INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2008)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1794

The return period of wind storms over Europe

Paul M. Della-Marta,[a]* Hubert Mathis,[a] Christoph Frei,[a] Mark A. Liniger,[a] Jan Kleinn,[b] and Christof Appenzeller[a]

[a] *Federal Office of Meteorology and Climatology, MeteoSwiss, Zurich, Switzerland*
[b] *Partner Reinsurance Company, Zurich, Switzerland*

ABSTRACT: Accurate assessment of the magnitude and frequency of extreme wind speed is of fundamental importance for many safety, engineering and reinsurance applications. We utilize the spatial and temporal consistency of the European Centre for Medium Range Forecasts ERA-40 reanalysis data to determine the frequency of extreme winds associated with wind storms over the eastern North Atlantic and Europe. Two parameters are investigated: 10-m wind gust and 10-m wind speed. The analysis follows two different view-points: In a spatially distributed view, wind-storm statistics are determined individually at each grid-point. In an integral, more generalized view, the wind-storm statistics are determined from extreme wind indices (EWI) that summarize storm magnitude and spatial extent. We apply classical peak over threshold (POT) extreme value analysis techniques (EVA) to the EWI and grid-point wind data. As a reference, a catalogue of the 200 most prominent European storms has been compiled based on available literature. The EWI-based return periods (RP) estimates of catalogue wind storms range from approximately 0.1 to 300 years, whereas grid-point-based RP estimates range from 0.1 to 500+ years. EWIs sensitive to the absolute magnitude of wind speed rank the RP of wind storms in the 1989/1990 and 1999/2000 extended winter season similarly to the RP derived from the distributed approach. The RP estimates derived from EWIs are generally higher when calculated using only land grid-points compared to RPs derived using whole domain. Both the uncertainties in EWIs and grid-point-based RPs show a greater dependence on the wind parameter used than on the uncertainty associated with the EVA for RPs less than 10 years, whereas for RPs greater than 10 years the effect of the different datasets is lower. The EWIs share up to approximately 50% of the variability of the local grid-point RPs. Copyright © 2008 Royal Meteorological Society

KEY WORDS return period extreme wind; wind storm; climatology; extreme value analysis; wind gust speed; wind speed; Europe

Received 28 December 2007; Revised 26 August 2008; Accepted 28 September 2008

1. Introduction

The exceptional severity of wind storms associated with cyclonic disturbances has long been recognized as a prominent feature of the North Atlantic and European climate (Lamb, 1991; Schiesser *et al.*, 1997). Strong winds and their associated influences on the sea state have been responsible for many shipping-related or coastal disasters (Lamb, 1991). Extreme winds over Europe also represent a major loss potential for reinsurance companies (MunichRe, 2000; SwissRe, 2000). This study aims at developing a continental-scale summary measure of the storminess of known severe winter wind storms in Europe during the past four decades. This measure shall be expressed in terms of the return period (RP) with which a storm of similar or greater intensity is expected in the area. The RP of storms can be a valuable measure in comparing actual and past events and in assessing their impact in terms of meteorological causes and societal vulnerability. This study was in part motivated by the needs of the reinsurance industry. In this sector, hindcast simulations with high-resolution limited area weather models are increasingly used as a source of information on the surface wind field (Weisse *et al.*, 2005; Leckebusch *et al.*, 2006; Walser *et al.*, 2006). In combination with damage models, hindcasts of past storms can be used for quantitative risk assessment. In such a procedure, however, estimates of RPs are needed for placing the limited number of hindcast cases into a climatological context, i.e. deriving the probability distribution of loss. Such a procedure is currently envisaged by the Partner Reinsurance Company (PartnerRe), building upon high resolution dynamically downscaled wind fields for approximately 100 storms of the past decades (Schubiger *et al.*, 2004; Turina *et al.*, 2004).

What is the RP of the storm Vivian? In its general sense, this is not a well-posed question. The meteorological storminess of any one storm is characterized by several parameters (e.g. location, path, maximum wind, spatial extent, duration), and the RP of a storm will inevitably depend on how different characteristics are combined into a scalar measure. Moreover, the RPs derived from continental-scale characteristics will be of limited representativity at the local scale. However, the

* Correspondence to: Dr. Paul M. Della-Marta, Partner Reinsurance Company, Bellerivestrasse 36, 8034, Zurich, Switzerland.
E-mail: paul.della-marta@partnerre.com

demand for a compound and continental-scale RP derives primarily from the ease of use in practical applications. Note that in many situations a fully quantitative climatological assessment simply cannot be afforded. For example, the reinsurance needs would require a continuous high-resolution hindcast (e.g. a high-resolution reanalysis) over several decades. In such situations, RPs are valuable summary measures, which help the selection of suitable study cases, which place those cases into a climatological perspective, and hence, support an overall risk assessment from selected scenarios. It must be kept in mind the inevitable uncertainty associated with the details of measuring storminess and the limited local representativity needs to be carefully considered in concrete applications.

In this study, we propose a generic procedure for quantifying the RP of storms over the European continent. The procedure relies on the definition of a scalar index (an Extreme Wind Index, EWI) that characterizes *storminess* in a two-dimensional wind field, and the subsequent analysis of the resulting index time series with methods of extreme value statistics. We apply this procedure to derive RPs for a set of prominent high-impact storm events of the recent decades, building on a coarse resolution reanalysis (the reanalysis ERA-40 of the European Centre for Medium Range Weather Forecasting, Uppala *et al*., 2005). In our study, we address the following specific questions:

1. How reliable are wind parameters in the ERA-40 reanalysis (spatial representativity, temporal homogeneity) for estimating RPs of high-impact storm events over Europe?
2. How sensitive are estimates of RPs to the details in the definition of the EWI?
3. What is the degree of uncertainty in estimating continental scale and grid-point-scale RPs for past storms? Which factors contribute most to the uncertainty?
4. How representative are continental-scale estimates of RPs as a measure of the local recurrence of a storm?

We chose to use the ERA-40 reanalysis dataset since it has the temporal and spatial homogeneity needed for a continental-scale overview of the European storm climate. Other possible data sources, such as *in situ* wind observations or derived wind from atmospheric pressure generally have limited spatial resolution and/or spatial representativeness and temporal inhomogeneities (e.g. Smits *et al*., 2005). Reanalysis datasets are generated by data assimilation systems used by state-of-the-art global weather forecasting models, which extend over several decades and provide physical consistency with all *in situ* observations available in the assimilation process.

Previous studies documenting the extreme wind climate of the North Atlantic and Europe use a number of different data and methodologies depending on the aim of the study. Those aimed at characterizing the absolute mean and extreme wind climate at a local or regional scales have analysed either *in situ* wind data, air pressure or wind speed derived from air pressure observations (e.g. Dukes and Palutikof, 1995; Kristensen *et al*., 1999; Sacré, 2002; Barring and von Storch, 2004; Alexander *et al*., 2005; Smits *et al*., 2005; Walter *et al*., 2006). Most continental-scale storminess studies have focused on either using air pressure observations (Lamb, 1991; Schinke, 1993; Kaas *et al*., 1996; Alexandersson *et al*., 1998), derived wind from air pressure observations (Schmith *et al*., 1998; Miller, 2003), sea level datasets (e.g. Bijl *et al*., 1999), derived wind sensors aboard satellites (e.g. Monahan, 2006) or use of reanalysis data (Yan *et al*., 2002; Pryor and Barthelmie, 2003; Pryor *et al*., 2006a; Yan *et al*., 2006; Seierstad *et al*., 2007). None of these studies have attempted to assign RPs to known historical storm events on the European scale.

Climate change has promoted a wide study of the potential impacts of the enhanced greenhouse effect on the frequency, duration and intensity of wind storms in a future climate compared to today. Held (1993) provides a good introduction to the expected response of large-scale climate to global warming. Recent studies (e.g. Knippertz *et al*., 2000; Rockel and Woth, 2007; Schwierz *et al*., 2008; Ulbrich *et al*., 2008) expect an increase in both the intensity and the frequency of high winds causing storms over Europe in 2071–2100 compared to today. Other studies show a more muted response of storminess in a future climate (Beersma *et al*., 1997; Bengtsson *et al*., 2006; Pryor *et al*., 2006b). In summary, the confidence in future wind-storm changes is low, although it seems likely (>66% chance) that there will be an increase in extreme winds over the North Atlantic and central Europe (Christensen *et al*., 2007). Thus far, observational evidence of an increased intensity of cyclones and their associated surface winds over the North Atlantic and Europe are not conclusive when considering long-term trends (e.g. the past 100 years). Some studies show an increase in storminess and extreme windiness during the period from about 1960 to 2000 in northwestern Europe (e.g. Alexandersson *et al*., 2000; Pryor and Barthelmie, 2003; Alexander *et al*., 2005), while during the same period other studies do not show significant trends (e.g. Raible, 2007; Raible *et al*., 2008) or artificial trends due to inhomogeneities (Smits *et al*., 2005). The increase in storminess in recent decades seems to be part of long-term multi-decadal trends (Kaas *et al*., 1996; Carretero *et al*., 1998; Schmith *et al*., 1998; Bijl *et al*., 1999; Jones *et al*., 1999; Alexandersson *et al*., 2000). We therefore approach our task of creating an extreme wind climatology without special attention to long-term non-stationarities.

We start with an overview of the data used in this study, where we focus on quality issues of the data used. The following section defines various EWIs. We then discuss the extreme value analysis (EVA), detailing the choices made. The main results, including the extreme wind climatologies and the RPs of prominent high-impact events are presented followed by some discussion and conclusions.

DOI: 10.1002/joc

2. Datasets

2.1. High-impact storm catalogue

In this study, we adopt a generic procedure for the estimation of storm RPs to a catalogue of high-impact winter wind storms, which have caused damage on the European continent or are otherwise well known in the reinsurance sector. We compiled our list starting with the comprehensive Lamb catalogue (Lamb, 1991) and supplemented this list with a number of other sources (Barkhausen, 1997; MunichRe, 2000; SwissRe, 2000; DWD, 2001; Heneka *et al*., 2006; KNMI, 2007). The list encompasses a total of 200 wind storms which occurred during the extended winter season from October to April over the period 1957–2002. The list consists of either a start and end date/time or simply a date/time for each stormy period. Both the duration and date/time of each storm is potentially biased to the time during which the greatest impacts were experienced over the area of interest of the authors of each source document. Therefore, the storm dates do not necessarily coincide with the maximum intensity of winds associated with the storm which has repercussions for the attribution of peaks in the index time series to the items in the storm catalogue (see Section 5). Also, the adopted catalogue is not necessarily a complete list of all possibly relevant storms. We expect that storms over sea will not be adequately represented. This is not critical for the results of our analysis, since the storm catalogue is merely used as an example application. In fact the omission of storms from the catalogue would not alter the RP results.

2.2. ERA-40 Reanalysis

The climatological basis for our analysis is the ERA-40 reanalysis of the European Centre for Medium Range Weather Forecasting (ECMWF, Uppala *et al*., 2005). It is the product of a comprehensive assimilation of surface, upper air and satellite observations into a global weather forecasting model. The ERA-40 encompasses physically consistent three-dimensional fields of atmospheric and surface parameters at 6-hourly intervals over the period September 1957–August 2002 (i.e. over 45 years). ERA-40 has a spatial resolution of about 1.125°.

For the purpose of this study, we use ERA-40 data over the North Atlantic and European sectors from 35 °W to 35 °E, and 35 to 73 °N for all the 45 extended winter seasons (October till April). This is known as the main stormy period over Europe and is reflected in the source documents used to compile our catalogue (see also Lamb, 1991). For technical reasons associated with the choice of analysis domain (see Section 3) we used a 0.5° grid version of ERA-40, encompassing a total of 10 857 (141 × 77) grid-points.

Two parameters from the ERA-40 dataset have been considered in this study: a 10-m wind gust (FG10, ECMWF parameter name) and a 10-m wind speed (WS10). The values of wind gust at each analysis time represent maximum gust values from the past 6 h. Hence, wind gust seems to be ideal to describe wind peaks at the surface, the ultimate cause of storm damages. However wind gust is a model diagnostic, and therefore, depends on the parameterizations in the numerical model underlying the assimilation system (see ECMWF, 2003 for the parametrization method, available at http://www.ecmwf.int/research/ifsdocs/CY23r4/index.html). In Section 2.3 we illustrate that wind gust shows unrealistic behaviour near coasts and steep topography which requires masking of certain areas. In contrast, the 6-hourly fields of WS10 represent instantaneous wind fields and do not necessarily sample maximum storm intensity.

2.3. Data quality and homogeneity

The quality of reanalyses depends strongly on the parameter, for example, temperature is well represented, even in mountainous areas (Kunz *et al*., 2007), other parameters like integrated water vapour (Morland *et al*., 2006) or precipitation might be less realistic in absolute terms. In particular, there can be serious biases in absolute values of wind near the surface, (Smits *et al*., 2005), and in some cases there are inhomogeneities (Bengtsson *et al*., 2004; Sterl, 2004; Smits *et al*., 2005). The known wind biases make it difficult to justify a climatology of absolute wind measures directly from reanalysis (Caires and Sterl, 2005). However, the limitations are primarily systematic and it can be expected that the relative ranking of storms in a reanalysis is reproduced much more reliably than the absolute wind. This is why the focus of this study is entirely in the frequency domain (i.e. RPs) where issues of data quality are far less serious.

Initial screening of wind fields in ERA-40 suggests that in many cases there are unrealistic values of wind gust. Compared to other areas, extremely high wind gust values were found in areas of steep orographic gradients. Figure 1 illustrates the case for storm Daria (26 Jan 1990). The original wind gust field from ERA-40 (Figure 1(a)) shows strong discontinuities and unrealistically high values over the Alps, coastal Scandinavia and parts of the Mediterranean (e.g. Greece). Such artefacts are not present in the wind-speed field (Figure 1(c)).

Areas with unrealistic wind gusts are almost identical for other storms and they are collocated with areas where the surface roughness, z_0 values are highest in the ERA-40 reanalysis wind gust parameterization (ECMWF, 2003). z_0 shows (not shown, see Della-Marta *et al*., 2007) a high contrast in values between ocean areas/smooth orography and areas of complex orography such as the Alps and the western coast of Scandinavia. Over land, surface roughness in ERA-40 is a fixed parameter that combines roughness lengths from land use and from sub-grid-scale orography (ECMWF, 2003). Apparently, the inclusion of sub-grid-scale orography has a high impact on the realism of wind gusts over complex orography. Note, the ECMWF has updated the wind gust parameterization of its operational forecast model in summer 2006. The parameterization now separates the two contributions to surface roughness with a vast improvement in the wind gust values.

DOI: 10.1002/joc

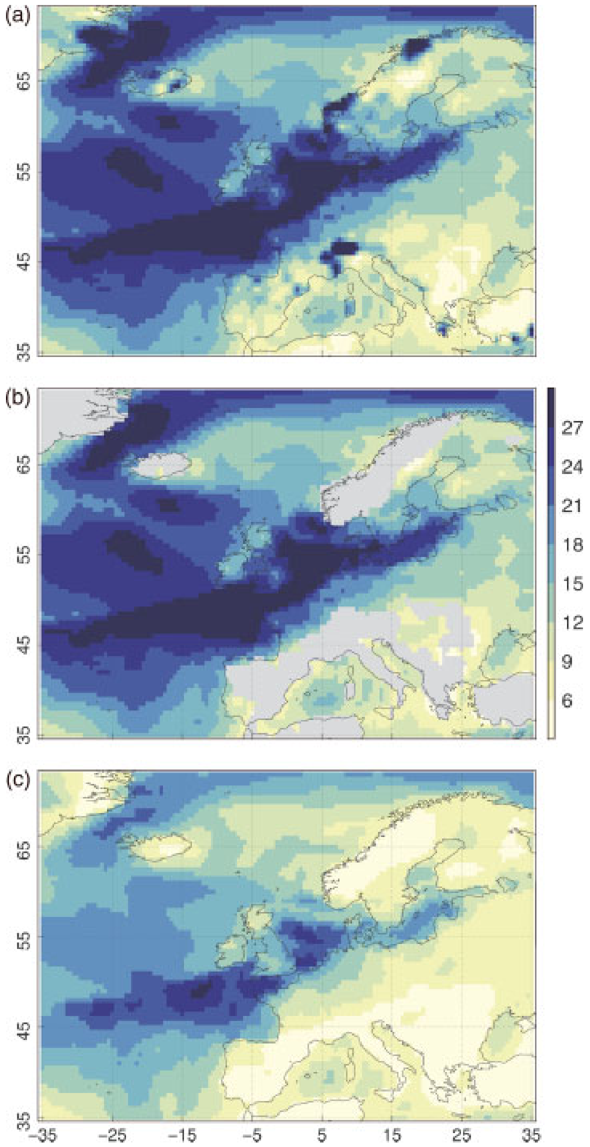

Figure 1. A comparison of the 72-hour maximum wind fields using ERA-40 data for the storm, Daria (26 Jan 1990) (ms^{-1}). (a) The ERA-40 FG10 field, unmasked, (b) FG10 masked, and (c) WS10. Grey areas denote masked values. This figure is available in colour online at www.interscience.wiley.com/ijoc

To avoid that our scalar EWI is dominated by wind gust values at a few extreme grid-points, we decided to mask out areas with unrealistic wind gusts. For this purpose we chose to mask grid-points where z_0 is greater than 3 meters and at grid-points where the elevation of the ERA-40 orography is greater than 700 m. (Figure 1(b)) and also Della-Marta *et al.* (2007)). These criteria are subjective, but they are motivated from a visual inspection of the wind gust fields for many extreme wind situations. Note that no masking has been applied to fields of WS10, hence the latter covers the wind conditions over the entire domain.

We also performed a basic check on the long-term homogeneity of the ERA-40 wind parameters. The availability of new observation systems in the past decades could in principal have affected the temporal homogeneity of the ERA-40. For this purpose, we have inspected seasonal time series of the mean wind gust and wind speed over the eastern North Atlantic. These time series (not shown) reveal a high correlation with the North Atlantic Oscillation Index (NAOI, typical r values of 0.8), an index which is independent of ERA-40 (Hurrell *et al.*, 2002) and measures the strength of westerlies over the North Atlantic Ocean and Europe (see e.g Appenzeller *et al.*, 1998; Wanner *et al.*, 2001). The correlations seem to be of a similar strength over the whole period, and the time series do not show unexpected trends or shifts. Although these elementary tests are no final confirmation of homogeneity, they do not reveal artificial temporal changes that would seriously affect our EVA of ERA-40-derived storm indices.

3. Extreme wind indices

Scalar indices (time series) are used to summarize a wind storm's magnitude and spatial extent. Each index uses the grid-point wind speeds from ERA-40 as the basis for calculation. A number of such different compound EWIs were analysed by PartnerRe which we used as a basis for the different indices presented below. Reinsurance companies often need a singular estimate of the frequency of a wind-storm event to estimate the expected frequency of an aggregated loss over a portfolio. In other words, they need a frequency estimate of the wind-storm event, and not only the frequency (RP) of wind speed (or wind gust) at a specific place. In this report, we only present a selected number of such indices which we determined to be independent enough and useful in the assessment of the RPs. Since the magnitude of EWIs is likely to be dependent on the area over which the indices are calculated, we decided to investigate using either *all* grid-points in the domain or *land only* grid-points in the domain. In all cases, except where specified, the FG10 data were masked as explained above. Where possible, we tried to take into account the unequal areas of each grid box by weighting of sums and multipliers by the cosine of the latitude of each grid-point. For each index we provide a brief rationale and their expected sensitivity. Detailed descriptions of each index in mathematical notation can be found in Appendix A.

$\overline{X}$: Mean wind. This index is a time series of the weighted (for latitude) mean wind speed calculated over a given area (either all grid-points or only those over land) in the units of ms^{-1}. This index is simple and is intuitive as a starting point for comparison with more complex indices described below. The index is likely to be sensitive to both the severity of the wind storm (at each analysis time) and its spatial extent.

***Q*95: The spatial 95% quantile wind.** This index summarizes the wind speed in the windiest 5% of (latitude weighted) area per unit time. At each point

DOI: 10.1002/joc

in time the wind speeds at all grid-points are ranked and the empirical 95% quantile chosen as the value for this index. In other words this index measures the lower bound of wind speed in the top 5% of the area considered. Therefore, this index is concentrated on measuring only the windiest area in the domain at any given time and is more likely to be an estimate of storm severity than $\overline{X}$ (units of ms^{-1}).

***Sw3q90*: Cube root of the sum of wind cubed above the domain climatological 90% quantile.** This index is motivated by considering the area and time integral of kinetic energy associated with extreme winds expressed as the Power Dissipation Index (Emanuel, 2005) non-dimensional units, (*NDU*). In our case, we ignore the time integral component but incorporate the area component by summing over the area of interest. Note that Emanuel has used the maximum sustained wind whereas our estimates are based on either maximum wind gust or instantaneous wind speeds. This index first calculates the 90% quantile based on all grid-points in the domain for the entire ERA-40 period. This ensures that only areas with absolute relatively high wind-speed areas are considered. Then the excess winds at all grid-points which exceed this threshold are cubed and summed. The cube root is then taken as a final step to help make the index less skewed. This index gives cubic weight to grid-point wind-speed exceedences, and therefore, should be sensitive to areas of high absolute magnitude wind speeds. This index is similar to that used in Klawa and Ulbrich (2003) but takes into account the full domain climatology. They use the cube of the excess above a local based percentile. The indices described below also show similarities to their index.

***Sfq95*: Sum of the fraction of wind divided by the grid-point climatological 95% quantile.** It is envisaged that this index summarizes the extremity of wind speed over a given area relative to the local extreme wind climate at each grid-point (*NDU*). The index first calculates the 95% quantile of wind speed at each grid-point in the chosen domain using all wind speed observations during the extended winter season over the ERA-40 reanalysis period. Then, at each time step and for each grid-point with wind above the local 95th percentile, the fraction of wind speed above the local 95th percentile is summed. This index is only sensitive to extreme wind speeds relative to local climate, and therefore, not sensitive to the magnitude of wind speed in absolute terms.

***Sfq95q99*: Sum of the fraction of extreme wind divided by the length of the distribution tail.** This index should also be sensitive to the relative extremity of local wind speed, however, unlike *Sfq95* this index has a normalizing factor which is proportional to the length of the tail of the local extreme wind distribution (*NDU*). This index should give equal weight to the winds in a storm region whether the storm be located over the sea or land, especially where we see a contrast in both the scale and shape of the local extreme wind distribution (see Figure 8(c) and 8(d)).

4. Return periods derived from extreme value analysis

The second step in our procedure consists in estimating the frequency distribution of scalar EWIs. The RP for a given storm is then specified as the exceedence frequency of the value of that storm's EWI. For this step of the procedure we choose classical techniques of EVA. These techniques are based on the asymptotic statistical behaviour of extreme values and they permit unbiased estimates of the tail of the EWI's distribution function. Fisher and Tippett (1928) and Gumbel (1958) have put forward the theoretical foundations and application principles, respectively, of EVA (an introduction to EVA is given in, e.g. Coles, 2001). The methods have also been widely used for the analysis of extreme wind speeds, and Palutikof *et al*., 1999 review procedures and applications in this particular context.

In our application of EVA we use the Peaks Over Threshold (POT) approach. In this approach, we consider independent exceedences of the EWI above a suitably chosen threshold and model their distribution with a Generalized Pareto Distribution (GPD). Our application essentially follows the classical procedure as described in Coles (2001, Chap. 4). Specifically we adopt the maximum likelihood principle for the estimation (MLE) of the GPD parameters, and (in slight deviation from Coles, 2001), assume a Poisson distribution for the events exceeding the threshold (see also Palutikof *et al*., 1999). More specific approaches were adopted in the declustering of threshold exceedences, in the determination of a suitable threshold, and in the calculation of sampling uncertainty. In the following subsections we describe these steps in more detail.

In order to introduce notation that will be used throughout the paper we rewrite the relevant formulae from Coles (2001) and Palutikof *et al*. (1999). The GPD can be written in terms of a generic variable x as:

$$G(x) = 1 - \left[1 + \frac{\xi}{\sigma}(x-u)\right]^{-\frac{1}{\xi}} \quad (1)$$

Conditional on $x > u$ and $\xi \neq 0$ where u is the selected threshold. The GPD is characterized by two parameters, ξ the shape parameter and σ the scale parameter. If $\xi > 0$ then the maximum of the GPD is unbounded, whereas if $\xi < 0$ then the tail has a finite extent, if $\xi = 0$ then the GPD reduces to the exponential distribution and is also unbounded in the limit $\xi \to 0$. Equation 1 can be rewritten in terms of probabilities which leads to the calculation of the N-year return level (RL), x_N which is exceeded once every N years (the RP) and is given by:

$$x_N = u + \frac{\sigma}{\xi}\left[1 - (\lambda N)^{-\xi}\right] \quad (2)$$

DOI: 10.1002/joc

Where λ is the mean number of threshold exceedences per unit time.

To find the RP of each catalogue storm we first rearrange Equation 2 in terms of N and substitute all the values of the EWI or the grid-point wind speed within a 72-h period centered on the date/time in the storm catalogue. We then take the maximum value of N as the RP of a particular storm.

4.1. De-clustering

The asymptotic theory underlying the POT approach requires that threshold exceedences are statistically independent and are from a stationary random process. Raw time series of an EWI are unlikely to satisfy this condition. Spanning over as much as seven months of the year we anticipate that the EWI time series are influenced by the annual cycle. Also, synoptic disturbances over Europe have a lifetime considerably longer than the 6-h time resolution of the index, which will reflect in serial correlation in the index time series. Indeed, serial correlation is obvious in the partial autocorrelation function of the EWI time series (not shown).

De-clustering is a pre-processing of the original time series which aims at extracting threshold exceedences that can be considered statistically independent. Most de-clustering methods are based on the estimation of a statistic called the extremal index θ. The extremal index can be thought of as the reciprocal of the limiting mean cluster size (Coles, 2001). In the presence of no autocorrelation (clustering) in the series then $\theta = 1$. Else if $\theta < 1$ then there is clustering in the data. In our de-clustering of the EWI time series we chose an estimator of θ proposed by Ferro and Segers (2003). Their extremal index is based on the inter-exceedence times and it represents the proportion of inter-exceedence times that may be regarded as times between independent clusters. They show that their estimate of θ has better de-clustering characteristics than other commonly used methods, e.g. the runs de-clustering. Moreover, their method has the advantage that it is automatic in the sense that θ changes with changes in the POT threshold. Hence threshold selection (see below) and de-clustering are actually linked together.

Ferro and Segers (2003) stipulate that their estimate of θ is only representative if it is calculated on a strictly stationary series. Analyses in our application to EWI time series shows that the performance of the de-clustering method is degraded by the contribution from the annual cycle in these time series. In order to compensate for this we take account of the seasonal cycle in terms of a threshold that varies across the season. The threshold for a particular calendar day is calculated from the 95th percentile of all four daily analyses time steps (i.e. from 4×45 values) for that particular day, and subsequent application of a smoothing spline operator to the resulting quantile time series. Note, the filtering of the annual cycle is only adopted to de-cluster compound EWI times series but not for the analysis of wind time series at individual grid-points, where the annual cycle was clearly less evident.

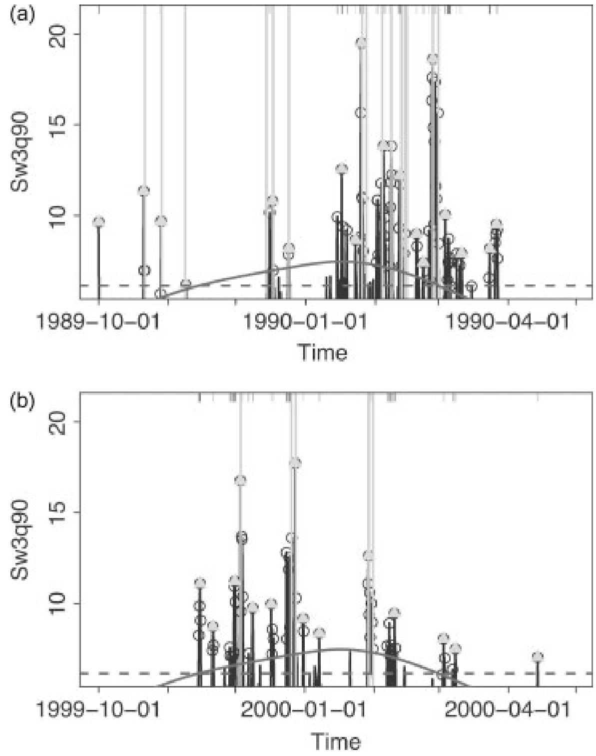

Figure 2. Examples of the EWI, $Sw3q90$ and the daily de-clustered POT series using the approach of Ferro and Segers (2003) for the extended winter season (ONDJFMA) of 1989/1990 (a), and 1999/2000 (b). The thin black line is the $Sw3q90$ index calculated over land using WS10. The circles indicate values of the index which exceed the daily threshold (solid black line). Grey filled triangles show the maximum value of the index within each cluster. Membership of POTs (circles) to a particular cluster are denoted by alternating light and dark grey bands on the top margin of the plot. The solid black line and dashed black line show the daily 95th percentile and the seasonal 95th percentile respectively. The vertical grey lines indicate the date of the storms in the storm catalogue.

Figure 2 displays results of the de-clustering method for index $Sw3q90$ (based on WS10) and for two extended winter seasons (1989/1990 and 1999/2000). The EWI time series exhibits prominent peaks which are mostly coincident with storms in the storm catalogue (vertical grey lines). The peaks seem to be superimposed onto a gradual annual cycle whose climatological evolution is indicated with the time varying threshold. The de-clustering technique seems to resolve this obvious dependence in that the clusters (see shading at the top of the panels) mostly contain periods with semi-contiguous threshold exceedences. For example, Daria and Vivian (Figure 2(a), 7th and 14th vertical lines from the left) are clearly in separate clusters, whereas the two storms Vivian and Wiebke (14th and 15th vertical lines) are closely related, and therefore, merged into the same cluster by the methodology. A similar example is evident in season 1999/2000 (Figure 2(b)) where the storm Anatol (1st vertical line) is clearly a separate cluster, whereas the storms Lothar and Martin (2nd and 3rd vertical lines) are merged. This has implications for the EVA such that

DOI: 10.1002/joc

only the maximum wind within a cluster (grey triangles, Figure 2) is used in the GPD model. Hence in this case the storms Lothar and Martin are treated as one storm. Note, that similar difficulties are encountered for storms in the catalogue when the corresponding EWI value is so low as to not exceed the threshold. A more detailed comparison between the Ferro and Segers (2003) de-clustering and the classical runs de-clustering (not shown) reveals that there is little sensitivity on the RP results to these two de-clustering methods (Della-Marta *et al.*, 2007).

Figure 3 depicts similar de-clustering results calculated from wind time series at two individual grid-points (a) north of the British Isles, (b) northeast of Europe. Extreme value analyses for individual grid-points is used for comparative purposes and for representativity tests (see later Section 5). The time series of WS10 show less serial correlation than the index time series. Catalogue storms tend to coincide with wind peaks (as is the case with the EWI, Figure 2), but there is obviously much weaker correspondence due to the more local character of time series at grid-points.

4.2. Threshold choice

The threshold of the POT analysis should be large enough to ensure near-asymptotic behaviour of the exceedences. There are several diagnostic means to estimate a suitable threshold. These diagnostics build upon the known threshold dependence of GPD parameters in the asymptotic tail (Coles, 2001). In this study, we have

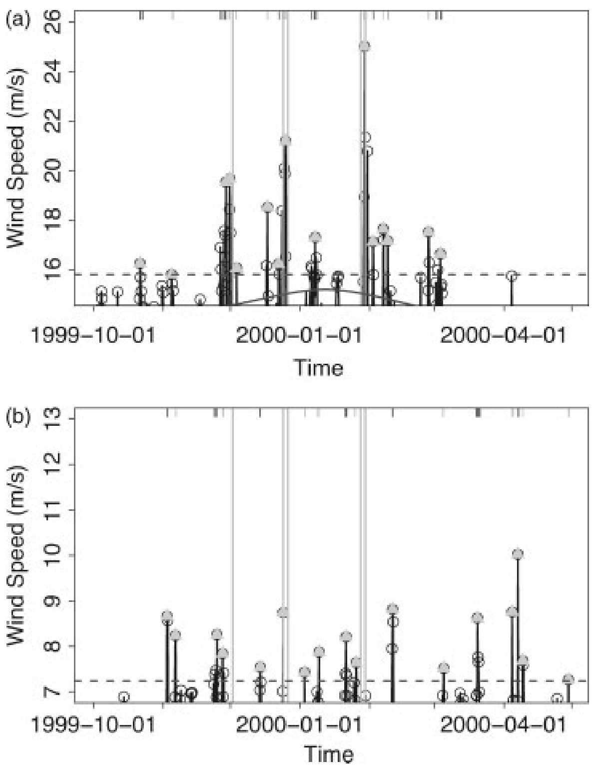

Figure 3. As for Figure 2 but for the grid-points 2.5 °W, 57.5 °N (a), and 30 °E, 67.5 °N (b). Note that the EVA was performed on raw wind values (units of ms^{-1}) at each grid-point.

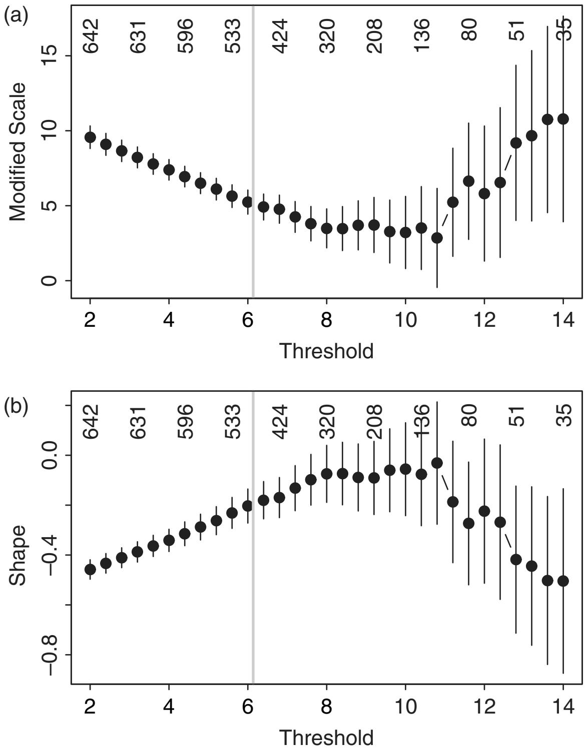

Figure 4. Modified Scale, $\sigma*$ (see Coles, 2001) (a) and the shape parameter, ξ, (b) diagnostic plots for selecting the fixed threshold *NDUs* above which the de-clustered POT are modelled using the GPD. This example is based on the de-clustered POT $Sw3q90$, WS10 using the land-only region. The vertical black lines denote the 95% confidence intervals calculated using the parametric resampling technique detailed in Section 4.3. The numbers aligned vertically in the top of the plot are the number of cluster maxima identified by the de-clustering technique. The lowest threshold value shown in each plot is the seasonal 70th percentile, and the vertical grey line is the 95% quantile threshold.

inspected the asymptotic independence of the GPD shape parameter and modified scale parameter upon threshold (Coles, 2001, chapter 4). These diagnostics are depicted in Figure 4 for index $Sw3q90$ (based on WS10). For threshold values smaller than about 6 *NDUs* there is clear dependence of the parameters on the threshold, but at larger values the variations are mostly contained within the uncertainty ranges of individual estimates. For the index under consideration a value of 6 corresponds approximately to the 95th percentile (vertical grey line). Note that we have constrained these results by the choice of daily quantile used to de-cluster the time series (see above), and hence, the de-clustering and threshold selection are not independent of each other. Inspection of similar diagnostics for other indices reveals that the 95th percentile is a generally acceptable choice of threshold. Based on these results, we have chosen the pertinent 95th percentile as the threshold for all indices. Clearly, there is no final proof that a particular choice is sufficient, since the assessment is also limited by sampling uncertainty. Some additional confirmation is, however, available from quantile-quantile plots (see example in Figure 5), which demonstrates a good fit of the theoretical distribution to the data. Additional checks are given in Della-Marta *et al*. (2007).

Threshold diagnostics for time series of wind at individual grid-points also show that a threshold corresponding to the 95th percentile is acceptable, i.e. asymptotic behaviour was assured only beyond that value. Figure 6 depicts the diagnostics for WS10 at one example grid-point. A value of 15.8 ms^{-1} corresponds to the 95th percentile. Consideration of several grid-points suggested this quantile as a reasonable setting for the POT analysis at individual grid-points. For the de-clustering, however, we kept the seasonally varying threshold at the 90th percentile which insured that only grid-points with an exceptionally strong seasonal cycle included a seasonally varying threshold.

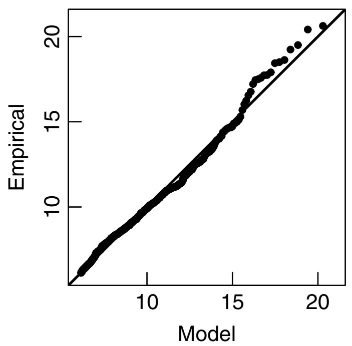

Figure 5. A quantile-quantile (*qq*) plot (NDU) of the fitted GPD to the de-clustered POT $Sw3q90$, WS10 for the land-only region.

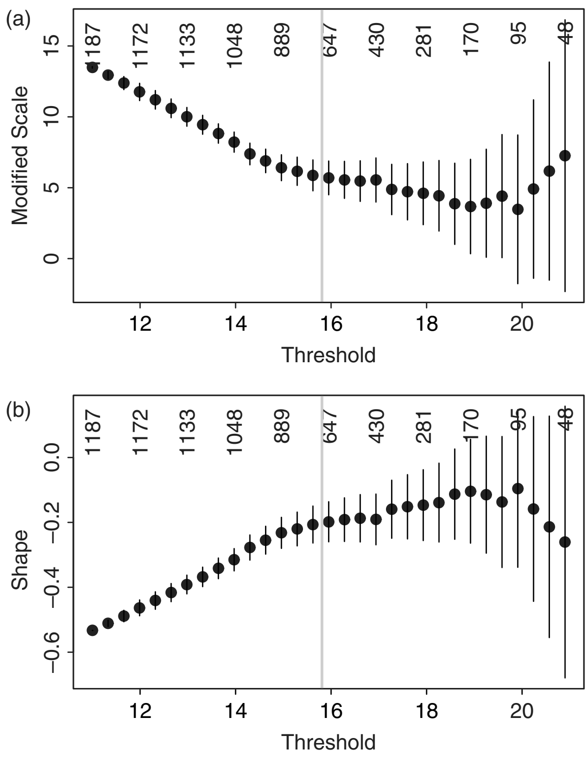

Figure 6. As for Figure 4 but for the declustered POT WS10 grid-point 2.5 °W, 57 °N.

4.3. Sampling uncertainty

Results from EVA can be subject to considerable sampling uncertainty. Classical procedures for estimating the uncertainty of GPD parameters (and related quantities, such as the RPs) encompass the asymptotic maximum likelihood confidence intervals (also termed the *delta method*, see Coles, 2001), and resampling techniques. The latter builds on repeated estimates with random draws from the dataset (non-parametric) or with random draws from the best estimate distribution (parametric). Both these techniques turned out to have limitations in our application (see comparison below).

In the present study, we calculate confidence intervals of RP estimates using the profile likelihood method. Profile likelihood confidence intervals are based on a likelihood ratio test. In practise, they are calculated from a projection of the likelihood surface on the respective parameter axis, which can be obtained from a sequence of numerical optimizations (Coles, 2001). The likelihood profile method is related to the delta method but exploits the true shape of the likelihood surface instead of approximating it around its maximum. As a consequence, calculating confidence intervals with the likelihood profile methods makes more efficient use of the data points in the sample.

A comparison of results from all three techniques is depicted in Figure 7 for index $Sw3q90$ (based on WS10). The figure shows the RL as a function of RP with pertinent confidence intervals from all three methods. The curvature of the best estimate (middle thick line) suggests that this index (like most others in this paper) exhibits a short tail behaviour and hence its distribution function has an upper bound. The best estimate of the upper bound is around 90 *NDU*, but there is considerable uncertainty about its value. According to the delta method (dash-dotted line) and the parametric resampling (dashed line) there is non-zero probability that the upper bound is 84

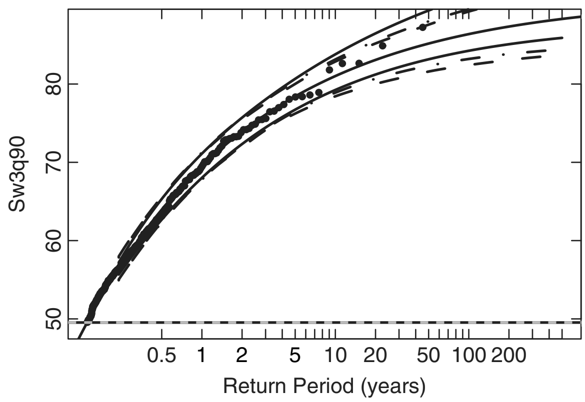

Figure 7. A RL / RP plot showing the GPD fit (centre solid black line) and a comparison of 95% confidence intervals using the methods outlined in Section 4.3. The index shown here is the $Sw3q90$, WS10 using all grid-points. Horizontal axis: RP (years), vertical axis: RL (NDU). Different estimations of confidence intervals: profile log-likelihood (solid black), parametric resampling (dashed black) and the delta method (dot dashed black). Note the log scale on the horizontal axis.

DOI: 10.1002/joc

NDU or lower (the lines for the lower confidence bound level off around this value). This is in contradiction with the actual dataset, which comprizes events with index values larger than 88 *NDU*. Obviously the upper bound of the distribution is necessarily larger than the largest data value. This is not reflected in the confidence intervals of the delta method and the resampling. Both these methods do not take explicitly into account the highest data values in the sample. In contrast, the confidence bounds calculated from the likelihood profile reveals an uncertainty range for the RLs which is shifted to larger values and is consistent with the data at large RPs. Some more general comparisons show that the likelihood profile method differs from the two other methods particularly in cases with a short (bounded) tail. Consideration of the true likelihood surface in the profile method makes more direct use of the data sample in estimating the sampling uncertainty and this avoids inconsistencies of confidence intervals with the actual sample.

Short tail behaviour was observed with most of the EWI considered in this paper. We therefore chose the likelihood profile method for estimating the sampling uncertainty of RPs (and distribution parameters). A slight modification was however applied for RPs near the threshold (typically for periods smaller than 0.2 years). For these cases, the likelihood profile was replaced by the uncertainty in the mean exceedence based on the underlying Poisson process.

5. Results

The Results Section is divided into three subsections. Firstly, the distributed (grid-point) and generalized (EWI) storm climatologies are presented in order to identify their similarities and differences. The second subsection is devoted to comparing the RPs of wind storms derived from the two different approaches answering research questions numbers two and three (Section 1). The final subsection is aimed at answering research question four on the representativity of the generalized storm RPs for estimating the grid-point storm RPs.

5.1. Distributed and generalized storm climatologies

We fitted a GPD distribution to each of the 10 857 grid-points over the domain to form an extreme wind climatology to be representative of the local extreme wind climatology. As described in Section 4.2 we used the same fixed threshold (95% quantile) in the grid-point GPD model as the EWI (95% quantile), based on the extensive diagnostic checks performed for individual grid-points. Generally, the GPD fits to POT series at individual grid-points, (assessed using *qq*-plots) is very good (not shown). We quantified this more rigorously using an Anderson-Darling (A2) goodness-of-fit test (Choulakian and Stephens, 2001) and find that 74% of the fitted GPDs passed this test at the 5% significance level.

Figure 8 provides a summary of the important parameters of the EVA at each grid-point. Figure 8(a) shows the empirically based seasonal 95% quantile threshold (u) of WS10. Generally, there are higher winds over the North Atlantic Ocean and the British Isles than over the north, east and south of the domain. The average number of extreme wind events per season, λ shows a band of lower values running from the southwest of the domain to the northeast of the domain, whereas in the northwest, the southeast as well as the far northeast there are higher values of λ (Figure 8(b)) corresponding to the major storm track and genesis areas of the Northeast Atlantic and the Mediterranean regions respectively. Values of λ range from approximately 6 to 25. The spatial distribution of σ, the scale parameter of the GPD (Equation 1), resembles the distribution of u, i.e. a wider GPD distribution over ocean areas compared to land areas (see Monahan, 2004, for a physical explanation). The shape parameter ξ is rather mixed and shows little spatial coherency. In some cases, individual grid-points have a slightly positive shape parameter indicating that the GPD has no upper limit. The extremal index θ (Section 4.1), a measure of the tendency for storms to cluster in time, is almost identical in appearance to the spatial distribution of λ. θ shows a wide area of the domain east of the main North Atlantic storm track where θ is lower than around 0.3 indicating that extreme wind events tend to form larger clusters that are separated by more time in this region compared to the western North Atlantic, south of the Alps and the eastern Mediterranean, in part confirming the storms clustering analysis of Mailier *et al.* (2006) for cyclone count-based statistics. This process is evident in Figure 3 which shows the de-clustered POT series for two points, one where there is clearly more clustering (Figure 3(a), higher λ and lower θ) and the other where there is less clustering of extreme winds (Figure 3(b), lower λ and higher θ). The RL for various RPs are shown in Figure 9. Note that the purpose of this figure is to compare the relative differences in the RLs for various RPs and not as a measure of the absolute magnitude of surface wind speeds. For each of the RPs a similar spatial structure of the extreme winds can be seen, with higher values in the far west of the domain and over ocean regions than over land. Relatively high values can be seen over the British Isles and the north coast of Spain as well as the western and northern coasts of western Europe with relatively lower values over Scandinavia, eastern and southern Europe.

The generalized wind-storm climatology is presented as a RL/RP plot in Figures 10 and 11. The RL/RP plot summarizes the fitted GPD (i.e. the extreme wind climatology) together with the estimates of the RP and RL of each of the catalogue wind storms which are above the chosen threshold. Using the $Sw3q90$ index calculated from WS10 over the whole domain we estimate a wide range of RPs (Figure 10, vertical grey lines) for the catalogue storms between approximately 0.2 and 18 years. These estimates are based on the GPD fit (black line) and not the cluster maxima (black dots). The curvature of the GPD fit is negative, indicating that the shape parameter is negative (ξ, Equation 2) and that there is a physical upper limit to the extreme process.

DOI: 10.1002/joc

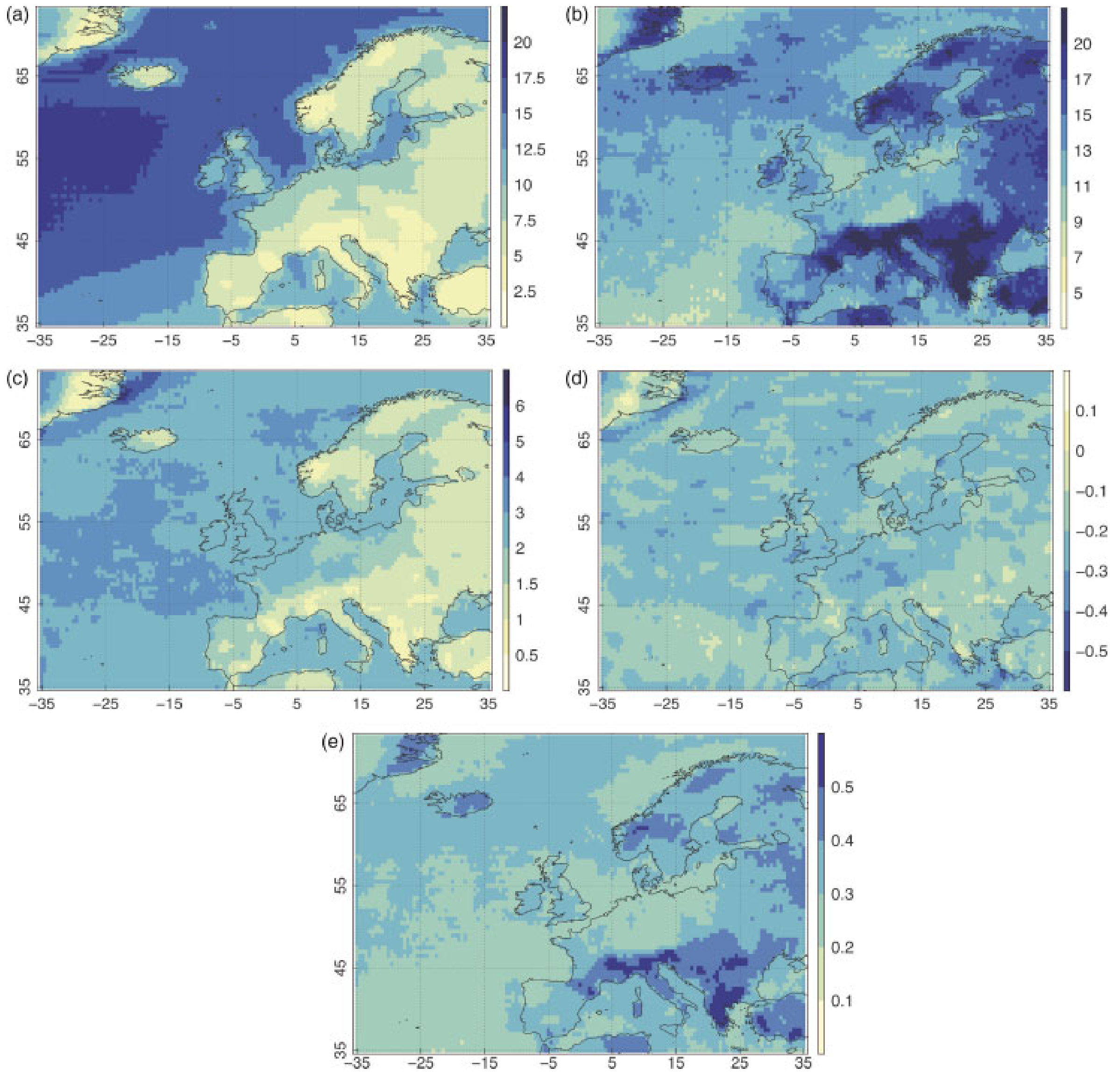

Figure 8. Parameters of the grid-point EVA (Section 4) for WS10. (a) The grid-point empirical 95% quantile threshold, u (ms^{-1}) and the MLEs of the GPD fit (Equations 1 and 2) , (b) the average number of extreme wind events per season after de-clustering, λ, (c) the scale parameter of the GPD, σ, (d) the shape parameter of the GPD, ξ , and the extremal index, θ (Section 4.1) (e). This figure is available in colour online at www.interscience.wiley.com/ijoc

This is in agreement with distributed storm climatology (see Figure 8(d)). Figure 11 shows the generalized storm climatologies of all five indices calculated using land-only grid-points. The climatologies all exhibit a negative shape parameter in the range $-0.23 \leq \xi \leq -0.10$ and each have a similar number of storm occurrences per season, λ in the range of $8.2 \leq \lambda \leq 11.0$. The scale parameter varies greatly due to the different units of each EWI. The quality of the GPD fit to each of the EWI was assessed using qq-plots (e.g. Figure 4) and the Anderson-Darling (A2) test. The quality of the fit assessed using this statistic was noticeably the best for the index $Sw3q90$ and the worst for the index $Q95$. Comparing Figure 11 with Figure 10 it is noticeable that there is a greater range of RP estimates for the catalogue storms. This reflects that the sampling of wind storms in the catalogue is biased towards wind storms which had an impact on the western European region. With this in mind, and returning to Figure 10, we see that there are a number of storms in the EWI (black dots) that do not coincide with the seven most extreme catalogue storms (grey lines). A sign that more intense wind storms have occurred in the northeast North Atlantic than documented in the storm catalogue.

5.2. Distributed and generalized return periods for prominent European wind storms

In this section, we present a comparison of the RPs of catalogue wind storms using the generalized and distributed wind storm climatologies highlighting the major sources of uncertainty. We start with an intercomparison of the EWIs using either WS10 or FG10. Figure 12 presents scatter-plots of the catalogue storm RPs estimated using FG10 versus the RPs using WS10 for each index for the land-only domain. On each plot there are

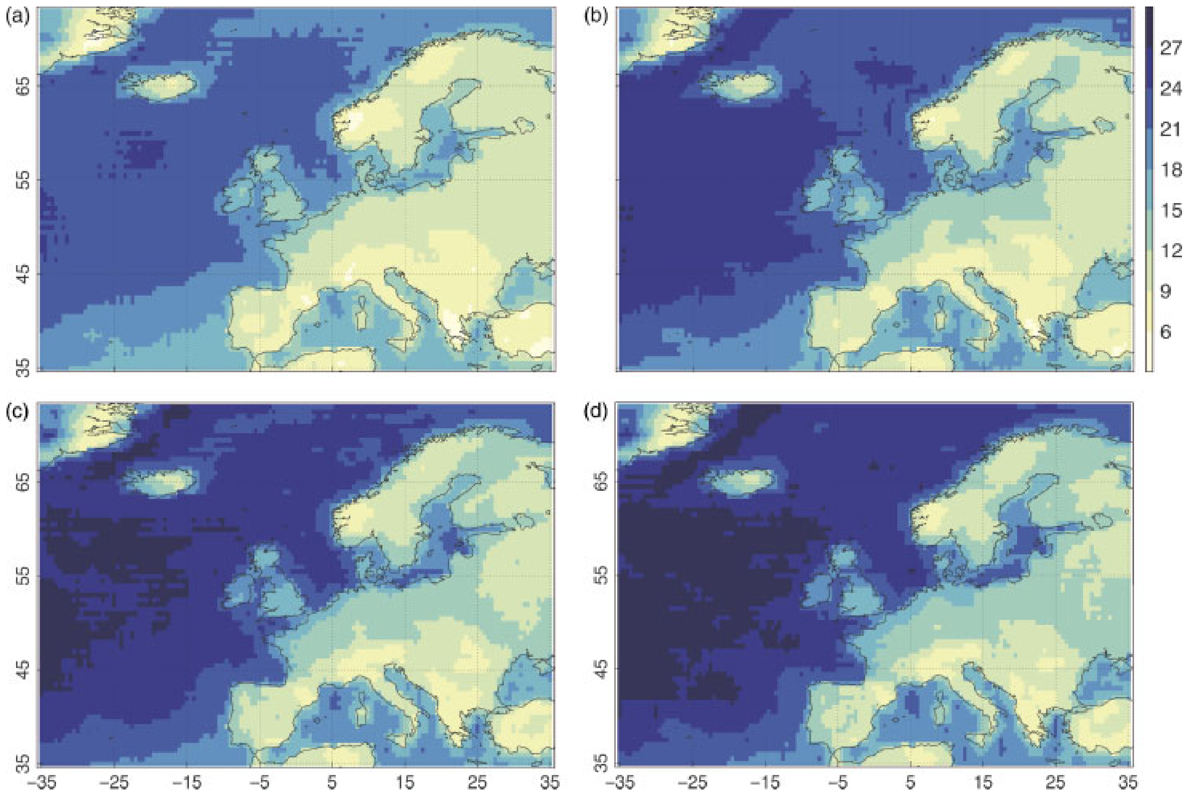

Figure 9. The RL (ms^{-1}) of WS10 at each grid-point for RPs of 1 year (a), 5 years (b), 20 years (c), and 50 years (d). Note that the actual RL magnitude is not representative of absolute (in-situ) wind speed at the surface. This figure is available in colour online at www.interscience.wiley.com/ijoc

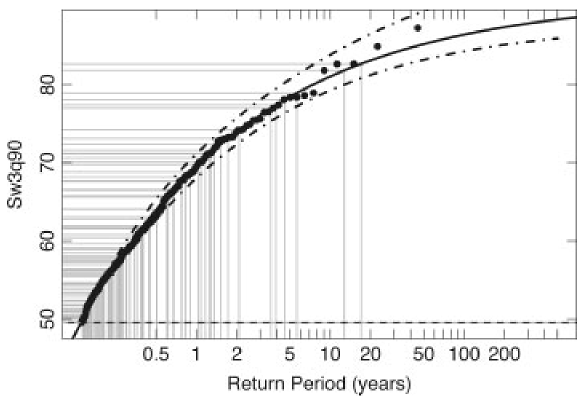

Figure 10. The RP (years) and RL (NDU) of the GPD fit (black line, Equation 2) of the EWI $Sw3q90$, using (WS10) over the whole domain. The black dots represent the maxima of the de-clustered POT series. Dashed dotted lines show the upper and lower bounds of the 95% confidence interval. The horizontal and vertical grey lines denote the RL and RP of the catalogue storms, respectively. The dashed grey line denotes the 95th percentile threshold above which the de-clustered peaks were chosen. Note the log scale on the horizontal axis.

200 points (representing the number of storms in the storm catalogue) together with the 95% confidence interval of each RP (based on the EVA). Generally, there are fair relationships between the RPs of the catalogue storms calculated using the extreme indices and FG10 and WS10 (indicated by the Spearman rank and Kendall Tau non-parametric correlation coefficients at the bottom of each sub-plot). In some cases the RP of each storm is not explained by the uncertainty of the EVA, as indicated by the lack of overlap between the error bars and the diagonal one-to-one line. For example, in Figure 12(a) there is a storm which has a RP of approximately 40 years using $\overline{X}$ and WS10, whereas the same storm using $\overline{X}$ and FG10 is only estimated to be a 2-year RP event. In some cases, the RPs are higher when calculated from WS10 over land compared to using FG10 over land (*cf* Figure 12(c), d and e). This may be explained by the reduction in the number of grid-points used to calculate the EWIs due to the applied mask. Upon further investigation it was found (not shown) that masking the WS10 as if they were FG10 improved the correlation coefficients for the EWIs, $\overline{X}$, $Sfq95$ and $Sfq95q99$ but not for the EWIs $Q95$ and $Sw3q90$. For example, in Figure 12(d) and (e) the low RP storms using FG10 are not visible when WS10 is masked. This result is reasonable given that the $Sfq95$ and $Sfq95q99$ indices are sensitive to local grid-point percentiles. When the majority of the wind storm 'foot print' is located over grid-points which are masked, then the resulting RP estimates are lower. Other differences in the RP estimates stem from more fundamental differences in the datasets than the result of masking. It is also evident that some storms have the lowest RP possible (approximately 0.1 years). Any storm which is not above the fixed POT threshold is given the same RP of λ^{-1} as explained in Section 4.3.

DOI: 10.1002/joc

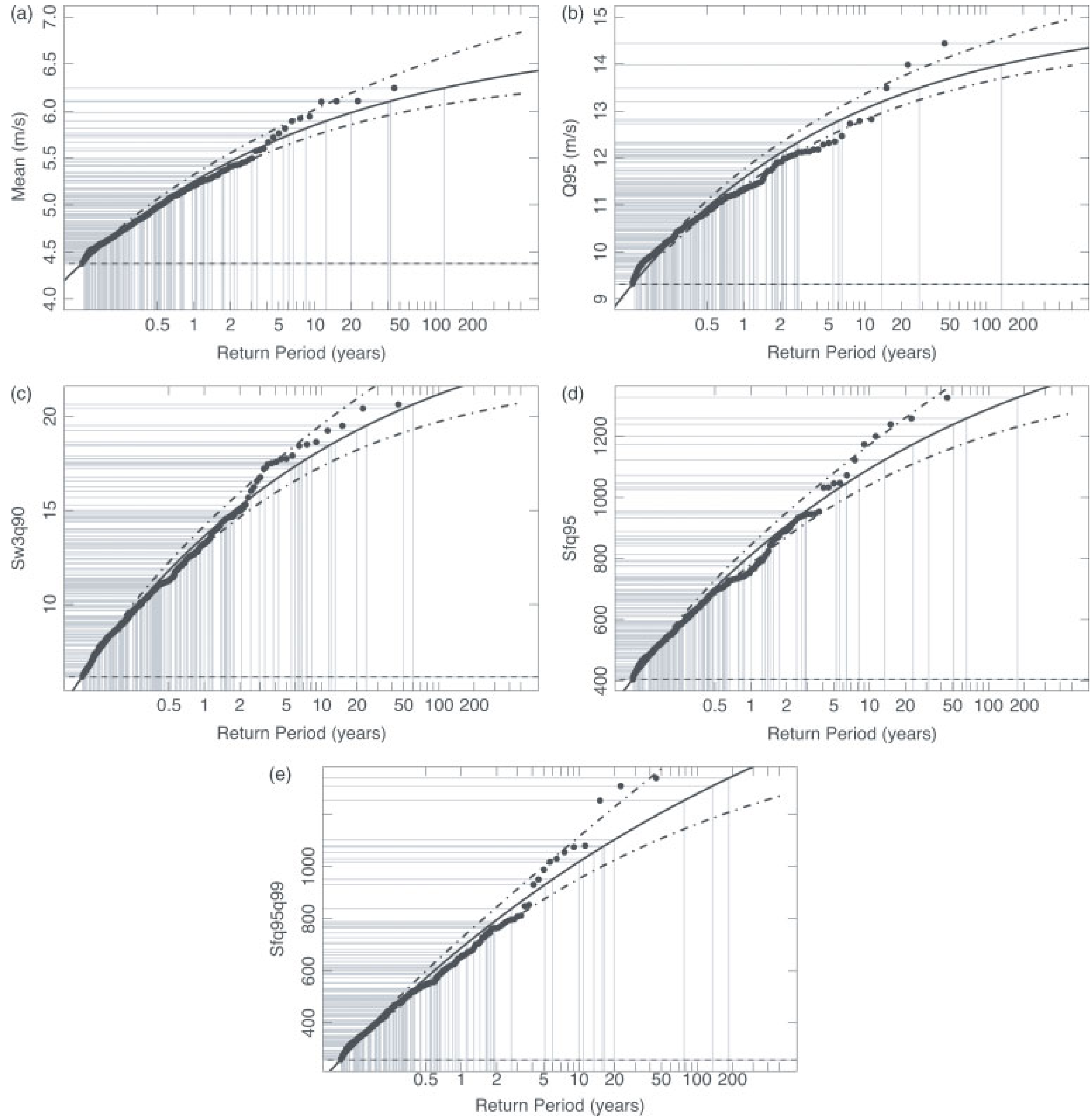

Figure 11. As for Figure 10, but the RP (years) and RL of the GPD fit (black line, Eqn. 2) of the EWI (a) $\overline{X}$, (b) $Q95$, (c) $Sw3q90$, (d) $Sfq95$, and (e) $Sfq95q99$ using WS10 over the land domain.

An intercomparison of catalogue storm RPs from different EWIs shows that with some EWI, data and mask combinations result in very similar RP estimates (in terms of rank correlations, not shown), while others are very different from one another. Inter-index comparisons indicate that $\overline{X}$, $Sfq95$ and $Sfq95q99$ are most highly correlated with each other for both datasets (FG10 and WS10) and for all masks. Whereas, correlations between RPs of $Sw3q90$ and $\overline{X}$ or $Sfq95$ or $Sfq95q99$ are lowest contrasting the differences in the sensitivities of these indices. Inter-dataset indices all show similar correlations (*cf* Figure 12). Note that the range of inter-index correlations ($0.49 <$ Spearman rank correlations $<$ 0.94) is larger than the range of inter-dataset correlations ($0.84 <$ Spearman rank correlations < 0.89), indicating that greater differences between RP estimates are due to the EWIs rather than the datasets. Some plausible explanation for these differences are presented later in Section 5.3 and is related to the sensitivity of the EWIs to the domain and storm catalogue.

Figure B-1 presents an overview of RPs of prominent wind storms from the storm catalogue calculated using WS10 (a) and FG10 (b). It has been placed in Appendix B since we envisage that it could be used as a reference for readers who wish to see individual storm RP estimates (which are not possible to derive from previous figures). For each storm above the threshold, the name (when known) and date, as well as the 95% confidence interval of the storm RP is shown using the EWI $Sw3q90$ (other EWIs are available upon request). The number of

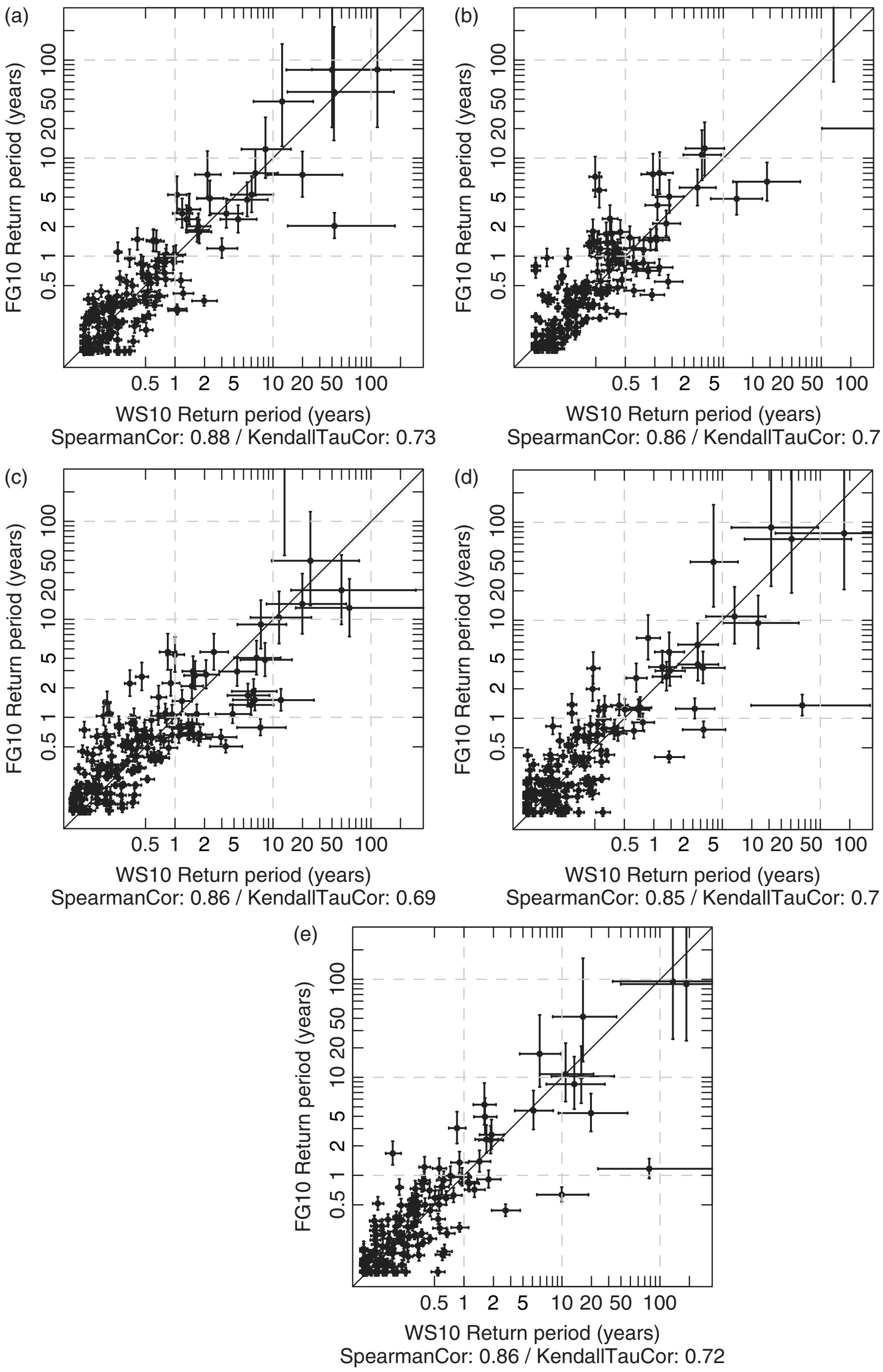

Figure 12. Comparison of RP (years) over land for the 200 catalogue wind storms calculated using FG10 and WS10 for each of the five EWI, (a) $\overline{X}$, (b) $Q95$, (c) $Sw3q90$, (d) $Sfq95$, and (e) $Sfq95q99$. Note the logarithmic scale. 95% confidence intervals for each of the RPs are denoted by the vertical (FG10) and horizontal (WS10) whiskers on each point. At the bottom of each sub-figure is the Spearman rank and Kendall Tau correlation coefficient.

catalogue storms above the 95th percentile threshold of the index $Sw3q90$ is 158 and 161 storms respectively. However, only the common storms above each threshold respectivley are shown for comparison purposes (153). Generally, the range of RP estimates is similar using either FG10 or WS10, and the RPs of individual storms are quite consistent with one another. In summary, the EWIs summarize the extreme wind climatology over a domain and offer a single estimate of the extreme wind RP of a storm which may be intuitively appealing for applications, such as in the reinsurance industry where a single estimate of the intensity of an event is needed to explain an aggregated loss over a portfolio. The EWIs are a spatial summary statistic, and hence, the RPs estimates are representative of the RP of a *storm*, that could have occurred anywhere over the chosen domain.

Similarly as for the EWIs, we investigated the effect of using different data to estimate the RP of the catalogue storms at the grid-point level (distributed view). We found that achieving a consistent estimate of RP for each

DOI: 10.1002/joc

of the catalogue storms at individual grid-points is as difficult as when using different EWIs and datasets. A possible reason for the lack of correspondence between these estimates is due to fundamental differences in FG10 and WS10. Figure 13 presents the spatial distribution of RPs of FG10 and WS10 associated with selected storms in the 1989/1990 and 1999/2000 seasons. Qualitatively the fields in Figure 13 are similar, but however, on a regional and grid-point level the differences are greater, especially for the storm Lothar (Figure 13(e) and (f)). Lothar was a very fast moving mesoscale cyclone and its RP estimates are aliased due to the sampling frequency of WS10 shown as 'islands' of higher RP winds. Recall that the 6-hourly values of FG10 represent the maximum wind gust during a 6-h period, whereas the 6-hourly values of WS10 are the instantaneous analysis values. Another point to keep in mind when interpreting this figure is that the individual grid-point RPs are not the instantaneous RPs at the date and time of the storm, but the maximum RP calculated using the 72-h maximum FG10 and WS10 centred on the storm date/time. This was done in order to ensure that we capture the RP of the storm and not just the RP at the time of analysis. The RP patterns for Vivian (Figure 13(a) and (b)) are qualitatively similar, however, the area of maximum grid-point RPs using FG10 is larger than with the corresponding WS10 analysis. Possible reasons for these discrepancies could be storm-specific dynamics (Wernli *et al*., 2002; Holton, 2004) which produced stronger gust speeds (and hence higher RPs) than are usually associated with the corresponding WS10, or

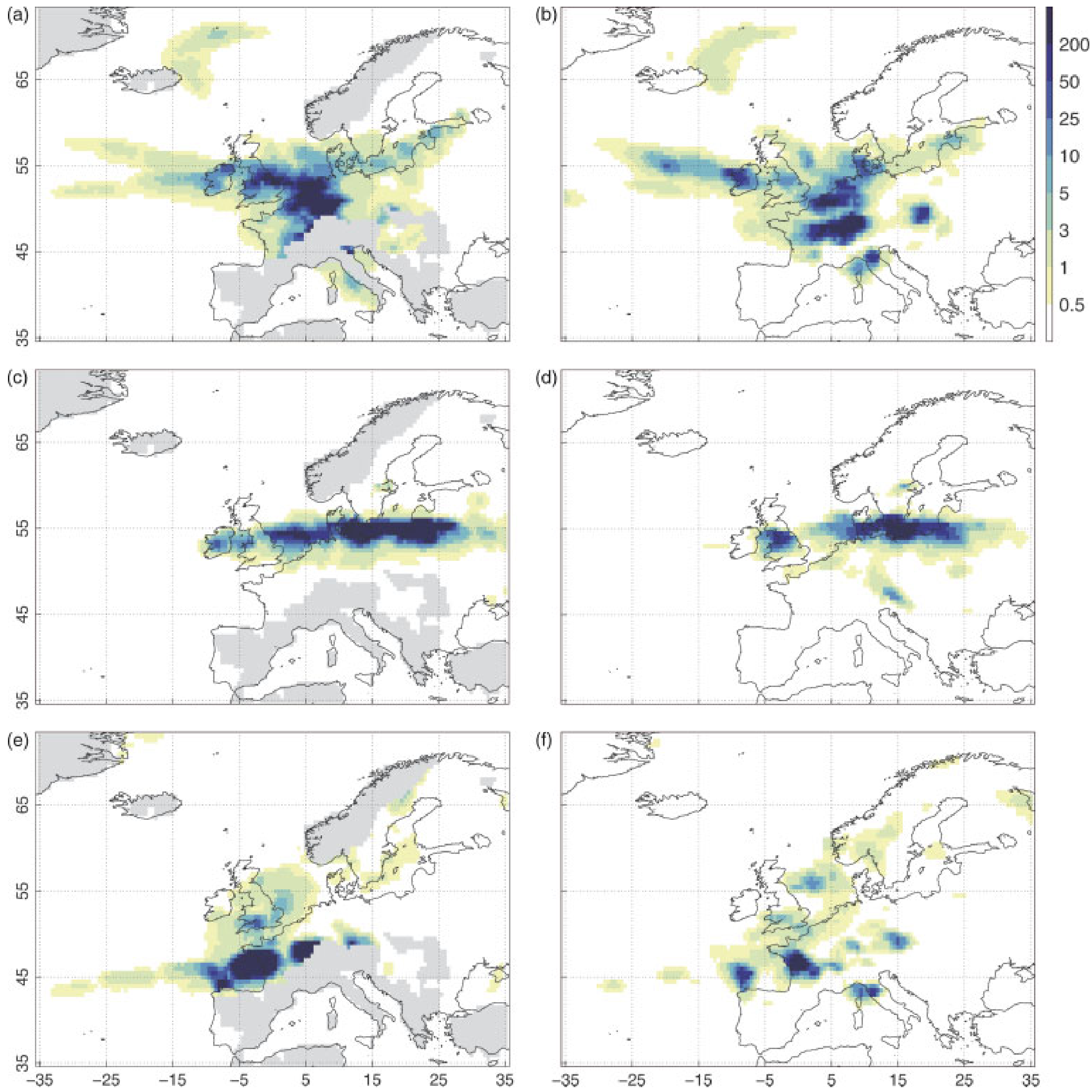

Figure 13. The RP (years) for each grid-point for the selected catalogue storms in the 1989/1990 and 1999/2000 October–April extended winter seasons estimated from FG10 and WS10, (a) Vivian: 26 Feb 1990 1200UTC, using FG10, and (b) using WS10, (c) Anatol: 3 Dec 1999 1200UTC using FG10, and (d) using WS10, (e) Lothar: 26 Dec 1999 0000UTC using FG10, and (f) using WS10. The RP scale is in the top right of the plot. Note that grey areas in (a), (c) and (e) denote masked values. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

remaining problems in the ERA-40 gust parameterization scheme. Storm-specific dynamics are likely to be the reason for this discrepancy since an investigation of the PartnerRe high-resolution wind gust field reveals (not shown) a qualitatively similar pattern to the RPs shown in Figure 13(a). In the other example shown, the storm Anatol has a very similar structure and magnitude of RPs (Figure 13(c) and (d)) using either FG10 or WS10. Another limitation of using WS10 (Figure 13(b) and (f)) is the lack of high RPs over the Alps. This could be due to the climatologically low wind speeds (see Figure 8(a) and Figure 9) at 10 m perhaps caused by problems in the parameterization of wind speed in areas of complex orography. It is clear from other examples (not shown) that the estimation of RPs of storms that are less than 72 h apart from each other are either very similar or very disparate (e.g. Lothar and Martin, separated by approximately 48 h) due to the criteria of taking the maximum grid-point wind RP over a 72-h period.

A qualitative comparison of RPs calculated from the generalized EWIs and the distributed grid-point approach demonstrates their utility in estimating the RPs of catalogue storms during the 1989/1990 and 1999/2000 extended winter seasons. Tables 1 and 2 summarize the RPs calculated using $Sw3q90$ and WS10 or FG10 over land and the range of grid-point RPs over land, respectively. The most severe wind storm in these two seasons (according to this index and dataset, Table 1), is associated with the storm Daria, with an RP estimated to be 24.1 years (see also Figure 14(b)). Its corresponding RP estimates using the same index, but FG10, is approximately 39 years (Table 2), and although these estimates differ they are well contained within each of the uncertainty estimates. In the grid-point analysis the same storm has produced local winds over land to be between 0.1 and 200+ years (see Figure 14(b)) with some of the individual grid-point RPs as high as 1500 years in the English Channel area, where the storm had its highest intensity. By comparing the value of the third quartile of grid-point RPs with the $Sw3q90$ RPs for all storms and both datasets (Tables 1 and 2), we can say that the EWI RP and the grid-point RPs are in qualitative rank-order agreement for most storms. The major exception to this rule is for the storm Vivian, where the EWI RP estimates differ greatly between datasets (13.1 years using WS10, and 374 years using FG10) but do not differ greatly comparing the third quartile of grid-point RPs (10.6 years using $WS10$ and 9.4 years using FG10). However, this result is not unreasonable given the large uncertainty in the EWI RP estimates and individual differences in storm dynamics represented by FG10 and WS10 as stated above. It is interesting to note that when using $Sfq95$ the six least frequent storms are all unnamed storms and different to those derived from $Sw3q90$ contrasting the differences between these indices. If we look in more detail to the storms Lothar and Martin we can see that the RP estimates differ substantially between datasets due to both the mask applied to FG10 and the aliasing of the signal due to WS10. While the grid-point RP analysis also exhibits dependence on the dataset at the grid-point level, qualitatively the pattern and magnitude of the RPs are similar (see Figure 13(a)–(d)). The RPs for storm Lothar over Switzerland (Figure 13(e) and (f)) are hard to compare with Albisser *et al.* (2001) due to the masking and aliasing effects of the ERA-40 data used here.

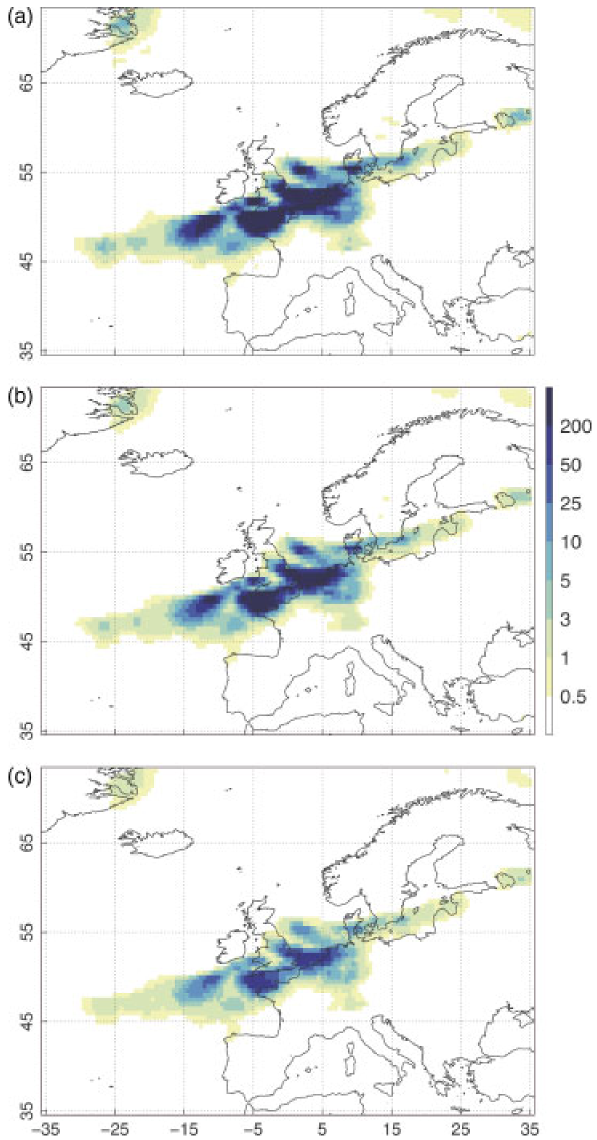

Figure 14. As for Figure 13 but the RP (years) of WS10 for each grid-point estimated for the storm Daria: 26 Jan 1990 0000UTC (b). In (a) and (c) are shown the upper (lower) bound of the 95% confidence interval of the RP (years). This figure is available in colour online at www.interscience.wiley.com/ijoc

5.3. Representativity of continental RPs

A basic evaluation of the ability of the EWI RP estimates to represent the grid-point wind-based RP estimates for the 200 wind storms has been performed using a Spearman rank (and Kendall Tau) correlation analysis. Figure 15 demonstrates that the rank of the RPs of $Sw3q90$ are most highly correlated with the rank of the RPs from the grid-point analysis over a region centred in the mid-western part of the domain over the North Atlantic Ocean (Figure 15(a)) with r values in the order

DOI: 10.1002/joc

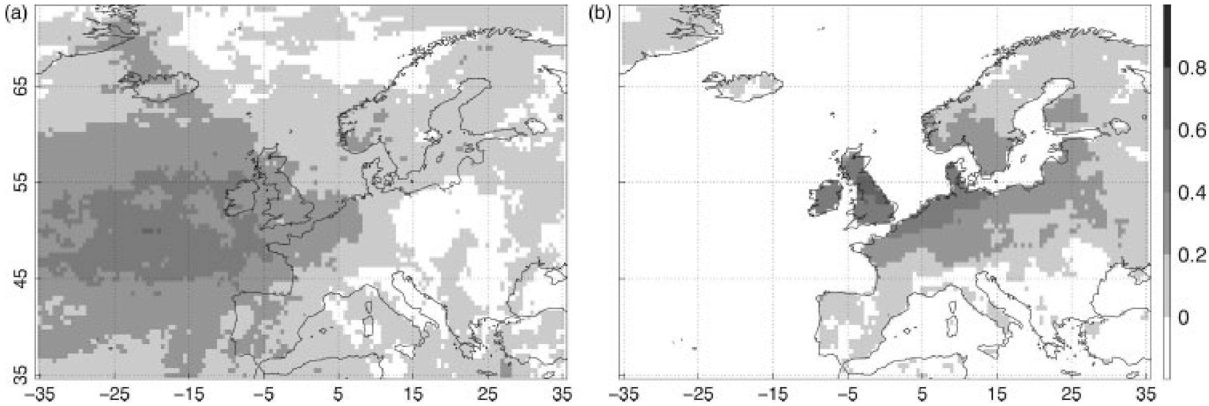

Figure 15. The Spearman rank correlation between the RPs of the 200 catalogue storms based on $Sw3q90$ of WS10 and the RP at each grid-point based on WS10, (a) using all grid-points, and (b) land grid-points only.

of 0.4–0.6. The rank correlation decreases radially from this point such that r over the European coast and further inland is between 0 and 0.3 implying that these indices only share a small fraction of variability of the local grid-point RPs. Results are better if we consider EWI using land grid-points (Figure 15(b)). Here, again we see a 'bullseye' centred on western central Europe. Within these regions, r ranges from 0 to 0.7 indicating that a reasonable amount of shared variability exists between grid-point RPs and the EWI RPs. For the other indices (except $Q95$) it is important to note that the centre of highest r over the land-only domain is located further east (east of Germany, not shown). This may explain why $Sw3q90$ and $Q95$ have a higher number of catalogue storms above the 95% quantile threshold (161 and 152, respectively) than the other EWI (145, 137 and 143 for $\overline{X}$, $Sfq95$ and $Sfq95q99$, respectively) since the region of highest sensitivity of the index also coincides with the region over which most of the catalogue storms have been selected.

6. Conclusions and discussion

We use state-of-the-art reanalysis data combined with extreme value analysis techniques to estimate the climatology of extreme winds and the recurrence frequency of prominent wind storm events over Europe during the period from 1957 to 2002. If we return to the original aims of the study which are reiterated below, we conclude the following from this analysis:

How reliable are wind parameters in the ERA-40 reanalysis (spatial representativity, temporal homogeneity) for estimating RPs of high-impact storm events over Europe?

- FG10 from ERA-40 should not be used in areas where the roughness length parameter is high (>3m), due to the boundary-layer physics in the ERA-40 model. Such areas should be masked from the analysis (*cf* Figure 1).WS10 from ERA-40 appears to be free from the problems associated with high wind gust values, however, 10-m winds over complex orography might to be too low (*cf* Figure 8(a) and Figure 9).
- WS10 (and any other analysed wind-speed parameters) represents the instantaneous wind speed at the reanalysis output time which results in aliasing of the wind speed and hence contributes to the unreliability of RPs in both the generalized and distributed approaches to RP estimation. FG10 represents the maximum wind gust during the previous 6 h and does not suffer from this problem.
- Both the generalized and distributed extreme wind climatologies demonstrate the expected land–sea contrasts and a negative shape parameter of the GPD.
- The distributed wind storm climatology shows an increased storm frequency and decreased storm clustering in the Northeast Atlantic associated with the storm track, and the Mediterranean region of secondary cyclogenesis confirming the results of Mailier *et al.* (2006) (*cf* Figure 8(e)).

How sensitive are estimates of RPs to the details in the definition of the EWIs?

- The EWIs $\overline{X}$, $Sfq95$ and $Sfq95q99$ result in similar RP estimates for individual storms. These indices should be used when more weight to local winds relative to their climatology is needed, and to find storms also affecting regions not subject to regular extreme wind events.
- The EWIs $Sw3q90$ and $Q95$ result in similar RP estimates for individual storms. These indices should be used when more weight to the absolute magnitude of a wind storm is needed, regardless of the local wind climatology.
- The RPs estimates from $Sw3q90$ and $Q95$ are more representative of loss figures published in MunichRe (2000) than the remaining EWIs. We note that the quality of the GPD fit to Q95 shown in Figure 11(b) is lower than other indices.
- EWI storm RPs are most sensitive to the domain over which they are calculated, and to a lesser extent, the definition of the index. Higher RPs result from using land-only grid-points than from using the whole domain. This is partly due to the storm catalogue being biased towards wind storms which had an impact on

DOI: 10.1002/joc

continental Europe, especially in the northwest, and due to the process of spatial averaging.

What is the degree of uncertainty in estimating continental-scale and grid-point-scale RPs for past storms? Which factors contribute most to the uncertainty?

- The largest uncertainties in RP estimates (both the generalized and distributed approaches) derive from the differences in the wind gust and wind-speed data (such as masking and aliasing, mentioned above) generally for RPs less than 10 years. This uncertainty tends to be comparable with the uncertainty associated with the GPD fit at higher RPs (*cf* Figure 12, Tables 1 and 2).
- The RPs calculated using either $Sw3q90$ or $Q95$ compared with either $\overline{X}$, $Sfq95$ or $Sfq95q99$ shows that the EWI definition plays a greater role in the RP differences than the wind parameter chosen.
- The spatial distribution of RPs for catalogue storms is generally similar using either WS10 or FG10. Physical processes leading to extremes specific to individual storms may not be adequately captured by WS10 and hence the calculation of RPs.
- The method of taking the maximum RP within a 72-h period prevents an accurate RP estimate for catalogue storms which occurred less than 72 h apart from one another. This aspect of the data processing prevents the estimates of RPs from storms such as Lothar and Martin from being independent. This method also potentially biases RPs to the time when the storm has its maximum intensity, and not necessarily at its peak impact.
- In some cases, the catalogue storm dates/times do not match the cluster maxima of the EWI or the grid-point wind exactly. This could be due to the fact that the catalogue storm dates were recorded when the storm had the highest impacts, and not necessarily when the storm had its highest intensity.

How representative are continental-scale estimates of RPs as a measure of the local recurrence of a storm?

- The generalized and distributed RP estimates share up to approximately 50% of common variations. The larger the domain considered the lower the effectiveness of the EWI at explaining local wind RPs. The $Sw3q90$ and $Q95$ indices have the highest correlations with local RP estimates due to these indices being based on an absolute wind-speed magnitude. Best correlations coincide with regions of climatologically highest magnitude winds (over land, *cf* Figure 9). The success of these indices is also partly due to the way in which the storm catalogue was created. The storm catalogue is biased towards storms which occurred in the region of interest of the source documents, which in this case, is mainly focused on northwestern Europe.

We can also state various conclusions associated with the chosen EVA methodologies: The generalized Pareto distribution is a robust extreme-value model of the de-clustered peak over threshold (POT) series obtained from EWI and grid-point wind speed. The POT series was obtained using an automatic de-clustering method which largely avoided the need for the specification of arbitrary parameters as is the case with more common de-clustering methods. To determine the uncertainty of the extreme wind climatology and catalogue storm RPs, the profile log-likelihood method is used as it gives more physically meaningful results than other common methods. The choice of de-clustering method (either the Ferro and Segers (2003) or runs-de-clustering (Coles, 2001), not shown, see Della-Marta *et al.*, 2007), has little effect on RPs of catalogue storms calculated using EWIs, however, the quality of the GPD fit can be affected depending on whether the seasonal cycle of the index is included in the de-clustering.

The question of determining the RP of a 'wind storm' is rather ill-posed since a wind storm has many degrees of freedom which are difficult to summarize in a generalized or a distributed approach. The EWIs are a spatial summary statistic, and hence, the estimates of RPs are more representative of the RP of a *storm* that could have occurred anywhere over the chosen domain. We conclude that it is very difficult to obtain a spatial summary statistic (EWI) which works equally well for each type of storm event in the context of trying to estimate a *storm* RP, although $Sw3q90$ seems to have higher representativity of the given catalogue storms than the other indices. On the other hand, the grid-point wind speed RPs also show dependency between datasets, and may not be used to estimate the RP of the *storm* given that no spatial dependence between grid-point RPs is taken into account.

We did not include any treatment of non-stationarity in our GPD model since there is little evidence of any long-term linear trends in storminess over Europe during the past 100 years (e.g. Alexandersson *et al.*, 2000) however, there remains decadal variability in the dataset (e.g. the more frequent and higher-intensity storms in the 1990s) which does affect the interpretation of RPs such as those listed in Tables 1 and 2. For example, both Daria and Vivian, two high-RP storms, occurred during the same season. Further analysis of the trends in the DPOT series of each EWI show (not shown) an increase (significant at 0.05 level) over land areas using $Sw3q90$, whereas, no other index indicates a trend during this period.

As more accurate and longer datasets of either *in-situ* wind data, pressure datasets or dynamically down-scaled reanalyses become available (Diaz *et al.*, 2002; Alexander *et al.*, 2005; Ansell *et al.*, 2006; Heneka *et al.*, 2006; Schwierz *et al.*, 2008) this analysis should be repeated in order to minimize the uncertainty in the RP calculations. We also recommend further research into the spatial structure of historical wind storm events through the application of advanced spatial EVA techniques (Coles, 2001). Given the lack of longer and more

DOI: 10.1002/joc

Table I. The RPs of some prominent wind storms in the storm catalogue during the 1989/1990 and 1999/2000 extended winter season estimated using the EWI $Sw3q90$ based on WS10.

Storm Name	Date/Time	Lower (Years)	RP (Years)	Upper (Years)	Region affected	Grid Point RP range: Min.	1st Qu.	Median	3rd Qu.	Max.
unknown	17-12-1989 12:00	0.31	0.36	0.40	GB, ES, PT, IT, SI, HR	0.1	0.42	0.71	1.98	200+
Daria	26-01-1990 00:00	9.72	24.10	75.80	GB, FR, BE, NL, LU, DE, DK	0.1	0.48	1.28	5.28	200+
Herta	04-02-1990 00:00	0.88	1.09	1.37	IE, GB, FR, DE	0.1	0.46	0.71	1.47	200+
Nana	12-02-1990 12:00	0.49	0.58	0.69	ES, FR, IE, GB	0.1	0.43	0.72	1.20	10.3
Vivian	26-02-1990 12:00	6.46	13.10	29.60	IE, GB, FR, BE, NL, LU, CH, DE, DK, SE, PL, LT, LV, EE, AT, CZ	0.1	0.61	1.88	10.60	200+
Wiebke	01-03-1990 00:00	3.64	5.97	10.10	FR, DE, AT, CZ, CH, IT	0.1	0.55	1.12	3.32	137.0
Anatol	03-12-1999 12:00	2.82	4.31	6.72	GB, NL, DE, DK, PL, LT, BY	0.1	0.48	0.86	2.46	200+
Lothar	26-12-1999 00:00	0.81	1.00	1.24	GB, FR, CH, AT	0.1	0.41	0.66	1.34	200+
Martin	28-12-1999 00:00	4.30	7.45	13.50	ES, FR, CH, AT, IT	0.1	0.43	0.74	2.71	200+
Kerstin	29-01-2000 12:00	0.57	0.68	0.81	GB, DK, SE, FI	0.1	0.41	0.63	1.02	122.0

Columns 3 and 5 show the lower and upper bound of the 95% confidence interval of the storm RP (Column 4). The 'Region affected' column shows the countries most affected by the storms assessed using the grid-point RP estimates. The country codes are standard ISO two-letter codes (see http://www.iso.org/iso/country_codes/). Columns 9–11 give an impression of the distribution of grid-point RPs over land areas within the domain, where Min. is the minimum grid-point RP, 1st (3rd) Qu. is the first (third) quartile, Median is the median, and Max. is the maximum grid-point RP where RPs greater than 200 years are replaced by '200+', since the error associated with these RP estimates is extremely large.

Table II. The RPs of some prominent wind storms in the storm catalogue during the 1989/1990 and 1999/2000 extended winter season estimated using the EWI $Sw3q90$ based on FG10. See Tab. 1 caption for more details.

Storm Name	Date/Time	Lower (Years)	RP (Years)	Upper (Years)	Region affected	Grid Point RP range: Min.	1st Qu.	Median	3rd Qu.	Max.
unknown	17-12-1989 12:00	0.42	0.49	0.57	GB, ES, PT, IT, SI, HR	0.1	0.38	0.71	1.15	4.73
Daria	26-01-1990 00:00	13.90	39.60	125.00	GB, FR, BE, NL, LU, DE, DK	0.1	0.53	1.88	6.79	200+
Herta	04-02-1990 00:00	0.55	0.65	0.77	IE, GB, FR, DE	0.1	0.40	0.60	0.90	3.20
Nana	12-02-1990 12:00	0.52	0.62	0.73	ES, FR, IE, GB	0.1	0.52	1.13	2.04	10.20
Vivian	26-02-1990 12:00	45.00	374.00	4000.00	IE, GB, FR, BE, NL, LU, CH, DE, DK, SE, PL, LT, LV, EE, AT, CZ	0.1	0.88	1.92	9.42	200+
Wiebke	01-03-1990 00:00	1.06	1.34	1.72	FR, DE, AT, CZ, CH, IT	0.1	0.52	0.90	3.11	200+
Anatol	03-12-1999 12:00	2.09	2.93	4.17	GB, NL, DE, DK, PL, LT, BY	0.1	0.52	1.28	4.52	200+
Lothar	26-12-1999 00:00	2.90	4.36	6.63	GB, FR, CH, AT	0.1	0.46	0.92	3.65	200+
Martin	28-12-1999 00:00	0.65	0.79	0.96	ES, FR, CH, AT, IT	0.1	0.42	0.78	2.21	200+
Kerstin	29-01-2000 12:00	0.84	1.03	1.29	GB, DK, SE, FI	0.1	0.46	0.69	1.10	83.80

Int. J. Climatol. (2008)
DOI: 10.1002/joc

reliable reanalysis or in-situ data, utilization of the ever-increasing amount of dynamical ensemble prediction data (van den Brink *et al.*, 2004a,b; Jung *et al.*, 2005; Frei *et al.*, 2006) is suggested as a means to build more realistic estimates of the frequency of wind storms.

In conclusion, the results demonstrate that significant progress has been made through the development of an extreme wind climatology based on a robust reanalysis dataset consistent in space and time. We have created a climatology of wind storm RPs based on EWIs together with estimates of their local wind RP using the grid-point approach. Considerable challenges remain both from the methodological treatment of spatial extreme events and from the limitations imposed by the currently available datasets.

Acknowledgements

We wish to acknowledge PartnerRe who substantially sponsored this research. We wish to thank Hervé Castella, Malcolm Haylock, Manuel Prechtl from PartnerRe, and Felix Fundel from MeteoSwiss, for useful discussions and inputs to this paper. Malcolm Haylock suggested the formulation of the index $Sw3q90$. Two anonymous reviewers provided detailed comments and suggestions which resulted in a substantially improved paper. Paul Della-Marta was partially funded by the 6th EU framework program ENSEMBLES contract number GOCE-CT-2003-505539, and by the Swiss National Science Foundation through the National Centre for Competence in Research Climate (NCCR-Climate) Sponsors: Reinsurance Company, Zurich, Switzerland 6th EU framework program ENSEMBLES contract number GOCE-CT-2003-505539 Swiss National Science Foundation through the National Centre for Competence in Research Climate (NCCR-Climate).

Appendix A

This appendix describes each of the Extreme Wind Indices (EWI) in mathematical notation for the benefit of readers who may wish to implement such indices. The indices are denoted in terms of a generic wind variable, W and could be substituted for either FG10 or WS10. Where possible, we tried to take into account the unequal areas of each grid box by weighting sums and multipliers by the cosine of the latitude of each grid-point. A rationale for each index is given in Section 3.

$\overline{X}$: Mean wind (units of ms^{-1}).

$$\overline{X}\,(t) = \frac{1}{N_{\kappa\delta}} \sum\nolimits_{x,y\in\delta} \kappa(x, y) w(x, y, t) \qquad \text{(A1)}$$

where κ are the individual grid-point weights which only depend on y, $\kappa\,(x, y) = \cos\,(latitude\,(y))$, $N_{\kappa\delta} = \sum_{x,y\in\delta} \kappa\,(x, y)$ and δ denotes the domain.

$Q95$: The spatial 95% quantile wind (units of ms^{-1}).

$$Q95(t) = F_*^{-1}(p) = \min\{w : p \le F_*(W)\} \qquad \text{(A2)}$$

where $p = 0.95$ and F_* is the latitude weighted empirical cumulative distribution function of $\{w(x, y, t) : (x, y) \in \delta\}$ where δ denotes the domain. The weighted cumulative distribution function is given by Horvitz and Thompson, 1952; R Development Core Team, 2005).

$$F_*\,(W) = \frac{1}{N_{\kappa\delta}} \sum\nolimits_{x,y\in\delta} \kappa(x, y) \mathbb{1}(w(x, y, t) \le W) \quad \text{(A3)}$$

Where κ are the individual grid-point weights and $N_{\kappa\delta}$ is given above and $\mathbb{1} = \begin{cases} 1 : w(x, y, t) \le W \\ 0 : otherwise \end{cases}$.

$Sw3q90$: Cube root of the sum of wind cubed above the domain climatological 90% quantile (non-dimensional).

$$Sw3q90(t) = \sqrt[3]{\begin{array}{c} \sum_{x,y\in\delta} (\mathbb{1}_{\{>\overline{q90}\}}(w(x, y, t) > \overline{q90}) \\ \cdot\kappa(x, y) \left(w(x, y, t) - \overline{q90})\right)^3 \end{array}} \qquad \text{(A4)}$$

where κ are the weights given above, the $\mathbb{1}_{\{>\overline{q90}\}} = \begin{cases} 1 : w\,(x, y, t) > \overline{q90} \\ 0 : otherwise \end{cases}$. The domain mean quantile function $\overline{q90}$ is given by:

$$\frac{1}{N_\delta} \sum_{x,y\in\delta} q90(x, y) \qquad \text{(A5)}$$

where $q90(x, y) = F^{-1}(p) = \min\{w : p \le F(W)\}$, $p = 0.90$, F is the empirical cumulative distribution function of $\{w\,(x, y, t) : t \in ONDJFMA\}$

$Sfq95$: Sum of the fraction of wind divided by the grid-point climatological 95% quantile (non-dimensional).

$$Sfq95(t) = \sum_{x,y\in\delta} \mathbb{1}_{\{>1\}} \left(\frac{w(x, y, t)}{q95(x, y)}\right) \cdot \kappa(x, y) \frac{w(x, y, t)}{q95(x, y)} \qquad \text{(A6)}$$

where κ are the weights given above, the $\mathbb{1}_{\{>1\}} = \begin{cases} 1 : \left(\frac{w(x,y,t)}{q95(x,y)}\right) > 1 \\ 0 : otherwise \end{cases}$. The grid-point quantile function $q95$ is given by:

$$q95(x, y) = F^{-1}(p) = \min\{w : p \le F(W)\} \qquad \text{(A7)}$$

where $p = 0.95$, F is the empirical cumulative distribution function of $\{w\,(x, y, t) : t \in ONDJFMA\}$

$Sfq95q99$: Sum of the fraction of extreme wind divided by the length of the distribution tail (non-dimensional).

$$Sfq95q99(t) = \sum_{x,y\in\delta} \mathbb{1}_{\{>0\}}\left(\frac{w(x,y,t)-q95(x,y)}{q99(x,y)-q95(x,y)}\right) \cdot \kappa(x,y)\frac{w(x,y,t)-q95(x,y)}{q99(x,y)-q95(x,y)} \quad (A8)$$

where κ are the weights given above, the $\mathbb{1}_{\{>0\}} = \begin{cases} 1 : \left(\frac{w(x,y,t)-q95(x,y)}{q99(x,y)-q95(x,y)}\right) > 0 \\ 0 : otherwise \end{cases}$. The grid-point quantile functions, $q95$ and $q99$ are given above.

Appendix B

In this appendix, we present two figures which show the RP of many well known European winter wind storms that occurred during the ERA-40 reanalysis period (1957–2002). A list of prominent winter wind storms was compiled from various available sources, and is detailed in Section 2.1.

References

Albisser P, Aller D, Bader S, Brnag P, Broccard A, Bründl A, Dobbertin M, Forster B, Frutig F, Hächler P, Hänggli T, Jacobi C, Müller E, Neu U, Niemeyer S, Nöthiger CJ, Quiby J, Rickli C, Schilling A, Schibiger F, Truog G. 2001. *Lothar Der Orkan 1999*, Eidg. Forchungsanstalt WSL, Bundesamt für Umwelt, Wald und Landschaft BUWAL, Birmensdorf, Bern.

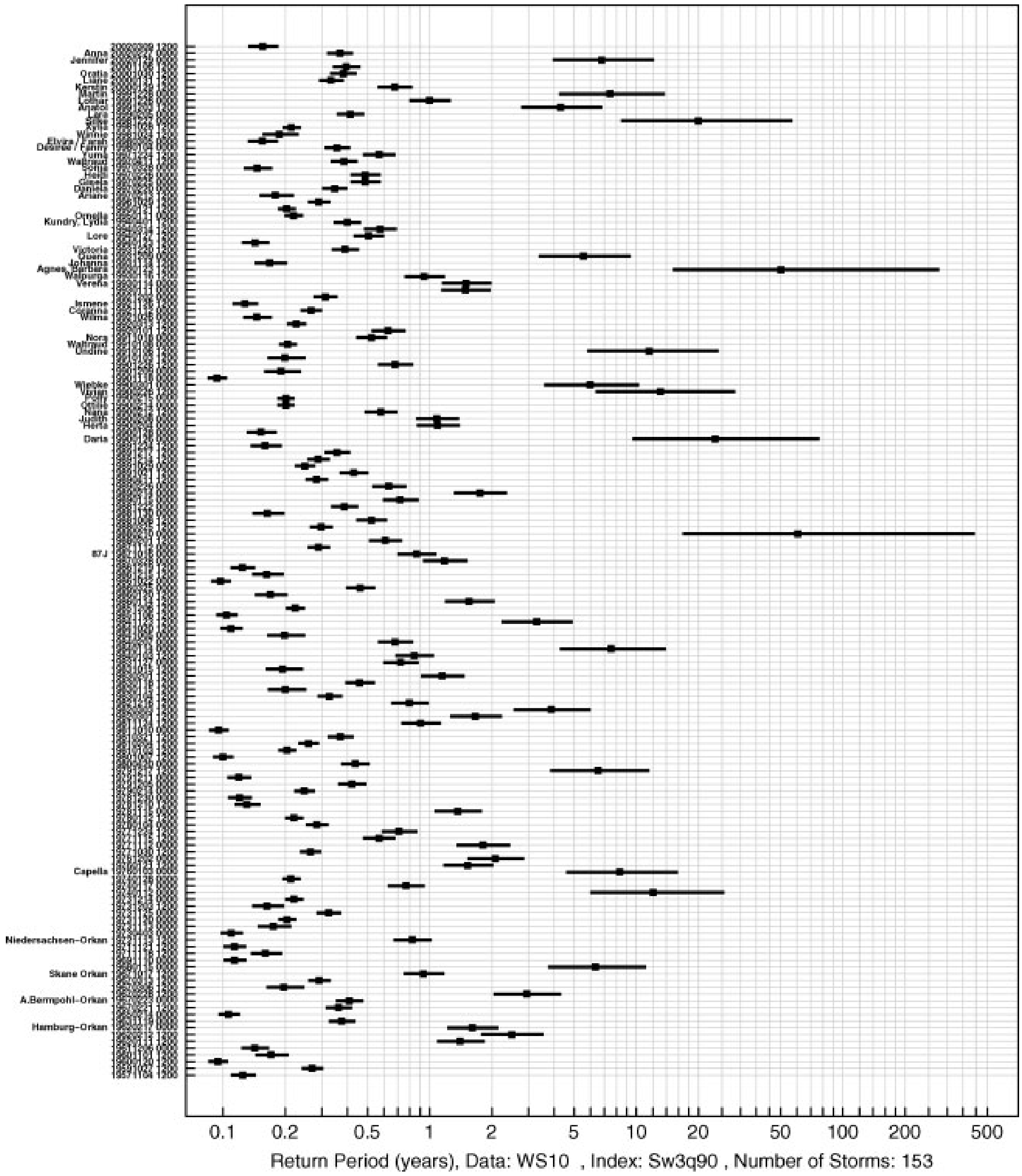

Figure B-1. The RP (years) of catalogue wind storms above the 95th percentile threshold for the EWI $Sw3q90$ calculated using only land grid-points and WS10 (a) and FG10 (b). On the vertical axis is the name of the storm (when known), and the date and time (UTC, see Section 2.1 for more details). The horizontal axis (RP, years) is logarithmic. The number of catalogue storms above the threshold matched in both datasets is shown in the bottom right of each sub-plot.

DOI: 10.1002/joc

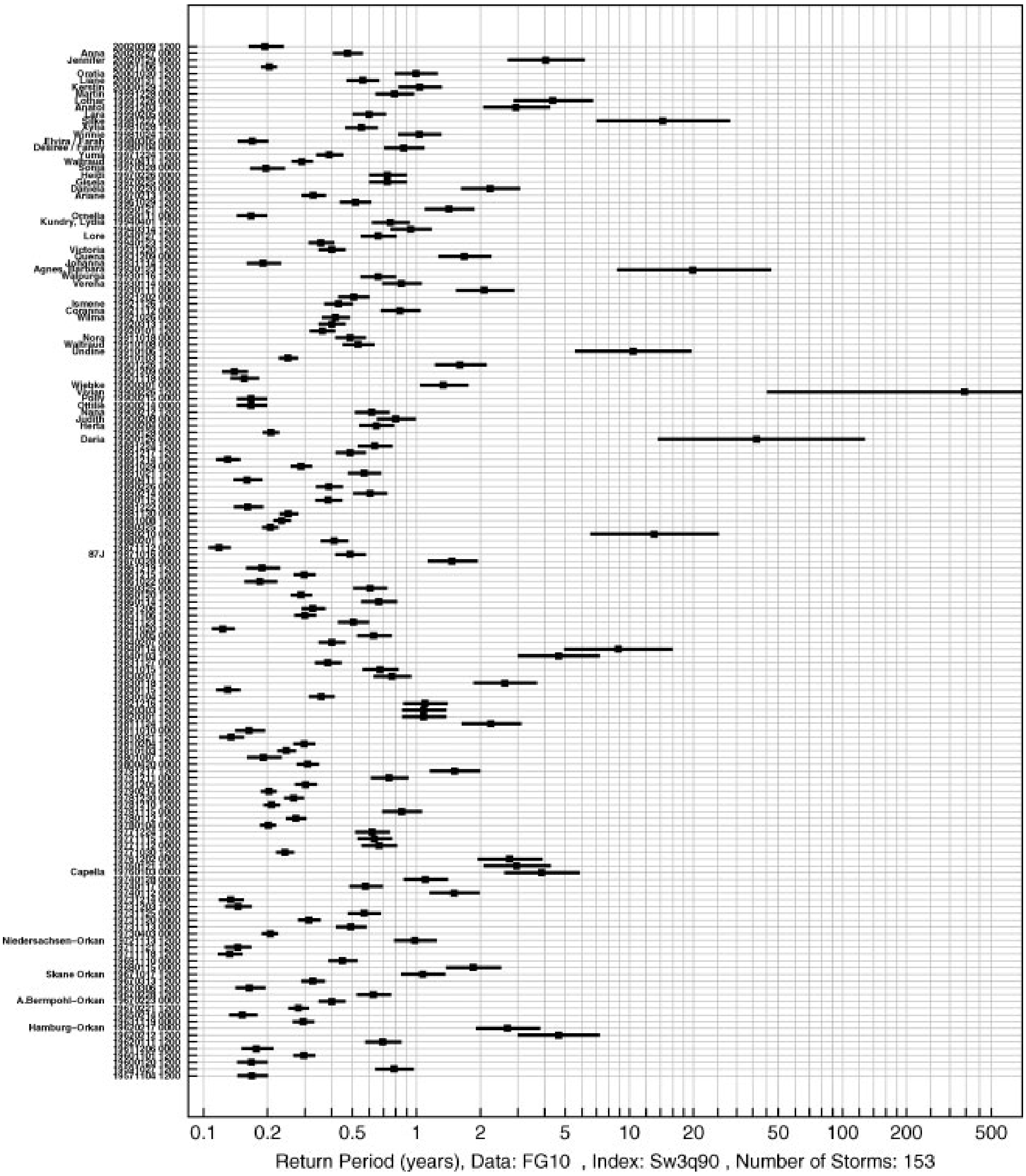

Figure B-1. (*Continued*).

Alexander LV, Tett SFB, Jonsson T. 2005. Recent observed changes in severe storms over the United Kingdom and Iceland. *Geophysical Research Letters* **32**(13): L13 704.

Alexandersson H, Schmith T, Iden K, Tuomenvirta H. 1998. Long-term variations of the storm climate over NW Europe. *Global Atmospheric and Ocean Systems* **6**: 97–120.

Alexandersson H, Tuomenvirta H, Schmith T, Iden K. 2000. Trends of storms in NW Europe derived from an updated pressure data set. *Climate Research* **14**(1): 71–73.

Ansell TJ, Jones PD, Allan RJ, Lister D, Parker DE, Brunet M, Moberg A, Jacobeit J, Brohan P, Rayner NA, Aguilar E, Alexandersson H, Barriendos M, Brandsma T, Cox NJ, Della-Marta PM, Drebs A, Founda D, Gerstengarbe F, Hickey K, Jonsson T, Luterbacher J, Nordli O, Oesterle H, Petrakis M, Philipp A, Rodwell MJ, Saladie O, Sigro J, Slonosky V, Srnec L, Swail V, Garcia-Suarez AM, Tuomenvirta H, Wang X, Wanner H, Werner P, Wheeler D, Xoplaki E. 2006. Daily mean sea level pressure reconstructions for the European-North Atlantic region for the period 1850–2003. *Journal of Climate* **19**(12): 2717–2742.

Appenzeller C, Stocker TF, Anklin M. 1998. North Atlantic Oscillation dynamics recorded in Greenland ice cores. *Science* **282**(5388): 446–449.

Barkhausen M. 1997. Untersuchung von Intensitäten und auftritts-wahrscheinlichkeiten aussertropischer Stüme in Mittel-und Westeuropa anhand historischer Sturmfelder (1967–1997), Master's thesis, Institute of Meteorology, University of Munich.

Barring L, von Storch H. 2004. Scandinavian storminess since about 1800. *Geophysical Research Letters* **31**(20): L20 202.

Beersma JJ, Rider KM, Komen GJ, Kaas E, Kharin VV. 1997. An analysis of extra-tropical storms in the North Atlantic region as simulated in a control and 2xCO(2) time-slice experiment with a high-resolution atmospheric model. *Tellus Series A-Dynamic Meteorology and Oceanography* **49**(3): 347–361.

Bengtsson L, Hagemann S, Hodges KI. 2004. Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research-Atmospheres* **109**: D11 111.

Bengtsson L, Hodges KI, Roeckner E. 2006. Storm tracks and climate change. *Journal of Climate* **19**(15): 3518–3543.

Bijl W, Flather R, de Ronde JG, Schmith T. 1999. Changing storminess? An analysis of long-term sea level data sets. *Climate Research* **11**(2): 161–172.

Caires S, Sterl A. 2005. 100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. *Journal of Climate* **18**(7): 1032–1048.

Carretero JC, Gomez M, Lozano I, de Elvira AR, Serrano O, Iden K, Reistad M, Reichardt H, Kharin V, Stolley M, von Storch H, Gunther H, Pfizenmayer A, Rosethal W, Stawarz M, Schmith T, Kaas E, Li T, Alexandersson H, Beersma J, Bouws E, Komen G, Rider K, Flather R, Smith J, Bijl W, de Ronde J, Mietus M, Bauer E, Schmidt H, Langenberg H. 1998. Changing waves and storms in the Northeast Atlantic? *Bulletin of The American Meteorological Society* **79**(5): 741–760.

Choulakian V, Stephens MA. 2001. Goodness-of-Fit Tests for the Generalized Pareto Distribution. *Technometrics* **43**: 478–484.

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon WT, Laprise R, Magana Rueda V, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P. 2007. *Regional Climate Projections in: Climate change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, New York; 847–940, chapter. 11.

Coles S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer: London.

Della-Marta PM, Mathis H, Frei C, Liniger MA, Appenzeller C. 2007. Extreme wind storms over Europe: Statistical analyses of ERA−40. Technical report, Federal Office of Meteorology and Climatology, MeteoSwiss, Arbeitsbericht 216.

Diaz H, Folland C, Manabe T, Parker D, Reynolds R, Woodruff S. 2002. Workshop on advances in the use of historical marine climate data. *CLIVAR Exchanges* **25**: 71–73.

Dukes MDG, Palutikof JP. 1995. Estimation of extreme wind speeds with very long return periods. *Journal of Applied Meteorology* **34**(9): 1950–1961.

DWD. 2001. Extreme wetter- und witterungsereignisse im 20. jahrhundert, Internet, URL http://www.dwd.de/de/FundE/Klima/KLIS/prod/KSB/ksb01/Extreme.% pdf.

ECMWF. 2003. IFS Documentation Cycle CY23r41, Part IV: Physical Processes, Technical report, European Center for Medium Range Forecasting, URL http://www.ecmwf.int/research/ifsdocs/CY23r4/index.html.

Emanuel KA. 2005. Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **436**: 686–688.

Ferro CAT, Segers J. 2003. Inference for clusters of extreme values. *Journal of The Royal Statistical Society Series B-Statistical Methodology* **65**: 545–556.

Fisher RA, Tippett LHC. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* **24**: 180–190.

Frei C, Scholl R, Fukutome S, Schmidli R, Vidale PL. 2006. Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research-Atmospheres* **111**(D6): D06 105.

Gumbel EJ. 1958. *Statistics of Extremes*. Columbia University Press: New York.

Held IM. 1993. Large-scale dynamics and global warming. *Bulletin of the American Meteorological Society* **74**(2): 228–241.

Heneka P, Hofherr T, Ruck B, Kottmeier C. 2006. Winter storm risk of residential structures-model development and application to the German state of Baden-Wurttemberg. *Natural Hazards And Earth System Sciences* **6**(5): 721–733.

Holton JR. 2004. *An Introduction to Dynamic Meteorology*. Elsevier: Burlington.

Horvitz DG, Thompson DJ. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**: 663–685.

Hurrell JW, Kushnir Y, Ottersen G, Visbeck M. 2002. *The North Atlantic Oscillation: Climatic Significance and Environmental Impact Geophysical Monograph 134*, ch. An Overview of the North Atlantic Oscillation. American Geophysical Union: Washington, DC; 1–36.

Jones PD, Horton EB, Folland CK, Hulme M, Parker DE, Basnett TA. 1999. The use of indices to identify changes in climatic extremes. *Climatic Change* **42**(1): 131–149.

Jung T, Klinker E, Uppala S. 2005. Reanalysis and reforecast of three major European storms of the twentieth century using the ECMWF forecasting system. Part ii: Ensemble forecasts. *Meteorological Applications* **12**(2): 111–122.

Kaas E, Li TS, Schmith T. 1996. Statistical hindcast of wind climatology in the North Atlantic and northwestern European region. *Climate Research* **7**: 97–110.

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

Knippertz P, Ulbrich U, Speth P. 2000. Changing cyclones and surface wind speeds over the North Atlantic and Europe in a transient ghg experiment. *Climate Research* **15**(2): 109–122.

KNMI. 2007. Zware stormen in Nederland sinds 1910, Internet, URL http://www.knmi.nl/klimatologie/lijsten/zwarestorm.html.

Kristensen L, Rathmann O, Hansen SO. 1999. Extreme winds in Denmark, Technical report, Riso National Laboratory.

Kunz H, Scherrer SC, Liniger MA, Appenzeller C. 2007. The evolution of ERA-40 surface temperatures and total ozone compared to observed Swiss time series. *Meteorologische Zeitschrift* **16**: 171–181.

Lamb HH. 1991. *Historic Storms of the North Sea, British Isles, and Northwest Europe*. Cambridge University Press: Europe.

Leckebusch GC, Koffi B, Ulbrich U, Pinto JG, Spangehl T, Zacharias S. 2006. Analysis of frequency and intensity of European winter storm events from a multi-model perspective, at synoptic and regional scales. *Climate Research* **31**(1): 59–74.

Mailier PJ, Stephenson DB, Ferro CAT, Hodges KI. 2006. Serial clustering of extratropical cyclones. *Monthly Weather Review* **134**(8): 2224–2240.

Miller C. 2003. A once in 50-year wind speed map for Europe derived from mean sea level pressure measurements. *Journal of Wind Engineering and Industrial Aerodynamics* **91**(12–15): 1813–1826.

Monahan AH. 2004. A simple model for the skewness of global sea surface winds. *Journal of the Atmospheric Sciences* **61**(16): 2037–2049.

Monahan AH. 2006. The probability distribution of sea surface wind speeds. part 1: Theory and seawinds observations. *Journal of Climate* **19**(4): 497–520.

Morland J, Deuber B, Feist DG, Martin L, Nyeki S, Kampfer N, Matzler C, Jeannet P, Vuilleumier L. 2006. The startwave atmospheric water database. *Atmospheric Chemistry and Physics* **6**: 2039–2056.

MunichRe. 2000. Winter storms in Europe (II) – analysis of 1999 losses and loss potentials, Münchener Rückversicherungs-Gesellschaft.

Palutikof JP, Brabson BB, Lister DH, Adcock ST. 1999. A review of methods to calculate extreme wind speeds. *Meteorological Applications* **6**(2): 119–132.

Pryor SC, Barthelmie RJ. 2003. Long-term trends in near-surface flow over the Baltic. *International Journal of Climatology* **23**(3): 271–289.

Pryor SC, Bathelmie RJ, Schoof JT. 2006a. Inter-annual variability of wind indices across Europe. *Wind Energy* **9**: 27–38.

Pryor SC, Schoof JT, Barthelmie RJ. 2006b. Winds of change?: Projections of near-surface winds under climate change scenarios. *Geophysical Research Letters* **33**(11): L11 702.

Raible C. 2007. On the relation between extremes of midlatitude cyclones and the atmosperhic circulation using ERA-40. *Geophysical Research Letters* **34**: L07 703.

Raible CC, Della-Marta PM, Schwierz C, Wernli H, Blender R. 2008. Northern hemisphere midlatitude cyclones: A comparison of detection and tracking methods and different reanalyses. *Monthly Weather Review* **136**: 880–897.

R Development Core Team. 2005. R: A language and environment for statistical computing, URL http://www.R-project.org., ISBN 3-900051-07-0.

Rockel B, Woth K. 2007. Future changes in near surface wind speed extremes over Europe from an ensemble of RCM simulations. *Climatic Change* **81**: 267–280.

Sacré C. 2002. Extreme wind speed in France: the '99 storms and their consequences. *Journal of Wind Engineering and Industrial Aerodynamics* **90**(10): 1163–1171.

Schiesser HH, Pfister C, Bader J. 1997. Winter storms in Switzerland north of the Alps 1864/1865–1993/1994. *Theoretical and Applied Climatology* **58**(1–2): 1–19.

Schinke H. 1993. On the occurence of deep cyclones over Europe and the North Atlantic in the period 1930–1991. *Beiträge zur Physik der Atmosphäre* **66**: 223–237.

Schmith T, Kaas E, Li TS. 1998. Northeast Atlantic winter storminess 1875–1995 re-analysed. *Climate Dynamics* **14**: 529–536.

Schubiger F, Lopez S, Turina A, Romaneessen E. 2004. Re-Analyse von extremen wetterereignissen mit dem LM ausgehend aus dem ECMWF ERA-40 datensatz.

Schwierz C, Köllner-Heck P, Zenklusen E, Bresch DN, Vidale PMW, Schär C. 2008. Modelling European winter wind storm losses in current and future climate. *Climatic Change*: 1–35, Submitted.

Seierstad IA, Stephenson DB, Kvamsto NG. 2007. How useful are teleconnection patterns for explaining variability in extratropical storminess? *Tellus Series A-Dynamic Meteorology and Oceanography* **59**(2): 170–181.

Smits A, Tank AMGK, Konnen GP. 2005. Trends in storminess over the Netherlands, 1962–2002. *International Journal of Climatology* **25**(10): 1331–1344.

Sterl A. 2004. On the (in)homogeneity of reanalysis products. *Journal of Climate* **17**(19): 3866–3873.

SwissRe. 2000. *Storm Over Europe – An Underestimated Risk*. Swiss Reinsurance Company: Zürich.

Turina A, Lopez S, Schubiger F, Aichinger M, Romaneessen E. 2004. High-resolution re-analysis of extreme weather events – part a: The hazard model.

Ulbrich U, Pinto JG, Kupfer H, Leckebusch GC, Spangehl T, Reyers M. 2008. Changing northern hemisphere storm tracks in an ensemble of IPCC climate change simulations. *Journal of Climate* **21**: 1669–1679.

Uppala SM, Kallberg PW, Simmons AJ, Andrae U, Bechtold VD, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van De Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Holm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**(612): 2961–3012.

van den Brink HW, Konnen GP, Opsteegh JD. 2004a. Statistics of extreme synoptic-scale wind speeds in ensemble simulations of current and future climate. *Journal of Climate* **17**(23): 4564–4574.

van den Brink HW, Konnen GP, Opsteegh JD, van Oldenborgh GJ, Burgers G. 2004b. Improving 10(4)-year surge level estimates using data of the ECMWF seasonal prediction system. *Geophysical Research Letters* **31**(17): L17 210.

Walser A, Arpagaus M, Appenzeller C, Leutbecher M. 2006. The impact of moist singular vectors and horizontal resolution on short-range limited-area ensemble forecasts for two European winter storms. *Monthly Weather Review* **134**(10): 2877–2887.

Walter A, Keuler K, Jacob D, Knoche R, Block A, Kotlarski S, Muller-Westermeier G, Rechid D, Ahrens W. 2006. A high resolution reference data set of German wind velocity 1951–2001 and comparison with regional climate model results. *Meteorologische Zeitschrift* **15**(6): 585–596.

Wanner H, Bronnimann S, Casty C, Gyalistras D, Luterbacher J, Schmutz C, Stephenson D, Xoplaki E. 2001. North Atlantic Oscillation – Concepts and Studies. *Surveys in Geophysics* **22**: 321–382.

Weisse R, von Storch H, Feser F. 2005. Northeast Atlantic and North Sea storminess as simulated by a regional climate model during 1958–2001 and comparison with observations. *Journal of Climate* **18**(3): 465–479.

Wernli H, Dirren S, Liniger MA, Zillig M. 2002. Dynamical aspects of the life cycle of the winter storm 'lothar' (24–26 december 1999). *Quarterly Journal of the Royal Meteorological Society* **128**(580): 405–429.

Yan ZW, Bate S, Chandler RE, Isham V, Wheater H. 2002. An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. *Journal of Climate* **15**(15): 2073–2088.

Yan Z, Bate S, Chandler RE, Isham V, Wheater H. 2006. Changes in extreme wind speeds in NW Europe simulated by generalized linear models. *Theoretical And Applied Climatology* **83**(1–4): 121–137.