

Article

The Compound Inverse Rayleigh as an Extreme Wind Speed Distribution and its Bayes Estimation

Elio Chiodo^{1,*}, Maurizio Fantauzzi¹ and Giovanni Mazzanti²

1

¹ Department of Industrial Engineering, University of Naples Federico II, 80125 Naples, Italy; e-mail: elio.chiodo@unina.it; fantauzz@unina.it

² Department of Electrical, Electronic and Information Engineering, University of Bologna, 40136 Bologna, Italy; e-mail: giovanni.mazzanti@unibo.it

*Correspondence: elio.chiodo@unina.it

Abstract: The paper deals with the Compound Inverse Rayleigh distribution, shown to constitute a proper model for the characterization of the probability distribution of extreme values of wind-speed, a topic which is gaining growing interest in the field of renewable generation assessment, both in view of wind power production evaluation and the wind-tower mechanical reliability and safety. The first part of the paper illustrates such model starting from its origin as a generalization of the Inverse Rayleigh model - already proven to be a valid model for extreme wind-speeds - by means of a continuous mixture generated by a Gamma distribution on the scale parameter, which gives rise to its name. Moreover, its validity to interpret different field data is illustrated, also by means of numerous numerical examples based upon real wind speed measurements. Then, a novel Bayes approach for the estimation of such extreme wind-speed model is proposed. The method relies upon the assessment of prior information in a practical way, that should be easily available to system engineers. In practice, the method allows to express one's prior beliefs both in terms of parameters, as customary, and/or in terms of probabilities. The results of a large set of numerical simulations - using typical values of wind-speed parameters - are reported to illustrate the efficiency and the accuracy of the proposed method. The validity of the approach is also verified in terms of its robustness with respect to significant differences compared to the assumed prior information.

Keywords: Renewable energy; Bayes estimation; Beta distribution; Lognormal distribution; Compound Inverse Rayleigh distribution; Extreme values; Safety; Wind Power

1. Introduction

In recent years, renewable energy has become one of the major research topics among power system engineers. Especially wind energy is considered one of the largest potential renewable energy able to generate a great amount of electricity with a low value of polluting emissions into the atmosphere [1-2].

The potential of wind energy is strictly related to the availability and the intensity of wind in a certain geographical area. These two characteristics depend on the climatic conditions as air temperature, wind speed, solar radiation, and other factors. Therefore, the wind electrical energy use needs a deep analysis of the wind conditions.

In particular, the electrical power produced by the wind energy is modelled as a statistical distribution of the wind speed (WS) random variable (RV); this determines that the estimation of wind electrical energy production is strictly related to the accuracy of the adopted wind speed distribution [3-16]. So, an accurate wind speed distribution modeling is the first step to achieve accurate wind energy production estimation. Moreover, as discussed in [17-32], the randomness of WS also has a great impact on the mechanical reliability of wind power systems, since extreme values of wind speed (EWS) may damage sensible components of the structures, such as towers and wind blades, so that EWS characterization also constitutes a basic tool for an efficient wind turbine design. Furthermore, as also pointed out in [15-19], values of WS that are greater than the "cut-off" value of the wind generator are generally undesirable, since the electric generator has to be disconnected from

the wind turbine to avoid damages to the electrical section of the wind power system; consequently, the “cut off” value of the generator must be chosen according to the characterization of EWS in the particular location, since it has a great impact on aggregate power production [1-4, 21-23]. So, EWS is also indirectly related to the electricity production, as discussed in [4], although the paper is primarily devoted to the issue of mechanical safety. Furthermore, as reported in [24], a reliable assessment of the expected EWS values is necessary – still in view of guaranteeing a desired safety level- also for the design of structures at exposed locations (e.g. bridges, radio masts), and not only for the wind turbines.

For such purpose, a new model, the Compound Inverse Rayleigh (CIR) distribution, is proposed and illustrated in the paper. Moreover, a novel Bayes method of inference for such model is adopted in the paper, in view of designing an efficient method of estimation.

Concerning the topic of estimation, it should be remarked that EWS characterization is also important in order to attain an efficient estimation of the upper quantiles of the wind speed. Indeed, wind power statistics is highly affected by the “extreme tails” of the probability distribution of the wind speed random variable [24-28]. The upper quantiles estimation of the wind can result very crucial, especially in view of the fact that such quantiles are to be transformed into the corresponding quantiles of the “Wind Power Density” (WPD), formulated in W/m² by the well-known cubic rule [1-4]:

$$\text{WPD} = \frac{1}{2} \rho W^3 \quad (1)$$

where W is the wind-speed in m/s; ρ is the air density (nominally 1.225 kg/m³ at 15.55 °C and 101.325 Pa). Indeed, it is known from basic statistical properties, a given p -quantile, w_p , of the WS is transformed into the p -quantile of the RV $Z=WPD$ by means of same cubic rule (1), thus emphasizing any possible bias in the quantile estimations. This holds especially in the cases of upper extreme quantiles (e.g., those corresponding to probability values $p=0.95$ or 0.99) which corresponds to very high values of power, around which the power curve is most sensitive to even very small variations of its argument, so that a small error in wind speed evaluation may cause a large error in wind power, because of the above cubic relation. That is one first reason why electric utilities need a careful wind speed statistics assessment.

With this aim, a new model, the Compound Inverse Rayleigh (CIR) distribution, is illustrated and proposed in the paper as valid model for extreme wind speeds. Such CIR model takes its name (as well as its derivation) from the more popular Inverse Rayleigh distribution [19, 33-35], which was proposed since a few decades for the purpose of modeling positive random variables, and especially for its applications in reliability and survival times studies. Such model also constitutes a particular case of the Inverse Weibull distribution [33] which, which – apart being rather popular for its reliability applications [36-38] - has been also adopted as a valid model for the characterization and estimation of EWS [6-11, 23]. The IR model is studied in this paper in relation to applications to EWS data analysis, by extending a similar approach promoted by some of the authors in a previous paper on the subject [19].

The following expressions for CIR, CDF and PDF are reported, while their derivation from the more known Rayleigh and Inverse Rayleigh models will be shown afterwards:

$$F(x|\eta) = \frac{1}{\left[1 + \left(\frac{\eta}{x}\right)^2\right]} \quad (2)$$

$$f(x|\eta) = \frac{2\eta^2}{x^3 \left[1 + \left(\frac{\eta}{x}\right)^2\right]^2} \quad (3)$$

where η is a positive scale parameter. It is apparent that η is the median value, or the “0.50-quantile” of the CIR model. Indeed:

$$F(\eta) = 0.5 \quad (4)$$

It is recalled that the p -quantile is a value x^* of a RV is such that $p100\%$ of the observed values of the RV fall below x^* , i.e. - in the case of a continuous RV as in the present case - it is the unique solution of $F(x^*)=p$, being $F(x)$ the CDF of the RV. So, the median value is equal to the 0.5-quantile, such that 50% of the observed values of the WS values fall below η . In particular, the upper quantiles of the WS probability density function (PDF) are a key topic of the present study.

The rest of this paper is organized as follows. In section 2, the CIR model is developed, and illustrated with some details also concerning its deduction from other models, or its similarities with other, more popular models, such as the Pareto one [24,33].

Then, Sec. 3 shows many applications of the CIR model to extreme wind speed real data. In Section 4, we introduce a suitable Bayesian approach [39-41] for the estimation of the SI, with reference to the CIR model. The method constitutes a "practical" Bayesian statistical inference approach [40] for the estimation of the EWS distribution from available data, by using a model quantile or the model CDF as the proper input for the prior assessment of the model. In order to assess the overall performance of the proposed Bayes estimation, a large series of Monte Carlo simulations have been carried out, whose results are illustrated in the successive section 5, on the basis of typical values of real EWS data. Such results assess the good performance of the proposed approach, especially when compared to "classical" estimation methods, such as the Maximum Likelihood estimation, the method of Moments, or the Quantile Estimation method [28-29,39-41]. Moreover, to highlight a special feature of the proposed method, in the final part of section 5 the robustness of the proposed estimation method is illustrated, by means of a further large series of simulations. Such "robustness analysis" clearly shows that the same estimation efficiency holds, at least in most of the examined cases, also when the true prior model is very different from the one assumed for the random parameter under study.

2. The Compound Inverse Rayleigh distribution as a Wind Speed Model

2.1 CIR model genesis

As stated in Sec. 1, the CIR distribution is obtained from a mixture of IR distributions, arising from the random variation of some related weather variables, such as temperature and pressure. The mixture of the shape parameter has been adopted also for the widespread Weibull model [29] in order to allow for the randomness of environmental conditions of the windfarm under study. Such continuous mixture generated by a probability distribution on a given parameter of the WS distribution allows to obtain a more flexible WS distribution. This method, when applied to the IR model, gives rise to the above introduced "Compound Inverse Rayleigh" distribution, leading to the analytical expression of its CDF and PDF reported in (2) and (3) respectively.

It is recalled that, given a positive random variable X (the EWS in this study), the IR Model CDF $F(x)=P(X\leq x)$, with positive shape parameter may be expressed by the following equation:

$$F(x|\alpha) = e^{-\left(\frac{\alpha}{x^2}\right)} \quad (5)$$

with the following probability density function (PDF):

$$f(x|\alpha) = 2\alpha x^{-3} e^{-\left(\frac{\alpha}{x^2}\right)} \quad (6)$$

In this paper, the so-called CIR model constitutes a generalization of the Inverse Rayleigh Distribution obtained by means of a continuous mixture of its parameter α . So, let the parameter α be considered – for the reasons above outlined - as an uncertain quantity, which is to be modeled in a Bayesian approach [39] through a Gamma PDF [33]:

$$g(\alpha|\delta, \xi) = \frac{\xi^\delta \alpha^{\delta-1}}{\Gamma(\delta)} e^{-\xi\alpha} \quad (7)$$

where ξ is the scale parameter and δ is the shape parameter of the PDF of α , both positive numbers, and $\Gamma(\cdot)$ is the Gamma special function defined as follows:

$$\Gamma(t) = \int_0^{+\infty} y^{t-1} e^{-y} dy \quad (8)$$

Let the EWS RV be denoted by X , and possessing the IR CDF of (5), conditional to a given value of the parameter α . Then, using the total probability theorem, the unconditional CDF of the RV X can be expressed, denoting (for covering also more general cases) $\beta = -2$, as:

$$F(x) = \int_0^{+\infty} g(\alpha) e^{-\alpha x^\beta} d\alpha \quad (9)$$

and, by means of some straightforward manipulations, as follows:

$$F(x) = \frac{\xi^\delta}{\Gamma(\delta)(\xi + x^\beta)^\delta} \int_0^{+\infty} [\alpha(\xi + x^\beta)]^{\delta-1} e^{-\alpha(\xi + x^\beta)} d[\alpha(\xi + x^\beta)]$$

So, equation (9) can be rewritten as:

$$F(x) = \frac{\xi^\delta}{(\xi + x^{-2})^\delta} \quad (10)$$

or, denoting by $\eta = \xi^{1/\beta}$, the following expressions for the CDF and PDF are obtained:

$$F(x|\eta, \delta) = \frac{1}{\left[1 + \left(\frac{\eta}{x}\right)^2\right]^\delta} \quad (11)$$

$$f(x|\eta, \delta) = \frac{2\delta\eta^2}{x^3 \left[1 + \left(\frac{\eta}{x}\right)^2\right]^{\delta+1}} \quad (12)$$

where η, δ are a scale and a shape parameters, respectively. The above model belongs to the family of "Inverse Burr" distributions [28-29,33]. Assuming $\delta = 1$, so that the mixing Gamma distribution becomes an Exponential distribution, the expression of the CIR distribution of (2) and (3) is obtained.

2.2 CIR model statistical parameters

In this sub section, some graphical examples are illustrated for the CIR model, as well as its comparisons with the IR and similar models.

By virtue of (3), the median value – or the 0.5-quantile - of the CIR distribution is simply:

$$m = \eta \quad (13)$$

Any p -quantile X_p may be expressed by means of the median value m by:

$$X_p = \frac{m}{k_p} \quad (14)$$

with $k_p = \sqrt{\frac{q}{p}}$ and $q = 1 - p$. Of course, if $p=0.5$, then $k_p=1$, and (14) coincides with (13).

The mean value is expressed by:

$$\mu = m \frac{\pi}{2} \quad (15)$$

2.3 CIR model examples related to WS distribution

In Fig.1, three typical CIR PDF curves corresponding to the analytical expression of (1)-(2) are shown, characterized by three different values of the parameter η (10, 25, 40). The values on the x-axis are WS values in m/s.

It may be shown indeed that the two models become close numerically for large values of the WS.

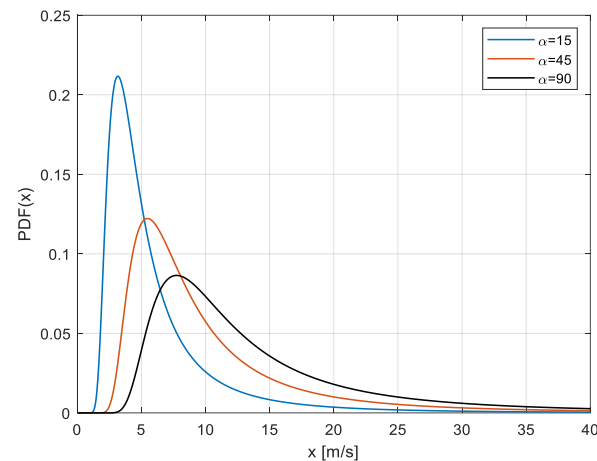


Figure 1. Various Compound Inverse Rayleigh PDFs, with three different values of the parameter η (the parameter $\alpha = \eta^2$ appears in the legend) The values on the x-axis are WS values in m/s.

It is simple to note that when the CDF argument diverges, or, in practice when the WS value x is sufficiently high, so that $x \gg \max(\alpha, \tau)$, the first-order Maclaurin series expansions of the IR and CIR CDF reported respectively in eqns. (2) and (4) show that $F(x) \cong G(x)$ when α and τ are not very different, since:

$$F(x|\tau) = \frac{1}{1 + \left(\frac{\tau}{x}\right)^2} \cong 1 - \left(\frac{\tau}{x}\right)^2 \quad (16)$$

and:

$$G(x|\alpha) = e^{-\left(\frac{\alpha}{x}\right)^2} \cong 1 - \left(\frac{\alpha}{x}\right)^2 \quad (17)$$

2.4 Comparison among similar WS models

It seems opportune to compare the CIR model with other similar models such as the IR and the Exponential model. Further on, also some hints at a possible approximation with the Pareto CDF – also widely adopted for extreme value data characterization – are illustrated in the present subsection.

First, the IR and CIR model are examined. It is interesting to compare, in order to get a better knowledge of the two models in their “extreme tails”, the quantiles of the two models. when they possess the same median or 0.50-quantile. So, let m be the common median value. The CIR model is expressible directly in terms of m , of course, being $m=\eta$, as follows:

$$F(x|m) = \frac{1}{1 + \left(\frac{m}{x}\right)^2} \quad (18)$$

Also the IR CDF (as well as the corresponding PDF) is expressible in terms of the median m , as follows, by denoting:

$$\psi = \log(2) = 0.6931 \quad (19)$$

$$F(x|m) = e^{-\psi \left(\frac{m}{x}\right)^2} \quad (20)$$

Such re-parameterization can be useful, as remarked in [37-38], in the phase of a proper model identification, since different probability models often imply similar estimated values of "central parameters" - such as the median or mean value - when fitted to the same data, but exhibit different "extreme" tails, i.e. extreme quantiles such as the ones corresponding to values of p equal to $p=0.95$ or $p=0.99$.

The generic p -quantile of the IR and CIR models - respectively denoted as X_p and X'_p , are simply expressible in terms of their median m as follows:

$$X_p = m \left[\frac{\psi}{K_p} \right]^{\frac{1}{2}} \quad \text{with } K_p = -\log(p) > 0 \quad (21)$$

$$X'_p = m [H_p]^{\frac{1}{2}} \quad \text{with } H_p = \frac{p}{1-p} > 0 \quad (22)$$

In Table I, three different p -quantiles (for $p=0.63, 0.95, 0.99$, in column 3,4,5) are reported for an IR model (row 1) and a CIR model (row 2) with the same median $m=7$ m/s. A not very large value of m has been chosen for sake of illustration, with the purpose to highlight that, also in such case, the extreme quantiles may sometimes be rather different. Moreover, the "large tails" of the above models under studies are also made more evident by reporting the corresponding quantities for a "classical" Exponential model, i.e. with CDF expressible as follows in terms of the mean value μ :

$$F(x|\mu) = 1 - e^{-\left(\frac{x}{\mu}\right)} \quad (23)$$

For a more complete comparison also the mean (expected) values of the three models are reported. The one of the Exponential model is reported as a parameter in (23), those of the IR model and the CIR model (as reported above) are respectively:

$$\mu = (\alpha\pi)^{\frac{1}{2}} = m(\Psi\pi)^{\frac{1}{2}} = 1.4757m \quad (24)$$

$$\mu' = m \frac{\pi}{2} = 1.5708 \quad (25)$$

For the Exponential model, it is well known that $m = \psi\mu$, and that the mean is equal to the p -quantile for $p=0.63$, which is the reason why such 0.63 quantile is reported, for easier reference with the other models. In all the three cases, the mean is higher than the median, as always happens with unimodal distributions with positive skewness (i.e., with "right tails"[33]).

Table 1. Some statistical parameters of the three models for WS with same median= 7 m/s.

Model	mean	median	0.63-quantile	0.95-quantile	0.99-quantile
<i>Exponential</i>	10.04	7.0000	10.04	30.25	46.51
<i>IR</i>	10.33	7.0000	8.574	25.73	58.13
<i>CIR</i>	11.00	7.0000	9.134	30.51	69.65

The curves of the PDF corresponding to the three models are shown in Fig. 2.

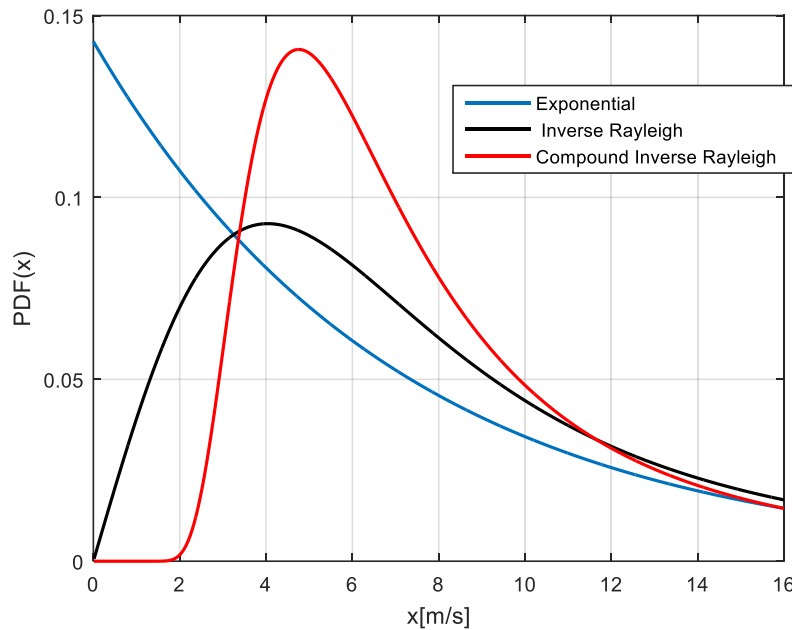


Figure 2. The curves of the three PDF corresponding to the models (the Exponential, IR, CIR) in Table I, with the same median value 7 m/s.

It is clearly apparent, from the values reported in Table I, as also from the curves, that, although possessing the same median (and about the same mean) - or more generally being similar in their central part" (including the 0.63-quantile) - the IR and CIR models behave quite different in their extreme right tails, the one of the CIR model being much larger than the one of the IR model. For instance, the 0.95-quantile has a 18% increase from the IR to the CIR model, while the 0.99-quantile has a 20% increase. Both models, of course, possess larger tails when compared to the Exponential model.

It is remarkable that the above right-hand sides in eqns. (16) and (17) can be seen as a particular case of the Pareto distribution (although only for very large x values), a very classical model for describing the extreme values distribution [31-33]. The Pareto CDF has indeed the following expression (which, if $h=2$, is the same expression of (14) and (15) for finite values of x):

$$F(x|\alpha, h) = 1 - \left(\frac{\alpha}{x}\right)^h, \text{ for } x > \alpha \quad (26)$$

A similar argument also applies to the *Inverse Log-logistic* and the *Inverse Burr* models [28-29]. (the latter is also known in literature as "*Dagum distribution*" [42]), whose CDF can be expressed respectively as follows:

$$F(x|\tau, \beta) = \frac{1}{1 + \left(\frac{\tau}{x}\right)^\beta} \quad (27)$$

$$F(x|\tau, \beta, \gamma) = \frac{1}{\left[1 + \left(\frac{\tau}{x}\right)^\beta\right]^\gamma} \quad (28)$$

being (β, τ, γ) all positive parameters.

The problem of the estimation of the CIR model is dealt by means of a proper Bayes inference approach in sec. 4. First, in the following sec. 3, many applications of the CIR model to real WS or EWS datasets are shown.

3. Application of the CIR model to real wind speed datasets

3.1. EWS datasets

The CIR model is applied to characterize real-world data collected at several locations in Greece in 2012 [25-26]. Data consists of hourly observations of wind speed, from which the EWS are extracted in terms of Period-Maxima (PM) EWS and Peaks-Over-Threshold (POT) EWS. PM values are the maximum WS values in a given time interval (one week for the present data set), while POT values are only the WS values which are greater than a given threshold value. Four locations, that significantly differ in terms of wind speed characteristics, are presented in this Section: Amintayo, Beui, Koilada, and Pontokomi.

In order to keep the dimension of data consistent, the threshold picked for building the POT EWS dataset is not the same across the locations. It is set at 10 m/s for locations Amintayo and Pontokomi, and 5 m/s for Beui and Koilada. With reference to the PM EWS dataset building, the period considered is 1 week for all the four locations. With this approach, the eight considered datasets have the size and the statistical characteristics presented in Table 2.

Table 2. Size and statistical characteristics of the datasets used in the experiments

Dataset	Size [-]	Mean [m/s]	Median [m/s]	Std. dev. [m/s]
Amintayo PM	34	6.39	4.05	4.27
Beui PM	34	6.66	6.15	3.04
Koilada PM	34	6.66	6.30	1.71
Pontokomi PM	34	8.24	8.00	2.12
Amintayo POT	119	12.21	11.50	1.99
Beui POT	296	7.32	5.90	4.03
Koilada POT	395	6.31	6.00	1.21
Pontokomi POT	52	11.02	10.75	0.78

The goodness of fit of the CIR distribution upon these data is compared to the goodness of fit of other benchmark distributions that are typically used for EWS characterization, i.e., the Inverse Burr (IB) distribution, the Inverse Weibull (IW) distribution, the Gumbel (GU) distribution [33], and the Inverse Rayleigh (IR) distribution.

To evaluate the performance of the CIR distribution and to compare it with the benchmark distributions, the outcomes of the Kolmogorov-Smirnov (KS) test [41-42], i.e., the KS test statistics (s_{KS}), is presented. The Adjusted Determination Coefficient (ADC) is presented too, in order to account for the number of parameters in each distribution [41].

3.2. Period-maxima EWS characterization

The results for the PM EWS characterization are shown in Table 3. Bold values indicate the best results for each location.

The CIR distribution has overall the best performance for the locations Koilada and Pontokomi, and it also shows the smallest KS test statistic for the location Amintayo (note that the CIR ADC for the Amintayo location is only slightly smaller than the greatest value achieved by the IW distribution). Nevertheless, the CIR distribution is not suitable to characterize the PM EWS at the Beui location, as it is unable to reach performance close to the IB distribution, which is the best pick for this location. In only one case (i.e., the GU distribution for the location Amintayo) the KS test was not successful.

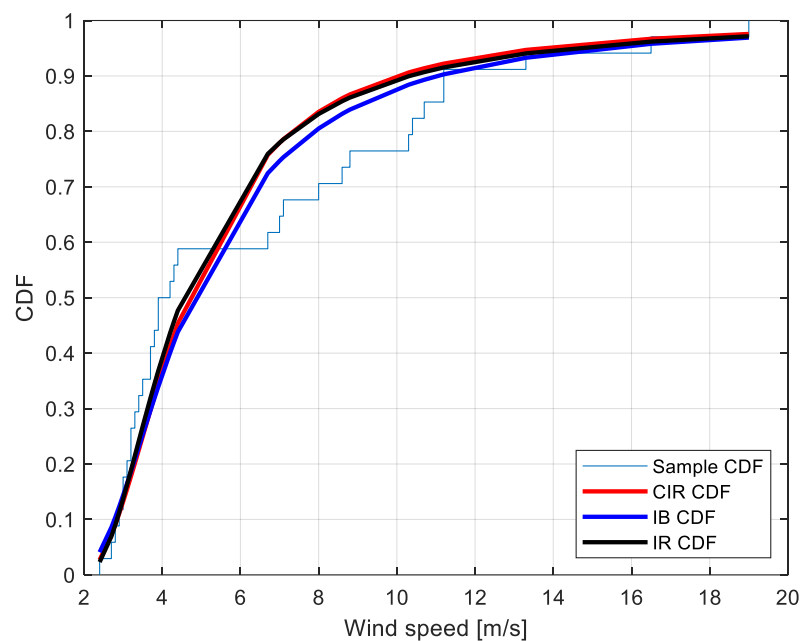
Table 3. Goodness of fit of the CIR distribution and of the benchmark distributions on the PM EWS datasets.

Bold values indicate the best results for each location.

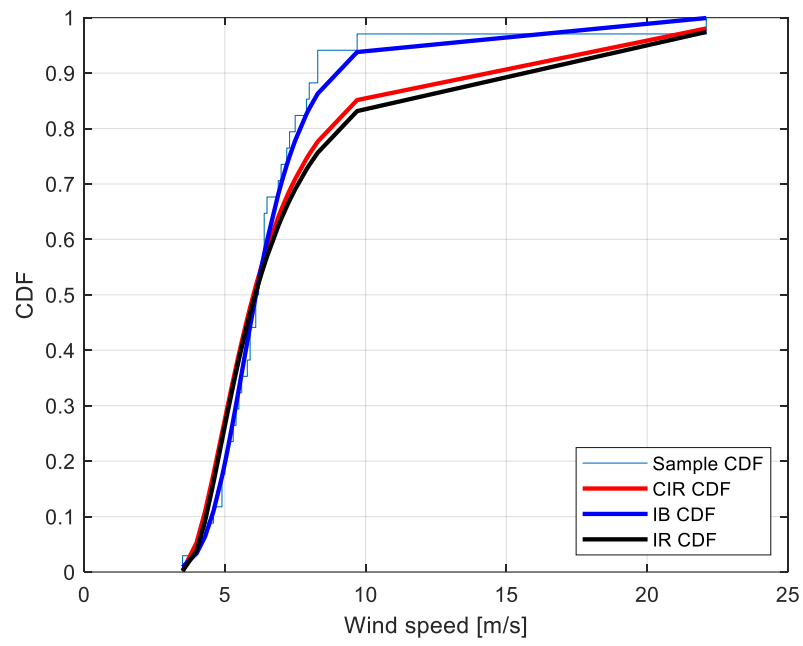
Distribution	Goodness of fit index	Dataset			
		Amintayo	Beui	Koilada	Pontokomi
CIR	s_{KS}	0.169	0.165	0.086	0.129
	ADC	0.906	0.937	0.983	0.964

IB	S_{KS}	0.170	0.079	0.100	0.130
	ADC	0.905	0.988	0.982	0.963
IW	S_{KS}	0.170	0.114	0.100	0.129
	ADC	0.908	0.973	0.983	0.960
GU	S_{KS}	0.246	0.117	0.129	0.135
	ADC	0.862	0.970	0.976	0.962
IR	S_{KS}	0.171	0.185	0.096	0.143
	ADC	0.912	0.923	0.980	0.957

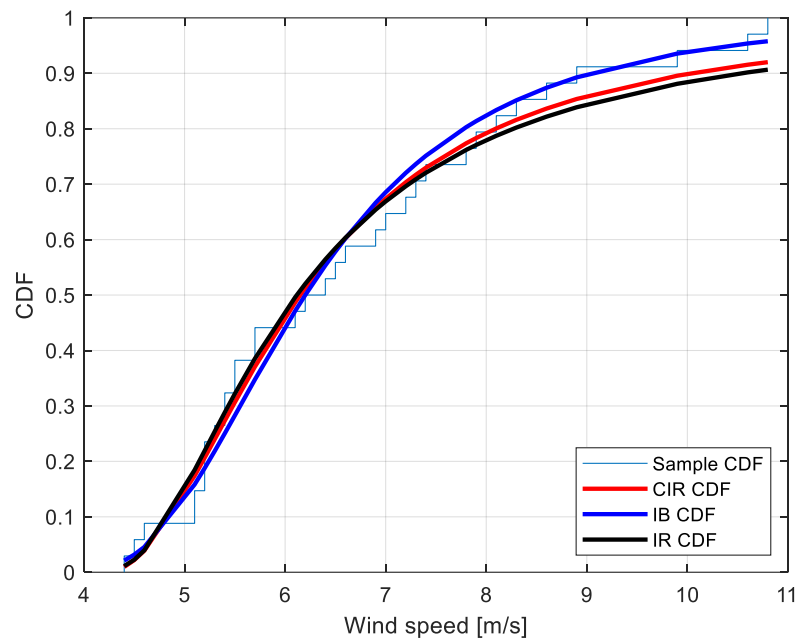
With the purpose of providing a graphical interpretation of the fitting, Figure 3 shows the sample CDF and the fitted CIR, IB, and IR distributions, for the Amintayo, Beui, Koilada, and Pontokomi locations, respectively.



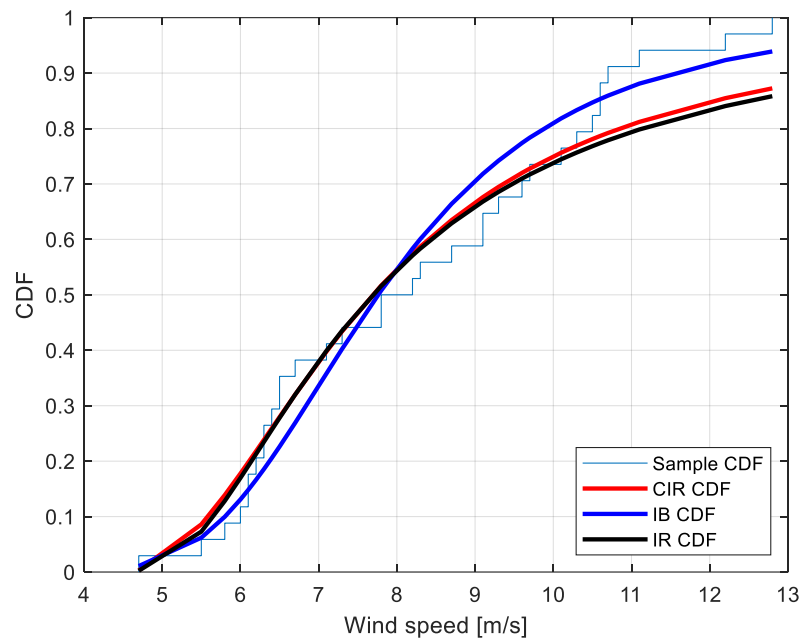
a)



b)



c)



d)

Figure 3. Sample CDF and fitted CIR, IB and IR distributions for the PM EWS at the locations:

a): Amintayo; b): Beui; c): Koilada; d): Pontokomi

3.3. Peak-over-threshold EWS characterization

The results for the POT EWS characterization are shown in Table 4. Bold values indicate the best results for each location.

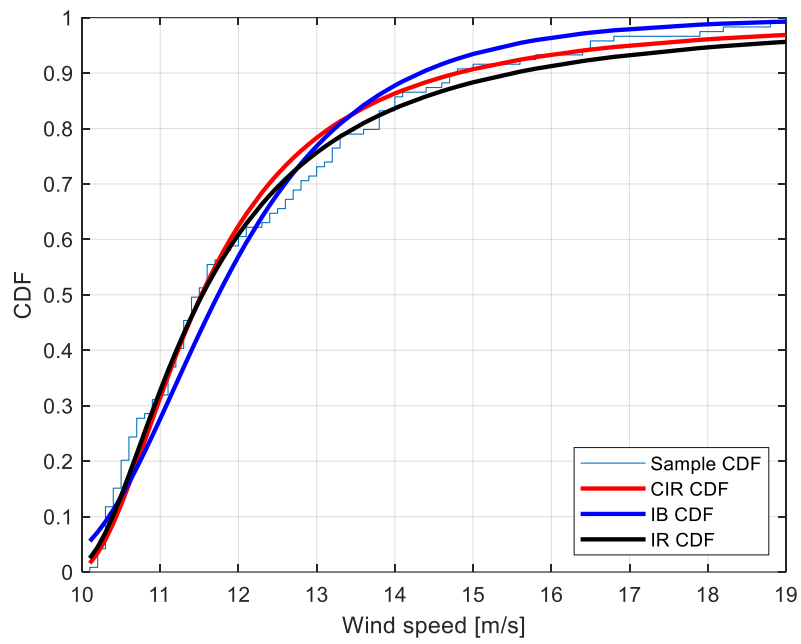
The CIR distribution always returns the smallest KS test statistic and the greatest *ADC*, in all the considered locations, denoting an overall better fitting than the fitting of PM EWS.

Table 4. Goodness of fit of the CIR distribution and of the benchmark distributions on the POT EWS datasets.

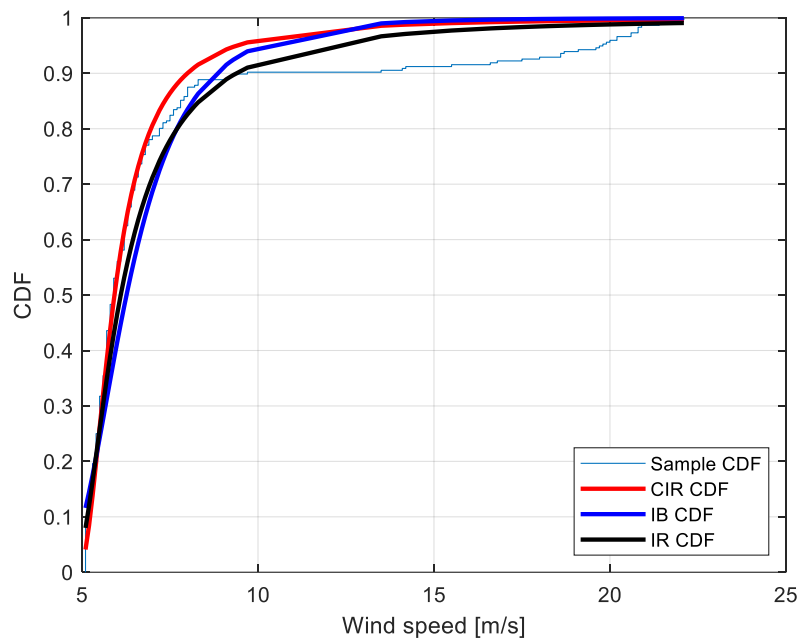
Bold values indicate the best results for each location.

Distribution	Goodness of fit index	Dataset			
		<i>Amintayo</i>	<i>Beui</i>	<i>Koilada</i>	<i>Pontokomi</i>
CIR	s_{KS}	0.086	0.084	0.091	0.107
	<i>ADC</i>	0.983	0.969	0.990	0.976
IB	s_{KS}	0.096	0.149	0.094	0.111
	<i>ADC</i>	0.973	0.895	0.979	0.963
IW	s_{KS}	0.096	0.149	0.094	0.111
	<i>ADC</i>	0.973	0.895	0.979	0.964
GU	s_{KS}	0.122	0.238	0.100	0.120
	<i>ADC</i>	0.963	0.738	0.967	0.959
IR	s_{KS}	0.089	0.102	0.104	0.177
	<i>ADC</i>	0.979	0.945	0.920	0.849

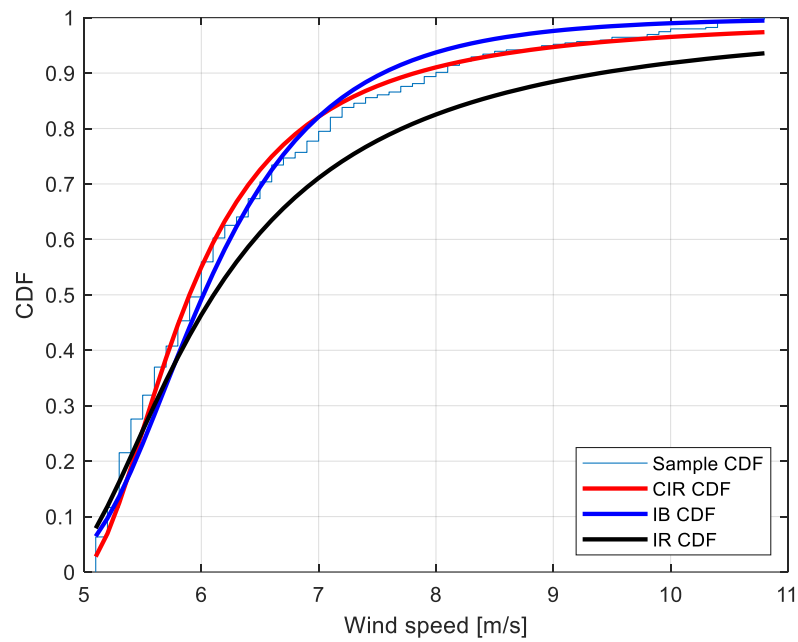
To provide a graphical interpretation of the fitting, Figures 7, 8, 9 and 10 show the sample CDF and the fitted CIR, IB, and IR distributions, for the Amintayo, Beui, Koilada, and Pontokomi locations, respectively.



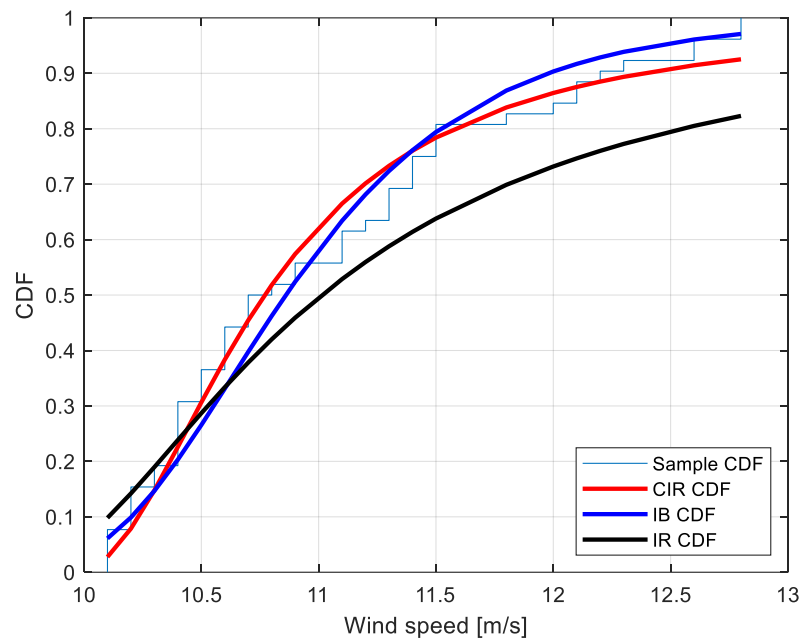
a)



b)



c)



d)

Figure 4. Sample CDF and fitted CIR, IB and IR distributions for the POT EWS at the locations:

a): Amintayo; b): Beui; c): Koilada; d): Pontokomi

4. A Bayes method of inference for the CIR model

4.1. Introduction

Since it is analytically straightforward, the classical statistical approach for the estimation of the CIR model is only hinted at in the next section, while in this section the Bayes method of inference is summarized, and then applied to the CIR model. As well known, in the classical statistical approach [41], featuring (among the others) the method of moments and the maximum likelihood (ML) method

of estimation, any parameter to be estimated is seen as a fixed but unknown variable. Instead, in Bayesian parameter estimation [39-41,43], each unknown parameter is seen as a random variable, so that it possesses a “prior” PDF representing the probabilistic information or the “belief” of the observer about the parameter.

Then, letting the parameter (possibly a vector) denoted by \mathbf{w} , its prior PDF $g(\mathbf{w})$ is integrated and updated with field data – denoted by \mathbf{D} – by the below reported well known Bayes’ theorem, which allows to obtain the the posterior distribution $g(\mathbf{w}|\mathbf{D})$ of \mathbf{w} , conditional to the observed data \mathbf{D} :

$$g(\mathbf{w}|\mathbf{D}) = g(\mathbf{w}) \frac{L(\mathbf{D}|\mathbf{w})}{C} \quad (29)$$

where:

- $L(\mathbf{D}|\mathbf{w})$ denotes the Likelihood (i.e. the joint PDF in the present case) of the data \mathbf{D} conditional to the parameter \mathbf{w} ;
- C is a constant (with respect to the parameter values), obtained by the “normalizing” integral:

$$C = \int_0^{\infty} g(\mathbf{w}) L(\mathbf{D}|\mathbf{w}) d\mathbf{w} \quad (30)$$

The above integral is a multiple one in the case w is a vector of parameters, say: $w = (w_1, w_2, \dots, w_p)$, so that $d\mathbf{w} = dw_1 dw_2 \dots dw_p$.

In a Bayesian analysis, the posterior distribution sums up the current status of knowledge about the unknown parameters of probability distributions, integrating the prior knowledge (before data observation) – or the initial belief about the parameter - represented by the prior distribution, and the knowledge produced by data in terms of the likelihood function of the observed data, which in a sense measures the information contained in the data. Once the posterior distribution, here in terms of posterior PDF $g(\mathbf{w}|\mathbf{D})$ of (29), has been deduced, there are various possible choice or the Bayes point estimate of the parameter w . A very common choice as the “best” Bayes estimate - in the mean square error sense – is the posterior mean of such PDF. As well known [39], such estimate minimizes the posterior mean squared error.

Another popular choice, here performed, is the so called “Maximum a Posteriori Probability” (MAP) estimation, in which the point estimate w° of the parameter w is the value which maximizes the posterior PDF $g(\mathbf{w}|\mathbf{D})$:

$$w^\circ = \operatorname{argmax} \{g(\mathbf{w}|\mathbf{D}), w \in S_w\} \quad (31)$$

where S_w is the parameter space, i.e. the set of all possible values that w can assume. Such choice, which by definition maximizes the probability of a correct estimation (or minimizes the estimation error probability), constitutes a “classical” method in Bayesian analysis and has gained new interest in last years, being also adopted in the framework of the recent *machine learning* methods [44-46].

Both the posterior mean and the MAP estimator, as will be apparent from the following equations, require numerical methods, but the MAP estimator is somewhat simpler to compute, being nonetheless very close to the posterior mean.

In view of the estimation of the CIR model, the parameter w to be estimated, appearing in eqns. (29)-(31), i.e. the “input” of the estimation process, may naturally be the parameter η of (2) – (4), also reported in the following eq. (32). The parameter η also represents the median of the CIR distribution. However, also any other parameter $\tau = \tau(\eta)$ related to the unknown parameter η may be chosen as the input parameter. It is recalled indeed that, in the Bayesian methodology, also any functions of the random parameter η are themselves RV, described by appropriate distributions. In particular here, once a given WS value x has been fixed, the attention may be focused on the inference on value of the CIR CDF at x :

$$F(x|\eta) = \frac{1}{\left[1 + \left(\frac{\eta}{x}\right)^2\right]} \quad (32)$$

Indeed, also the CDF – as well as other parameters (e.g. moments and quantiles) - may be considered as the input of the estimation process. It is remarked that, although formally equal to eqn. (2), the expression in (32) should not be seen as a function of the WS x , which is instead a fixed constant there. This is indeed a function of the unknown parameter η , thus constituting a new parameter, on which it should be easy to assess a prior PDF. For mathematical convenience, alternatively, by introducing the new RV:

$$Y = \frac{\eta^2}{x^2} \quad (33)$$

the above expression of the CIR CDF can be re-parametrized (omitting for clarity the dependence from x , which is a constant, so writing F instead of $F(x)$):

$$F = \frac{1}{\left[1 + \left(\frac{\eta^2}{x^2}\right)\right]} = \frac{1}{[1 + Y]} \quad (34)$$

Of course, assessing a prior PDF on the random parameter η implies – being x a given constant - assessing a prior PDF on the random parameter Y , as well as on F and other related parameters; of course, also the converse is true. The choice of the most convenient way depends on the kind of information that the observer possesses.

For instance, a reasonable prior assessment could be the one assessing the probability that the WS is higher than a given value x : this seems indeed a realistic piece of information, which should be available to the engineer based on past WS data. By the above relation, such "exceedance probability" is expressed by:

$$S = 1 - F = \frac{Y}{[1 + Y]} \quad (35)$$

being $S(x)=1-F(x)$ the "Survival function" of the WS RV. It is interesting to remark that this can be considered as a "un-safety index", accounting for the most severe EWS amplitudes as disturbances for wind tower safety.

In summary, for the above estimation, two methods are considered in the paper in order to establish an appropriate prior distribution as illustrated in the following sub-sections, i.e. assessing assigning a Lognormal prior distribution to the parameter η (sec. 4.1.), or a Beta prior distribution to the parameter S (sec. 4.2). Although, in general, the two prior assessments (i.e., assessing a prior distribution to the parameter η or a prior distribution to the parameter S) can be made equivalent by appropriate choice of the relevant prior PDF parameters, choosing one or the other as the basic input information has some peculiar analytical aspects that are highlighted in the following, since – for instance - a Lognormal prior distribution on the parameter η does not imply a Beta prior distribution on the parameter S , but a so-called "Beta-Lognormal" distribution discussed in the following, which can nonetheless approximated to certain extent by a Beta distribution.

Finally, it is remarked that the choice of assessing a prior distribution to the parameter S is in accordance with the so-called "practical" Bayes estimation, which is widely adopted nowadays in various applications after being first proposed in [36] in the framework of reliability analyses.

4.2. "First approach": choice of a Lognormal prior distribution the parameter η

Let us examine the first choice, for which a Lognormal prior PDF is suggested for its great flexibility and for its easy transformation into the corresponding prior PDF for the above introduced RV Y and S , when needed. Let η be a Lognormal RV, with the following PDF [33]:

$$g(\eta; \beta, \gamma) = \frac{1}{\beta\eta\sqrt{2\pi}} e^{-\frac{[\ln(\eta)-\gamma]^2}{2\beta^2}}, \quad \eta \geq 0 \quad (36)$$

The function $g(\eta)$ is zero for $\eta < 0$; parameters γ, β are respectively the scale and the shape parameters of the Lognormal PDF. Then, by virtue of (33), also Y is an opportune Lognormal RV, because of known properties of the Lognormal model [33] (among which, the fact that – given a Lognormal RV Y - any power function, Y^k is still a Lognormal RV, whatever the value of the real constant k). The parameters of the PDF of Y are simply deducible from those of η . Then, the RV S of (35) possesses a known PDF, which can be expressed as follows. First, the new RV :

$$T=1/Y \quad (37)$$

is introduced. It is again a Lognormal RV with PDF, here denoted by $h(t)$, simply expressible in terms of the Lognormal PDF of (36):

$$h(t) = g(\eta; \beta, -\gamma) \quad (38)$$

Then, the PDF of S is easily expressible in terms of the one of T , since:

$$S = \frac{Y}{1+Y} = \frac{1}{1/Y+1} = \frac{1}{1+T} \quad (39)$$

Then, denoting by $q(s)$ the PDF of S , it can be expressed in terms of the above PDF $h(t)$ of T by means of a "Beta-Lognormal" PDF (introduced in [47]), which – by well-known rules of PDF transformations, is given by:

$$\begin{cases} q(s) = \frac{1}{s} h\left(\frac{1}{s} - 1\right) & \text{for } 0 < s < 1 \\ q(s) = 0 & \text{elsewhere} \end{cases} \quad (40)$$

Of course, since $F=1-S$, also the prior PDF of F is simply expressible. The above PDF should be conveniently approximated by a Beta PDF, as discussed in [47]. Indeed, it is known that a Beta RV, say Z , may be characterized, under some hypotheses, as the ratio $Z=U/(U+V) = 1/(1+W)$, being U and V independent Gamma RV [33,43], and so $W=V/U$ the ratio of two independent Gamma RV. Because of the known similarity between the Gamma and the Lognormal PDF (for given common values of mean and standard deviation, [46]), it may be expected that also the PDF of Y in (39) may be at least approximated by a Beta PDF, as in our application was successfully verified. Indeed, in the following section 5 it is shown that the estimation results are not so different when assuming a Lognormal PDF on η (as discussed here), or when assuming a Beta PDF on S , as discussed in the following sub-section.

4.3. "Second approach": choice of a Beta prior distribution the parameter S

With the purpose of assessing the prior information on S , a Beta PDF [33,39,41,43] may be successfully adopted, as generally done for RV in $(0,1)$ [39], due to its great flexibility in this interval. The Beta PDF for S , $f(s)$, with positive parameters (p,q) , can be expressed as follows:

$$f(s; p, q) = \begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{y^{p-1}(1-y)^{q-1}}{\Gamma(p)\Gamma(q)} & (0 < y < 1) \\ 0 & \text{elsewhere} \end{cases} \quad (41)$$

It is noticeable that in such case also $F = 1 - S$ is a Beta RV [33], whose PDF can be obtained from (41) simply swapping the roles of parameters (p,q) . Such choice also implies that the RV Y of (35), which is obtained as a function of S (as simply deducible from (35) itself) by:

$$Y = S/(1 - S) \quad (42)$$

is a RV with support $(0, \infty)$, which possesses a so-called "Beta Prime" distribution [33], with the same positive parameters (p,q) of the Beta PDF of S . This means that the PDF $u(s)$ of Y can be expressed as follows:

$$u(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{y^{p-1}}{(1+y)^{p+q}} \quad (0 < y < \infty) \quad (43)$$

Such PDF of Y implies in turn an analytical expression also for the parameter η , which is also a RV with support $(0, \infty)$, and by means of (33) is proportional to the square root of Y :

$$\eta = x\sqrt{Y} \quad (44)$$

This means that the PDF of η can be still expressed, with some adequate transformation, in terms of the Beta Prime distribution of (43). I.e., given the PDF $u(y;p,q)$ of (43), the implied prior PDF of η , say $v(\eta)$, is expressed by:

$$v(\eta; p, q) = (2\eta/x^2)u(\eta^2/x^2; p, q) \quad (0 < \eta < \infty) \quad (45)$$

The applications of the above equations (36) and (45), expressing respectively the prior PDF of η according the first approach and the second approach, have been performed in the course of the computations of the following section, as described afterwards.

Then, both in the case of the first and the second approach, in order to obtain the Bayes estimate (be it the posterior mean or the MAP estimate), the posterior PDF shall be computed according (29), i.e. by multiplying the above prior PDF ((36) or (45), according to the case under study) and the Likelihood function $L(D|w)$. Such Likelihood function, which can be denoted by $L(D|\eta)$, since in this case is the only parameter to be estimated, is the joint PDF of the CIR RV which constitute the sample of n WS values (x_1, x_2, \dots, x_n) , i.e. it is the product of n PDF $f(x_j|\eta)$ $j=1..n$, such as the one in (3), evaluated in x_j ($j=1..n$), being n the sample size.

As long as the MAP estimate is concerned, it isn't important to evaluate the constant C in (29), since its value does not change the maximum of the above product $g(\eta) L(D|\eta)$. However, the MAP estimate must be obtained numerically (and the same would be true also for the posterior mean or other possible estimates).

5. Numerical application for estimation efficiency assessment of the proposed approach

5.1. Introduction: illustration of the Bayes estimation results

The present section illustrates a series of numerical applications for assessing the estimation efficiency of the proposed Bayes approach based on a large set of simulated experiments. All the relevant results are expressed in terms of the estimation of the only model parameter, η , here considered as an unknown. Nonetheless, as it was illustrated in previous sub-section 4.3, in some cases, according to the “second approach”, the primary source of (prior) information will not be considered the information on the parameter, but the information on a given values of the S. This efficiency assessment is carried out in “absolute” terms, as well as in terms of ‘relative efficiency’ [41-43-46] of the Bayes approach compared to the “classical” approach. As is well known [41,43-46], the classical approach is generally based on the aforementioned ML estimation method, which in the present case can be easily derived from the more general one regarding the EWS model presented in [29]. However, as discussed in [29], such MLE procedure yields sometimes some convergence problems. In such cases, simpler ways to attain parameter estimates are the Method of moments estimation [41] (briefly, “moments estimation”) and the “quantile estimation” (QE) [29] methods. In the case of the CIR model, the moment estimation procedure consists in solving a simple system of one equation, obtained equating the first sample moment (or the “sample mean”, i.e. the average value of WS in the sample) to the theoretical mean value. The QE procedure consists in equating a given sample quantile to the analogous theoretical quantile. Since the parameter η is equal to the median of the CIR model, it is convenient in this case to adopt a QE method by choosing the sample median. As intuitive from previous equations expressing the statistical parameters of the CIR model (showing that the mean and the median are related by a linear relationship), it may be expected that the moment estimation and the QE yield very close estimates for η , as was verified in the course of the computations of the present section. Moreover, in all the applications here presented they resulted practically equal to the MLE. So, for brevity and computational simplicity in the following the value of the QE estimate is reported as the “classical” estimate.

As hinted at before, the Bayes MAP estimates have been obtained, in each simulation, numerically (since an analytical solution of the above equations is not available), on the basis of the performed simulations which are described in the following sub-sections, respectively for the first and the second approach (i.e. the choice of a Lognormal prior distribution on the parameter η , or the choice of a beta prior distribution on the parameter S, as explained in section 4).

The results are reported with the following subdivision:

- 1) **Case A** (in sub-section 5.2), based upon the adoption of a Lognormal prior PDF on parameter η , (i.e. the median),
- 2) **Case B** (in sub-section 5.3), based upon the adoption of a Beta prior PDF on parameter S.
- 3) **Case C** (in sub-section 5.4), in which a “robustness analysis” is illustrated,

The three cases (A,B,C) are further subdivided into more sub-cases (A1-A3, B1-B3, C1-C4), as explained in the following .

5.2. Case A: Bayes estimation results for the “first approach” (choice of a Lognormal prior distribution on the parameter η).

Tables 5-7 illustrate the results of 3 sets of simulation cases, among the many more performed, in which a Lognormal prior PDF $g(\eta; \beta, \gamma)$ as in (36) is assumed for η , with prior mean $E[\eta]=11.5\text{m/s}$, and three different values of prior coefficient of variation (CV), namely the ratio between standard deviation and mean value), i.e. the following CV values:

$CV[\eta]=0.150$ (case A1); $CV[\eta]=0.150$ (case A2); $CV[\eta]=0.050$ (case A3),

Such three different assumptions yield indeed the three sub-cases A1, A2, A3, reported below together with the corresponding values of the above introduced Lognormal prior PDF (β, γ) of (36), namely:

- Case A1: $CV[\eta]=0.150$ ($\gamma=2.4312$; $\beta=0.1492$);
- Case A2: $CV[\eta]=0.100$ ($\gamma=2.4374$; $\beta=0.0998$);
- Case A3: $CV[\eta]=0.050$ ($\gamma=2.4411$; $\beta=0.0500$).

These three cases are selected to highlight different degrees of prior uncertainty; indeed, a greater prior CV involves a greater degree of prior uncertainty, thus the estimates are expected to be less efficient when examining the results from case A1 to case A3, as will be confirmed to some extent by the results hereafter. Out of the above cases, case A1 is to be considered the “worst” case since it possesses the highest CV value. In Fig. 5, the Lognormal prior PDF of case A2 for η is shown.

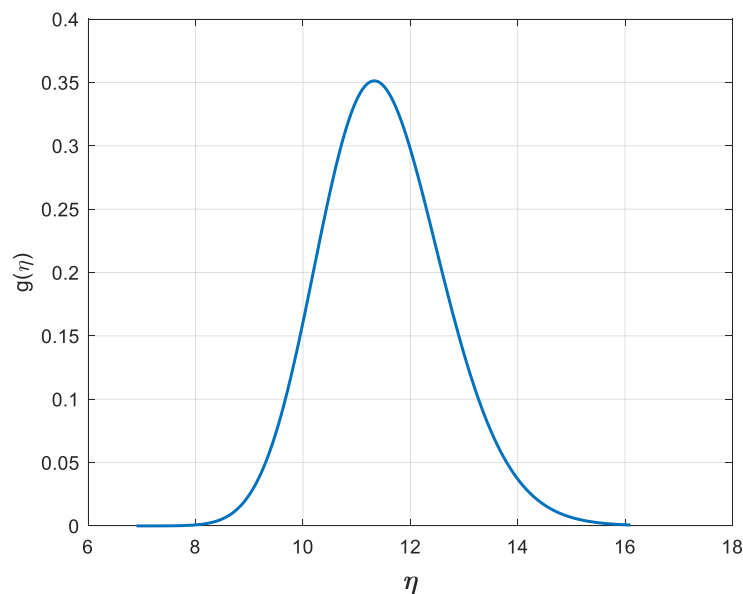


Figure 5. Lognormal prior PDF of η of case A2, with mean 11.5 m/s and $CV=0.10$ (as in equation (36), with parameter values: $\gamma=2.4374$; $\beta=0.0998$)

For all the considered cases, Monte Carlo simulations (each one performed with $N=10^4$ simulated samples of various sizes) have been used as a tool to evaluate the efficiency of the Bayes estimates, relying on the well-known basic statistics listed here below [39,41,43], all referred to the estimation of η :

- Mean Square Error of the Bayes estimator (BMSE);
- Mean Square Error of the classic estimator (CMSE);
- Relative Efficiency of the Bayes estimator with respect to the Classic estimator ($REFF=CMSE/BMSE$).

It is recalled that the classic estimator may be here computed as the MLE or the QE estimator, since the two values are in practice always very close; the figures in the tables are referred to the latter, which has the advantage of easier and faster computation. The most popular quantitative way to assess the ‘relative efficiency’ of the Bayes estimator compared to the classic estimator is the REFF

index: the more REFF overcomes unity, the more efficient is the Bayes estimate compared to the classic estimate. Furthermore, the meaningful “accuracy indexes” listed here below were also considered as a quantitative way to assess the “relative bias” (or the “precision”) of the Bayes estimator [39], which should be of course as small as possible:

- Mean Relative Error of the Bayes estimator (BMRE);
- Maximum Relative Error of the Bayes estimator (BMAXRE).

In all simulated samples for the estimation of η , the estimates are carried out for different values of sample size n - as usually happens in reality - which can be classified as small, i.e. $n=5,10$, and moderately large, i.e. $n=20,30$.

Table 5. Estimation efficiency: numerical results for case A1

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0004	0.0010	1.5276e-05	0.1528	2.4256
10	5.6830e-04	4.7757e-08	7.7401e-06	0.1193	6.1251
20	7.0528e-05	2.2098e-04	7.2437e-06	0.0724	3.1332
30	1.6313e-05	3.7336e-05	3.6917e-06	0.0369	2.2887

Table 6. Estimation efficiency: numerical results for case A2

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0001	0.0043	7.7945e-06	0.0732	76.198
10	0.0002	0.0099	1.1932e-05	0.1193	61.125
20	0.0001	0.0003	7.6365e-06	0.0764	3.6823
30	0.0001	0.0001	6.9416e-06	0.0694	2.3670

Table 7. Estimation efficiency: numerical results for case A3

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0001	0.0043	6.0268e-06	0.0603	15.040
10	0.0002	7.7193e-06	1.8654e-06	0.0187	1.7576
20	2.0628e-05	8.5296e-05	4.1176e-06	0.0412	4.1350
30	3.5564e-05	6.0538e-05	6.9416e-06	0.0491	1.7022

It is apparent from the above tables that, as desired, the indexes BMSE, BMRE and BMAXRE are very limited, while the REFF index is always larger, often much larger than 1 (unless the sample size and/or the prior CV are relatively high), accounting for a very efficient estimation procedure. A more detailed discussion of the reported results is postponed to the end of next sub-section, in order to examine also the cases (B1, B2, B3) relevant to the “second approach”.

5.3. **Case B:** Bayes estimation results for the “Second approach” (choice of a Beta prior distribution the parameter S).

According to the same methodology as in the previous sub-section, tables 8-10 illustrate the results of 3 sets of simulation cases, among the many more performed, in which the before discussed choice of a Beta PDF on the Survival function $S = S(x)$, at $x=11.5$ m/s is adopted. All the cases reported here are relevant to a prior mean $E[S] = 0.5$ at $x=11.5$ m/s (which is –in a probabilistic sense – to a certain degree equivalent to assume a prior mean of the median having value 11.5 m/s, as in cases A). As in case A, the same three different assumptions on the prior CV are held, originating the three sub-cases B1, B2, B3, reported below together with the corresponding values (p,q) of the beta prior PDF $u(p,q)$ of (43):

- Case B1: $CV[S]= 0.150$ ($p=q= 21.722$);
- Case B2: $CV[S]= 0.050$ ($p=q= 49.500$);
- Case B3: $CV[S]= 0.025$ ($p=q= 199.500$).

In Fig.6. the Beta prior PDF of case B1 for S is shown.

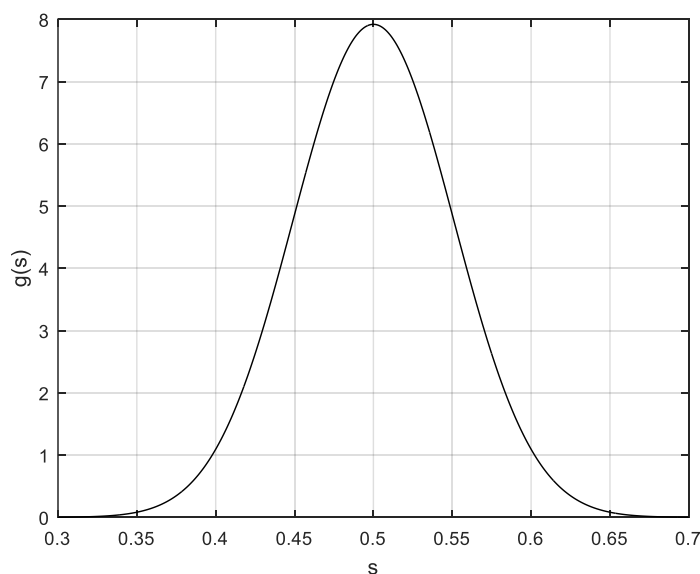


Figure 6. Beta prior PDF of S of case B1, with mean 0.5 and $CV=0.15$ (as in equation (43), with parameter values: $p=q= 21.722$)

Again, the same indices (BMSE, etc) are reported as in cases A1-A3. Such indices are all still referred to the estimation of η , for an easier comparison with previous results.

Table 8. Estimation efficiency: numerical results for case B1

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	1.6344e-04	0.0047	1.0991e-05	0.1099	28.941
10	4.1878e-05	8.4174e-05	5.1433e-06	0.0514	2.0100
20	0.0004	7.8416e-04	1.6035e-05	0.1604	1.9604
30	2.8932e-05	7.8054e-05	7.0802e-06	0.0708	1.1454

Table 9. Estimation efficiency: numerical results for case B2

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0001	0.0026	1.6415e-05	0.1641	50.603
10	0.0001	0.0037	8.2551e-06	0.0826	45.659
20	2.8932e-05	0.0010	4.8385e-06	0.0484	35.507
30	6.9708e-06	2.3737e-04	2.2241e-06	0.0222	34.051

Table 10. Estimation efficiency: numerical results for case B3

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	4.6991e-06	3.9250e-04	0.0004	0.0192	83.527
10	6.6727e-06	3.8225e-04	2.2827e-06	0.0228	57.286
20	1.9640e-05	7.3067e-04	4.0137e-06	0.0401	37.204
30	1.1367e-05	2.9914e-04	2.8232e-06	0.0282	26.317

As a general comment to the results reported in tables 5 through 10, first a look should be given to each single table, examining the results referred to the various sample sizes. In any of the three cases referred to in tables 5 through 7 for case A, and in tables 8 through 10 for case B, the results about the efficiency of the proposed Bayesian approach are very satisfactory. Indeed, the Bayes estimate errors (in terms of BMSE) are reasonably limited, for each value of n . Moreover, the relative efficiency (REFF index) with respect to the classic estimate is always larger than 1, not only for small sample sizes ($n = 1, 3$), as expected, but also for sample sizes as large as $n = 30$ (of course, as well known [39], the REFF index generally decreases with n).

Looking in succession to the three tables, the efficiency of the Bayes estimates is seen to increase from case A1 to A3, and from case B1 to B3. This was to be expected, since the prior precision (as measured by the prior CV) increases as the CV decreases. Nonetheless, the results show that also in cases A1 and B1, the “worst” cases here reported (i.e. those implying the highest CV value), the results denote an excellent efficiency. This is true both in absolute terms and with respect to the classic estimation results. Moreover, the estimation efficiency seems generally greater, in most cases, for case B than for case A.

Similar considerations, as the ones regarding the estimation efficiency; may be made with respect to the precision of the estimates, described by the indices BMRE and BMAXRE, which are always very small; this feature appears to a certain extent independent from the sample size.

5.4. Case C: Bayes estimation results relevant to a robustness analysis

The numerical assessment of Bayes approach efficiency is closed by a “robustness” analysis [39] with respect to the choice of the prior distribution (but always retaining, except in the last case, the same values of the assumed prior mean and CV). The purpose of a robustness analysis is to certify, to a certain degree, that the performance of the estimation method is not so much sensitive with respect to “wrong” prior assumptions, in a sense which will be better clarified afterwards.

In the cases examined here, and illustrated by the last four tables 11-14, it is hypothesized the Bayes computations are performed assuming a “wrong” prior distribution, be it the Lognormal model of case A for η or the Beta model of case B for S: instead, a different prior distribution generates the true values of η or S. In detail, the following more cases are illustrated in Tables 11-14, indicated overall as “Case C”, with the following subdivision:

- *Case C1*: The prior distribution of η is Normal, with the same mean (11.5 m/s) and CV=0.15 of the assumed prior distribution of η , which is the Lognormal prior PDF of case A1;
- *Case C2*: The prior distribution of η is Uniform, with the same mean (11.5 m/s) and CV=0.15 of the assumed prior distribution of η , which is again the Lognormal prior PDF of case A1. This implies that the prior distribution of η is Uniform in the interval (a,b), with: a=8.5122; b=14.4878;
- *Case C3*: The prior distribution of S is Uniform, with the same mean (0.50) and CV=0.15 of the assumed prior distribution of S, which is the Beta prior PDF of case B1. This implies that the prior distribution of S is Uniform in the interval (a,b), with: a=0.3701; b=0.6299;
- *Case C4*: The prior distribution of S is Uniform in (0,1), while the assumed prior distribution of S which is the Beta prior PDF of case B1. This implies that the prior distribution of S has the same mean (0.50) of the assumed prior, but a quite different CV (0.5477).

The four cases are representative of a wide range of cases in which the assumed prior PDF is a wrong one, and in particular, in the last case C4, a very wrong assumption is hypothesized, since the Uniform PDF in (0,1) is very different from the Beta PDF; the Uniform PDF possesses a mean value 0.5 and, as well known [33], a standard deviation equal to $1/\sqrt{12}$, and thus a CV equal to $2/\sqrt{12} \approx 0.5477$.

Table 11. Robustness: numerical results for case C1

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0013	0.0031	2.4785e-05	0.2479	2.4347
10	0.0003	6.2142e-04	1.1813e-05	0.1181	2.0714
20	0.0012	0.0028	2.2463e-05	0.2246	2.2980
30	3.9400e-05	2.7720e-04	1.6371e-05	0.1637	1.3560

Table 12. Robustness: numerical results for case C2

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0003	0.0053	1.8553e-05	0.1855	15.739
10	0.0001	0.0013	7.9363e-06	0.0794	13.576
20	0.0001	0.0003	6.9166e-06	0.0692	4.2522
30	0.0003	0.0002	1.4438e-05	0.1444	0.5367

Table 13. Robustness: numerical results for case C3

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	5.9213e-05	8.4936e-04	7.2905e-06	0.0729	14.344
10	0.0007	0.0014	1.8835e-05	0.1883	1.9937
20	0.0004	0.0006	1.4905e-05	0.1490	1.3132
30	5.2346e-04	7.5096e-04	1.701e-05	0.1701	1.4346

Table 14. Robustness: numerical results for case C4

n (sample size)	BMSE	CMSE	BMRE	BMAXRE	REFF
5	0.0012	0.0146	4.456e-05	0.4456	11.698
10	0.0007	0.0010	3.0160e-05	0.3016	1.4499
20	0.0021	0.0017	0.0001	0.6971	0.8146
30	0.0006	0.0002	3.0335e-05	0.3034	0.3938

It is remarked that any any case the less favourable prior PDFs are adopted as the reference models, i.e. those of cases A1 and B1, possessing the highest values of the CV.

The reported results show that the efficiency and the precision of the of the estimates is not so much sensitive with respect to “wrong” assumptions on the prior distribution, at least when the values of the assumed prior mean and CV are the “correct” ones. Indeed, the REFF index is less than 1 in only three cases: one is relevant to the case C2 with sample size $n=30$, the other two are relevant to the case C4 with sample size $n=20$ and $n=30$, which may be not surprising given – in such cases - the large difference between the assumed prior model and the true prior model, and the high value of CV and sample size. Also in such cases, nonetheless, the very low values of the BMSE should be noticed.

It is remarked that the above terms “wrong” and “correct” are to be interpreted with a certain degree of caution, since - in the framework of the Bayes inference - the prior assumptions have the meaning of personal or subjective beliefs, based upon the experimenter’s experience or past data, and should not be justified. Nonetheless, in order that the results could be, in a sense, also accepted in the framework of an objective estimation philosophy [39], the robustness appears to be a very desirable property, also aimed at circumventing possible criticisms that an “ad hoc” prior assumptions is made with the purpose of obtaining the desired results.

Finally, the proposed Bayes estimation method appears to be both efficient and robust. It is remarked that the proposed method of estimation appears to be both simple and very practical, since it only requires some prior guess on the median of the EWS distribution, or on the probability that the wind speed is larger than a given value, and that such choice could be expressed by very simple and flexible prior distributions, such as the Beta and the Lognormal models.

5. Conclusions

In the paper, the Compound Inverse Rayleigh model has been motivated for the probabilistic characterization of extreme values of wind-speed, also illustrating its validity in interpreting real wind-speed data sets. Moreover, the statistical estimation of the model has been developed by means of a novel Bayes estimation method, whose results have been illustrated in the second part of the paper, on the basis of a very large set of numerical simulations. Such Bayesian approach constitutes a “practical” statistical method for the estimation of the above distribution parameter from available data, proposing two alternative, simple prior models, the Beta and the Lognormal models, for the prior information assessment of a model quantile or of the model Survival Function.

Finally, the feasibility, efficiency and precision of such Bayes estimation method have been illustrated by numerical examples and a broad amount of Monte Carlo simulations performed relying on typical values of real extreme wind speed data. Such results prove the good performance of the proposed approach, especially with respect to classical estimation methods,

It is deemed that the proposed model is a valid tool to accomplish both the necessity of an efficient assessment of wind tower safety, and the necessity of a valid characterization of the statistical parameters of wind-power generation, which is also dependent on the above extreme values, e.g. in view of the establishment of the cut-off threshold. The two aspects are in fact related to each other since the crossing by part of the WS of the speed cut-off threshold determines a nihil generated power, and at the same time may pose a severe threat to the mechanical safety of the wind structures .

Before introducing its practical adoption to real data sets, and then its estimation, the model has been deduced analytically as a mixture of Inverse Rayleigh models (a mixture which originates the model's denomination), also discussing with some detail its comparison with other more popular distributions adopted for extreme wind-speeds, such as the Pareto, Gumbel, Inverse Weibull, Inverse Burr and more.

In section 5, different performance indexes, such as the MSE, the "Relative Efficiency", the “Mean Relative Error” and the “Maximum Relative Error” indices have been evaluated by means of a large set of numerical simulation, with the purpose of a thorough assessment of the performances of the estimation method, with very satisfactory results. In the final section, a numerical analysis is devoted to the robustness assessment of the methodology (i.e. using prior distributions different from the ones chosen here), again with good results., showing that the estimation efficiency holds, at least in most of the examined cases, also when the true prior model is rather different from the one assumed for the computations.

Nomenclature

<i>ADC</i>	Adjusted Determination Coefficient
<i>CIR</i>	Compound Inverse Rayleigh distribution
<i>CV</i>	Coefficient of variation
<i>CV[Z]</i>	Coefficient of variation of a RV Z
<i>CDF</i>	Cumulative distribution function, $F(x)$
<i>EWS</i>	Extreme wind speed
$E[Y]$	Mean (expected) value of the RV Y
<i>GU</i>	Gumbel Distribution
<i>IB</i>	Inverse Burr distribution
<i>IR</i>	Inverse Rayleigh distribution
<i>IW</i>	Inverse Weibull distribution
<i>KS</i>	Kolmogorov-Smirnov test
<i>PDF</i>	Probability density function, $f(x)$
<i>PM</i>	Period Maxima
<i>POT</i>	Peaks-Over-Threshold
<i>m</i>	generic median
η	median of the Compound Inverse Rayleigh distribution
<i>WS</i>	Wind speed (RV)
<i>QE</i>	Quantile estimation
<i>ML, MLE</i>	Maximum Likelihood, Maximum Likelihood estimate
<i>RV</i>	Random variable
$S, S(x)$	Survival function, $S(x) = 1-F(x)$
μ	mean value (expectation) of a RV
σ^2, σ	Generic variance and standard deviation of a RV
ζ°	denotes a Bayes estimate of the parameter ζ

Author Contributions: Conceptualization, E.C.; methodology E.C., M.F. and G.M.; validation, E.C., M.F. and G.M.; formal analysis, E.C., M.F. and G.M.; investigation, E.C., M.F. and G.M.; writing—original draft preparation, E.C.; writing—review and editing, E.C. and M.F.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to Dr. Pasquale De Falco for helping the authors in finding the datasets for the application of Section 3, and performing the computations of the relevant statistics.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jain, P. *Wind Energy Engineering*; McGraw-Hill: New York, NY, USA, 2011.
2. Letcher, T. *Wind Energy Engineering: A Handbook for Onshore and Offshore Wind Turbines*; 1st ed.; Academic Press/Elsevier: Amsterdam, The Netherlands, 2017.
3. Kantar, Y.M.; Usta, I.; Analysis of wind speed distribution. *Energy Conversion And Management* **2008**, *49* (2), 962-973.
4. Dorvlo, A.S.S. Estimating wind speed distribution. *Energy Conversion and Management* **2002**, *43*, 2311-2318.
5. Akdağ, S.A.; Dinler, A. A new method to estimate Weibull parameters for wind energy applications. *Energy Conversion and Management* **2009**, *50*, 1761-1766.
6. Villanueva, D.; Feijoo, A. Wind power distributions: a review of their applications. *Renewable and Sustainable Energy Reviews* **2010**, *14*, 1490-1495.
7. Wang, J.; Qin, S.; Jin, S.; Wu, J. Estimation methods review and analysis of offshore extreme wind speeds and wind energy resources. *Renewable and Sustainable Energy Reviews* **2015**, *42*, 26-42.
8. Jung, C.; Schindler, D. Global comparison of the goodness-of-fit of wind speed distributions. *Energy Conversion and Management* **2017**, *133*, 216-234.
9. Jung, C.; Schindler, D. Wind speed distribution selection – A review of recent development and progress. *Renewable and Sustainable Energy Reviews* **2019**, *114*, 109290.
10. Ahsan ul Haq, M.; Srinivasa Rao, G.; Albassam, M.; Aslam, M. Marshall–Olkin Power Lomax distribution for modeling of wind speed data. *Energy Reports* **2020**, *6*, 1118-1123.
11. Zhang, Y.; Wang, J.; Wang, X. Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews* **2014**, *32*, 255-270.
12. Zhang, Y.; Pan, G.; Zhao, Y.; Li, Q.; Wang, F. Short-term wind speed interval prediction based on artificial intelligence methods and error probability distribution. *Energy Conversion and Management* **2020**, *224*, 113346.
13. Tizgui, I.; El Guezar, F.; Bouzahir, H.; Benaïd, B. Wind speed distribution modeling for wind power estimation: case of Agadir in Morocco. *Wind Engineering* **2019**, *43*, 190-200.
14. Bofinger, S.; Luig, A.; Beyer, H.G. Qualification of wind power forecasts. In Proceedings of the 2002 Global Wind Power Conference, Paris, France, 2-5 April 2002, pp. 1-5.
15. Tavner, P.; Edwards, C.; Brinkman, A.; Spinato, F. Influence of wind speed on wind turbine reliability. *Wind Engineering* **2006**, *30*, 55-72.
16. Carta, J.A.; Ramirez, P.; Velazquez, S. A review of wind speed probability distributions used in wind energy analysis. Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews* **2009**, *13*, 933–955.
17. Cheng, E.; Yeung, C. Generalized extreme gust wind speeds distributions. *Journal of Wind Engineering and Industrial Aerodynamics* **2002**, *90*, 1657–1669.
18. Yan, Z.; Liang, B.; Wu, G. Wang, S.; Li, P. Ultra-long return level estimation of extreme wind speed based on the deductive method. *Ocean Engineering* **2020**, *197*, 106900.
19. Chiodo, E.; Di Noia, L.P. Stochastic extreme wind speed modeling and Bayes estimation under the inverse Rayleigh distributions. *Applied Science* **2020**, *10*(16), 5643.
20. Diriba, T.A.; Debusho, L.K. Modelling dependency effect to extreme value distributions with application to extreme wind speed at Port Elizabeth, South Africa: a frequentist and Bayesian approaches. *Comput. Stat.* **2020**, *35*, 1449–1479.
21. Kasperski, M. Specification of the design wind load. *Journal of Wind Engineering and Industrial Aerodynamics* **2009**, *97*, 335–337.
22. Akgül, F. G.; Şenoğlu, B.; Arslan, T. An alternative distribution to Weibull for modeling the wind speed data: Inverse Weibull distribution. *Energy Conversion and Management*, **2016**, *114*, 234-240.
23. Chiodo, E.; Mazzanti, G.; Karimian, M. Bayes estimation of Inverse Weibull distribution for extreme wind speed prediction. In Proceedings of the 2015 International Conference on Clean Electrical Power (ICCEP), Taormina, Italy, 16-18 June 2015; pp. 639-646.
24. Gross, J.; Heckert, N.; Lechner, J.; Simiu, E. A study of optimal extreme wind estimation procedures. In Proceedings of the 9th International Conference on Wind Engineering, New Delhi, India, 9-13 January 1995; pp. 69–80.
25. Efthimiou, G.C.; Hertwig, D.; Andronopoulos, S.; Bartzis, J.G.; Coceal, O. A statistical model for the prediction of wind-speed probabilities in the atmospheric surface layer. *Boundary Layer Meteorology* **2017**, *163*, 179-201.
26. Efthimiou, G.C.; Kumar, P.; Andronopoulos, S.; Giannisi, S.G. Prediction of the wind speed probabilities in the atmospheric surface layer, *Renewable Energy* **2017**, *132*, 921-930.
27. An, Y.; Pandey, M.D. A comparison of methods of extreme wind speed estimation, *Journal of Wind Engineering and Industrial Aerodynamics* **2005**, *93*, 535–545.
28. Chiodo, E.; De Falco, P. The inverse Burr distribution for extreme wind speed prediction: genesis, identification and estimation. *Electric Power Systems Research* **2016**, *141*, 549-561.

29. Chiodo, E.; De Falco, P.; Di Noia, L.P.; Mottola, F. Inverse Log-logistic distribution for extreme wind speed modeling: Genesis, identification and Bayes estimation. *AIMS Energy* **2018**, *6*, 926-948.
30. Gross, J.; Heckert, A. Lechner, J.; Simiu, E. Novel extreme value estimation procedures: application to extreme wind data. In *Extreme Value Theory and Applications*; Galambos, J., Lechner, J., Simiu, E., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1994; pp. 139–158.
31. Heckert, N.; Simiu, E. Extreme wind distribution tails: A peaks over threshold approach. *J. Structural Engineering* **1996**, *ASCE 122*(5), 539-547.
32. Heckert, N.; Simiu, E.; Whalen, T. Estimates of hurricane wind speeds by peak over threshold method. *J. Struct. Eng.* **1998**, *ASCE 124*(4), 445–449.
33. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*, 2nd ed.; Wiley: New York, NY, USA, 1995; Volumes 1 and 2.
34. Soliman, A.; Amin, E.A.; Abd-El Aziz, A.A. Estimation and prediction from Inverse Rayleigh distribution based on lower record values. *Appl. Math. Sci.* **2010**, *4* (62), 3057-3066.
35. Feroze, N.; Aslam, M. On posterior analysis of inverse Rayleigh distribution under singly and doubly type II censored data. *International Journal of Probability and Statistics* **2012**, *1* (5), 145–152.
36. Erto, P. The inverse Weibull survival distribution and its proper application. *ArXiv* **2013**, preprint arXiv:1305.6909.
37. Chiodo, E.; Mazzanti, G. Theoretical and practical aids for the proper selection of reliability models for power system components. *International Journal of Reliability and Safety* **2008**, *2* (1/2), 99-128.
38. Chiodo, E.; Mazzanti, G. Mathematical and physical properties of reliability models in view of their application to modern power system components. In *Innovations in Power Systems Reliability*; Anders, G.J., Vaccaro, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 59–140.
39. Press, S.J. *Subjective and Objective Bayesian Statistics: Principles, Models, and Applications*, 2nd ed.; Wiley: New York, NY, USA, 2002.
40. Erto, P. New practical Bayes estimators for the 2-parameters Weibull distribution. *IEEE Trans. on Reliability* **1986**, *31* (2), 194-197.
41. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Duxbury Press: Pacific Grove, CA, USA, 2001.
42. Kleiber, C. *A guide to the Dagum distributions*. Springer, New York, 2008
43. A. Papoulis, S.U. Pillai, "Probability, Random Variables, and Stochastic Processes". McGraw-Hill, New York, USA, 2002.
44. DasGupta, A. *Probability for statistics and machine learning: fundamentals and advanced topics*. Springer Science & Business Media, 2011.
45. Murphy, K.P. *Machine learning: a probabilistic perspective*. MIT press, 2012.
46. Dangeti, P. *Statistics for machine learning*. Packt Publishing Ltd, 2017.
47. Battistelli, L.; Chiodo, E.; Lauria, D. Bayes assessment of photovoltaic inverter system reliability and availability, 2010, *Proceedings of the 2010 International Symposium on Power Electronics Electrical Drives Automation and Motion (SPEEDAM)*, Pisa, Italy, 14-16 June 2010; pp. 1-5.



© 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).