Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2018-102-AC2, 2018

Natural Hazards and Earth System Sciences Discussions

Open Access

Interactive comment on "On the relevance of extremal dependence for spatial statistical modelling of natural hazards" *by* Laura C. Dawkins and David B. Stephenson

Laura C. Dawkins and David B. Stephenson

l.c.dawkins@exeter.ac.uk

Received and published: 10 August 2018

Key:

- **Reviewer's comment**
- *Our response*
- Additional/edited text in the manuscript

0.1 Observational errors

- **The analysis is performed on mean winds from a numerical model which are post-processed to gusts using a highly simplified model given on Line 124. These estimates of gusts will differ considerably from the true storm-max gusts experienced at sites and the influence of this error on results has potential to be significant. Therefore, errors in estimated gusts need to be quantified, and their impacts on results should be measured and presented to readers.**

  A comparison of MetUM-derived estimates with weather station data would provide a realistic measure of observational uncertainty. I suggest the rms difference in max storm gust between the authors' dataset (grid-cell encompassing Heathrow) and observed gusts for Heathrow is computed using the top N storm max gusts *observed* at Heathrow, where N is approx. 50 to focus on tail extremes. GSOD is a free source of observed weather for many stations, including Heathrow.

  Thank you, this is a very valid point that should be discussed and explored. We have addressed this point in Section 2 by first reviewing the rigorous evaluation of the Met UM footprints in Roberts et al. (2014), and presenting your suggested comparison with the GSOD data. The GSOD weather station within our London grid cell is London City, hence we have used this station rather than Heathrow. We have added this to the end of Section 2:

  Using model generated windstorm footprints for representing historical storms has benefit in terms of spatial and temporal coverage, however these estimated maximum wind-gust speeds will inevitably differ from the those observed at nearby weather stations. For example, as noted by Roberts et al. (2014), several alternative methods for parameterising wind gust speeds are available (see Sheridan (2011) for a review), which can lead to large differences in estimated gusts (10-20ms^{-1}). The validity of simplistic gust parameterisation stated above was evaluated by Roberts et al. (2014), who found an overestimation in the effect of surface roughness at stations greater than $\sim$ 500 metre altitude, leading to underestimation of MetUM modelled extreme winds in these locations. In addition, within this thorough evaluation of MetUM windstorm footprints, Roberts et al. (2014) identified a slight underestimation in extreme wind gust speeds greater than ~ 25ms^{-1}. This was found to be due to a number of mechanisms including the underestimation of convective effects and strong pressure gradients, leading to the underdevelopment of fast moving storms (Roberts et al. (2014)).

[Figure 1 in attachments] (a) The relationship between MetUM windstorm footprint wind gust speeds in the London grid cell and the corresponding observed wind-gust speeds at the London City weather station within the Global Summary Of the Day dataset, and (b) the same relationship for the 50 must extreme windstorm events at the London City weather station.

To explore the possible discrepancy in the MetUM windstorm footprint wind gust speed relevant for this study, we extract daily maximum observed wind gust speed recorded at the London City weather station (the station located within the London grid cell used throughout this study) from the Global Summary Of the Day (GSOD) data repository ($https://data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod$), and, for each of the 6103 windstorm events in our dataset, find the maximum observed gust in the 3 day period centred on the same date as in the MetUM model generated footprints. A comparison of the observed and MetUM modelled footprint wind gusts in London is presented in [Figure 1 in attachments] (a), indicating a general overestimation in modelled wind-gust speeds below 25m^{-1} and a slight underestimation for wind-gust speeds above 25m^{-1}, reflecting the findings of Roberts et al. (2014). [Figure 1 in attachments] (b) presents this same relationship for the 50 most extreme events in the observed dataset, highlighting this underestimation of modelled extreme wind-gust speeds. For example, during the windstorm in which the highest observed wind-gust speed of 34.98 ms^{-1} occurs at the London City station, the MetUM model produces a maximum wind-gust speed of just 23.88 ms^{-1}. Indeed, the root mean squared difference between the observed and modelled footprint wind-gust speeds for these 50 extreme events is 4.57ms^{-1}, giving an indication of the model uncertainty in representing extreme windstorm footprint wind-gust speeds.

The discrepancy in model generated wind-gust speeds compared to the observations could lead to differences in results, namely the identification of the extremal dependence class between locations. To explore this possibility we repeat parts of the analysis presented in the following sections using this GSOD data, discussed further in Section 3. It should also be noted, however, that the station data should not be treated as the true state of the world, since a number of factor, such as instrumental inaccuracies, lead to observational error. A full exploration and quantification of the observational uncertainty present in the MetUM model as well at the observations themselves is beyond the scope of this study, however this discrepancy should be kept in mind when interpreting results.

- **The confidence intervals in Figure 4 of original manuscript are based on sampling error and need replaced to include these estimates of observational error for each MetUM storm max gust. Figures 3 and 5b would also benefit from the inclusion of estimates of uncertainty in plotted values, due**

to both observational and sampling errors.

While this would be a very beneficial addition to the paper, as expressed above, we feel that the full exploration and quantification of the observational uncertainty present in the MetUM model (as well at the observations themselves) is beyond the scope of this study, and we are unsure of how to translate the RMS difference into confidence intervals. We do, however, agree that the effect of the model bias on the results of the analysis should be explored. We have therefore reproduced the plots in Fig. 2 (a) and Fig. 3 (a) of the original manuscript using GSOD data for stations within the London and Amsterdam model grid cells. We have added this plot in the supplementary material and refer to it at the end of Section 3.2:

As noted in Section 2, the discrepancy in model generated wind-gust speeds compared to observations could lead to differences in the identification of the extremal dependence class between locations. To explore this discrepancy we extract daily maximum wind gust speeds at Amsterdam Schiphol Airport (weather station within the Amsterdam grid cell) from the Global Summary of the Day (GSOD) data set, and calculate windstorm footprint maximum wind gust speeds as was done for London City. [Figure 2 in attachments] (a) in the Supplementary Material presents a comparison of these footprint maximum wind gust speeds for London City and Amsterdam Schiphol Airport, the observation equivalent of Fig. 2 (a), and [Figure 2 in attachments] (b) in the Supplementary Material presents the empirical extremal dependence coefficient $\chi(p)$ for $p \in [0, 0.4]$ calculated for this pair of observation stations, the observation equivalent of Fig. 3 (a). This comparison indicates that footprint wind gust speeds at the London and Amsterdam weather stations are, in fact, less related in the extremes since, the most extreme events in each location do not coincide. This is reflected in the empirical extremal dependence coefficient in [Figure 2 in attachments] (b), which deceases more rapidly than in Fig. 3 (a) as $p \rightarrow 0$. This difference may be due to the underestimation of extreme wind gust speeds in the MetUM model, as discussed in Section 2 and Roberts et al. (2014), resulting in the modelled footprints missing the most extreme winds and hence producing wind gust speed that are more similar, and less extreme, in both locations. The empirical indication of asymptotic independence is, however, consistent for both observations and MetUM modelled windstorm footprints.

[Figure 2 in attachments] (a) Scatter plot comparing observed windstorm event maximum wind gust speeds (ms^{-1}) at London City and Amsterdam Schiphol Airport GSOD weather stations, (b) extremal dependence measure $\chi(p)$, for $p \in [0, 0.4]$, for observed windstorm footprint wind gust speeds at London City GSOD weather station paired with Amsterdam Schiphol Airport GSOD weather station.

0.2 Events analysed in Figure 4, and interpretation of results

- **Fig 2 indicates approx. 25 points above quantile=0.99, which suggests that the quantile=0.5 in Figure 4 is based on over one thousand storm events in a 35 year period. The inclusion of about 30 events per year on average will contain many breezy days. These data points are potentially misleading to include, because the spatial structure of days with weak winds is likely to be substantially different from the spatial structure of severe events producing tail winds. I request that Figure 4 is re-drawn using data from quantile=0.9 and upwards. This would still include weak winds from minor cyclones, but is a step in the right direction, while maintaining sufficient sample sizes. The conclusions to be drawn from Figure 4a should be reviewed in a revised version of manuscript. First, the results in Figure 4a indicate rising values of the coefficient of tail dependence for quantile**

thresholds above 0.9, towards a value of unity for the highest quantile. Given the aim to capture behaviour in the limit as p tends to 0, it seems unsafe to conclude that London-Amsterdam has tail independence. Second, the inclusion of observational uncertainty (point 1 above) will broaden the confidence limits which may require a new interpretation of results.

Thank you for this feedback, however we feel there must be some confusion in the interpretation of the coefficient of tail dependence (η). This parameter is equivalent to the scale parameter of a Poisson process (or the shape parameter of a GPD (see Ledford and Tawn (1996)). Therefore η is a parameter of a model for the joint excesses of the pair which determines the asymptotic behaviour, and should be chosen such that the model is stable above the threshold (similar to when choosing the threshold in the Generalised Pareto Distribution (GPD) model), not an asymptotic measure of dependence like χ and $\bar{\chi}$, which we are interested in as $p \to 0$.

We have now made this interpretation of η clearer by editing the paragraph after Eqn. (4):

We fit this model to the pairs London-Amsterdam and London-Madrid, requiring the specification of the high threshold, w, above which the Poission process model is fit. As discussed by Ferro (2007), this threshold selection is a trade-off between being low enough that enough data is attained to ensure model precision, but high enough that the extreme-value theory that motivates the model provides accurate estimates, suggesting we should select the lowest level at which the extreme-value approximations are acceptable (Ferro (2007)). In a similar way to choosing the appropriate threshold when fitting a Generalised Pareto Distribution (see Coles (2001)), empirical diagnostic plots can be used to inform this selection. For example parameter stability plots, in which the estimated model parameters and mean excess should be constant above the chosen high threshold; and quality of fit plots, in which for this model, the transformed excesses, $(Z - w)/\eta$, should be exponentially distributed if an appropriately high threshold has been chosen (see Ferro (2007) for more details).

Here, the 85$\%$ quantile of the structural variable T is selected, based on these diagnostic plots (examples for these plots for London-Amsterdam are presented in [Figure 3 in attachments] in the Supplementary Material). This choice is similar to the 0.88% and 0.9% thresholds selected in the applications of Ferro (2007) and Ledford and Tawn (1996) respectively.

[Figure 3 in attachments] Threshold selection diagnostic plots for the Ledford and Tawn (1996) model: (a) Quantile-Quantile plot comparing the transformed excesses, $(Z - w)/\eta$, with the Exponential distribution (with rate parameter equal to 1), where w is the selected 0.85$\%$ quantile of the structure variable T, (b) stability plot for the mean excess as a function of threshold w.

The diagnostic plots in Fig. 4 are equivalent to those shown in Figures 3 and 4 in Ledford and Tawn (1996) and Figure 2 in Ledford and Tawn (1997). In both of these papers they use the range (0.5-1). We agree that a threshold as low as 0.5 is most likely too low to give an appropriate extreme-value theory model, however, we would like to keep this range on the plots to be in line with the aforementioned papers. We have added a clarification of this replication of their approach by editing the paragraph after Fig. 4:

As in Ledford and Tawn (1996, 1997), here this is done by observing the proportion of time $\eta = 1$ is within the profile likelihood confidence interval for η, when estimated using w in the interval of $0.5 - 1$ quantile of T. The pair

(X, Y) are said to be asymptotically dependent if $\eta = 1$ is contained within these confidence intervals for a majority of the range of w, and asymptotically independent otherwise.

0.3 Section 3.4 conjecture

- **The conjecture to explain the tail independence in section 3.4 begins by representing storm winds as isotropic turbulence. It is standard to represent storm winds as the sum of a mean wind and a smaller turbulent contribution. This is also consistent with the MetUM model gust dataset used by the authors (description around Line 124). The authors then assume that gusts at two locations are bi-variate normal. While the turbulent contribution to winds at two distinct locations might be bi-variate normal at any instant in time, the gusts analysed for tail dependence are the maximum gust over the whole storm. The storm-max gusts between neighbouring locations are expected to have strong tail dependence since they would have very similar mean wind and max gustiness from isotropic turbulence.**

  Thank you for raising this interesting issue, which we had not addressed in the turbulence discussion. This paragraph of explanation has now been added after the first paragraph in Section 3.4:

  It is useful to first consider the more tractable problem of dependency in simultaneous wind speeds rather than maximum wind speeds over a given time period. The dependency between maximum gust speeds over 3 days will not generally be less than the dependency between simultaneous wind gust speeds because maximum wind gusts for a storm do not occur at the same time at different locations. However, for locations that are close to one another, maximum gust speeds for fast moving extreme storms will occur within a short time window (e.g. within around 3 hours or less for extreme storms over the UK) and so simultaneous results become more relevant.

- **Regarding the text on Lines 254-258: McNeil et al. (2005) showed that if correlation is less than one, then the coefficient of upper tail dependence equals zero, their Example 5.32. McNeil, A J, Frey, R, and Embrechts, P: Quantitative Risk Management Concepts, Techniques, Tools. Princeton University Press, 2005**

  Thank you for this reference. We have now added this in line 274:

  So what can be deduced about the extremal dependence class of wind speeds from such turbulence models? Firstly, as shown in Example 5.32 of McNeil et al. (2005), since the individual velocity components are bivariate normal, the individual velocity components are asymptotically independent at different locations e.g. $u_1 = u(s_1)$ and $u_2 = u(s_2)$ are asymptotically independent when s_1 differs from s_2, and likewise for $v(s)$.

0.4 Section 4 on losses

- **Lines 277-283: the authors state a simple loss function provides better storm loss estimates than the Klawa and Ulbrich (KU) loss function. There are various reasons why this judgement on loss functions is misleading. Besides the minor fact that the two articles quoted excluded population weighting hence did not test the KU loss function, there is a more significant issue that 'better' is defined in non-standard and highly specific**

terms as 'a subset of 23 significant storms completely contained in highest quantile of all storms'. Further, there is much published work on how loss severity is a function of wind speed, and KU's loss function certainly captures this effect more accurately than a step function.

The 'conceptual loss function' used by the authors could be more accurately described as areal frequency of loss occurrence, and ignoring loss severity. Its usefulness in estimating total loss is an interesting result, since it suggests the area of storm above a loss-causing threshold is the dominant contributor to total storm loss. I suggest the authors describe their loss function in more specific terms as 'areal frequency of loss' in the text.

Further, if the authors wish to retain text comparing a step function to the superior KU loss function, then the authors should include more information for readers: errors in loss estimates depend on wind speeds, loss functions and exposure density, and the success of the simplest loss function over KU in tests performed by Roberts et al. is very likely due to its relative insensitivity to errors in other components of their loss estimates, such as estimated gusts. This helps resolve the dilemma of a rapid growth of loss with windspeed indicating KU, while a less sophisticated testing framework indicated a step function.

We agree that this introduction of the loss function needs to be more detailed and balanced. In combination with this comment and those of Reviewer 1, we have decided to restructure Section 4 and introduce a more generic form for the loss function. In doing so, we discuss the KU loss function along with other functions and give a more detailed justification of our chosen loss function, incorporating your comments and concerns as caveats. After the first paragraph

Printer-friendly version

Discussion paper

of Section 4 we have edited:

Similar to other natural hazard loss models, in the absence of confidential insurance industry exposure and vulnerability information, it has become common in the literature to define conceptual windstorm loss as a function of the footprint wind gust speeds (see Dawkins et al. (2016) for a review). While these conceptual windstorm loss functions vary in the detail of their composition, it is possible to express most in a general form, for the pair (X, Y), as: L(X,Y) = g[V(X)e(X)H{X-U(X)} + V(Y)e(Y)H{Y-U(Y)}] where V is a function the wind gust speeds characterising the magnitude of the hazard, e represents exposure (e.g. population density), U quantifies a high threshold of the wind gust speed above which losses occur, H is a Heaviside function such that $H\{m\} = 1$ if $m > 0$ and $H\{m\} = 0$ otherwise, and g is an additional function applied in some cases to reduce skewness. For example, in the widely used and rigorously validated conceptual loss function of Klawa and Ulbrich (2003), $V(X) = (X - x_{0.98})^3$, $U(X) = x_{0.98}$ (where $x_{0.98}$ is the 98% quantile of X) and $e(X)$ is represented by the population density at the location (with equivalent expression for Y), while Cusack (2013) used a loss function in which $V(X) = (X - x_{0.99})^3$, $U(X) = x_{0.99}$, the 99% quantile of X, and $g[\cdot] = \sqrt[3]{\cdot}$. See Table 2.1 in Dawkins (2016) for a summary of previously published conceptual loss functions in terms of the components of Eqn. 0.4.

More recently, Roberts et al. (2014) presented an exploration of the success of a number of these conceptual windstorm loss functions in representing insured loss throughout the European domain, based on the same data set as in this study, with the aim of developing a method for selecting extreme storms for the eXtreme WindStorms (XWS) catalogue. While there is much published work on the existence of a relationship between loss severity and the magnitude of the wind, in particular the cubed excess wind as used in the loss functions of

Klawa and Ulbrich (2003) and Cusack (2013), Roberts et al. (2014) found that a conceptual loss function representing just the area in which the windstorm footprint exceeds a high loss threshold (i.e. $V(X) = 1$ and $e(X) = 1$ in Eqn. 0.4) to be more successful at characterising a subset of extreme windstorms known to have caused large insured losses. It should be noted however, that this exploration did not include population density within the Klawa and Ulbrich (2003) loss function, and was therefore not a direct comparison of this measure. In addition, an alternative subjectively selected subset of extreme storms may have given an alternative result, and the success of this simplistic 'areal frequency of loss' function in representing losses in this climate model generated data set of windstorm footprints may be due to its relative insensitivity to errors in other components of the loss estimates, such as estimated gusts, and may not perform as well as other loss functions if applied to wind gust observations.

However, following the results of Roberts et al. (2014) in the context of this data set, and in line with Dawkins et al. (2016), within this study we propose a similar threshold exceedance conceptual loss function. Roberts et al. (2014) used an exceedance threshold of 25ms^{-1} while Dawkins et al. (2016) used a threshold of 20ms^{-1}, as is commonly used by German insurance companies (Klawa and Ulbrich (2003)). Here, similar to Klawa and Ulbrich (2003) and Cusack (2013), we propose a locally varying wind gust speed quantile threshold, accounting for local adaptation to varying wind intensity.

- **The over-estimation of joint loss probabilities in the maps in Figures 7e $\&$ f are explained as a mis-specification of asymptotic dependence (lines 308-309). However, it could be due to a too high estimate of the dependence parameter r. Could the authors include in the text the group of**

data used for estimating the dependence parameter? *We have now edited the paragraph beginning on line 223 in the original manuscript to include this information:*

The Gumbel bivariate copula model characterises asymptotic dependence with the degree of dependence quantified by parameter r. For each pair of locations, this parameter is estimated via maximum likelihood using the `copula` R package. The Gaussian bivariate model characterises asymptotic independence with dependence parameter ρ, here, for each pair of locations, represented by the Spearman's rank correlation coefficient. Both models are fit to the full bivariate data pair, as presented in Fig. 2. For the Gumbel model the data is transformed to uniform margins using the empirical distribution function. The same transformation is made for the Gaussian model, followed by a transformation to Gaussian margins using the standard normal distribution function. The parametric forms of $\chi(p)$ and $\bar{\chi}(p)$ for these two opposing models are shown in Table 1. In Fig. 3, the Gumbel model is calculated as in Table 1, however, since the closed form definition for the Gaussian model in Table 1 only holds for the limit $p \to 0$, for this model $\chi(p)$ and $\bar{\chi}(p)$ are estimated as the median of 1000 parametric bootstrap simulations from the associated bivariate normal distribution.

0.5 Technical Corrections

- **There are many instances of 'apposing' when 'opposing' may be more appropriate?**

Thank you, we have now corrected these.

References

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Cusack, S. (2013). A 101 year record of windstorms in the Netherlands. *Climate Change*, 116:693–704.

Dawkins, L. C. (2016). *Statistical modelling of European windstorm footprints to explore hazard characteristics and insured loss*. PhD thesis, University of Exeter, College of Engineering Mathematics and Physical Sciences.

Dawkins, L. C., Stephenson, D. B., Lockwood, J. F., and Maisey, P. E. (2016). The 21st century decline in damaging european windstorms. *Natural Hazards and Earth System Science*, 16:1999–2007.

Ferro, C. A. T. (2007). A probability model for veryfiying deterministic forecasts of extreme events. *Weather Forecasting*, 22:1089–1100.

Klawa, M. and Ulbrich, U. (2003). A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Natural Hazards and Earth System Sciences*, 3:725–732.

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.

Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society*, 59(2):475–499.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press.

Roberts, J. F., Dawkins, L., Youngman, B., Champion, A., Shaffrey, L., Thornton, H., Stevenson, D. B., Hodges, K. I., and Stringer, M. (2014). The XWS open access catalogue of extreme windstorms in Europe from 1979 to 2012. *Natural Hazards and Earth System Science*, 14:2487–2501.

Sheridan, P. (2011). Review of techniques and research for gust forecasting and parameterisation: Forecasting research technical report 570. Technical report, Met Office.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2018-102, 2018.

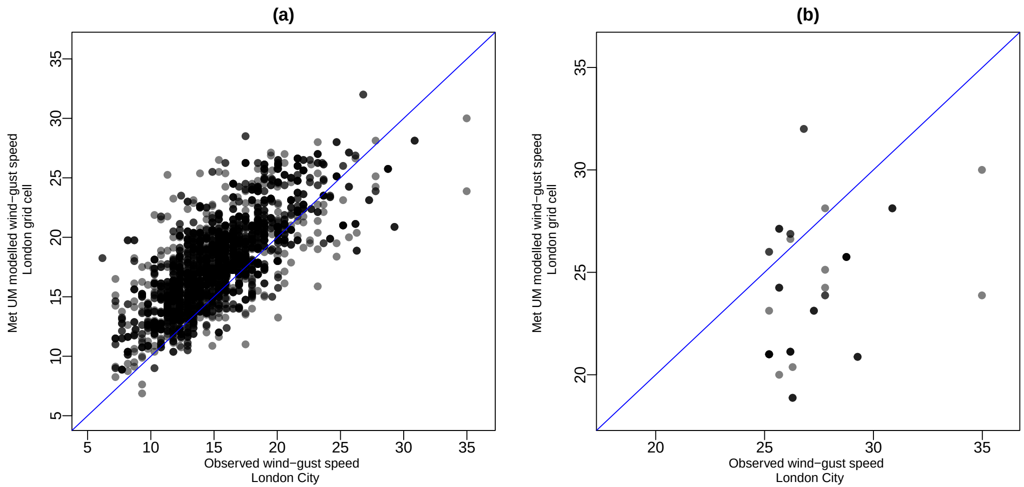
(a)
(b)
Met UM modelled wind−gust speed
London grid cell
Observed wind−gust speed
London City
5
10
15
20
25
30
35
Met UM modelled wind−gust speed
London grid cell
Observed wind−gust speed
London City
20
25
30
35

Fig. 1.

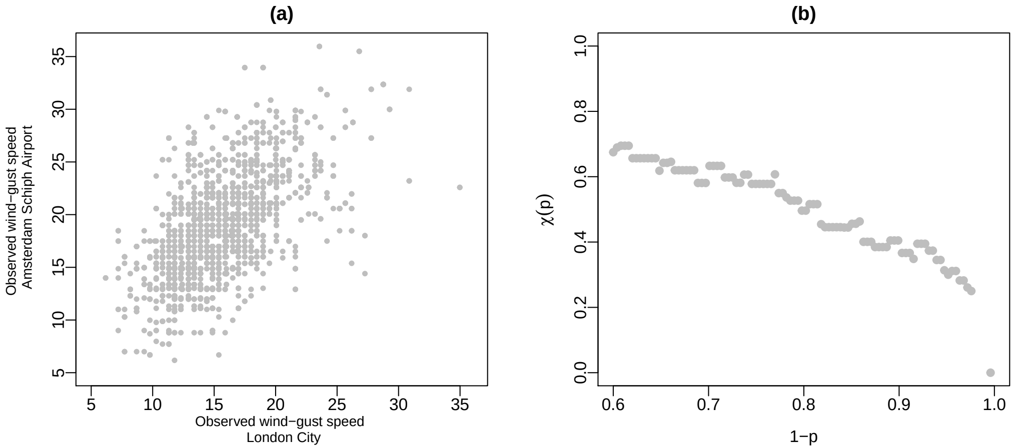
(a)
Observed wind−gust speed
Amsterdam Schiphol Airport
Observed wind−gust speed
London City
5
10
15
20
25
30
35
(b)
χ(p)
1−p
0.0
0.2
0.4
0.6
0.8
1.0
0.6
0.7
0.8
0.9
1.0

Fig. 2.

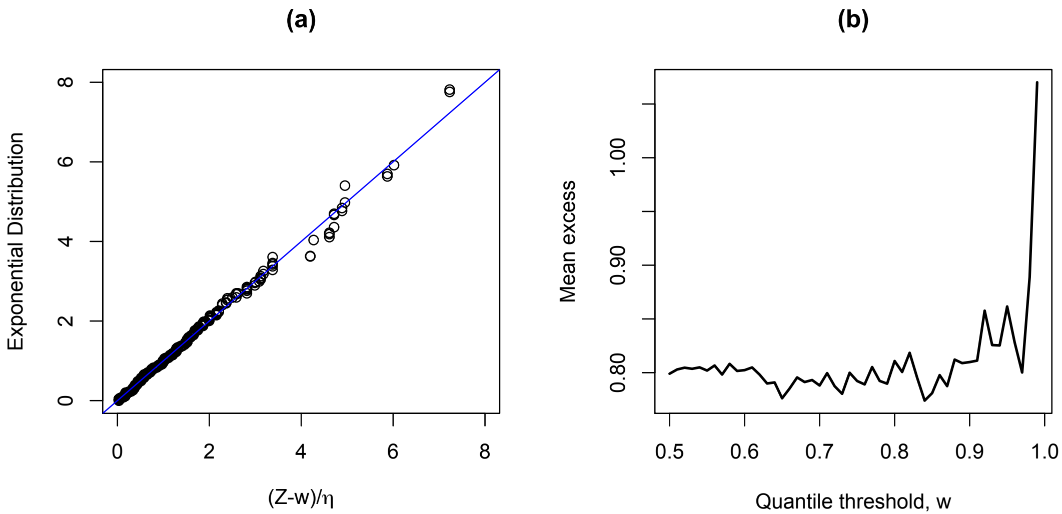
(a)
Exponential Distribution
(Z-w)/η
0
2
4
6
8
(b)
Mean excess
Quantile threshold, w
0.80
0.90
1.00
0.5
0.6
0.7
0.8
0.9
1.0

Fig. 3.