Extreme wind events in a complex maritime environment: Ways of quantification

Platon Patlakas [a], George Galanis [a,b], Nicolas Barranger [a], George Kallos [a]

[a] University of Athens, Department of Physics, Atmospheric Modeling and Weather Forecasting Group, University Campus, Bldg. PHYS-V, Athens 15784, Greece
[b] Hellenic Naval Academy, Section of Mathematics, Hatzikiriakion, Piraeus 18539, Greece

A R T I C L E   I N F O

A B S T R A C T

The rising demand of energy consumption in isolated locations such as in islands leads in the expansion of on and off-shore wind farms. The optimization of the structural design of wind turbines for such applications requires a risk analysis that is made by using the definition of return periods of extreme events with respect to the lifespan of wind turbines. This work is focusing on the estimation and the analysis of extreme wind speeds by means of the corresponding return periods based on two methods: the Peaks Over Threshold and the Annual Maxima. In addition, different methodologies and tools are tested in order to achieve more accurate results. The data used for the application are both: observations (measurements from Met Stations located on Greek islands) and modeling (a 10-year model hindcast database). The sensitivity test results were used to adjust the methodologies and make 50-year extreme wind speed maps for Northeast Mediterranean (focusing on the sea and the islands). The outcome should be used as a guide for on and off-shore wind energy applications and other construction activities.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

In the structural design of wind turbines, probabilistic approaches of risk assessment are adopted in order to optimize the constructions in terms of profit and durability and avoid time and cost overruns that can compromise the economic viability of the project. To this end different approaches are used for estimating conditions that contribute to or form potential threats for wind turbines such as extreme wind speed. Such conditions can be evaluated through risk analysis, a methodology that combines the magnitude and the likelihood of occurrence of an extreme event.

In this direction, risk can be expressed through the concept of return period that is a statistical estimator for extreme phenomena reoccurrence based on data of shorter range. Although, there are different approaches proposed for the estimation of the magnitude and reoccurrence interval of events, Annual Maxima and Peaks Over Threshold methods (Coles, 2001) meet great acceptance for their effectiveness. Cook (1985) suggested that for Annual Maxima method, extreme wind speed is often well represented by Gumbel distribution. The same author (1982) used the dynamic pressure to achieve a faster convergence and better distribution fitting. A more recent study was held by Larsén et al. (2011) where an extreme wind speed atlas is created based on the principles of Generalized Extreme Value (GEV) theory and the Annual Maxima (AM) method. Peaks Over Threshold (POT) methodology is employed for studies based on smaller time series and the use of exponential is

supported (Abild et al., 1992). As in the first case, the wind speed square is found to fit better, especially in areas with low wind speeds and in cases where the wind speed distribution is not skewed enough for an exponential quick convergence to the distribution tail (Caires and Sterl, 2004; Galambos, 1987; Cook, 1982). These extreme value analysis methods are also used to more targeted studies of extremes based on similar characteristics such as the year season or the direction (Cook, 1982). The necessity for bigger datasets that do not violate the principles of Extreme Value (EV) theory led to the introduction of other methodologies such as the Method of Independent Storms (MIS) (Harris, 1998) and the EV theory based on the largest annual events (Smith, 1986). At the same time different approaches are proposed by Lopatoukhin et al. (2000) for the estimation of extreme wind wave heights such us the Initial Distribution Method (IDM). Breivik et al. (2014) studied wind and wave extremes using large ensembles and computed a non-parametric Direct Return Estimate (DRE) from the tail of the fitted distribution function. This was used for the estimation of the 100-year marine wind speed over the Globe.

The purpose of this work is to study ways of estimating the likelihood of occurrence of extreme wind speed events and to apply them over the area of Greece. More specifically, different methods and tools are applied over multiple datasets and the convergence (or not) of the results is further discussed. Through this procedure the uncertainties on the estimation of extreme

winds are presented and the appropriateness of the methods/tools used is studied.

Towards this direction, the two approaches mentioned above and used are the Peaks Over Threshold (POT) and the Annual Maxima method (AM) (Palutikof et al., 1999; Larsén et al., 2011). The data employed for the application consists of both measurements from nine different stations and modeled time series. The selected area is characterized by a complex land–water distribution and the existence of several islands (smaller or bigger) where wind energy production is considerable. The reason for this selection is the local climatic and geographic characteristics. More specifically, an almost constant Boundary Layer (BL) depth ($\sim$300 m) is observed over the sea during day and night. This BL depth can change dramatically from night to day over islands that can lead to strong vertical wind components (over the islands) during day-time especially on the warmer period of the year. In addition, close to laminar flow conditions are observed near the water surface while more turbulent over the land and downwind. Moreover, strong dominant effects can take place at the wake part of the islands and at the same time wind shading can be the case in areas surrounded by island clusters.

The modeled time series are extracted from a database created by the Atmospheric Modeling and Weather Forecasting Group (AM&WFG) of the University of Athens within the framework of Marina Renewable Integrated Application Platform project (Kallos et al., 2012) (MARINA – http://forecast.uoa.gr/proj_marina.php). The advantage of the use of the database is the fact that a wide area is represented and wind speed is recorded in five levels within the boundary layer.

2. Materials and methods

Two of the most known methodologies to estimate return periods, are the Block Maxima and the Peaks Over Threshold method. A quick description of them is provided below.

2.1. Block (Annual) Maxima method

The Block Maxima method uses the GEV theory (Jenkinson, 1955). For this application the time series are divided in same-size blocks and the maximum value of each block is used to create the dataset for the application.

The choice of the block size is of major importance since a very small can lead to overestimation and increased bias. On the other hand, very large blocks will lead to smaller datasets, large variability (Coles, 2001) and rather unreliable estimation. These reasons led to the use of annual blocks (Annual Maxima) because shorter periods may violate the principles of the GEV theory (Coles, 2001). The sample created by selecting the annual maximum values, is used to fit a distribution that belongs to the GEV family.

It is widely accepted that wind speed is well described by the Weibull distribution (Hennessey, 1977), while the extremes (AM) are often approached by the first type of GEV (Cook, 1985). The later, combined with the fact that Gumbel's Probability Density Function:

$$F(x) = \frac{1}{\alpha} e^{-z - e^{-z}} \tag{2.1.1}$$

where $z = \frac{x - \beta}{\alpha}$, β = location parameter, α = scale parameter, requires the estimation of only two parameters, led to this selection.

The estimation of the parameters of the fitting distribution is based on two methods. The first one is the Maximum Likelihood (ML) Method (Cramér, 1946; Hazewinkel, 2001) and the second is the Method of Moments (MoM) (Cramér, 1946; Kendall and Stuart, 1987). Using the ML Method, the location (β) and the scale (α)

parameter can be estimated through the numerical solution of the following equations simultaneously:

$$\tilde{x} - \frac{\sum_{t=1}^{n} x_t \exp(-x_t/a)}{\sum_{t=1}^{n} \exp(-x_t/a)} - a = 0 \quad \text{and} \quad -a \cdot \log\left[\frac{1}{n}\sum_{i=1}^{n} \exp\left(-\frac{x_i}{a}\right)\right] - \beta = 0 \tag{2.1.2}$$

where $x_1, ..., x_n$ is a random sample, $\tilde{x}$ is the sample mean and α, β the scale and location parameter respectively.

Using the MoM, the location (β) and the scale (α) parameters can be calculated by:

$$\alpha = \frac{s \cdot \sqrt{6}}{\pi}, \quad \beta = \tilde{x} - 0.57721 \cdot a \tag{2.1.3}$$

where $\tilde{x}$ and s are the sample mean and standard deviation, respectively.

In order to verify the appropriateness of the distribution selection, the raw data under study (modeled or observed) is compared with the corresponding values of the theoretical distribution. There are different approaches either graphical or analytical such as Probability plots (P–P plots), Quantile plots (Q–Q plots) (Coles, 2001) and the Kolmogorov–Smirnov test (Marsaglia et al., 2003).

The next step is to estimate extreme wind speed (U_T) with the preferred return period (T) through the relation $F(U_T) = 1 - (1/T)$ leading to the following results (Palutikof et al., 1999):

$$U_T = \begin{cases} \beta + \frac{\alpha}{k}\left\{1 - \left[-\ln\left(1 - \frac{1}{T}\right)\right]^k\right\} & k \neq 0 \\ \beta - \alpha \ln\left[-\ln\left(1 - \frac{1}{T}\right)\right] & k = 0 \end{cases} \tag{2.1.4}$$

where α, β and k are the scale, location and shape parameter respectively.

The extreme wind speed uncertainty is normally distributed and expressed through the 95% confidence interval that equals to $1.96 \cdot \sigma(U_T)$, where $\sigma(U_T) = \pi \cdot \alpha \sqrt{\frac{1 + 1.14k_T + 1.10k_T^2}{6n}}$, n is the number of maxima, $k_T = \frac{\sqrt{6}(\ln T - \gamma_E)}{\pi}$ and γ_E is the Euler's constant.

For the successful implementation with respect to the principles of Extreme Value theory, events should be independent and identically distributed (Palutikof et al., 1999). It is also assumed that a stationary extreme wind speed climate characterizes the study area. The main disadvantage regarding the AM method is that only one value per year is used. This reduces the amount of the analyzed data significantly. For this reason, the original time series must be large enough. Cook (1985) suggests the use of 20 years of data for reliable results, and argues that the method cannot be applied to time series of less than 10 years.

2.2. Peaks Over Threshold method

To overcome the above mentioned shortcomings, a second approach for the estimation of return periods has been used through the Peaks Over Threshold method that is based on the Generalized Pareto Distribution (GPD) that is used to estimate the values exceeding a threshold.

The great advantage of POT method is the utilization of more data for the application that can be achieved also by smaller time series. For this reason, in contrast to AM, a period of 5–6 years is statistically adequate (Coles and Walshaw, 1994).

The first step for creating the dataset is to apply a high threshold and form wind speed clusters above it. The problem that arises with the selection of the threshold is similar to the block selection for the Block Maxima. Low thresholds may lead to violation of the asymptotic behavior of the distribution, while high will create fewer exceedances and will lead to an increase of variance. Therefore, the threshold should be high enough so as to converge to GPD and avoid the coexistence of different

populations of extremes. At the same time it must be sufficiently low in order to create a dataset big enough for a better distribution parameters estimation (Abild et al., 1992).

The climatic characteristics of the study area are of major importance for the application and should be taken into consideration before the selection of the threshold (Caires and Sterl, 2004). Independence between the events is critical and even high thresholds cannot ensure it. This is the reason why minimum separation time between the events should be established. For European climates the separation time can be set at 48 h (Cook, 1985; Gusella, 1991) while Walshaw (1994) uses 60 h for Sheffield wind data.

The next step is to select the peaks of the clusters and subtract these values from the threshold. The created data (exceedances) is used to simulate the distribution. For high thresholds, the number of exceedances per year (crossing rate) is low and Poisson distributed while the total dataset is well approximated by the exponential distribution (Palutikof et al., 1999) whose Probability Density Function is given by:

$$F(x) = ae^{-ax} \tag{2.2.1}$$

where α is rate parameter and x represents the sample.

The exponential distribution fit to the exceedances is obtained with the same techniques (ML and MoM) in order to achieve comparable outcome. However, in this case the two methodologies converge and the parameter of the exponential distribution (rate parameter) is given by $\alpha = 1/\rho$, where $\rho = \frac{\sum_1^n x_i}{n}$ is the mean of the sample.

At this point a goodness-of-fit test is necessary for checking the suitability of threshold selection in line with the parameters estimation techniques. A graphical technique is the so called Conditional Mean Exceedance (CME) graph that is also known as residual life graph (Davison, 1984; Ledermann et al., 1990). Walshaw (1994) proposed a different approach of CME named reclustered excess graph. An analytical way to determine the appropriate threshold is the well-known Kolmogorov–Smirnov test (Marsaglia et al., 2003).

For the calculation of the extreme wind event U_T with return period of T years (T-year event) the threshold crossing rate is necessary. The T-year event can be calculated for different values of the shape parameter (k) and rate parameter (α) (Abild et al., 1992):

$$U_T = \begin{cases} \xi + \left(\frac{a}{k}\right)\left[1 - (\lambda T)^{-k}\right] & , k \neq 0 \\ \xi + a\ln(\lambda T) & , k = 0 \end{cases} \tag{2.2.2}$$

Assuming that the crossing rate (λ) follows the Poisson distribution, it can be calculated by $\lambda = n/M$ where n is the total number of exceedances above threshold ξ, and M is the length of data in years. For a Poisson simulation, uncertainty can be determined using the variance that is given by: $\sigma(U_T) \approx \left(\frac{a}{\sqrt{\lambda L}}\right)\sqrt{1 + \ln^2(\lambda T)}$, where L is the length of data in years.

The T-year event can be assumed to be normally distributed (Kite, 1975). Thus, the 95% confidence interval can be estimated as $1.96 \cdot \sigma(U_T)$ as in the AM method.

2.3. Statistical analysis

Despite the fact that the main objective of this study is to present ways of estimating return periods of extreme wind events, a statistical analysis and a comparison between the model and the measurements will provide useful information and support to the final results. For this reason different statistical indexes and graphs were estimated for the data. The Coefficient of Determination (R^2)

is a number indicating the fit of the modeled data and the measurements, being calculated by:

$$R^2 = 1 - \frac{\sum_i^k \left[\mathrm{obs}(i) - \mathrm{for}(i)\right]^2}{\sum_i^k \left[\mathrm{obs}(i) - \frac{1}{n}\sum_i^k \mathrm{obs}(i)\right]^2} \tag{2.3.1}$$

where "for" denotes the modeled values, "obs" the corresponding observations and "k" is the size of the sample. Bias and Normalized Bias provide information about the systematic deviations between the two data sets while the Root Mean Square Error (RMSE) takes also into consideration non-systematic errors. Following the same terminology, these indexes are estimated by the relations:

$$\mathrm{Bias} = \frac{1}{k}\sum_{i=1}^k \left[\mathrm{for}(i) - \mathrm{obs}(i)\right] \tag{2.3.2}$$

$$\mathrm{Normalized \quad Bias} = \frac{1}{k}\sum_{i=1}^k \left|\frac{\mathrm{for}(i) - \mathrm{obs}(i)}{\mathrm{obs}(i)}\right| \tag{2.3.3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{k}\sum_{i=1}^k \left[\mathrm{for}(i) - \mathrm{obs}(i)\right]^2} \tag{2.3.4}$$

2.4. Data used

The estimation of return periods was based on data obtained from both stations and atmospheric model simulations. Particularly, the measurements are derived from ground stations and refer to a five-year period (2006–2010). The selection of the location of each station was made on a way to cover a large part of the marine area of Southeastern Europe with different climatic characteristics and complex geomorphological distribution. This area is affected by a trade wind system called "Etesians" and deep cyclogenetic activity moving from Central to East Mediterranean. Concerning the West coast of Greece, the islands of Corfu and Kefalonia were used. For the region of the Aegean Sea, the stations of Skyros, Chios, Mykonos, Milos and Santorini were used. For Crete, the used stations (Souda and Sitia) are located at the western and eastern part of the island respectively. The data consists of measurements recorded every three hours. For each measurement, the average wind speed of the last ten minutes (out of the three hours period) is used following the World Meteorological Organization (WMO) instructions.

The location of the Meteorological Stations is provided in Table 1 and illustrated in Fig. 1.

The second source of data to be processed is the database generated under the framework of MARINA Platform project. One of the major objectives of this project was the development of a wind-wave-ocean current database that covers NE Atlantic and Mediterranean. For the construction of this database, the atmospheric model SKIRON was used (Kallos et al., 1997; Spyrou et al., 2010) in combination with the 3-D data assimilation system, LAPS, (Albers, 1995) to simulate the meteorological fields. For waves, the

Table 1
The locations of the Meteorological Stations used.

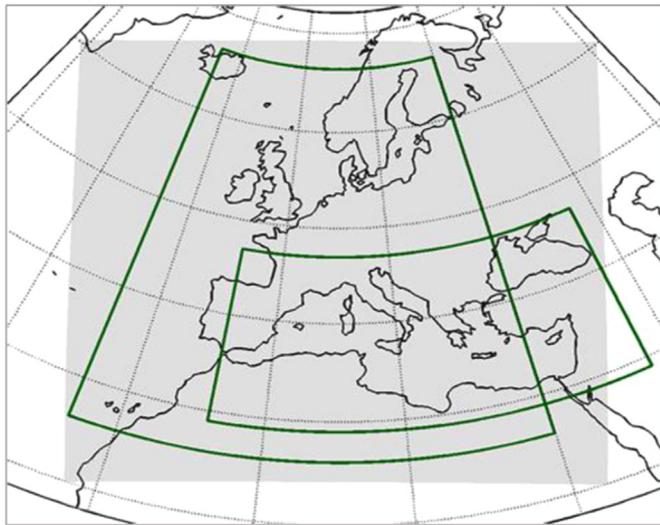| Station	Latitude	Longitude	Altitude (m)	Station code
Argostoli	38.18	20.48	25	16685
Chios	38.97	24.48	5	16706
Corfu	39.62	19.92	4	16641
Milos	36.72	24.45	183	16738
Mykonos	37.43	25.35	123	16750
Siteia	35.20	26.10	114	16757
Skyros	38.97	24.48	22	16684
Souda	35.53	24.12	150	16746
Santorini	36.42	25.42	37	16744

Fig. 1. The Met-Stations used for the study.

Fig. 2. The Skiron (gray-shading) – WAM (green lines) coupling area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

WAM model was used (Hasselmann et al., 1988; Bidlot et al., 2007; Galanis et al., 2011; Bidlot, 2012).

SKIRON is a limited area atmospheric modeling system based on the non-hydrostatic form of the primitive equations of motion, continuity and energy conservation. It uses full radiative transfer and surface energy budget parameterization. Arakawa E grid is used for horizontal coordinates and Eta for the vertical. For more details someone can look at Spyrou et al. (2010) and references therein.

SKIRON and WAM have been used in various applications worldwide and have been evaluated against various datasets. More specifically, SKIRON and WAM models were used in a number of projects related to wind and wave energy applications with remarkably good performance. ENVIWARE, POWWOW, ANEMOS, MARINA PLATFORM, WAUDIT and IRPWIND are few of them. More information related to the projects and publications can be found in http://forecast.uoa.gr/oldproj.php. Indicative publications are: Galanis et al. (2012), Papadopoulos et al. (2002), Zodiatis et al. (2003), Janeiro et al. (2012), Dykes et al. (2009),

Korres et al. (2002), Stathopoulos et al. (2013) and references therein.

The models have been run at a high spatial resolution of $0.05° \times 0.05°$ latitude/longitude covering a large part of Europe (Fig. 2).

SKIRON uses 45 levels in the vertical on a telescopic distribution (from surface to 50 hPa with more layers near the ground), and a time step of 15 s. The initial and lateral boundary conditions are prepared with the use of ECMWF $0.5° \times 0.5°$ gridded fields in combination with the global network surface and upper air observations and the 3-D data assimilation model LAPS. The lateral conditions are updated every 3 h. The geomorphological datasets that were used for the atmospheric and wave model are $30'' \times 30''$ global elevation, $30'' \times 30''$ land use and vegetation cover, $2' \times 2'$ soil classification and $1' \times 1'$ bathymetry. The SST fields used are derived from NCEP with a resolution of $0.5° \times 0.5°$. For ocean circulation, the results from the global model Hybrid Coordinate Ocean Model – HYCOM (Chassignet et al., 2003) have been interpolated from the original grid of $0.07° \times 0.07°$ to the model domain of SKIRON and WAM. The produced output is available for the period 2001–2010 on daily base.

3. Results

3.1. Statistical analysis

The results of the statistical analysis reveal the differences between the model and the station data. It is obvious that in three cases, there are higher systematic errors (Chios, Skyros and Siteia). In Table 2, more statistical indexes are provided for the stations under consideration.

The observed differences can be attributed to a number of reasons including basic issues related to mesoscale modeling algorithms (spatial and temporal resolution), topographic representation etc. On the other hand, observations at discrete locations include various sources of errors related to spatial representation, instrument inaccuracies etc. In the present case, the observations are recorded as discrete values of 10-min averages (in knots) and stored in integer form. Converting knots to m/s for the comparison with the model results an error, that may affect the distribution fitting procedures discussed in the following sections, may appear.

Table 2

Statistical indexes of the Met-Stations used. JFM denotes January, February and March, AMJ denotes April, May and June, JAS denotes July, August and September and OND denotes October, November and December.

	Argostoli	Corfu	Milos	Mykonos	Santorini	Siteia	Skyros	Souda	Chios
R^2	0.396	0.377	0.433	0.493	0.448	0.503	0.537	0.393	0.505
R^2 for JFM	0.483	0.446	0.451	0.533	0.460	0.512	0.489	0.455	0.487
R^2 for AMJ	0.372	0.392	0.395	0.440	0.415	0.495	0.560	0.412	0.444
R^2 for JAS	0.311	0.327	0.387	0.484	0.380	0.465	0.554	0.268	0.450
R^2 for OND	0.400	0.371	0.438	0.493	0.460	0.507	0.538	0.402	0.519
Bias	−0.246	−0.475	−0.560	0.224	0.699	−1.012	1.041	0.076	1.645
Normalized bias	0.008	−0.064	0.013	0.134	0.219	−0.086	0.347	0.100	0.676
RMSE	2.051	2.097	2.945	2.806	2.588	2.735	2.778	2.353	2.758
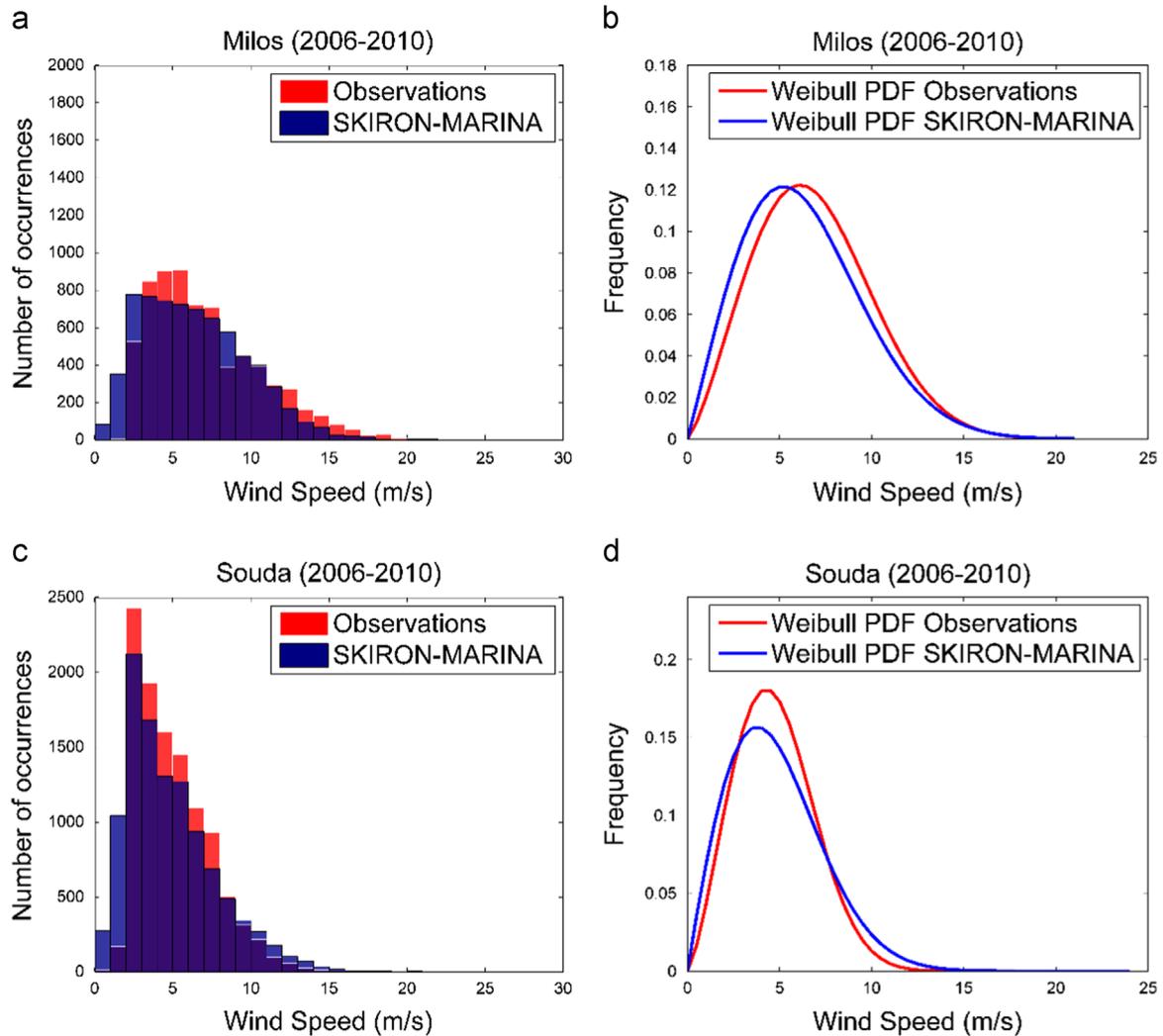
Fig. 3. Histograms (left- a, c) and Weibull PDF plots (right- a, d) of wind speed for Milos (up- a, b) and Souda island (down- c, d) respectively.

For the study of return periods, the above statistical comparison does not give all the information needed since there is no focus on the extreme wind speed behavior that is more important. A graphical representation of the distribution of the datasets provides additional information for the wind speed variability (see Fig. 3).

As the histograms and the Probability Density Function (PDF) of the observed and modeled values show, the model underestimates the wind speed (see the light skewness to the left). However, what is not obvious in all cases, is the fact that the model overestimates the extreme wind speeds. This is depicted in the scatter plots of Fig. 4a and b (values of wind speed over an imposed threshold of 20 m/s). This range of values will be utilized for the estimation of return periods.

The model simulates quite successfully the temporal evolution of maximum wind speeds. However, for most of the stations, the modeled peaks tend to be higher than the corresponding measurements. This means that there will be an overestimation concerning the annual maximum values of the sample as well. Having in mind that these values constitute the data for the study of return periods, an overestimation of the model estimation of extreme wind events is expected. The latter is not inconsistent with the purpose of the study. Instead, in such applications that
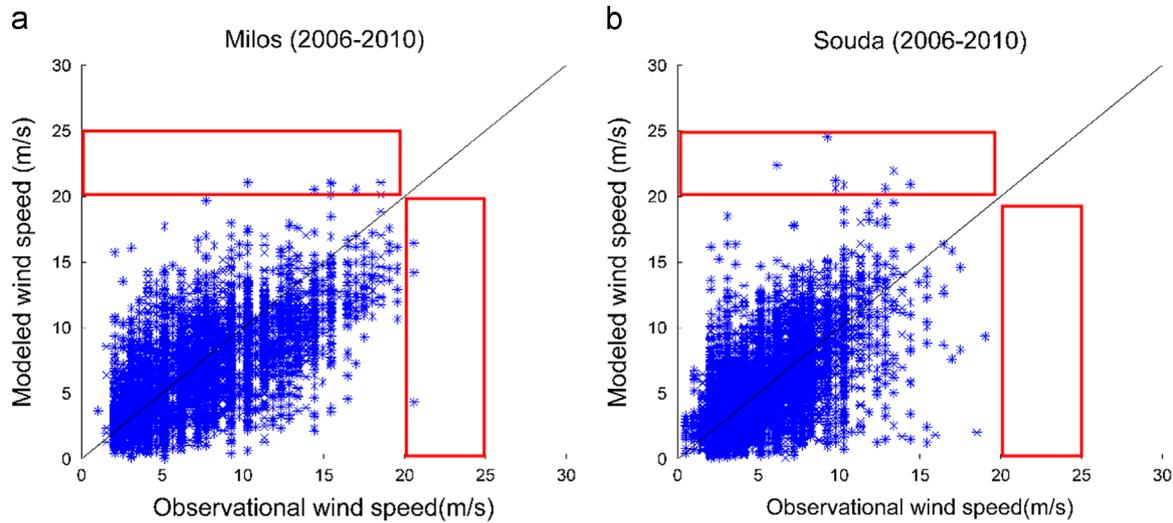
Fig. 4. Scatterplots of the modeled wind speed and the observations for Milos (left-a) and Souda island (right-b). Wind speed time series for a short period for both datasets (bottom).

can be used by construction industry, underestimation of return periods will be more problematic.

In any case and in order to avoid the deviation observed, a local adjustment of the model outputs would be useful. Such optimization can be held by various techniques such as Kalman filter (Galanis et al., 2006; Louka et al., 2008; Kalnay, 2002) that is considered to be a good tool to reduce bias. For the application of this technique, observations are needed to be combined with model results and train the system so as to minimize the corresponding biases. The main advantage of this methodology is the easy adaptation to the dynamic nature of systems and the short training period needed. However the need of observations constrains the applicability of the method in certain areas and timescales.

3.2. Extreme wind events and return periods estimation

The procedure developed for the estimation of return periods of extreme wind speed is based on Annual Maxima and Peaks Over Threshold methods. Each method was tested both for wind speed and wind speed squared. In addition, the Method of Moments (MoM) and the Maximum Likelihood Method (ML) were used for the parameter estimation for the distribution fitting. The goodness-of-fit was based mainly on the Kolmogorov–Smirnov test.

For each one of the nine locations/islands that are mentioned earlier, the input data was separated in four different datasets. The first dataset is the 10-year time series derived by the MARINA database (SKIRON-MARINA 2001–2010). The second is obtained from the same database but for a five-year period (SKIRON-MARINA 2006–2010) (corresponding to the same time period with the observations). The third is the observations (Station with missing values 2006–2010). The fourth is the modeled time series that corresponds to available data of the station. (SKIRON-MARINA database with missing values 2006–2010).

The methodology for the estimation of return periods is separated according to the parameter distribution fitting tool. Beginning with the MoM, the AM method can be utilized only for SKIRON-MARINA 2001–2010 data as it is the only dataset that covers the minimum time length required by the methodology. As it is seen in Fig. 5, the Gumbel distribution does not provide a very good fit for the station of Milos. This is mainly due to the small number of values of the input data.
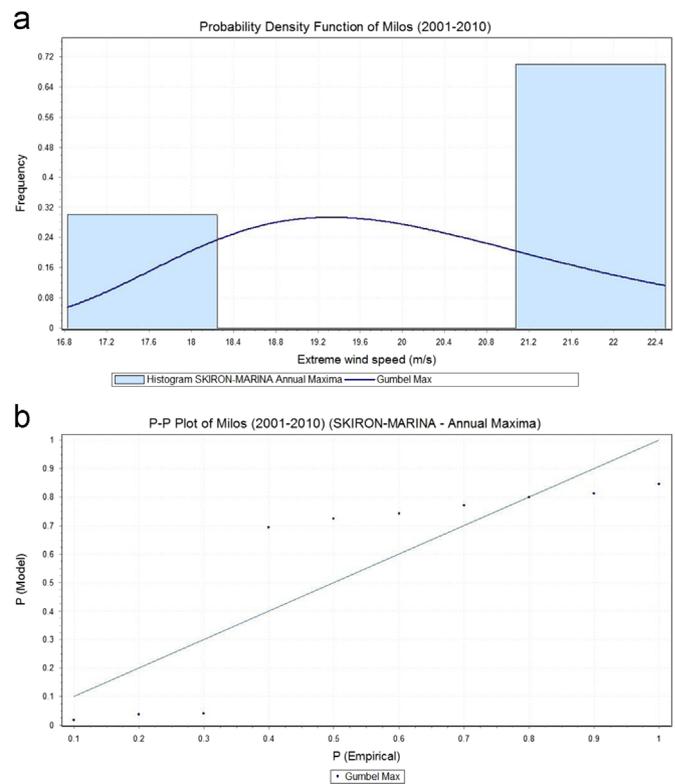

Fig. 5. Gumbel PDF of Annual Maxima (top-a) and P–P plot for the same fit (bottom-b) for the island of Milos.

For the same station the extreme wind speeds and the corresponding return periods have been estimated by applying the MoM method and the results are illustrated in Fig. 6.

The POT method was applied to all datasets. The threshold was selected between 98.5% and 96% quantile and the separation time for the independence of events was set to 48 h. The Kolmogorov–Smirnov tests have been employed to identify the exact threshold within the mentioned range that provides the optimum fit for each grid point. As it can be seen in Fig. 7 and the following graphs, the exponential distribution fits well to the 10-year SKIRON-MARINA dataset for the station of Milos.

Based on the estimated parameter of the exponential, the extreme wind speed and the corresponding return periods with
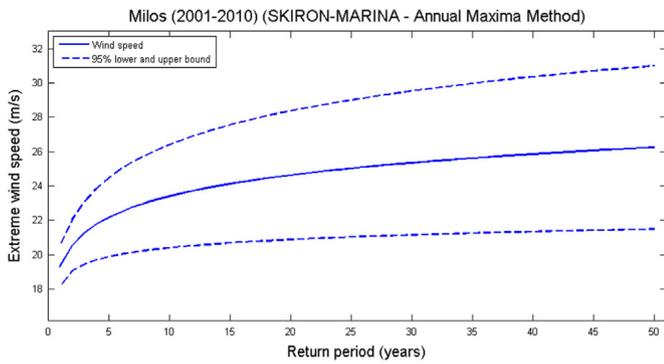
Fig. 6. The extreme wind speeds and the corresponding return periods for up to 50 years using AM (MoM method – Milos).
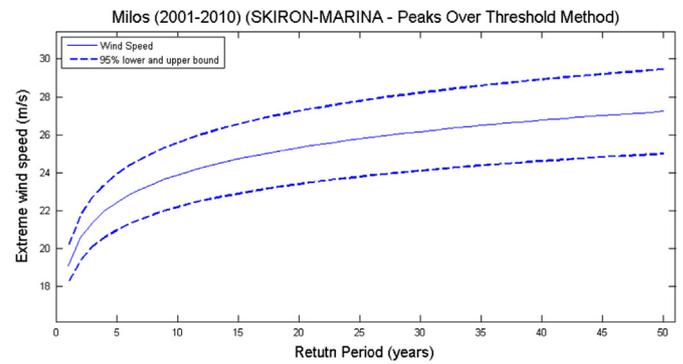

Fig. 8. The extreme wind speeds and the corresponding return periods for up to 50 years using POT (MoM method – Milos).
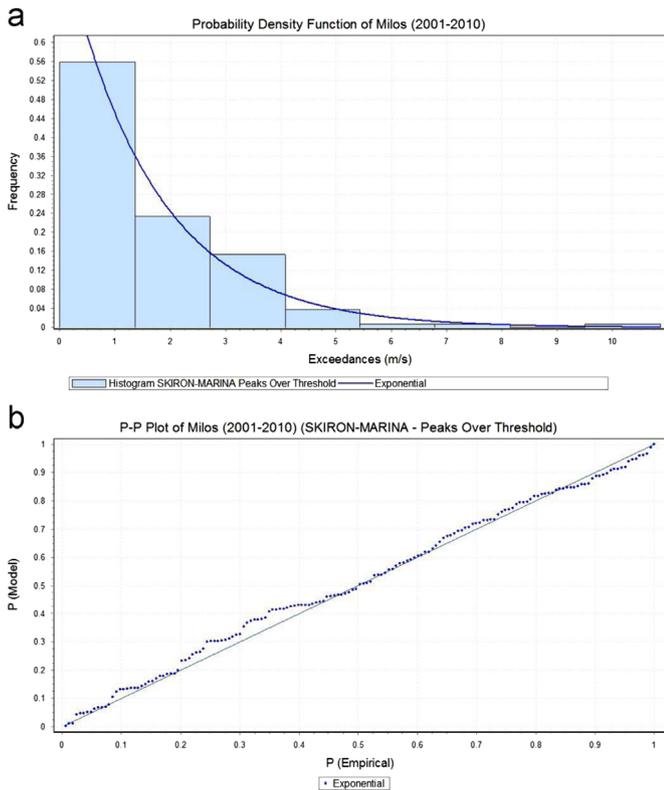

Fig. 7. Exponential PDF of exceedances (top-a) and P–P plot for the same fit (bottom-b) for the island of Milos.

Table 3
Annual Maxima of Milos modeled wind speed for the timeperiod 2001–2010.

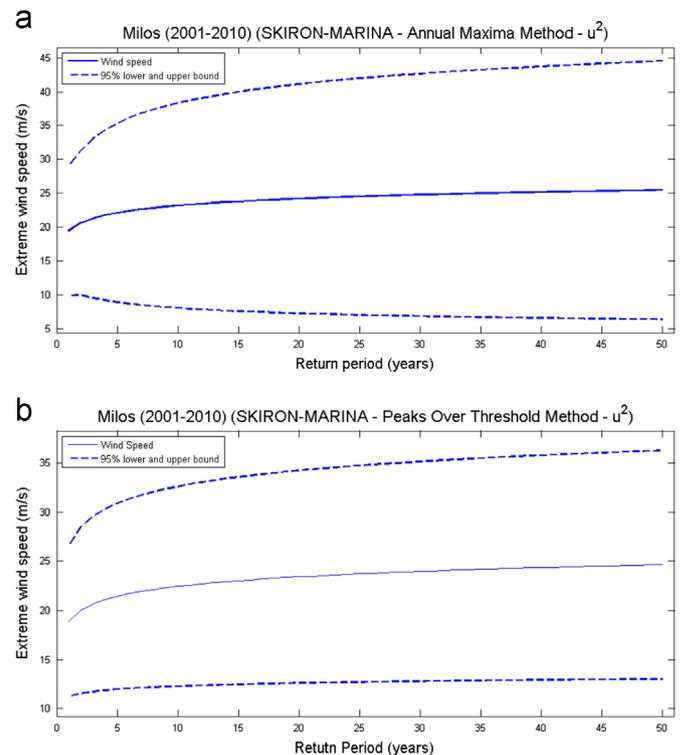| Year	Maximum wind speed (m/s)	Year	Maximum wind speed (m/s)
2001	17.22	2006	21.46
2002	17.26	2007	16.83
2003	21.70	2008	21.31
2004	22.09	2009	22.49
2005	21.09	2010	21.96

Fig. 9. The extreme wind speeds and the corresponding return periods for up to 50 years using AM (top-a) and POT (bottom-b) based on wind speed squared (MoM method – Milos).

the 95% upper and lower bounds have been estimated. The results are illustrated in Fig. 8.

As an additional information that can be used in conjunction with the above discussion, the maximum wind speed values for each year are presented in Table 3.

Despite the fact that the length of the time series is small, the estimated parameters of the AM method have led to acceptable results since the POT method outputs converge to those of AM.

The same procedure was followed for the case of wind speed squared values as input data but the methods did not provide acceptable results because of the radical increase of the corresponding confidence intervals as illustrated in Fig. 9.

Similar results have been achieved for all the datasets and this is the reason why they are not further discussed.

The next technique to be tested is the ML method. For the use of ML method a similar procedure was followed. Concerning the

application to wind speed squared, the results were similar to the use of MoM.

Extreme wind events, defined with the use of the direct wind speed values that are calculated by the AM method are illustrated in Fig. 10. The results of the POT using MoM and ML are identical. This is something expected because both methods end up to the same equations for the estimation of the parameters of the
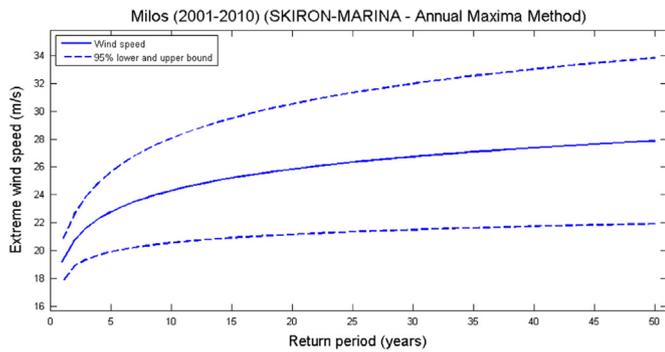
Fig. 10. The extreme wind speeds and the corresponding return periods for up to 50 years using AM (ML method – Milos).

Fig. 11. The extreme wind speeds and the corresponding return periods for up to 50 years for all datasets and methodologies (Milos).

exponential distribution. For this reason the extreme wind speed plot using the ML and the POT method is not included.

The obtained results are considered as quite satisfactory since the values of these different approaches are close enough, a fact that ensures the convergence of the methods and reduces the corresponding uncertainty.

Figs. 7 and 10 depict that both results are characterized by good exponentiality. For better understanding the behavior of the methodologies tested here, a common representation has been prepared and illustrated in Fig. 11. More specifically, the extreme wind speeds with return periods up to 50 years as calculated for both fitting methods (ML, MoM), both methodologies (AM, POT) and all data sets have been plotted.

From this Figure it is clear that the model data series provide higher extreme wind speed for almost all the return periods and the applied methodologies. This is something that it was also discussed previously and illustrated in Fig. 4. Comparing the results obtained from the estimators MoM and ML on the same dataset (2001–2010 SKIRON-MARINA) and methodology (AM) it is clear that the first provides a slight underestimation (maximum underestimation of 6.17%). However, this underestimation is within the confidence intervals of both curves. Similar behavior has been observed in most of the stations used in the analysis. The small underestimations of MoM for almost all the sites and the fact that ML method is easily adaptable to include effects of covariates, or other influencing factors has led us in the selection of the latter as fitting methodology. This is also in agreement with previous work of other researchers (see Katz et al., 2002; Smith, 1989; Zhang et al., 2004). For this reason the discussion will be limited to the results obtained by the ML fitting methodology. The outcomes of this analysis in all nine stations and methodologies are summarized in Table 4.

Beginning with the confidence interval estimation, it is obvious that for the POT method, smaller time series result to higher confidence intervals. This is reasonable and expected since the number of data on the second and third case diminishes. In addition, the magnitude of the confidence intervals is highly associated with the intensity of the estimated extreme wind speeds. The AM method is based on the highest wind speed occurred within each year that means there are only 10 values used for the application. This leads to higher confidence intervals ranging from 16.51% to 22.78%.

According to the same table, the 50-year extreme wind speed estimated by applying the POT method on 5 and 10-years long time series deviate by about 5% maximum that is within the confidence intervals. This means that the used time series have adequate length for such applications.

To explore the influence of missing data or the sampling frequency of the time series on the return period estimation, the results obtained by SKIRON-MARINA 2006–2010 and SKIRON-
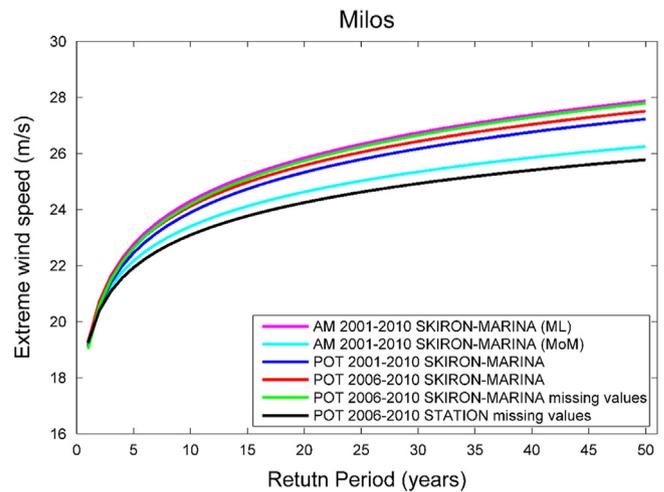
MARINA 2006–2010 paired (missing data) are compared. It was found that the highest difference was 9.41% and the lowest 0.35% in Chios and Corfu respectively as seen in Table 4. Obviously this difference is strongly related to the sampling frequency of the time series because of the changes on the distribution upper tails.

Another considerable remark is that the POT methodology applied in all datasets from the same source gives deviation in the extreme wind estimation that is always within the confidence intervals. This is a good indicator of the robustness of this approach.

Continuing with the comparison of POT and AM for the estimation of the 50-year extreme wind speed, the variations observed in Table 4 can be attributed to the amount of data taken into consideration by each method. Based on this, there are two extreme cases to point out. The first is when the study area is characterized by low wind speeds with a relatively small number of some specific strong wind events. Such a behavior depends on both the climatology of the region and synoptic scale characteristics. In this case, the application of the AM will lead in overestimation as compared to the POT method that is not the normal case. The opposite is true when an area is characterized by high wind speeds encountered frequently. This is a result of the thickness change of the probability distribution tail, fitted to the data.

Regarding the measurements, a constant extreme wind speed underestimation is evident with only three cases (Milos, Siteia, Souda) to be within the confidence interval limits. This was already mentioned previously but here is worth mentioning another influencing factor. More specifically, the observations are measured in knots and are rounded to their integer part. The conversion to S.I. metric system led to an equivalent discretization that causes a distribution fit misbehavior. It is remarkable that in every case the distribution fitting barely passed Kolmogorov–Smirnov test. These fitting issues are obvious in the P–P plots of Fig. 12.

Another important influencing factor in the analysis of observational time series is the quality of the measurements (e.g. missing values, instrument malfunctioning, quality controls applied for corrections etc.). For most of the cases such errors that affect the values used for the extreme wind estimation cannot be easily detected and any further modification is rather subjective.

Concluding, the spatial and temporal resolution are always affecting the results and the credibility of the model and problems of systematic or not biases may appear especially over complicated topography. However, numerical models provide today a very good alternative to observations especially over areas that are not

Table 4
Fifty year extreme wind speed and the corresponding confidence intervals for the all the stations.

|		Argostoli	Chios	Corfu	Milos	Mykonos	Santorini	Siteia	Skyros	Souda
POT	SKIRON-MARINA (2001–2010)	24.92 ± 2.19	24.13 ± 1.99	28.45 ± 2.83	27.22 ± 2.22	26.17 ± 1.86	32.25 ± 2.79	26.01 ± 2.34	30.88 ± 2.86	25.63 ± 2.02
	SKIRON-MARINA (2006–2010)	25.49 ± 3.19	22.91 ± 2.33	29.45 ± 4.18	27.50 ± 3.15	26.10 ± 2.60	31.69 ± 3.77	26.70 ± 3.70	29.94 ± 3.89	25.89 ± 2.89
	SKIRON-MARINA (2006–2010 paired/missing values)	25.66 ± 3.84	21.86 ± 2.30	28.55 ± 4.23	27.79 ± 3.78	27.45 ± 3.19	33.21 ± 4.57	25.59 ± 3.56	31.36 ± 4.39	26.14 ± 3.07
	Station (2006–2010 missing values)	19.53 ± 1.87	17.41 ± 1.57	21.56 ± 2.09	25.77 ± 2.93	20.88 ± 1.52	20.43 ± 1.43	24.91 ± 2.39	25.64 ± 2.68	23.10 ± 2.08
AM (ML)	SKIRON-MARINA (2001–2010)	23.53 ± 5.36	26.59 ± 5.81	22.82 ± 3.98	27.87 ± 5.95	29.53 ± 6.15	31.66 ± 6.39	27.09 ± 4.86	29.99 ± 4.95	27.76 ± 6.11
AM (MoM)	SKIRON-MARINA (2001–2010)	22.80 ± 4.82	25.49 ± 5.00	23.03 ± 4.12	26.25 ± 4.75	28.85 ± 5.64	31.79 ± 6.49	25.56 ± 3.75	29.06 ± 4.27	26.79 ± 5.41
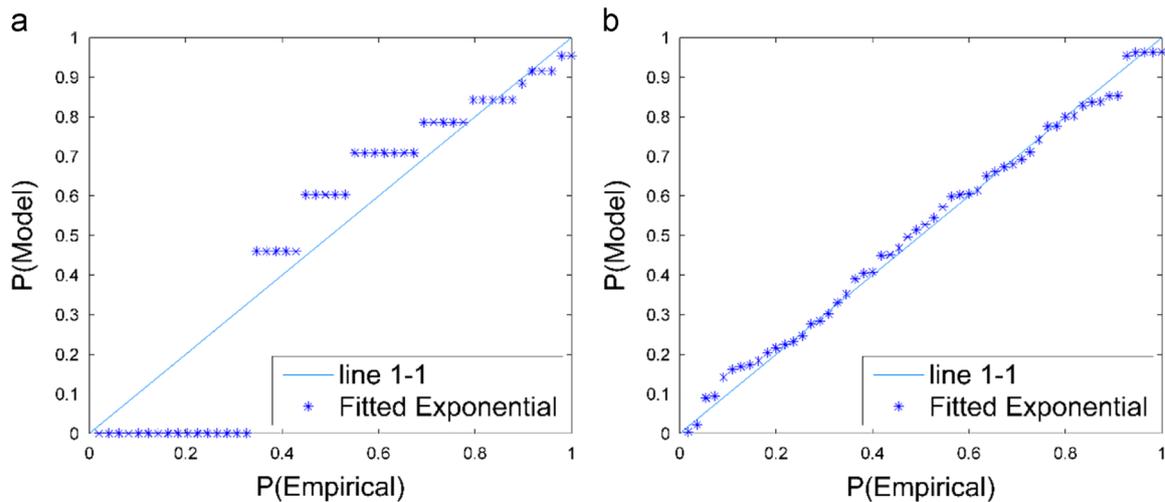
Fig. 12. P–P plots of the exponential fit on observations (left-a) and modeled wind speed (right-b) exceedances.

covered by the network of met stations. In the presented work, in order to deal with problems relative to the model's accuracy due to the above issues, high resolution – in time and space – models have been adopted. However, there are always cases where local effects cannot be fully resolved. One of the main objectives of this paper is to examine the discrepancies that could emerge by the utilization of NWPs in contrast to observations for the estimation of extreme values and estimate extreme winds in areas where observations are insufficient.

3.3. Extreme wind atlas

Based on the previous discussion, the SKIRON-MARINA database is considered as an appropriate dataset for extreme wind speed analysis. The purpose of this chapter is to apply the proposed methodology for mapping the extreme wind speed with a return period of 50 years. For this purpose the 10 m wind fields of the SKIRON-MARINA 2001–2010 database were utilized. The map covers the area between 19E–29E and 34N–41N longitude and latitude respectively that includes the marine area surrounding Greece and the islands of the region of interest. The results presented, provide information regarding the likelihood of strong winds events, that can be used in risk analysis in constructions such as offshore and onshore wind farms, oil platforms etc. In each case and especially in terrestrial applications, this information should be accompanied with higher resolution analysis simulations to capture local features.

Two extreme wind speed Maps (Atlas) accompanied with their confidence intervals have been prepared following two different methodologies namely AM and POT.

For AM the annual maximum values have been used (10 values) while for POT a threshold with the range 96–98% quantile was employed. The threshold selection was based on the optimum exponential distribution fit to the exceedances.

Figs. 13 and 14 illustrate the 50-year extreme wind speed and the associated confidence interval respectively by applying the POT methodology. Similarly, Figs. 15 and 16 illustrate the 50-year extreme wind speed and the associate confidence interval respectively by applying the AM technique. As it can be seen, both techniques identify the highest values in the same locations. However, POT estimates in general higher values over larger areas as compared to AM.

Another interesting point is that the POT methodology has in general smaller confidence intervals as compared to AM. This is something that has been mentioned also in previous chapters of this work.

The first map (Fig. 13) illustrates an extreme wind speed range from 10–11 m/s to 38–39 m/s while for the second (Fig. 15) is slightly lower. The lowest values are found in areas where the wind speed is generally lower due to topographic features such as the Gulf of Corinth, the Dardanelles and the lee side of the islands. A characteristic behavior is observed in Cycladic islands, where the shading caused by clusters of islands reduces the likelihood of extreme events occurrence. Larger values are observed mainly
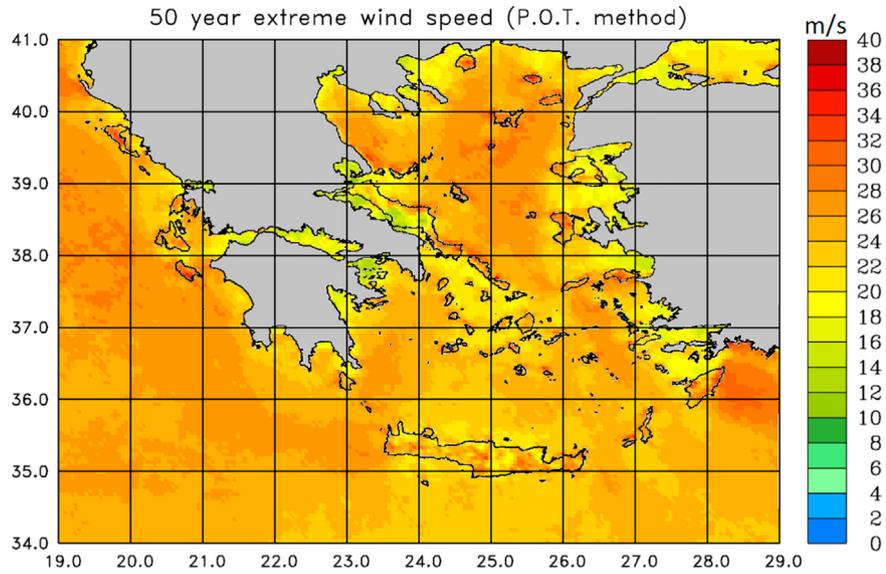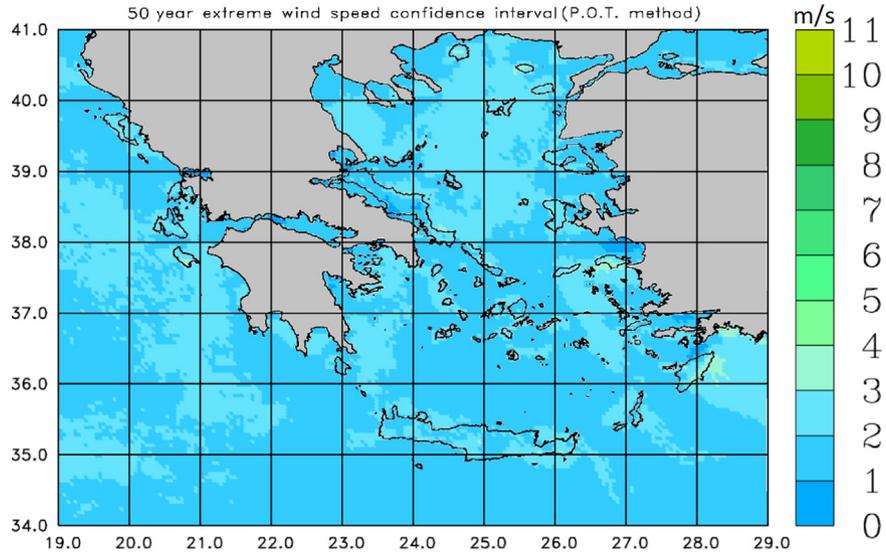
Fig. 13. Fifty year extreme wind speed atlas using POT.

Fig. 14. Fifty year extreme wind speed confidence interval using POT.
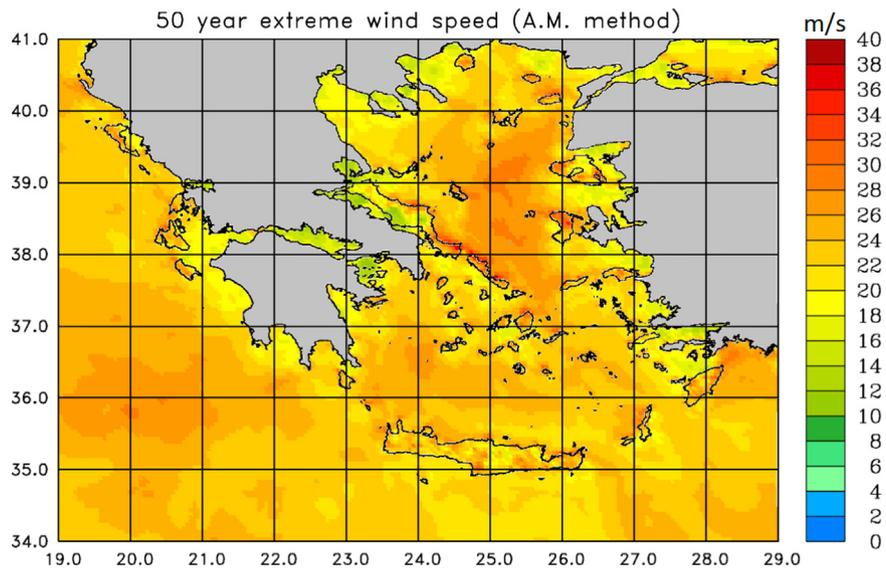

Fig. 15. Fifty year extreme wind speed atlas using AM.
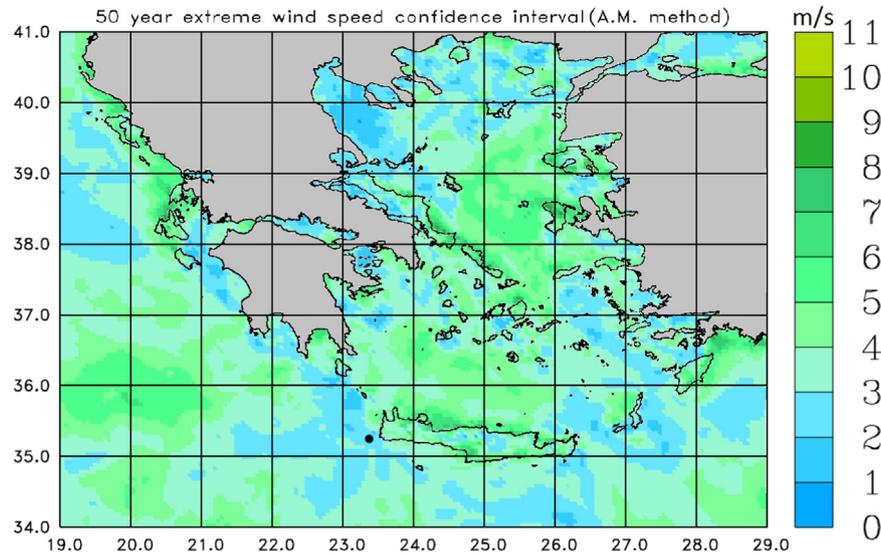
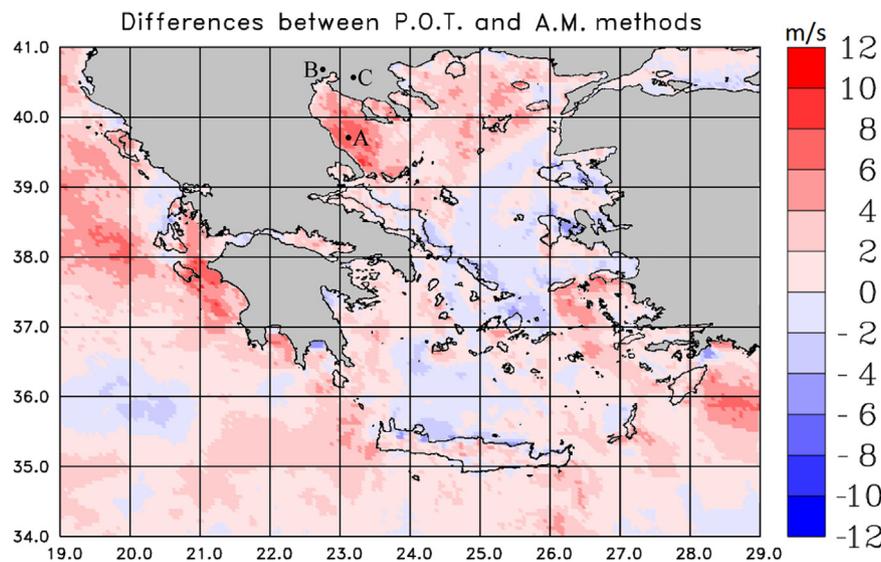**Fig. 16.** Fifty year extreme wind speed confidence interval using AM.

Fig. 17. Fifty year extreme wind speed differences of AM and POT (red color represents POT overestimation).

offshore due to the absence of land obstacles (see Ionian Sea), on the wind sides of the Aegean islands, at higher altitudes as well as in areas with local features that affect the flow field.

Each country issued certain constrains and regulations in general related to construction (on and offshore). In almost all cases a uniform threshold is implied. For example, according to the 2009 Greek National Annex of ELOT-EN-1991-1-4 (http://portal.tee.gr/portal/page/portal/SCIENTIFIC_WORK/scient_typopoiisi/eurocodes, http://www.eurocodes-online.com), the reference velocity, which corresponds to the extreme wind speed calculated here, is set to 33 m/s for the islands and the coastal zones. It can be easily observed that using POT method, this threshold is surpassed in many cases. More specifically, concerning the Ionian islands, higher wind velocities are found on the west part of them (windward side). Similarly, in the Aegean Sea the northern parts of the islands are mostly affected. Analogous distribution was found by following the AM methodology.

As mentioned earlier, the extreme values obtained by using the two methods exhibit differences, which in some cases are significant. These are mainly attributed to the way the input data is created in each case, the size of the data used and the climatic

characteristics of the study area. Fig. 17 illustrates the difference in the results of the two methods. Red represents cases where the POT method overestimates the extreme wind speed relatively to AM, while blue represents the AM overestimation as compared with POT.

Based on Fig. 17 and on the differences that are observed, information about wind speed characteristics of an area can be extracted. A characteristic example is the case of Thermaikos Gulf (NW AEGEAN, point A in the Fig. 17). The climate of the area is influenced by both synoptic and local conditions and is characterized by strong winds under certain circumstances. Vardaris and Hortiatis are two cases of winds that affect and form the local flow pattern. Vardaris is a dry and cold northerly wind that is channeling along a valley (point B in Fig. 17) and ends up at Thermaikos Gulf. It has a high frequency and occurs mainly in winter. At the same region, Hortiatis is a powerful easterly wind that blows from the nearby mountain (point C) during winter. These wind systems of this region form general conditions characterized of moderate to high frequency of appearance. Taking this into consideration, it is obvious that the POT method could be considered as more suitable for this area since it takes into account

more input data. These characteristics lead to the conclusion that the use of POT method will produce higher extreme wind speed values as compared to AM in this case. This comes also into agreement with Fig. 17.

4. Concluding remarks

In this study an effort was devoted to analyze the frequency of extreme wind speed events in terms of return periods. Different techniques have been tested and implemented accordingly. Two basic methodologies have been selected, namely the Annual Maxima (AM) and the Peaks Over Threshold (POT). They have been applied to observational and modeled datasets. The analysis results are related to the climatic characteristics of the study area.

The uncertainties that arise from the different approaches of estimating extreme winds have been studied and quantified through multiple sensitivity tests. The convergence of the theoretical probability distribution to the sample, for both methods, has been tested for wind speed and wind speed squared. For the same purpose, two parameter estimation techniques have been applied, namely the Method of Moments (MoM) and the Maximum Likelihood (ML). The results are considered as satisfactory by using the wind speed. The opposite is true when using the wind speed squared.

The impact of the length of the time series has been investigated for the POT method. It was found that the deviation of the outcome is within the confidence intervals. Based on these results we can say that time series with length of at least five years and sampling frequency of one hour is considered as adequate to estimate the 50-year return periods. However, a 10-year period is desirable.

The influence of missing data and sampling frequency of the time series on the return period estimation was found to be critical. According to these findings, the results are strongly related to the sampling frequency of the time series because of the impact on the probability distribution upper tails.

The use of the POT method to estimate extreme wind events resulted in overestimation when model datasets were used as compared to observational input. This is partly due to the structure of the observation data and the resolution capabilities of the model.

The application of the AM method led to comparable results with POT. However there are cases where the two methods differ significantly. These differences are mainly associated with the climatological wind patterns of the study area, the frequency and intensity of extreme wind events.

The ML method was used along with AM and POT to derive the 50-year extreme wind speed for the broader maritime area of Greece (including the Ionian and the Aegean seas). The mapping of the results from each approach provided valuable information that can be used in projects and actions that are affected by non-frequent wind speed events. In addition, the differences that can be found in certain areas are indicators of the local climate and denote the necessity of both procedures.

Finally, the discussed methodologies can be considered as useful tools on defining extreme wind characteristics and impose the necessary regulations for wind energy and/or construction activities.

Acknowledgments

References

Abild, J., Andersen, E.Y., Rosbjerg, D., 1992. The climate of extreme winds at the Great Belt, Denmark. J. Wind Eng. Ind. Aerodyn. 41–44, 521–532.

Albers, S.C., 1995. The LAPS wind analysis. Weather Forecast. 10 (2), 342–352.

Bidlot, J., Janssen, P., Abdalla, S., Hersbach, H., 2007. A revised formulation of ocean wave dissipation and its model impact. ECMWF Tech. Memo. 509. ECMWF, Reading, United Kingdom, 27 pp. Available online at: ⟨http://www.ecmwf.int/publications/⟩.

Bidlot J.-R., 2012. Present status of wave forecasting at ECMWF. In: Proceedings of ECMWF Workshop on Ocean Waves. 25–27 June, ECMWF, Reading, United Kingdom, 2012.

Breivik, Ø., Aarnes, O.J., Abdalla, S., Bidlot, J.-R., Janssen, P.A.E.M., 2014. Wind and wave extremes over the world oceans from very large ensembles. Geophys. Res. Lett. 41 (14), 5122–5131.

Caires, S., Sterl, A., 2004. 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 Data. J. Clim. 18, 1032–1048.

Chassignet, E.P., Smith, L.T., Halliwell, G.R., Bleck, R., 2003. North Atlantic simulations with the Hybrid Coordinate Ocean Model (HYCOM): impact of the vertical coordinate choice, reference pressure, and thermobaricity. J. Phys. Oceanogr. 33, 2504–2526.

Coles, S.G., Walshaw, D., 1994. Directional modelling of extreme wind speeds. Appl. Stat. 43, 139–157.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. London, Springer-Verlag, p. 208.

Cook, N.J., 1982. Towards better estimation of wind speeds. J. Wind Eng. Ind. Aerodyn. 9, 295–323.

Cook N.J., 1985. The Designer's Guide to Wind Loading of Building Structures. Part 1: Background, Damage Survey, Wind Data and Structural Classification. Building Research Establishment, Garston/Butterworths, London. 371 pp.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton New Jersey.

Davison, A.C., 1984. Modelling excesses over high thresholds, with an application. In: Tiago de Oliviera, J. (Ed.), Statistical Extremes and Applications. Reidel, Dordrecht, pp. 461–482.

Dykes, J.D., Wang, D.W., Book, J.W., 2009. An evaluation of a high-resolution operational wave forecasting system in the Adriatic Sea. J. Mar. Syst., 78; , pp. S255–S271.

Galambos, J., 1978. The Asymptotic Theory of Extreme Order Statistics, 2nd ed. New York, Krieger 414 pp.

Galanis, G., Chu, P.C., Kallos, G., 2011. Statistical post processes for the improvement of the results of numerical wave prediction models. A combination of Kolmogorov–Zurbenko and Kalman filters. J. Oper. Oceanogr. 4 (1), 23–31.

Galanis, G., Louka, P., Katsafados, P., Pytharoulis, I., Kallos, G., 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. Ann. Geophys. 24, 2451–2460.

Galanis, G., Hayes, D., Zodiatis, G., Chu, P.C., Kuo, Y.H., Kallos, G., 2012. Wave height characteristics in the Mediterranean Sea by means of numerical modeling, satellite data, statistical and geometrical techniques. Mar. Geophys. Res. 33, 1–15.

Gusella, V., 1991. Estimation of extreme winds from short-term records. J. Struct. Eng. 117, 375–390.

Harris, R.I., 1998. Improvements to the 'Method of Independent Storms'. J. Wind Eng. Ind. Aerodyn.

Hasselmann, S., Hasselmann, K., Bauer, E., Bertotti, L., Cardone, C.V., Ewing, J.A., Greenwood, J.A., Guillaume, A., Janssen, P., Komen, G., Lionello, P., Reistad, M., Zambresky, L., 1988. The WAM Model – a third generation ocean wave prediction model. J. Phys. Oceanogr. 18 (12), 1775–1810.

Hazewinkel, M.(Ed.), Maximum-likelihood method, Encyclopedia of Mathematics, 2001, Springer, ISBN 978-1-55608-010-4.

Hennessey Jr., J.P., 1977. Some aspects of wind power statistics. J. Appl. Meteorol. 16, 119–128.

Janeiro, J., Martins, F., Relvas, P., 2012. Towards the development of an operational tool for oil spills management in the Algarve coast. J. Coast. Conserv. 16 (4), 449–460.

Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) of meteorological elements. Q. J. R. Meteorol. Soc. 81, 158–171.

Kallos, G., Galanis, G., Spyrou, C., Kalogeri, C., Adam, A., Athanasiadis, P., 2012. Offshore energy mapping for Northeast Atlantic and Mediterranean: MARINA PLATFORM project. Geophys. Res. Abstr. 14 EGU2012-10767.

Kallos, G., Nickovic, S., Papadopoulos, A., Jovic, D., Kakaliagou, O., Misirlis, N., Boukas, L., Mimikou, N., Sakellaridis, G., Papageorgiou, J., Anadranistakis, E., and Manousakis, M, 1997. The regional weather forecasting system Skiron: an overview. Proceedings of Symposium on Regional Weather Prediction on Parallel Computer Environments, 15–17 October, Athens, Greece, pp. 109-122, 1997.

Kalnay, E., 2002. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge, Cambridge University Press, p. 341 2003.

Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. Adv. Water Resour. 25, 1287–1304.

Kendall, M.G., Stuart, A., 1987. The Advanced Theory of Statistics. Oxford University Press, Inc., New York, NY, USA, ISBN:0-195-20561-8.

Kite, G.W., 1975. Confidence limits for design events. Water Resour. Res. 11, 48–53.

Korres, G., Lascaratos, A., Hatziapostolou, E., Katsafados, P., 2002. Towards an ocean forecasting system for the Aegean Sea. Glob. Atmos. Ocean. Syst. 8 (2–3), 191–218.

Larsén X.G., Badger J., Hahmann A., Ott S., 2011. Extreme wind atlas from the selective dynamical downscaling method. In: Proceedings of European Wind Energy Conference and Exhibition. 2011.

Ledermann, W., Lloyd, E., Vajda, S., Alexander, C. (Eds.), 1990. Handbook of Applicable Mathematics, 7. Wiley-Interscience, Chichester, p. 479.

Lopatoukhin, L.J., Rozhkov, V.A., Ryabinin, V.E., Swail, V.R., Boukhanovsky, A.V., Degtyarev, A.B., 2000. Estimation of extreme wind wave heights. World Meteorological Organization, LCOMM Tech. Rep. No. 9.

Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., Kallos, G., 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. J. Wind Eng. Ind. Aerodyn. 96, 2348–2362.

Marsaglia, G., Tsang, W., Wang, J., 2003. Evaluating Kolmogorov's distribution. J. Stat. Softw. 8 (18), 1–4.

Palutikof, J.P., Brabson, B.B., Lister, D.H., Adcock, S.T., 1999. A review of methods to calculate extreme wind speeds. Meteorol. Applic. 6 (1), 119–132.

Papadopoulos, A., Kallos, G., Katsafados, P., Nickovic, S., 2002. The POSEIDON weather forecasting system: an overview. Glob. Atmos. Ocean Syst. 8, 219–237.

Smith, R.L., 1986. Extreme value theory based on the r largest annual events. J. Hydrol. 86, 27–43.

Smith, R.L., 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Stat. Sci. 4, 367–393.

Spyrou, C., Mitsakou, C., Kallos, G., Louka, P., Vlastou, G., 2010. An improved limited area model for describing the dust cycle in the atmosphere. J. Geophys. Res. 115, D17211. http://dx.doi.org/10.1029/2009JD013682.

Stathopoulos, C., Kaperoni, A., Galanis, G., Kallos, G., 2013. Wind power prediction based on numerical and statistical models. J. Wind Eng. Ind. Aerodyn. 112, 25–38.

Walshaw, D., 1994. Getting the most from your extreme wind data: a step by step guide. J. Res. Natl. Inst. Stand. Technol. 99, 399–411.

Zhang, X., Zwiers, F.W., Li, G., 2004. Monte Carlo experiments on the detection of trends in extreme values. J. Clim. 17, 1945–1952.

Zodiatis, G., Lardner, R., Georgiou, G., Demirov, E., Manzella, G., Pinardi, N., 2003. An operational European global ocean observing system for the eastern mediterranean levantine basin: the Cyprus coastal ocean forecasting and observing system. Mar. Technol. Soc. J. 37 (3), 115–123.

Web references (10/09/2015)

http://forecast.uoa.gr/proj_marina.php.

http://forecast.uoa.gr/oldproj.php.

http://www.eurocodes-online.com.

http://portal.tee.gr/portal/page/portal/SCIENTIFIC_WORK/scient_typopoiisi/eurocodes.