scientific reports

OPEN

Multivariate sea storm hindcasting and design: the isotropic buoy-ungauged generator procedure

Gianfausto Salvadori[1]✉, Giuseppe Roberto Tomasicchio[2], Felice D'Alessandro[3], Letizia Lusito[2] & Antonio Francone[4]

The present work provides indications for assessment of wave climate and design of structures at sea at ungauged sites, both critical issues in Ocean sciences. The paper is of methodological nature and of global worldwide applicability. It shows how suitable wave hindcasting relations can be exploited in order to provide sea storm scenarios at an ungauged (Target) location useful for design purposes: in particular, only geographical information and the knowledge of another gauged (Source) buoy are used. Several are the novelties introduced. (i) New hindcasting relations are derived. (ii) A full statistical model is set up for the Target area, whereas traditional hindcasting simply transfers time series from a gauged to an ungauged site: this gives the possibility to appropriately deal with design and hazard assessment at the Target location. (iii) The multivariate behavior of non-independent random variables is properly modelled by using the Theory of Copulas. As an illustration, a number of case studies is investigated, involving four pairs of buoys which, given their positions and exposures, are representative of a wide variety of sea states and conditions, as well as of different wave generation mechanisms.

Oceanographers and Maritime Engineers are primarily concerned with wave conditions at a specific location, in order to provide indications useful for the assessment of wave climate and the design of structures at sea at ungauged sites. The present work proposes a novel methodology to produce wave data at ungauged sites, based on historical data observed at a gauged site. In particular, the problem tackled here is: *what information about the sea state can be provided for design and hazard assessment at a Target location, if that site is ungauged?*

Hindcasting methods first appeared during World War II, when it became crucial to "translate" weather forecasts into expected wave conditions[1–3]. Briefly, wave hindcasting consists of using historical wind records and geographical information about the sea area of interest, in order to predict wave heights, periods and directions according to, e.g., wind duration and strength, and fetches: see, e.g., the SMB empirical method[4,5]. Later, more refined hindcasting guidelines were provided by[6] exploiting the results of the JONSWAP experiment[7].

In general, wave hindcasting techniques require minimal input information and computational effort, and provide ready and handy estimates. In turn, they may represent a valuable alternative/complementary approach with respect to, e.g., global wave models such as those developed at NOAA/ECMWF[8,9]—which, however, remain somewhat limited concerning the processes involved in the wave generation. Note that the outputs of hindcasting procedures can always be checked against global wave models.

After about 50 years since the JONSWAP trial, here the corresponding wave hindcasting relations take a new life. The approach proposed entails several novelties—see "Methods". On the one hand, a full (statistical) model is developed for the ungauged Target site—instead of simply generating a limited data set of the Target variables of interest. On the other hand, the case of non-independent variables is properly dealt with, by using the Theory of Copulas. Furthermore, new hindcasting relations are derived.

In the present work, an original *isotropic Buoy-Ungauged Generator* procedure (hereinafter, *iBUG*) is outlined, in order to provide valuable hindcasting estimates of the sea state at an ungauged (Target) buoy, only exploiting some information given by another (Source) gauged buoy and the Effective Fetches (a geographical information

[1]Dipartimento di Matematica e Fisica, Università del Salento, 73100 Lecce, Italy. [2]Dipartimento di Ingegneria dell'Innovazione, Università del Salento, 73100 Lecce, Italy. [3]Dipartimento di Scienze e Politiche Ambientali, Università degli Studi di Milano, 20133 Milan, Italy. [4]Dipartimento di Ingegneria Civile, Ambientale, del Territorio, Edile e di Chimica, Politecnico di Bari, 70125 Bari, Italy. ✉email: gianfausto.salvadori@unisalento.it

always available at any site). The procedure is of global worldwide applicability, and works under minimal (and "natural") assumptions, such as that the climatic event generating the sea storms has the same (comparable) features at both the Source and the Target sites. Note that also the predictions of global wave models implicitly thrust a "homogenous" sea state within each cell of the simulation grid used.

Here, a pragmatic approach is adopted. The idea is to check whether the iBUG procedure is able to provide valuable Target estimates of the design values of some relevant variables associated with given Return Periods (hereinafter, *RP*). The rationale is that such values are those used in practice for the design of structures at sea and the assessment of the corresponding hazards. In particular, in the present work, the Significant Wave Height *H* (in *m*) and the Sea Storm Duration *D* (in *h*) are used as the variables characterizing a sea storm—see "Materials".

For this purpose, several buoys, subject to different sea state conditions and exposures, are considered on a pair-wise basis, and the following calculations are carried out.

1. Firstly, one buoy (say, A) is regarded as the Source of information (i.e., it is gauged), whereas another buoy (say, B, the Target one) is assumed to be ungauged (i.e., no data collected at the corresponding site are used to compute the iBUG estimates). Then, the iBUG procedure is utilized to set up a statistical model at B only using (i) the information available at A (i.e., the local observations of the variables *H* and *D*), and (ii) the ratios of the Effective Fetches at A and B. Finally, iBUG sea storm scenarios are generated at B via Monte Carlo techniques, and suitable *H* and *D* design values (associated with given RP's) are compared with those actually observed at B.
2. The role of the buoys A and B is swapped: B is taken as the Source buoy, and A becomes the Target buoy. Then, the same calculations as above are carried out.

Most importantly, note that *only* the Source data, and the information about the Source and Target Fetches, are used to compute the iBUG Target estimates, since the Target buoy is assumed to be ungauged. This corresponds to a situation of "complete ignorance" of the sea state at the Target site, typical of many practical cases. The "double-blind" estimating approach used here (i.e., A vs. B, and B vs. A) gives the possibility to carry out an accurate roundtrip survey of the performance of the iBUG technique.

Results and discussion

In this section, the outcomes of the analyses are illustrated and commented. Hereinafter, the following acronyms are used: MC for Monte Carlo, GoF for Goodness-of-Fit, and CI for Confidence Interval.

In general, waves generated in deep water can be fetch-limited, duration-limited, or constitute a fully developed sea[10,11]. In order to thoroughly illustrate the applicability of the iBUG procedure, in the following four pairs of buoys A and B are considered: given their positions and exposures, these are representative of a wide variety of sea states and conditions, as well as of the wave generation mechanisms mentioned above. The features of the data sets are briefly introduced in "Materials", and fully presented in the accompanying Supplementary Information (SI) document. The pairs investigated are the following: the corresponding maps are plotted in Fig. 1.

1. *Gibraltar (Mediterranean Spain).* The *Puertos del Estado* buoys "6100417" and "6100198".
    URL: https://www.emodnet-physics.eu/Map
2. *North Sea.* The buoys located at *Elbe* (Federal Maritime and Hydrographic Agency-BSH, Germany) and *Fino* (R&D Centre, Kiel University of Applied Sciences, Germany).
    URL: https://www.emodnet-physics.eu/Map
3. *Bay of Biscay (Atlantic Spain).* The *Puertos del Estado* buoys "62024" and "62025".
    URL: https://www.emodnet-physics.eu/Map
4. *Hawaii (USA).* The *NOAA* buoys "51003" and "51004".
    URL: www.ndbc.noaa.gov

The figures presented in the sequel all have the same structure. Firstly, seven design Return Periods of practical interest are fixed: namely, 10, 20, 35, 50, 75, 100, and 200 years. Secondly, the two variables *H* and *D* are considered, and for each variable, 99% Confidence Intervals of the corresponding design values associated with the RP's mentioned above are drawn. In particular: (i) *black* lines and markers are CI's constructed using the Target data (the buoy B); (ii) *blue* lines and markers are CI's constructed via Traditional Hindcasting techniques using the Source data (the buoy A); (iii) *red* lines and markers are CI's constructed via the iBUG procedure using the Source data (the buoy A). In cases (i) and (ii), the CI's are Bootstrap ones, while the iBUG CI's are computed via Monte Carlo simulations—see "The iBUG procedure in practice". Should two confidence intervals intersect, then it would not be (statistically) possible to reject the Null hypothesis that the corresponding point estimates of the design values are compatible (comparable). Below, the four case studies are presented and discussed.

The Gibraltar case. The plots of interest are presented in Fig. 2. Considering the Source "6100417" buoy vs. the Target "6100198" buoy, the iBUG CI's always intersect the Target ones, whereas the Traditional Hindcasting CI's fail (at least) for RP's larger than 20 years, both for *H* and *D*. Instead, swapping the two buoys (i.e., Source "6100198" vs. Target "6100417"), there are no failures, although the Traditional Hindcasting CI's show kind of a limit approximation considering *H* and very large RP's.

The North Sea case. The plots of interest are presented in Fig. 3. Considering the Source "Elbe" buoy vs. the Target "Fino" buoy, the iBUG CI's always intersect the Target ones, whereas the Traditional Hindcasting CI's
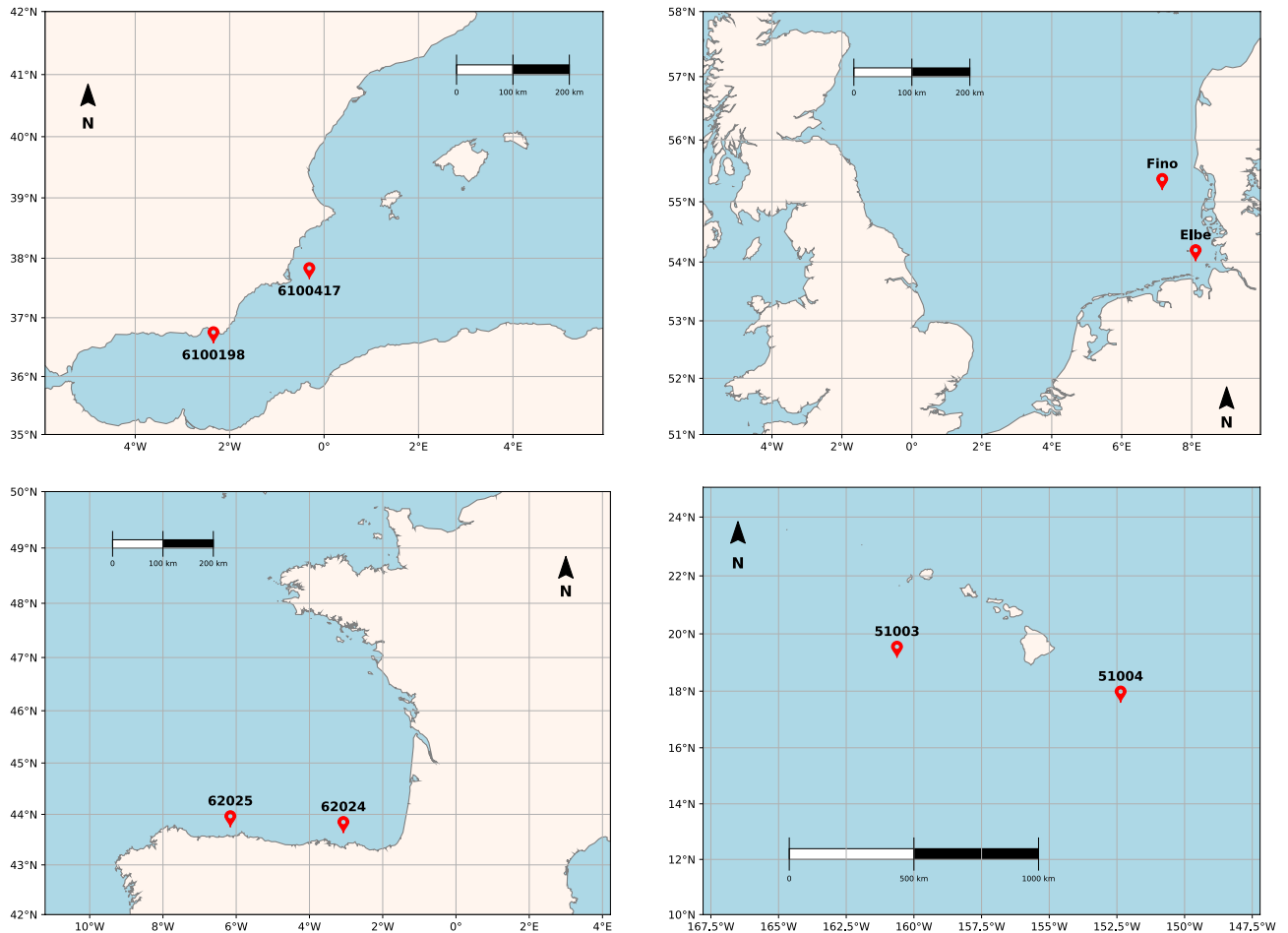
Figure 1. Maps. Locations of the pairs of buoys investigated. (*topleft*) The Gibraltar (Mediterranean Spain) case: buoys "6100417" and "6100198". (top right) The North Sea case: buoys "Elbe" and "Fino". (bottom left) The Bay of Biscay (Atlantic Spain) case: buoys "62024" and "62025". (bottom right) The Hawaii (USA) case: buoys "51003" and "51004".

fail (at least) for RP's larger than 20 years, considering the duration D. Instead, swapping the two buoys (i.e., Source "Fino" vs. Target "Elbe"), there are no iBUG failures, whereas the Traditional Hindcasting CI's fail, considering H and D, (at least) for RP's larger than 20 years .

The Bay of Biscay case. The plots of interest are presented in Fig. 4. Considering the Source "62024" buoy vs. the Target "62025" buoy, the iBUG CI's always intersect the Target ones, whereas the Traditional Hindcasting CI's almost always fail, considering the duration D. Similarly, swapping the two buoys (i.e., Source "62025" vs. Target "62024"), there are no iBUG failures, whereas the Traditional Hindcasting fails considering the duration D.

The Hawaii case. The plots of interest are presented in Fig. 5. Considering the Source "51004" buoy vs. the Target "51003" buoy, the iBUG CI's always intersect the Target ones, whereas the Traditional Hindcasting CI's almost always fail considering H. Instead, swapping the two buoys (i.e., Source "51003" vs. Target "51004"), there are no iBUG failures, whereas the Traditional Hindcasting CI's almost always fail considering H.

Interpretation of the results. As already mentioned above, the case studies considered in this work are representative of a wide variety of sea states and conditions, as well as of wave generation mechanisms.

The Traditional Hindcasting approach is nothing more than a *deterministic* technique: given a set of N data at a gauged Source buoy S, only a set of N estimates can be generated at a Target ungauged site T. Moreover, the Traditional Hindcasting may only yield design estimates based on a single sample (i.e., the data available at the Source buoy): in practice, only a single Target sea state scenario can be generated.

On the contrary, the iBUG procedure transfers a full statistical model at the target site T, i.e. it constructs a suitable Probability Space for the ungauged buoy, which also accounts for the dependence structure (i.e., the multivariate Copula) ruling the joint dynamics of the variables at play: note that this would not be possible via Traditional Hindcasting, being a deterministic univariate technique without any link to Statistics.
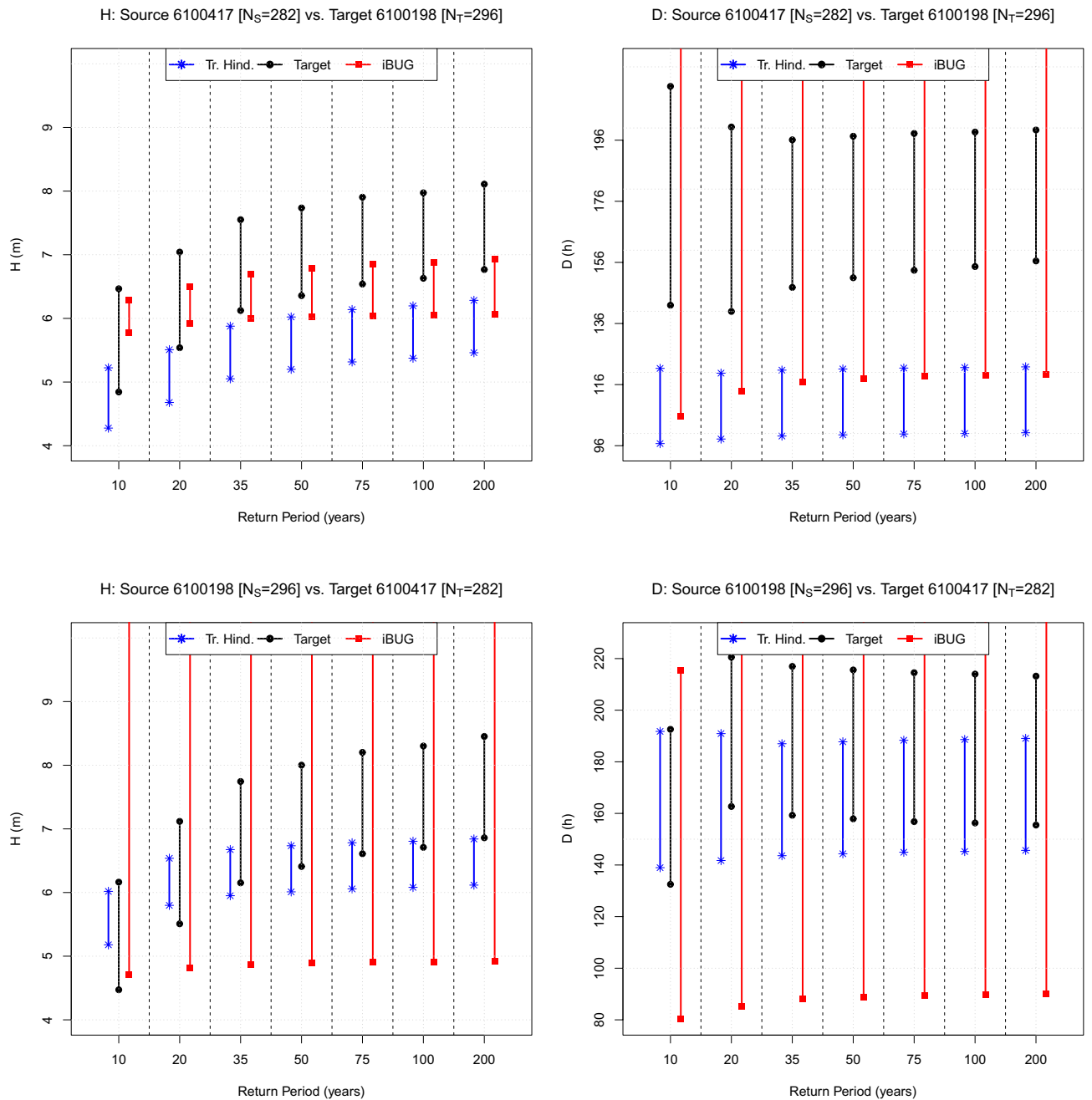
Figure 2. Comparison of Design values: the Gibraltar case. The plot shows the CI's of the Target design values (black) associated with the RP's indicated on the horizontal axis. The blue and red CI's represent, respectively, the corresponding estimates computed via the Traditional Hindcasting and the iBUG techniques. The variables of interest (H and D) are indicated in the title, as well as the labels of the buoys considered and the Source (N_S) and Target (N_T) sample sizes.

Most importantly, the statistical model used by the iBUG procedure may generate a number of possible (independent) Target scenarios via MC techniques: this entails that the Probability Space associated with the Target sea storms can be more appropriately and accurately explored, especially considering rare extreme events associated with large design Return Periods (difficult to observe via a single sample).

It is also interesting to note that, in general, the Monte Carlo iBUG CI's are larger than the corresponding Target and Traditional Hindcasting ones, constructed via Bootstrap techniques. The reason is that the latter are based on a single sample of limited size *N*, whereas the iBUG CI's represent the information extracted from many (independent) samples of size *N*—here, 100,000 MC simulations. Clearly, this gives the possibility to better account for the uncertainty and randomness of the Target sea states.

Overall, the results shown in "The Gibraltar case"–"The Hawaii case" indicate that the iBUG procedure is able to construct sea storm scenarios always compatible (comparable) with the sea state observed at the Target
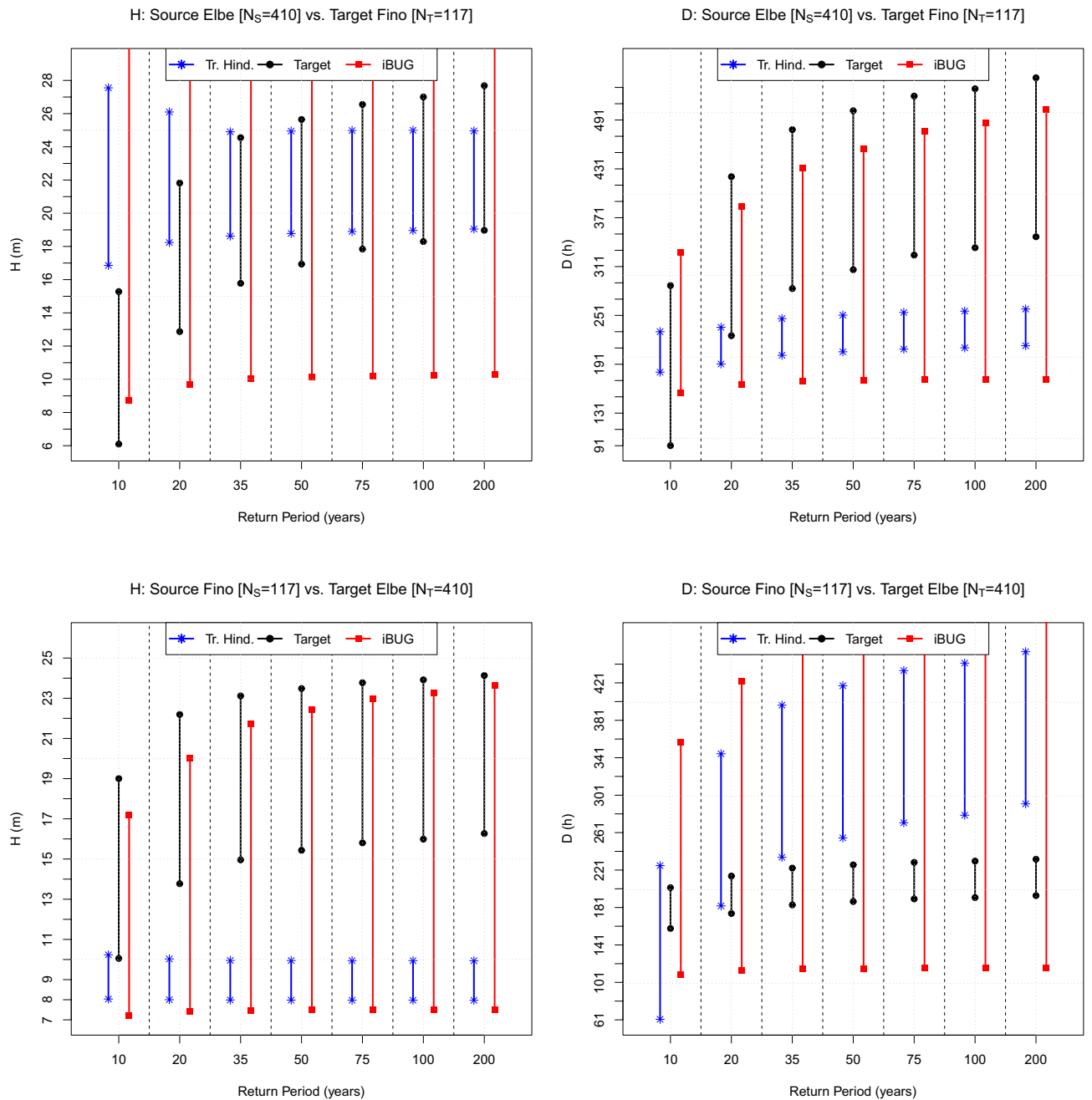
Figure 3. Comparison of Design values: the North Sea case. The plot shows the CI's of the Target design values (black) associated with the RP's indicated on the horizontal axis. The blue and red CI's represent, respectively, the corresponding estimates computed via the Traditional Hindcasting and the iBUG techniques. The variables of interest (H and D) are indicated in the title, as well as the labels of the buoys considered and the Source (N_S) and Target (N_T) sample sizes.

site (it is successful in 100% of the cases), whereas (as expected) the Traditional Hindcasting often fails: actually, in 9 cases out of 16, more than 50% of the tests, and especially considering Return Periods larger than 20 years (i.e., the ones of interest in practical applications).

Materials

The available sea storm Time Series provide information about the significant wave height and the wave direction. In order to prevent problems with possible serial dependencies and correlations, in the present study an event-based approach is used (adopting a maritime version of the *Run Method* outlined in[12]), by considering suitable Sea Storm occurrences to describe the sea state.

Here, a sea storm is characterized via the Significant Wave Height H (in m) and the Sea Storm Duration D (in h), as well as a Direction parameter θ (in $deg N$): thus, the relevant triples are of type ($H, D; \theta$)—however,
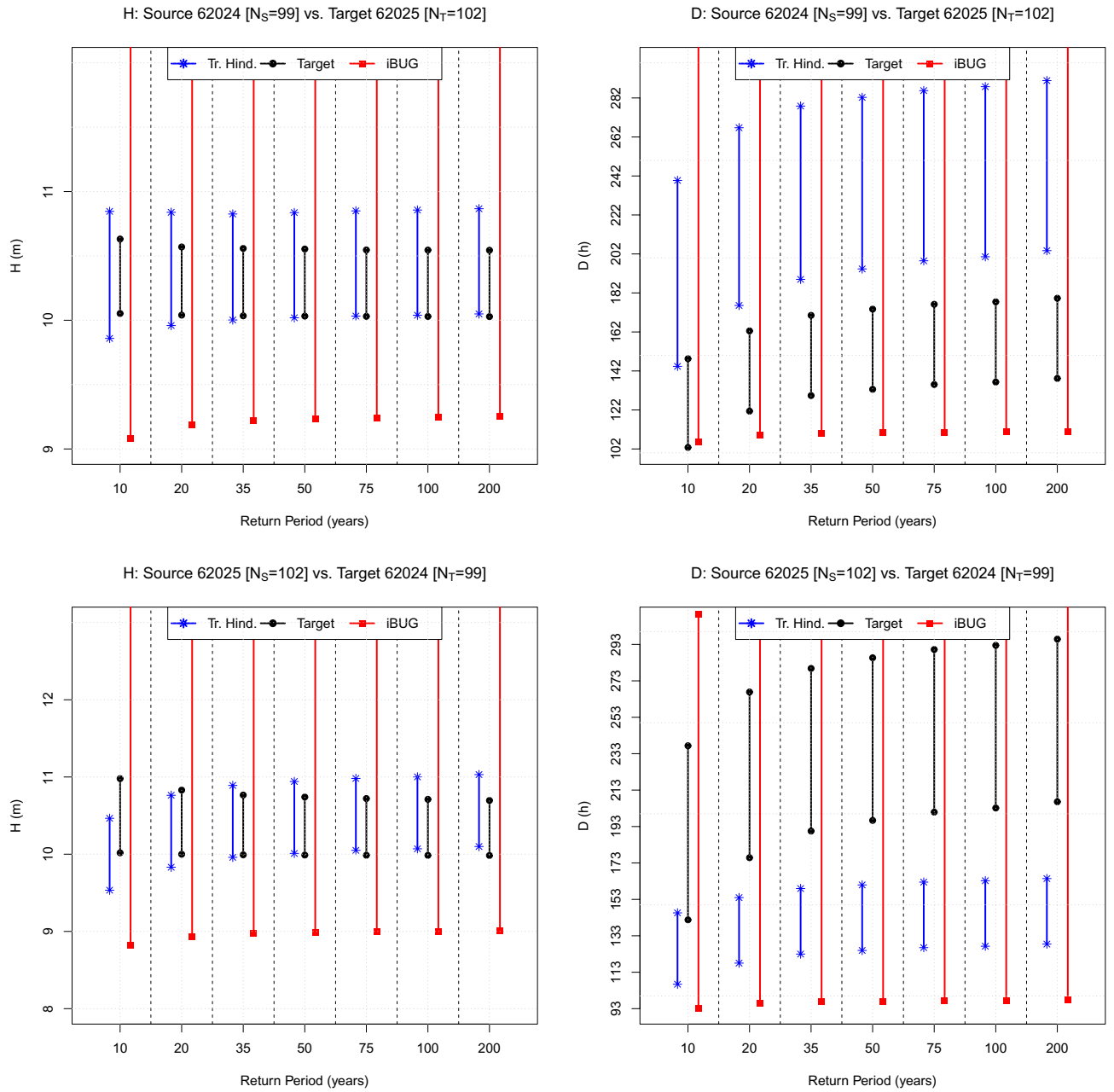
Figure 4. Comparison of Design values: the Bay of Biscay case. The plot shows the CI's of the Target design values (black) associated with the RP's indicated on the horizontal axis. The blue and red CI's represent, respectively, the corresponding estimates computed via the Traditional Hindcasting and the iBUG techniques. The variables of interest (H and D) are indicated in the title, as well as the labels of the buoys considered and the Source (N_S) and Target (N_T) sample sizes.

exploiting the multivariate approach outlined in this paper, other sets of variables can be considered, e.g. by using the Wave Period P, such as $(H, P; \theta)$, $(D, P; \theta)$, or $(H, D, P; \theta)$.

Practically, a sea storm starts when H crosses upwards a given threshold H_0, and ends when H persists below the same level for at least I_0 hours: for a graphical illustration see Fig. 6, showing a standard Triangular Wave Model[13] used to identify a sea storm sequence.

Here, the threshold H_0 corresponds to a suitable (large) percentile p_H of the empirical distribution of wave heights, in order to focus the analyses on extreme events. The choice of the percentile represents a trade-off between (i) the purpose of investigating strong sea storms (and hence p_H should be as large as possible) and (ii) the resulting size N of the extracted sample of sea storms (which should also be as large as possible for carrying out valuable statistical surveys): clearly, increasing p_H reduces N, and vice-versa. In the present work, the following values of p_H have been used: 95% for the Gibraltar case, 90% for the North Sea case, 95% for the Bay of Biscay case, and 75% for the Hawaii case.
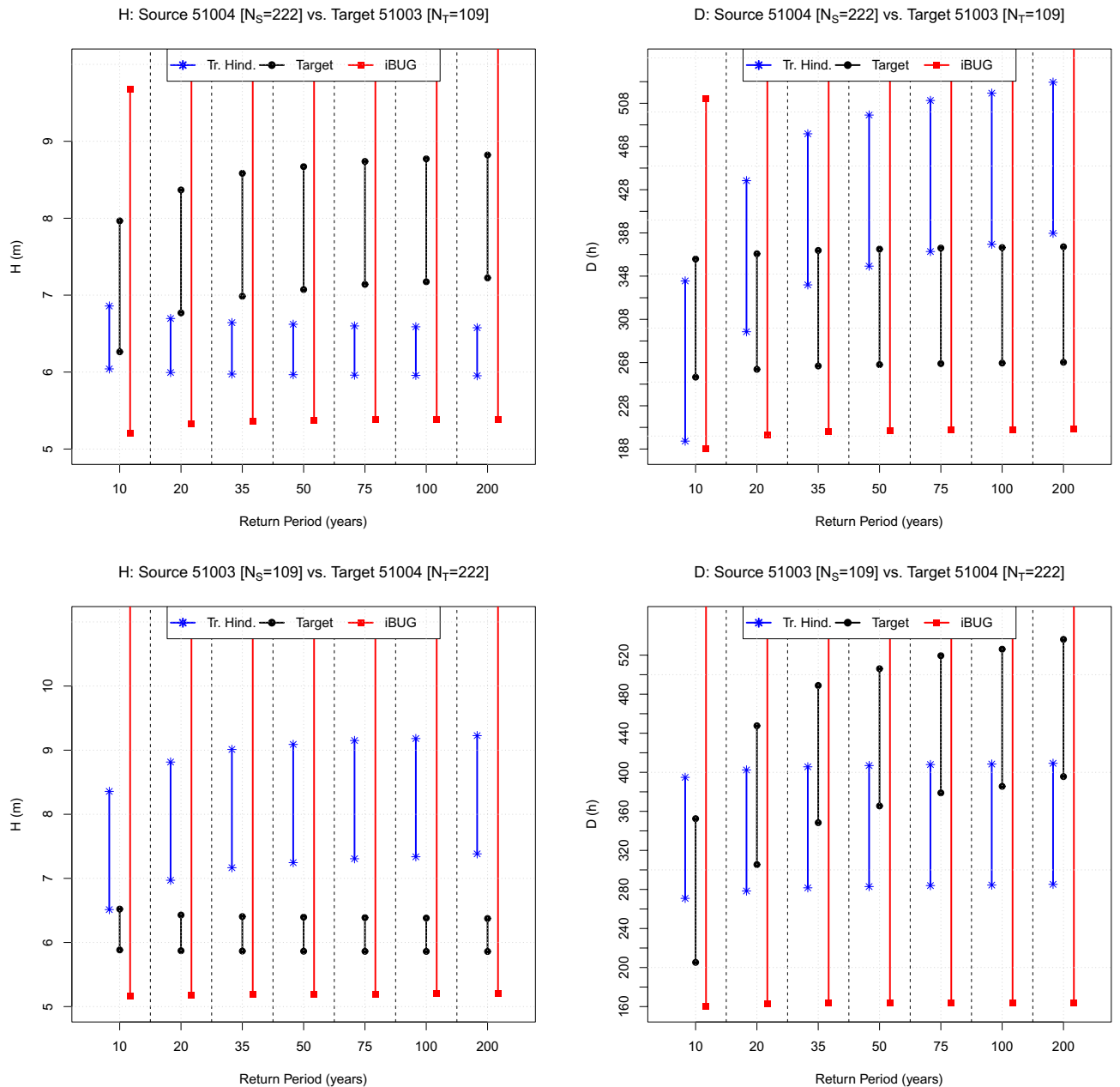
Figure 5. Comparison of design values: the Hawaii case. The plot shows the CI's of the Target design values (black) associated with the RP's indicated on the horizontal axis. The blue and red CI's represent, respectively, the corresponding estimates computed via the Traditional Hindcasting and the iBUG techniques. The variables of interest (H and D) are indicated in the title, as well as the labels of the buoys considered and the Source (N_S) and Target (N_T) sample sizes.

Also, a minimum inter-storm temporal distance I_0 is used to temporally separate successive sea storms: this guarantees the physical independence of the events of interest. Actually, according to the tests provided by[14,15] and implemented in the R package npcp, the distributions of the variables H and D extracted show no temporal Change-Points (distributional stationarity). Here, $I_0 = 48$ h for the Hawaii (accounting for the specific meteorology of the site), and $I_0 = 24$ h in all the other cases.

Once the thresholds H_0 and I_0 are chosen, the duration D of a sea storm is immediately computed, while the significant wave Height is taken as the maximum H during the whole sea storm, in order to investigate its extreme dynamics—and, correspondingly, θ is taken as the direction associated with the maximum Height recorded.

As anticipated in "Results and discussion", in the present study four pairs of buoys are investigated. The accompanying SI shows the time series of H and D, the corresponding univariate fits, the fits of bivariate families of Copulas, and the Effective Fetches—see Sects. S.1–S.4. Here, eight standard univariate laws have been tested (namely, Weibull, Exponential, Gamma, Lognormal, Chi-square, F, GEV, GPD), as well as a dozen of copulas families and their survival variants (namely, Frank, Clayton, Gumbel, Joe, Tawn, Galambos, Husler–Reiss, t-Student). As a result all fits, both univariate and multivariate, are valuable, since the MC p-values of the GoF
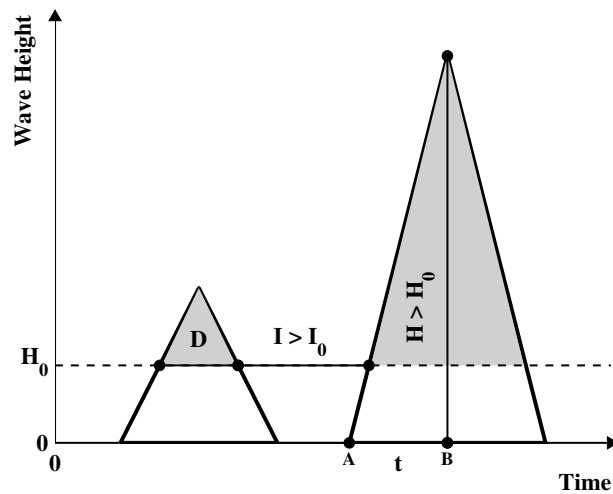
Figure 6. Triangular sea storm model. Graphical illustration of the Triangular Sea Storm model. Shown are: the wave height threshold H_0, the inter-storm period I larger than the threshold I_0, the sea storm duration D, a sea storm wave height H larger than H_0, and the approximate duration t of the climatic event generating the sea storm (the temporal segment $\overline{AB}$)—see text. The shaded regions indicate successive triangular sea storms.

tests are large. It is worth noting that, in all cases, H and D are fairly well associated (concordant): in fact, the Kendall's τ and Spearman ρ measures of association[16] are large, with negligible p-values—this corresponds to the failure of a test of independence for the pair (H, D).

Methods

In this section, the mathematical framework adopted is thoroughly outlined. The present work is of *methodological* nature and of global worldwide applicability.

Multivariate hindcasting. For the sake of illustration, here the (multivariate) analyses will involve the variables H and D. However, the same procedures can straightforwardly be used by considering other ones. The general idea (and novelty) is to make use of the observations available at a gauged Source buoy S, and then exploit the hindcasting relations provided by the JONSWAP experiment[7], in order to construct a multivariate statistical model (accounting for the dependence between H and D) at an ungauged Target site T. Most importantly, such a model can be used to simulate suitable time series and scenarios of sea storms at T (see "Multivariate hindcasting"): clearly, the latters can be of any desired size, even much larger than the size of the sample observed at S—e.g. appropriate for specific design and hazard assessment purposes at T (see below).

The basic assumption of hindcasting procedures is that the climatic event generating the sea storms has the same (viz., "comparable") features at both the Source and the Target site. In turn, the wave climate at the Source buoy S is supposed to be the same as—or comparable to—the (unobserved) one at the Target buoy T. In practice, such a condition/constraint can only be approximately satisfied: thus, the hindcasting outcomes must be taken with care. Here, the inspiring engineering principle is that a crude prediction at the ungauged site is always better than no information at all.

Below, we use Copulas as a basic tool for dealing with a multivariate framework. Indeed, a convenient way to model the joint behavior of a set $X_1, \ldots, X_d$ of random variables is to utilize Copulas. A copula is a multidimensional function describing, and mathematically formalizing, the dependence structure ruling the joint random dynamics of the X_i's. Note that, in Literature, it has been increasingly recognized that the use of a multivariate framework is practically mandatory in Maritime Engineering, given the fact that often (if not always) the variables at play are non-independent—see[17–19], as well as the typical case studies presented in "Materials".

Indeed, recent advances in Mathematics show how copulas may represent an efficient tool to investigate the statistical behavior of dependent variables. For a theoretical introduction to copulas, see[20–22]; for a practical approach, see[16,23,24]. In particular, elementary guidelines for using copulas in maritime/hydrological applications are outlined in[18,19,25–28]. Freeware R routines are provided and illustrated in[29,30].

According to Sklar's Theorem representation[31], the joint distribution

$$\mathbf{F}_{1,\ldots,d}(x_1, \ldots, x_d) = \mathbf{P}(X_1 \leq x_1, \ldots, X_d \leq x_d)$$

of the X_i's, with marginal distributions F_i's, can be written as, for all $(x_1, \ldots, x_d) \in \mathbf{R}^d$,

$$\mathbf{F}_{1,\ldots,d}(x_1, \ldots, x_d) = \mathbf{C}_{1,\ldots,d}(F_1(x_1), \ldots, F_d(x_d)), \tag{1}$$

where $\mathbf{C}_{1,\ldots,d}$ is the copula associated with the random vector $(X_1, \ldots, X_d)$. Practically, a multivariate model $\mathbf{F}_{1,\ldots,d}$ can be readily constructed by (i) fitting suitable univariate laws F_i's on the marginals X_i's, and (ii) fitting an appropriate copula $\mathbf{C}_{1,\ldots,d}$ on the observed $(X_1, \ldots, X_d)$'s.

Hindcasting based on JONSWAP formulas. Based on the ratio of Effective Fetches at a Source and a Target site, the JONSWAP hindcasting formulas[11] make it possible to determine the wave climate at a site where no observations are available. Clearly, the outcomes of this methodology (hereinafter, *Transfer Method (TM)*), as well as the ones of traditional hindcasting methods, must be taken with care.

Let θ be a given direction: for instance, $\theta = 0°$ N indicates the North. The TM consists of the following directional relations, where the subscripts "S" and "T" label, respectively, the Source and Target variables of interest:

$$H_T/H_S = \varrho_\theta^{1/2}, \tag{2}$$

$$P_T/P_S = \varrho_\theta^{1/3}, \tag{3}$$

$$t_T/t_S = \varrho_\theta^{2/3}, \tag{4}$$

where t is the duration of the climatic event generating a sea storm, and

$$\varrho_\theta = F_T(\theta)/F_S(\theta) \tag{5}$$

is the ratio of the Target $F_T(\theta)$ and Source $F_S(\theta)$ Fetches associated with the direction θ.

Incidentally, as a novelty, the use of a Triangular Wave Model (see "Materials" and Fig. 6) may provide an estimate of t. In fact, it may be reasonable to assume that t equals half the duration of the triangular sea storm: viz., the storm "grows" until the wind feeds it (i.e., the temporal segment $\overline{AB}$ in Fig. 6), and then fades away. In turn, via the elementary proportion $t \div H = D/2 \div (H - H_0)$, an approximate value of t can be computed as

$$t = \frac{D\,H}{2\,(H - H_0)}. \tag{6}$$

As a consequence, exploiting Eqs. (4)–(6), a (new) rough estimate of the Sea Storm Duration ratio can be derived as

$$D_T/D_S = \varrho_\theta^{2/3}. \tag{7}$$

Similar relations can be gleaned by using different Sea Storm Wave models.

The iBUG procedure. As a novelty with respect to traditional hindcasting methods, here we do not simply transfer data bases from a Source to a Target site. Rather, by using Eqs. (2)–(7), we show how to convey a full statistical model from a (gauged) Source buoy S to an (ungauged) Target one T. In particular, here we focus the attention on the pair (H, D): however, there are no theoretical limitations to consider other sets of variables.

The advantages of transferring a full statistical model, instead of a single set of data, are obvious: in fact, once the (multivariate) statistics of the sea storms at the Target buoy T is set up, it is then possible to generate Target scenarios and samples of any suitable size (e.g., via MC simulations—see below), and thus provide useful information for design and hazard assessment at the ungauged site.

Before proceeding, the following important point needs to be stressed. In the present work, we adopt an *isotropic invariance principle*: viz., the Source marginals and copula do not depend upon the direction θ. Instead, the Target distributions will be ruled by the directional Fetch ratios ϱ_θ's—see, below, Eqs. (8), (9), and (11). Actually, there are no theoretical reasons for requiring such a directional distributional invariance, rather only empirical ones. In fact, without such an assumption, specific univariate laws and copulas should be fitted for each θ. Clearly, this makes little sense in practical applications (even introducing a coarse discretization of the directions θ's) since the directional sample sizes would almost surely be insufficient to provide a statistically significant number of observations, both for fitting the Source marginals and copula, and for carrying out univariate and multivariate Goodness-of-Fit tests. Practically, the iBUG procedure starts with an unique Source statistical model, and "drives" it to the Target buoy by exploiting the directional information provided by the Fetches' ratios.

Now, denoting by G_S^H and G_S^D, respectively, the univariate marginal distributions of the Source sea storm variables H_S and D_S, it is immediate to calculate the corresponding laws G_T^H and G_T^D of the Target variables H_T and D_T:

$$G_T^H(x; \theta) = \mathbf{P}_\theta(H_T \leq x) = \mathbf{P}\left(\varrho_\theta^{1/2} H_S \leq x\right) = G_S^H\left(x\,\varrho_\theta^{-1/2}\right), \tag{8}$$

$$G_T^D(y; \theta) = \mathbf{P}_\theta(D_T \leq y) = \mathbf{P}\left(\varrho_\theta^{2/3} D_S \leq y\right) = G_S^D\left(y\,\varrho_\theta^{-2/3}\right), \tag{9}$$

for all $x, y \in \mathbf{R}$. Evidently, given a direction θ, the Target distributions are re-scaled versions of the Source ones via the Fetch ratio ϱ_θ.

Then, by using Sklar's Theorem, Eqs. (8)–(9), and the isotropic invariance, the bivariate probability distributions $\mathbf{F}_S$ and $\mathbf{F}_T$ of the pair (H, D) of, respectively, the sea storms at the Source and Target buoys, can be written as

$$\mathbf{F}_S(x, y) = \mathbf{C}\left(G_S^H(x), G_S^D(y)\right), \tag{10}$$

$$\mathbf{F}_T(x, y; \theta) = \mathbf{C}\left(G_S^H\left(x\,\varrho_\theta^{-1/2}\right), G_S^D\left(y\,\varrho_\theta^{-2/3}\right)\right), \tag{11}$$

for all $x, y \in \mathbf{R}$, where $\mathbf{C}$ is the copula associated with the Source pairs (H_S, D_S)'s. Evidently, one of the advantages of transferring a full statistical model is that it may provide a complete set of information concerning the joint random behavior of the sea storms at the Target buoy.

Estimates of design values. Traditionally, in Maritime Engineering (as well as in Terrestrial Hydrology) design values are computed via a Return Period approach. First, a design RP is chosen (e.g., 20, 50, or 100 years, depending on the structure of interest), and then the associated design value(s) of the variable(s) of interest are estimated using the available data. In the following, the general notion of RP introduced in[25,27] is adopted. Given an event E, the associated RP T_E is

$$T_E = \mu/\mathbf{P}(E), \qquad (12)$$

where μ is the mean inter-arrival time of the events E's in the time series considered. Note that μ provides the time unit (e.g., years) in which T_E should be expressed.

Considering events like $E = \{X > x^*\}$—the ones of typical interest in applications—where x^* is a suitable threshold for the variable X (e.g., X could be the Significant Wave Height H or the Duration D), it is easy to compute design values x^* once a design RP T^* is assigned. In fact, Eq. (12) would become $T^* = \mu/\mathbf{P}(X > x^*)$, and inverting it yields

$$x^* = F_X^{-1}(1 - \mu/T^*), \qquad (13)$$

where F_X is the distribution function of X.

The iBUG procedure in practice. Below it is shown how to simulate Target sea storm scenarios via the iBUG technique.

1. A set of design Return Periods T's is fixed: here, T equals 10, 20, 35, 50, 75, 100, 200 years.
2. The original Source data are used to extract a set of N_S (directional) sea storms $(H_S, D_S; \theta)$'s.
   Note that, adopting a Traditional Hindcasting approach, and using Eqs. (2) and (7), then N_S estimates of the pairs $(H_T, D_T; \theta)$'s at the Target site would be given by

$$\begin{cases} \tilde{H}_{T,i} = \varrho_{\theta_i}^{1/2} H_{S,i} \\ \tilde{D}_{T,i} = \varrho_{\theta_i}^{2/3} D_{S,i} \end{cases}, \qquad (14)$$

   with $i \in 1, \ldots, N_S$, and no further steps would be required.
3. Suitable univariate marginals G_S^H and G_S^D, and a copula $\mathbf{C}$, are fitted over the Source data base, irrespectively of the different directions θ's associated with the sea storms, due to the isotropic invariance. As a statistical law for the directions θ's, the corresponding empirical distribution function $\hat{F}_\theta$ is used.
4. A set of N_S pairs (u_i, v_i)'s is simulated from the Source copula $\mathbf{C}$. Similarly, a corresponding independent set of N_S directions θ_i's is extracted from $\hat{F}_\theta$.
5. Via Sklar's Theorem, and the Probability Integral Transform (i.e., by inverting G_T^H and G_T^D), a sample (x, y)'s of N_S iBUG Target pairs $(H_T, D_T; \theta)$'s is generated as

$$\begin{cases} x_i = \left(G_T^H\right)^{-1}(u_i) = \varrho_{\theta_i}^{1/2} \left(G_S^H\right)^{-1}(u_i) \\ y_i = \left(G_T^D\right)^{-1}(v_i) = \varrho_{\theta_i}^{2/3} \left(G_S^D\right)^{-1}(v_i) \end{cases}, \qquad (15)$$

   with $i = 1, \ldots, N_S$, where $\left(G_\bullet^H\right)^{-1}$ and $\left(G_\bullet^D\right)^{-1}$ denote, respectively, the inverses of the distribution functions $G_\bullet^H$ and $G_\bullet^D$. Thus, the pairs (x, y)'s represent a possible Target sea storm scenario according to the iBUG method.
6. Steps 4–5 are repeated K times, in order to provide a suitable number of independent Monte Carlo simulations: K may be chosen according to the desired accuracy of the iBUG estimates at the Target site for design and hazard assessment purposes (e.g., 10^5 for large Return Periods, as in the present work).
7. Finally, using Eq. (13), and the K independent iBUG samples generated at Step 6, a set of K design values x^*'s and y^*'s is computed for each design RP T chosen at Step 1. In turn, appropriate Confidence Intervals can be constructed for the iBUG estimates. Similarly, via suitable Bootstrap techniques[32–34], apt CI's can also be set up for the corresponding Target and Traditional Hindcasting design values.

Data availability
The data used are available at the sites indicated in the paper, and specific requests may be sent to the Corresponding Author.

References
1. Sverdrup, H. V. & Munk, W. H. Empirical and theoretical relations between wind, sea and swell, 27, 823–827. *Trans. Am. Geophys. Union* **27**, 823–827 (1946).

2. Sverdrup, H. V. & Munk, W. H. *Wind, Sea and Swell: Theory of Relations for Forecasting.* (U.S. Navy Hydrographic Office, Washington, publication no. 601 edn, 1947).
3. Kinsman, B. *Wind Waves: Their Generation and Propagation on the Ocean Surface* (Prentice-Hall, Englewood Cliffs, 1965).
4. Bretschneider, C. L. The generation and decay of wind waves in deep water. *Trans. Am. Geophys. Union* **33**, 381–389 (1952).
5. Bretschneider, C. L. Revisions in wave forecasting: deep and shallow water. in *Proceedings of the 6th Conference on Coastal Engineering (Gainsville, Palm Beach and Miami Beach, Florida)*, 30–67 (1958).
6. CERC. *Shore Protection Manual.* (Department of the Army, U.S. Army Corps of Engineers, Washington, 1984).
7. Hasselmann, K. *et al.* Measurement of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). in *Ergänzungsheft zur Deutshen Hydrographischen Zeitschrift Reihe A(8), Nr. 12* (Deutsches Hydrographisches Institut, Hamburg, 1973).
8. Tolman, H. L. A third-generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents. *J. Phys. Oceanogr.* **21**, 782–797 (1991).
9. The WAMDI Group. The wam model—A third generation ocean wave prediction model. *J. Phys. Oceanogr.* **18**, 1775–1810 (1988).
10. Kinsman, B. *Wind Waves: Their Generation and Propagation on the Ocean Surface* (Dover Phoenix Editions, Dover, 2002).
11. Kamphuis, J. W. Introduction to coastal engineering and management. in *Advanced Series on Ocean Engineering*, Vol. 30 (World Scientific, 2010).
12. Yevjevich, V. *An Objective Approach to Definitions and Investigations of Continental Hydrologic Droughts.* (Colo. State Univ., Fort Collins, Colo., Hydrology Paper, 23 edn., 1967).
13. Boccotti, P. Wave mechanics for ocean engineering. in *Elsevier Oceanography Series*, Vol. 64 (Elsevier Science, Amsterdam, 2000).
14. Holmes, M., Kojadinovic, I. & Quessy, J.-F. Nonparametric tests for change-point detection à la Gombay and Horváth. *J. Multivar. Anal.* **115**, 16–32 (2013).
15. Kojadinovic, I. *npcp: Some Nonparametric CUSUM Tests for Change-Point Detection in Possibly Multivariate Observations.* R package version 0.1-9 (2017).
16. Salvadori, G., De Michele, C., Kottegoda, N. & Rosso, R. Extremes in nature. An approach using copulas. in *Water Science and Technology Library Series*, Vol. 56 (Springer, Dordrecht, 2007). ISBN: 978-1-4020-4415-1.
17. Salvadori, G., Tomasicchio, G. R. & D'Alessandro, F. Multivariate approach to design coastal and off-shore structures. *J. Coast. Res.* **65 (Special Issue)**, 386–391. https://doi.org/10.2112/SI65-066.1 (2013).
18. Salvadori, G., Tomasicchio, G. R. & D'Alessandro, F. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. *Coast. Eng.* **88**, 1–14. https://doi.org/10.1016/j.coastaleng.2014.01.011 (2014).
19. Salvadori, G., Durante, F., Tomasicchio, G. R. & D'Alessandro, F. Practical guidelines for the multivariate assessment of the structural risk in coastal and off-shore engineering. *Coast. Eng.* **95**, 77–83. https://doi.org/10.1016/j.coastaleng.2014.09.007 (2015).
20. Nelsen, R. *An Introduction to Copulas* 2nd edn. (Springer, New York, 2006).
21. Joe, H. Dependence modeling with copulas. in *CRC Monographs on Statistics & Applied Probability* (Chapman & Hall, London, 2014).
22. Durante, F. & Sempi, C. *Principles of Copula Theory* (CRC/Chapman & Hall, Boca Raton, 2015).
23. Genest, C. & Favre, A. Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.* **12**, 347–368 (2007).
24. Salvadori, G. & De Michele, C. On the use of copulas in hydrology: theory and practice. *J. Hydrol. Eng.* **12**, 369–380 (Special Issue: Copulas in Hydrology). https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(369) (2007).
25. Salvadori, G. & De Michele, C. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resour. Res.* **40**, W12511. https://doi.org/10.1029/2004WR003133 (2004).
26. De Michele, C., Salvadori, G., Passoni, G. & Vezzoli, R. A multivariate model of sea storms using copulas. *Coast. Eng.* **54**, 734–751 (2007).
27. Salvadori, G., De Michele, C. & Durante, F. On the return period and design in a multivariate framework. *Hydrol. Earth Syst. Sci.* **15**, 3293–3305. https://doi.org/10.5194/hess-15-3293-2011 (2011).
28. Salvadori, G., Durante, F., De Michele, C., Bernardi, M. & Petrella, L. A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. *Water Resour. Res.* **52**, 3701–3721. https://doi.org/10.1002/2015WR017225 (2016).
29. Hofert, M., Kojadinovic, I., Maechler, M. & Yan, J. *copula: Multivariate Dependence with Copulas.* R package version 0.999-18 edn (2018).
30. Hofert, M., Kojadinovic, I., Mächler, M. & Yan, J. *Elements of Copula Modeling with R* (Springer, New York, 2018).
31. Sklar, A. Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231 (1959).
32. Efron, B. Better bootstrap confidence intervals (with discussion). *J. Am. Stat. Assoc.* **82**, 171–200 (1987).
33. DiCiccio, T. & Efron, B. Bootstrap confidence intervals (with discussion). *Stat. Sci.* **11**, 189–228 (1996).
34. Davison, A. C. & Hinkley, D. V. Bootstrap methods and their application. in *Cambridge Series in Statistical and Probabilistic Mathematics*, Vol. 1 (Cambridge University Press, Cambridge, 1997).

Acknowledgements

Author contributions

G.R.T., F.DA, L.L., and A.F. conceived the investigation and prepared the data; G.S. sketched the theoretical framework and carried out the statistical analyses. All Authors surveyed the results and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-77329-y.

Correspondence and requests for materials should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.