the currents were not fed into the wave model, though the hindcast data were widely validated and proved to be reliable.

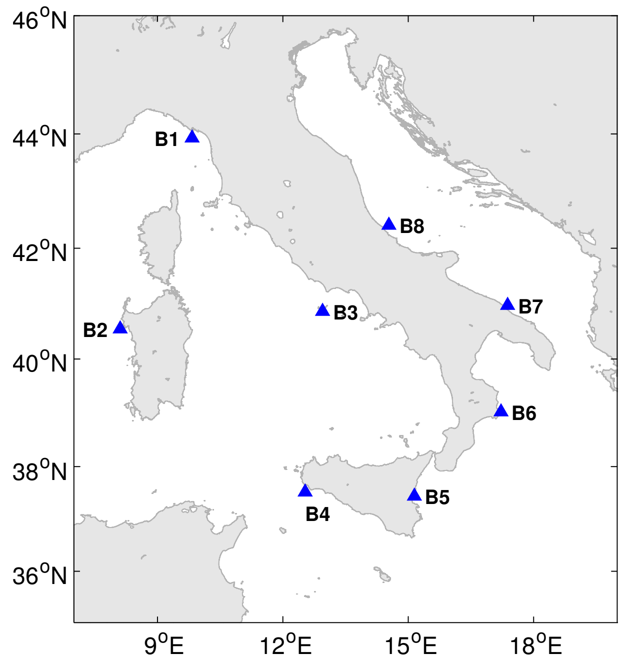


Figure 1. Study area and investigated locations with their respective codes

CODE	LON	LAT	DEPTH [m]	NAME
B1	9.8278	43.9292	83.8	La Spezia
B2	8.1069	40.5486	99.7	Alghero
B3	12.9500	40.8667	242.0	Ponza
B4	12.5333	37.5181	90.8	Mazara del Vallo
B5	15.1467	37.4400	65.4	Catania
B6	17.2200	39.0236	611.7	Crotone
B7	17.3778	40.9750	80.0	Monopoli
B8	14.5367	42.4067	55.8	Ortona

Table 1. Lon/lat coordinates and depths of the hindcast locations employed in the study (reference system: WGS84)

2.2 Extreme events selection

For each location, wave height peaks were selected through a Peak Over Threshold (POT) approach, and in particular by using a time moving window. This approach works as follows. First, the whole series of H_s is spanned through a time moving window of given width; second, when the maximum of the data within the window happens to fall in the middle of the window itself, it is retained as a peak; finally, in order to get rid of the peaks which are not related to severe sea states, a first H_s threshold is chosen and only peaks exceeding this threshold are retained for further analysis.

In this study, for each location the width of the moving window was set equal to one day, meaning that the inter-arrival time between two successive storms is at least equal to one day. The threshold was fixed as the 95^{th} percentile of the resultants peaks. This ensured to maintain a uniform approach for all the locations, efficiently capturing the different features of the local wave climates. Beside the significant wave heights, we retained as well the waves mean incoming directions corresponding to the peaks (θ_m), which were used for analysing the outcomes of the clustering algorithm. Finally, for each peak we extracted the mean sea level pressure field (MSLP) and surface wind fields for several time lags (0, 6, 12 24, 36, and 48 hours earlier with respect to the peak's date), over the whole MR. Wind fields were used to classify the selected peaks due to their parent WP, as described in the following section; MSLP fields were used instead for the post-processing and climatological analysis of the results.

2.3 Extreme events classification: definition of weather patterns

The classification of extreme events is based on surface wind fields ($\bar{u}_w$) observed in the whole MR during the hours before and concomitant to the time of the peaks. In order to define the spatial and temporal domains to be taken into account, we looked at the correlation maps between the wind velocities and the H_s peaks for different time lags. Correlations were evaluated over a sub-grid of the atmospheric hindcast, with nodes spaced of 0.5° both in longitude and latitude. To compute the correlation between H_s and $\bar{u}_w$ series is not straightforward, as the former variable is scalar and the latter is directional. To tackle this issue, we followed the procedure suggested by Solari and Alonso (2017). Given a time lag Δt, the wind at every node (i, j) is defined by its zonal and meridional components $(u_x, u_y)_{(i,j,\Delta t)}$; the correlation between H_s peak series and the time lagged surface wind speed at any given node is then estimated as the maximum of the linear correlations obtained by projecting the wind speed series in all the possible directions:

$$\rho_{(i,j,\Delta t)} = \underset{0 \leq \theta < 2\pi}{max} \{\rho\left(H_s; \bar{u}_{(i,j,\Delta t,\theta)}\right)\} \tag{1}$$

where $\rho_{i,j,\Delta t}$ is the resulting correlation for node (i, j) at time lag Δt, ρ refers to linear correlation function, $u_{(i,j,\Delta t,\theta)}$ is the surface wind speed projected along direction θ according to Eq (2):

$$\bar{u}_{(i,j,\Delta t,\theta)} = u_{x(i,j,\Delta t)} cos(\theta) + u_{y(i,j,\Delta t)} sin(\theta) \tag{2}$$

in this way not only a maximal correlation is obtained for every node, but also the direction corresponding to the maximal correlation, estimated as:

$$\hat{\theta}_{\rho(i,j,\Delta t)} = \underset{0 \leq \theta < 2\pi}{\text{argmax}} \{\rho\left(H_s; \bar{u}_{(i,j,\Delta t,\theta)}\right)\} \tag{3}$$

The correlation maps computed with Eq. (1), allowed to evaluate the spatial domain and the time lags for which $\bar{u}_w$ is significantly correlated to (i.e., directly affecting) the resulting wave peaks at a given location.

Once the spatial and time domain of the wind fields producing the peak wave conditions were defined, the wind fields were used for clustering and classifying the extreme events. To this end, the *k-means* algorithm was used, fed with the normalized wind fields. *k-means* is aimed at partitioning a N-dimensional population into k sets (clusters) on the basis of a sample, in order to minimize the intra-cluster variance (MacQueen et al., 1967). The normalization of the wind fields sought to reduce the influence of the intensity of the wind speed on the classification, so that only the spatial form of the field and its time evolution were taken into account. Note that the values of H_s did not play any role in the classification of the peaks but the identification of the point in time of the wind fields.

2.4 Analysis of the WPs climatology

Once the wind fields, and therefore peak H_s series, were grouped into k clusters, the MSLP corresponding to the events within each cluster were averaged and the position of the lowest pressure was recorded, for all the Δt taken into account. This allowed us to track the paths of the averaged low pressure systems corresponding to each cluster, defining in turn the respective WP.

At a second time, the dynamics of the systems were compared with those of the cyclones typically detected in the Mediterranean sea (Trigo et al., 1999; Lionello et al., 2016), while the frequency of occurrence of the events of different clusters were compared with the outcomes of Sartini et al. (2015). The number of clusters needed to group the series into was defined for every location by looking at the outcome of the cluster analysis: when an increase from k to $k+1$ clusters did not further lead to a new clearly differentiated WP, the research was stopped and k was used for the cluster analysis of that particular location.

2.5 Extreme value analysis

The EVA were performed independently over the subsets of H_s peaks resulting from the cluster classification. We followed the methodology proposed by Solari et al. (2017), where the threshold for the POT analysis was estimated as the one maximizing the p_{value} of the upper tail Anderson-Darling test. Then, once the threshold is estimated (i.e. a subset of the peaks within a given WP is defined), the three parameters of the GPD are estimated through the L-moments method, and the return-period (T_r) quantiles of the variable under investigation are computed. In this work, in order to estimate the overall T_r-H_s curve and its confidence intervals from the GPDs fitted to each WP, a bootstrapping approach was implemented (see Algorithm 1). Algorithm 1 shows a pseudocode summarizing the bootstrapping procedure. First, N_{boot} series of H_s, each N_{years} long, are generated for every WP. Second, the series generated for the different WPs (N_{WP} per location) are combined in order to obtain one single N_{years} long series for each of the N_{boot} simulations. Third, an empirical relation between T_r and H_s (i.e. an empirical cumulative distribution function or ECDF) is estimated from each one of the N_{boot} series. Lastly, expected value and confidence intervals of H_s are estimated from the N_{boot} ECDFs for several return periods.

The method assumes a Poisson-GPD model for each WP and that the realizations of different WP are independent from each other. This independence hypothesis was evaluated by estimating the correlation between the annual number of peaks associated to each WP.

Algorithm 1 Computational scheme of empirical extreme values curves and its confidence intervals

for $j \leftarrow 1, N_{boot}$ **do**
  for $i \leftarrow 1, N_{WP}$ **do**
    $\{H_s\}_{i,0} \leftarrow random(GPD, \hat{\theta}_{(i,0)}, N_i)$
    $\hat{\theta}_i \leftarrow fit(GPD, \{H_s\}_{i,0})$
    $\{N_{simu}\}_i \leftarrow f_i(\hat{\lambda}_i, N_{years})$
    $\{H_s\}_i \leftarrow random(GPD, \hat{\theta}_i, \{N_{simu}\}_i)$
  end for
  $\{H_s\}_j \leftarrow [\{H_s\}_i, i = 1 \ldots N_{WP}]$
  $H_s \leftarrow sort(\{H_s\}_j)$
end for
$H_s = \hat{f}(T_r) \leftarrow \{f_i\}$
$C.I. \leftarrow \{f_i\}$
Comments:
N_{boot} is the number of bootstrapping repetitions
N_{WP} is the number of WP
$\hat{\theta}_{(i,0)}$ are the parameters of the GPD estimated from the original sample
N_i is the length of the original sample of H_s peaks within the WP
$\hat{\lambda}_i$ is the yearly number of events of the ith WP
N_{years} is the number of years simulated; it must be larger that the maximum return period to be analized
N_{simu} is the numer of events in N_{years} obtained for the ith WP
$H_s = f(T_r)$ is the empirical Hs-Tr curve

For a given location, the overall work-flow can be summarized as follows:

- selection of a series of H_s peaks through a POT approach
- selection of the wind field data to be employed in the clustering algorithm (i.e. Δt and spatial domain)
- classification of the H_s peaks due to the $k - means$ algorithm
- definition of a suitable number (k) of WPs
- averaging of the MSLP corresponding to the peaks of each WP, for each Δt taken into account
- performing EVA over the single subsets
- computation of the overall long term distribution through a bootstrapping technique

This methodology was applied to all the hindcast locations shown in Fig. 1. In this paper, for the sake of clarity, just the results of the locations B4 (Mazara del Vallo) and B7 (Monopoli) are shown and discussed. Indeed, just one WP was necessary for classifying the peaks in La Spezia and Alghero (B1 and B2), thus no further analysis were performed. Among the locations left, we simply selected the locations furthest from each other in the East-West direction. All the results related to the other locations can be found in the supplementary material.

3 Results and discussions

Once the series of H_s peaks is selected, the first step of the proposed methodology requires to define the domains of $\bar{u}_w$ in time and space due to the outcomes of the correlation analysis between the two parameters. Figure 2 shows the correlation maps for B7 for different time lags, along with the directions leading to the maximum values of correlation in each node. It is interesting to see how $\hat{\theta}$ for Δt=0 hours are distributed along the nodes characterized by similar values of ρ. Actually, even though the values of $\hat{\theta}$ come from a purely statistical analysis (i.e., they were computed with Eq. (3)), their spatial distribution follow that of a typical cyclone. Velocities happen to be uniformly oriented along the nodes characterized by the higher values of ρ, close-to-tangential to a circle centred on the nodes showing instead lower values of ρ. This allows to get a first insight on the predominant process most likely affecting the wave climates of the investigated locations, as it will be discussed further on this paper. On the contrary, the analysis of the correlations between H_s and $\bar{u}_w$ reveals that the areas characterized by similar values of ρ are not uniformly distributed in the neighbourhood of the points considered. It is therefore difficult to uniquely contour the nodes to be taken into account for the successive analysis. As regards the time step, correlations rapidly decrease for Δt longer than 12 hours, after which no evidence of significant ρ can be observed. The same outcomes apply for all the other locations taken into account. Results suggest that the events shall be linked with broader circulation patterns. In view of the above, it was decided to refer to the whole MR and time lags of 0 and 12 hours for the purposes of peaks classification.

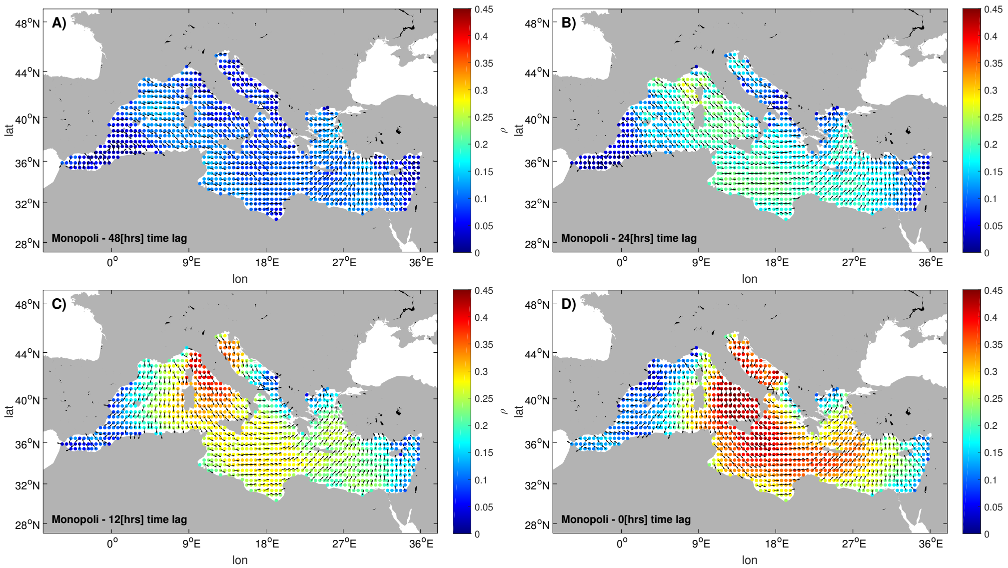


Figure 2. Correlations between H_s and $\bar{u}_w$ in location B4 for different time lags. Panel A): Δt equals 48 hours; panel B): Δt equals 24 hours; panel C): Δt equals 12 hours; panel D): Δt equals 0 hours

Once the spatial and temporal limits of the wind fields to be used in the $k-means$ were defined, we performed a sensitivity analysis over the resultant average MSLP belonging to each cluster, among the total amount of tested clusters (k). Two clusters were needed to detect different systems at all the sites but Alghero and La Spezia, where ~~a single pattern was enough to describe~~ the local extreme waves could be related to a single pattern. For the other locations, to increase k did not lead to systems significantly diverging from those already defined. The reason for such a small number of resulting WPs has to be found in the nature of the H_s employed in the analysis. Indeed, these are associated to extreme sea states, which are most likely driven by atmospheric phenomena developing along well-defined and fixed tracks.

Figures 3 and 4 show the averaged pressure fields corresponding to both the clusters and Δt in B4 and B7 hindcast points. Here, two main systems can be clearly distinguished: i) a low moving SW-NE towards the Balkan area (say WP#1) and ii) a low moving NW-SE (henceforth referred to as WP#2).

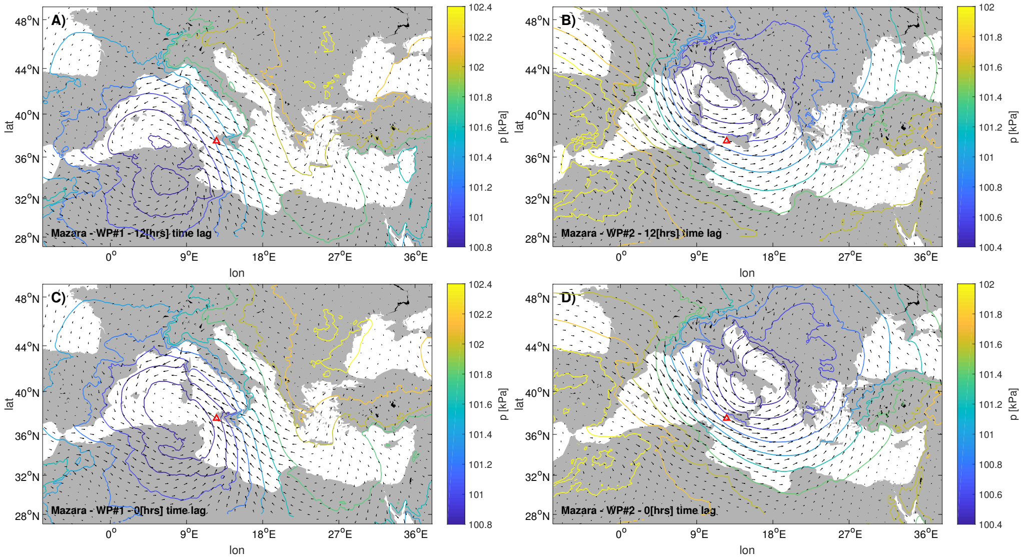


Figure 3. Average MSLP for the H_s peaks in Mazara del Vallo (B4). Panel A): WP#1, Δt equals 12 hours; panel B) WP#2, Δt equals 12 hours; panel C): WP#1, Δt equals 0 hours; panel D): WP#2, Δt equals 0 hours

Let us first focus on WP#2. Low pressure moves SE from the central Europe, crossing the Adriatic sea and decreasing its intensity once it gets to the south Balkan area, where it stops until it finally dissolves. As regards WP#1, the cyclogenesis most likely takes place in the east area of north Africa, with cyclones first approaching the west coastline of Italy while moving NE. The paths of WP#1 and WP#2 show evident similarities with well known cyclones typically forming and departing from two of the most active cyclogenetic regions in the MR, respectively the lee of the Atlas mountains and the lee of the Alps (Trigo et al., 1999). Actually, WP#2 seems to follow the Genoa system path that usually moves down to the Albanian and Greek coasts, while that of WP#1 ~~is characteristic of~~ shows characteristics similar to the Sharav depression, moving north-eastward toward the Greek region (Flocas and Giles, 1991; Trigo et al., 1999, 2002). The same paths characterize the lows of the WPs detected in all the investigated sites (see the supplementary material). A summary of the lows position at the two different Δt can be appreciated in Fig 5, reporting the tracks of both the WP#2 and WP#1 systems in all the sites but B1 and B2; indeed, in the latter ~~case, no results for B1 and B2 are available, since all the peaks happened to belong to a single WP.~~ cases the analysis of the MSLP fields did not allow to identify two separated systems.

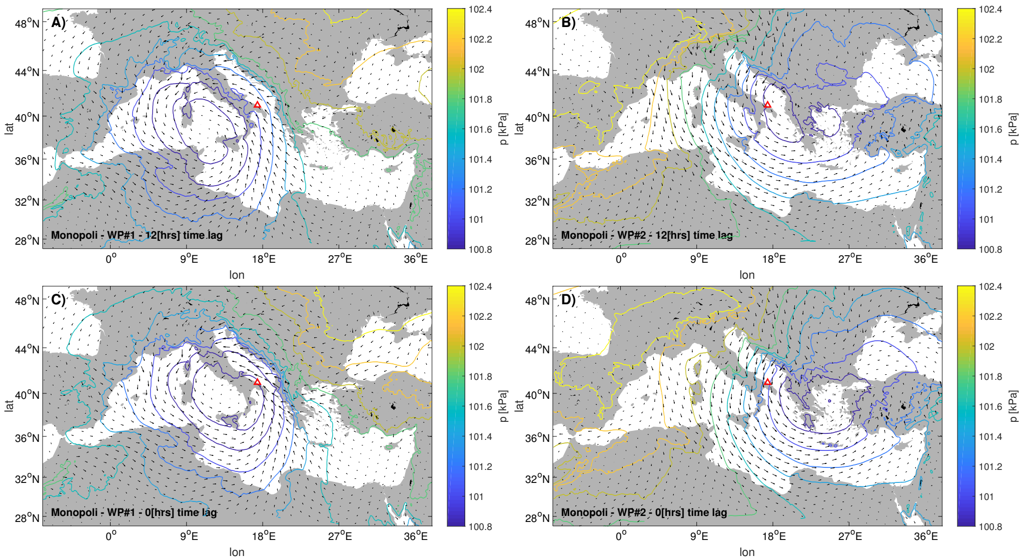


Figure 4. Average MSLP for the H_s peaks in Monopoli (B7). Panel A): WP#1, Δt equals 12 hours; panel B) WP#2, Δt equals 12 hours; panel C): WP#1, Δt equals 0 hours; panel D): WP#2, Δt equals 0 hours

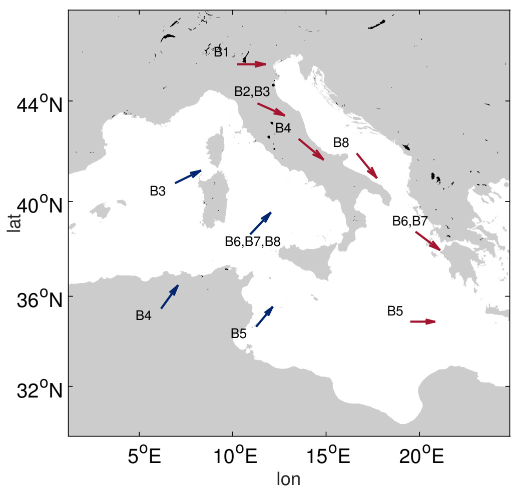


Figure 5. Time evolution of the center of the low pressure for the different WP identified for each buoy. In blue WPs with lows traveling northeastward (WP#1); in red WPs with lows traveling east-or southeastward (WP#2)

It is interesting to see how the WP#2 low get across the investigated sites in a precise chronological order, crossing first the northernmost locations and then those next to the south Balkan area where the cyclone actually ends its run. As such, we took as a reference four buoys affected by WP#2 at different times, evaluating the time lag between their respective peaks generated by such system. We considered the extreme series of B4, computing for each storm the time lag between peak at the buoy and peaks occurring in B2 and B1 (occurring earlier) and B6 (occurring later); distributions of the time lags are shown in Fig. 6.

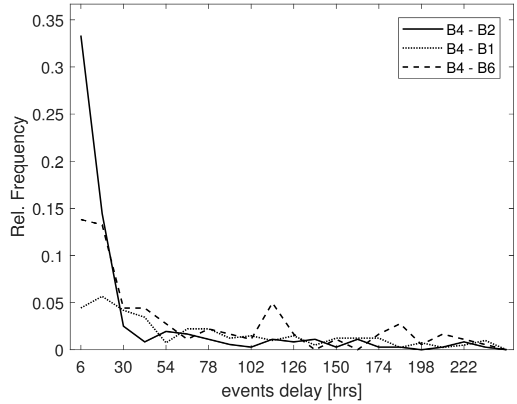


Figure 6. Relative frequency of the events' delay for B2, B1 and B6. Referring series is that of B4

Looking at Fig. 6, it is evident how the majority of the events' delays between B4 and the other investigated points fall within about forty hours. We therefore evaluated the average MSLP fields for time lags up to 48 hours before and after the storms occurring at the reference location, tracking the low pressure centre evolution. Results show how the cyclone runs out in a couple of days, the center of the low travelling for approximately 1600 kilometres, with a resulting speed of $\sim$33 km h^{-1} (see Fig. 7). Both the lifetime of the identified cyclone and the speed it moves at are compatible with the features of the cyclones most frequently encountered in the MR (Lionello et al., 2016). Unfortunately, as regards WP#1, an equally clear path cannot be detected, ~~as~~ since the average lows apparently arise simultaneously in most of the points taken into account. Indeed, the MSLPs related to this pattern show a higher variability with respect to those characterizing the events of WP#2, thus they would require further deepening and more detailed investigations. ~~We therefore~~ Therefore, we did not ~~further~~ characterize the mean evolution of the MSLP fields ~~characterizing~~ related to the events of WP#1.

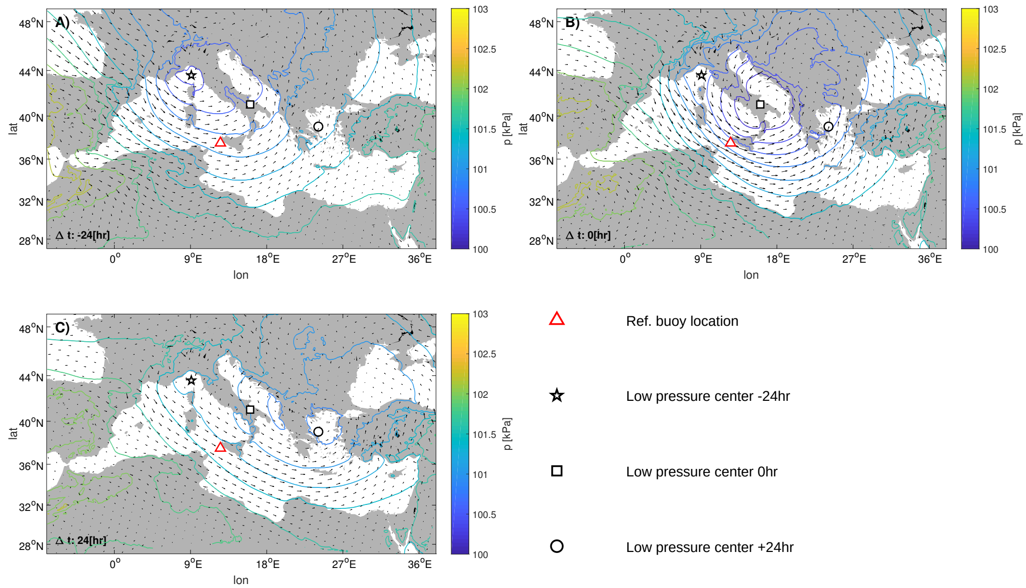


Figure 7. WP#2: average MSLP evolution with respect to the reference dates of the events in B4 (underlined with the red triangle)

The characterization of the two systems is reflected in the frequency of occurrence of the storms, with peaks belonging to different subsets showing as well distinctive seasonality. From the results shown in Fig. 8, it can be noticed how WP#2 peaks mainly occur in winter, whereas the events of WP#1 are characterized by two milder intra-annual peaks of occurrence, spread among the spring and autumn months. The intra-annual cycle of the WP#1 events further suggest a direct link with the Sharav cyclones, which show similar seasonal fluctuations; as regards the WP#2 peaks, even though the storms of the Genoa low are more uniformly distributed along the year, the most intense events precisely occur during winter, as it happens in the above mentioned locations (details can be found in Lionello et al., 2016).

Figure 9 summarizes the results of the monthly frequency of occurrence of the extremes in all the investigated locations, grouped according the WP. For each location, frequencies were normalized in the 0-1 space with respect to the total amount of peaks, in order to be able to compare outcomes defined over different ranges.

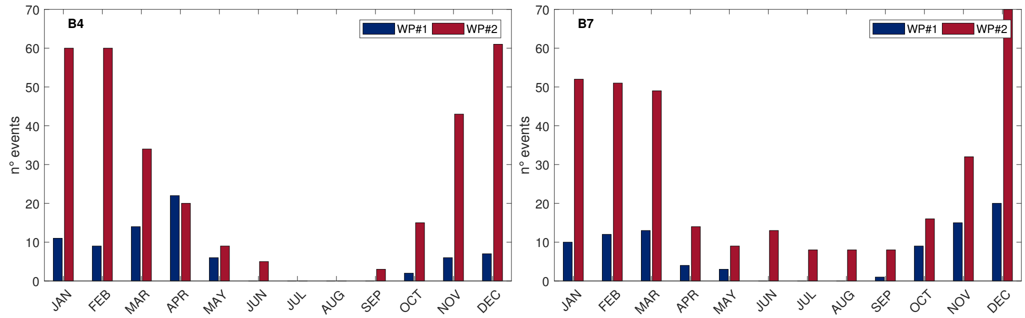


Figure 8. Monthly number of events for different WP. Left panel: B4; right panel: B7

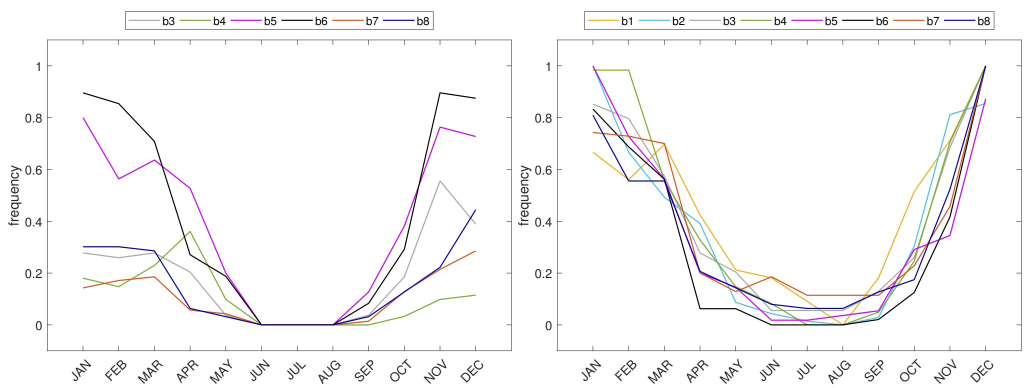


Figure 9. Normalized monthly frequency of occurrence of the extremes. Left panel: events belonging to WP#1; right panel: events belonging to WP#2.

It can be noticed how the relative weight of WP#1 events on the overall peaks distributions increases moving south. Points in the northernmost locations (B1 and B2) apparently are not influenced by the WP#1 system, and their extreme waves are actually linkable just to WP#2. Moving to the southernmost locations (B5 and B6), no significant differences in the seasonality of the two patterns' events can be appreciated; in these two buoys none of the two systems shows a prevailing frequency of occurrence for the induced storms.

The aforementioned behavior may be justified by looking at the location of the investigated points (see Fig. 1): lows moving north-eastward directly run over B5 and B6, marginally affect B7, B8, B4 and B3, whereas they do not interest at all B1 and B2. However, position of the buoys with respect to the cyclones' paths is not the only relevant variable. Indeed, local bathymetry and prevailing fetch characteristics may result in storms having, on average, unique characteristics, e.g. ~~Indeed,~~ it may happen that peaks related to the same WP show distinct frequencies of occurrence, for instance the events related to WP#1 in B6 and B7. In this latter case, the predominant parameter seems to be the fetch length, which for B7 is very limited with regards to the NE incoming waves (precisely related to the first weather circulation pattern).

The WPs that are defined in the present study show common characteristics with those qualitatively identified by Sartini et al. (2015) in a seasonal variability analysis of extreme sea waves. In particular, the same WP was identified in B1 and B2, linking the extremes with the Gulf of Genoa cyclogenesis WP type. As Sartini et al. (2015) noted, even if cyclones in the Genoa Gulf are a constant feature over the whole year, those connected at the extreme events are the ones characterized by the lowest value in MSLP. These findings (i.e., higher values during the winter rather than in spring and summer time) were observed in the southern Thyrrenian Sea as well (for instance in B4) and in the Central Thyrrenian Sea (B3). In the latter case, we found a more marked seasonal variation of the extremes, as the two resultant WPs are well separated between winter and autumn. Intra-annual variability was observed also for B7, while the analysis carried out by Sartini et al. (2015) did not reveal this kind of behaviour; analogously, B5 and B8 buoy revealed the presence of two distinct WPs, while the previous analysis identified just one cyclogenesis system. These differences may be justified in the first place by the different peaks selection (a moving window for this study, while Sartini et al., 2015, use a partial duration series approach). Moreover, evaluation of the seasonality follows completely different algorithms: the present study directly applies a clustering technique over the selected peaks, while the former analysis used the peaks to model a time dependent distribution, further characterizing their seasonality on the basis of the values of the distribution coefficient aroused from the fitting procedure.

Another interesting outcome regards the characteristics of the extremes differentiated due to the parent clusters following the $k-means$ algorithm. Fig. 10 shows the covariates H_s and θ_m of the selected peaks belonging to each WP.

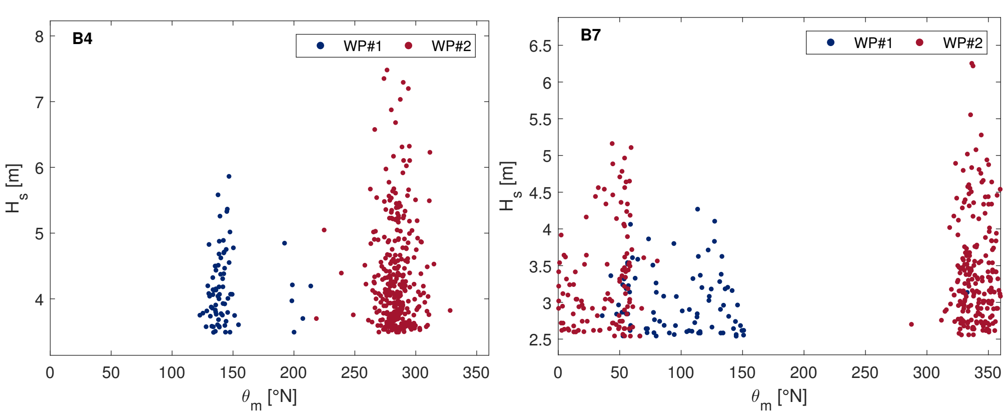


Figure 10. Scatter plot of H_s and θ_m due to different WP. Left panel: B4; right panel: B7

Looking at the scatter plots, it is evident how the proposed methodology allows to differentiate the sea storms according to the directional frames of the local wave climate. This particularly applies for locations characterized by a bimodal distribution with respect to the waves' incident direction (like B4, where is clear the presence of two peaks corresponding to SE and NW).

On the other hand, when the waves' direction is more uniformly scattered over a given sector, it is not unusual that different climate forcing result in the same wave direction; this can be noticed in B7, where θ_m-H_s scatters belonging to different patterns are partially overlapped, thus a directional classification is not straightforward and most probable would not be able to differentiate the storms due the cyclones generating them.

Finally, Fig. 11 shows the T_r-H_s curves, comparing the results obtained directly from the whole set of peaks with those obtained from the single-WP distributions, combined by means of Algorithm 1 (i.e., omni-WP curves).

The omni-WP curves show a striking agreement with those carried out through the analysis of the whole dataset without WP classification; in both the locations B4 and B7 it can be even appreciated a narrowing of the confidence intervals, meaning a reduction in the total variance for the long-term estimates. Actually, the curve related to B4 show a slight deviation between the two approaches ($\simeq$ 30 cm over the 200 year wave); however, such small magnitudes imply relative errors of $\simeq 4\%$ with respect to the omni-WP curves, and can be therefore considered as an uncertainty inherent in this kind of computations (see e.g., Borgman and Resio, 1977, 1982, stating that reliable long term estimates can be carried out just up to three times the years the length of the original dataset).

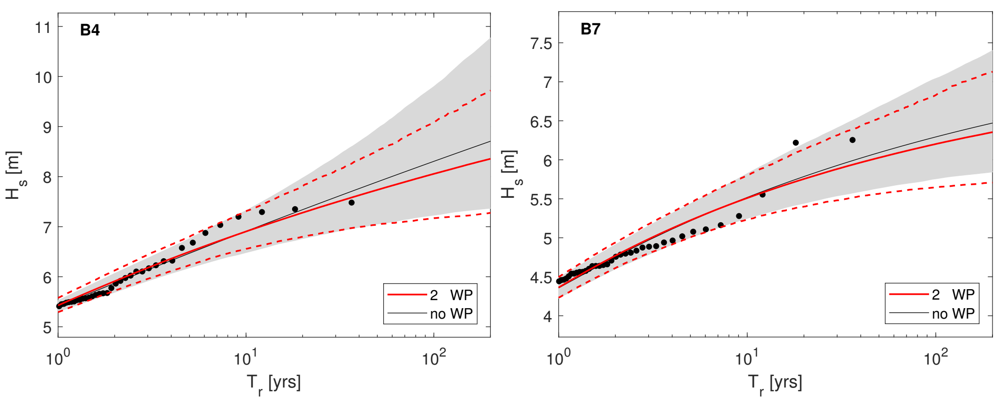


Figure 11. Omni-WP extreme value distributions of H_s obtained from the whole set of peaks (black) and from combining single-WP distributions (red), along with 90% confidence intervals (grey shadow and red dashed lines, respectively). Left panel: B4; right panel: B7

These results seem to reinforce the validity of developing the omni-WP curve under the hypothesis of independence between the events of different WPs. In this case, the independence hypothesis was to some extent corroborated by the low correlations values attained by the intra-clusters annual frequency of occurrence for all the locations: -0.17, 0.23, -0.13, -0.11, -0.34, 0.13 for B5, B6, B4, B7, B8 and B3 respectively).

4 Conclusions

Results on the analysis of the extremes reveal how ~~different weather circulation patterns may lead to~~ sea storms having, on average, unique characteristics, can be driven by different weather circulation patterns. Such homogeneous features (both in terms of frequency of occurrence and significant wave heights) suggest that the extreme distributions of H_s shall be singularly evaluated for each WP, the starting datasets being homogeneous and independent with respect to each other.

The methodology here introduced allows the classification of extreme wave events (or other oceanic variables) into homogeneous subgroups according to the atmospheric processes most likely generating them. By focusing solely on extreme events, the proposed classification method results in a reduced number of circulation or weather patterns, which facilitates their physical interpretation as well as their linkage with the climatology of the area.

The method, as presented here, does not contemplate the inclusion of trends and inter- or intra-annual cycles. However, the extension of the methodology in this direction is straightforward, as the methods previously developed for non-stationary analysis (see e.g., Sartini et al., 2015) could be applied without major complexities to each of the subsets that are obtained from the classification. On the contrary, it is not obvious how to proceed in the selection of the number of patterns to be considered. Here this number was chosen following a qualitative analysis of the results, which was viable for the case study analysed. However, this approach is not always feasible, and sometimes it might be necessary to (at least) resort to a sensitivity analysis of the results, as long as a quantitative methodology is not available for the definition of the number of clusters.

For the analysed locations, the proposed methodology led to the identification of two well known cyclonic systems, characteristic of the atmospheric circulation on the Mediterranean Sea, as the possible origin of the extreme wave events affecting the Italian shores. When extreme events are classified according to their meteorological origin, there is a great confidence of working with homogeneous samples, thus being in compliance with the main hypothesis underlying the EVA. Nevertheless, the practical consequences of the classification in terms of the omni-WP extreme distribution are relatively limited: although in some cases a narrowing in the confidence intervals was achieved, in general the distribution obtained by following the two different approaches was very similar. However, it must be kept in mind that the classification could facilitate other aspects of the analysis not included in this work, ~~such as~~ as an instance the multivariate analysis of extreme events. Indeed, in such a framework, to classify the wave fields according to the wind velocities lead to clusters of T_p consistent with those of H_s, as the latter parameter is closely tied to the former one (especially in case of extreme sea states).

Code and data availability. The algorithms used in this work were developed in Matlab®. The codes and the hindcast data employed are available upon request. Please contact FDL at: francesco.deleo@edu.unige.it.

Author contributions. FDL developed the algorithms and wrote the paper. SS coordinated the work and contributed to developing the codes and writing the paper. GB provided the data, analysed the results in the frame of the MR climatology, and revised the writing of the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Mobility of Francesco De Leo has been financed by the University of Genoa through the program "Fondo Giovani 17/18".

Authors deeply acknowledge Francesco Ferrari of the University of Genoa for the help in processing the meteorological hindcast.

References

API: API RP 2A-WSD: Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms-Working Stress Design, American Petroleum Institute, 2002.

Bencivenga, M., Nardone, G., Ruggiero, F., and Calore, D.: The Italian data buoy network (RON), Advances in Fluid Mechanics IX, 74, 321, 2012.

Borgman, L. and Resio, D.: Extremal prediction in wave climatology, in: Proceedings of Ports, ASCE, New York, vol. I, pp. 394–412, 1977.

Borgman, L. E. and Resio, D.: Extremal statistics in wave climatology, Topics in ocean physics, 80, 439–471, 1982.

Camus, P., Menéndez, M., Méndez, F. J., Izaguirre, C., Espejo, A., Cánovas, V., Pérez, J., Rueda, A., Losada, I. J., and Medina, R.: A weather-type statistical downscaling framework for ocean wave climate, Journal of Geophysical Research: Oceans, 119, 1–17, https://doi.org/10.1002/2014JC010141, 2014.

Camus, P., Rueda, A., Méndez, F. J., and Losada, I. J.: An atmospheric-to-marine synoptic classification for statistical downscaling marine climate, Ocean Dynamics, 66, 1589–1601, https://doi.org/10.1007/s10236-016-1004-5, http://dx.doi.org/10.1007/s10236-016-1004-5, 2016.

Cassola, F., Ferrari, F., and Mazzino, A.: Numerical simulations of Mediterranean heavy precipitation events with the WRF model: analysis of the sensitivity to resolution and microphysics parameterization schemes, Atmospheric Research, 164-165, 210–225, 2015.

Cassola, F., Ferrari, F., Mazzino, A., and Miglietta, M.: The role of the sea on the flash floods events over Liguria (northwestern Italy), Geophysical Research Letters, 43, 3534–3542, 2016.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values (Springer series in statistics), Springer-Verlang, 2001.

Cook, N. J. and Miller, C. A.: Further note on directional assessment of extreme winds for design, Journal of Wind Engineering and Industrial Aerodynamics, 79, 201–208, https://doi.org/10.1016/S0167-6105(98)00109-3, https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-198/, 1999.

Dangendorf, S., Wahl, T., Nilson, E., Klein, B., and Jensen, J.: A new atmospheric proxy for sea level variability in the southeastern North Sea: observations and future ensemble projections, Climate Dynamics, https://doi.org/10.1007/s00382-013-1932-4, http://link.springer.com/10.1007/s00382-013-1932-4, 2013.

DNV: DNV-RP-C205: Environmental Conditions and Environmental Loads, Det Norske Veritas, 2010.

Flocas, A. and Giles, B.: Distribution and intensity of frontal rainfall over Greece, International journal of climatology, 11, 429–442, 1991.

Folgueras, P., Solari, S., and Losada, M. Á.: The selection of directional sectors for the analysis of extreme wind speed, Natural Hazards and Earth System Sciences, 19, 221–236, 2019.

Forristall, G. Z.: On the use of directional wave criteria, Journal of waterway, port, coastal, and ocean engineering, 130, 272–275, 2004.

Guanche, Y., Mínguez, R., and Méndez, F. J.: Climate-based Monte Carlo simulation of trivariate sea states, Coastal Engineering, 80, 107–121, http://www.sciencedirect.com/science/article/pii/S0378383913000963, 2013.

Holt, T.: A classification of ambient climatic conditions during extreme surge events off Western Europe, International Journal of Climatology, 19, 725–744, https://doi.org/10.1002/(SICI)1097-0088(19990615)19:7<725::AID-JOC391>3.0.CO;2-G, 1999.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J., and Tveito, O. E.: Classifications of atmospheric circulation patterns: recent advances and applications., Annals of the New York Academy of Sciences, 1146, 105–52, https://doi.org/10.1196/annals.1446.019, http://www.ncbi.nlm.nih.gov/pubmed/19076414, 2008.

ISO: ISO 19901-1:2005: Petroleum and natural gas industries - Specific requirements for offshore structures - Part 1: Metocean design and operating considerations, International Organization for Standards, 2005.

Lionello, P., Trigo, I., Gil, V., Liberato, M., Nissen, K., Pinto, J., Raible, C., Reale, M., Tanzarella, A., Trigo, R., Ulbrich, S., and Ulbrich, U.: Objective climatology of cyclones in the Mediterranean region: a consensus view among methods with different system identification and tracking criteria, Tellus A, 68, 2016.

MacQueen, J. et al.: Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1, pp. 281–297, Oakland, CA, USA, 1967.

Mentaschi, L., Besio, G., Cassola, F., and Mazzino, A.: Developing and validating a forecast/hindcast system for the Mediterranean Sea, Journal of Coastal Research, SI 65, 1551–1556, 2013.

Mentaschi, L., Besio, G., Cassola, F., and Mazzino, A.: Performance evaluation of WavewatchIII in the Mediterranean Sea, Ocean Modelling, 90, 82–94, 2015.

Pringle, J., Stretch, D. D., and Bárdossy, A.: Automated classification of the atmospheric circulation patterns that drive regional wave climates, Natural Hazards and Earth System Science, 14, 2145–2155, https://doi.org/10.5194/nhess-14-2145-2014, http://www.nat-hazards-earth-syst-sci.net/14/2145/2014/, 2014.

Pringle, J., Stretch, D. D., and Bárdossy, A.: On linking atmospheric circulation patterns to extreme wave events for coastal vulnerability assessments, Natural Hazards, 79, 45–59, https://doi.org/10.1007/s11069-015-1825-4, 2015.

Rueda, A., Camus, P., Méndez, F. J., Tomás, A., and Luceño, A.: An extreme value model for maximum wave heights based on weather types, Journal of Geophysical Research, 121, 1–12, https://doi.org/10.1002/2015JC010952, 2016.

Sartini, L., Cassola, F., and Besio, G.: Extreme waves seasonality analysis: An application in the Mediterranean Sea, Journal of Geophysical Research: Oceans, 120, 6266–6288, 2015.

Skamarock, W.: Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Tech. Note, Tech. rep., NCAR/TN-475+ STR, 2009.

Solari, S. and Alonso, R.: A new methodology for extreme waves analysis based on weather-patterns classification methods, Coastal Engineering Proceedings, 1, 23, 2017.

Solari, S., Egüen, M., Polo, M. J., and Losada, M. A.: Peaks Over Threshold (POT): A methodology for automatic threshold estimation using goodness of fit p-value, Water Resources Research, 53, 2833–2849, 2017.

Trigo, I. F., Davies, T. D., and Bigg, G. R.: Objective climatology of cyclones in the Mediterranean region, Journal of Climate, 12, 1685–1696, 1999.

Trigo, I. F., Bigg, G. R., and Davies, T. D.: Climatology of cyclogenesis mechanisms in the Mediterranean, Monthly Weather Review, 130, 549–569, 2002.

Yarnal, B., Comrie, A. C., Frakes, B., and Brown, D. P.: Developments and prospects in synoptic climatology, International Journal of Climatology, 21, 1923–1950, 2001.