Coastal Engineering 167 (2021) 103896

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: http://www.elsevier.com/locate/coastaleng

Non-stationary extreme value analysis of sea states based on linear trends. Analysis of annual maxima series of significant wave height and peak period in the Mediterannean Sea

Francesco De Leo [a,b,*], Giovanni Besio [b], Riccardo Briganti [c], Erik Vanem [d,e]

[a] Civil and Environmental Engineering Department, California Polytechnic State University, San Luis Obispo, CA, USA
[b] Department of Civil, Chemical and Environmental Engineering, University of Genoa, Genoa, 16145, Italy
[c] Faculty of Engineering, University of Nottingham, Nottingham, NG7 2RD, United Kingdom
[d] DNV GL GTR, Høvik, Norway
[e] Department of Mathematics, University of Oslo, Oslo, Norway

ARTICLE INFO

Keywords:
Non-stationary extreme value analysis
Trend analysis
Significant wave height
Peak wave period
Environmental contours
Mediterranean sea

ABSTRACT

Non-stationary Extreme Value Analysis (NEVA) allows to determine the probability of exceedance of extreme sea states taking into account trends in the time series of data at hand. In this work, we analyse the reliability of NEVA of significant wave height (H_s) and peak period (T_p) under the assumption of linear trend for time series of annual maxima (AM) H_s in the Mediterranean Sea. A methodology to assess the significance of the results of the non-stationary model employed is proposed. Both the univariate long-term extreme value distribution of H_s and the bivariate distribution of H_s and T_p are considered. For the former, a non-stationary Generalized Extreme Value (GEV) probability is used, and a methodology to compute the parameters of the distribution based on the use of a penalty function is explored. Then, non-stationary GEV is taken as a reference to compute the Environmental Countours of H_s and T_p, assuming a conditional model for the latter parameter. Several methods to compute linear trends are analysed and cross-validated on the series of AM H_s at more than 20,000 hindcast nodes. Results show that the non-stationary analysis provides advantages over the stationary analysis only when all the considered metrics are consistent in indicating the presence of a trend. Moreover, both the univariate return levels of H_s and bivariate return levels of H_s and T_p show a marked dependence to the time window considered in the GEV distribution formulation. Therefore, when applying NEVA for coastal and marine applications, the hypothesis of linear trend and the length of the reference data used for the non-stationary distribution should be carefully considered.

1. Introduction

Non-Stationary Extreme Value Analysis (NEVA) has reached prominence in water engineering due to the importance of including climatic trends in traditional Extreme Value Analysis (EVA). Coastal engineering makes no exception in the relevance of this topic (Corbella and Stretch, 2012; Lucio et al., 2020). Design parameters, such as the design significant wave height with a prescribed return period, are directly influenced by the existence of trends. In fact, the time-dependency of the statistical distribution of a certain environmental stochastic process determines the time-dependency of the parameters of the associated extreme value distribution (Renard et al., 2013). For example, the shape, scale and location parameters in the Generalized Extreme Value (GEV) probability distribution, can be considered varying in time due to an underlying trend (Coles et al., 2001).

However, the condition of time-dependency of all three parameters is often simplified, and the location parameter only is modelled as a linear function of time, while the scale and the shape parameters are kept constant. This is motivated by the need of having very long time series to reliably modelling the time variation of the latter two parameters. This assumption was used in a number of applications (Barbero et al., 2017; Cheng and AghaKouchak, 2014; Luke et al., 2017; Mentaschi et al., 2016; Vanem, 2015); all of the cited works used the model:

$$\mu = \mu_0 + \mu_t t, \tag{1}$$

* Corresponding author. Civil and Environmental Engineering Department, California Polytechnic State Univeristy, San Luis Obispo, CA, USA.
E-mail address: fdeleo@calpoly.edu (F. De Leo).

https://doi.org/10.1016/j.coastaleng.2021.103896
Received 28 October 2020; Received in revised form 8 March 2021; Accepted 27 March 2021
Available online 7 April 2021

where μ is the location parameter of non-stationary GEV (hereinafter N-GEV) probability distribution, μ_0 the stationary part, μ_1 the time-varying one and t is the reference time used to estimate μ. This model assumes that the centre of the distribution at hand is, on average, shifted by a fixed ratio over time. Implicitly, it assumes that the trend of the population of the extremes is linear (Coles et al., 2001). Notwithstanding the popularity of this model, there is no analysis on the correlation between μ_t and the magnitude of the trend to determine whether the two are consistent regardless of the significance of the trend.

The magnitude and significance of linear trends was investigated for the annual maxima (AM) of significant wave height (H_s) in the Mediterranean Sea (MS) by De Leo et al. (2020), who assessed the use of a linear slope modified according to the model of Sen (1968) and Theil (1992) (referred to as b) to compute the magnitude of trends. To this end, De Leo et al. (2020) evaluated b against the Mann-Kendall test (Mann, 1945; Kendall, 1955) and the Innovative Trend Analysis (Şen, 2011, 2013), neither bound by the assumption of a linear trend, and showed that the outcomes of all the trend metrics are remarkably related. This finding is not obvious: indeed, while the use of b allows to compute the magnitude of a linear trend, the Mann-Kendall test, and in particular its p_{value} (henceforth referred to as p_{MK}), only allows to assess the significance of a monotonic trend. That is, if p_{MK} attains close-to-0 values, then the *null hypothesis* that an either positive or negative monotonic trend characterises the data at hand cannot be rejected; on the contrary, close to 1 values indicate that the *null hypothesis* should be rejected. In case of the time series of AM H_s in the MS, the vast majority of relevant (negligible) trends in terms of b also resulted to be significant (not significant) according to p_{MK}. Therefore, b may be used to get a first insight into the magnitude of a long-term trend, and to validate the hypothesis that the parameters of the selected extreme value distribution are linearly varying in time. Based on these grounds, the results of De Leo et al. (2020) were here employed to assess the reliability of the estimates of μ_t computed on the same AM H_s series using the Maximum Likelihood Method (MLM). The Mann-Kendall p_{MK} and the slopes of Theil-Sen (Sen, 1968; Theil, 1992), referred to as b_{AM}, were considered to evaluate the consistency among the different methods at more than 22,000 hindcast locations. As a second step, the return levels of H_s were computed at specific sites according to both EVA and NEVA, and results were compared.

For the determination of the N-GEV distribution parameters it was taken into account that the MLM could fail to provide reliable Maximum Likelihood Estimators (MLE) parameters, especially when the shape parameter (ξ) attains values lower than −1 (Smith, 1985). In such a case, alternative methods for the computation of the parameters should be employed; for instance the method of L-moments (Hosking and Wallis, 2005) is a suitable alternative for stationary EVA. However, in case of N-GEV distribution, no explicit expressions for the computation of the parameters have been derived yet, thus the use of L-moments is not straightforward. As a possible alternative, a penalized version of the MLM, referred as PMLM, can be used. This method takes advantage of a penalty function (P_f) to constrain the likelihood to a subset of possible values of ξ (Coles and Dixon, 1999). In the framework of met-ocean parameters, Mackay et al. (2011) considered a penalty function to derive return levels of significant wave heights, using the estimators proposed by Coles and Dixon (1999). In the present work, we present an alternative formulation for P_f, which allows to avoid spurious estimates of ξ, while it does not affect the N-GEV parameters in case that the MLM can be reliably applied.

EVA in coastal engineering is further complicated by the fact that in understanding extreme sea state conditions the knowledge of H_s is not sufficient and information on a spectral period (e.g. the peak period T_p) is also needed. A possible way of characterizing joint extremes of H_s and T_p is by way of bivariate quantile curves corresponding to OR and AND exceedances, respectively, as outlined in e.g. Vanem (2020). Among the other methodologies to take into account both parameters in extreme value analysis, the Environmental Countour Method, first proposed by Haver (1985) and Haver (1987) and subsequently followed and extended in many works such as Haver and Winterstein (2009); Haselsteiner et al. (2017), defines extreme conditions depending on H_s and T_p by computing contour lines. These lines represent the Environmental Contours (ECs) of a given exceedance probability, which essentially indicate the probability of exceeding a hyperplane (or a tangent line in 2-dimensional cases). Different contour methods differently defines these hyperplanes, for instance in the variables space (Huseby et al., 2013, 2015), or in a transformed standard normal space (Haver and Winterstein, 2009); the differences between these approaches are discussed in Vanem and Bitner-Gregersen (2015) and Vanem (2017). In general, the Environmental Contour Method requires a joint extreme value probability for the two parameters, which can be obtained, for example, starting from a GEV distribution for H_s and assuming a conditional probability distribution for T_p. Starting from this distribution, one of the methods available to compute the ECs is the Inverse First Order Reliability method or "IFORM" (Haver and Winterstein, 2009; Leira, 2008).

While the cited literature analyses traditional EVA, the analysis of the ECs with NEVA is still very limited. Examples of bivariate analysis of H_s and T_p can be found in Vanem (2015), who took advantage of the theory of copulas; and Huseby et al. (2013), who computed ECs based on a Monte Carlo approach and modelling the effect of a long-term trend in H_s by modifying the parameters of a 3-parametric Weibull distribution. The present work seeks to provide a novel, efficient and straightforward method to compute ECs for bivariate analysis of H_s and T_p in the framework of NEVA. For this reason, ECs were computed instead assuming a distribution of T_p conditioned to the N-GEV model applied to the series of AM H_s. As in case of the univariate analysis, the return levels following both the EVA and NEVA approaches were computed and compared.

In summary, there are still several open questions on the optimal use of NEVA for engineering purposes, above all in the assessment of extreme sea state in a basin such as the MS in which the level of significance of the trends found varies greatly (De Leo et al., 2020).

In this respect, this work aims to trace a work-flow for a solid exploratory data analysis, that can be used for engineering practice and that helps to prevent the misuse of non-stationary models (Serinaldi and Kilsby, 2015).

The article is organised as follows. After this Introduction, Section 2 introduces the data employed in this study, along with the methodology used to compare the stationary and non-stationary analysis. Results are presented and discussed in Section 3, followed by Section 4, which provides conclusions of this research.

2. Data and methods

This work used the wave hindcast data provided by the Department of Civil, Chemical and Environmental Engineering of the University of Genoa (Mentaschi et al., 2013, 2015), already validated and used in several works on the MS (Besio et al., 2017; Ferrari et al., 2020 among others) The hindcast is developed over the whole MS (see Figs. 3–5) at a resolution of about 10 km both in longitude and latitude, resulting in 22,373 nodes. At each node, the time series of H_s and T_p over the period 1979–2018 were retained, and the AM H_s were selected for the analysis in the paper. First, EVA and NEVA were performed on the selected AM H_s series, using both a three parameters GEV distribution and a non-stationary four parameters N-GEV distribution. In particular, the N-GEV distribution depends on the location parameters μ_0 and μ_t, that are related to the parameter μ by Eq. (1), and scale and shape parameters that are considered constant in time. The stationary GEV distribution is recovered when $\mu_t = 0$.

Subsequently, the results of De Leo et al. (2020) for b_{AM} were correlated to the estimates of μ_t computed on the same AM H_s series. p_{MK}

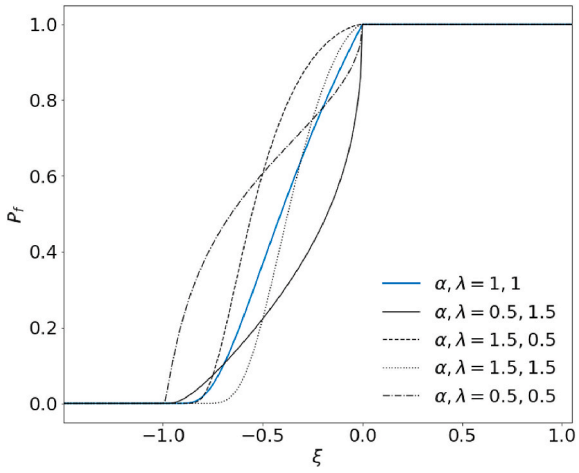

Fig. 1. P_f profile for different pairs of α and λ. Highlighted in blue is the curve for α and λ both equal to 1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

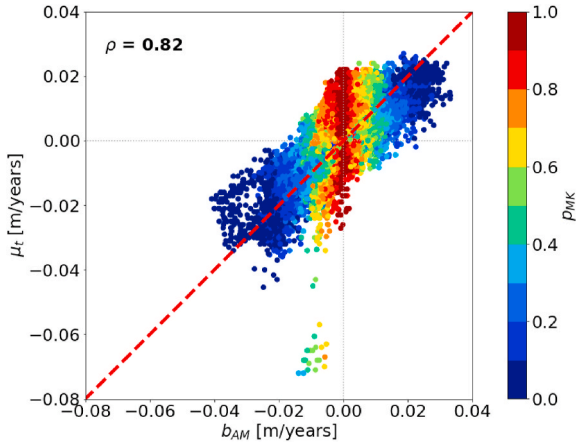

Fig. 2. Correlation between the series of b_{AM} and μ_t

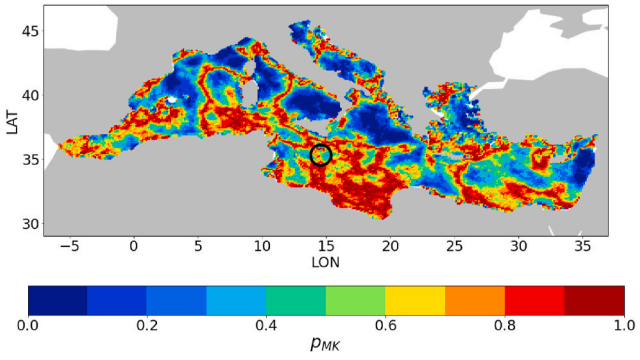

Fig. 3. Spatial distribution of p_{MK} over the Mediterranean Sea. The black circle delimits the hindcast locations showing diverging trend metrics (same in Figs. 4 and 5).

was used to verify the consistency between the two different parameters. Seven locations among the computational nodes were examined and different cases showing b_{AM} and μ_t either in agreement or not were analysed in depth. The significance of trends identified by b_{AM} and μ_t for these locations was assessed by looking at the 95% confidence interval (CI) of the two parameters. For b_{AM} CI was computed with the model of Hollander et al. (2013); for μ_t CI was computed using the maximum

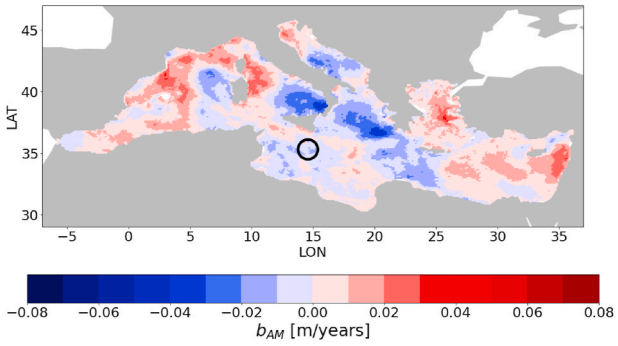

Fig. 4. Spatial distribution of b_{AM} over the Mediterranean Sea.

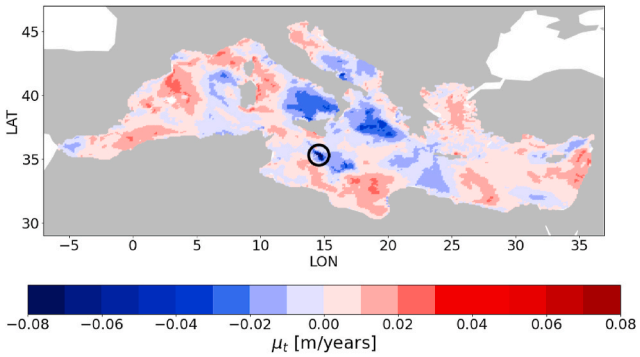

Fig. 5. Spatial distribution of μ_t over the Mediterranean Sea.

likelihood estimation of the N-GEV distribution parameters.

The Aikake Information Criteria (AIC, Akaike, 1998) was employed to compare the GEV and the N-GEV distributions, to select the best performing distribution among them. The AIC allows to select the best performing extreme value distribution when the parameters are estimated via the MLM. In particular, for a j^{th} ($j = 1, 2, ..$) distribution, the AIC_j statistic is computed as:

$$AIC_j = -2l_j + 2k_j, \tag{2}$$

l_j being the maximum of the log-likelihood function and k_j the number of parameters of the j^{th} distribution; for instance, $k = 3$ and $k = 4$ for the GEV and N-GEV distribution respectively. The distribution with the lowest value of *AIC* is the one to be preferred (Sartini et al., 2015).

Subsequently, in order to evaluate the possible implications of accounting for intra-period trends in the determination of design H_s, return values H_s resulting from both the EVA (using the GEV distribution) and the NEVA (using the N-GEV distribution) approaches were computed. Return values are defined as the values that a variable, in this case H_s, attains in relation of the so called return period (T_r), i.e. the level that is expected to be exceeded on average once in T_r years (Coles et al., 2001). For the N-GEV distribution, the following equations for the computation of the return levels of a variable z generally apply:

$$\begin{cases} z = \mu - \dfrac{\sigma}{\xi}\{1 - [-\log(1-P)]^{-\xi}\} \text{ for } \xi \neq 0 \\ z = \mu - \sigma\log[-\log(1-P)] \text{ for } \xi = 0 \end{cases} \tag{3}$$

in this case, z refers to H_s, σ and ξ are the scale and shape parameters, respectively. P is the cumulative probability, associated to the return period considered. μ is defined by Eq. (1) and, since when $\mu_t = 0$ the stationary GEV distribution is recovered, Eq. (3) can be used also for the EVA.

As mentioned in Section 1, at the hindcast nodes where MLM could not provide sound estimates for the N-GEV parameters it was introduced

the use of a penalty function. The method to compute the Penalized Maximum Likelihood Estimators (PMLE) uses a penalty function (P_f) as follows (Coles and Dixon, 1999):

$$L_{pen}(\mu,\sigma,\xi) = L(\mu,\sigma,\xi) \times P_f \tag{4}$$

where L is the common likelihood function and L_{pen} is the corresponding penalized version. The chosen P_f in this work reads:

$$P_f(\xi) = \begin{cases} 0, & \text{if } \xi \le -1 \\ \exp\left[-\lambda\left(\frac{1}{1+\xi}-1\right)^{\alpha}\right], & \text{if } -1 < \xi < 0 \\ 1, & \text{if } \xi \ge 0 \end{cases} \tag{5}$$

where α and λ are coefficient that allow to tune the shape of P_f. Therefore, the use of $P_f(\xi)$ allows restrictions on the ξ values. In the original formulation by Coles and Dixon (1999), the range on ξ was defined in order to obtain values of ξ strictly lower than 1, consistent with the hypothesis used to derive analytical expressions for the GEV distribution weighted moments (Hosking and Wallis, 2005). In this work, $P_f(\xi)$ was applied instead within the [-1,0] range, to avoid spurious estimates of the N-GEV distribution parameters owing to MLE of ξ lower than −1. Fig. 1 depicts the profiles of P_f used in this research according to different choices for α and λ.

Finally, this analysis was extended to the bivariate distributions of H_s and T_p and the ECs for the H_s-T_p pairs were computed.

According to the IFORM (Haver and Winterstein, 2009; Leira, 2008), a conditional log-normal distribution for T_p was considered, whose return levels are defined as:

$$\begin{cases} T_p = \exp\left[\mu_{T_p|H_s} + \left(P^* \times \sigma_{T_p|H_s}\right)\right] \\ \mu_{T_p|H_s} = a_1 + a_2 H_s^{a_3} \\ \sigma_{T_p|H_s} = b_1 + b_2 \exp(b_3 H_s) \end{cases} \tag{6}$$

$\mu_{T_p|H_s}$ and $\sigma_{T_p|H_s}$ indicate the mean and standard deviation of T_p with respect to H_s, respectively, while P^* is the cumulative probability in a transformed standard normal space, which has to be transposed in the real H_s-T_p space (details can be found in Vanem, 2016, 2018). a_i and b_i have to be estimated from the whole time series of H_s and T_p: first, the support of H_s needs to be divided in subsets of given width, for which expected value and standard deviation of the corresponding T_p are computed, and subsequently used to fit the curves of Eq. (6). In this study, bins of 0.5 m were selected according to the suggestion of Vanem (2016). The ECs computed according to this method correspond to the probabilities of exceeding hyperplanes in a transformed standard normal space.

3. Results & discussion

3.1. AM trends in the Mediterranean Sea

The correlation between b_{AM} and μ_t for the AM series of H_s is shown in Fig. 2, where the color of each marker is scaled according to the p_{MK} of the respective series.

Both b_{AM} and μ_t allow to estimate a linear trend in the data; however, while the latter corresponds to the expected value (i.e. average) of a trend, the former corresponds to the median trend, and it is therefore less sensitive to possible outliers in the data. Despite the differences characterizing the two metrics investigated, for the analysed AM H_s series they result to be highly correlated, as shown in Fig. 2. Indeed, it is noticed how the Spearman's correlation coefficient $\rho = 0.82$, indicating a high correlation between the two metrics, i.e. low (high) values of b_{AM} are most likely to occur when low (high) values of μ_t are observed. Although the cloud is well aligned with the first/third quadrant bisector, a slight bias is recognisable, for which μ_t tends to be higher than b_{AM} of about 0.1 mm/years. In general the signs of b_{AM} and μ_t agree for higher values of the two parameters, while for low trends, showing values of p_{MK} close to one, signs are often opposite, i.e. the points are in the second or fourth quadrants. There are also locations characterized by consistent trend signs for both statistics, but with significant variation in magnitude. However, also this condition generally occurs when p_{MK} attain values close to one, indicating that the hypothesis of underlying trends should be rejected. Finally, there are a few locations whose AM H_s series show remarkable negative trends in terms of μ_t, although they are not characterized by equally relevant trends according to p_{MK} and b_{AM}.

The aforementioned correlation can be also appreciated by looking at the spatial distribution of p_{MK}, b_{AM} and μ_t in the MS, as shown in Figs. 3–5, respectively. By comparing the maps, it is noticed that b_{AM} and μ_t are similarly distributed over the basin, consistently with the correlation shown in Fig. 2. Also, the nodes showing close to zero trends are mainly located in transition areas between positive and negative trends and they are the nodes with values of p_{MK} close to one. The southmost portion of the central Mediterranean basin shows the lowest trend values and the highest p_{MK} values. Interestingly, the nodes where significantly diverging trend metrics are all found lie in a well limited area, bordered with the black circle in the figures.

3.2. Analysis on the trends significance

Based on the results illustrated in Section 3.1, seven hindcast points were chosen for an in-depth analysis. The points are marked in Fig. 6, along with their identifying codes. For the sake of clarity, the points are labeled with numbers from 1 to 7, and the respective grid nodes locations can be found in Fig. 7. In the following the node identification numbers will be used.

The locations are representative of the variety of results for b_{AM} and μ_t. In particular, the results for Point_005284 (1) and Point_010830 (2) show evidence of consistently positive and negative trends according to the metrics examined, respectively. Point_007803 (6) and Point_009108 (4) show consistent trends signs, but significantly different intensity; the same consideration applies for Point_015293 (7) and Point_009107 (5), but in these cases b_{AM} and μ_t also indicate opposite trends. Finally, Point_015123 (3) is taken into account as it belongs to the few locations significantly departing from the main cloud in Fig. 2. Note that while Point 5 and Point 4 are both geographically close and similar in trend metrics, on the other hand Point 3 and Point 7 are geographically close but with μ_t being significantly different. The trend metrics corresponding to the selected locations are summarized in Table 1. In the table the

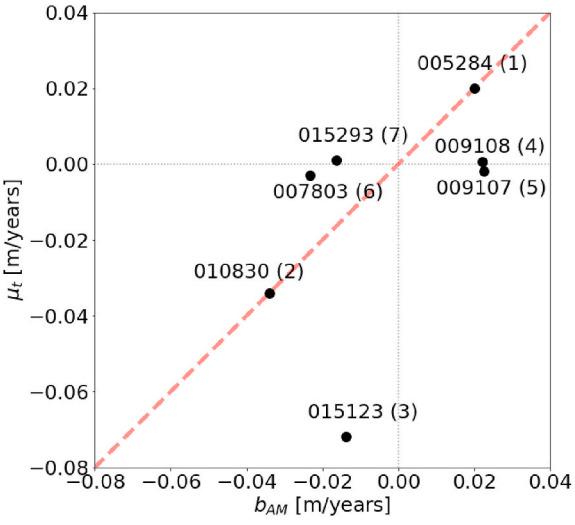

Fig. 6. Hindcast locations retained for the comparison between EVA and NEVA approaches.

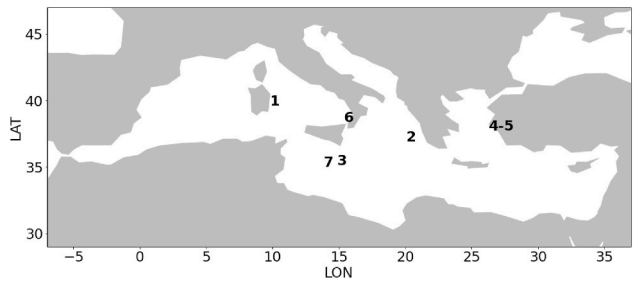

Fig. 7. The Mediterranean Sea with the locations of the grid points used for the detailed analysis. Points 009180 (4) and 009107 (5) are placed together as they lie next to each other.

subscript 2.5% and 97.5% indicate the 95% CI for the parameter used.

From the analysis of CI and p_{MK}, Point 2 is the only location where p_{MK} attains a value significantly close to zero and no changes of sign occur between the 2.5% and the 97.5% levels of μ_t and b_{AM}, indicating a robust consensus on the presence of a trend. On the contrary, in case of Point 1, CI limits of μ_t and b_{AM} are of opposite sign, although this point shows values of the two parameters among the highest encountered in the whole MS for upward trends, albeit with $p_{MK} > 0.05$. No significant trends with respect to μ_t are found also for Point 7 and Point 5; and for Point 6, and Point 4, though according to p_{MK} strong evidence of trend exist. In these two cases, an alternative trend model rather than Eq. (1) could be considered. Finally, as for Point 3, b_{AM} and p_{MK} suggest that no trend exists, whilst for μ_t it was not possible to compute reliably the CI. As a matter of fact, for this location MLM fails to estimate the N-GEV distribution parameters, owing to an estimate of ξ lower than -1. This

Table 1
Trend metrics for the hindcast locations considered; b^*_{AM} and μ^*_t are expressed in [m/years].

Point	$\mu_t^{2.5\%}$	μ_t	$\mu_t^{97.5\%}$	$b_{AM}^{2.5\%}$	b_{AM}	$b_{AM}^{97.5\%}$	p_{MK}
1	−0.0090	0.0200	0.0490	−0.0070	0.0200	0.0450	0.1210
2	−0.0560	−0.0340	−0.0120	−0.0620	−0.0340	−0.0050	0.0290
3	-	−0.0720	-	−0.0460	−0.0140	0.0130	0.3890
4	−0.0090	0.0010	0.0100	0	0.0220	0.0540	0.0460
5	−0.0140	−0.0020	0.0100	−0.0020	0.0220	0.0520	0.1160
6	−0.0210	−0.0030	0.0150	−0.0540	−0.0230	0.0020	0.0750
7	−0.0310	0.0010	0.0330	−0.0420	−0.0160	0.0160	0.2940

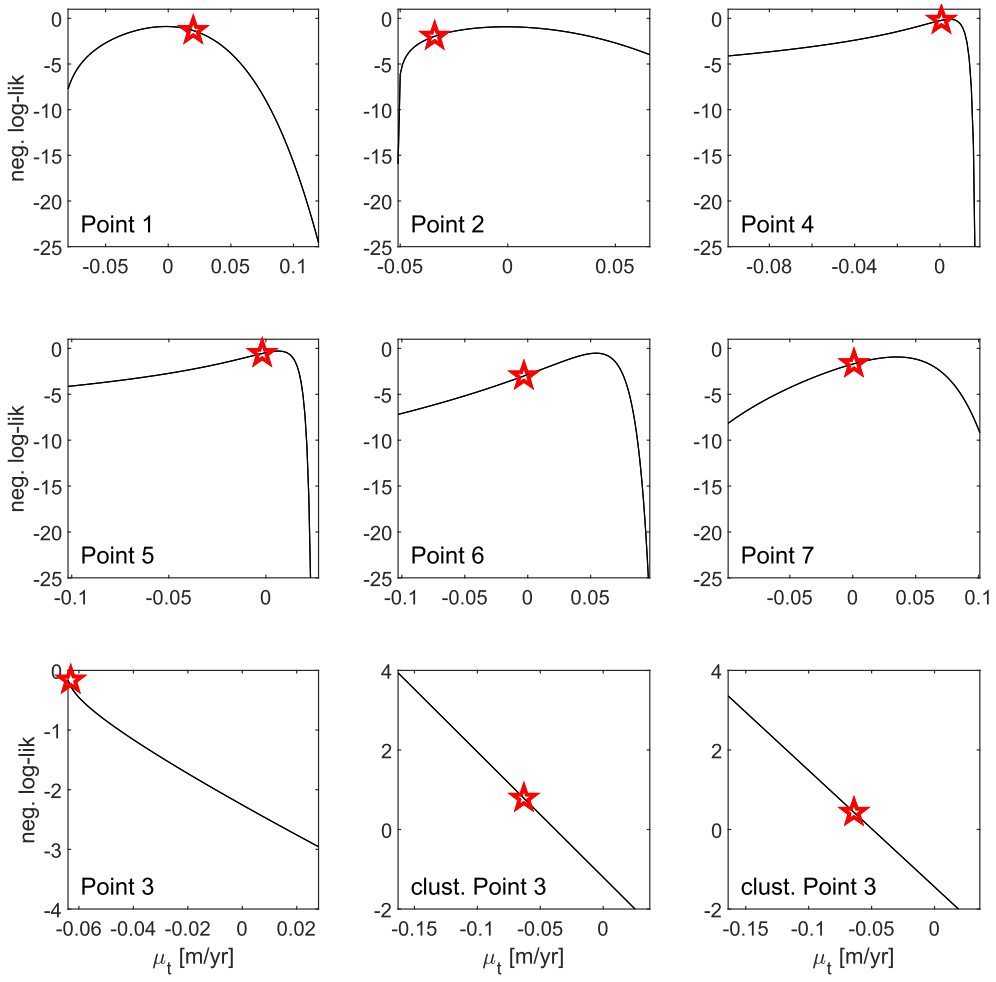

Fig. 8. Negative log-likelihood function with respect to μ_t for the investigated locations. The optimal estimates of μ_t are highlighted through a red star. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

can be appreciated by looking at the log-likelihood functions for μ_t at the investigated locations in Fig. 8. Shown in the panels are the curves of the negative log-likelihood for varying μ_t and values of the other parameters fixed to the respective maximum likelihood estimates.

It is shown that the estimates lie close to the maxima for all the locations but Point 3, where no local maximum can be found. The bottom panels of Fig. 8 show the likelihood functions also for two other locations showing trend outcomes similar to those of Point 3 and located as well within the black circle in Figs. 3–5. Also in these cases, no local maxima could be found in the likelihood functions, thus the PMLM was used, and results are presented further on Section 3.5.

3.3. Goodness of fit of the GEV and N-GEV distributions

When the MLM succeeded in computing the N-GEV distribution parameters, the AIC was used as Goodness-Of-Fit (GOF) test to compare GEV ($k = 3$) and N-GEV ($k = 4$) distributions results. Results are reported in Table 2.

The AIC statistic reveals that the N-GEV distribution performs better only in case of Point 2, i.e. where all the metrics employed agree on the significance of the trend. For all the other locations, the corresponding series would be better modelled by a GEV distribution according to the GOF test, as AIC attains lower value compared to the N-GEV distribution. However, in case of Point 1, the values of AIC related to the different models are very similar. At this location, also the CI of μ_t and b_{AM} are characterized by different orders of magnitude, with 2.5% and 97.5% levels indicating opposite trends in the orders of mm/year and cm/year, respectively (see Table 1). This indicates a milder variation in the CI with respect to all the locations but Point 2, and it is consistent with the outcomes of the AIC test. Even though a non-stationary analysis would be completely justified only in case of Point 2, for the sake of comparison the return levels of H_s following the EVA and NEVA approaches were computed also for Point 1. The AM H_s series at the two locations are shown in the Supplement (Fig. S1).

3.4. Return levels of H_s with EVA and NEVA approaches

The return levels of H_s were computed by Eq. (3), and the results presented in Fig. 9 refer to a reference time $t = 40$ years (see Eq. (1)); here the CI were computed using the Delta method (Coles et al., 2001). The colors of the curves and the shaded areas follow the palette used in Figs. 4–5 to indicate positive and negative trends (red and blue, respectively).

In case of Point 1, the CI of the EVA approach is narrower than the CI of the NEVA one, although the lower bounds of the two approaches practically overlap. Moreover, the return levels of H_s following the N-GEV distribution lie within the CI computed according to the stationary analysis, except for T_r lower than ~ 4 years. This indicates that the EVA and NEVA approaches do not lead to significantly different results. On the other hand, in case of Point 2 the return levels obtained with the NEVA approach are remarkably lower than the corresponding stationary results: the estimated trends are outside the CI of the alternative model (i.e. the GEV distribution return levels curve is outside the N-GEV distribution CI and vice versa). This applies regardless of the T_r taken into account, and it indicates that the different approaches yield significantly different results.

Table 2
AIC statistics for the hindcast locations considered.

	GEV		N-GEV	
Point	neg. log-lik	AIC	neg. log-lik	AIC
1	55.32	116.64	54.36	116.72
2	57.23	120.46	54.46	116.92
4	47.17	100.33	47.16	102.31
5	51.05	108.11	51.01	110.01
6	46.81	99.63	58.64	125.29
7	57.30	120.61	57.30	122.60

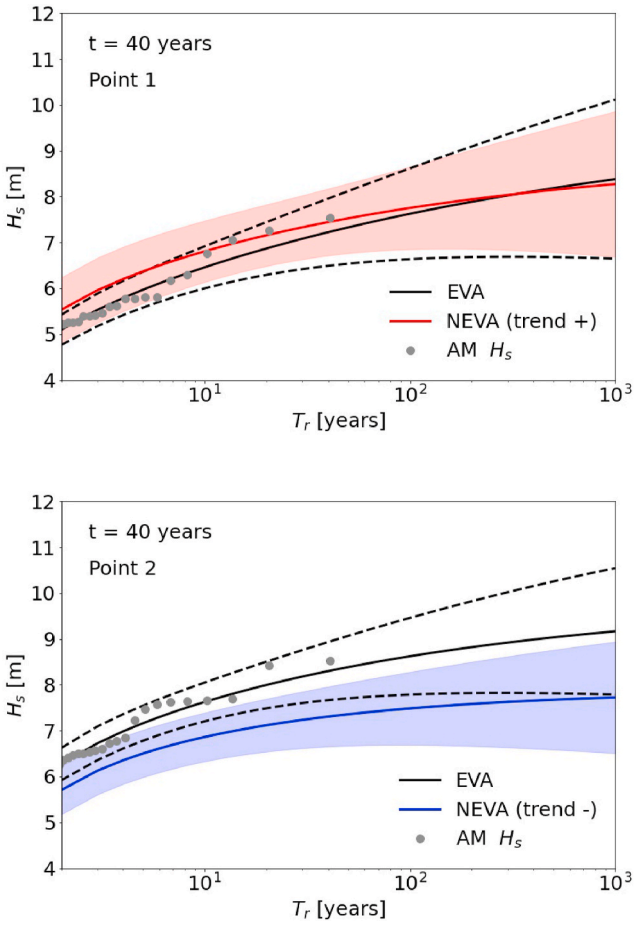

Fig. 9. Return levels of H_s according to EVA and NEVA approaches and reference time t equals to 40 years. Top panel: Point 1; bottom panel: Point 2.

Besides, the analysis of the return levels of H_s at Point 1 shows that including a trend in the extreme value analysis of a time series may lead to unexpected results. In particular, the differences in the shape parameters between the GEV and the N-GEV distributions imply the intersection of EVA and NEVA curves for T_r of about 300 years, meaning that the non-stationary analysis leads to smaller, i.e. less conservative, H_s, notwithstanding the underlying upward trend.

However, it needs to be pointed out that the intersection between EVA and NEVA curves at Point 1 is shifted to higher values of T_r if the reference time t in Eq. (1) is increased. This can be appreciated by looking at Fig. 10, which shows the return levels of H_s for T_r equal to 100 years, and reference time t varying between 0 and 100 years.

Here, it is evident how, in case of a positive trend (Point 1), H_s with $T_r = 100$ years resulting from NEVA is higher than the EVA predicted value for $t =$ around 33 years. On the other hand, when negative μ_t is considered (Point 2), EVA and NEVA results intersect at t lower than 10 years, owing to the higher magnitude of the trend. These results need further clarification. To adopt a model like that of Eq. (1), implies the assumption that μ_t is able to summarize the trend over the entire period in which the variable at hand is defined. In other words, once T_r is fixed, the corresponding H_s varies depending on the year of the computation. For example, for the hindcast data in this work, trends were computed over the 1979–2018 period. Therefore, $t = 40$ years indicates that the trend is referred exactly to the last year of the series, while lower values of t indicate that the computation of the return levels is carried out before 2018. On the contrary, to raise the value of t beyond 40 years implies to assume that the same trend will also characterize the future

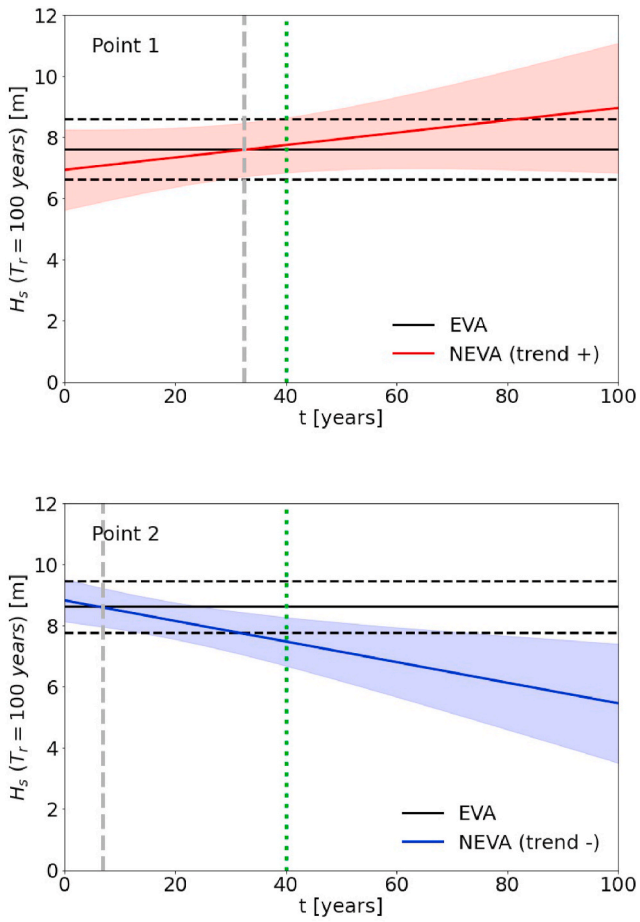

Fig. 10. Return levels of H_s for $T_r = 100$ years at varying reference time t. The green dotted lines refer to t equal to 40 years (i.e. the reference time used to compute the return levels of H_s in Fig. 9); the gray dashed lines indicate the values of t at which the EVA and the NEVA curves intersect. Top panel: Point 1; bottom panel: Point 2. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

extremes with respect to the last year of the hindcast data at the time of the analysis, i.e 2018. However, such hypothesis cannot be taken as granted. In principle, one should be careful with extrapolating estimated trends beyond the period of available data. Even if a significant trend is detected in the historical data, there is no reason to believe that this trend will continue into the future, and such extrapolation would be associated with large uncertainties. Hence, a more appropriate way of predicting future wave climate could be to make regression on some carefully selected covariates, for which there exist reliable future projections. For example, sea level pressure projections were used to estimate future wave climate in Wang et al. (2004) and Wang and Swail (2006), and regression on levels of atmospheric CO_2 were used in Vanem et al. (2014). Another alternative is to utilize physical wave models forced by output from global and regional climate models in order to predict future wave climate, as illustrated in Aarnes et al. (2017) and De Leo et al. (2021). Nevertheless, it might still be tempting to extend historical trends to illustrate the possible effect this could have on the future wave climate, acknowledging the large uncertainties this involves. In case of Point 2, a trend ≈ -0.03 m/years cannot be expected to indefinitely affect the data, as this would eventually result in negative values of AM H_s, which obviously would not have any physical meaning. Analogously, at Point 1 a monotone upward trend would eventually lead to unreliably high extreme waves. Based on these grounds, the trends computed at Point 1 and Point 2 were compared to the trends computed on the AM H_s series projected in the 2006–2100 period at the same locations (De Leo et al., 2021). At Point 2, b_{AM} was found to be equal to −2 mm/years, with $p_{MK} \approx 0.06$. These metrics indicate a future downward trend which is consistent to the trend computed on the 1979–2018 a.m. H_s, although the latter is one order of magnitude higher (i.e. more negative). At Point 1, the trend of the projected AM H_s was also found to be negative, thus not being consistent with the trend computed from the historical data. However, these trends are not significant on either case, according to the confidence intervals of b_{AM}; this can be appreciated in Fig. S2 in the Supplement, where the location with significant trends on both the historical series and the future projections are highlighted, and further confirm how at this location the outcomes of the non-stationary analysis must be carefully assessed.

3.5. Analysis of the use of the PMLM

For the locations where the N-GEV distribution parameters could not be computed via the MLM, i.e. the locations highlighted with black circle in Figs. 3–5, the PMLE described with Eq. (4) and Eq. (5) was employed. In this case, α and λ were set both equal to 1, but other possible values for these parameters could be investigated, and the best setting for P_f should be evaluated on a case by case basis. Yet, this choice allowed to compute estimates of μ_0, μ_t, σ, and ξ capable to fit the N-GEV distribution to the series of AM H_s where the MLE failed, as shown by the empirical frequency distributions in Fig. 11 (gray bins). For these particular locations, the pdf curves resulting from the MLE of the N-GEV distribution parameters (i.e. the yellow curves in the panels) show a strict upper limit that has no physical meaning. This is due to spurious estimates of ξ being lower than −1 according to the MLE, and is partially overcome through the use of the penalty function. Indeed, as shown by the panels of Fig. 11, the pdf's resulting from the use of $P_f(\xi)$ (highlighted with the green color) are not bounded by an upper limit, and better fit the frequency bins of the observed data.

In addition, the use of a PMLE yields less negative values of μ_t. This is in turn reflected in trends that are not significant in the vast majority of the locations under investigation, with the CI of μ_t showing a change of sign between the upper and lower bounds, as shown in Fig. 12. These results are more in line with the estimates of p_{MK} and b_{AM} than the trends related to the MLE for the same locations. On the contrary, the estimates of the N-GEV distribution parameters at Point 1 and Point 2 are not significantly affected by the use of $P_f(\xi)$, as shown in Table 3.

3.6. Environmental contours for the joint H_s - T_p probability

Finally, the ECs for Point 1 and Point 2 are presented in Fig. 13, for $T_r = 100$ years and reference times equal to 1, 40, and 100 years. Also in case of the bivariate analysis, it can be noticed how the contribution of the non-stationary term increases according to the reference time t, i.e. the ECs at Point 1 (positive trend) are shifted to high values of H_s and T_p for increasing t. Conversely, in case of negative trends (Point 2), the ECs tend to lower H_s-T_p pairs for increasing values of t. This further illustrates how NEVA strongly depends on t. Results also show that for the bivariate analysis, the shift between the EVA and the NEVA in ECs is far more relevant for Point 2 than for Point 1, i.e. where the underlying trend is higher and significant. However, in the latter point NEVA still selects a more severe set of conditions, although depending on t this might not have engineering relevance.

The analysis of the ECs leaves room for further considerations. First, it must be mentioned that a bivariate extreme sample is not always unambiguously defined. When the block maxima approach is employed, one choice is to consider one of the variables as the primary variable as done in this research, but other choices could be made (e.g. component-wise maxima, samples that maximizes some function of the two variables, etc.). Second, when a conditional model is employed, it is reasonable to assume the non-stationarity only in the marginal of the lead variable. However, to assume a conditioned distribution of T_p with

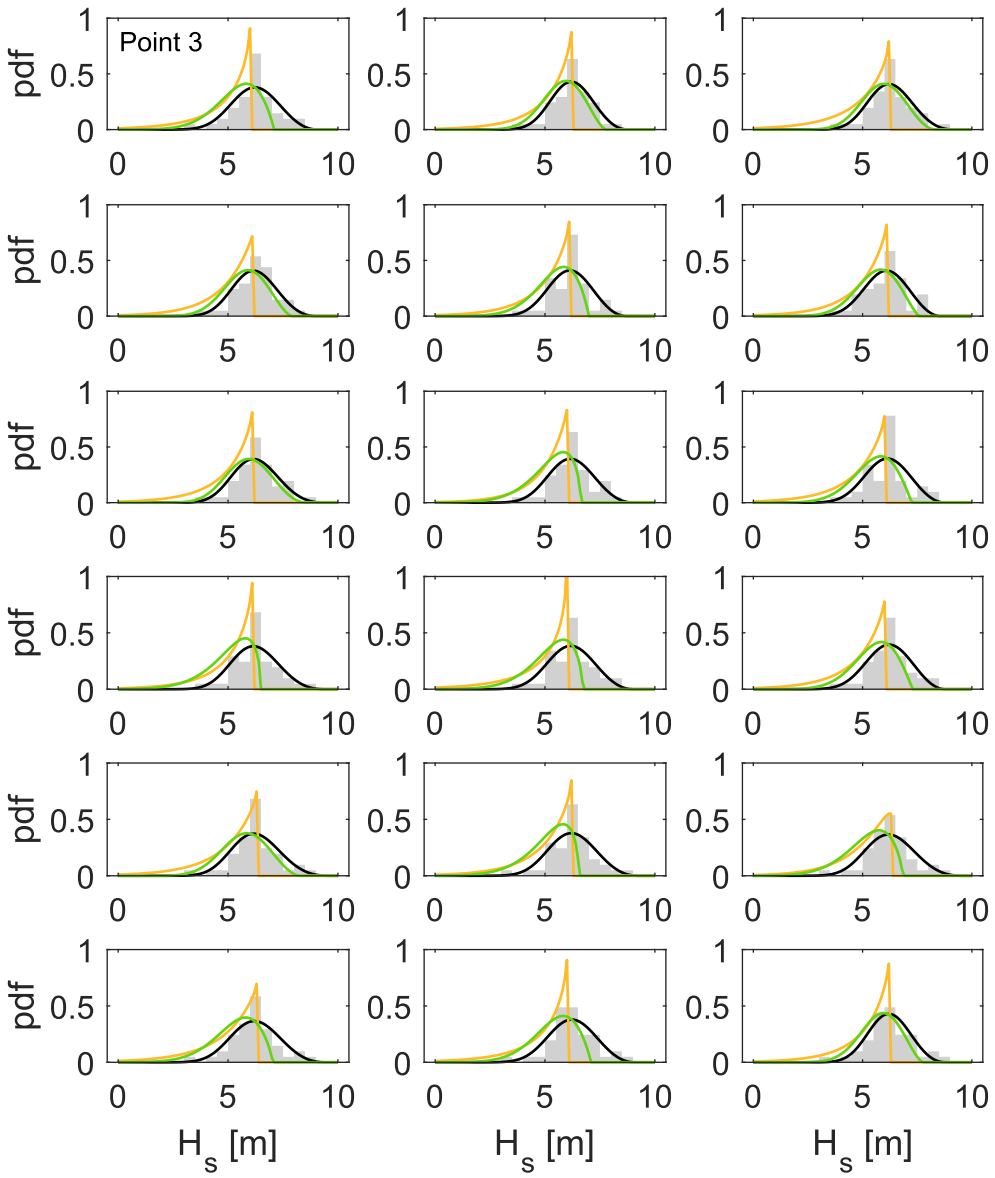

Fig. 11. Pdfs of the GEV (black curves) distribution with the N-GEV distributions resulting from MLE (yellow curves) and PMLE (green curves) for the cluster of points around Point 3 diverging from the main cloud in Fig. 2. Empirical frequencies of AM H_s are shown with the gray bins. t is equal to 40 years. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

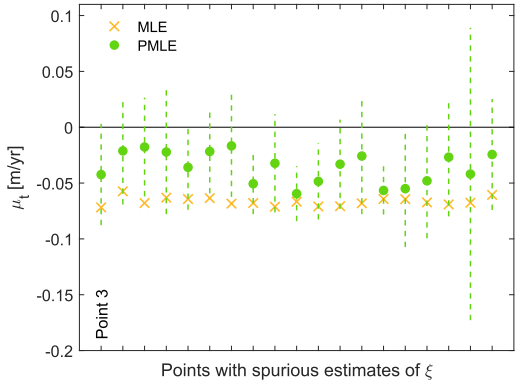

Fig. 12. Comparison of μ_t resulting from MLM and PMLM in the framework of NEVA. CI of the μ_t estimates are highlighted with the dashed lines.

Table 3
Estimates of the N-GEV distribution parameters using the MLM and the PMLM for Point 1 and Point 2.

		MLE	PMLE
Point 1	μ_0	4.39	4.41
	μ_t	0.02	0.02
	σ	0.93	0.92
	ξ	−0.25	−0.23
Point 2	μ_0	6.70	6.67
	μ_t	−0.03	−0.03
	σ	1.02	1.00
	ξ	−0.40	−0.38

respect to the non-stationary distribution for H_s, implies to assume that their variations in time are consistent. In fact, these parameters are not perfectly correlated, although in case of extreme events they are closely tied. Therefore, such hypothesis should be checked thoroughly to avoid

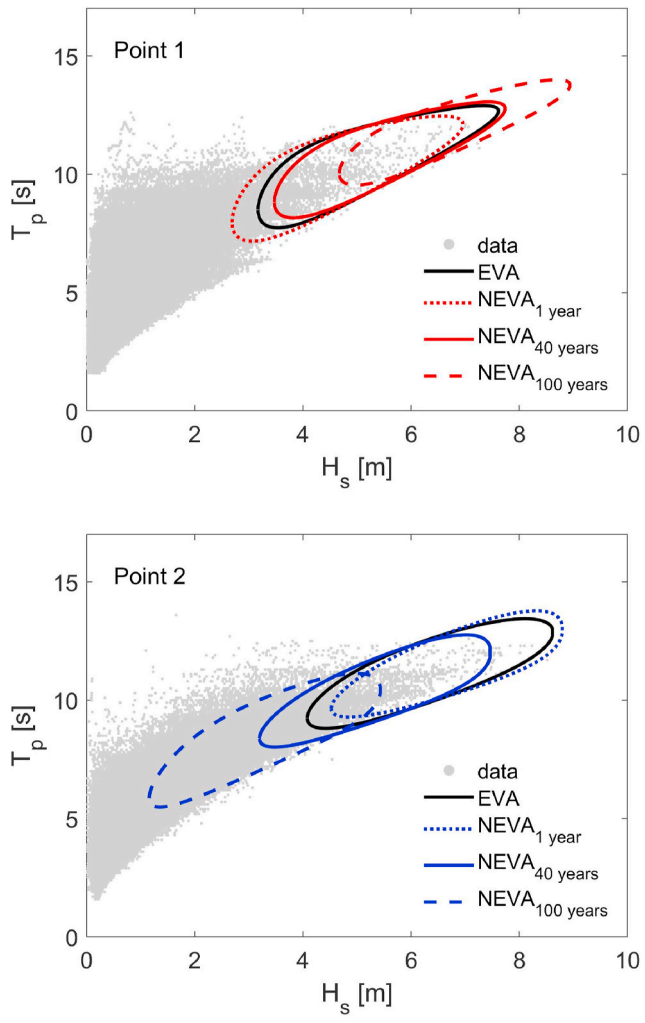

Fig. 13. IFORM ECs for $T_r = 100$ years according to EVA and NEVA approaches at Point 1 and Point 2 for different values of t.

that expected variation in the wave parameters were erroneously taken into account. In case of the two points here analysed, the trend analysis of the maxima T_p concurrent to the respective AM H_s reveal similar pattern between the series. That is, a fully consistent trend can be pointed out only in case of Point 2, as shown by the results summarized in Fig. S3 and Tab. S1 in the Supplement. On the contrary, if the AM T_p series were independently selected, no significant trend could be highlighted at neither locations.

Finally, it must be recalled that the ECs shown in Fig. 13 provide the exceedance probabilities for the selected subsamples, i.e. the AM. If the initial datasets were selected using the peak over threshold or the total distribution approaches instead of the AM method, this would possibly result in a different shape of the ECs related to the low values of the H_s-T_p pairs. In other words, the same values of H_s and T_p may be related to different joint probabilities depending on the initial selection of the extreme subsample, as shown by Vanem (2019). Therefore, care should be taken in reading the ECs and selecting the sampling method more suitable for the engineering application at hand.

4. Conclusions and final remarks

In this paper the use of the Non-stationary Extreme Value Analysis (NEVA) was analysed with the aim of developing of a methodology to establish when NEVA is to be preferred to traditional Extreme Value Analysis (EVA) for engineering applications. The Mediterranean Sea (MS) is ideal for such an investigation because of the very variable level of significance of the slope of the linear trends of Annual Maxima (AM) of the significant wave heigh (H_s) over the basins. The analysis of the correlation among the slope of the linear trend (b_{AM}), the time-varying part of the location parameter (μ_t) of the Non-stationary Generalized Extreme Value (N-GEV) probability distribution, and the p-value of the Mann-Kendall test (p_{MK}) allowed to conclude that only when μ_t and b_{AM} indicate a consistent trend above the confidence intervals (CI), the performance of NEVA is superior to that of EVA. This is quantified by the Aikake Information Criteria (AIC), which is higher for the N-GEV distribution only when the consistency of the three parameters occurs. In the MS, this occurs in computational nodes such as Point 2, which show significant negative trends, and is consistent with the projections of future wave climate and trends expected in the basin (De Leo et al., 2021).

In conclusion, it is suggested to first compute both b_{AM} and μ_t and check that both have consistent sign within the respective CI; when this occurs together with low values of p_{MK}, NEVA might be preferred to traditional EVA. The paper has also shown that N-GEV distribution parameters can be obtained using the Penalized Maximum Likelihood Method (PMLM) that provides Penalized Maximum Likelihood Estimators (PMLE) when the Maximum Likelihood Method (MLM) fails to compute the required parameters. On the other hand, the use of PMLM in the form here proposed does not affect the computation of the N-GEV parameters when the MLM is successful. This result indicates that PLME may be used as a general method for this type of applications, however the generality of this conclusion needs to be further verified.

The computation of the return levels of H_s is central to marine and coastal engineering practice. This work showed that return levels obtained with NEVA are particularly sensitive to the reference time t used in the N-GEV distribution (see Eq. (1)), this was further shown by drawing the Environmental Contours (ECs) to describe the return levels of pairs H_s and peak period (T_p). As expected, the difference in return level increases with t in the univariate N-GEV distribution, while the effect of t is more complex when ECs are used. Results show very well that the inclusion of the more extreme combinations of H_s and T_p within the chosen EC is deeply influenced by the choice of t. The sensitivity of NEVA to t prompts the need of carefully considering two factors when applying NEVA in coastal and marine applications. The first is the duration of the time series used to infer the trend, the second is the validity of the assumption of linear trend for the whole duration of the time series available, above all for multi-decadal time series of met-ocean variables that are nowadays available.

CRediT authorship contribution statement

Francesco De Leo: Conceptualization, Data curation, Formal analysis, Software, Investigation, Methodology, Writing – original draft, Writing – review & editing, Visualization. **Giovanni Besio:** Supervision, Data curation, Writing – review & editing. **Riccardo Briganti:** Conceptualization, Investigation, Supervision, Writing – original draft. **Erik Vanem:** Conceptualization, Formal analysis, Software, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The computation of GEV and N-GEV parameters took advantage of the *ismev* package within the R environment https://cran.r-project.org/web/packages/ismev/ismev.pdf. The annual maxima H_s and T_p for

Point 1 and Point 2 and a snippet of the code to compute and plot the ECs are available for download at the following repository: 10.5281/zenodo.4148732.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.coastaleng.2021.103896.

References

Aarnes, O.J., Reistad, M., Breivik, Ø., Bitner-Gregersen, E., Ingolf Eide, L., Gramstad, O., Magnusson, A.K., Natvig, B., Vanem, E., 2017. Projected changes in significant wave height toward the end of the 21st century: Northeast a tlantic. J. Geophys. Res.: Oceans 122, 3394–3403.

Akaike, H., 1998. Information theory and an extension of the maximum likelihood principle. In: Selected Papers of Hirotugu Akaike. Springer, pp. 199–213.

Barbero, R., Fowler, H., Lenderink, G., Blenkinsop, S., 2017. Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions? Geophys. Res. Lett. 44, 974–983.

Besio, G., Briganti, R., Romano, A., Mentaschi, L., De Girolamo, P., 2017. Time clustering of wave storms in the Mediterranean Sea. Nat. Hazards Earth Syst. Sci. 17, 505–514.

Cheng, L., AghaKouchak, A., 2014. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. Sci. Rep. 4, 7093.

Coles, S., Bawa, J., Trenner, L., Dorazio, P., 2001. An Introduction to Statistical Modeling of Extreme Values, vol. 208. Springer.

Coles, S.G., Dixon, M.J., 1999. Likelihood-based inference for extreme value models. Extremes 2, 5–23.

Corbella, S., Stretch, D.D., 2012. Predicting coastal erosion trends using non-stationary statistics and process-based models. Coast. Eng. 70, 40–49.

De Leo, F., Besio, G., Mentaschi, L., 2021. Trends and variability of ocean waves under rcp8. 5 emission scenario in the mediterranean sea. Ocean Dynam. 71, 97–117.

De Leo, F., De Leo, A., Besio, G., Briganti, R., 2020. Detection and quantification of trends in time series of significant wave heights: an application in the mediterranean sea. Ocean. Eng. 202, 107155.

Ferrari, F., Besio, G., Cassola, F., Mazzino, A., 2020. Optimized wind and wave energy resource assessment and offshore exploitability in the mediterranean sea. Energy 190, 116447.

Haselsteiner, A.F., Ohlendorf, J.H., Wosniok, W., Thoben, K.D., 2017. Deriving environmental contours from highest density regions. Coast. Eng. 123, 42–51.

Haver, S., 1985. Wave climate off northern Norway. Appl. Ocean Res. 7, 85–92.

Haver, S., 1987. On the joint distribution of heights and periods of sea waves. Ocean. Eng. 14, 359–376.

Haver, S., Winterstein, S.R., 2009. Environmental contour lines: a method for estimating long term extremes by a short term analysis. Trans. - Soc. Nav. Archit. Mar. Eng. 116, 116–127.

Hollander, M., Wolfe, D.A., Chicken, E., 2013. Nonparametric Statistical Methods, vol. 751. John Wiley & Sons.

Hosking, J.R.M., Wallis, J.R., 2005. Regional Frequency Analysis: an Approach Based on L-Moments. Cambridge university press.

Huseby, A.B., Vanem, E., Natvig, B., 2013. A new approach to environmental contours for ocean engineering applications based on direct Monte Carlo simulations. Ocean. Eng. 60, 124–135.

Huseby, A.B., Vanem, E., Natvig, B., 2015. Alternative environmental contours for structural reliability analysis. Struct. Saf. 54, 32–45.

Kendall, M.G., 1955. Rank Correlation Methods.

Leira, B.J., 2008. A comparison of stochastic process models for definition of design contours. Struct. Saf. 30, 493–505.

Lucio, D., Tomás, A., Lara, J., Camus, P., Losada, I., 2020. Stochastic modeling of long-term wave climate based on weather patterns for coastal structures applications. Coast. Eng. 161, 103771.

Luke, A., Vrugt, J.A., AghaKouchak, A., Matthew, R., Sanders, B.F., 2017. Predicting nonstationary flood frequencies: evidence supports an updated stationarity thesis in the united states. Water Resour. Res. 53, 5469–5494.

Mackay, E.B., Challenor, P.G., Bahaj, A.S., 2011. A comparison of estimators for the generalised pareto distribution. Ocean. Eng. 38, 1338–1346.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica. Journal of the Econometric Society 245–259.

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

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

Mentaschi, L., Vousdoukas, M., Voukouvalas, E., Sartini, L., Feyen, L., Besio, G., Alfieri, L., 2016. The transformed-stationary approach: a generic and simplified methodology for non-stationary extreme value analysis. Hydrol. Earth Syst. Sci. 20, 3527–3547.

Renard, B., Sun, X., Lang, M., 2013. Bayesian methods for non-stationary extreme value analysis. In: Extremes in a Changing Climate. Springer, pp. 39–95.

Sartini, L., Cassola, F., Besio, G., 2015. Extreme waves seasonality analysis: an application in the Mediterranean Sea. J. Geophys. Res.: Oceans 120, 6266–6288.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379–1389.

Şen, Z., 2011. Innovative trend analysis methodology. J. Hydrol. Eng. 17, 1042–1046.

Şen, Z., 2013. Trend identification simulation and application. J. Hydrol. Eng. 19, 635–642.

Serinaldi, F., Kilsby, C.G., 2015. Stationarity is undead: uncertainty dominates the distribution of extremes. Adv. Water Resour. 77, 17–36.

Smith, R.L., 1985. Maximum likelihood estimation in a class of nonregular cases. Biometrika 72, 67–90.

Theil, H., 1992. A rank-invariant method of linear and polynomial regression analysis. In: Henri Theil's Contributions to Economics and Econometrics. Springer, pp. 345–381.

Vanem, E., 2015. Non-stationary extreme value models to account for trends and shifts in the extreme wave climate due to climate change. Appl. Ocean Res. 52, 201–211.

Vanem, E., 2016. Joint statistical models for significant wave height and wave period in a changing climate. Mar. Struct. 49, 180–205.

Vanem, E., 2017. A comparison study on the estimation of extreme structural response from different environmental contour methods. Mar. Struct. 56, 137–162.

Vanem, E., 2018. A simple approach to account for seasonality in the description of extreme ocean environments. Marine Systems & Ocean Technology 13, 63–73.

Vanem, E., 2019. Environmental contours for describing extreme ocean wave conditions based on combined datasets. Stoch. Environ. Res. Risk Assess. 33, 957–971.

Vanem, E., 2020. Bivariate regional extreme value analysis for significant wave height and wave period. Appl. Ocean Res. 101, 102266.

Vanem, E., Bitner-Gregersen, E.M., 2015. Alternative environmental contours for marine structural design—a comparison study. J. Offshore Mech. Arctic Eng. 137.

Vanem, E., Huseby, A.B., Natvig, B., 2014. Bayesian hierarchical spatio-temporal modelling of trends and future projections in the ocean wave climate with a CO2 regression component. Environ. Ecol. Stat. 21, 189–220.

Wang, X.L., Swail, V.R., 2006. Climate change signal and uncertainty in projections of ocean wave heights. Clim. Dynam. 26, 109–126.

Wang, X.L., Zwiers, F.W., Swail, V.R., 2004. North atlantic ocean wave climate change scenarios for the twenty-first century. J. Clim. 17, 2368–2383.